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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2016 (21.07.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/115179 A1

(51) International Patent Classification:
C12Q 1/00 (2006.01)

(21) International Application Number:
PCT/US2016/013113

(22) International Filing Date:
12 January 2016 (12.01.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/102,347 12 January 2015 (12.01.2015) US

(71) Applicants: **MASSACHUSETTS INSTITUTE OF TECHNOLOGY** [US/US]; 77 Massachusetts Avenue, Cambridge, MA 02139 (US). **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [—/US]; 17 Quincy Street, Cambridge, MA 02138 (US).

(72) Inventors: **SHAREI, Armon, R.**; 333 Highland Ave, Somerville, MA 02144 (US). **LAJOIE, Marc**; 2331 Minor Ave E, Seattle, WA (US). **JENSEN, Klavs, F.**; 103 Outlook Drive, Lexington, MA 02421 (US). **LANGER, Robert, S.**; 98 Montvale Road, Newton, MA 02459 (US).

(74) Agents: **BEATTIE, Ingrid, A.** et al.; Mintz Levin Cohn Ferris Glovsky And Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

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

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

Published:

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

— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: GENE EDITING THROUGH MICROFLUIDIC DELIVERY

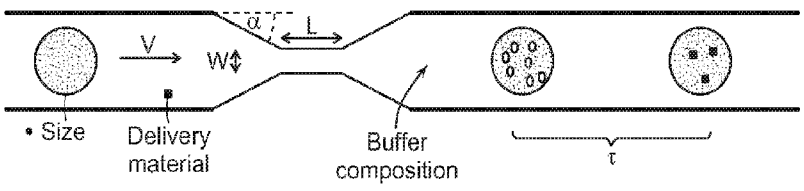


FIG. 2A

(57) **Abstract**: Gene editing can be performed by introducing gene-editing components into a cell by mechanical cell disruption. Related apparatus, systems, techniques, and articles are also described. The methods and systems of the invention solve the problem of intracellular delivery of gene editing components and gene editing complexes to target cells. The results described herein indicate that delivery of gene editing components, e.g., protein, ribonucleic acid (RNA), and deoxyribonucleic acid (DNA), by mechanical disruption of cell membranes leads to successful gene editing. Because intracellular delivery of gene editing materials is a current challenge, the methods provide a robust mechanism to engineer target cells without the use of potentially harmful viral vectors or electric fields.

WO 2016/115179 A1

GENE EDITING THROUGH MICROFLUIDIC DELIVERY

RELATED APPLICATIONS

[0001] This application claims the benefit of priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No: 62/102,347, filed January 12, 2015, which is incorporated herein by reference in its entirety.

STATEMENT AS TO FEDERALLY-SPONSORED RESEARCH

[0002] This invention was made with Government support under Grant Number R01GM101420-01A1 awarded by the National Institutes of Health. The Government has certain rights in the invention.

TECHNICAL FIELD

[0003] The subject matter described herein relates to gene editing by introducing gene-editing components into a cell by mechanical cell disruption.

REFERENCE TO THE SEQUENCE LISTING

[0004] This application incorporates-by-reference nucleotide and/or amino acid sequences which are present in the file named "38172-510001WO_Sequence_Listing_ST25.txt," which is 517 kilobytes in size, and which was created January 12, 2016 in the IBM-PC machine format, having an operating system compatibility with MS-Windows, which is contained in the text file filed January 12, 2016 as part of this application.

BACKGROUND

[0005] Genome editing technologies, such as clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR associated protein 9 (Cas9) and transcription activator-like effector nucleases (TALENs), have shown much potential in their ability to change the genetic code of cells. These technologies could thus enable novel insights in drug discovery and lead to the development of next generation gene therapies. Gene editing complexes, which include a protein

component and a nucleic acid component, e.g., deoxyribonucleic acid (DNA) and/or ribonucleic acid (RNA) cannot readily cross the cellular membrane. Thus, delivery of such complexes has been a challenge.

SUMMARY

[0006] The methods and systems of the invention solve the problem of intracellular delivery of gene editing components and gene editing complexes to target cells. The results described herein indicate that delivery of gene editing components, *e.g.*, protein, ribonucleic acid (RNA), and deoxyribonucleic acid (DNA), by mechanical disruption of cell membranes leads to successful gene editing. Because intracellular delivery of gene editing materials is a current challenge, the methods provide a robust mechanism to engineer target cells without the use of potentially harmful viral vectors or electric fields. Moreover, the scalability and relative simplicity of the process make it suitable for broad adoption. The strategy and methods are suitable for genome engineering applications in research and therapeutics.

[0007] Accordingly, a method for delivering a protein-nucleic acid complex into a cell is carried out by providing a cell in a suspension solution; passing the solution through a microfluidic channel that includes a cell-deforming constriction; passing the cell through the constriction such that a pressure is applied to the cell causing perturbations of the cell large enough for said protein-nucleic acid complex to pass through; and incubating the cell in a complex-containing solution for a predetermined time before or after the cell passes through the constriction. An exemplary protein-nucleic acid complex comprises gene editing components. For example, the protein-nucleic acid complex comprises a Cas protein (such as a Cas9 protein) and a guide RNA (gRNA) or donor DNA. In other examples, the protein-nucleic acid complex comprises a TALEN protein, Zinc-finger nuclease (ZFN), mega nuclease, or Cre recombinase.

[0008] The methods and system is generally applicable to cytosolic delivery of complexes, *e.g.*, a protein – protein complex, small molecule + RNA complex, etc.

[0009] A variety of target cells types are processed in this manner. For example, the cell comprises a mammalian cell such as an immune cell (*e.g.*, T cell) or a stem cell such as a hematopoetic stem cell.

[0010] The microfluidic system may include a plurality of microfluidic channels. Each of the microfluidic channels of the plurality defines a lumen and is configured such that a cell suspended in a buffer can pass through the lumen. In some embodiments, microfluidic channels include one or more cell-deforming constrictions. In some embodiments, the diameter of the constriction is a function of the diameter of the cell. Thus, there may be many microfluidic channels within a microfluidic system of the invention. For example, the microfluidic system may include a plurality of the microfluidic channels arranged in parallel, *e.g.*, 2, 5, 10, 20, 40, 45, 50, 75, 100, 500, 1,000 or more.

[0011] Microfluidic systems having a plurality of parallel microfluidic channels allow for the high-throughput delivery of payloads to cells. Many cells can be passed through each parallel channel one after the other. It will be understood that, depending on context, a reference to a "cell" herein may refer to more than one cell.

[0012] The diameter of the constriction is chosen depending on the dimensions of the cell type to be treated. In some embodiments, the cell may be primarily compressed by the fluid flow. In some embodiments, the diameter is less than the diameter of the cell. For example, the diameter of the constriction may be substantially or about 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 20-99% of the diameter of the cell. For example, the constriction is substantially 20-99% of the diameter of the cell, e.g., a diameter of the constriction is substantially 60% of the diameter of the cell. Non-limiting examples of the diameter of the constriction include substantially or about 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 2-10μm, or 10-20μm. Different lengths of the constriction are also possible. Non-limiting examples of constriction lengths include substantially or about 10, 15, 20, 24, 30, 40, 50, 60, 70, 80, 90, 100, 10-40, 10-50, 10-60, or 10-100μm.

[0013] The subject matter described herein provides many technical advantages over methods that deliver components (or nucleic acids encoding the components) of the complex piecemeal. Most gene editing systems require complex formation to occur inside the cell, which may be an inefficient process. Advantages of delivering the RNA and Cas (such as Cas9) in complex form (or other protein/nucleic acid gene editing composite assemblies) include better efficiency and specificity compared to other methods. By obviating the need for mRNA or DNA manipulation to express the Cas protein (such as a Cas9 protein), one can reduce how much time the cell spends exposed to the protein thus reducing the chance of off-

target effects. Having the editing components delivered to the cell in complexed form also eliminates/minimizes the risk of the Cas (such as Cas9) complexing with other RNA strands in the cell and cleaving the wrong sites. For example, the RNA alone may be detected by intracellular and extracellular Toll-like receptor (TLR) and pattern recognition receptors, prompting an interferon response or other antiviral pathways. The complexed form does not interact with these pathways and can thus avoid undesirable side effects.

[0014] By complexing *in vitro* prior to delivery into a cell, one can precisely control the Cas (such as Cas9) and gRNA complexing reaction thus ensuring optimal functionality, while complexes forming in the cytosol may not be as efficient. For example, delivering the complex cytosolically ensures simultaneous interaction of Cas (such as Cas9) and gRNA with the target DNA. The complexes formed *in vitro* and delivered to the cell as described herein are fully functional and ready-to-go upon gaining access to the cytoplasm of the target cell.

[0015] The approach described here is relevant to any protein + RNA/DNA based system to guide the nuclease as the delivery process is independent of the exact size and composition of the complex and because complex formation of the editing materials occurs and is controlled *in vitro* under their optimal conditions.

[0016] Implementations of the invention may also provide one or more of the following features. Deforming the cell includes deforming the cell for substantially or about 1 μs to 10 ms, e.g., 10 μs, 50 μs, 100 μs, 500 μs, and 750 μs. Incubating occurs for 0.0001 seconds to 20 minutes or more, e.g., substantially or about 1 second, 30 seconds, 90 seconds, 270 seconds, and 900 seconds.

[0017] The pressure and speeds at which a cell is passed through a microfluidic channel may also vary. In some embodiments, a pressure of substantially or about 10-35psi is used to pass the solution containing a cell through a microfluidic channel. The speed may be adjusted for a variety of reasons, including to improve viability of the treated cells while maintaining high payload delivery. In some embodiments, the cell passes through the microfluidic channel at a speed of substantially or about 300mm/s, 400mm/s, 500mm/s, 600mm/s, 700mm/s, 800mm/s, 900mm/s, 100-300mm/s, 200-700mm/s, 250-400mm/s, 1-1000mm/s, 1m/s, 2m/s, 3m/s, 4m/s, 5m/s, 6m/s, 7m/s, 8m/s, 9m/s, 10m/s, 0.01-5m/s, 5-10m/s, or 0.01-10m/s. Where the cell is a plurality of cells, substantially or about 5, 10, 20, 30, 40, 50, 60,

70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 90-95, or 80-99% or more of the cells may be viable after passing through the constriction. In some embodiments, the cells are viable for at least about any of one hour, two hours, three hours, four hours, five hours, six hours, seven hours, eight hours, nine hours, ten hours, twelve hours, eighteen hours, twenty four hours, or forty eight hours after passing through the constriction.

[0018] In some examples, a device of the invention includes 2, 10, 20, 25, 45, 50, 75, 100 or more channels. In some embodiments, cells are moved, *e.g.*, pushed, through the channels or conduits by application of pressure. In some embodiments, a cell driver can apply the pressure. A cell driver can include, for example, a pressure pump, a gas cylinder, a compressor, a vacuum pump, a syringe, a syringe pump, a peristaltic pump, a manual syringe, a pipette, a piston, a capillary actor, and gravity. As an alternative to channels, the cells may be passed through a constriction in the form of a net. In either case, the width of the constriction through which the cells traverse is 20-99% of the width or diameter of the cell to be treated in its unconstricted, *i.e.*, suspended, state. Temperature can affect the uptake of compositions and affect viability.

[0019] In certain embodiments, a temperature of 0 to 45°C is used during cell treatment, e.g., 0-25°C. In various embodiments, the methods are carried out at room temperature (*e.g.*, 20°C), physiological temperature (*e.g.*, 39°C), higher than physiological temperature, or reduced temperature (*e.g.*, 0.1°C), or temperatures between these exemplary temperatures (*e.g.*, 0.1 to 40°C).

[0020] In some embodiments relating to immune cells, treatment of unstimulated T cells, B cells and/or monocytes is carried out at temperature of 4-8°C, e.g., on ice. In another example, dendritic cells, activated T cells, and/or activated B cells are treated using the device at temperatures of 20-25°C, e.g., at typical ambient room temperature.

[0021] In some embodiments, following controlled injury (e.g., perturbations) to the cell by constriction, stretching, and/or a pulse of high shear rate, the cells are incubated in a delivery solution that contains the complex that one wishes to introduce into the cell. Controlled injury may be characterized as small, *e.g.*, 200 nm in diameter, perturbation in the cell membrane. The recovery period for the cells is on the order of a few minutes to close the injury caused by passing through the constriction. The delivery period comprises 1-10 minutes or longer, *e.g.*, 15, 20, 30,

60 minutes or more, with 2-5 minutes being optimal when operated at room temperature.

[0022] In some embodiments of the device and methods described herein, passage of stem cells or progenitor cells such as induced pluripotent stem cells (iPSCs) through a constriction channel does not induce differentiation, but does reliably induce uptake of compositions into the cell. For example, gene editing compounds are introduced into such cells without complications associated with the method by which the factor(s) was introduced into the cell.

[0023] The size and duration of temporary perturbations in cell membranes can be modified by adjusting various factors, such as the diameter of cell-deforming constrictions and the speed at which cells pass through the constrictions. Disclosures regarding the size and duration of perturbations provided herein should not be interpreted as limiting. Non-limiting descriptions of perturbations and recovery are provided in Sharei et al., (2014) Integr. Biol., 6, 470-475, the entire content of which is incorporated herein by reference. In some embodiments, the perturbations of the cell membrane may be characterized by a maximum diameter of substantially or about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm. In various embodiments, perturbations of the cell membrane having a maximum diameter of substantially or about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm persist on the cell membrane for at least substantially or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 1-10 minutes or more (11, 13, 15, 18, 20 minutes or more).

[0024] In various embodiments, the diameter is less than the diameter of the cell. For example, the diameter of the constriction may be substantially or about 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 20-99% of the diameter of the cell. Non-limiting examples of the diameter of the constriction include substantially or about 4, 5, 6, 7, 8, 9, 10, 15, 20 4-10μm, or 10-20μm. Different lengths of the constriction are also possible. Non-limiting examples of constriction lengths include substantially or about 10, 15, 20, 24, 30, 40, 50, 60, 70, 80, 90, 100 10-40, 10-50, 10-60, or 10-100μm.

[0025] Many cells are between 5-20 μm in diameter, e.g. unstimulated T cells are 7-8 μm in diameter. For example, the diameter of the constriction portion is 4.5, 5, 5.5, 6, or 6.5 μm for processing of single cells. In another example, the size/diameter of the constricted portion for processing of a human egg is between 60

μm and 80 μm, although larger and smaller constrictions are possible (diameter of a human ovum is approximately 100 μm). In yet another example, embryos (*e.g.*, clusters of 2-3 cells) are processed using a constriction diameter of between 12 μm and 17 μm. In a non-limiting example relating to unstimulated T and B cells, the device comprises a constriction having a length of about 10, 15, 20, 25, 30, or 10-30 μm, a width of about 3, 3.5, 4, or 3-4 μm, a depth of about 15, 20, 25, or 15-25 μm, and/or an about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 5-15 degree angle. Examples of microfluidic devices useful for delivering payloads into immune cells are described in PCT International Patent Application No. PCT/US2015/058489, Delivery of Biomolecules to Immune Cells, filed October 30, 2015, the entire contents of which are incorporated herein by reference.

[0026] In addition to single cells, even very large cells, e.g., eggs (approximately 200 μm in diameter), clusters of cells, e.g., 2-5 cell clusters such as an embryo comprising 2-3 cells, are treated to take up target compositions. The size of the aperture is adjusted accordingly, i.e., such that the width of the constriction is just below the size of the cluster. For example, the width of the channel is 20-99% of the width of the cell cluster.

[0027] Cells or cell clusters are purified/isolated or enriched for the desired cell type. Dendritic cells or other cells, e.g., immune cells such as macrophages, B cells, T cells, or stem cells such as embryonic stem cells or iPS, used in the methods are purified or enriched. For example, cells are isolated or enriched by virtue of their expression of cell surface markers or other identifying characteristics. Dendritic cells are identified and isolated by virtue of their expression of the β-intergrin, CD11c or other identifying cell surface markers. With regard to cells, the term "isolated" means that the cell is substantially free of other cell types or cellular material with which it naturally occurs. For example, a sample of cells of a particular tissue type or phenotype is "substantially pure" when it is at least 60% of the cell population. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99% or 100%, of the cell population. Purity is measured by any appropriate standard method, for example, by fluorescence-activated cell sorting (FACS).

[0028] Payload compositions such as polynucleotides, polypeptides, or other agents (e.g., Cas9 and gRNA) are purified and/or isolated. Specifically, as used herein, an "isolated" or "purified" nucleic acid molecule, polynucleotide, polypeptide,

or protein, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. Purified compounds are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. For example, a purified compound is one that is at least 90%, 91%, 92%, 93%, 94%, 95%, 98%, 99%, or 100% (w/w) of the desired compound by weight. Purity is measured by any appropriate standard method, for example, by column chromatography, thin layer chromatography, or high-performance liquid chromatography (HPLC) analysis. A purified or isolated polynucleotide (ribonucleic acid (RNA) or deoxyribonucleic acid (DNA)) is free of the genes or sequences that flank it in its naturally-occurring state. Examples of a an isolated or purified nucleic acid molecule include: (a) a DNA which is part of a naturally occurring genomic DNA molecule, but is not flanked by both of the nucleic acid sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner, such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Isolated nucleic acid molecules according to the present invention further include molecules produced synthetically, as well as any nucleic acids that have been altered chemically and/or that have modified backbones.

[0029] Complexes are prepared from purified modules or components, e.g., purified protein(s) and purified nucleic acids (RNA and/or DNA). Thus, the ratios of the components are controlled and tailored as desired to achieve a desired gene editing result. The present method is particularly suitable for delivery of sensitive payloads, e.g., protein-RNA/DNA complexes, e.g., complexes that are 40 kDa, 50 kDa, 75 kDa, 100 kDa, and up to 120, 130, 150, 200, 250, 300 kDa or more.

[0030] Surfactants (*e.g.*, 0.1-10% w/w) are optionally used (e.g., poloxamer, animal derived serum, albumin protein) in the flow buffer. Delivery of molecules into cells is not affected by the presence of surfactants; however, surfactants are optionally used to reduce clogging of the device during operation.

[0031] In some aspects, the device is made from silicon, metal (e.g., stainless steel), plastic (e.g., polystyrene), ceramics, or any other material suitable for forming one or more appropriately sized channels or conduits. In some aspects, the device is formed of materials suitable for etching micron scaled features and includes one or more channels or conduits through which cells pass. Silicon is particularly well suited, because micro patterning methods are well established with this material, thus it is easier to fabricate new devices, change designs, etc. Additionally, the stiffness of silicon can provide advantages over more flexible substrates like Polydimethylsiloxane (PDMS), e.g., higher delivery rates. For example, the device includes 2, 10, 20, 25, 45, 50 75, 100 or more channels. The device is microfabricated by etching the silicon. Cells are moved, *e.g.*, pushed, through the channels or conduits by application of pressure. A cell driver can apply the pressure. A cell driver can include, for example, a pressure pump, a gas cylinder, a compressor, a vacuum pump, a syringe, a syringe pump, a peristaltic pump, a manual syringe, a pipette, a piston, a capillary actor, and gravity. As an alternative to channels, the cells may be passed through a constriction in the form of a net. In either case, the width of the constriction through which the cells traverse is 20-99% of the width or diameter of the cell to be treated in its unconstricted, *i.e.*, suspended state.

[0032] Various implementations of the invention may also provide one or more of the following clinical and research capabilities. Quantitative delivery of gene-editing complexes or components thereof to cell models for improved screening and dosage studies can be achieved. The method could be deployed as a high throughput method of screening protein activity in the cytosol to help identify protein therapeutics or understand disease mechanisms. The devices and techniques are useful for intracellular delivery of gene-editing complexes to a specific subset of circulating blood cells (e.g. lymphocytes) or even whole blood; high throughput delivery of complexes or components thereof into cells, especially oocytes and zygotes; targeted cell differentiation by introducing gene-editing (optionally together with genetic material such as donor DNA) to induce cell reprogramming to produce iPS cells; delivery of DNA and/or recombination enzymes into embryonic stem cells for the development of transgenic or mutant stem cell lines; delivery of DNA and/or recombination enzymes into zygotes for the development of transgenic or mutant organisms; dendritic cell (DC) cell activation; iPS cell generation; creating mutations in normal or diseased cells (such as cancer cells) to study the contribution of one or

more genes to cellular function and/or disease; and stem cell differentiation. Skin cells used in connection with plastic surgery are also modified using the devices and method described herein. Methods of delivering gene-editing proteins disclosed herein may also be used to generate CAR-T cells or to genetically modify hematopoietic stem cells (HSCs) for treating genetic and other diseases. In embodiments relating to HSCs, a subject may receive an autologous, syngeneic, or an allogeneic edited HSC. In various embodiments, cells of a subject may be ablated before the subject receives a gene-edited cell. For example, bone marrow cells of a subject may be ablated with radiation or chemically before the subject receives a gene-edited HSC. In some embodiments, a gene associated with beta thalassemia or sickle cell anemia is edited using a method or composition disclosed herein. Cells processed *ex vivo* or *in vitro*, *i.e.*, outside of the body of a subject, in accordance with the invention are useful for subsequent administration to a subject in need of treatment or diagnosis of a pathology. In alternative embodiments, *in vivo* cell processing is carried out.

[0033] In various embodiments, the SHP2 gene is edited/mutated to reduce the activity thereof or knock out or reduce SHP2 expression. In such embodiments relating to gene editing in T cells, the T cells become less responsive to immunosuppressive signals and have increased activity toward tumors. In such embodiments, the T cells may be more responsive to tumor antigens and more effective at treating cancer.

[0034] Aspects of the present subject matter relate to the rapid and transient delivery of protein-protein as well as protein-nucleic acid complexes, e.g., gene-editing complexes to cells. A nucleic acid component of the complex comprises a deoxynucleic acid (DNA), ribonucleic acid (RNA, *e.g.*, mRNA, gRNA) or other double-stranded or single stranded nucleic acid compounds, respectively. For example, the delivery of a gene-editing complex (e.g., a ribonucleoprotein (RNP)) may achieve gene editing faster than if an expression vector encoding components of the gene editing complex (*e.g.* a Cas protein and a gRNA) was delivered to the cell. For example, the gene may be edited (e.g., mutated or replaced) in the cell 0.1, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, or 0.1-4 hours sooner than in a corresponding cell that has received microfluidic or electroporation-mediated delivery of an expression vector that encodes gene editing complex components.

[0035] Each embodiment disclosed herein is contemplated as being applicable to each of the other disclosed embodiments. Thus, all combinations of the various elements described herein are within the scope of the invention.

[0036] Related apparatus, systems, techniques, and articles are also described.

[0037] The details of one or more variations of the subject matter described herein are set forth in the accompanying drawings and the description below. Other features and advantages of the subject matter described herein will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

[0038] FIG. 1A and 1B are schematic diagrams of a microfluidic system in which cells are exposed to the delivery material (payload) after passing through the constriction;

[0039] FIG. 2A and 2B are schematic diagrams of an embodiment of a microfluidic system in depicting parameters such as channel depth, width, and length.

[0040] FIG. 3A and 3B are a series of flow cytometry plots (FIG. 3A) and a bar graph (FIG. 3B) of the recombination efficiency for K562 reporter cells that had genetic editing material delivered to cell cytoplasm using the microfluidic device illustrated in FIGS. 1-2. For the reporter used, there is a frame-shifted GFP gene in the cell line. To perform the editing the Cas9 gRNA complex and a donor oligonucleotide would be delivered. The complex would cut near the GFP site and the oligonucleotide would insert itself into the cut site. Successful insertion of the oligo would correct the gene and result in GFP expression which is what is seen in this figure. Thus, in this assay something turned on upon gene editing as opposed to turned off. A gene would be expected to be turned off in instances where a gene (or depending on the context, a nucleotide or portion thereof) was being deleted in the absence of a donor oligonucleotide.

[0041] FIG. 4 is a series of FACs plots showing β_2 microglobulin component of MHC class 1 (B2M) expression vs. delivered dextran for four different cell populations obtained using FACS. The delivery of the RNP using the 30-4 chip at 90 psi results in a 54.4% reduction in B2M expression as compared to the endocytosis control whereas the 10-4 chip at 90 psi results in a 25.2% reduction in B2M expression. B2M expression on the CAS9 control is not significantly different

than the endocytosis control. The longer constriction chip results in more delivery of the RNP complex and a larger reduction in B2M expression.

[0042] FIG. 5 is a series of FACs plots showing reduced B2M expression in a dose dependent manner determined by FACS as a measure of functional editing with the indicated conditions.

[0043] FIG. 6 shows a FoxP3 genomic sequence running from the first sheet of FIG. 6 to the last sheet of FIG. 6 (SEQ ID NO: 56). Exons belonging to FoxP3 are shown in underlined and highlighted letters. Other exons within this region that do not belong to FoxP3 are shown in non-underlined highlighted letters.

[0044] FIG. 7 shows a FoxP3 translated amino acid sequence (SEQ ID NO: 57). Alternating exons are underlined and non-underlined. Bold with italics indicate a residue overlap splice site.

[0045] FIG. 8 shows a SHP1 genomic sequence running from the first sheet of FIG. 8 to the last sheet of FIG. 8 (SEQ ID NO: 58). Exons belonging to SHP1 are shown in underlined and highlighted letters. Other exons within this region that do not belong to SHP1 are shown in non-underlined highlighted letters.

[0046] FIG. 9 shows a SHP1 translated amino acid sequence (SEQ ID NO: 59). Alternating exons are underlined and non-underlined. Bold with italics indicate a residue overlap splice site.

[0047] FIG. 10 shows a SHP2 genomic sequence running from the first sheet of FIG. 10 to the last sheet of FIG. 10 (SEQ ID NO: 60). Exons belonging to SHP2 are shown in underlined and highlighted letters. SEQ ID NO: 60 is also as follows:

AGGCTCAAGCAATCCTCTCACCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCGCGCCA
CCACGCCCGGCTAATTTTTGTATTTTTTGTAGAGATGGGATTTCACTATTTTGCCCGGGC
TGGTTCCCAACTCCTGGACTCAAGCGATTCGCCCGCCTCAGCCTCCCAAAGGGAAGTGCT
GGGATTTCAGGCGTGTGCCACCGCTCCCACCCCAAAGTAGTATTTATTGTAATTATTATT
ATTATTTTGAGACGGAGTCTCGCTCTATTGCCAGGCTGGAGTGCAGTGGCGCGATCTCGG
CTCAATGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCTTCAGACTCCCAAGCAG
CTGGGACTACAGGCGCCCCCCACCACGCCAGGCTAATTCTTGAATTTTTAGTGGAGACGG
GGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCTTGACCTCGTGATCCGCCCACCTCG
GCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCAGCCTATTATTATTTTT
TTAGGCAGTGTCTTGCCCTGTCGCTCAGGGTGTAGTGCAGTGGCGTGATCACGACTCACT
GCAGCCCCGACTTCTCGGGCTTAAGTTATCTTCCCGCCGCAGCCTCCACGCCCGGTTAGT
TTTTTGCATTTTTTGTAGAGATGAGGTCTTGCTTTTTTGCCCAGGCTGGCCTCGAACTCC

TTGGCTTAAGCGAACCTCTTGCCGCAGCCTCCCAAAGTGTTGGGATTACGGGCGTGAACC

ACCGCGCCCAGCCTACTATCTTTATCTTACAG**AAAGAAAAGAATGGAGGAAACCGAGGCT**

CGGAGACAGTAGGTAATTTCCCCAAGGTTCCACAGCTAATGAGTGGAGCGGCGATTTGTG

GAACGAAATGAATGAAATCGATGTGGCAGCGGGCCCGGACGGGTCGGTGGCGTAGACGCG

GAGCGCGCAGCTCACACCTGGCGGCCGCGGTTTCCAGGAGGAAGCAAGGATGCTTTGGAC

ACTGTGCGTGGCGCCTCCGCGGAGCCCCCGCGCTGCCATTCCCGGCCGTCGCTCGGTCCT

CCGCTGACGGGAAGCAGGAAGTGGCGGCGGGCGTCGCGAGCGGTGACATCACGGGGGCGA

CGGCGGCGAAGGGCGGGGGCGGAGGAGGAGCGAGCCGGGCCGGGGGGCAGCTGCACAGTC

TCCGGGATCCCCAGGCCTGGAGGGGGGTCTGTGCGCGGCCGGCTGGCTCTGCCCCGCGTC

CGGTCCCGAGCGGGCCTCCCTCGGGCCAGCCCGATGTGACCGAGCCCAGCGGAGCCTGAG

CAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGAGGAACATGACATCGCGGAGGTGAG

GAGCCCCGAGGGGCCCGGCGCGGGCCTCGGCCCGGCCACCGCCGCGTTCGGTTAGCCCCG

TCCGGAAGGGGGCGCCCCGGCCGGGCTTCGGGCTCCCGCCCCGGGTCGGGGTTGGGGGCC

GGTTCCCTCCTCGTCCCCTCGCCCTCCAGGGGCCGGGGGCCGGCCCCACCGCGCCCCCAC

CCCTCGGGTCCCCATTCATTTCCTGCCTCCCCGAGTTCCGGCTGCGGCAGCCCCGGGGAT

GCCCGTCAGGCCCGGGGCAGGTAGAGCCGCCGAGGGAACCACGGGTGCCAGCGGCCAGGC

TCAGCGCCGCATTCCTGACCCATTGCCTCATGAGAATTGCCTCATGGTGATTCCGAAATA

ACCCTGCTCACTTGGGGAGGCTCCTTGGGACACGAGAGGGGAGTTGCGCGGGGCCGGGCC

CCCAGTGGTCTAGTCGTTCTGGCTCACTGTGCCACTTTCGTGCATTTGGGGACTTCACGC

AGGACCCCTGACCCTTTTATATGCCTCTTTGTGTCTTCTTTTCCTCCTACCCCTCACGTG

CCAGAAATGGAAAAACTGACTGTATCTGCAGCCACTAGAAGTATTTCCTTCCTCTGCGAT

CTTCGCTTTGGGAGATGGAAAGGAAGGGAGCCGCATCTCGTTATTTAATCCTTCACTGCA

ACCTTAACAGTCAGGTCACTTTACTGGTACCCGTTTTATGGATGAGGAAACCGAGGCCCA

GAAGCAACATGCTAGTAAATGACAAGATTTGAAACTTAGGAGGATTAGTGAGTTAATGAG

ATCCTTTGAAAGGTCAGGGTAATACTACTACTAATAGCTAACATTTGCTTAGTTCTGACC

ACAGCCCTATCAGATGGCTACTATTATCCCCATTGTAAAGATGAGTAAACCGAGTTTCAG

AGGTTAAGTAAATTGCCTAACCTCACAGCTAGTAGGTGGTGGAGACAGAATCCCTACTTT

TAATCACTATGTTGCTTCTATTATTTTGTAACTATTGCTAACCATTTGTAAGCCTTAATT

TTGTTGTCAAACAGTAGTGTGACCTGTTGTTTTCAGATAGTGATCCTGCTATTTTGTATA

GTCACTCTATATACCACTCACACTTAAGACCCATTGTCTATTCTTTTCCATGATTGTTCA

ATTATGGTCACTGTCTCAGACATTTAAAAAACGATTCAAGCTATTGAGGCTATTTGAATG

AGATTTTCTTTTCTTTTTTTCTTTTTTTTTTTGGAGACGGAGGCTCACTCTGTTGCCCAG

GCTGGAGTGCAGTGGCGCAATCTCGGCTCACCACAATCTCCGCCTCCTAGGTTCAAGCGA

TTCTCCTGCCTCAGCCTCCCAAGTAACTAGGACTACAGGCGCACCACTATGCCCGGCTAA

TTTTTGTATTTTTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGTCTCAAACTCC

TGACCTCGTGATCCGCCCGCCTTGGCCTCCCAAAGTGCTGGAATTACAGGCGTGAGCCAC

CGTACCCAGCCTGAATGAGATTTTTCAAAATATTAGGAATGTCTCCTCCAAACACACCTG

GCATGTTATTCATACATGGATCTGGAATTTAAAAAGGGGAGAAAAAGAAAACTGAGAACT
CGTAGGAAGTGAGTGACTTGGACAGGTCGGTTGGCAAGTGCTTACAGATCTGGGTAATAT
ATAACTGCATTTCAACAGAACAGTGTATAGCCTCAAATGTTCTAATTCTTTAGGGAGCTT
TTAAATAAACAGTTGTCTATTCTTTAATCTGTCAAATAGTCATTGAGCCTTTTGTTCCTG
GTGTCTGCTCTTCCAGACAAGTAAGGATCTGCTGCTTTAGGAGACATCAGACGGGGCTGG
GGGTTGGGAAAAGGTCTGGGTAGTAATAGACCCTACATTGTCCAGTTTGTTCATTTAGAA
GCATAGAAGTGTGGGCATAGTCAAAGTAGCAAGTGGTAAAGATGACAGTTTGAAATGGAG
TAATTCCTTCTCCCCTCCAGCCCTGGTATTATGCACCACCCAAAAAGCCGGGTTATGAAC
ATAATACACATAATTTTGAATGATTCATTATTTTTTGGATTATAAGCCTGTTTTATTTGT
TAACCAGCCTTAATGAGGTATAAATGACATGCAATTAATTGCATATATTTAAATGTACAA
TTTGATCAGTTTTGACATACATATACACTTGGGAAACCACCACCATAGTCAAGATAATGA
ACACATCTATCACCCCTGGTAATTTTGCCTTATGTTCTTTATAATCCTTCCTTTGTTCTT
AGGCAGCCACTATTCTGCTTTCTGTCACTATGTATTAGTTTGCATTTCCTAGAATTTTAT
TTTTAAAAATTTTAAAATTGTTTGAATAGAGATGGGGTCTCACTGTGTTGCCCAGGGCAG
TCTCAAACTCCTGGGTTCAAGTGATCCTCTCACCTTGGCCTCCTGAAGTGTTGGGATTAT
AGGCATGAGACACCCTGCCCAGCCCTAGAATTTTATTATTATTGTTATTATTGTGTTTTT
TTGAGATAGGGTCTCACTTTGTTGCCCAGGCTGGAGTGCAGTGGTGCAATCACTGCAGCC
TTGTTTTCCTAGGCTCAATCCATCCCCCCTCCTCAGCTTTCCGGTTACTGGGGCTACAGG
TGTGCACCACCACACCCGGCTAATTTTTGTATTTTTTTATAGAGACAGGGTTTTGCCATG
TTGGCCAGGCTGGTCTCAAACTCCCGGGCTCAAGCGATCTTCCTGCCTCGGCCTCCCAAA
GTGCTGGGATTACAGGCATGAGCTATTGCGTCCCGCCTTCAAATTACTTTAACCTAGTAT
TAATTCATTCAACAGGAAGTTAATGAGCCAGGCAGGATAAAGCAGTAAGATAGGAAAATA
TTGCTATTTTCATGGCTGAGAGAGAGCAGACAAACACATGACTAAATAGGGCAATTTCAG
GTAGTAATAAATTCTAGGAGGGAAAAAATCCCACAGAAATGTGAGGATGGGAGAATGCAG
TTAGTTTTGATAGGTGGTTTAGAGAAGGTGATCGTGTGAGCTGACACCTGAATGACAATT
AGTAGTCTGAATTTTGTTTTGCTTAATTATCAAAATAACTCCTCTTGGGTTCGGCTTTTA
TATGCATCCAGTAATTAAAATGTAAGTATATTCAATGTACTGATATCTCTCAGCATCATA
GGTAGGAAAACTAAGGCATTCAGCAATTAAGTGACTCCTCCCTTGATCATGTAGCAGTGA
TAGTACTGGATTTAGATTTTGAGGTTGCTTCTCTGCCCTTTTCTGCCTTTGTGAAACCAA
CAAAGCTGCCTGTATTTTCCAACTCTTCCTTCAGCATGTGGTACCTCCTTTACATCTGTT
TTTGTTGCTCTGAAATCCATACGCGACGATGAGCTGAGAGGGGCAGAAAATTGAGCTTGT
TCTGAGACTGGAGGCTTTTGGTTTATCTCTTGCAGGTCAAGTACATTTTGTCCTGGGCTC
TCCCTGGTGGCCACGTTTGTTTATCTCCTGCGGGAGTAAATAAACTTGCCTTGCTGAAAA
ATAACAGTTCTGTGTCTTTGCAGTGGAAACTGGGATGTCTTTATTAACGTTAGGTCCTGA
TGTAAGGCCAAGTTTTTGGTTAGAGTTGCTCAAGTGCAGAGGCCACTGCTAAGATGACTT
ACCCCTCGTGTCCATGGTCAATGTGGAGACTGTTATGAGTGGCACATGATGCTGGAAAAG
CAGAGCCAACTCATGTTTGTAATTGTCCTAGCAGGCCGTGGTGTACTTTGTTAGGCAGCC

ACAGAACAATAGAGAAACTCAGCTTATTCCCCTTCCCTCTGGGAAACACAGACAGTACTT
GCCATCCAACGCCAATGTTTTTAAGGAAGAAAGAGGCAAAAAGTGATGTTGGCAAGGTCT
CTGGGAGTTGTGGACCCCAACCAAGGATTGGAGACCCTGAAATGGATTCAGATGCCCTAA
AATGCAGCCCAGTTCATTACTATGAATTTTGGAGGACTTTGTGCCTTGAGCAAATGTGTA
TATGTGACGCTCTTTGACAACACTGAAATAGGAAAAATACTATCCATGTTCGCGAGGAGC
ACTGAATTTAGAGAGGGAGACAGACTTTTATGCCAGCATCAAATGAATTTGATAAAGCTA
GTACCAAAATGAAATTTGAAATTTTTTTTTTTTGAAATAGAGTCTTACTCAGTCACCCAG
GCTGGAGTGCAGTGATACAATATTGGCTCACTGCAACCTCCACCTCTTGGGTTCAAACAA
TTCTTGTGCCTCAGTCTCCTGAGTAGCTGGGATTACAGGTGCGTGCCACCATGTCTGGCT
AATTTTTATATTTTTAGTAGGGATGGGGTTTCACCATGTTGGCCAGGCCGGTCTTGAACT
CCTGGCCTCAAGTGATCTGCCCACCTTGGCCTTCCAAAGTGCTGGGATTATAGGCATGAG
CTACCACACAAGCCTGAAATTTGAAATGTATTGGTATAGAATATACTGTTTAGAATGTAT
GTGTATATATGTATATTTGTATACTCATATAAACACAAATACACATTGTATGTGTTTCTG
TAATATGTATATCTGTCTACACATACATGTATATACACACATACAATGTCTTTTTTTTTT
TTTTTTTTTTTTGAGACAGGGTCTTACCCTGTTGCCCAGGCTGGAGACTGCAGTGGCATA
ATCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTCAAGTGATCCTCCCATCTCAGCCTCC
TGAGTAGCTGGGACTGACTACAGGCACGTGGCATCAAACTTGTCCAATTTTTCTATTTTT
TTGTAGAGTTAGGGTCTTGCTCTGTTGCCCAGGCTGGTCTCAAATTCCTGGGCTCAAGCT
GTCTGCCTGCCTCGGCCTTCCAAAGTACTAGGATTACAGATGTGAACCACTGTACCTGGC
CTTTACAATGTCTATTTTAAAGATAATGGTTCAAGTTTTTATCATCCCACTGGCCTACTC
TAATGAAACATCTATCCATTCATTGAAGAATTATTTATGGTGGGATAACTCTGTGCCAGG
TACCGTGCTAGGCATTGAGTATTCCAGGTTTTAGGAAACAGCACATGCAAAAGTGCTGAA
GTGGGAGAAGATCTCGGAGTGATTGAAGGCTAGGAGAGAGCAAGTGTGGGAGCTGTGAGG
CTGGGAAGGTGGGAGGTAGGTGGGAGCAGACCACATAGGGATTCTTAATGTCTTTAGTGT
CATGTGGACCATGGAGAGGAGTGTAGATTGTATTTTTAGAGCAATGCAAAATCATAGAAG
GATGTGATCGGGGGAGTGGCATGAGCTGATCTATTTAAAAATATTTCTCTGGCTGCTGTG
AAGGAAGGATTGTAGGAGGCAGGAGTAGATTCAGGGAGATGAGACAAGTGATGAGAGAGG
CTTTGAACTTGGGTAAAAGTAGTTTGTGGAAAGTCTTTTTTGGAGGTAGTTTTTGTTTAT
TGCCTTGTCATCAAAGCAGAGATGCTGACCAATGAAACTCCATGAGAAAATAGTGATTTA
TAAAGACATATCTATGCACTGCCATTAAAAAGCTGCTTGGAAAAAAAGGATAAAAAGCTG
CTTTAACAACTTTTTTTTTTGAGATGGGGTCTTACTCTGTCACCCAGGCTCACGACCTCA
GCTCACTGCAACCTCTGCCTCCCAGGCTCAAGCATTCTCCCACCTCAGCCTCCCGAGTGG
CTGGGACTGCAGGCACACGCCACCATGTCAGGCTAATTGTGTGTGTGTGTGTGTGTGTGT
ATGTGTGTGTGTGTGTGTGTGTGTGTGTGCTGGGACTGCAGGCACACACCACCATGTCAG
GCTAATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA
TGTAGAGATGGGGTTTTGCCATGTTGCCCAGGCTGGTCTCAAAATGTTGCCCAGGCTGGT
CTCAAACTCCTGAGCTCAGGTGATCCACCCGCCTCGGCCTCCAAAGTGCTGGAGATTACA

GACGTGAGCCACTGTGCCCACCTAACAACTTTAAAAAAATTTTGACATTTAGTAGGATAT
TTATTGCATTATTGTTGAGATGGCAAAATATTGGAGACAACTGAAATGTTCATCAGTGGG
GGGGGCTAGTTAAATGAAATACAGTGTAGCATGCATTAGAACACTTTTCAAGAATTTAAC
TTTTTTTGTAGCCTTTTACTTATAATGCTTGTCCCTATTGATGCCTTTTTTTTCAGCATG
ACTTACTCTTTTACTATAGGATATTAAAATTTAATTAGATTAGAAATGAGGAATATTCTT
GTAATCTGTAGAAAGTAACAAACTATAAACTTATTCCCCAAGAACAAATATAATAATTTT
TCTGGAGTAGCAGGTAAGAAAGATATAAATTTATATGTATACAAGAAACTGAAATTAGAC
TTTATACATTTAAAGGTTACAAGTGCAGTTTTATTACATGAATGTATTATCCAGCATTGA
AGTCTGGGCTTTTAGTGTAACCAGCACCTGAATAACATACATTGTACCCATTAAGTAATT
TCTCATCCCTCAAACCCCTCCCACCCTGAAATTAGACTTTGGATCCCTAGTTTAAATTCC
ACCCCTCTCTTTTTTTTGAGACAAGGTCTCACTCTGTCACCCAGGCTGGAGGGCAATGTTG
CAATGATAGCTTACTGTAGCCTCAACCTCCTGGGCTCAAGGGATACACCCTCCTCAGCCT
CCTGAGTAGCTGGAACTGCAGGCGTGCACCACCACATTCAGCTAATTTTTTGATTTTTTT
ATAGAGATGAGGTCGGAACTCCTGGGCTCAAGCGATTCTCCCCAAGTGCTGGGGTTACAC
ACATGGGCCACTGCCCCCAGCCTAAACCTCCTTTCTCAGTATAGCAGCCTTGAGATGAAG
TTCCTGAAATTACTGGCCAGCTTGACTGTTTCCCCACATCACTGGAGGAGGGGGATGCAT
AGATAAAACAAAATATTCAGCATCATTGTATTTTCTTTTTGTTTCATCAGCATCTTTTTT
TAAAACTCACTTGACATAAGTCCCTAGCCTCAAAGAGTAAAGCCTTTGCAGAATCTGCAT
TCAGATTTCGGGTGTGATTTCCTGACAGATAGTTCAGGTTTGTAAACTCTTTTTTTTTTC
TTTGAGACAGAGTTTCACTCTTGTAGCGCAGGCTGGAGTGCAGTGGCACCATCTTGCCTC
ACTGCAACTTCTGCCCCCTTGATTCACGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCATGCGCCACCACACCTGGGTAATTTTTGTATTTTTAGTAGAGATGGGGT
TTCACCATGTTGGCCAGGCTGGTTTTGAACTCCTGACTTCAGGTGATCTACCTGCCTCAG
CCTCCCAAAGTGATGGGATTACAGGTGTGAGCCACCGCAGCCGGCCAAAACTTTGTTTTT
TTTCCTCTTTTTGTTGCTGAGAAATGTAAACTCTTACAGACACAAATTATGTCTCCCATT
TTTTAAAACCCACTCAACACAGGGGTCATGTGTAATAGGCCCTGGAGCTTATTTTAGACA
TTGATTTGAGGCTCTTTTCCCCAAGTGCTGGTTTGTGTGTGTGTGTATGTGTGTGTAAGT
CTTTCTATGAGATGAGTGGTACCTACCTGGGCTGTGTGATCTTTTTTATTTTATTTATTT
TATTTTTGTAGATACGAGGTCTCACTATGTTGCTCAGGCTGGTCTTGAACTCTGGGGCTC
AACCTATCCTCCCTCCTTGGCCTCCTAGAGTGCTGAGATTACAGGTGTGAGCCACTGCAC
CTGGCCAGCGATCCTTAATAAATATAGATAATGGCCGGGCGTGGTGGCTCACACCTATAA
TACCAGTACTTTGAGGGGCCGAGGCTGGCAGGTCACCTGAGCTGAGGAGTTTGAGACCAG
CCTGGGTAACGTGGGTGAAACCCTGTCTCTACAGAAAATAGAAAAATTAGCCAGGTGTGG
TGGTGCATGCCTGTAGTCACAGCTACTTGGGAGGTTGAGACAGGAGAATTGCTTGAACCT
GGAAGGTGGAGGTTGCAGTGAGCCGAGATCGTGTCTTTGAACTCCAGCCTGGGTGACAGA
GTGAGACCTTGTCTCAAAAAAAAATATAGATATAGGCTGGGCGTGGTGGCTCACACCTGT
AATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCAGGAGGTCAGGAGATCGAGACCAT

CCTAGCTAACATGGTGAAACCCTGTCTCTACTAAAAATACAAACAATTAGCCAGGCCTGG
TGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCC
GGGAGGTGGAGGTTGCAGTGAGCCGAGACTGTGCCACTGCCCTCCAGCCTGGGCGACAGA
GCGAGACTCTGTCTCAAAAAAAAAAAATCTATATATCTATATATCTATATCTATATAGAT
ATAGATATAGATAATGCCAGATGATGGCTGGTTAGAAGGGATTGTCAGGGGCTGGCAGGT
TTTGCAGGTGTTAGAATGAGCAAGATGAGGAGAAGGATGCTTACTTCCCTCTCCTTGTAA
CTCTCTACCCCCTCCCCTCAGTGTTTTTTTATTTTTATTTTTATTTATTTATTTTTTTTG
AGACAAGGTCTTGCTCTGTCACCCACACTGGATTGCAGTGATGCAATCATAGCTCATTGA
AGCCCAAACTCCTGGGCTCAAGTGATCCTCTTGCCTCAGCCTCCCAAGTAACTGGGACCA
CAGGTGCGTACAACTATGCCCAGTTAAGTTTTTCATTTTTTATACAGACGGGGTCTTGCT
ATGCTGTCCAGGCTGGACTTGCACTTCTGGCTTCAAGTGATTCTCTTGCCTCAGTTTCCC
AAAGTGCTGGCATTATGGGCATAAGCCACTGTGCCTAGCCCATCAGTGTCTTTTTATCCT
TTACTCCTATCAAAATTCATTCACTCAGCAGCCATTGATCAAGTGCCTACTATATACATG
TTGAGGACTGGAAATTTATTTGTCTCTTCTCATCTTATCTGGACCCTCTGTGTTAATTGT
AATTAACTGTAATCATTCTGTATTAATTGTAATAAACTTGTTGATAAACTCAAATGAGGC
CATACCGTTTTGCCACTTCCCCTCCTTCCAGGTTATATGGATGTACTTACATTGCAGGTT
TCATTTGTTGGTTCAGTTTTTAAACTAAGCCCTATTGTGTCAAATTATGCTAGGTGTGAG
ATGGGGAGTTCAAGCTGTGTGTTGTCTTTTTTTTTTTTTTTTTTTTTGCCTCACTTACTA
ATATACAAGCGCTTATAACCTTTGAGGCTGGCCCTATACATTAAGATTTTTATTAATTCC
ACTGTTCTTTATCTTCTCTTACTAAGTTCTCAGGGTCGAATGAACTCTAACTGCTCCTTG
CTAGTGATAAGCAAGTTGCAAATTACAGAATTGTCAGTGATTGAATACACGTATTAAACC
TGTAACTGGGAAGCATTTTTGGTAATTATGAATACTTTTGGAAAAAAAAAAGCTATGGAA
GGAAAGTTTAAAATCTACGAAAGCTCAAGTAGATGGTCATGGAATAGCTATTTCAATTTC
TAACTATATATTACTTATTTATTTATTTATTTTTGAGACGGAGTTTAGCTCTTGTTGCCC
AGGCTGGAGTGTAATGGCGTGATCTCAGCTCACTGCAACCTCCACCTCCCGGGTTCAAGC
TATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTATAGACATGTGCCACCACGCCAGG
CTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCACATTGGTCAGGCTGGTCTCGAAC
TCCCAACCTCAGCTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGA
GCCACCGCGTCCGGCCTCTTAACTATTGTTTGAAATAATGTAGAGACAGCTCCAGAGCCA
TGAAGAAGTGTATGAAGAAGCAGTGTTAGCTTAAATGACATACATGTCACAATTGCCTAT
GTGAAACTATCATAATTATGCATGAGAAGTATCTATCCTGCATAACCTCCACCAATAATA
ATAATGTTAATAATAGTGAAAACTAATGTTTATTAAGTCCTTACTGTCTCCAGCCTCTGT
GCTAAATACTGGTTACTAAGTTTCCCTGAAAATACTATTCTCATCTGTTTGTTCTTAATA
ACAGGATAGCATAATTGTAAGTTGTAAATGAAATAATACAGTTTATGTAATAAAAGGGTA
AAAGAGAAGACCACCTACCTTATCTTCTGTTGCTGATCTGGATGGATGTAGGTGGTGTTT
ACCTAGTTTCACCTTTGGCAGTTGAAACTACTTTTTTTTTTTTTTTTTTTTTTTTAAGA
GACAGGGTGGGCCAGGCGCAGTGGCTCACGCCTGTAATCCCCGCACTTTGGGAGGCTGAG

GCGGACAGATCACTTGAGGTCAGAAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCT
GTCTCTACTAAAAATACAGAAAAATTAACTGGGTGTGGTGGTACACACCTGTAATTCCAG
CTACGTGGGAGGCTGAAGCAGGAGAATCGCTTGAACCCGGGAGTGGAGGTTGCAGTGAGC
TGAGATTGTGCCACTGCACTCCAGCCTGGGTGACAGAGCAGGACTCCGTCTCAAAAAAAA
AAACAACAACAAAAAAAGAAATTTTTAGAAATATGAGATGACAGCAAGAATGAGGGTATT
AAAAAGAAATTTTTAGAACTAAATAGCAGAATGTAATGGTGAAAAGTTTGATTTCTCAAG
TCTGCTTTGCACACAGGCATGTGGCAAACATTCAGTAAGTATAGCTGTAATTTTAACCAG
CTGTAATGTATAATAGCCAACATATCACATTTTTCTTTTTTCTTTTTTGAGACAGAGTCT
TGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCACCATCTCGGCTCACTGCAACCTCTGCCT
CCTGAGTTCAAGTGATTCTTGTGCCTCAGCCTCTCAAGTAGCTGGGATTACAGGTGTGTG
CCACCACACTCGGCTATTTTTTGCATTTTTAGTAGAGATGGGGCTGGTCTTGAACTCCCA
GCCTCAGGTGATCTGCCTGCCTCAGCCTCCCAAAGTGCTGAGATTACAGGTGTGAGCCAC
AGCGCCTGGCCATATATTGCTTTTTTCTTATTATCAGAGCCAGTTCATAATTGTGGAAAA
ATAGTGTTTGTAACAATGTAAGTATGGATAAATCATCTTTTTAATTTTGTGATTCATATA
GGTTTGTTGTTGTTGTTGTTTTGTTTTTATCTTGAGACAGAGTCTTGGTCTGTCACC
CAGGCTGGAGTGTAATGGCACAACCATGGCTCACTGCAGCCTCAGATGCCTGGGTTCAAG
CAATCCTCCCGTCTCAGCCTCTAGAGTAGATGGGACCACAGGTGTGGGCCACCATGCCTG
GGTAATTACAAAACTTTTTTTTTTTTTTCTAGAGATGAGGTCTCACTATGTTGCCCAGGC
TGGTCTCAAACCTTTGACCTCGCTTCAGCCTTTAGAGTAGCTATGACTATAGGCATGTGC
CATCACCCAGCTAATTAAAATTTTTTTCTTTTTTTTTTTGGTGGAGATGCGGTCTTACT
TTGTTACCCAGACTGCAAGTTAGTTTCAGATATCAACATTTGGTGTTTCCAAATGCACGG
GGAGGCTTTGGAGCAAGTTTTTGGCTCATATGCATAGGTGTCCTAGACATTCACTTTGCA
AATTCTTATTAAAATGACTACAGTAGCATACAGATAGGGAAAAATATCCTTGTCAGTACC
ACCGATTGGGTGAGAAGAGACTGTATATTAAAAACAATGACCATCTTTTTGCCACATAAA
TTGCTGGTGGGGCCAGTTTGAAGAGGGCTTTGTCAGCTGCCTTCTGCCTCTTCCTCTTGA
GTACGTGGAGTTGGAGTCATCCTTGACAGCCTCCTGTTGACACCACCCGGGTCACAGATG
TGAAACTGTGTGGATGTAGGAGAGAGCAGTGATGGGGCTTACCCCAAGGTTGCTCTTCCT
TCCCTCTGGCCACAAATGTTTAGTAAGGAACTGCTCTGTATTAACCATTTGCTAGGGGCT
GCAGATACGGTGGTGAAGAAATAGACATGTTCCTACTCGGGATGCTGAGGTGGGAGGATT
GCTTGAGCCCAGGAGTTGGAGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCTG
GGGGACAGAGCGAGACCCTATCTCTAAAAAACAATAAAAGAAATAGATGTGTCCTTCACC
CTCATGGAACTGCCAGTCTAGCCTTCAACCTGGTGACTGTAGAAATGTGTGATTAGATGC
TATATTGCCATGTTGAGTGTCACCCCTGAGAAGCAGGGTTTTTTTTGAGAAGGTAGGATG
GGGGATCTGACTGTGGGACCACCAGAGGGAAAAGCACATGTAAAAGCTGCGTGTACCAAC
TGGAGGAAATCGGAGACGTGATCAGAGAACCAGAGTCAACCAGGGGCCATGCCGTACAGG
GTCCTGTTAAGATCTGTGACTTTTTTCTAAACGTTTTCTTCTGGATAACATCTAAATTTC
TAGTTCCAAATGTGAAACTCCAAGGGCGTTCTGTGCTAAACATTTTGCATGTATTAATTA

ATTTCCACCACACAACATTGCTGTGAATTAAGACAGTTTCTAAGCATGGCAAGAAACCCA
GAAATCATAATGGAAAAATCTGATAAATTTAACAATGCCAACATGAACCTCTGTAGGAAA
AAAAATACCACAGACTAAAAAGGGGGGAAAAAAACCAGAGACAAATATTTGCAACACATA
CAGTAAAGGGTAATTTTCTGGTTATATCAAGAGCTCCTACAAATCAGTAAGAAAAAAAAT
CTAATAGGAAATGAGCAACGACAAACTGACAACTCATAGAAAAGGAAACACAAGTGGTCT
GAAAACATGAAAAAGTGCTCAGTCTCACAAAGAAATGCAAACTAACATGGTACCATTTTC
CATTAATCAGATAGACAAAGATGAAAGAGTTTGGTAATGTATGTAGTATTGGCACAAGTG
AGGGAAAACAGGGGATTTCACACTCTATGCCCGTCCAAACCAGTACCTTATTTTGAGGGT
GGTTTGACAATATTTGTCAAAATAAAAAAATTATATATAGTCATTTGCCACATAATGATG
GTTCAGTTGATGATGGACGGCATACATAATGGTGGTCCCATAAGAATATAATGGGCTGGG
TGCAGTGGCTCTCACCTGCAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATTGCCTG
AGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTGCTAAAAACA
TACAAACAATTAGCCAGGCATGGTGGCGGGTGCCTGTAATCCCAGCTACTCAGGAGGCAG
AGGCAGGAGAATCGCTTGAACCCGGAAGGCGGAGGTTGCAGTGAGGTGAGATTGGGCCAC
TGCACTCCCATCTAGATGACAAGGCAAAACTCCATCTCAAAAAAAAAAAAAAAAAAGAAT
ATTATGGGCCCAGCCACAGTGGCTCACACCTGTAATCCCAGTACTTTGGTAGGCCAAGGC
AGGAGAATCATTTGAACTCAGGAGTTTGAGACTAGTGGGGACAACATAGCAAGACCCCAT
CTCAAAAAAAAAAAGATTATGGTGGAGCTGTCCTGTATAGACATACCATTTTTAACTTTTT
TTTTTTTTGAGATGGAGTCTTGCTGTGTCACCCAGGCTGATGTGTAGTGGCGTGATCTGG
GCTTACTGAAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCTTCCTGAGTA
GCTGGGACTGCAGGCGCAGGACACCATATCTGGCTAATTTTTATATATTTAGTAGAGATG
GGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCCGCCTGCC
TCAGCCTCCCAAAGTGCTGGGATTACAGGCATTAGCCACCATTTACAGGCACCTGGCCAC
CATTTTTAATCTTTTATATTGTATTTAAACTGTACCTTTTCTATGTATGGATGTGTTTAG
ATACACAAATACCATTGTGTTACAGTTACTTACAGTATTCAGTACAGTAGCATGCTGTAC
AGGTGTGTAGCCTAGGAGCAATAGGTTATACCATATAGCCCAGGTGTGTAGTAGGCTCTG
CCATCTAGGTTTGTGTAAGTACGCTCCATGATGTTACCACAGTGACGAAATCGCCTAATG
ATGCATTTCTCAGAACATATTCCTGTTGTTAAGCAATGCATGACCGTATCTTGACAAAGC
CATTTTATTTCTAAAACTTTAATTTTACAGATTTATTTGTAAAAGTATGTAAAAATGATT
GTAAAGGATATGTTCTGCTGCATTATTTGTAATAACAAAAAACCAGAGGATAACATAAAT
GTCCTATAAGAAGGGTTAGATTATGGATGGCACATTCATACAATGGGGTATTATGTAGCC
ATTGAATAAAAGGGTACTGGCTGGGCGCAGTGGCTCATGCCTATAATCTCAACACTTTGG
GTGGCCAAAGAAGGAGGATTGCTTGAAGCCAGGAGCTTGGGGCCAGCCTGGGCAACATAG
CAAGACCCTATCTCTACAAAGGAAAAATAAAACAATTAGCCAGGTTTGGTATTGGACACC
TTCATGGTCCCAGCTACTGAGGAGGCTGAGATTGGAGGGATCGCTTGTGCCTGGCAGGTT
GAGGCTGTAGTGAGCCATGATTGTGCCACTGCACTCCAGGCTGGGAGATAGAGTGGGACC
CTATCTCAAAAAAACAAAAACAAAAACAAAACCTCCTGTAAAATGTCAAGAAGTCCTAGA

TGTGGGCCAGGTGTGGTGGCTCACACTTGTAATCCCTGCACTTTGGGAGGCTGAGGCCAG
GAGTTTGAGACCAGGCAGAGCAAGATAGCAAGACTCCATTTCTACAAAAAATAAAAAAAA
TTAGTTGGGCATAGTGGTGCATTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGTGGGAG
GATTGCTTGAGCCTGGGAGGTTGAGGCTGCAGTGAGCCATGATCACACCTCTGCACTCCA
ACCTGCGCAACAGAGTGAGACCCTGTCTCTAAAAACAACAACCAAAAAAACCCAGCAAAG
TACTGATAAAGATCTTTGGCTGGGCGCAGTGGCTCACACCTGTAATCCCAACACTTCAGG
AGGCTGAGGCGGGCAGGTCACAAGATCAAGAGATCAAGACCATCCTGGCCAACATGGTGA
AACCCGGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCGTGCACCTGTAGTC
TCTGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGTGGCAGAGGTTGCAG
TGAGCCGAGATCACGCCACTGCATTCCAGCCTGGCGACAGAGCAAGACTCCGTCTCAAAA
AAAAAAAAAAGAGAGAAAGATCTTCAAGTTGTAGTATGTGAAAAAATCAGGGTGTAAAAC
AAGAGAATCCCATTTGTGTGTGTGTCGAGTGTGTTTCACACAGGCTCAGAGGGAGTAGTG
TGTATATGCACATGAACATACGTGTCAGTGTATATATGTATATATACAAGGTTGTGGGTT
TGTTTGTTTTTTTTGAGACAGAGTCTTACTCTGTTGCCCAGGCTGGGGTGCAGTGGTGCA
ATCTTGACCCACTGCAACCTTCACCTCCCAGGTTCAAGTGATTCTTGTGCCTCAGCCTCC
CAAGTAGCTGAGACTACAGGCACGCACCACCATGCCCAGTTAATTTTTGTATTTTTAGTA
GAGATGGGGTTTCATCATGTTGCCCAGGCTGGTCTGGAACTCCTGGCCTCAAGTGCTCTG
CCCGCCTTGGCCTCCGAAAGTGCTGTTGCCCAGGCTGGAGCTCAGTGGCACAATCGCAGC
TCACTGCAACCCCGACGTCCCAGGCTCAGGCAATCTTTCCGTCTTAGCTTCCCAAGTAAC
TGGGACTACAGGTGTGTGCCATCAATGCCCCACCAATTTTTTAATTTTTTGTAGAGATGG
GGTTTCCCTACGTTGCCCAGGCTGATCTTGAACTCCTGGTCTCAAGCAATCCTCCCACCT
CAGCCTCCCAAAGTGCTGCGATTACAGGTGTGAGCCACCTTGCCCTGCCCTGTACAAAGA
TCTGCATAAAAGCAGTTAATAATACTATGTTTGAGGCTGCCATCACAGGGGTGAGGTCAA
GGACAAGTGTGAGAAATTCTTTTAGAATCTATTTTAAAAAAAGAAGAGATGACAGTGGTG
ACAGTCAGGGAACAGATAAGCAGGTAGATTGTGGGGGTCTAGGCTGTCTAACTGGTGTTT
AAAATGAAGCAACCGCTGAGCCTGCTGTATTTCATTTAATGGAGACTAGTAAAACAACAG
CCAGAAATTCTTCACTTTCCATCTAAGAGAGGCAAAAGTTATTTTCCCTTCAATAACCTG
GGACTGTAGGATTAAGGTTTTTTTTTTTTTTTTTTTAAATACTACAATATGACTACCAGT
ATAATTTAAAAATGATTAGAATTCTATTTGAGTAAGAAATAGGTGTCTGCCTGAAGTAGA
CAGTCACTGAAGTCACTAAGTGGCAAAAGACAGAAAAAAAATTGAAAGTAGGAAACAATC
AGCAGATATGATACCAAACATGAGCTGTCAGTGATAATGGATTAAGTCCTTCAATAATGG
CTGAGCCAGATGGAATTAAAAGAAAAAATCCAGGCCGGGCATGGTGGCTCACACCTGTAA
TCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATCACTTGAGTCCAGGAGTTTGAGACCAG
CCTGAACAACATAGTGGGACCCCATCTCTATTTTATAAAAATATTTTGAAAAAAGAAAAA
AAAATTCAGTTGTGTTCTGCTTTAAAAAGACAAATTGGCACAGAATGTCAAAGAATAAAT
AAAACAAACATGGGCAAAAGAGATTCAGGTGGTACCAATATCGGGCTAAGTAGCATTCAA
GATAAAGATTATTAAATAATAAGTTAGTTAATACTAGAGTAATTGCATATTAATGAAACA

TAATCTATGGTAGAGATATTATAGTCAATAATTGTTTTATGTATTCATTAAGGTAACAAC
AAGCAAACAAGCTTTAATAGTTTTAAATGCTTTATATGCTTTATAGTTCTTTTATGTGCA
TTAATTCATTAATTCTCATTTCCTATGAGGTAAACACTATTATTATCCACATTTTACAGA
TGTAAAAACCGAAGCAGAGAGATTAATTAGCTTGCCCAGGAGATGTGGCATTCTGGGATT
TGAGACAGTGGTTTGGCTCTGTAGGTTGCTTCAATAACCAAGAGATGCTTCAAATCAGAT
TTTTAAAATATGTTTTTCAGAAGCATTTTCCTGATACTTCTCCCCTTACATGGGTGTTAG
TCTTTTGGGTTGAAAAACATGAGTAAGTGCTAGAAGAGCAAAATATGCATCCAGATTTAA
TAGTATGTCTGTTTTTCTGAGCCTTGGCATTTCATTGCTTTTATAATAGAAATGAAGGCT
TTTTTTTTTTTTTGGCTGAGAATAGCACTGAACTCAGTGGGAGGGACTGTGGGTTGTAAG
TTGTCCGCCTCTGAATGGAGTTGAATTTAAGTTTCTTGGTTTCCAAAGAATGATTGATTT
AAAGACCCTCAAATTGCAAGTTAGAACTGACTTCAGTCCTTGAGGTTTTTTACCATTTAA
TGAATAATTAAATTTATGGTAATAAATGGTAATAAATGGTAAAAATGGTAATAAATTTTA
CCATTTAATGAATTTTTCTTAAAAAGCAATTGAATTGTTGATGAAAGGTGATGTTAAAAT
TATCCCAGATTTATCAATCTTTTTTTTATTGCCCCTGGATTTTGAGTCATAGAAAGCCTT
TCCTTATTCTAAGGTTAACAAGACATTCACCCATGTTTTCCTCTAGTATTGCATTGTTTC
ATCTTTTACGTTTATTATTTATTTTATTTTATTTTTTTGAGACAGGGTCTCACTGTGTCA
CTCAGGCTGGAGTGCAGTGGAATGATCTTGGCTCACTGCAGCCTCTGCCTCCCGCCTCCC
GGGTTCAAGCGATTCTGCTGCCTCGGCCTCCCAAGTAGCTGGGATTACAGGCACCTGCCA
CCGCGCCTGGCTAATTTTTGTATTTTTTTTTTAGTACAGATGGGGTTTTGCTGTTGGCCA
GGCTGGTCTCGAACTCCTGACCTTAAGTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGG
GATTACAGGCATGAGCCACCGTGCCCGGCCTAAAATTTATTCTGATATGTGATATGATGT
ATGGTTCTAACTACTTTGTTACGGTGCATTATTTTCTAAATGTGGTATTGGATTCTTTTA
TATTTTGTTTAGAAGTTCTGCATCAATATTCATGAGTACCATTGGTCTCTGTTGTTTTTC
TTGTGCCATCTTTATTGGTATAGGTATCAGTGTTATATTTAGTTTGTAAAAGGAAGTTGG
AAGTTTTCCTTTCTTTTTAGTACTCAGGAATGATTTTAAGAATTGAGACTATTTGGTCTT
TGAAGGTTTGGTAGAAGTCCATTGGGAATCCATCTGGGCCTGGTGATTTTCTGTGCGGTA
GTTCCTTAATTGTTTTCCCTATTTTTTCTTATTTTTAATCAGGTAGCCTCTGAACCAGAA
TAGGTTCAGAGAGGCTCCCTCTATTTTTTTTAATACAAGTTGGTCTGCCTAAGTTTTCTT
ACTCTAATGGGTTAATTTTTGTAGACTGCATTTCCCTGAAAAATTACACGTTTGTTCTAG
GTTTTCTGACTTATTTCCACAACTTTTTAGTCTTTCCCCCTGGAATCATGCCCCTTTCCA
TAAACAGGACTCTGATGTACCTGAAGTATTTTCACACTTCGGGTGGACTTTCTGTTTCTG
GGGGTGGTTTTAGAGCAATTTTAGGCCTGCCACTAGCTACCCTGTTCTCTACACCATGCT
GTTTTTCTCAGAATGCTCTTCTTTTGCACAAAGGCTTGGAGTAGGAGGTTGAGCAGTCAC
TCACTGACGTTTGGTATATTTTCTTTTTTTTGCTTACAGGTAATCTGGAAGTTTGGGCAT
TCTCTTTAAGTTGAGGGTGTGGTTTTCATGTCATTTTATTTGTTTATTGTTTTCTTGTGT
GTGTTTCTTAGAGACAGGGTCCCACTCTTGCCCTGGCTGGAGTGCAGTGGCGTCTTGATC
ATAGCTTACTGCATCCTCAAGCTGCTGGGCTTAGATGAACCTCCCACCTCAGCCTCCTGA

GTAGCTGGGACTACAGGAGCACACCACCATACCTAATTTTTTTTTTTTTTGAGACGAAGTC
TTGCTCTGTCCCCCAGATTGGAGTGTAGTGGTGCAATCTCGGCTCACTGCAACCTCTGCC
TCCCGGGTTCAAGCGATTCTCTCACCTCAGCCTCCCGAGTAGCTGAGACTGCAGGTGCAT
GCCACCATACCCGGCTAATTTTTGTATTTTTTAGTAGAAACAGGGTTTCACCATGTTGGC
TAGGCTGGTCTCAAACTCTTGACCTCAAGTGATCCACCCACCTTGGCCTCCCAAAGTGCT
GGGATTACAGGCTTGAGCCACTGTGCCTGGTCCCTGGCTAATTTTTAATTTTTTTGTAGA
GATGGGATCTTGCTATGTTGCCCAGGCTGGTCTTGAACACCTGGCCTTAAGCAATCCTCC
CACCCTAGCCTGCCAAAACACTGGGATTTACAGGCATGAACCATTGTGCCTGGCTTGTTT
TGTTTTTAATTCTATGTTGTTTTTGAAGGATGTATGGGGAGAGATGGATTTAGGCAATCA
TCGTTGTCCTTGGCTACCTGAAAGTCCAGGCACTCTTCTAGATACTTTATAAATATTAAC
TCATTTTATCCTCTCAACAACACTATGACATGGGTACTGTTACACCTTCCATTTTATAGG
ACTTAACAGAGAGGTTAAATATGTAGCCCAGGGTCACAGAGAGCTGGGCTTCAGACCAAG
ACAATCTGGCACCAGAGTCTATGTGGCTACCCCTAAGGCTTTGCCACCATGTGTTAGTGA
TTCTCAGCCTGTCATTTGGGGAGGGGATTGCCCTTTTTTTTAAACTTTTTAAAAAATTTA
TTCTTATTTTATTATATTTTTGAGACAGAGTCTCCCTCTTTTGCCGAGGCTGGAGTGGAG
TGGTGTGATTTCAGCTCACTGTAACCTCTGCCTCTGGGGTTCAAGTGATTCTCATGCCTC
AGCCTCCCAAGTAGCTGGGATTACAGTTGCCAGCCACCATGCCCAGCTAATTTTTGTATT
ATTATTATTATTATTTGAGACGGAGTCTCGCTCTTTTGTTCAGGCTGGAGTGCAGTGCTG
TGATCTCGGCTCTCTGTAACCTTCGTCTCCTGGGTTCAGGTGATTCTCCTGCCTCAGCCT
CCGGAGTAGCTGGGACTATAGGCGCGCACCACCATACTTGGCTAATTTTTTGTATTTTTA
GTAGAGACGGGGTTTCACTATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGAT
CTACCTGCCTTGGCCTTCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATGCATGGCTG
GATTGTCCTTTTTTAAAAAAAAAAACAAAAACAAAAAAAAAAACCCAAACCATAAACCCA
ATATTCTGAAAGATTTGGTCTCCACACCTGTGTTATATAATAATTAGTTTTTCCATTTTT
TTCCTCTTGGTAGAAGGCACATATGCCACTCAGTTTCCAGTTGCCACACCCAATTAACAT
AATTGTTTTGCAGCCAAAAGCAAAAGAGAGTTGACATTTTAATTAGCTTATGTAGGTAGA
CAAATTGAGGCCTAATGTAAGAGTTTCATTATACCTTTTTGAAAAACTATAAATAGCTAG
AAGCCAGTTGTCATTACTTTTTGATTCCTTAGAATTCTGGGCATCTTTCATCTGGAACCA
CAGATGAAAGAAGCTGCAAGGAAGGATTTTTTTTCTTAACGGAATAGTTTAACCATTCTG
AATGCAAAAGTATTGGATGCTAGAATAATAGGTATCACATAAATTGAGGTTGACGTTTTC
CCGGGTGAAATTCTATTCTGTCTCAATTTTCCTTTTTTTTTTGAGACGGAATCTTGCTCTG
TCGCCCAGGCTGGAGTGCAGTGGCATGATCTCGGCTCACTGCAAGCTCCACCTCCTGGGT
TCATGCCATTTTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGGGCCTGCCACAAC
ACCCAGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCCCAGGATGGTCTCAATCT
CCTGACCTCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCCGGGATTACAGGCGTGAGCC
ACTGTGCCTGGCCTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGACAGAGTCTCGCTTTG
TTGCCTAGGCTGGAGCGCAGTGGCATGATCTCAGCTTATTGCAACCTCCGCCTCCCGGGT

TCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTATCTGAGATTACAGATGTGTGCCACCAT
GCCTGGCTAATTTTTGTATTTTTAGTACAGATGAGGTTTTGCCATGTTGCCCAGGCTGGC
CTCAAACTCCTGACCTCAGGTAATCCTCCTGCCTCAGCTCTTCCCAAAGTGCTGGGATTA
TAGGCATGAGTCACCGGGCCCAGACTCAATCTTCTGACAAGCTCTCAGAGAGAGTAAAAA
GCAAATGAATATTTCATTATTTTGATCTGAGCTTTACGATTTTTCTTTTCTTTTCTTTTT
TTTTTTTTTTTGAGATGGAGTTTTGCGTTGTTGCCCAGGCTAGAGTGCAGTGGTGGCGAT
CTTGGCTCACCGCACCCTCCGCTTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCTG
AGTAACTGGGATTACAGGCATGCGCCACCATGCCCGGCTGATTTTGTATTTTTAGTAGGG
ACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTTAAGCTCCCGACCTCAGGTGATCCACCT
GCCTCGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCACCTTGCCCAGCCTTTTTTT
TTTAAATCTGAGAAGAGGTCTTGCTCGATTGCCTAGGCTGGAGTGCAGTGGTGCGATCTC
TGCTCACTGCATTCTCTGCCTCCCAGACTCAAGCAATCCTCCCACCTTAGCCTCCTGAGT
AGCTGGGACTACAGGCATATGCCACCACACCTGGCTAATGTTCGTATTTTTTTGTAGAGA
CAGGGTTTTGCCATTTTGCCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCTCCCA
CCTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACTGTGCCTGGTCTCCTTCAC
TGTTGTAAGATACTTGAATTGGGTCAATATTTGTGGAGAAGTCTCTTAAAAGTTCACTTG
ATTGTCAGTACTAGAACTCTACATTTAATATTGACATATTCCTGGGAGCATTTCAGAGCA
TTCTATTAGCTTAGAAAGGTCCAGGATAATTTGACTTTAGAAGTTACTGTTACCATGAAT
CTCAATGACTTTTGAAATCCATGAAGAATATCTTTTTTTTTTTTTTGAGACGGAGTCTCA
CTCTGTCGCCCAGGCTGGAGTGCAGTGGTGATCTGGGCTCACTGCAAGCTCCGCCTACTG
GGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACATGCCAC
CACGCCTGGCTAATTTTTTTGCATTTTTAGTAGAGAGGGGGTTTCACTGTGTTAGCCAGG
ATGGTCTCGATCTCCTGACCTTGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGCGCCTGCCCAAGAATATCTTTTTGCTGGTAACTAGAGAGGACT
CCTCTGAAGCAGATGCCATTCATGATGGATTTCATCATTTATGGGTTTTAAAAAACATTT
TATTTTGAAATAATTTCAAATTTAAATAAGAGTTGCAAAATAGTACAAATAATTCGTGTT
AACTTTTCATCCAGATTTACAAGTCAACCTTATACAGGTTGAGTATCCCTTATCCAAAAT
GCTTGGGACCAGAAGTGTTTTGGATTTCAGATTTTTTCGAATTTTGGAATATTTTTATTA
TATACTTAAGCATCTCTAATCCCCAAATCTCAAATCTGAAATATCTGAAATGCTATGATG
AGCATTTCCTTTGAGTGTTATGTGGGCACTTTTTAAATTTATTTAATTAATTTATTTTTT
GAGATGGAGTATTGCTCCATCACCCAGGCTGGAGTGCAGTGAGCGATCTTGGCTTATTGC
AAACTTCACCTTCTGGGTTCAAGTGATTCTCCTGCCTCAGCCCCCTGAGTAGTTGGGACT
ATAGGCGCTTGCCACCACGGCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCAC
CGTGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGGTCCACCTGCCTCCGCCTCC
CAAAGTGCTGGGATTACAGGAGTGAACCACCGCGCCTGGCCATGGATTTTGCAGCATTTT
AGATTTGGGATACTCAACCTGTACCATGTTTACTCTCTCTCCTCTCTCTCTCTCTTTT
TATATATATATATATATATATATATATATATATATATATATATATATATAAATTATATATAC

ACTACACATATATGTATGTATATGTATGTATTTTATATATAAAATACATATCTACATATA
AAATACACATGTATATATACATGTGTACATATATGTGTCTCTATATTTAAGTTTTGTTGG
AACCACTTGAGGGTAAGTTGCAGACATGGCGTCTCATTGCTCCAAAATACTTCAGTGTGT
ATTTCTTAAATACAAGGACACTTGGTTACATAACCACAGTATATCACCAAATGTATATTA
TAACAAGACTACCATCAAATCCTTATATCTCTTTCAAATTGTTTTAGTAATATCCTTATA
GCAAAAGACAAAACAACAACAAAAACTGTTCCCTTTTATTTTGTTTGTTTTGGTCCATTA
TATGTCCAGGTTATGCATTAATGCATTGTGTTACTTGCTAAGTCTTGTTACTGGCCTTTA
ATTAGGATATTTCTTTGCATCCCGCCAAACTCCTCTTCATGGTTGTATCTTTTTTTTTTT
TTTTGGAGATGGAATTTTGCTTATGTTGCCCAGGCTGGAGTATAATGATGCGATCTTGGC
TCACTGCAACCTCCGTCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAAC
TGGGATTGCAGGCCTGCGCCACCTTGCCCAGCTAATTTTGGAATTTTGTGAGACGGGGTT
TTGCCATGTTGGTCAGACTAGTCTCGAACTCCTGACCTCATGATCCGCCCGCCTTGGCCT
CCCAAACTGTTGGGATTACAGGTGTGAGCCACTGTGCCCGGTCTTTTTTTTTTTTTTTTT
GAGACAGGGTCTTATTCTGTTGCCTGGCCTGGAGTGCAGTGGTATGATCTTGGCTCACTG
CAACCTGGACCTCCTGGGCTCAGGCGATCCTCCCACCTCAGCCTCCTTAGTAGCTGGGAC
TATAGGCACACACCACCATGCATGGCTAATTTTTATATTTTTTTGTAGAGACTGGGTTTC
GCCATGTTGCCCAAGCTGGTCTTGAACTCCTGGGCTCAAGTGATCCACCTGCCTTGGCCT
CCCAAAATGCTAGGATTACAGGTGTAAGCCACTGCGCCTGGCCCTAATTTTTGCATTTTT
TGTAGAGATGGGGTTTCACTATATTGCCCAGGCTGGTCTTGAACTCCTGGGCTCAAGTGA
TCTTCCCATCACAGCCCCCTAAAGTGCTGGGATTATAGGCGTGAACCACTGTGCCTGGCT
GAGGATTAAGTTTCAACCTCAGGGGAGCGGCATTCAAACTATAGCATTGTCCTTTAGTGA
CTGGCTTAGTTCACTTAGAATGTTTGTCTATTCATCCATCTATAGACACTGTTTTCTTTC
ACCTTTTGGCTTTGCAAATAATGCTGCTGTGAATATGAGTTATAGAAAAATACCAATTTG
AATCCGTGTTTTCAATTACTTTGAGTATATACCTGGAAGTGGAATTTCTGGATCATATGG
TACTTCCAAGTTTTTTTTTTTCTTTTTTGAGACAAGGTCTCACTCTGTCACCCAGGCTG
GAGTGTAGTGGCACGATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCT
CCTGCCTCAGCCTCTCAAGTAGCTGGGATTACAGGCACGCGCCACCACGCCCAACTAATT
TTGTATTTTTAGTAGAGATGGGTTTCTCCATGTTGGTCAGGCTGCTCCCGAACTCCCGAC
CTCAGGTGATCTGCCTGCCTCAGCCTCCCAAAATTCTGGGATTACAGGTGTGAGCCACCG
CACCTGGCCTCCATGTTTCAATTTTTAAACAAACAATTAGTTAAAAAAATAGGAAACTAA
GAGAATGAACTATTTCCTGTTTTATTCAGTGGGTTATAATCTGTTACTATCATTGTTTAT
TTTGAGGTACAAATTGTCCCTACTTTGGCCAGCAGAGGATCCTGCAGTTTGTCTCCTGTG
TCCTTTTCATAGCTCCTTGTTGGAACTCTTACTGGCCCACAATAGGATGTTCCAAGTTCA
TCTTCTTACTTTTACTGCCCCAACGCTGGGATCAGCCATTTCTTCAAGGAGGCCAGTTCC
TTTCATTGGAGAATGGAAAACCCAATATGTAGAAACCAAGATAGAGGTGTTAGGTGTGAT
TGCTACTGGAGTGTCATTGCTTCCAAACCCTTTCAGAAGAGACCTAGGAAATGTGTGTGT
GTGTGTATATATATATGTGTGTGTGTGTGTGTATTCATAAAAGCACATACACATACACAT

ACCCCGAAGCATGTATTTCTGTATTATTATTATTTTTTTGAGATGGAGTCTTGCTCTGTC
GCCCAGGCTGGAGTACAGTGGCACGATCATGGCTCACTGCAACCTCTGCCTCCTGGATTC
AAGCAATTCTCCTGTCTCAGCCTCCTGAGTAGCTGGGATTACAGGTGTCCACCACCACGC
CCACCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCACATTGGCCAGGATGGTCT
TGAACTCCTGACGTCAAGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTATAGG
CGTGAGCCACTGTTCCCATCCAGAAGCATACATATCTATTTCTATATCTACATTTCTGTC
TTTACATGTATATATTAAAAATTACAGTTTGCACTAATACCTCCAATTACAATCTAACAT
CATGGGATTTATTCTGGCTTTCTCCCTTCTCATATTTGTGTCTCCCCAACAGTGAGAAAC
CTGGCTTGCTATCCTCAACATGGTAACTTATTTATTAAGAAACTTATTCTTTTTTTTTTT
TTTTTTCTGAGATTGAGTTTCGCTCTTGTTGCCCAAGCTGGAGTGCAGTGGTGTGATCTT
GGCTCACCGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCTTCTCAAGT
AGCTGGGATTACAGGCATGCACCACCATGCCCAGCTAATTTCGTATTTTTAGTAGAGATG
GGTTTCTCCATGTTGGTCAGGCTGCTCTGGAACTCCCGACCCCAGCTGATCTGCCTGCCT
CGGCCTCCCAAAGTCCTGGGATTACAGGCGTGAGCCACCGTGCCCTGCCTCTAGTTTATT
TATTTTTATTCCATGTGCTCAGTCTTGCGAGCACGTGGTCTGTTTTCTTGGGCCTGGCCC
CCTCAGTGCACTGTCTTAATACCCTAGCCCCCAGTCCCTCTGATCATATCCCCAGACACC
CCTACTGAATCCCAGGTCTCTACCAAGGGAAAGGCAGGGAGGAGGCATTGACCAAGGAGA
AGAGGGGGAAGGGACAGGGAAGGTCTTGATTTGTATTTTCTAAAATTTTCTACTCTGCTC
ATAATGCGTCTTAGCTGTGTTGTTGTGGAAAGTAGTGCTGACAGTGTCTTGTTTTTTTAT
TACTTACTTTGTCTTTCTTTTTAAG**ATGGTTTCACCCAAATATCACTGGTGTGGAGGCAG**
AAAACCTACTGTTGACAAGAGGAGTTGATGGCAGTTTTTTGGCAAGGCCTAGTAAAAGTA
ACCCTGGAGACTTCACACTTTCCGTTAGGTAAGTTGGAATGAAAAGAGAGGATCCTGAGA
GTGTTTTCTAGGTAGGAAGTGGTAAAACCATGCTTGGATAGCTTGCTGCCTGCATTTCGA
GTTTGAAGGCCTTATCTGAGCCCTGGGCTGCCTTCAGGGTTTGGGGAGTGGCCTCCTGGA
CATTTAGCAGAAGAGGAGTAAGGAGGGCCCTTCTTCTCCCTCTGAGACCTCATGGAAGGT
GAGTTGGAGCAGGTCATAGAAGTTCTTAAGCCCTCCAGTGCTTGAGACTTGTTCCACACA
TCTTGAACCTGGTTTCTGCATTTTTCTTTTCCTTCCTGTTGATTTATTTAAAAATTTTAT
TTCTTTTCAATTTTTTTTTTTTTTTTAAATAGAGGTGGGATCTTCCAATGTTGGCCAGGTT
GGCCTTGAACTTCTGGCCTCAAGCAATCCTGCCTCGGCCTCCCAAAGTGTTAGGATTACA
GGCGTGAGCCACTATGCCTGGCCTTCTTTTTTTGAGACAAGCTGTTGCTCTGTTGCCCAG
GCTGGAGTGCAGTGGTACGATCACAGCTTACAGCAGCCTTGAACTCCTGGGCTTAAGTGA
TCCTCCCGCCTCAGCCTCCCGGGTAGCTGGGACTCCAGGCTTGTGCCACCATGCTCAGCA
TTTTTAAAAAATATTTTTTGTAGAGATGAGGTCTCACTGTATTACCAAGGCTGATCTTTA
ACTCTTAGCCTCAAGTGATCCTCCTGCCTCAGCCTCCCAAAGTGTTGGGATTACAGGCAT
GAGCCACCACACTCAGACTTTGTTGACTTCTTAATAAGAAAAATACTTGTTAAGAGTTTC
TTCAGATCACTTTCCTTTATCAACAAGTAAAACATGACTGAGGAAGTTGTGGTCCCCTTT
GCTTCCCTGCCCAGGCCCGTTTCCCTCCCTCTTTCCCCAGAGGAAACCACCAAGAGGTTG

GCATATATTCTTCCTGAACGTGTTTTTATAGTTGTACTGCACTTGTACTGTGTATGAACA
ATATAAAGTTGGTTTGTGTGTTTAAAAAATTCACATACATGGATTTATAATGTATGTATC
ATTTTGCAACTTAAAAATTTTTTTTTGAGCTCCATGCTGATTGATAACGATCTATTTTTT
TTTTTTGAGATGGAGTTTCAGTCTTATTGCCCAGGCTGAAGTGCAATGGCGTGATCTCAG
CTCACTGCAACCTCAGCCTCCTGGGTTCAAGCTATTCTCCTGTCTCAGCCTCCGGAGTGG
CTGGGATTACAGGTGCATGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAGATGG
GGTTTCACCATGTCGACCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCTGCCTGCCT
TGGCCTCCCAAAGTGCTGGAATTACAGGCATGAGCTACCATGCCTGGCCTTTTTTTTTTT
TTTTTTTTGAGACAAAGTCTTGCTCTTTTTCCCAGGCTGGAGTGCAGTGGCCACAATCTT
GGCTCACTGCAACCTCTGCCTCCTGAGTTCAAGCAGTTCTCCTGCCTCAGCCTCCTGAGT
AGCTGGGATTACAGACATGTACCACCATGCCAAGTTAATTTTTGTATTTTTTGTAGAGAC
TAGGTTTTACCATGTTGGCCAGGCTGGTCCTGAACTCCTGACTTAAAGTGATCCATCTGC
CTTGGCTTCCCAAAGTGCTGGGGTTACAGGCATGAGCTATCGCGCCTGGCCTGAGAAATC
TCATTCTTACTCCTACTCCCTTGCACACTATCTCCATTCTGTAGGTAGCCATTTCTATTA
ATTTCTTGTTTACCCTTCTGTGTTTCTTTCATTCTTTTTCTTTTTTTCTTTTTTTTTTTT
GAGACAATCTTGCTCTGTTGCCCAGACTGGAGTGCAGTGGTGTGATCTTGGCTCACCGCA
ACCTCCACCTCCTGGGTTCAAGTGATTTTCATGACTCAGCCACCTAAGTAGTTGGGATTA
CAGCGCCTGGTGTACACTACCACACCCAGCTAATTTGTGTATTTTTAGTAGAGATGGGGT
TTCACCATGTTGTCCAGGCTAATCTCCAACTCTTGGCCTCAAGGGATCTGCCTGTCTCAG
CCTCCCAAAGTGCTGGGATTATAGGCATGAGCCACCATGCCTGGCCCTATGTTTCTTTTT
ATAAAAATAAGCAAATTAATATTTTTATTACTATTTTCCTTTTATTTTTACACATCAAGT
AGAACATTAAATATATTTCTCTGTAATTTTTTTCAGTTACCTAAATCTTTTAGTGATCTC
TCTCATCTTTTTAATCAGCTGGATCGCATTCTATCATGTGAATATTTTATAACTTCTATA
TACTGTCACCAGCAGGTAGCGATTTAGTTGTGTCTAATATTTTAAAATGATATATAATGC
CTCAATGAATATAGTAACCTTTTGCATATATTGTTTTGTGCTTTGGGATAACACTACCTC
GTATTGGAAACTGTGTCATTACATGTGTCTTTAAAATTACATGTGTCTTTTTATTTTTAT
TTTTATTTTTTTTGAGTGGGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAGTGGTGAG
ATCTCGGCCGACTGCAACTTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
CCAGTAGGTGAGATTACAGGTGCCTGCCACCACGCCCAGCTAATTTTTGTATTTTTAGTA
GGGACGGGGTTTCACCATGTTGGCCAGGCTGGTATCGGTCTGCTGACCTCAGGTGATCCT
CCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGACGTGAGCCACCATGCCTGGCCATCA
CTTTTTTTTTTTTTCTTAATTGCTGCATAGTGGCCGGGCACAGTGGCTCACGCCTGTAATC
CCAGCACTTTGGGAGGCCAAGGCAGGCGGCGGATCATGAGGTCAGGAGACCAATACCATC
CTGGCTAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAAATTTAGCTGGGCGTC
GTGGCGGGCGCCTGTAGTCCCAGCTACTTGGGAGGTTGAGGCAGGAGAATGGTGTGAACC
CGGGACGTGGAGCTTGCAGTGAGCCAAGATTGCACCACTGCACTCCAGCCTGGGTGATGG
AGTGAGACTCTGTCTCAAAAACAAACAAACAAACAAAAAAATTGCTGCATAGTATTCCAT

TGTATGAGTAGTAACACAACAATTTTTATAATGCATAGTATTCCATTGTATGAATAGTAA
TGTAGCACTATTTGTTTATACATTTTTATGATTAAAAAACAAAATGTTTTTCTATTATGA
ATAAAGTGGCAATGAATATTTTTGTACAAGTGTTTTGGTAGCTATACAGTTATTGTCACT
TAATATATGCAATTCGATAGGCCAGTCATTCAAAATAGAAGATATACAAGGTAGGCCGGG
CGTGGTGGCTCACGCCTGTAATCTCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTG
TGGTTAGGAGTTTCAGACCAGCCTGACCAACATGGAGAAACCTCATCTCTACTAAAAATA
CAAAAGTAGCTGAGCGTGGTGGCGCATTCCTGTAATCCCAGCTTCTTGGGAGGCTGAGGT
AGGAGAATCACTTGAACCTGGATTTATAATGTATGTAAATCCACCGCGAAGGTTGCGGTG
AACCGAGATCACGTCATTGCACTCCAGCCTGGGCAATAAGAGCGAAACTCCATCTCAAAA
AAAAAAAAAAAAGATATGCAAGGTAAAGATACTAATAAAGACCTTTGTGTTGAGTTGGTT
GACATGTGGTTATTTCACCCATCGTATTTCTTATAGGGAATAGGTAAATTCGTTCCTTGG
GTTTCTTTCAACACTTAGGTAAAATCCGACGTGGAAGATGAGATCTGATTTTACTGGTGT
AACTCTTTATTTGTCCCCTTGCCTCCCTTTCCAATGGACTATTTTAG**AAGAAATGGAGCT**
GTCACCCACATCAAGATTCAGAACACTGGTGATTACTATGACCTGTATGGAGGGGAGAAA
TTTGCCACTTTGGCTGAGTTGGTCCAGTATTACATGGAACATCACGGGCAATTAAAAGAG
AAGAATGGAGATGTCATTGAGCTTAAATATCCTCTGAACTGTGCAGATCCTACCTCTGAA
AGGTCAGTAACATTTTAGTGACCACAAAGTCTGCTGCTCCCTTGTGCCCTGAGTGTCAGA
AATGCATGACGGTCTGTGTATGACTCTCTGACTCCAAAGGCTTGTGACTGTTTTTTGAGC
TGTAATCTTTAAAGAATTACTAAAGTGAGACTAATAGCATCAAATTATTTTCAGAGTACC
TTTTTCCTGCAAAAGTTTTAATCAGTGTTACTTACACTCATCCTATAGGGGTTGCATACC
ATTCCTGCATATACTTGGTACGTGTATTAGTTTTAAGACTTATTGAACTTCAGCAGATAA
TCTTTGAGAGTTATTAGAGGAAAACAAATGATAATGGAGACACCAAAATAGCAGCAGTTT
TCTATGGTGGCTCTCGACCAGTTATTCAGCAATGTCACCAACAGATGTCAGTTTAAGCTC
AGAAGTGGAAAAGCAGAGAGCTCAGAGGGTCAGCTTTTTCATCAGTTCTTTTAATGTTAT
CACCACAATTATGTGAGAATGACCTTGCTTAGAGAAAATTATGTTATTTTCGAGATCTTT
CCCCCTGTGTTGGAACTAGGCTGATGAAAGCATGGGCTTGACTTATTTATTGATTGTATT
CGTTTTGTACATTCCCAATCTCCTCTCTGACTTGGTGCAAATTCAGGATCTCTTAGTTAG
TTTGTATATTTTGTGTCTTCAGGTATGATTTTTTCAGCTTATACCTTTATGTCAGTGCTA
TTATGTGCTGATAATTTGTTTCTCTAGCTACCACCGTAGCTTCAGGCAAAAGGCTGTCAG
CCAACTCTGTACAGTTTATTTCTAAATTTTACTGTTTTCAGTTGAGTATGGATGAAGAAT
AACTCAAAGTTTATTCTTTTGATGATGAGCCCTTAACACCACCTGCCATGATAGTACTTG
CTTTCTGACCAAGATCCTGAGGGAAAAAGCCACTTTATTATTAGAACTATGTTAAGATGC
TTCCCAAAAAACATGGAGCAGTATTGTCTCAAAGTCTGTCCTTGGATGGCTTTGGATGCC
TACATCAGGACTGTCTGATGTGCTGGTTAAAATGCAGATTCCTGGGCCTCATTCAGACTT
ACATGTATTGATATTGCTGGTTGTGGAGCCTGGGAATTCATATTTTTAGCAAAATCCCTC
ATTTTTACTCCAAGTCTTATGTGCATTATACAGTTTGAGATGATCACCCAGGATATAGTC
CAAAGACACTGGAGGCTGTTGAAGTATAGGTTGTATATATGGAAAAGGTTGGAATGTTTG

AATTAATTTATAATGAAGATCCTTTTTAATTGAGTGTTCACATGCCAAGGCAAGGACAAA
CATTCAAAATGATTTTCTGTCTCTGTTACAACTTTTTCTTTCTTTTTTTTAATTTATTTA
TTTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTCAAGTGACGCGATCTCGGCTCA
CTACAACCTCCGCCTCCCAGATTCAAGTAATTCTCTTGCCTCAGCCTCCCGAGTAGCTGG
GACTACAGGCATGTGCCACCATGCCCAGTTAATTTTTGTATTTTTAGTAGAGACAGGGTT
TTGTCATGTTTGCCAGGCTGGTCTCAAACTCCTGAACTCAGGTGATCCGCCCACCTTGAC
CTCTCAAAGTGCTGGGATTATAGGCGTGAGCCACCGTGCCTGTCTCTATTACAACTTTTT
ATTACAACTTCTTTATTTTGACTTTATTTTTACAAATTATTTATTTATTTTTTTTGAGAT
GGAGTTTCGCTCGTCACCCAGGCTGGAGTGCAATGGTGCGATCTCAGCTCACTGCAACCT
CCGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGG
CACTTGCCACCACACCCGGCCAATTTTGTATTTTTAGCAGAGACAGGGTTTCACCATGTT
GGTCAGGCTGGTCTCGAATTCTTGACCTCAGGTGATCCACCTGCCTCGGCCTCCCAAAGT
GTTGGGATTACAGGCATGAGCCACCACGTCCGGCCGACTTTTATTTTTTTTTCTTGAGAC
AGGGTCTTGCTCTGTCACCCAAGCTGGAGTGCGGTGGCATGATCATAGCGCACTGCAGCC
TCGACCTCCTGGACTCAAGTGATCCTCCTGCCTCGGCCTTGTGTATAGCTGGGATTACAG
GCAGTTGCCACCATGCCAGGCTAATTTTTAATTGTTTTGTGAAGATGGGGATTTCACTGT
GTTGCCCAGACTGGTCTTGAACTCCTGGCCTCAAGTGATCTTCCTGCCTTGGCCTTCCAA
AGTGTTGGGATTACAGGCATAAGCCACTATGCATGGCCTGTAACTTCTTTAAATGGCTAT
AATTAAACAGTTGGTCCTTTTAAGATTGGGCAATGGACGAATGGCAAATTGCATTTTTAA
AAGAGGAGGGATTTAAAAAAAAACAGGAAAGATTGGGGCATTTGTCTCTAAAGGACTGTG
GACTCATTTAAGAAGTTTAGTGGTCATTCTTACCATCTTTGTGGTTTTTCCTGCCTGCAT
GGGATGCAGATTTTCTGTCTCAGGTGGGATTGATCAATCCCTTGGAGGAATGTGTCTACT
TTTTAATTGTGTTTAGGAGAGCTGACTGTATACAGTAGTTTTGTGAAAGAACAACATGAA
CCCATAGTAGAGCTAAATTCTTTTTTATTTTTTAAAAACTTTAG**GTGGTTTCATGGACAT**
CTCTCTGGGAAAGAAGCAGAGAAATTATTAACTGAAAAAGGAAAACATGGTAGTTTTCTT
GTACGAGAGAGCCAGAGCCACCCTGGAGATTTTGTTCTTTCTGTGCGCACTGGTGATGAC
AAAGGGGAGAGCAATGACGGCAAGTCTAAAGTGACCCATGTTATGATTCGCTGTCAGGTA
AATCTCCAGTTGAAAAATGGGTCTGGCAAGATGTTACCTTTGGGTGATTTTTCTGCTGAC
AGAAGACAGACACCATTACATTCAAAGTCAGATTGTCTTTTATTTATTTATTTATTTATT
TATTTATTTGAGACAGGGTCTTGCTCTATCACCTACAGATGGGGTTTCACCACGTTGGGT
CTGGTGACCCAAATCTTTGGGTGATTTTTCTGCTGGAAGAGGACAAACACCATTACATTC
AAAGTCAGATTTTCTGTTTTTTTTTTTTTTTTGTTTTTGTTTTTTTAATATTCATTTGTT
TATTCATTTGAGACTGGGTCTTGCTCTGTCACGCAGGCTGGAGTGCAACCTCCCTGGGCT
CAGTTGATCTTCCCTCAGCCTCTTGAGTAGCTGGGACTACAGGTGTGTGCCACCATGCCC
AGCTAGTGTTTGTATTTTTTGTGGAGATGGTGTTTTGCCGCATTGCCCAGTGTGGTCTTG
AACTAGTGCTCAAGAGGCCTGCCTCCTTCAACCTCTCAAAGTGTTAGGATTACAGATGTG
AACTACTGTGCCTGATCCAAAGTCAGATTTTCTTTGCTTACTTAGTCAAGTTCGTCTATG

CTTTTATTATACTTAATATATTAGTATAGTTACTGTATTAGTATATTAGCATATTTAATA
TATTATTATACTTATCATACTTGAGTATATTGAGTATATTTACACTTTTAGTATATTTGT
ATACACACACCACATTTTTATTATTTATCTTTTTTTTGAGACAGAGTCTCCCTCTGTCTC
CCAGGCTGAAGCACAGTTGGCTCACTGCAACCTCTGCCTCTTGGGCTCAAGTGATTCTCG
TGCCTCACCCTCCTGAGTAGCAGGGATTACAGGTGTCCACCACCAAGCCTGGCTAATTTT
TGTATTTTTAGTGGATATGGGGTTTTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGAC
CTCAAATGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGAATTACTGGCGTGAGCCACTG
CACCCAGCCTATTATCTGTCTTTTGATGGACATTTAAGTTGTCTCTATATACTAGCTATT
GTGAATAATGCTGCAGTGAACATGAGAGTGCTTGAAAACACTAATGTAACATAAAGGTAA
CAAATAATAAATGTCATGTGTTTATCTTGAAAG**GAACTGAAATACGACGTTGGTGGAGGA**
GAACGGTTTGATTCTTTGACAGATCTTGTGGAACATTATAAGAAGAATCCTATGGTGGAA
ACATTGGGTACAGTACTACAACTCAAGCAGGTGAGCAGATTGGAAAGCTCAAGCTTTCTC
CTTAAAAACTTAAAACAAATCCTAATAGAGAATTTTGCAAACATACAGAGGTAGACAGAA
TAGTATCATCAGCCTCCATGTACCCATTGCAGCTTCAACTATCAAATCTTTTTTTTTTTT
TTTTTTTTTGAGACAGTCTTACTCTGTCACCCAGTCTGGAGTACAGTGTTGCAATCTTGG
CTCACTACAACCTCTGCTTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAG
CTGGGACTACAGGTGCCCACCACCATGCCCGGCTAGTTTTTGTGTTTTTAATAGAGATGG
GGTTTCACCATGTTGGCCTGGCTGGTCTTGAATTCCCGACCTCAGGTTTTCTGCCCGCCT
TGGCCTCCCGAAGTTTTGGGATTACAGGCGTGAGCTACCACGCCCGGCCCTAAATCTTTT
CTTATTATGATTCCACTCACTGACTGCCGCTATAGTACTTGGAAACATATTCCAGATTTA
TATTATTCCCATATTTATCTGTAAAAGGCATTACAGAGGTTCTTTTTTTTTTTTTTTTTT
TTTGAGATGGAGTTTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGTGTTCTTGGCTCA
CTGCAACCTCTGCGTCCCGGGTTCAAGAGCTTCTCCTGCCTCAGCCTCCTGAGTAGCTGG
GATTATAGGTGGTGCCACTACACCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTT
CACCATGTTAGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCTGCCTGCCTCAGCC
TCTCAAAGTGCTGGGATTATAGGCATGAGCCACTGCATCTGGCCTAAGGCTGTACAGAGT
TTTAAAGCAAGTTTTCATTATAGATCCACTTCTGGTTACCTTTAGGTAACCTCACTTATT
CACTTTGGCATTGTTGCTATTTCAAATTTCACCTTTATGATAGTGGAAAATGATATAATC
TCTCTAAATAATGTGGTCTATTCATAAAGAAAAATAGGCTTGAATTTATATCAGCAGAGT
AAAGTGTATGTGAAGACTGAAGAAAGATACATTTTCTGGCTGAACAGAAAACACGGTGAA
ACGATTTGAAAACTTTTATTGTGAATTACAGGGTCCTATGAACCCTCTGTCCGTGCCTTT
ATGAATATCAACATAGACATGTTTTTTTTTTTTTTTTTGCATTAACACCGTTTTCTGTAA
TATTTTCTTTATTTTACATCAACTGCTGTACTCGATCAG**CCCCTTAACACGACTCGTATA**
AATGCTGCTGAAATAGAAAGCAGAGTTCGAGAACTAAGCAAATTAGCTGAGACCACAGAT
AAAGTCAAACAAGGCTTTTGGGAAGAATTTGAGGTAAGTTATTAAAAAACTGTTTTTACG
TGAGTTGTTATATCCTATTTTTAGTGGAGGAGAAGTTGCTCTTGTGTTTGGAATTGGACC
TGAGAGACTTGAAACTGACGTCCTTTTTTAATTCGGCCATTGATTGACACGGAGCAAGTT

GCTGAGAGGGCTTCTTCGAAACAGAAGAGCATTGTGTTCTGAGGGAAGGGAGTTGGCAGT
GAGTAGTCAATGGATGTGCTAGCCGCTCCATTTGGCTCTTTTGGTTTGGACTGGTGGCAA
AATCTCAGAGAAACAAAAGGATCTAATTTCTTCGAAAGATTTCCAGCATGCACTGGGGTC
TTTAGAAACAATCTATAGCCTTAGTGCAGCAAATGAGTATGAGTAAAAGAGAAACACCTT
GTGGTGGCTTTTTTTTTTTTTTTTTGAGACAGGGTCTCGCTCTGTCGCCGAAGCTGGAG
TGTAGTGGCGTGATCTCGGTTTACTGCAGCCCCGTCCTCCCTGGGCTCAAGTGATCTTCC
CATCTCAGCCTACTGAGTAGCTGGGACTACAGGCACATGCCCCTATGCCTGGCTAATTTT
TGTATTTTTGGTAGAGATGAGGTTTTGCAGTGTTGCCCAGGCTGGTCTTGAACTCTTGGG
CTCAAGTGATCCTCCTACTTAAGCTTCCCGAGTAGCTGGGACTACAGGCACACGATACCA
TGCCCATCTAATTTTTGTATTTTTTTGTAGAGATGGGGTTTTGCAGTGTTGCCCAGGCTG
GTCTTGAACTCTTGGGCTCAAGTGATCCTCCAGCTTTGACGTGCCAAATGTGGTGGCTTT
AATTTCAGAGTTCAAATTGATAACTCTGGTAAGTTAAGTGAACTGATTTCTTTTTTTTTT
AAATTATTTTTGTTGATTATACTTTAAGTTCTGGGATATATGTGCAGAACGTGCAGGTTT
GTACATAGGTATACATGTGCCATCATGGTTTGCTGCACACATTAACCCATCATTTAGGTT
TTAAGTCCTGCATGCATTAGGTGTTTGTCCTAATGCTCTCCCTCCCCTTTAATGCATCAG
TGAAAAAGTGATGATAGGCTGGGCGTGGTGGCTCACTCCTGTAATCTCAGCACTTTGAGA
GGGTGAGGCAGGTGGACCACTTGAATCCAGGAGTTTGCCCCCATCCCCAGACAGTGTGTG
TGATGTTCCCCTCCCTGTGTCCATGTGTTCTCATTGTTTGGTTTTCTGTTCCTGTGTTAG
TTTGCTGAGAATGATGGTTTCCAGCTTCATCCATGACCCTGCAAAGGACATGAACTCATT
CTTTTTTTATGGCTGCATAGTATTCCATGGTGTGTATGTGCCACATTTTCTTTATCCGGT
CTATCATTGATGGGCATTTGGGTTGGTTCCAAGTCTTTGCTATTGTAAATAGTGCTGCAA
TAAACATATGTGTGCATATGTCTTTATAGTAGAATGTTTTATAATCCTTTGGGTATATAC
CCAGTAATGGGATTGCTGGGTCAAATGGTATTTCTGGTTCTAGATCCTTGAGGAGTCACC
ACACTGTCTTCCACAATGGTTCAACTAATTTACACTCCCACCAACAGTGTAAAAGCATTC
CTATTTCTCCACATCTTCTCCAGCATCTGTTGTTTCCTGACTTTAAGTGAACTGATCTCT
TTCCTGAAACTAACTTGGGTTGGAGAATGTCCCTGATGGGAATGTGCTGTGTTCCCATTG
CACTCTTCTATATCACTTACCCATTGACAATGTGATCTCTTTCATTTTCTCCTCATCCAT
TTGACAGAAAACTTCAAAAACAAGGATTCTGGCATATTTACCTTTGCAGTTGTCCCCAGC
ATGTAGCACGGTGCCTAGTACACAGAAGAAACTCCATAAATGTTTGTTGAATGAGATTTA
CATTTAACTCATGTTTACATCATTTTATTTTCCTGTTCTGTTTTATGGGAATGATTATTC
TATGCTTTTTGAGGACTACAATTTATAAATATTTGTGGATTGAATGAATAAGTGAATACT
GGGCAAATAAAGTCCTTTTAGCCAGAGTATGTCTGAACAACTTGCTGAGATAGATATGAT
TTCCCATTTTCCAGCTGAGGGGCCTAAGGGAGGTTAAGTAAATTATTCAATCTTCATACC
ACAGTTTTTGTTTTGTTTTGTTTTGTTTTTTTTCCTCCTGAGACAGAGTCTCACTTTGCT
GCCATACTGGAGTACAGTGGTGCAATCATAGCTCACTGCAGCGTCCAACTTCTGGGCTCA
CGCCATCCTCCCACCTCAGCCTCCTGAGTAGCTGGTACTACAGGTGTGCACCACCATAGC
CGGCTAATTTTTCATTTTTTGTAGATATGGGGTCTCACTGTGTTACTCAGGTTGGTCTTG

AACTTCTGAGCTCAAACAATTCTCCTGTCTTGGCCTCTCAAAGTGTTGGGATTACAGGTG
TGAGCCACTGTGCCCGGCCCATACCACAGATATTGATTGAATTCCAGCAGTGGGGAGGAG
TGTGGAATAGAACATTCTCAGTCCTTGCTCAACATTACTGAACAGAGACTTGAATTTGAG
TTTATTCTCTCATCCCAGGCTTCGCGTTAGGCTCTGAAGACACTAGTGAACAAGACAGAC
AGGGTTACTGCCTTTAAAGGGAGCTTTTAGTTGAGAGAAGGAAAACAGTGATGAAAAGCA
TCAGTGAAAAAGTGATGATAGGCTGGGGCGTAGTGGCTACTCCTGTAATCTCAGCACTTT
TAGAGGGTGAGGCAGGCAGCTCACTTGATTCCAGGAGTTTGAGACCAGGCTGGGCAACAT
GGTAAAACCCCGTCTCTACAAAAAATACAAAAAGTAGCTGGGTGTGGGGGTGCGCACCCA
CAGTCCCAGCTACTCTGGGGGTTGAGGTGGGAGGATTGCTCGAGCCTGGGAGATTGAGGC
TGCAGTGAGCTGAGATCACGTCACTGCTCTCCAGCCTGAGCAACAGAGCCAGAACCTGTC
CCAAAAAAAAAAAAAAATTGATGATAAACATAGTGAGACAGAATTTTGAAATCTCAGCCTC
ACTGTTGCCTTCCTTGTCCCCTGCCTGCCTAAATAATAAAAGGCAGCATTTCAGCAGTCA
TTCATTTCATTACTTTCACTTCATTTCACCTTCATAAAGCCTCATGAGGTAAGATGGGAA
GATACAGAAGTTTTAGAAACCGCTCATCAAAATTGAATGGAAAGCCGATTGTTCCAAAAC
TTTTTAGTGTGGAAAATTTCTATTATATGCAAAAGTAGAGAGAATGGGATAGTTATAGCA
GTATACCTGACACCCAGCATTAACAACTGTTGATAATATGGCCAATCTTTTTCGACTCTG
CCCCACTCACTTCCCCAGCCCTGACTTGTCTTGAAGCAAATACTTTTTTTTTTTTTTTGA
GATAGAGTTTTGTTTTGTTTTGTTTTTTGTTTTTGAGATGGAGTCTCACTCTGTCCCCCA
AGCTGGAGTGCTGTGGCTTGATCTTGGCTCACTACAACCTCCGCCTCCTGGGTTCAAGTG
ATTCTTGTGCCTCAGCCTCCTGAGTAACTGGGATTACAGGTGTGTACCACCATGCCCAGC
TAATTTTTGTATTTTTAGTAGGGACAGGGTTTTCACTATGTTGGCCACGCTGGTCTCAAA
CTCCTGACCTCAGGTGATCCGCCTGACTTGGCCTCCGAAAGTGCTGGGATTGTAGGTGTG
AGCCACTGCTCCCGGCCTTGAAGCAAATCTTAACACATCATTTCGTCTGTAACTATTTTA
TTTCAAAAAATTATAACCTGAATAGCATTATCATATCTAAAACTATTAACAGTATTTCCT
TAATATTAACACATATCAGTCACATTTTCCTGATTGCTACACACACACACACACACACAC
ACACACACACACTTGCAATTTGTGTTTTTTTCTTTTTAGATGGATCTCACTCTGTTGCCC
AGGCTGGAGTGCAATGGTGCATTCTCAGCTCACTGCAACCTCCACCTCCTGGGCTCAACT
GATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACCTCACCTGG
CTAGTTTTTGTATTTTTAGTAGAGGTGGGGTTTCACCATGTTGGCCAGGTTGGTCTCAAA
CTTCCGACCTCAGGTGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATG
AGCCACTGTGCCCAGCAGCAATTTGTTTGAATTGGGAGTGCTTTCTTCCACCTTGATTAT
GAAAAAATTTCAAATGTGTATAAAACAGATTCATATAAAGGATCCTGATATGCCATTATC
AGCTTTATCAATTATCCCTGTCATCATATTTTTTATTTATAAATATTTCAATATTTGTGG
AATCCTTAAAAATGCATCACATAACCCAACATTGTTCATATTATACCAATTGTCTTATAA
TTTAAAAATATTTTGTTCAATCATTTTTCAGATAAGCTTCACACACTGTGGTTGGCTAAG
TCTCATAATATTTCTGTTGTAAAAATCTTAAGTCTGGGCGTGGTGGCACACGGCTGTCAT
TCCAGCACTTTGGGAGGCTGAGGTGGGCGGATCACGAGGTCAAGAGATCGAGACCATCCT

GGCCAACATGGTGAAACCCGGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTAGT
GCGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGAA
GGTGGCAGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTAGAGACAGAGTGCG
GCTTCATCTCAAAACGAAACAAAACAAAACAATCTTAAGTCTCTTAGAATACTTTGATGC
CCCTTCCATCTCTCTTTTTCTGTCTTCCTTCCCCCTCTCCCTGTCTTTTCTGCTGTTGAA
GAAAGCAGATCATTTGTCCTGAGAGTTACTTATAGTCTGAATTTTGCTGAGTGCCTCTCT
GTGGTGGACTTAAGCATGTATCCATCCCTTATATTTCTTGTAAGTTGATATATCTAGAGA
CTTCATTGGATACAAGTTTTCTTTGGCAAGATAGCATGTATGGTGGTGTATCAGGAGGTG
TTTATGTCCTGTTGTTTCTTCTCTGATTTTCTTAGCAGCTCCTGATCATTATTACTTAGA
TCCATTAATTCATAAGGGACTATATGGTAGTGATATTGTAATTTTATCATTCTTCTTCAT
TTGTTAGGTTGGCATATTTCTATAAAAAGCTTTTCATCGCCGAGGGTTGATTTTTTCCTT
CTTACTAAGCAGTTTTCTTTTCTTTTTCTTTTTTTTTTTTTTTGAGGTAGGTCTCACTGTG
TTGCTCAGGCTGGTGTGCAGTGGCGCAAACACACAGTTGCGAACTCTTGGGCTGAGGTGA
TCCTCCTGCCTCAGTTTCCTGTGTAGTTGGGACCACAGGTGCATGCCACCATGCCTGGCT
AATTTTTTGATTCTTTTGTAGAGATGAGGTCTCACTTTATTTCCCAGGCTGGTCTTGAAT
GTCTGGGCTCAAGCAATCTTTCTACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACAT
ACCACCATGCCCAGCTAATTTTTTAATTTTTATTTTTAGTAGAGATGTGGTCGTATTATG
TTGCTCAGGATGGTCTCGAACTGCAGAGCTCAAGTGATCCTCCTGCCTCAGCCTCCCAGT
GTGCTGGGATTATAGGTGTACTACAGGCAAGAGCCAATGAGCCTGGTCAGATTTTTTTTT
CCTGATTTGAAATCTGTTATGGGTTCAATTGATACTTCCAAATCAAACTCAGGGTTTCAG
GATTTTTACTAACCTCATTGATCTTACCCATGTATCTCCTTTCTCTAATGCCAAAAATCC
TACTTCTTGAAGCCATAATAAGATTATTCATTTGTTTTATCCCACATTACACACAACAAT
CTTAGAATAATGACTTCCCAATAATATGATTACTGAAAACAGTTTAATTTTTTTTGCGCT
TTTCAAAAAAATCCTTCAGAGATGTGTAGTCAAGTTACTGTATTCTGCTGGGCACAGTGG
CTCACGCCTATAATCCCAGTACTTTGGGAGGACAAGAAGGGAGGATCGCTGGACCTCAGG
AGTTTGAGACCAGCCGGGGCAATATAGTGAGACCCTGTCTCTACAAAAGAAAATTAAAAA
TTAACCAGACATGGTGGCATGTCCCTATAGTCCCAGCTATTGAGAGGCTGTGGCGAGAGT
AGGCTTAAGCCCAGGAGTTTGAAGCTGCAGTGAGATACGATTGTGACACTGTACTCTAGG
GTGACAGAGCAGGGACCCTGTTTTTAAAAAAAAAAAATGAAAAAACTTCCTGTGCCTTAG
ACTCATTTGTAATCGTCCTTCTCTCTGTGTGGCTATATGCTAACTGGGTATATGGTTAGT
TTATTTGTTTCATTTAAAAAATCTCTTTCTGTTAAGTTTTATTTATAATTACACAAATAC
TGGCTTTGATAGTCAAATTGAAAAAACAAAGTGTATTCAAAGAAGTCTACCTTCTATCCT
TGTCCTTTCCTATGTTTTAGCCATAGTATAAAAAGTTATGGTTTATCATTATATTTCAAA
AATATAAGAAGATATTCCCATATCCCACTTTTTCTTAAACAGTAGCATAACTTTACATAC
TTTTTTCTAACCTTGCTTTTTTAAATATCCTGGACATCCTGGATATCCATAATAGTGTCT
AGAGATAGTCTTCATTCTTTTTTTACTGTATAGTAATCCACTGTGTACTTGTACCATAGT
TTATTCAACCTATTGATGGGCATTTGGGTAGTTTCCAAATGTATCACAGAGAGGATTACA

GTGAATAGCCTTGTGTATGCATCCTGCTTTACTTTTGCTGACTACTGGTAATATTAACAT
TTTTTATGTTCTGTATTTAAAAAATGGTGGTTATTATTCATCTATAACTTTTATTATACA
TGACTTTGGTTAGCATGCTTTAACCTTTTAGCATAACATTTGCAAGCTACTTGTTTTAAT
TAAAATTTTGGTTAAATGTAAAAAATAGTGAGCTATTTTGTAATCTAGATTCAATAGAAT
CTTATACTTCCTTTACAAATGATAGCTGAGTTGATCATTTGTGTAAATGACTGTGAACTT
AAAAATTACAGCATTTTTTAAAATAAATTTTTTTAACATTTTAAAATTATTTAAAATAAT
AGACACACAAAGTAAAAAGAGAAGAAAAAAAAAAAGAGACAGGGTCTTGCTATGTTGCCCA
GGCTGGTCTCAAACTCCCAGGCTCAAATGATCCTCCTGCCTTGGCCTCCTAAAGTGTAAG
CCACCACACTTGGCAAAAATTAGTTTCTTTAAAACAAAAACATTACAGGTTATCTGGTAC
CATGGTAGCTTCTTTAACACTAGGTTCACTTAGAACAAAGCTTAGGAACAAAGTCAGACT
TTCACAAAGAGCTTGTGTGGCAATGGGGTATTTTTTGCAAATTCCATTGGTGGGGTCAAG
ATGTGAGTTTAGAAGGAACTCTTAGCCTGACTCTTCTGGCCATGGAAAAAGATGGTTGCT
TCTAAATGCTGACCTGGTGATTTTACACTGTCACATCTCAAATTGTGGTCATCTTTTATA
CATTATTAACAACAAAAGGGAAAAATTGAGTTGACTTTAAGAGGAAGTGGAAAATAACGA
GATCACATCTGTACTCTACAGGCTCTCCACAGAGGTCAGACTGAGGTGGTAAAATTGTTG
TGCACTAAATTAGGGCATTAACGTTTCATGGAAACTGAAGCTATATCTAAATAGCTGATG
GCCTGCTTTCTAGATCTCCTATATACCTGCTTCTCAAATTCAGTCTGTTTTAAAAAATTG
CCCTTTGAGGTTGGAACCAGCGAAATAAGGCTGAAAACAGAATAAGCCATTATTGAAAAA
ATTAGGAACTTGGAAGCAGATACTCATAATCTAAATCCTCTGAAGCTAAAGTTTGATCCA
CAATAGCAAAGCATTATCATTTTAGTGATTGTACCTTAGTTGTTTCCTGGCAGGTGATAA
ATTTGGGATCACTTTCTTCTTACAGTGTGCTCTGATAGTCTTTAAAACAAACCAGAGCTC
TAAATTGTAATGCCATTGGTAATTTAACTCTGATTTGTCTCTATGCCTGTCTCCTGGTGT
TCTGTAAAATTCTACACGTCATTTCAGGTATCACTATCCAGAAGACGTTACTTTTGCCTT
TGATGCACTTTAAAATGTGAAGTCTCTTGTGAAGCTCTTTGGTTATTTTCTCCTTTGCTG
CTGAAATAAATTCAGGTTGATGATTTTCTTGTAGGATATGTTGTGTGATCTAGACATTGC
AAACCCAAGTCTTTGATTTTTTTTTCCCTACAGATTGCCTGTTTCTTTTTTATTTTAATT
TTTATTAGTTATTATTATTTTTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCA
GAGGTGTGATAGCTCACTGCAACCTCCACCTCCCGGGTTCTTGTGCCTCAGCCACCCAGG
TAGCTGGGATTACAGGCACGTACCACCACTCTCAGCTAATTTTTTTGTATTTTTAGTAGG
GATGGGATTTCTCCATGTTGGCCAGGCTGATCTCAAACTCCTGACCTTAAGTGATCTTCC
TGCCTTGGTCTCTGAAAGTGTTGGGATTACAGGTGTGAGCCACTGTGCCTGGCCAGTTAT
TAATTTTTTTAAAGAGATGGGGTCTCACTATCTTGCCCAGGCTGGAGTGCAGTGGCTCTT
TACAGGCACTGTTGTAGTGCACTGCAGCCTTGAACTCCTGGGCTCAAGTGATCCTCCTGA
GAGGCTGGAATTACAGGCACACACCACTGTGTCCAACAGATTGCCCATTTGTGATCTGTG
TAAATATCTCTCACTTCCTGCAGTATCTCTGCTCAAGAATGTAAAGAGATGGATAATATT
TTTAGATTTGTTGAAACAAAGTAAAGTTCTGCTCAAATGAGAATGACACTAACTAAATGA
AAAGGCCGGTTATAATTCTGTAATTTTGTGCCTGCAATGTGTGTGTTATTGTACACTTGA

ATCGGCCCTGTGCATTGTGGCGAGGTGCATATTGCATGGTTGTATTGAAAAGGTGCTTGG
GCCGGGCGTGGTGGCTCACACCTGTAATCCCAGCAATTTGGGAGGCTGAGGCAGCTGGAT
TACCTGAGGTTAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCTGTTTCTAGTA
AAAAATACAAAAAATTAGCTGGGTGTGGTGGTGGGTGCCTGTAATACCAGCTACTAGGGA
GGCTAAGGCAGGGAGAATTGCTTAAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTG
TGCCACTGCACTCCAGCCTGAGTGTATCACAAAAAAAAAAAAAAAAGGTTTTTGCCCTCT
CTCTGTGCCTGCTGCTCCCTGTTGAGTCCTATAGGCCTGAGCTGCCAGGGGGTACTGTGG
GCTGAGACTGGACATTGCAACCGACTGCAAGGCACCGTGGGACCCAGGTTGTGGATGGAC
TGTCTCTCGGGCTTTCTTCTTTCCATTCATCTTCCTCCTCTAACTCCCCTCTGTATCCAG
TATCCTTGCTCTCCATACACCTGCTTCATTCTTTTTCCTTCAGTAGATTTTTCTGCTTCT
TGACTTACAAACCCTACTTCTAGCCCCTTTCAGATATTGAAACTAGCAACTTTCAGGCTT
TGTACCAAAGTCTCAGAGATTCTCATTGACTCGGATGCCATCCATCTCTAGTCCAAAGAA
CAATGTCAAGGACATGAACATGTGGAACAAAAGTGTCTGCTGTGGACACCTTTGGGGAGA
AATAGTTTTCAGTGATGAGGGTTGTAGTGAGTTGGGCAGATATCCCAAAAATATCTGCCA
AAAACTATAGACACTTCTGGTTGCAGTGACTTATTCCTTCCTTCATTCAGCAAATACTGA
TTGAACACCGACTGTATGTCTGGATCTATTCTAGGTTTTGGGGGTGGAGCAGTGAACAAA
TCAGTCTTTATCTTTATAGAGTGTACAGTCAAGTGGGAGAGACAGGCAGTAAACAAAGAA
ACAGTTCAATATTCAATCTGTGAGATGGTGATAAGTGCTACAGAGAAAACAAACTAGTGT
AAGATAAAAAGGGTGTTTTGATAGGCCTTTACTATTTAGGTCTCTTTGATAAGGTGGCAT
TTGAACAAAGCTCTGAAGGAAATAATGGAGCCAACCATGCATATAACCTCAGGGAGAACA
TTCTAGGTAGAGGGAACAGCAAGTGCAAAGGCCCTGAAGTGGGGGTTTGTTTACCTTGTT
GCACAATCTGCACACAGGCCAGTACAATTGGAATGGATGGGAAATGTAAAAGAGAGAAGT
TGAAAAGGCCAGGTGCAGTGGCTCATGCCTACAATCCCAGCATTTTGGGAGGCTGAAGTG
GGAGGAATTTGAGATCAGCCTGGGCAACAGAACCAGACCTCGGGCTAATTTTTGTATTTT
TAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGATCTCAAACTCCTGACCTCAGGTG
ATCCTCCTGCCTCAGCCTCCCAAAGTGCTAGGATTACAGGTGTGAGCCATGGCCCCCAGC
CGTATCTTTGTCTTAAAAAGTAATCTCTGTGCTTGGTAGGCCAAGAATTTAAAATATAAA
AAATTTAAGAAAGAAAAAAAATAAGTAAAGTAACTATACAGGTTGGTCTGGCCGTAATGG
TGAGTGTCATTATTTTTCTTCCCTAGGTATTTTGGCTCTGTTGCTCAGAGCAGTGCAGGC
GAAATGGTCATTAGGGCATCGTCATGGTGCCTGGGGATGCCTGGCTCAGCCAGTTTATTT
TCTGTCTGCCTCTCTCCTTGGTCCTTTTCCTCCACTTTCATTCATGAAATTCTAGTCAAG
AGCTGGGTCCAGTGGTTTTCAATCCAAGGGCTTTGGAAGCCTCTGGGGTCTATTTTGGTC
ATTGCAGTCACTGGGCTGCTGCTCCTGGCATTTAGGTTGGCAGGGGTCTGGGCTGGGAAG
CAGGAATGTTCAGTGGCCATAAATGTAAGGGTTGGTCTTACATTTACATAAGGGAGACAA
TGAAAACTTAACTCCTCCACAGTAGTGGAGTAGTGCCGTTGGGTACTCACAGTCAGTAGT
GCCGTTGGGTACTCACATGTACAACATGGATCAGGACATTGACTTTCTGTGGATACCTTT
TAATAGTTTATTAGATGTGTTAGGCTGTTTTGCACTGCTCTAAAGGAATATCTGAGTCTA

GGTAATTTATAAAGACAAGAGGTTTAATTGGCTCATGGTTCTGAAGGCTGTACAAGCATG
GCTCCAGCATCTGCTTCTGGTGAGGGCCTCAGGAAGCTTCCGGTCATAGTGGAAGGCAAA
AGGAGGGCAGACGATCACATGGCCGGAGTGGTGGCAAGGGTGGGGTGGGAGCCACGCTCT
TTTTTTAATTTTATTTTAATTTGAGACAGTGTCTCACTCTTTTGCCCAGCCTGGAGTGCA
GTGGCGTGATCTCAGCTCACTGCAGCCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCT
CAGCCTCCTGAGTAGTTGGGACTACAGGCGCGCATCACAATGCCCAGCTGATTTTTGTAT
TTTTAGCAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGGACTCCTGATCTCAA
GTAATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCAC
GGCCACCACACTGTTTTAAACAACCAGATTGCACGTGAACTTAGAGTGAGAACTCACTGT
GAGGATGGCACCAAAACATTCATGAAGGATCCACCACCTTCCTTTAGGCCCCACCTCCAA
CACTGGAGGTCATATTTCAACTTGAGATTTGGAGGGGACAGACATCCAAACCGTATCATT
AAATTTAATAGTTTTATGCAGTTTTTTTGGCTCTAGATCTGTTTAGACTCCTGCAGTCAG
GTGTCTGTAACTAGCCTCTGGTCCTTTTTGAGAGTTCACAGTTTGGTGCAAACCCTTTGG
ATGTATTATTTGGGAAAATGGGATATCTGGCAGCCTGTGTCCCTGCTTTACATTATCCTT
TTTGCTGCCTGCCCCAAGCCTCCTCATTAGCATCCCTGCCAAGGCCAGTGGAGAAGGATG
GAGATGCGGTGACATTCAGCTTGACAGGTCATTAGCAGCTTTTGTGCCCTAGGGACTGCT
GGTGGGAGGGAGGTTGTGGAAGATAAACCCTGACAGGAATGTATTCTCCTCGAGGGCAGG
GTTTATTTGATATTTTTCTGGAGCTTAGAACCATAAGCCTGGTGCTGGGGAGGAAGCGCC
CTTAGCATTTGGTAGCCTCTGTGGGCAGAGCATGGAAAGTCACAACTTCTGAATTGTTTG
TATTTTCAGTCTCACTCTAGATGGATGGCATCTTCTGCTATGGGAAATGAAATATGTTTA
GGCAACTTGAGTCCCAGGTGCAGATGAGGCTGGGCTAATTGGTGCACTAGGGAAGGAGCC
GGGGGAGAGATGTGCTGTTAGCTATTATCAATCTGTGACAACTGTCAGCTGCTGGCAGTT
AGCACCCACCTGAGCCTGGGATGCAGGGGTGCCTCTCCTGTCCTCTGTGGAAGCCTCTGG
ACCCAGCAGCCATCTTGACTGTGCACTGTTCAAGCCCCAAGTCCGCCTGGAAGAGGTGAT
TGAGAACTTACTGCAGGATAAGGAAAGCGCAGGACAGGTGCAGTGGCTCACGCCTGTAAT
CTCAGTGCTTTGGGAGGCTGAGGCCGGAGGAGGGCTGGAGTCCTTGAGTGCGAGACCAGC
CTGGGCAACATAGTGAGACCCTGTCTTTACAAAAAGGAAAAGAATTAGCCAGATGTGGTG
GTGCGTGCCTGTAGTCCCAGCCACTCAAGAGGCTGAGGTGCGAGGATCACTTGAGCCCAG
GAGTTTGAGGTTACAGTGAGCTATGATCATACCACTGCATTCCAGCCTGGGTGAGAGAGC
ATGACTCTGTCCCAACAACAAAAAAAAAAGATTAAGGGAAGCCTCTGGCAGACCTGATGAT
GGGTGGCCCAGCCAAAATGAGTATTGATGAGGATTTCCCTGGTCTGGAACTCTGAATTTA
GTCTGGCAAAGTATTCCCTTTGTGTTGTGAGATGATTCTTGGTGTTACCCCATCACGGTA
GGTAAGATGAATTAGCAAATGAGAAAGGCTTTCTCTTTTTCATCCTTATCTAGTCCGTAG
ATGAAGCCTGAAGAAGGTCTCCATATGGTAGTAGTAAGTGTTTAACATCTACCTCTAACA
CTTGCCTGTGTCTTTTTTTTTTTGCAAAGCCTCAGGAATGCCCCAGTATCTAGGTAGAAT
TTGATAATATTTCATTTTTGTTATATTCCCTTTTCTGTTTACCTTCTATATACAGCAAAA
TGAAAAAATTTTTAAAATTTGTGCAAGTAAGGGCAATTTCTTTTTTCTTTTTCTTTTTTT

TTGAGACAGGGTCTTGCTCTGGCACCCAGGCTGGAGTGCAGTGACACAATCTCGGCTCAC
TGCAACCTCTGCTTCCTGGGTTTAAGCGATTCTCCTGCCTCAGGCTTCCAAGTAGCTGGG
ATTACAGGTGCCTGCCACCACTCCCAGCTAATTTTCATATTTTTAGTAGAGACCAGGTTT
TGCCATGTTGACTGGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCATCCACCTTGGCC
TCCCAAAGTGCTGGGATTATAGGCTTGAGCCACTGGGCCTGGCTGAGGCAGTTTCTTTTT
GAAATATATTTTGTGAAGGAGAAAAAGAGGAGTTCAGTTTAAAGAAACAAATGACATAAG
AGGTGGTATGCAGAGATGCCAAAGCATCTTGAAGGTGCTTTTTTTTTTTGGAAACAGAGTC
TTGCTTCATTGCCCAGTCTGGTCTGCAGTGGTGCAATCATGGTTCCCTGCAGCCTTGACC
TTCTGGGCTCAAGTAATCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGATGCAT
GCCACTATGTCTGGCTAATCTTTAAATTTTTTGTAGAAGCCAGCTCTCACCATATTGCCC
AGGCTGGTCTTGACCTCCTGTCCTCGAGCAAAAATACCGATTTTGATTAAGTCTGGGGTA
GGACCTGGGGCTGGGATTCTAACCAGCTCCCAGGTGGTGCTAATGCTGCTGGTCTACAGA
CCACACGTGGAGTAGCCAGTGTAGAGTTCATGTAGCAATAGTGATGTCATAGAAATAGCC
AGTATCTGTATACTTGCTTTGTTGTATGTCACGCACTGTATAGTGATGTACATGCATCTC
ATTTGACCCTCACCCCGCCCCTTTGGGGGGTAGAAAGGATTGTGCTCATTTCACACTCAAG
GAAACTGAGGCACAGACAGGCAAAGTAGCTTGGCGAAACAGAAAGGAACTTAGAGGCAGG
CCCTGATTAGCTCAGAGACTAGAAGGCCTTGTGCGTCATCCTGAACAGCTTGGACTTGAT
CTTGAAGGTGGAGGGAGAAATTGAAGGGTAATTAAACAGGAACTGTAGGAAATTCACCTT
GCATAGTGATTGCTTTGGCCACGTGTGCCCTGCCACCGCCCCCCCACCTCAGTGAAGTGT
CATGCGAAGTTGGGTTCGTAAATGAAGGCCCGAATGCTTTCCTGACAAGTTTGTTTTAAA
TCAAGCTGCTAATTAGTCCCAGTCCCCCTCCCCCGGTATGTATTTTTTTGTTGATGTCGT
TTCACTTCATTTAGTTGAAGTGATTGATTCAGTTCAGTGTTTGAACTTCTTTTTGAACCT
CACCTTAATAACCTGTCTAAACATCAAGGTTAAACCTTCTTGCTAACACAGCAGTATTGC
TTGGTAAGACTGGCTCACAGTCCAAGGAAATGCTTGCCCAGAGAGGGCAAACTGCCTTAA
CTCCTTAACCTGAGCTCATTAAAAAAAATTCAAATGACTGATTCCTTGTCACAGTTCTAC
CTACATTGTTTTTATTTTTGTCCAGGTTTCAGCTAGTTAAATGCTTTTGTGATGAGCTTA
TGTCCAGGCTGAAGGTTGCATTTTGAAACTGAGCGTCAAATACCAATTTAAAGTCCAGAC
CTTTACACTTGTGAAATTCAGATAAATGAAATGGAAATAAAACAGGGCTGCTGTGTTGTG
AAATATGACTGTGTTTTTCCTTGTAGGACTCTTTGAGGGTAGCCATTTTGGCATTTTATA
TATAAATTTTCTTTTCTTAGCCTACCTTTTACTTTCTTGATTTGCCTATTTGTGATTTCC
CATTAAACACTAGGCTTTTTGTAAACCAATTATCCCTTGAAATTGACTTTTTTTTTTTTT
GAGACAGGATCTTGTTTTGCCACACAGGCTGGAGTGCCGTGGCTCCATCATATGATAAAC
AGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACCCTGTCTTATTTAAAACA
AAAAAAGAAGAAGAAAAAAAGAATATAGATCACAGCTGTTATTTGTATATGCTACGCCAA
TCCTTGTTGGGTTTCATTCTTTATAATTGTTATTTTTAAAGATTTTTCTTATGAATATTC
TATTGTTTCATTGTAGAAAATTTAAGGGAGAACACAGTGGGAAAAAAAAAACAAGAAAAG
GACTTCATAATCCTGCTACCCTGGGAGAAAAAAAAAATCACCATTACCTATTTGGTTCTT

CTCCCACTTTTTTTTTTTTTCGAGATGGAGTCTCCCTTTGTTACCCAGGCTGGAGGGCAGG
GACGTGATCTTGGCTCTCTGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCGTGCCTCA
GCCTCCCGAGTATCTGGGATTACAGGGGTGTGCCATCACACCTGGCTAATTTTTGTATTT
TTAGTAGAGACGGGGTTTTGTCATGTTGGCCAGGCTGGTTTGTTGGCCATGTCTGGTTTT
TTGTCATATTGGCCAGTCTGTTTGTCATGTCAGGCTGACATGTTTTGTCATGTTGGCCAG
GCTGGTCTTTAACTCCTGACTTCAGGTAATCCTGAAGTGCTAGGATTATAGGCGTGAGCC
ATTGCACCTGGCCTTCTGCCTTTTTTTTAAAGAAAAAAAATTAAAACATTTTTTTCTTTT
TAAGATAGCGTCTCATTTTGTTGCCCAGGCTGGTCTTGAACTCCTGGGCTCAAGTGATCC
TCCAGCCTCAGCCTCTGGAGTAGCTGGGACTACAGATGCACATCATGGTGTCCTTATGCC
ATTTCTTTTGTACGTAGGTGAATGCAAGTGTATGATTACATCATATGCTATTTTGGAGGT
TTGACTTTCTTTTCACTTTCATCATCTTTCCAAGGTGTTATTTTCCTAGTACATCTTTTT
AAATGGACATAGAACATTCTTTTGTATGAACAAACAATAGTTTTATTTAGGCGGTCCTTT
CCTGTTGGACATTTATATTATTTTCAGCATTTCTCCACAGTTGTTGCAGCATTCAGATGA
ACCTTCTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTTTCGCCCAGGCTGGAGTGCAG
TGGCACAATCTCTCCTCAAGTGATTCCTGTGTCACCCTCCCACGTAGCTGGGATTACAGG
TGCCCATGTCTGGCTAATTTTTGTGTTTTTGGTAGAGCTGTGGTTTTACCATGTTGGCCA
GGCTGGTTTCGAACTCCTGCCCTGAAGTGATCTGCCCACCTCAGCCTCCCAAAGTGTGGG
GATTACAGGTGTAAGCCATCACGCCTGACCCAGATGAACATTCTTGTAGCTATCGCACAC
AATTCTGAACATTTCCTAGGATGAATTCCTTAAAGAAGTAATGCTGATCCAGGCTTTTTT
CTTTTTCTGTGACTCTTTGACACGTAATAATATTGACTTTTCTTTCTTTCCAG**ACACTAC**
AACAACAGGAGTGCAAACTTCTCTACAGCCGAAAAGAGGGTCAAAGGCAAGAAAACAAAA
ACAAAAATAGATATAAAAACATCCTGCCCTGTAAGTATCAATATTCCGCTCAGTAATAGT
CACTCTTGGAGATTTTGATTCCTAGCACCTCTGTACCTTTCCTCAGGGTCGTGTGCTCTT
GTTAGCACATCGGAGGCCTTAGCTTCTTTAATTGCAAGCAGTTTCCAAAATAATCAACCA
TGGTGGGTGTTGATGACTTCATTCACTGAGCTCCCGTGATGCTGATTACTGAGTAAAGTT
GCCACTAGGTGGCTTTGTCTGTGGTTGGTTCCTTCTGTTAATTAATTTTCTGTCTGCCCA
AGATAGATCATCTCAAGGCTTGGGATCTCTCAGTGTCAGGGACCTTAGGGTGCCAGATTT
GTGTCTTGACTCCTCCTCACTGGGCCTGTGAGTCCTGGGTAAGGCCTGCCTCCTTTCTGG
GACTCAGTTCCCTTAAGTGGGAAACAGACAAACACCTCCTGAGGGCTCCTAGAACTGTTC
TGCTTGCTGATCCCCTGAGCTCAAGTTACTGGAGAAAGGGTATATACCTAAACTGCTCAG
AAGAAGACTTTGTGGGCCGGGCGCAGTGGCTCACACCTGTAATCCCAGCACTTTCGGAGG
CCGAGGCAAGCGGATCACCTCTGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAA
ACCCCATCTCTACTAAAAATACAAAAATTAGCCATATGTGGTGGTGTGCGCCTGTAATCC
CAGCTACTCGGGAGGCTGAGGCGGGAAATTGGTTGAACCCAGGAGATGGAGGTTGCAGTG
AGCCGAGATGTGCCATTGCACTCCAGCCTGGGTGACAAGAGCAAAACTCCGTCTCAAAAA
AAAAAAAGGAAGACTTTGTGAATATTCGCAAAGCTGTAAAGCTGTACCTTTCAATTTTTT
TTTGAGACATAGTCTCACTCTGTTGCTCAGGGTGCAGTCACAGCTCACTGTAGCCTCAAC

CTCCTGGGCTCAAGCGATTCTCCCACCTCAGCCTCCTGATTAGCTGGGACAATAGGCAGG
CACCAGTACACCTGGTTGATTTTACAGTTTTTCTGTAGGCCGGCGCAGTGGCTTACGCCT
GTAATCCCAGCACCCTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTTAGGAGTTCGAGA
GTAGCCTGGCCAACATGGTGAAACCCCATCTCTATTAAAAATTACAAAAATTAGCTGGGC
GTGGTGGTGGATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCTGAGGCAGGAGAATC
GCTTGAACCTGGGAGGCGGAGGTTGCAATGAGCCGGAGGTGCTATGTGCACCACTGCACT
CCAGGCTGGGCGACAGAGTGAGACTCTGTCTCAAAACAAAAAACGATTTAAAAAATAATA
AAATTTTTTCTAGGGCGGGGTCTCCCTATGTTGCCCAGGCTGGTCTTGAACTCCTGGGCT
CAAGTAGTCCTCCTGCCTCAGCCTCCCAAACTGTTGGGATTACCAGTGCAAGCCATTGTG
CCTGGCTGTACCTTCTGTAACACCCAAATGCCACCTGGCAAAGCCCAAGTTGAATCATGA
GGAAAAAAGGCCTGGAAGGATGTAGACCTTCCTTTTTTCTACTTATTTATTTATTTATTT
TTGAGATAGGGTCTTACTCTGTTGCCCAGGCTGGAGTGCAGTGGCATGATCATGGGTCAC
TGCAGCCTCAACCTCCCGGGCTCAAGTGGTCCTTCCCACCCCAGCCTGCAATGTAGCTGG
GACTACAGGCATGTGCTACCATGCCCAGCTAATTTTTGTATTTTTTGTAATTATTTTTTT
TGTAGAGACAGGGTTTCGTCATGTTGCCTAGGCTGGTCTCGAATTCCTGGGCTCAAACGA
TCTGCCTGCATCGGCCTCCCAAAGTGTTGGGATTACAGGTGTGAACCACTGTGTCTGGCT
ATATCTTCTGTAACACCCAAATGCCACCAGGCAAAGCCCAAGTTGAACCAGGAGGGAAAA
AGGCCTGGCAGGATGTAGGCCTTGCATGAGGATCTCAGAAACTGCACTAAACCAGTCACA
GTTCCTCTCTCCCGAGGTCTAACTCTATGCTGAACTCTTTGCATTTTTATCTCACTTAAT
CCATATCACATGCACAGGAAGGAAGCATTCGTAGTATCCTGGTTTCCTAGACCATTTTAG
CAAGGTTATAAGTGAAGGGGAGTGGGTGGGAGAACTGGCACTAGAGCCCCCAAAGTCACT
GTTCTTAGCACCACTCTAATGCATGGGGTTCTCCATTGATGTGCTATGCAAGGCAGTGCA
CTGAGGAGAAAGGAAGGAACATTTACAACTTCTCTTTATTTATATCCTGTCCCTAAAAAA
AAAAGAAAAAGAAAAATTTGTCTGAGGCCTAGATTGATTGCAGGGAGTGCATAATGTTTT
ATTGATTGATTGATTGATTGTATATAGAGATGGGGGGTCTCACTATATTGCCCAGGCTGA
TCTCGAACTCCTAGGCTCAAGCAATCCTCCTGCTTTGGCTTCCCAAAGTGCTGGGATTAC
AGGCATGAGCGACTGCACCTGGCTATGCATACTATATTTATCCAACTTACAAATAAGGCT
TGCTTGCCTGTAGTGCATATGTGTATACATTTCAGCATAGAAAAACTGTGTGATTGGGGG
TTGTGATCAAATTTGGAGAGCATTGCTCTCATGTCTTATCAGGTCAGAGTCATTTTGTCA
AATCTTGTAAACCATTCTTTGTGTGTGTCTATGCATGAAACATAGTCTTTCTCTTTCTGC
ATGCATATGTACATATACATGGTATATATGTATATCATATCTACATGGATATTGTAATGT
ATATGTATGAGGATGGGGGAAAGTGGAGACATTTGTAATACTGAGAAAAGGCAGTGAGGA
ATTTGCAGAGAAGCAGTTTGAGCTGTAGCATGGTACTAGTGACCTTGAGGAAGCCTTATC
CTTTTTTTTTGGAATTTATTTTTTCAATTTTTAGAAATAGACAAGAGTTTCTCTATGTTG
CCCAGGCTGGTCTTGACCTCCTGGGCCCAAACTATCCTCCTGCCTTGGCTTCCCAAAGTG
CCAGGATTACAGGTGTGGACCACCATGCCTGGCCACCTTGTCCTTTCTATGTCTAAGTTG
TGACATCTGCTCAGGGGTCAGGTGGTATTAAATGGTATAAAATGTATGGGAAAGTGAAGG

GATCAATGGTATGCAGTATCTAAATAGAATATCGCTTTTTCCTCCCTTAAAGGTCTCATT
CAGATGTTTCCTCTGATGAACATCTCATTTCCTTAAAGATGAGGAGTCTGAAGCAAAAAA
GACATTATTCTTTTAAGACACATGGCTGTCTTACTAATTCCCATTGCAAAATATGTTGTT
TAGGTAGAGCACTCAGATTTTTATACGAATAATAGACTTTTGTACAGAATTTGGACAGTT
GATACTATCAGAGCCTTGTGATATTCCACTGCATTATGCTTCACTAAAAAATACCTGGCT
GGGTGCGGTGGCTCACAACTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGCAGATCAC
CTGAGGTCAGGAGTTCAAGATCAGCCTGGCTAACATGGCAAAACCCCATCTCTACTAAAA
ATACAAAAATTAGCCAGATGTGGTGGCACGCTCCTGTAATCCCAGTTACTCAGGAGGCTG
AGGTATGAGAATTGCTTGAGCCCAGGAGGCAGAGGTTGCAGAGAGCCGAGATAGTGCTAT
TGCACTCCAACCTGGGTGACAGAGGAAAACCCTGTCTCAAAAAATAAATTTAAAACAACA
ACAACAACAACAACAAAAACCCCTCTTTATTATGGAAATTTTCAAATATATTCAAGAGCA
TAAAGAACCCACATGTACCCATCACCCAGCTTCAACAATTATCAACTCATGCCCAGTCTT
GGTTTCATCTATACTCTGATCCACATCTCCTCTCTCCTTGAATTATTTTGAAGCCCATCT
CAGACATCATGTCATATATGTATACTTCAATCTTCTTTTTTTTTAAAACTCCCCCTCCCC
TTTTCTTTTTTCTTGAGACTGTGTCTCACTCTGTCATCCAGGCTGGAGTGATCTTGGCTC
ACTGCAATGTCCGCCTCTCGGGTTCAAGCGATTTTTGTACCTCAGCCTCCCTAGTAGCTA
GGATTACAGATGTGGACCAACATGCCTGGCTAATTTTTGTATTTTTAATAGAGACAGGGT
TTTGTCATGTTGGCCAGGCTGGTCTTGACCTCCTGACCTCATATGATCCACCTGCCTTGG
CCTCCCAAAGTGCTGAAATTATAGGCCACTGCGCCCAGCCCAAAATTTCTTGGTTTGAAA
TAATTTTGGAACTCATAAGAAGTTACACATATAGTAGAGAGAATTTTCTTGTACCTTCTC
TGAGCTTCCTATATACCCAATGATAACATCCTATATACCCATAGTATATGATCAAAACTA
GGAAATTGTGAAGATGGCATTTTGAGACATCAGGCAGTGTTCACGTTACTGTTTTGCTTA
CCTGGGCTTTAATTTTTATGTGTTTTTTTTTCAATCATTGAATGAACAAAACTTGGACTA
GGCTGGGGAGTAACTGATTTGAACTGTTTTTTCCTGAAGCAGTCCAGGACTTATGTGACC
GTGGTCTCTTTTTCTTCTAG**TTGATCATACCAGGGTTGTCCTACACGATGGTGATCCCAA**
TGAGCCTGTTTCAGATTACATCAATGCAAATATCATCATGGTAAGCTTTGCTTTTCACAG
TGTTTTCTGACCATACATTTCTAGCCTATTTTTGTATTTTAAATCCTTCCTCATGTCCTG
AAAGTAACTTTAAGGTGTTTGAAGGATTTTCTTCCTAAATTTCTAG**CCTGAATTTGAAAC**
CAAGTGCAACAATTCAAAGCCCAAAAAGAGTTACATTGCCACACAAGGCTGCCTGCAAAA
CACGGTGAATGACTTTTGGCGGATGGTGTTCCAAGAAAACTCCCGAGTGATTGTCATGAC
AACGAAAGAAGTGGAGAGAGGAAAGGTAAATCACAGAAACTTCTTTTCTGCTAAACTGTT
TTTAAAGTATCAGACATGTCAGATTGGCCATGTTTAGGAATTGAATAAATGAATTAAGCT
TACTGTAACTGATTCTCTGGAAAAAAGGGACTAGGAGAAATTTGATTATGTTATTCCTTG
GTGTAGTTTTCTTTATGTTTCTTCTGCTTGGGATTTGTTGAGCTTCTTGGCTCCATGGAT
TTGTAGTTTTCCTTAAATTTGGATAATGTTCAGTCTTAGTTTCTTCAGATACATATCCTG
GGCTGGGCATGGTGGCTCATGCCTGTAGTCCCAGCACTGTGGGGTGTTGAGGTGGGCGGA
TCACTTGAGGTCAGGAGTTTGAGACCAGCCTGGGCAATGTAGTAAGACCCCATCTCTTAA

AAAAAAAAAATGTACCCTGCACAACCTTGTCCTAGGACAGCAGTCATACGTGTATTAGAC
TACTTGAAGTTGTCTCATAGCCCACTGATACTTGGTTTATTTTATTCAGTTTTTTCTCCC
CGTGTTTCATTTCGAATAGCTTCTTTTGCTATGTCTCCAAGTTAATCTTCTGCAATATGT
CATCCGCTCTTAATCCTATCCAGAGTATTTTTCATCACAGACATTGTATTTTTCATCTCT
AGAAGTGTTAATGTCATCTATAGCTTTCCTTTTAACATGTGTAGCATTTTCCTTACCTTT
TGAATGTATGGAGTATTTCTGTTGTTGTTTTTTGTTTTGTAGAGACAGGGTCTCGGTCTG
TTGCCCAGGCCGGAGTGCAGTGGCATGATCTCAGCTCACTGCAGCCTCTGCCTCCCGGTT
CAAATGATTCTCATGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGTGCGTGCCACCACG
CCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGTTGGCCAGGCTGGTT
TTGGAACCCCTGAGCTTAGGTGATCCACCTTCCTTGACCTCCCAAAGTGTTGGGATTATA
GGTGTGAGCCACCATGCCTGGCCATGTTGTCTGTTTTAATTAACTCTGCCTAACTGTCCT
CCCAAATGGTTGCTGCAGTGCTCACTCCCACCAGCAGCACCTGCCTAGGACTCATTACTC
CATACTCTTCAAGACACTTCAGATTAAAAAAATAAATTGTAACACCCCACACCTACAGAA
GAGCGGACAGATCTTATTGAGTGACAGCCCTCTGTGTTATCTCAAAGTGAGCCCACCATG
GTGGTTTTTTTTTTAAATATGGAAAAGTTCTGTGTTTTTGTTTGTGTTCTAGTGAAAGTT
CTTTTTTAGATATCCTTTAATTGGTTTATATAAGATTTTATGTGGAATGTAGCAGTCATA
CCTATAAATTAAACCTAAGGCAGATGGAGAACTTTGGAGTTGAGCCTTCCTACTGTAATT
TTCATATTGGATGTGAAGGGCAGTGTGATTTTCATAAGACTTTCATTGTTGTACTCCTAG
TTGGTATACTTCTGAATACCTTTGAGGCCAGTTCTGGTCATCGTGAAACAAAGGTTTCCT
TCAGCAAATGCCTGTGGTAACATTAGGTGTTCTTGAATTAATGGACCAATGAAAACATCT
TTGTAGTTTCTGCTTCAGGCAAGGGTTTTTTGCCCTAAATGTGGATAGGAAGAATGAAGC
CCTTCATCCTCCTTTTTGCCTGATTATAGCTATAGGAGGTTCACCTGTTCTCAGAAGACA
TGAGGATTGTGAAGAGAGGGGTCTTGTGTTGCTTCAGAGGAATCAGTATCAGTCCCTTTC
AGAAGCTCTCCTGGATAGACAGGCATTAGGGCCAAATCACTCTGCCCCACCCCTCACCAC
CATGTCCTACTCTCTGCTCCCTGTCTCATTCTTCCTCTTTACTTTGGTGGTGCCGAGAGG
ATGACATGATGGGTATTGATTCTCTCCACAGACCTTTCTGACATCCTACTTTCAGTATCC
CCCCAGTGCACAGAAGACAAGCCAGACTGTGGACTGTGTTTGATTCCTGGGCTCTATTTT
AAAAGACAGTGTATTAGTTCTCACATTTTAGAATTTGTTTGCCAAGGTTTCCACGGGAGT
TTAGAAACTAGGGGGAGGGCTGATGTTTAAAGTTAGCTAAAATGTTCTTTTCAGGGTCAT
GATTTAATTTTATATTCTCTGGTGAGTTCCCTATAGTGACTGGGAGCAGTCCTCAGTCTT
GATTGGCCAGTGACAGCATAGAGTACAATTAATATTAGGAGTGCTCATTTGGGGAAACTA
AAATTTGCATCAAATCTGTCAGAGGTGTTTGGATCTACAAAATACCGGAGGGAAAGCTGA
ATTGAGAATCATAATAAATAAAAGACCACATCGTTCTTTTTTTTTTTTTTTTGGGACT
GTATCTTGCTCTGTCACTCAGGCTGCAGTGCAGTGGCACTATCTTGGATCACTGCAGGCT
CCGCCTCCCGGATTCAAGCGATTTTCCTGCCTCAGTGCCTGAGTAGCTGGGATTACAGGC
GTGTGCCACTACACCTGGCTAATTTTTGTAATTTTAGTAGAGACAGGTTTCACCATGTTG
GCCAGGCTGGTCTCAAACTCCTGGCCTCAAGTGATCCACCCGGCTTCCCAAAGTGCTGGG

ATTACAGGCGTGAGCCACTGCGCCCAACCAAGACCACATCCTTTTATTGAACGTTCCTCC
TACCATGTTTTCTTTTTTCTTTCAATTAATCATTGACTCATTGACTCTCACTGTTGATGT
CTGTAGCTGCTCTCTTATTTCCAGTTTTATAGCTGTAAATTTCTCTGTCTTCCTAAGATA
CAAGGTAAATTTCTCTTGCTGATATTGGTGGTTTTGGAAAGTGAGTGGTGTGGATGACTG
CCCAGAAAACAACAGAACACAAAAGCATTCTCTGCCCAGAACACATCACCAAATAGATAC
AAACTCATCTCTTACTGAGTGAAATAGCTTCCTTTTTGGCAGCAAGAATGATTTTCTTGG
TGCCATATTTTTCAATCCGCCTGCTCTTGAAGCCAGCAGCTATTGCAGACTTGGCATTCC
CAGGCACCCAGTTAAGGGAAAGTGACGTGTAGAGGAGGTATCAGATGGGTCTGGATATAG
AAAAAGCAGCTGGTTCAAAACCCCATGGGCTGCCTTTCTGTGATAGAGTTATTCACACTT
GGGTTAGATAAGGCACAGAGTCCTCCTACACTGGTGCGGAAATGAAACAGACAGTCTGGC
TCGTTGGGCAGCCTAGCCTCCTCCAGAATCTGTGCTTGCCTTCCCTATGGAGTGACTGGT
AGATCTTAGAATTCAGACCTCAGTGGTTGCTAGCCAGCACTCTCACATTGGTTGGTCCTT
CTCTCTGCATCTTTGATTCTTTAGAGATAGATAAACCAAGCACCGACTCTCCTTTGACAT
GTGCTTGGAACAGACACCTGCACGAGCTGCCTTTCTCCTCCCACTTCTGCCTGGTCTTCC
AAACACCTGCTTTTCTTGTTTGAACTCTTCCTTTTTTTTTTGAGACAGAACCTCTCTCTGT
CACCCAGGCTGGAGTGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCAGGTT
CAAATAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCTGCTATCACG
CCTGGCTAATTTTTGTATTTTTAGTAGAGACACGGTTTCACCATTTGGCCAGGTTGGTCT
CAAACCTCTGGTCTCAAGTGATCTGCCCGCCTCGGCCACCCGAACTGCTGGGATTACAGG
CATGAGCCACTGCGCCCCAGCTGATTCTTTACAGATAAACAAACATTGACTCTGCTTTGA
CATGTGCTTGGATCAGGTAACTGCACCAGCTGCCTTTCTCCTCCCACTTCTGCCTGGTCC
TCCGAATGCCTGCTTTTCTTATTTGAACTCTTCTGTCCTTTTCTGAAAACCTAACAGATG
CGAAACAGGCCATTTTCCATGTTGGTGGTTATTAAGCAAGACTTGAACATTTGTTTGTTG
CTTGTTTAGGCTTTTATTTCAGAGTTCACAGAATTAACTTTCTTTTTTTCTGATCTCTTC
CAG**AGTAAATGTGTCAAATACTGGCCTGATGAGTATGCTCTAAAAGAATATGGCGTCATG**
CGTGTTAGGAACGTCAAAGAAAGCGCCGCTCATGACTATACGCTAAGAGAACTTAAACTT
TCAAAGGTTGGACAAGTAAGTATATTGTCGTATTCTAGAGACTTTGGGAACTGTTGATGG
TGTGTAGGAATTCAGGGTCTTGCCGTTACTCATGTTTGCATACATGCATGCATTCGCTCA
CTCATTGATTCAGTAGCCATTTATTAGCTTCCTTCTATGTGCCAGGTACAGTTTAAGCAG
TACTGGTACATTGTGAACAAGGCAGGTAGTGTTCCTGCCCTCATCGAGCCTAGGGAGATA
GACAATTTAAAAACAAATAACTGGCCAGGCGCCGTGGCTCAGGCCTGTAATCCCAGCACT
TTGGGAGGCTGAGGTGGGTGGATCGCTTGAGCCGGGGAGTTCGAGACCAGCCCTGGGTGG
GAGACTGGGATAGGGTGACCTGAGTGGCTACAAGGTCTGTTAGGAGGCCTCCGCAGGGGC
CTATGTTGATGGCCTCCTCTCCAAGTATCCACAGACTTCAGCAGTTGTTCTTTTTTGTTC
CTTCCTTTGGAATGGAATATTATATAAAATGGCAGAATAAACTGGAAGAGAAGCAGTAGA
TGTGAGAGGTGCCGGGGGGTGAAGTCTGCAGGATGTGGGGATTGTTTGGCTTTTGGAGGA
GGAAGGAGGGATTCAAGACACATTGTAGAGGTTTGAGTCTGAGCGGACAGTGGTGCTGTG

GCAGACACCACAAAAGCTGGAAGGAGAACTGATGTGGGCAGTGATTTGTTTTCTTCTGGA
TGTGTTCAGCTGGGCATCTGAACAGTCATGTGGACATTCATCTATTCATTCAGAGATATT
TGTTCAATGACCTCTTGGTTCCTGGCACCATGCTGCTTGCTGGAGATAGAGCTGGGGAAC
AAAACAGATGGAATCCCTGCACTCCCAAGTGTACACTATACTGGCCAGTAATCTACCAGC
CCAGTAATTGCACATATAAATATATCATTATAAACTGTAATCAGGGCTAGAAAGAAAAAA
TGCAGGAGTTTAGGGTTCATTTGGAGGGGGAAGGGACTTTTTTTTTTTTTTTTTTGAAAC
AGAATCTTGTTCTGTCACCCAGACTGGAGTGCACTGGTGCATTCACGGCTCACTGCAGCC
ACAACCTCCTAAGCTCAAGTGATCCTCTCACCTCAGCCTCCCATGTAGCTGGGGGCTACA
GGTGTGTGCCACCATGCCCACCCAATTGTTAAATTTTTTATAGAGACGGTTGTCTCATTA
TGTTGCCCAGGCTGGTCTTGAACTCCTGGGCTTAAGCGATCCTGCTGCCACATGCAGCCT
CCCAAGGTGCTGGAATTACAGGCGTGAGCCAGCGCACCCGGCCAAGGGAGGGGAGGTTCT
TAAGGCATAGGGAACAATGTGTTTGAGTCAGCAAAGGAGGTTGTGGGGGTTTGTCCTAAG
TGTGGTAAGCAGCCAGAGTTGGATTTAAGTTTTTAAGAGATTCCCCTCCACCCTGTAGAG
ACTGGAGGGGGCAGGAGTTGTTCTAGGGATTAGGACCAATTTGGAGGTAGTGCAGCCGTC
AGAGTAAAAAATAATAGGGATTGAACTAGGCCAGTGCCCAGGGTGCCTGAAAGAAGAGGA
CCCAGTAGAGCTGACTGGAGGCAGACATGCAGGGATTCAGTGAAGGAGTGTACCAAGGGC
GAGGGTGGTGTGCAGGGTGACTGGCAATTTTCTAGCTTGAGAAAGGTCCGGGGGGATGGC
AGTGGAGTTGAGGAAGCTGGGAGGATCAAGGACCTTTTTGTGAACACACAAAGTTTGAGA
TGCCTTGGACACATTGAAGTGGAGCGGTCAGGGAGGCAAGGGTGGAGGTGGGATGCGGAG
GGGAGGTGGGATGCAGAGCGTCGTGGATGGATCAGTTTTGCTCGATAGAGGGACATGTTT
TTCTGTGGCAACAGGAGGGCAAAAGGAGAAGGTGGCCACAGATGCCGGTAGATGAGCTGA
GAGTGATTGTATTCCCTATCCTCTCGGAAGCTTGAGGCAAGGCCATCAACAGACAATCAG
AGGGAATAAGAAGAGATAGAATATATGAAGAAAGGGAGAAAAGATGAAATCGTAATTGTG
TAGCAGGGCAAGAAGTCCAGAAATTTCTGTGCTGTGCCAAGTTCCCAGTTGAGGCGGTGA
ACATGAAAATATACTGATACCCATTGCCTGGTTTTTCTCCAAGGACACTTGGCTCCTAGG
GCACAAAACAGAAAGTACGTGGTTTGTCCAGGCCGAGGGCTTTGCATAGTTGCAGTGGAT
GGAGAGGAGGTCAAGGAATGGAGGCACATGGTAGAGAGAGACTGTCCCCAGAGCACGGGG
ACTCCTGGCCGGATGAGGGGGACAGGGGCAGGAGGAGGCAGGTGGAAAGTAGAGGGAGGG
CTCAGTGGTCTGGAGGCTACAGGAAGTGACGGGGGGACCAGAAGGAGCTGGAAACCAGTG
TGGTTGTGGCCCAGGGTGGGATGTTTGGATTTCTGATGTCAGAGAGGGTCCAGTCCTTCT
GATGATGGGGAGGGGTGGAGGCTGAATCTATGGTAGAGATAGTGAGAGGAACTGGAACAA
TGTAGCTGTCAAGTGGAAATGGGAGAAAGGGCTGGGCGTGGTGGCTCACGCCTGTAATCC
CAGCATATTGGGAGGCTGAGGCAAGAGGATCGTGTTAGCTCAGGAGTTCTGGGCTGCATT
GAGCTGTGATTGTGCCACTGCACTCCAGCCTTGGCAACAGAGTGCCCAGTTAAAAATAAA
AATAAAATAAAATAAAAAAATTAAAAAAAAAAGAAGAAGAAAAAAGAGAAAAGTGTCCTT
TTACATCCCTTTTAAAAATGTCACTTAAGGCTGGGCAAAGTGGCTCATGCCTGTAATCCC
TGCACTTTGGGAGGCTGAAGTGGGTGGATTACTTGAGGTCAGGAGTACAAGACCAGCCTG

GCCAACATGGCGAAACTCCTTCTCTACTAAAATTAGCTGGATGTGGTACATGCCTGTAGT
CCCAGCTACTCGGGAGTCGAGTCTGAGGCCCAAGAATTGCTTGAATCGGGGAGGCGTAGG
TTGCAGTGAGCTGTGATCAGGTCACTGTGCACCAGCCTGGATGACAGAGTGAGACTCTGT
CTCAAAAAAAAAAGTCACTTAGCTTAGATTGTCTCTACATATATAGGAAGAAGATGTAGG
AATGAATGGTGCTGCTACAATTACGTCATCTGGATAGACCCAGAAACATGATACTTTTTG
GTTTTCTGTAGCCTTGGTGCCATTGTTGATCTTTATTAATTATCATTATCCTCAAAATAG
CCATAATGTGCTGAGTCTCTTCCTATTTGCTGGGCAGAGGCTGAGTATTTCAGCGAGCTC
ACTGAGTCCTTAAAATTGCATTATGATAGAGAGAAAGAGATTATTATTTGCATTTTGCAA
AATGAAGAAATTGAGGTTTAGAGATACCCAAGGGCCACGTGAGTGTGAGTGCCTGGAATT
GGAGCCTAAATCTAGTCATCTGATAGCAAAGCCTGTTTTCTTATCTGCTTTGCATTAAAT
ATAAGTTTAAAATAGAACAATACTGGCCAGGCTGGGTGGCTCACGCCTGTAATCCCAGCA
CTTTGGGAGGTCGAGGCAGGCAGATCACCTGAGGTCAGGAGTTTGCAACCAGCCTGGCCA
ATATGGCGAAAGAAACCCCATCGCTACTAAAAATACAAAAATTAGCCAGGCATGGTGATG
TGTGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCTTGAACCCGGGAG
GCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGTAAG
ACTCTGTCTTGGAAAAAAAAAAAAAAAAGAATGATACTATAGTCTGTGTTTATATGGTGG
GGAAGGTTGAGTATCAAAAAAATAACAAAGAGGAATGAATGTCTTAAGTGAATGCCTGTT
TCCCCATCTGCTTCCTCTTCTGCTGGGAGGAGAGACCTGGATCCCTAGAGGTTTCAGTTG
CCTCCAGAGCTGAGTGCCACAGGGATGCAGGGGAATAGGGATGTTACCTGTCGCTGGTAA
TTCAGAGAGATGATTCAGGGTATAGTTACCTGAAAGAACAAATTGCCATGCCAGACGTCT
TGGTTCTTATGACAGAGGCAAAGAGTTGCCTCCAGGATTGCCCAAAAGGAGACGAGTTCT
GGGAACCTCACGAAGAGGACCTTTCAGTGGAACCTGGGGAGATTCTCTTCCTCTCCATTG
GATTTAGGAAAGCTTAGAACCGGGTGATTCCTCAACCTCTTGATTTATTTAATTCTTTTC
TGG**TTTTTCTTGGCTCTACTCCAGGGGAATACGGAGAGAACGGTCTGGCAATACCACTTT**
CGGACCTGGCCGGACCACGGCGTGCCCAGCGACCCTGGGGGCGTGCTGGACTTCCTGGAG
GAGGTGCACCATAAGCAGGAGAGCATCATGGATGCAGGGCCGGTCGTGGTGCACTGCAGG
TGACAGCTCCTGCTGCCCCTCTAGGCCACAGCCTGTCCCTGTCTCCTAGCGCCCAGGGCT
TGCTTTTACCTACCCACTCCTAGCTCTTTAACTGTAGGAAGAATTTAATATCTGTTTGAG
GCATAGAGCAACTGCATTGAGGGACATTTTGATCCCAAGGCATATTTCTCCTAGACCCTA
CAGCACTGCCATTGGCCATGGCCATGGCAACATGCTCAGTTAAAACAGCAAAGACTAAGT
CAGCATTATCTCTGAGTCCACCAGAAGTTGTGCATTAAACAACTTCATCCTGGCTCTGCA
GTTTCTCCTTATTCTTCATGATGTTTGCTTTGTAGCTGTTGACTGCTTTGTAGGTATTGA
GGTGGTGGGGGTGTGGTGGAAATAGGCCTGACTCTTGAGGATCCCTTAAGTCATTTTTGC
TTGGTTCTCTTTTTCCTTCTTTTCTTCTACTCTTCTATGATTCATCTCTTTGATTGTGAT
TCTGTTCTCTCTCTCTCTCTCTCTTTTTTTTTTTTCGTTTTTGAGACAGAGTCTTGTTTT
GTTGCCCAGGCTAGAGTGCAGTGGTGCCATCTTGGCTCACTGCAACCTCCGCCTCCCGGG
TTCAGGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATCTGACACTA

CGCCCGGCTAATTTTTGTATTTTAATAGAGACAAGGTTTTGTCATGTTGGCCAGGCTGGT
CTCGAACCCTTGACCTCAGGTGATCCACCTGCCTTGTCCTTCCAAAGTGCTGGGATTACA
GGTATGAGCTACCATGCCCGGCCCATTCTGTTCTCTTCTACCATAAATATATTTCTCCCC
TAACACTATATTTGTTTGCTTCACAAGATTCCAGCTGCTTTTCCACCAAGGCCTTTGATG
GAAGCTGTGCTGTGACCTCTGTAATGAGTCTGTGGGCTGCTGATTCTCCAGTTTGGGCTT
CATGATTATACTGGGGAATATTGGGTTTCCTAAATCTCATTCATTTCTTGGGCAAGTAGA
TATATGTGAAAGTGTTTATTTGTCCAGTTGTTAAAGAAGCTACCATTTATTGAGCCAGCC
TCTGAGCACAATGTTTTTTGTTTTGTTTTGTTTTTAATTTTTAAAATTATTTACTTCTTC
TATTTCAATAACTTTATTATTATTATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGC
TAGAGTGCAATTGAGCGATCTTAGCTCACTGCAACCTCTGCTTTCTGGGTTCAAGCAATT
CTCATGTCTCAGCCTCCCGAGTAGCTGGGATTACTGGTACGTGACAACATGCCTGGCTAA
TTTTTGTGTTTTTTAGTAGAGACGAGGTTTTGCTATGTTGGCCAGGCTGGTCTGGAACTCC
TGGCCCCAAGTGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGTATTATAGGTGAGAGCC
ACTGCGCCCGGCCCTCTTTCAGTAATTTTGATGTATTTTTTTGTATATGATTCCTGTTTC
ATTCTGTCCAACCAGCACTCTGTATGGTATGTGCTGTTGTCCCCATTTCACAGATGCAGA
AATTAAGGGTCAGAGAGGTTAAGGGACTTACCTCAGGCACGTTGTACTGGAGAAGCTGAA
CTCCAAGAGCAGGTTTGGGCTGACTCCAAAGCCCTATGCTTTTTGCCAACATATTTTCAA
ACATAAATAGACAATTTTATAAATAGCTCCAAAGAGTAGACATTGTTTCTGTTGATATTA
ATGGCTTGGTTTTGAGTCTGAAACCCCCATGAATGATTCTGTTGTCCCTGCTTTTTGTCC
TTCTGCCCGCAG**TGCTGGAATTGGCCGGACAGGGACGTTCATTGTGATTGATATTCTTAT**
TGACATCATCAGAGAGAAAGGTGGGTCATCTGGTGGGCAAGAAGCGACAGTTTCTGTTTT
TAGTTTATGGAAGGAAAGTGCTCACGAAAACAGTCTGGGGAAGAGAGGTTGAATGGGAAA
ATTCTTTCACAAAAATCTGGGCTGAAGACTTCAGTGTGTCTGCCTGAGAACAGAAGTGAC
ACTATTTGAGCTTTTGGCATAAAATGAAGTCTAGGAGCTGCAGAACCCACTGCCATGGCC
TTTTGTTGCATACACAGTGGTGGTCTCTATCCAGCCACCTGACCTTGTTTACAGTATGGG
GTGATTTGTTGGCAAGTGAGGGAATCCTGACTTCTGCCACTTCGTTATTTATGTAGTCTT
CTGGGATCATTGGTATTGGTCAGAAGTTCAACACTGTAGCCATTGCAACATGCTCAGTTA
AAACAGCAAAGACTAAATTAGCATTGTCTCTGAGTCCACTAAAAGTTGTGCATTAAACAA
CTTCATCCTGGCTCTGCAGTTTCTCTTTATTCTTCATGATGTTTCCTTCGTAG**GTGTTGA**
CTGCGATATTGACGTTCCCAAAACCATCCAGATGGTGCGGTCTCAGAGGTCAGGGATGGT
CCAGACAGAAGCACAGTACCGATTTATCTATATGGCGGTCCAGCATTATATTGAAACACT
ACAGCGCAGGATTGAAGAAGAGCAGGTACCAGCCTGAGGGCTGGCATGCGGATTCTCATT
CTCTTGCTAGGCCTCTTGGATACGCTCTCCTTTTGAGCAGGAGGACAGGCTCTGATAGAC
AACTGTTTGATTTCGGAATGGGAAACAAACTCCCAACTAAAAGGGCCTCTGGAAACTGTC
AATTATTCTCCACTTCTCAGCTCTGATTTTTCACTGCAGAGGAGCTTAGGGAAGGGCACC
ATCCTATCAGCCTGGCCTGCCAGATTGAAGAACTGCCATGCAGAAAGGTTCTGATGTTCT
CAGGCTCATGTGGCAAGCGTAAAACTCAAAGCCTTGAAGTTTCTAGCCTGTTCCAGCCTT

GATCCAGGCCATGTTTATCCTGATTCCATCCTTTAAAACGAATGCCTCACTCTTAATAGC
GCACGGCAGTTTGAACCACTAATTTGGTCGAGTTGGAAACAGTGAAATTTCAATTTTAAT
AAGCTGTGCATAATGAAGAGGAATGTGGAATTGGAGCCTTTCCATCTGAAGCTATTCATA
ACAGGCACAAAGCTGAGTTAATTAGGAATATGCTGAGATGAAGGAAATGAGGAGAGCTGC
TCTTTTGGGGGCTGTGCTTCTCTCCCCAACCCCTCAACCCCATTGCCATGCTGCAGATGG
GGTGGTGTCTAAACATCAGTGGCGAGTGCCTGCATTACTCTGCTCGTTGCCTTCCAGAGA
ACTCAGCTTCTCCAAATGCTGAGCTCTTTTCAGAATGGGACCTGCCACCAGTATTTGAAA
GATTTCTAGCCTAGCAGAACAGCAGCCACGTTATCAAAGTTTGGTTGGCCAAAGGAAGGT
ACTTGCTAATTAGTTTAGTAGGTTTTCAGTCCGCACAGACATACGGGATTGTTTTATTGT
ACATAGACATCTTCAGAAACAGTGTATGTATAGAAATGTAAGGTCAAAATTTGAACCTCA
GTGCTTTAAATCTGAATTTGTATTAACTGATATGAAATATTTAGACGGTTACTTTATTTT
ATATCTGTCTTCCATTATACTTAATTTGGCTCAAGAATAGTTAGGCAAAAAGTTGCCCAA
AGAGAAGGATCTCCTAGTAAATACAAAGAGAATGTAACATAGTTGCTACAAGTTGGAGCA
TGTTCAGGGATGTCTTTTTTTTTTTTTTTTTTGAGAGAGAGGTCTCTCTCTGTTGCCCA
GGCTGGAGTGCAGTGGTGTAATCATGGCTCACTGCAGCCTCAATCTCCCAGGCTTAAGCG
ATCCTCCCACCTCAGCCTCCCAAGTAGCTGGGACTATAGGCATGCGCCACCACACCTAGC
TAATTTTCGCATTTTTTGTAGTGTCACAGTTTCGCCATGTTGCCCAGGCTAGTCTCGAAT
TCCTAGGCTCAAGCAGTGCTTCTGCCTCAGCCTCTCTGAGTAGTTAGGACTACAAATTTG
TGGCTCCATGCCCGGCTAATTTTTTTATCTTTATTTTGTAGAGACAAGGTCTCACTGTGT
TGCCCAGGCTAGTCTTGAACTCCTGGGCTCAAACAACCCTCCCACTTTGGGTTTCCAAAG
TGCTGGGATTACAAGTGTGAGCCACTGAGCCCAGTGACCTCTGGGTTTTAAAAATGTGTA
GGCTTCAATTATTTATTTTAAAAAATGAAATCCTGCAATATATAGTTTTCTGCGTTGTGT
GGTTTGAATCAATCTGGGAACTGGCTTGCTGGCTGATTGTGGTAAAGTAAGAAGTACTTA
ATTTAGTAGAAAGTTTAAATGGCAGACATAACATTAAACCCAGCTGATTTATAAATGAAG
CAAAAGAACAAAACTCATTCAGGATAATTGGTTATTCTAAAATACAGTCATTTCTAAAAT
TATGAAGTGTTCAGGACCTTTGGGAGTGAAAGAATTTGCTAAAGAAGGATCAGTGAAAAA
AAGGAATGATGGGTGAAGAGCTGTGGAGAAGGAAGAGAAGAAACAGCACAAGGAAGGAAG
AATATAAAATCAGATGTGGGAATCCAGGGGAAAGTGCAAACGAAGCAAGATTGAGAAAAT
TCTCAAGTTTTTATAAACAGTTCTCACACTCTGCCAGTTCCTTGGAGGTAGACTTTTTTG
TTAACTTCCAACTACAGTAGTGAAAAAAAAAAAAAAACCCTCAAATTTGCAAAAGCAGTC
TGTGGAATTTTCTTTACCCAGCTTTCCTGACTGTTAACTTTTTAGCACACTTAACTTTAT
CATTCGTTTATTCTCTCTGTTTAAAATTAAAAATGTAAATTTTAAAAAGTAAAATGTTTG
TTGGTTACAAACATTTATACCCCTTTGTCTCTAAATATCATTTCATTTTAAAAAATGAAT
AATCTAAGCCTACACATTCTAAAATGTGTATATTTTCTAAAAATAAGGGCATTCTCTTAC
ATAACCAATGTCACAATTATTTGATACAGTGATCAAAATCAGGAAACTAACATTGATATA
ACACTATTATCTAACCTACAGACCATCTTCAAATTTTGTCCTGCTAGTATCTTTTATGGG
TCCAGGGTCACACAGTGCATTTGGCTATAATGTATCTTTTTTCTCTTTTTTTGAGACAGG

GTCTCACTTTGTTGCCCAGGTTGGAGTGCAGTGGTGCAATTATGGCTCACGGCAGCCTTG
ACCTCCTTGGGCTCAGGTGATCCTCCCACCTCAGCCTCTCGAGTAGCTGGAGACCACAGG
TGTGCACCACCATGCCTGGCTAAGTTTTGTATTTTTTGTAGAGATGGAGCTTCGCCGTGT
TGCCCCGGCTGGCCTTGAACTCCTGGGCTCAAGTGACCCTCCCGCCTTGGCCTCCCAAAG
TGCTGGGATTACAGGCGTGAGTCACCACACCTGGCCAGTTATTAGTATGTTTAGTCTCTT
TAATCTGGAACAGTTTCTCAGTCATTCTTTATTTTTCATGACCTGGATGTTTTTGAAGAG
TTTAGGCCAGCTATTTAGCAGAATGCCTTTCAGTTTGGATTTGTCCAGTGTTTTCTCTTG
ACTATATTCTAGTCATGCATTTTTGGCAGGACTGTCACAGAAATGTTGTTGTAGTCTTCT
TAGTACATCACATCAGGTACACACTGTTGATCTGATTCATTACTAGTGGTGTTAACTTTG
ATCACTTGAATAAGGTGGTGTCTGTCAAATTTGTCCACCGTAAAGTTACTTGAGCAAAAC
GTAGCTGGGACTACAGGCGTAGCAAAAAATGTAGCAAAAAGTAGTATTTTTGCTACATTT
TTTTTTTAGGAACAAAGTATTTTTCCCTTTTAAGTTAATCTCTTGTCCATAAAGTTATTA
TTTTTCCCTTTTAAGTTAATATCTTGTGGGTAGATACTGGAGACTGCGTAAATTACCTAT
TTCTCATAATACTTTTTTTTTTTTTGAGATGGAGTCTCGCACCGTCTCCCAGGCTGGAGT
GCAGTGGTGCAATCTCGGGTCACTGCAAGCTCCACCTCCCGGGTTGACGCCATTCTCCTG
CCTCAGCCTCCCAAGTAGTTGGGACTACAGGCGCCCGCCATCACACCTGGCTAATTTTTT
GTATTTTTAGTAGAGACGGGGTCTCACCGTGTTAGCCAGGATGGTCTTGATCTCCTGACC
TTGTGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGTCACTGCGC
CCGGCTCTCATAATACTTTTTGCCTACTAATTTTATATTCATTGATTAAATTCTTGCCTG
AAAAAATTATTACTGTGGTATTTGCCAAATGGCAATTTTCTGTTTCCATCATTGCCTTTC
CCCCGCTTTTAAAAGTATAAGTGACAAAGAAAAACTGTATATAAAGTGTACACCATGATA
TTTTGATATATGTATACTTTGTGAAATGATTATCAAAATTGAGTTAAATAATGCATCCAA
CATCTCAGTTACTTTTTTTTTTTTTTGAGACAGAGTCTTGGTTTGTCACTAAGGCTGGAG
TGCAGTGCCACAATCTCGGCTCATTACAACCTCCACCTCCCAGGTTCAAGTGATTCTCCT
GCCTTGGCCTCCCCAGTAGCTGGGATTACAGGTGCCCACCATCACACCCGGCTAATTTTT
GTATTTTTAGTAGAGGTGGGGTTTCACTACGTTGGCCAGGCTGGTCTCGAACTCCTGACC
TCAAATGATCCTCCCGTCTCAGCTTTCCAAAGTGGTGGGATTACAGGCGTGAGCCACTGT
GCCCGGCCACTCTTAGTAAATTTTAAGTGTACATTTTTTTTTTTTTTTTTTGAGATGGA
GTCTCACTTTGTCACCCTGGCTGGAGTGCAGTGGCATGATCTTGCCACACTGGAACCTCT
GCCTCCTGGGTTCATTCAGGTGCTTCTCCCACCTCAGCCTCCCAAGTAGCTGAGACTACA
GGTACCCGCCACCATGCCTGGCTAATTATTGTATTTTTAGTAGAGATGGGGGTTCACCAT
GTTAGCCAGGCTGGCCTCAAACTCCTGACCTCAGGTGATCTACCCACCTCGGCCTCCCAA
AGTACTGAGATTACAGGCATGAGCCACCACACCCAGCCACATTACGTTAGTATTAACTAT
AATCACCATGCTGTACATTAGATCTCCAAAATGTATTCATCTTATGTAACTTCAAGTTTG
TACCCTTTGACCAAAGTCTCCTTGTTTTCCCTACCCCCAACCCCTGGTAATCACTGCTTT
AATCTCAGTTTTTATGAGTTTGACTGGTTTAGATTCCACATACAAATGAGATCAGGCAGT
GATGGTTTATTTCACTTAGCATAATGTCATCCATGTTCTTGCAAATGACAGGATTTTCTT

CTTTTTAAAACTAATATCCATGCTGGACACGGTGGCTCATGCCTGTAATCCCAGCACTTT
GGAAGGCTGAGGAGGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACAT
GGTGAAACCCCATCTCTACCAAAAATATAAAAAATTAGCTGGATGTGGTGGCGCACACCT
GTGATCCCAGCTACTTGGGACACTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCGGAGG
TTGCAGTGAGCCAAGATGGTGCCACTGCACTTTAGCCTGGATGTTGATGTTGTTCCACTT
GTTTATTTTTATTTTGTTCCCTGTGCTTTTGGTATCAAATCCTAAAAACCATTGCCATGA
CCATTGTCATGTTACTTTCCCCATATGCTTTCTTCTAGAACTTTTAAGGTTCATCATTCC
CTTTTCTGTTTTTAGTTGCAAGCCTACTATAAGGAAGGGCTTTTCTTTCTTCCTTATTTA
TTTATTCATGTCTATCAGAATGGGCACCTTACTACTATTTTTGTTGTTATTGCTTGAATT
GACTTGAATTTGGCTAGTGGAAACCTTTTCAGATCGGGTACTCTGTCCTTTTGATCTCTT
TCCATTTTCAAGCACTTCTTTAGACTTAAGATGGTCTAGGCTCATCTTCTCCTTTCCCAG
CCATTTTTCAAAGGAACCTGATTCCTTTTAGTGAAGAGCAGTATTTTGAAACCAAGATCT
GGGCACTGGGTCTACTTGTTTGTACTGGTACAGTGTTCTTTGAATTGCTAATTAGCTGAT
CAATTACTGCTCTATTTGAGTTCCCTCTTTCTAAAACCTCACATATGTGTACAGACGGTC
CCTGACTTATGATGGTTCGACTTATGATTTTTGATTTTATGATGGTTTGAGAGCAATACA
TCCATTCTGTTTTTCACTTTTCATTCAACACTTTATTTTAAAATAGGGATTGTGAGATGA
TATTGCCCACGTGTAGGCTAATGTAAGTGTTCTGAGCACGTTTAAAGTAGGCTAGGCTAA
GCTGTGGTGTTTGGTAGGTTAGATATGTTAAATGCATTTTCGACTAGTGATATTTTCAAC
TTATGATGAGTTTATTGGGATGTATCCCCATAAAGTCGAGGAGCATTATACATATCTCTG
TATAACAGAGTGAGTTCCTTATACCTTTCATCCACTTTCCCCTGAAGTTAACATTTTACC
TAACCATGATACATTTATCAAAACTAAAACATTAACATCAATACATTGCTATTAACTAAA
CTAGAGTTTAATTGGATTTTGCCAGTTTTCCAATGAATATCCTTTTTCTGTTCCTTGATC
CAATTCATGGTCACACACTGAGTTTGGTCACTTGTCACTGTAGTCTTCTCCAATCTGCGA
CAGCTTCTTAGGCTTTCCTTGTTTTTCATGTACTCTTGACGATTTTTAAGAGTACTGGTC
AGATATCTTGTAGGATATCCCACAACTTGTGTTTAATCTTATGTTTTCTCATGATTAGAC
TTGAGTAATGGATTTTTGGGAAGAATACCACAGAGGTATATTGTTAAGTGTTCTCATCAC
TTGGAGGTAAATGTTATCAACATGGCCTGGTGATGTTAAACTTGTCAGTTTGTTTAGTTA
GTATCTGCCAGATTTTTCTCACTGCATAATTACAAATCCTCCTTAACTTATGATGGGGTT
ACAGCCTGATAAGCCCATCATAAATTGAAAATATCATAAGTCAAAAATGCATTTAATGCA
TCTAAACTACTAAACATCACAGCTTAGCCTAGCCTGCCTTGAACGTATTCAGGACACTTA
CATTAGCCTACAGTTGGGCAAAATCATCTCATGGGAAGCCTGTTTTATAATGTGTTGCAT
ATCTTATGTAATGTGTTGAGTACTGTACTCAGAATGAAAAACAGAAGGGTTGTATTGCTT
TTGCACCATCATAAAATCAAAAAAACCATAAGGCAAACCATCATGAAGTTGGGGACTGCC
TGTACTTTTTTCCTCTTTCCCTGTTCAATTCCTTGGAAGAAAGTCATTTAGTTCAGACCA
TACTCAAGAAAAGGGAAATAAAGCTCCATCTCTTGGAGCTTAATTGAAACTGGAATGACT
AGTTTCTATATACATTATTTAGAATCCTTTTGTAAGAAAGATTTGTTCCTTCTCTCCATT
TATTTATTCCATTATTTATATTGATAGAGACGCATGTACATTTATTTTATACTTTGGGTT

ATAATCTATTTTTCTTGCTCAAATTGTTACAGCTTTGGTCACTGGGAGGTTCTTCAGATT
GGCTCCTGTGTCATTTGACATGTCCCCACCCTCTCGTTTCTGAGTACTTCTCTACTTTGG
CATTACAAAAGATGTTCCAGGCTCCTCTTATATTTTTCCCTGCCGCAGCCCTAGAATCAT
CCATTTTTCTATGGTGCCCTGGTTCCTTTTACTTTAGATGGGGGTTTAGAAACCAATCTG
GGTGTTGGGTGTGCTCATTGCTACTGGAATCACTGCTTCTAGGCCCTCTCAGCAGATAGA
GCTAGAAAACATATGGCTGTATATGAATCCATGGATTCATATATATCTATAATTGTTTTC
TGTATCTGGCCATCTATATATATATTAAGCTAAACATGAATTCATACTGATGTCTCAGAC
TCGAATCCATTGCCGCAGGGCTCATTCTTGCCTTCCTCTTGCTTATTTGTGACTTCTTTC
TCTAACAGGGAGAAACCCCAGTCTCATTATCACCAACCTATCTACTCATTTGTTCAACCC
TGGTATAGGTGTAAAGTAGTTTCAGAATTACTAACCTATACCCATGTGAGAATTGTATTT
GCACTTCTTGTTTGAAGGAAATACATACAACACAGGTAGCGTCTCTACACTTCAGTATAC
AGAGATCTGAACAGTGTTCTCTCTGAGTGAATCATATTGCAGGACAGAAATTACTTTTAA
AAATTCTGTAATGGGTCAGGCCTATAATCCTAGCACTTTGGGAGGCTGAGGTGGGCAGAT
CACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAAAATGGTAAAACCCCATCTCTACAA
AAAATACAAAAATTAGCCAGGCGTAGTGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGG
CTGAGGCACGAGAATCACTTGAACCTGGGAGGCAGAGCTTGCAGTGAGCTGAGATTGAGC
CACTGCACTCCAGTCTGGGCGACAGAGCGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAA
AATTCCATAATGATAGCAGAGCTGGAATAGAAATGGGATTGCACAGGCTGAATCTGAGTT
GTTGCAACAGTAAACGAGCAAGATTTAAACTGGCCTTGTGTAGCACTTGCTATTTGGCTC
CTCATATTTTATTAGACGCTTATTCTTTTTTGTTTGGTGTCATTCCTTTGAGAAATATTT
GAGTGCCTTTTCTGTTGCAGACATTGATTAGATGCTGAGGTTGTAACAATGAAGAAGATA
GCCATCGCTGTTGCCTCATGGAACTGAAGTTTTACTAGATGTAAAATTTGAGTTAACATG
AGGCCGTGCCCCTATGTGCCCTATTGTTTCTTCACACAGCTCCCTTCATCTCCTTGGTCC
AATGAAAAGGTTTTTTCATACTTGTTCATTCATTCCTGCATTAATTAAAGTAGGTTGTAC
TGTGCCAGGCACTGGGAATATTTAAGTAGTTGTGTTCCTGAATTGGAAATGAATCCAGCA
TGGTTGGAGTAGAAGGAGCTGGGGGGCAATGTGGAGTGTGATGGGGAGATTGGAAAAGTA
AGCTGAGACCAGATTTTTCAGTTTGGAGGGAGAGGTGGGCCTTGTAGGCCATATTACAGA
TTGTAGACTTTATTTGGAGGGACATGGAAGTCATTGAGGAGTCTGAAGCAGGGGAATGAC
ATAAAAAGATCCTCATTTTAGGCCGGATGTGGTGGCTCACGCCTGTAATCCCAGCACTTT
GGGAGGTTGAAGTGGGTGGATTGCTTGAGGCCAAGAGTTTGAGACTAGCCTGGGCAACAT
GGTGAAACCCTGTCTCTATCAAAAATACAAAAATTAGCTGGGCATGGTGGCTCACACCTG
TAGTCCCAGCTACTTGGGAGGCTGAGGCATGAGAATCGCTTGAACCCGGGAGGCAGAGAT
TGCAGTGAGCCGAGATTGTGCCACTGCATTCCAGCCTGGGTGACAGAGTGAGACTTCGTG
TCAAAAAAAAAACAAAAAACCCCTCATTTTGAAAGGGAACCCTGGCTTGAGGGTGAAGAA
TGGGTGGGCACTAGGCTAGAGCAGCTGCAGGGTCAGTGAGGAGCTGCCGCAGTGCTGCAC
GTGAGAACCCGTCATGGTTTGGTCAGGGTGGGCAGGACTGACAGTGAGCACAGAGCGAAG
TAAAACCAGCAAAATTTCATGATTGGATAGTGGAAGGAATCATGGTGTTTGTAGTCTTCA

AATGTGAACCCAGAGTGCACTGGACAAGTAGTCTAGGCTGCTCTGTAACCAAGGCAAGTG
TTTTCATTTTACCCTCTCTTCCTGCTCTTGGCCTTTGGATTTTTTGTAATTTAAGGTTTA
TGAATGTAATCAGTTACTTAACATGGAAAGATACTTAATACCAGATGATTTTGGAGTCTT
GTGATCAATACCTTCTCTCAATCTTGGGTGTGTGTCAGTTGGCAAGGCCATAAAATTTGT
TATAAACATTGCAGAAGGCTTGGTTACTGTGCTGTGACGTTGAATTTGGGTGGAGATAGA
TCAATTTCAGTTGATTTTCTAGGCTTCAGAAACACATTACCCTCTACTCCACAAACACAA
ATCAAAACAAAACAATCCCTATTCCCTGAGCATTTCTCTTGATCTATAACACAGCCTGGG
CTGTCACAGTACTAAGACAAGCCCATCTGATTTGTGAGTCAGTTTTATTTCTTGGTCTTC
TACATAAGCTAAAAAGTTTCAACATTTTAATGCTTTTCCTTGGATTCCTTTGAGTCATTG
AAGTAATTCCTGTTTCATTTGTACTAATTATTCCACACTAGAAAATTCTGTTGTAATCAC
TTTATGTATTAATAGAAATACTGATTTTTATTTTCAAGGAAGTATTGAGTAGGGAGGGGG
AAATAGGGATTTGCTGTTCAATGGGTATAGAGTTTCAGTAATACAAGACAAAAAACTTCA
GAGATCTTCTATACAGCAGTGGGTATATAGTTAACAATACTGCACATCTAACAGTTTGTT
AAGAGGGTAGATCTCATGTCATGTGTTTTTAAAAATTGCTTTTAAAAAAAGTATCGAGTA
AAAAAGCAGTTTTACTCCTCAGTTTCTATTTATATTTAAAATTTTTATTTAAAAAGTGAG
TTGAGATTTTTAAACCTCAGGATAAGTTTTATTTTTTAAAAAATTTATTTTTTATTATTT
TTTGAGATGGAGTCTCACTCCATCTCAAGTCACCCAGGCTGGAGTGCAGTGGTGTCTTGG
CTCACTGCGACCTCTATCTCCCAGGTTCAAGTGTTTCTGCTGCTTCAGCCTCCTGAGTAG
CTGGGATTACAGGTCTGCACCACCACGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGG
GGTGTCACCATGTTGGCCAGGTTTGTCTTGAACTCCTAACCTCAAGTGACCACCTGCCTT
GGCCTCTCAAAGTGCTGGGATTACAGGTATGAGCCACAGTGCCCGGCGGGATAAGTTTTA
AAATAATATTCTCTGCTGGCTGGGCATGGTGGCTCATGCCTGTAAACCCAGCACTTTGGG
AGGCTGAGGCAGGAGCATCACTCGAGGCCAAGAGTTTGAGACCAGTCTGGGCAACATAAT
GAGACCCCCTCTCTACAAAAAATAAAAAAAATTTGGCTGAGTGTGGCATGTTCCTGTAGC
TATCGGGAGGCTGAGATGGGAGGATTGCTTGAGCCCAGGAGTTTGAGGCTGCAGTGAGCT
ATGATTGCACCACTGCGCTCTAGTCTGGGTGACAGTGTGAGACCCTGTCTCTTAAAAAAA
AAAAAAAAAAAGGCCAGGCACAGTGGCTCAGGCCTGTAACCCCAGCACTTTGGGAGGCCG
AGGCGGGTGGATCACTTGAGGCCAGGAATTTGAGACCAGGCTGGCCAACATGATGAAACC
CCGTCTCTACTAAAAATACAAAAATAAGCTGGGTGTTGTGGTGCACACCTGTAATCCCAG
CTACTTGGGAGGCTGAGGGAGAGAATTGCTTGAACCTGGGAGGCAGAGGCTACAGTGAGC
CGAGATCACACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAACAA
CAACAACAAAAAAACCAAATGTTCTTGCCAATTCTTCCATTTAATATTTAATTTTGAATT
ATATTGTATCTTTCTAAGGATTGTTTCTTATATAAGCAAAGATTTTTCAGTGCTAAACAT
TTACGACTGCTATTCAGAAATGGTTATTTACAAGTCTTTTTGTTTTAAGAAAATGGCTGT
TCAAAAAATTAAAATAGTATATAAACCAAACAAAATATTTTTGCTTTGGATGTCTGTTTT
GCAGCTTCTTCCCTACACTATAAGTTCTTACTGACTGCTTTATCACTTAATAAATTGGTT
TGGCTACTTTAACAGAGGCAAATAGTATCAGGCAAAAAATTATTTTTTATTTTTATTTTT

TGAGACAGTCTCACTCCATCACCCAGGCTGCAGTGCAGTGGCCTGATCTTGGCTCACTGC
AACCTCCACCTCCCAGGTTCAAGCGATTCTCATGCCTCAGCCTCCTGAGTAGCTGGAATT
ATAGGCATGCACCACCACACTCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTGC
CATGTTGACCAGGCTAGTCTTGAACTCCTGACCTCAAGTGATCCATCTGCTTTGGCCTCC
CAAAGTGCTGGGATAACAGGCATGAGCCACCATGCCCAGCCCTATTTTTTATTTTTTAGA
GATGGGTCTCGCTTTTTAGAGATGGGTCTTGTTGCCCAGGCCAGAGTGCAGTGGTGCGAT
CATAGCTTACTGCAGCCTTGAATTCCTGGGCTCAAGCAATTCTCCTGCCTCAGCCTCCCG
AGTAGCTGGGACTACAGGCCTGTGCCACCAGGCCTGGCTTGTACATTAGTATTTGATATG
GCTACCCTAAGGGCAATCCTATAGTGAAGTCAACATTAGATAATGATGCTCATCTGATGG
ATTAGATTTTCAGAGTTGGCTGTTTCCAGGTGCCTATAGGAGTAGAAAAGGGTGACAAAC
CTCCTAACTAGATGTCCTACCAAATATAGTTCACTCCACATCTGAGATGAGACTGCATGA
CTGCTGGTTTTCTTTGCCTTTTCCCCCCCAGGGTATCATCAGAACCAAAAATAAAGTTTT
AAAGGTGGGTCAGGTGTGTGTTGGCTCATGCCTGTAATCCTAGCACTTTGGGAGGCTGAG
GCAGGTGGATCATCTGAGCTCAGGAGTTCAAGACCAGCCTGGCTAATAACATGGTTAAGC
CCCATCTCTACTAAAATACAAAAAGTTAGCTGGGCATGGTGGTGGGCACCTGTAATCCCA
GCTACTCAGGAGGCTGAGGCATGAAAATCGCTTGAACCCCAGAGGCGGGGGTTGCAGTGA
GCCGAGATCATGCCACTGCACACTAGCCTGAACAACAGAGCAAGGCTCTGTCTCCAAACA
AACAAAAAATGGTGCCAGAGTCTTTTCCAGGGCTGAGGGGAGATACAATGAAGTGTGTTAT
TTTTTCTGATAAGAGTGCTACCATCTTTCATTCTTGTGTGCCATTTCTAGTTGGGGTGAA
TTTGTTTTCGGAGTTCCTTTCCCAGCTGTTTGCCTGAAAAACCATGAAATGTGTTCCACA
TGAACTATGAAATGATTAGATGCTAATGTGGCAAAGAAAGTGTGAATTCTCTTGTAGAAA
CAGGGACATTTGGTTCGGTACAGTAAGTTGTTAATGCGTGACTCTGTGCTTTCAAATTCT
GTGGTTCAAAAGTACTTTTCACTCCTACTGTGTATTTACCTTGAGAAGGTGAATCCCCTA
ACAATTTGGTCAATGTATCAGTATTCTCAACCCGTCTATCAATTTTTTTTTCTTTCTCCC
TCTTTTTTCTTTTTTTGGGCAAAATACCTTTTTTGCTTTTTATCCCCTTAAAATAACCAT
TGTCCCTCACATGTGCACTCTTCCAAATTTCAG**AAAAGCAAGAGGAAAGGGCACGAATAT**
ACAAATATTAAGTATTCTCTAGCGGACCAGACGAGTGGAGATCAGAGCCCTCTCCCGCCT
TGTACTCCAACGCCACCCTGTGCAGAGTAAGTAGTGCTGAAGGAAATTCTTTTTACCTGG
TCATGGTGGTTTAAAAAGGTTTAAAAAACAAAAACAAAAACAAAACACAAGTTTGTAGCA
CATGCCTTTCACTGGTGCACGTTCCTGTTGCCCTACTGTTAGTGTATCTGTGACTGGTGA
TATCTATTGATTGTGTTAATGCTATCTCAACCACGTTTTAATTTTCCTAAGCTGGCCAGG
CACGGTGGCTAACGCCTGTAATCCCAGTGCTTTGGGAGGCCGAGGTTCATGGATTACTTT
GAAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAAT
ACAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAATCCCAGCTACTCAGGAGGCTGAG
GCAGGAGAATCGCTTGAACCCAGGAAACGGATGTTGCAGTGAGCCGAGATCATGCCACTG
CACTCCAGCCTGGGCGATAGAGTGAGCCTCTGTCTAAAAATAAAATAAAATAAAATAAAT
TCCTAAACTGAAGGCTGACTGCTATGCTAGCTAGGATTATATGGGATTTTAAGTATATCA

AGTGGTGGTTCTCCAAGAAGAATCTAATTTTTCTTTTGATGGGCTGGGGATTGTAACAAA
GGAAGGTCATATGTCTTAATGATGTGTTAAGGCTCTTTGCAAAATCAAAGTAAATAAATT
GACCACTAATGTGTCAGCCCAGCCATGTTCTGCTCATTTGCCACCAGTCAACAGAAATCT
ACTTTGGGTGTTTAAACCAGGAGTCAGCAAACTACAGCTCACAAGGCCAGATGTGGGCCA
TGGCCTGTTACTGTATGGCCTGTTAATGGTTTTAAAGGGTTGTAAAACAAAAGAACACAA
AACAAAGACCCAATAACAAAACAAAGCCCGAAGAATAATATGCGACAGAGACCATGTATG
GCATATAGAGCCTAAAATACTGACTCTCAAGCCCTTCCCAGAAATCCTTCCCGACTCCTT
GTTGAAAACACGGTAGGAAAGCATTTGTCAAATTGAGGATATGAATAGCAATTGTAAGTT
ATTATTTTTCTATATATTCGAAAGTCACTTGCTAGTATAACATTTACCTTTTATTTTTCC
CTAAGAATCTTCTCTCTGTTTGCTTTCGACATGGATTTTTAAACCCCTGCAGATTTTAAT
ATTCTATATAAATGTTTTAGGTGGCATATATGAGGTTTGTATTAACATTTGCTTTCTATT
TAACATTGAAATGAAATTATACAGCAGAGGTATTTTCTCGTCCAAGTTGCCACTTCTTTC
TATCTTTTTTCTTTTCTTTCCCAGTGGACTGCCTGGGAAAATTGATATTTTAAATTGCTC
TCTGCAATAATTTGCAATGGAACTGGAATGCCAGGGTTCTGAGTCCTTGCCAGACAGCTC
GTCCCTCCTGTTGGCATGACTGAGTCAGCTGTCATGATTCCCTCAGTACCAGTGGCATGC
CTGTGACAGACAGCCTGTCTGCCTTTCATTCCCGTCGTCTCCCTTGTAGGGTTCAGATCC
AGGATACACTGGTCCTGGAGCCCCTCTCAGCCTGGCACCCACAGCTGCTGGGTTCCTTAC
TCTCCTGGACTGCTCTGATGTCATCTCCCTGCTCAGCAGAAAGAAGTCTGGGATCTTGAT
GCTTTGGCCCTCTGTCCTAGGCCCTAAACCACCCATTGCCCTTCACATAACCTGAGCTGG
GGCTAAATAGATCTCTCATCACTGCCTGCCTGCTCCTGTATTTTCCCTTCTTGGAGCTTT
TGCCTGTTCAGATCCCTCTACTGGAAATTAATAGGATTTCATTCTATGTGTGCATTTCCA
ACCTTTCTTCACAGTGCGATCCAAATGCCTCATCCTACAGGCCTCCTTAAAACAACCTGC
TTTCTGCCAGACCCCAGGGAGCACCAGGACTTGAGGCTTTTATTGCACTTCTGTTGTTTT
TTTGAGATGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGATCTCTGCTCA
CTGCAACCTCCATCTCCCGAGTTCAAGAGATTCTTCTGCCTCAGCCTCTCAAGCAGCTGG
GACTACAGGCATGTGCCATGACACCCGGATAATTTTTGTATTTTTAGTAGAGACGGGGTT
CACCATATTGGCCAGGCTGGTCTCAAACTCCTGACCTCGTGATCCACCCACCTGGGCCTC
CCAAAGTTCTGGGATTACAGGCGTGAGCCACCATGCCCAGCGTTATTTCACTTCTGCCTC
TGTAATTATATTGCTGTATGGCTATCTCTTCTCTCCCTGGGAATGTCAGGTCCTAGGCAC
AGGAACTGTGTCTGTACCATATCTGGTGCCCAAAGAATGTAGTATGTGTTTTATAGATAT
CATGTAAGCTTAAACAGCGTGGTCTACATTTTTGTAAATGTCTTTCTTTTTCTTTTCTCT
CCAG**AATGAGAGAAGACAGTGCTAGAGTCTATGAAAACGTGGGCCTGATGCAACAGCAGA**
AAAGTTTCAGATGAGAAAACCTGCCAAAACTTCAGCACAGAAATAGGTATTTAAATGCAA
GTGCTCTATTGGTTAATTGTTTATATAATTGGCAGTATTTTTAAGCAGGCAAGCAATTTG
GGAATGTTTTAGCAAAGTGTACCATAATTGAGTTTTACAAACCAGGCTCCTTTTTCCTCT
CCCTGTACTTCTTTTTCCAAGATGGTTTTAGTTTAGAGTTCATTAAACATTAAAATCAAA
CACAGAATTAATTCTGCATGAGGCAAGGCTAGCACTTATTCCAGAGAAATGGCTGATACT

GGTGGTAGAGTGCAGGTATCACTGTTCCTGCAATTTTTATTAGAGTTGGTTAGCCCAGGC
TGTGCTGGGGGATGATCTGTAGGGATCTGGGAAGCATCGGGACTCAGCACTGGGTGGTTG
GGAGTCAGGAAGCCTGAGTTCTCATTTCAGTCAGTCTCTGACCAACTGTGTGGCATGGGG
TGCTAGACCACTTGGCTGCCGACTGGGTCACCGACATCCCTTCCAGCTCTGCTGCTGGAA
ATTCATCTCTCCCATATGTTGCCTCCCCATCAATTACGTTTTTTAAGTGTGACCCAAGTA
TATGATGTATGTTTTCATGATAAATTAGAAACTTATCTGGGCATGGTGGCTCATACCTGT
AATCCCAGCACTTTGGGAGGCTGAGGTGGGCGGATCACCTGAGGTCAGGAGTTCGAGACC
AGCCTGACCAACTAAAATAGTAGAGACCAACCCGTCTCTACTAAAAATAGAAAATTAGCT
GAGCATGGTGGTGCATGCCTATAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAGGCAG
CGGTTGCAGTGTGCCAAGATCGCGCCATTGCACTCCACCTGGGCCACAAGAGTGAAACTC
CATCTCAAAAAAAAAAAAAAAAAAAAAAAACTCAGTGTCAGTATTTCATGTCGAAATTC
CACTTCAATGGGTAGTGTAGTTAAAAGCTCTAAGTCTACCTTAAAATCACCTAATGCTTT
GTTAAGCTTTTAGATATATGTTCCTTAAAAACTCTTAACTTATTTCTTCCCCAG**ATGTGG**
ACTTTCACCCTCTCCCTAAAAAGATCAAGAACAGACGCAAGAAAGTTTATGTGAAGACAG
AATTTGGATTTGGAAGGCTTGCAATGTGGTTGACTACCTTTTGATAAGCAAAATTTGAAA
CCATTTAAAGACCACTGTATTTTAACTCAACAATACCTGCTTCCCAATTACTCATTTCCT
CAGATAAGAAGAAATCATCTCTACAATGTAGACAACATTATATTTTATAGAATTTGTTTG
AAATTGAGGAAGCAGTTAAATTGTGCGCTGTATTTTGCAGATTATGGGGATTCAAATTCT
AGTAATAGGCTTTTTTATTTTTATTTTTATACCCTTAACCAGTTTAATTTTTTTTTTCCT
CATTGTTGGGGATGATGAGAAGAAATGATTTGGGAAAATTAAGTAACAACGACCTAGAAA
AGTGAGAACAATCTCATTTACCATCATGTATCCAGTAGTGGATAATTCATTTTGATGGCT
TCTATTTTTGGCCAAATGAGAATTAAGCCAGTGCCTGAGACTGTCAGAAGTTGACCTTTG
CACTGGCATTAAAGAGTCATAGAAAAAGAATCATGGATATTTATGAATTAAGGTAAGAGG
TGTGGCTTTTTTTTTTTTCTTTTTTCCAGCCGTTGACCAATTATAGTTCGGCTGTTGACT
GAGAAGTTTGTGGTGGGAAAACGTTTGCCATATTTTCTTTGCATTTGAATAATTGTCTTG
TACTTAGAAAAAAGGCGTCTATGAATGACCAGTGTTTTTGGTCGCCAAATGTTGCTGACA
AACTTATCCCAAAACTTTAGTGGCTTAAAAAAACCTGCCCCCAACTGTTAGTCAATCTGA
GCTGGGCTCAGCTGGGCTGTTCTTCTGCCAGCCTGCAGGTGGCCACTCATGTGGTCAGCA
GGTCGGCGGAGAGACTGGGATGGCTGGGCTTCTCTCTCTGCCTGCAGTCCTGAGTCTCTC
CTTCTTCGTGTAGTCTCTTTCAGTGGCCTGGCTGGCAGGGTAGCTAGACCTCTCACATGC
AGCTCAGAGCTCCCAAGAGCTCAAAAGCAGAAATGGCCAGGCCTTCTGAAAACTTAAGTC
CAGAATTGTCACAGTGTCCCTTCTACTTCCCTCTATTGATGATGATGATGATGATGATGA
TGATGATGATGATGATGATGGTTTTTTCTAATCAGAAGAAAGCTGGGGTATGCCCTC
TACTTACTAAACAAGTCACAAGCCCAGCTCAGATTCAAGAAAAGGGTGTGAAGTAGAGGT
GCAGTTAAGTGGGGGGCCACTAGTCTAACAGACGGTCACAACCAGTGCCATGGAAAACCA
AGGATATTAGCAAAAGCAGAAGTTGCTAGTGACCTTGGGAAGCCGAAGCTGCTTACAGTA
GCTGGGACAAGCTGAAAGTCAGACTAAGAAATAAAGAGAGGGCCTTCAAGAAGCTTCCTG

AATGATTTCTGCTAGCCCTGAGCCTATTTTTGGAACCAGCACTTGGGGAAACTGATCTTG
TGAGGATGGATGTGTTTAGGGACACAGGGCTTTTGAGAGCAGCACCACCCCACTGGGGCA
TCCCCAGACTTGGGAAACGTGACTCTTTCTTAATGCCACTGGGTTTTAGTCAGGCCACAG
TGAGAAGGAACAGCCCTAACAGGCCTCCAGCCAGGTTGAATGAGCTCATTTTTGTTGTAG
CCAACCAGTAAGATTTGCTAATGTTCTACATTAAGTGCCTTCTCCAAAGACATCCCTCTT
TGCCTCATATGTTGAATCATCCAGTGCGGATATTTCAATGAAAATATCATTGGTTGACTT
TTGTGATGGTAATAATGCTATGGCATCTTTGCCATGAAGTTGTGGCCTCCTTGGATTCTT
CTGACTTTGGCTTCTGAAAGGAAGGCCTAGATCCAGCCCTGGTGGTAGTTCCTTTCTGAG
GTCTCTCAGTCCCTTGAGACTTTGGGGTAGTTTGGCTGCCATTCTCACTGACAAAATGTA
TATCAGCCCCCACCTCCACCCCCCAATATTCCTTGAACTTTGAATTGCTTCAGAACACAG
GTGTGGCCTGAAGGTATTCCCTTATTAGGGAAGTGTCACTGCTGTCTTCTAGTCAAACTT
GTAAAGAAAAAGATTCCAGTTCAGTATTTGCAGCAAGAAGCTTGAATGCTGTTCTTTTTA
TCGCATTGTTACATCGACTCATTCTCCATTTTGCTTTGGTTTTGTCTTGACTTGACTTGA
CTTTGGGGGTAAAGTCTTTCACCAGCACACAAGAGTTTGATTGTACAAATATATCTTCTG
CATTAACATCTCTGCCTGTTGCTTAAGATCAGTTGCTTTTATACTCAGAATGGAAATACC
TGATCTTGGCTAGTTTTGTTATAAGATATTGATTTCATTTAGATTTCCCTCCACGAGGTC
AGCAAACTATCATGTTCTTATGTAAACTTAGGCCAAGGCCAGAGTTATCATAGTCCCTAG
GTTGCTACGGCTTATCATGTGCTTGGTAAAAGGTGATCGCAGGTTCTCAGACGAGTTTAC
TTTACATGAGATGGAATCAGGCAGAGAGGCTGGGATGATGGAGAAAGCTCGAGGTGAAGT
TTTAAAAAAAAAGTTGTGGAAAGGAAAGTTCCAAAGAGGTGGTTTCTGAGGAAGTCAGAG
CGCCCAGGGCCAGAGCAGTCAGTAATGGGTGAATGAGGTTGTTTGGAAAGTCGGTGTGAC
AGACACATGGATGCCATCTACTTCTAGGTTGCTGGTGGGTATTAAATATGCACAATATTC
CATAGCTCACTGAGGATTTTAAAATTATAAGCATAGGATTTTATATTTTGGGGTGAAAGA
ATTATCTGGCACATTAGGTATTGGAGTTTAAAAAAAAAGCCAAATTTCACAGTCTTAATA
ACTTTTTTTTAAAAAAAACTAAAAGGCGCTTCATGTCCAGTGTGTGGCCCTTCTGAAACTT
ATGGTCATCTCTCCCACTGAAACCAAGGTCTTTTCAAATGTGGCTAAATGGGGATGAGGA
GACACGGGTAGGACTTTCTTGGTGTGTGTGCATTCTTTAAAGAGCCAAGTTGCTTCGGGG
AAACAGCCAGGAAAATGGTCAAGATTATTTTTAGAGGTTATTTTATTGGGGATTTTAAGA
ACTAATAACATCTTGAGTTATTTTTAATTCAGGGGGATGTGGAAAGGTTTGCAATTGTCA
AGTGTTTTGTTGTAGCTTAGTATCCATAAGGGAAACTTAGACTATAGACATAACTACAAA
GCCAGTGCAGCTTTTGTTTTCTGTATGTTGTTGGGGGATCAACTTTCACACATAGCAAGC
ACATGGCCTCCCTGATGTCAGGATGCCTTTGTTAGGATCTGTATTTGCCCTTAATTTTGT
TGAAATCTTTTTTCCTTCTTCCTCTTGAAAAGTTCCAAAATATAGTTTATTGTATCTTTC
ATCACTAAAAATTTGTTCCTTTTTCACTATGGGCAGTTCACACAAGGCAAAAACTATTGA
ACAGTTGGTTTTAGTGTGTTGTATAACTTTGCTGTATATCAAACTAATTTTGACAAGTTT
TCATCCTAAGCCTCAAATCATGTAATTAATAATTTGCCTGTTTATTTATGACCTAATTGT
GATTCTTTTATTAATAAAAGCTAATGGGAAAAGGATCCCTGATTAAGCTGATGACTAGAC

CTACAATTAATTTTCCTGCAGTATATGAAGTATTGTACCAGAGTATTAAAAGATATGTAA
TATTTTATTGATAAATCTATCCTTTAAAAGGAATACGTTTTAGGATGTCATCATTTTGAT
GTGAATCATGTAAATGTTGATAATATGCTGTTTATTATACATTTAGTGTTTCAAGAGATT
CACTTAATTGCCTTTTTGCCCACGTATATTATGTAGTCTATTTGCAACTGTTCTTAAAAA
AATGACATTAAAAGAATAGTTTATGTAGAGAAACATTAGTGGATGTTAATTGTCTCCCCA
CCTATATTTATGGGTGTTAGCGCAACTGCTTTGCTAGTTGCAAAGCTGTATTATCAGAGT
AAAAGTGTATTTGTAAACTGTATGGGAACTAAAAATTAGGAATAAAACCATTTTCTTATA
TGATGGCATTTGTCGTTTGCTTCATCAGAAATGTCCAGGAAAAAAATGGGATTATTGGTC
ACTCCACCTCTCACACTGGCAAAATACTGACATTTAGCAGCTCTTATCTAGAAGTGACTT
GGAACATAGAATAAAGGCATGAGTTCCTGAAGAATTCATTGAGTGTTTCCTGTAGAAATA
GCTTTAGGAGATAGGGAGTTCTATCTGGGAGAACATATGAGTAACTCAAGAGTAAAAAGT
ATAGTCTGTGTAAACTATAGAAGAAATGCTGGGCATGGTGGCGCGCCCCTGTAATCTCAG
CTACTTGGAGGCTGAGACGGGAGGATTCCTTGAACCCAGGAGCCCAGGAGTTTTAGACCA
GTCTGGGTAACATAGTGAGACCCTTTCTCACCTACTCTCACTGCATGCCCCCCAAAAATA
TATATGTGCGCGCACGCGCGCGCACACACACATACACACACACACACACACACACACACA
CAGAGGAAATTGTTAGAAAACACACAGAACTGAATGTAAATAGTATTAGGTGGGAATAAG
AAGTAAAGGGATGGTAAGGAGGCTTGGAGGAGGAGTAAATTATCTGCTATGGGACATCAG
CTC

[0048] FIG. 11 shows a SHP2 translated amino acid sequence (SEQ ID NO: 61). Alternating exons are underlined and non-underlined. Bold with italics indicate a residue overlap splice site.

[0049] Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION OF THE INVENTION

[0050] Protein/nucleic acid complexes or assemblies are difficult to manipulate due to their fragility and requirement for structural integrity (e.g., 3-dimensional conformation. Genome editing technologies, such as clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9, transcription activator-like effector nucleases (TALENS), and others, have shown much potential in their ability to change the genetic code of cells. However, their activity is highly dependent on structural and conformational integrity.

[0051] Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are examples of a class of gene editing tools. These chimeric nucleases are composed of programmable, sequence-specific DNA-binding

modules linked to a nonspecific DNA cleavage domain. ZFNs and TALENs enable a broad range of genetic modifications by inducing DNA double-strand breaks that stimulate error-prone nonhomologous end joining or homology-directed repair at specific genomic locations (Gaj et al., 2013, Trends Biotechnol. 31(7):397-405; hereby incorporated by reference).

[0052] Numerous publications describe the use of viruses, mRNA and plasmids to code for the Cas9 and/or gRNA and/or donor oligonucleotides (where relevant). Compared to these methods, the methods and systems described herein utilize a different strategy – delivery of the protein itself, complexed with the guide RNA. This process is fundamentally different, because it does not require the cell to translate/transcribe anything for the editing to work. For that reason, and because the protein/guide complex has a short half-life, the approach of the invention results in faster and more efficient editing with fewer off-target effects.

Other approaches such as liposome mediated protein delivery, microinjection, and cell-penetrating peptides (CPP).

[0053] One example of liposome mediated protein delivery uses GFP fused to Cas-9, has also been used (Zuris et al., 2015, Nature Biotechnology 33:73–80). The GFP is capable of complexing with conventional lipofection agents (e.g. lipofectamine) due to charge interactions and appears to mediate a gene editing response. The main advantages of the invention relative to this approach are: 1) does not require a fusion protein 2) does not require lipofection agents which can have toxicity, endosome escape problems, and issues/problems translating to primary cells.

[0054] Microinjection mediated complex delivery is characterized by extremely low throughput and can be difficult to implement for most mammalian cell types. The latter drawback is highlighted by the fact that the work was done with embryos, i.e., cells that are much larger than a fibroblast or a T cell (cells which are desirable target cells for gene editing endeavors). By enabling high throughput and translatability to smaller primary cells, *e.g.*, fibroblasts, T cells, stem cells, the methods described herein have a big advantage.

[0055] A CPP-based strategy does not involve a complex. One example of such as strategy is described in Ramakrishna et al., 2014, Genome Res. 24(6):1020-7. CPP mediated delivery of individual components is also associated with drawbacks. Conjugating a CPP to the guide and Cas9 requires extra modification that

may inhibit function, limit scalability. CPP mediated delivery is known to go through endocytosis and is inefficient or ineffective in many primary cells (particularly immune cells).

Target cells and payload compositions

[0056] Any gene can be manipulated using the gene editing strategies described. Some target genes/proteins are particularly relevant in clinical disease and thus gene editing of such target genes/proteins is useful for therapy. Examples include C-C chemokine receptor type 5 (CCR5): prevent human immunodeficiency virus (HIV) infection; major histocompatibility complex class I (MHC-I): reduce graft vs. host disease; cluster of differentiation 1 (CD1): reduce graft vs. host disease; programmed cell death protein 1 (PD-1), programmed death-ligand 1 (PDL-1), Cytotoxic T-Lymphocyte-Associated Protein 4 (CTLA-4), interferon-regulatory factor (IRF) protein family, TLR protein family, pattern recognition receptors (PRRs): modulate immunity to enhance or dampen effector/antiviral responses; forkhead box P3 (FoxP3): eliminate Treg mediated tolerance; cluster of differentiation 80 (CD80), cluster of differentiation 86 (CD86) and other costimulatory molecules: knockout costimulation abilities to promote tolerance; T cell receptor (TCR), B-cell receptor (BCR): eliminate endogenous TCR or BCR to allow for engineering of T cells and B cells with desired specificity; oncogenes (e.g., Kras, Myc, Tp53): cancer therapy. In another example, targeting transcription factors is used to change cell fate, e.g., delete FoxP3 to remove Treg type function. Delete nuclear factor-kappa B (NF-kB), t-bet, Eomesodermin (Eomes), etc. to alter T cell differentiation.

[0057] A protein coding sequence for forkhead box P3 (FoxP3) is as follows:

[0058] ATGCCCAACCCCAGGCCTGGCAAGCCCTCGGCCCCTTCC
TTGGCCCTTGGCCCATCCCCAGGAGCCTCGCCCAGCTGGAGGGCTGCACC
CAAAGCCTCAGACCTGCTGGGGGCCCGGGGCCCAGGGGGAACCTTCCAG
GGCCGAGATCTTCGAGGCGGGGCCCATGCCTCCTCTTCTTCCTTGAACCCC
ATGCCACCATCGCAGCTGCAGCTGCCCACACTGCCCCTAGTCATGGTGGC
ACCCTCCGGGGCACGGCTGGGCCCCTTGCCCCACTTACAGGCACTCCTCC
AGGACAGGCCACATTTCATGCACCAGCTCTCAACGGTGGATGCCCACGCC
CGGACCCCTGTGCTGCAGGTGCACCCCCTGGAGAGCCCAGCCATGATCAG
CCTCACACCACCCACCACCGCCACTGGGGTCTTCTCCCTCAAGGCCCGGCC
TGGCCTCCCACCTGGGATCAACGTGGCCAGCCTGGAATGGGTGTCCAGGG

AGCCGGCACTGCTCTGCACCTTCCCAAATCCCAGTGCACCCAGGAAGGAC
AGCACCCTTTCGGCTGTGCCCCAGAGCTCCTACCCACTGCTGGCAAATGGT
GTCTGCAAGTGGCCCGGATGTGAGAAGGTCTTCGAAGAGCCAGAGGACTT
CCTCAAGCACTGCCAGGCGGACCATCTTCTGGATGAGAAGGGCAGGGCAC
AATGTCTCCTCCAGAGAGAGATGGTACAGTCTCTGGAGCAGCAGCTGGTG
CTGGAGAAGGAGAAGCTGAGTGCCATGCAGGCCCACCTGGCTGGGAAAA
TGGCACTGACCAAGGCTTCATCTGTGGCATCATCCGACAAGGGCTCCTGCT
GCATCGTAGCTGCTGGCAGCCAAGGCCCTGTCGTCCCAGCCTGGTCTGGC
CCCCGGGAGGCCCCTGACAGCCTGTTTGCTGTCCGGAGGCACCTGTGGGG
TAGCCATGGAAACAGCACATTCCCAGAGTTCCTCCACAACATGGACTACT
TCAAGTTCCACAACATGCGACCCCCTTTCACCTACGCCACGCTCATCCGCT
GGGCCATCCTGGAGGCTCCAGAGAAGCAGCGGACACTCAATGAGATCTAC
CACTGGTTCACACGCATGTTTGCCTTCTTCAGAAACCATCCTGCCACCTGG
AAGAACGCCATCCGCCACAACCTGAGTCTGCACAAGTGCTTTGTGCGGGT
GGAGAGCGAGAAGGGGGCTGTGTGGACCGTGGATGAGCTGGAGTTCCGC
AAGAAACGGAGCCAGAGGCCCAGCAGGTGTTCCAACCCTACACCTGGCCC
CTGA (SEQ ID NO: 62)

[0059] Src homology region 2 domain-containing phosphatase-1 (SHP1) is also known as tyrosine-protein phosphatase non-receptor type 6 (PTPN6). A protein coding sequence for SHP1 is as follows:

[0060] ATGGTGAGGTGGTTTCACCGAGACCTCAGTGGGCTGGAT
GCAGAGACCCTGCTCAAGGGCCGAGGTGTCCACGGTAGCTTCCTGGCTCG
GCCCAGTCGCAAGAACCAGGGTGACTTCTCGCTCTCCGTCAGGGTGGGGG
ATCAGGTGACCCATATTCGGATCCAGAACTCAGGGGATTTCTATGACCTGT
ATGGAGGGGAGAAGTTTGCGACTCTGACAGAGCTGGTGGAGTACTACACT
CAGCAGCAGGGTGTCCTGCAGGACCGCGACGGCACCATCATCCACCTCAA
GTACCCGCTGAACTGCTCCGATCCCACTAGTGAGAGGTGGTACCATGGCC
ACATGTCTGGCGGGCAGGCAGAGACGCTGCTGCAGGCCAAGGGCGAGCC
CTGGACGTTTCTTGTGCGTGAGAGCCTCAGCCAGCCTGGAGACTTCGTGCT
TTCTGTGCTCAGTGACCAGCCCAAGGCTGGCCCAGGCTCCCCGCTCAGGG
TCACCCACATCAAGGTCATGTGCGAGGGTGGACGCTACACAGTGGGTGGT
TTGGAGACCTTCGACAGCCTCACGGACCTGGTGGAGCATTTCAAGAAGAC
GGGGATTGAGGAGGCCTCAGGCGCCTTTGTCTACCTGCGGCAGCCGTACT
ATGCCACGAGGGTGAATGCGGCTGACATTGAGAACCGAGTGTTGGAACTG

AACAAGAAGCAGGAGTCCGAGGATACAGCCAAGGCTGGCTTCTGGGAGG
AGTTTGAGAGTTTGCAGAAGCAGGAGGTGAAGAACTTGCACCAGCGTCTG
GAAGGGCAGCGGCCAGAGAACAAGGGCAAGAACCGCTACAAGAACATTC
TCCCCTTTGACCACAGCCGAGTGATCCTGCAGGGACGGGACAGTAACATC
CCCGGGTCCGACTACATCAATGCCAACTACATCAAGAACCAGCTGCTAGG
CCCTGATGAGAACGCTAAGACCTACATCGCCAGCCAGGGTTGTCTGGAGG
CCACGGTCAATGACTTCTGGCAGATGGCGTGGCAGGAGAACAGCCGTGTC
ATCGTCATGACCACCCGAGAGGTGGAGAAAGGCCGGAACAAATGCGTCC
CATACTGGCCCGAGGTGGGCATGCAGCGTGCTTATGGGCCCTACTCTGTG
ACCAACTGCGGGGAGCATGACACAACCGAATACAAACTCCGTACCTTACA
GGTCTCCCCGCTGGACAATGGAGACCTGATTCGGGAGATCTGGCATTACC
AGTACCTGAGCTGGCCCGACCATGGGGTCCCCAGTGAGCCTGGGGGTGTC
CTCAGCTTCCTGGACCAGATCAACCAGCGGCAGGAAAGTCTGCCTCACGC
AGGGCCCATCATCGTGCACTGCAGCGCCGGCATCGGCCGCACAGGCACCA
TCATTGTCATCGACATGCTCATGGAGAACATCTCCACCAAGGGCCTGGAC
TGTGACATTGACATCCAGAAGACCATCCAGATGGTGCGGGCGCAGCGCTC
GGGCATGGTGCAGACGGAGGCGCAGTACAAGTTCATCTACGTGGCCATCG
CCCAGTTCATTGAAACCACTAAGAAGAAGCTGGAGGTCCTGCAGTCGCAG
AAGGGCCAGGAGTCGGAGTACGGGAACATCACCTATCCCCCAGCCATGA
AGAATGCCCATGCCAAGGCCTCCCGCACCTCGTCCAAGAGCTTGGAGTCT
AGTGCAGGGACCGTGGCTGCGTCACCTGTGAGACGGGGTGGCCAGAGGG
GACTGCCAGTGCCGGGTCCCCCTGTGCTGTCTCCTGACCTGCACCAACTGC
CTGTACTTGCCCCCCTGCACCCGGCTGCAGACACAAGGAGGATGTGTATG
AGAACCTGCACACTAAGAACAAGAGGGAGGAGAAAGTGA (SEQ ID NO: 63)

[0061] Src homology region 2 domain-containing phosphatase-1 (SHP2) is also known as tyrosine-protein phosphatase non-receptor type 11 (PTPN11). A protein coding sequence for SHP2 is as follows:

[0062] ATGACATCGCGGAGATGGTTTCACCCAAATATCACTGGT
GTGGAGGCAGAAAACCTACTGTTGACAAGAGGAGTTGATGGCAGTTTTTT
GGCAAGGCCTAGTAAAAGTAACCCTGGAGACTTCACACTTTCCGTTAGAA
GAAATGGAGCTGTCACCCACATCAAGATTCAGAACACTGGTGATTACTAT
GACCTGTATGGAGGGGAGAAATTTGCCACTTTGGCTGAGTTGGTCCAGTA
TTACATGGAACATCACGGGCAATTAAAAGAGAAGAATGGAGATGTCATTG

AGCTTAAATATCCTCTGAACTGTGCAGATCCTACCTCTGAAAGGTGGTTTC
ATGGACATCTCTCTGGGAAAGAAGCAGAGAAATTATTAACTGAAAAAGG
AAAACATGGTAGTTTTCTTGTACGAGAGAGCCAGAGCCACCCTGGAGATT
TTGTTCTTTCTGTGCGCACTGGTGATGACAAAGGGGAGAGCAATGACGGC
AAGTCTAAAGTGACCCATGTTATGATTCGCTGTCAGGAACTGAAATACGA
CGTTGGTGGAGGAGAACGGTTTGATTCTTTGACAGATCTTGTGGAACATTA
TAAGAAGAATCCTATGGTGGAAACATTGGGTACAGTACTACAACTCAAGC
AGCCCCTTAACACGACTCGTATAAATGCTGCTGAAATAGAAAGCAGAGTT
CGAGAACTAAGCAAATTAGCTGAGACCACAGATAAAGTCAAACAAGGCT
TTTGGGAAGAATTTGAGACACTACAACAACAGGAGTGCAAACTTCTCTAC
AGCCGAAAAGAGGGTCAAAGGCAAGAAAACAAAAACAAAAATAGATATA
AAAACATCCTGCCCTTTGATCATACCAGGGTTGTCCTACACGATGGTGATC
CCAATGAGCCTGTTTCAGATTACATCAATGCAAATATCATCATGCCTGAAT
TTGAAACCAAGTGCAACAATTCAAAGCCCAAAAAGAGTTACATTGCCACA
CAAGGCTGCCTGCAAAACACGGTGAATGACTTTTGGCGGATGGTGTTCCA
AGAAAACTCCCGAGTGATTGTCATGACAACGAAAGAAGTGGAGAGAGGA
AAGAGTAAATGTGTCAAATACTGGCCTGATGAGTATGCTCTAAAAGAATA
TGGCGTCATGCGTGTTAGGAACGTCAAAGAAAGCGCCGCTCATGACTATA
CGCTAAGAGAACTTAAACTTTCAAAGGTTGGACAAGGGAATACGGAGAG
AACGGTCTGGCAATACCACTTTCGGACCTGGCCGGACCACGGCGTGCCCA
GCGACCCTGGGGGCGTGCTGGACTTCCTGGAGGAGGTGCACCATAAGCAG
GAGAGCATCATGGATGCAGGGCCGGTCGTGGTGCACTGCAGTGCTGGAAT
TGGCCGGACAGGGACGTTCATTGTGATTGATATTCTTATTGACATCATCAG
AGAGAAAGGTGTTGACTGCGATATTGACGTTCCCAAAACCATCCAGATGG
TGCGGTCTCAGAGGTCAGGGATGGTCCAGACAGAAGCACAGTACCGATTT
ATCTATATGGCGGTCCAGCATTATATTGAAACACTACAGCGCAGGATTGA
AGAAGAGCAGAAAAGCAAGAGGAAAGGGCACGAATATACAAATATTAAG
TATTCTCTAGCGGACCAGACGAGTGGAGATCAGAGCCCTCTCCCGCCTTGT
ACTCCAACGCCACCCTGTGCAGAAATGAGAGAAGACAGTGCTAGAGTCTA
TGAAAACGTGGGCCTGATGCAACAGCAGAAAAGTTTCAGATGA (SEQ ID
NO: 64)

[0063] Other targets include areas of the genome that can have a plasmid or donor DNA inserted into them so that the target cell can express a new gene, e.g. a

recombinant TCR, a recombinant BCR, Chimerica Antigen Receptor, fluorescent protein, reprogramming factors.

[0064] In some embodiments, a genomic sequence is edited in a coding region. In certain embodiments, a genomic sequence is edited in a non-coding region.

[0065] In various embodiments relating to FoxP3, a genetic region upstream of FoxP3 may be edited. In such embodiments a region where a transcriptional repressor of Foxp3 might bind is edited. For example a site about 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 0.5-2.5, or 0.5-5kb upstream of the FoxP3 transcriptional start site may be edited.

Treating Subjects

[0066] Aspects of the present invention relate to editing the genomes of a plurality of a subject's cells. In various embodiments, cells are removed from a subject, receive a gene-editing complex using a method of the present subject matter, and then are reintroduced back into the subject. For example, mutant cells may be produced in a process involving delivery of a gene-editing complex as described herein. The mutant cells may be heterozygous or homozygous for a mutated allele a gene involved in a disease. In certain embodiments, the mutant cells are null for the gene involved in a disease.

[0067] Cells, such as stem cells (e.g. hematopoietic stem cells) from bone marrow, or circulating immune cells in whole blood, may be treated using methods and devices described herein. Cells may be genetically modified to reduce the expression of a receptor for a pathogen (such as a viral or bacterial pathogen) or a toxin (such as a microbial pathogen toxin). Since a gene-editing protein complex or components thereof may be directly introduced into target cells without the need for expression, no transgene delivery is necessary. This approach has important advantages over traditional gene-therapy approaches, which suffer from aberrant expression, insertion, and silencing, as well as variable delivery of transgene copy number.

[0068] In one embodiment, a gene-editing complex that targets the C-C chemokine receptor type 5 (CCR5) gene is introduced into a blood (such as a CD4+ T cell) or bone marrow cell (such as a hematopoietic stem cell) of a subject who is infected with human immunodeficiency virus (HIV). The gene-editing complex may be designed to mutate the CCR5 gene such that cells receiving the gene-editing complex no longer express CCR5 or express CCR5 at a reduced level. In one

example, hematopoietic stem cells expressing a version of CCR5 that binds HIV (or that produce progeny that express the CCR5) are removed from the subject, modified to no longer express a version of CCR5 that binds HIV, and then are transplanted into the subject. In another example, CCR5-expressing CD4+ T cells of the subject receive a gene-editing complex using methods and devices described herein such that the CD4+ T cells no longer express a version of CCR5 that binds HIV. The modified CD4+ T cells are then returned into the subject. Such treatment of the CD4+ T cell may be performed in whole blood from the subject. In these and other embodiments, bone marrow cells or blood cells are modified to no longer express a version of C-X-C chemokine receptor type 4 (CXCR4) to which HIV binds. Similarly, cells of a subject may be modified to have reduced CCR5 expression to treat or prevent an infection associates with *Yersinia pestis* (bubonic plague) or *Variola major* (small pox).

[0069] Subjects, other than humans, containing cells modified by methods and devices disclosed herein are also provided. Such subjects include non-human vertebrate, amphibian, mammalian, and primate subjects. Non-limiting examples include *Danio* sp., *Fugu* sp., *Xenopus* sp., *Mus* sp., *Rattus* sp., and others.

Introducing Gene-Editing Proteins and Complexes into Cells

[0070] The delivery of pre-formed protein complexes allows for the study of cellular processes without genetic modification of the cells being studied. The present subject matter is useful for delivering protein complexes and gene editing complexes to cells, including CRISPR.

[0071] The advantages of delivering protein complexes using the methods and devices described herein include the controlled and temporary introduction of test agents for the study of cell and protein complex function. Since transgene expression and cellular assembly of complex components is not needed, the timing and ratios (protein:RNA) of complex function can be controlled. Additionally, the transient nature of delivery enables the observation of changes due to temporary function, rather than prolonged expression which may result in off-target or secondary effects. From an *in vivo* homing perspective and a gene expression format, microfluidic delivery has far fewer side effects (10-fold) on treated cells rather than electroporation.

[0072] For example, microfluidic delivery results in fewer aberrant and non-specific gene expression changes compared to electroporation. Additionally, the

structural and functional integrity of microfluidically squeezed cells is preserved compared to electroporation-mediated cargo delivery. As an example, an increased number of T cells exposed to microfluidic delivery *ex vivo* (and then introduced into a subject in *vivo*) home to lymph nodes compared to T cells that have undergone electroporation. Cells (e.g., T cells) treated by electroporation and then administered into a subject are more likely to be cleared from the subject compared to cells treated by microfluidic delivery. Such clearance is related to altered/aberrant gene expression following electroporation that marks such cells for destruction or clearance by the body.

Target cells and Payload Compositions

[0073] Any eukaryotic, e.g., mammalian such as human, cell can be processed using the microfluidic device to alter the cell membrane for introduction of protein/nucleic acid complexes or assemblies into the cytosol of the target cell. Exemplary target cells include Lymphocytes/Immune cells: DCs, B cells, T cells, Natural killer cells (NK cells), neutrophils, basophils, eosinophils, innate lymphoid cells, monocytes, macrophages, hematopoietic stem cells, common lymphoid progenitor cells; Stem cells: Embryonic, mesenchymal, induced pluripotent; Other primary cells: Fibroblasts, hepatocytes, cardiomyocytes, neurons, epithelial, epidermal, endothelial, pancreatic islet cells; as well as Cell lines, e.g., cell lines for disease studies: T cell clones, Jurkat cells, HeLa cells, Human Embryonic Kidney 293 (HEK293) cells, U2OS cells, Chinese Hamster Ovary (CHO) cells. Prokaryotic cells can also be processed. The dimensions of the constriction of the device are tailored depending on the cell type to be processed.

[0074] In some embodiments, the cell is a prokaryotic cell. In other embodiments, the cell is a eukaryotic cell. Non-limiting examples of eukaryotic cells include protozoan, algal, fungi, yeast, plant, animal, vertebrate, invertebrate, arthropod, mammalian, rodent, primate, and human cells. The cell may be a cell, *e.g.*, of a unicellular organism or a multicellular organism. The cell may be, *e.g.*, a primary eukaryotic cell or an immortalized eukaryotic cell. In some embodiments, the cell is a cancer cell. In certain embodiments, the cell is other than a human cell. For example, a composition for treating cancer and/or a method of treating cancer or preparing a composition for treating cancer, comprises treating immune cells using the gene-editing methods described herein to reduce the expression/production of immune

suppressing signals from tumor cells. An example includes reduction or SHP-2 knockout for increasing immune activity towards tumors.

[0075] In various embodiments, a cell may be in a mixture of two or more cell types or a plurality of cells may be a mixture of two or more cell types. A mixture of cell types may be a co-culture of multiple cell types (such as two or more of those disclosed herein) or a mixture of cell types that naturally occur together, such as in whole blood.

[0076] In some embodiments, the cell is a peripheral blood mononuclear cell. In various embodiments, the cell suspension comprises a purified cell population. In certain embodiments, the cell is a primary cell or a cell line cell.

[0077] In some embodiments, the cell is a blood cell. In some embodiments, the blood cell is an immune cell. In some embodiments, the immune cell is a lymphocyte. In some embodiments, the immune cell is a T cell, B cell, natural killer (NK) cell, dendritic cell (DC), Natural killer T (NKT) cell, mast cell, monocyte, macrophage, basophil, eosinophil, or neutrophil. In some embodiments, the immune cell is an adaptive immune cell such as a T cell and B cell. In some embodiments, the immune cell is an innate immune cell. Exemplary innate immune cells include innate lymphoid cells (ILCs; ILC1, ILC2, ILC3), basophils, eosinophils, mast cells, NK cells, neutrophils, and monocytes. In some embodiments, the immune cell is a memory cell. In some embodiments, the immune cell is a primary human T cell. In some embodiments, the cell is a mouse, dog, cat, horse, rat, goat, monkey, or rabbit cell.

[0078] In some embodiments, the cell is a human cell. In some embodiments, the cell suspension comprises a cell other than a human cell or a non-mammalian cell. In some embodiments, the cell is a chicken, frog, insect, or nematode cell.

[0079] Any physiologically-compatible or cell-compatible buffer system can be used as a solution to bathe/incubate the cells and process the cells through the device. For example, phosphate buffered saline (PBS), Opti-MEM®, Roswell Park Memorial Institute (RPMI), Dulbecco's Modified Eagle's Medium (DMEM). A reduced serum or serum-free media or buffer composition is preferable. The buffer or medium is chosen based to maintain and preserve the health or viability of the target cell and/or the effect on gene expression. For example, in some cases the presence of calcium in the buffer is desirable to promote or support mRNA expression.

[0080] Payload compositions include a protein-nucleic acid complex or assembly. Exemplary complexes include components or modules of a gene editing system as described above, e.g., nuclease/guide nucleic acid combination or assembly. For example, gRNA:Cas9 molar ratio ranges from 1:100,000 to 100,000:1, e.g., a preferred range, 1:10 to 10:1, e.g., 1:1 or 1:2, 2:1. Complex concentration in the buffer to facilitate delivery (molar concentrations) typically ranges from 100mM to 1nM, e.g., 10uM to 100nM. Complexes can be mixed with cells before going through constriction or afterwards.

Microfluidic Delivery of Gene Editing Complexes

[0081] In order to effect gene editing manipulations, Cas protein (such as Cas9 protein), guide RNA and donor DNA can be delivered to a cell through mechanical deformation using, for example, a microfluidic platform (e.g., as described in U.S. Application Publication No. 20140287509, filed April 17, 2014; PCT International Application No. PCT/US2014/051343 filed August 15, 2014; PCT International Application No. PCT/US2015/060689 filed November 13, 2015; and PCT International Application No. PCT/US2015/058489 filed October 30, 2015, each of which is hereby incorporated by reference).

[0082] FIGS. 1-2 illustrate an example microfluidic system that can be used for the delivery of genome editing protein, RNA, and DNA. The microfluidic system 5 includes a channel 10 defining a tubular lumen. The microfluidic channel 10 includes a constriction 15 that is preferably configured such that only a single target cell 20 can pass through the constriction 15 at one time. Preferably, the cells 20 pass through the channel 10 suspended in a solution buffer 25 that also includes delivery materials 30, although the delivery materials can be added to the solution buffer 25 after the cells 20 pass through the constriction 15. As the cell 20 approaches and passes through the constriction 15, the constriction 15 applies pressure (e.g., mechanical compression) to the cell 20, squeezing the cell 20 (e.g., shown as cell 20_1). The pressure applied to the cell by the constriction 15 causes perturbations (e.g., holes) in the cell membrane (e.g., cell 20_2). Once the cell passes through the constriction 15, the cell 20 begins to uptake the material in the solution buffer 25 through the holes, including the delivery material 30 (e.g., cell 20_3). The cell membrane recovers over time, and at least a portion of the delivery material 30 preferably remains trapped inside the cell.

[0083] In some embodiments, the device comprises a constriction length of about 5 μm to about 50 μm or any length or range of lengths therebetween. For example, the constriction length ranges from about 5μm to about 40μm, about 5μm to about 30μm, about 5μm to about 20μm, or about 5μm to about 10μm. In some embodiments, the constriction length ranges from about 10μm to about 50μm, about 20μm to about 50μm, about 30μm to about 50μm, or about 40μm to about 50μm. In some embodiments, the constriction depth ranges from about 2um to about 200um or any depth or range of depths there between. For example, the constriction depth ranges from about 2μm to about 150μm, about 2μm to about 100μm, about 2μm to about 50μm, about 2μm to about 25μm, about 2μm to about 15μm, or about 2μm to about 10μm. In some embodiments, the constriction depth ranges from about 10μm to about 200μm, about 25μm to about 200μm, about 50μm to about 200μm, about 100μm to about 200μm, or about 150μm to about 200μm. In some embodiments, the angle of the entrance or exit portion of the constriction ranges from about 0 degrees to about 90 degrees or any angle or range of angles therebetween. For example, the angle is about 5, about 10, about 15, about 20, about 30, about 40, about 50, about 60, about 70, about 80, or about 90 degrees or more. In some embodiments, the pressure ranges from about 50psi to about 200psi or any pressure or range of pressures there between. For example, the pressure ranges from about 50psi to about 150psi, about 50psi to about 125psi, about 50psi to about 100psi, or about 50psi to about 75psi. In some embodiments, the pressure ranges from about 75psi to about 200psi, about 100psi to about 200psi, about 125psi to about 200psi, about 150psi to about 200psi, or about 175psi to about 200psi. In some embodiments, the device comprises a constriction width of between about 2μm and about 10μm or any width or range of widths therebetween. For example, the constriction width can be any one of about 3μm, about 4μm, about 5μm, about 6μm, or about 7μm.

[0084] The data described below was generated using the following materials and methods. Complexes were made as follows: Mix 10μl of 1mg/ml nuclear localization signal (NLS) tagged Cas9 protein with 5μl of 1mg/ml guide RNA. Incubate on ice for 20 min to allow complexes to form. For delivery, target cells are suspended at 10 million cells/ml in serum-free media. Cells and Cas9-gRNA complexes are mixed immediately before device treatment such that complex concentration is ~0.15mg/ml. Cells are treated by the device using pressure, temperature, chip design and buffer conditions specific to the target cell type. For

example, for primary human T cells, pressure is approximately 100psi, on ice, through a 30μm length, 4μm width constriction. After a 2 min incubation post-treatment, cells are diluted in media and washed to remove undelivered complexes. Cells are then cultured to allow for gene editing to occur (e.g., 1, 2, 5, 12, 24 hours or more (for non-clinical applications, timeframe depends on assay readout, e.g., 24 hours or later). For clinical use, e.g., for patient therapy, the cells could be injected back into patient immediately after device treatment. Optionally, the cells are incubated in vitro for a time (e.g., 1, 2, 5, 12, 24 hours or more) prior to injecting the cells into a patient recipient. Temperatures, concentrations, iterations of the molecules vary depending on the target cell type.

[0085] FIG. 3 is a series of flow cytometry plots and a bar graph of the recombination efficiency for K562 reporter cells (American Type Culture Collection (ATCC)® CCL-243™, bone marrow-derived cells derived from patient with chronic myelogenous leukemia; lymphoblast morphology) that had genetic editing material delivered to cell cytoplasm using the microfluidic device illustrated in FIGS. 1-2. At top are flow cytometry plots of K562 reporter cells that had a protein Cas9, site-specific gRNA, and donor oligonucleotide delivered. A CRISPR complex is ~150kDa. These reporter cells had an mCherry gene and thus would normally appear in Q1 of the plots. If the site-specific DNA cleavage and insertion of donor oligonucleotide is successful it would lead to expression of green fluorescent protein (GFP), *i.e.*, cells would appear in Q2. At bottom is a quantification of recombination efficiency based on flow cytometry for multiple device conditions as compared to endocytosis (inlet) and untreated (NC) controls. Delivery of Cas9 protein, guide RNA and donor DNA by cell squeezing led to successful changes in the genome of reporter cell lines. These data indicate that delivery of gene editing components (in the form of a complex or assembly) by mechanical cell disruption leads to effective changes in the genome.

[0086] Delivery of TALEN proteins or mRNA, zinc finger nucleases, mega nucleases, Cre recombinase or any other enzyme capable of cleaving DNA can also be delivered to the cytoplasm of a cell by mechanical disruption of the cell membrane. An exemplary TALEN genome-editing system, including exemplary TALEN proteins, is described in Ding et al., (2013) Cell Stem Cell, 12, 238–251, the entire content of which is incorporated herein by reference. Ding et al., (2013) Cell Stem Cell, 12, 238–251 describes non-limiting examples of generic TALEN amino

acid sequences to recognize 15 base pair sequences. Non-limiting examples of generic TALEN amino acid sequences are:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPSRVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGEL
RGPPLQLDTGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALESIVAQLSR
PDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRIGERTSHRVAGSQLVKSELEEKK
SELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDY
GVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKA
QLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

(SEQ ID NO: 65)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAVDLRTLGYSQQQQEKIKPKVRSTVAQHH
EALVGHGFTHAHIVALSQHPAALGTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGEL
RGPPLQLDTGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVAIASXXGGRPALESIVAQLSR
PDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRIGERTSHRVAGSQLVKSELEEKK
SELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDY
GVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFKGNYKA
QLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

(SEQ ID NO: 66)

[0087] In SEQ ID NOS: 65 and 66, to recognize C: XX = HD; to recognize T: XX=NG; to recognize A: XX = NI; and to recognize G: XX = NN except in the last position where XX = NK. Underline indicates N-term and C-term of TAL effector.

[0088] Although a few variations have been described in detail above, other modifications or additions are possible. For example, genetic editing material can include TALEN proteins, TALEN mRNA, zinc finger nucleases, mega nucleases, Cre recombinase or any other enzyme capable of cleaving DNA delivered to the cytosol by mechanical disruption of the cell membrane.

Delivery of RNA and Cas9 in Complex Form

[0089] The results achieved were surprising in view of numerous factors that could potentially have impeded successful gene editing by microfluidic delivery of the gene complexes. For example, the Cas9-gRNA complex may have caused a Toll-like receptor (TLR) mediated or other PRR (pattern recognition receptor) mediated response that would have inhibited gene editing function and/or survival but this potential problem was not observed. Since the complex is not guaranteed to be stable once it enters the cytoplasm, it could have been degraded and rendered non-functional, but surprisingly, the delivered complexes were still able to edit.

[0090] The integrity of the gene editing complex was preserved using microfluidic based, cell-squeezing delivery to the cell. The complex does not have the same physical/chemical properties as a gRNA alone or protein alone and thus it was uncertain if the delivery process would behave the same in the context of delivering a complex vs. its individual components. Complexes are larger and less stable than their constituents. Complexes may fall apart due to, e.g., shear forces. Additionally, complexes may not survive membrane transit or in the cytosol because some other elements may break the complexes up before they are functional or have an opportunity to act on cellular targets. Complexes also have a different charge distribution which may affect the ability of a complex to be delivered. Shape and thus transport properties can also change compared to complex constituents. The delivery methods successfully preserved the structural and functional integrity of the complexes.

[0091] The shear forces involved with the delivery process could potentially have disrupted the Protein/gRNA complex and rendered it non-functional but surprisingly the delivery system was effective to introduce the complexes into the cell and the gene editing still worked. It was also not obvious that the complex would still have the appropriate nuclear localization behavior as compared to an uncomplexed Cas (such as Cas9) protein alone with NLS; however, the behavior and function was preserved throughout the process as demonstrated by the gene expression results described above.

[0092] The CRISPR-Cas system is known in the art. Non-limiting aspects of this system are described in U.S. Patent No. 8,697,359, issued April 15, 2014, the entire content of which is incorporated herein by reference.

[0093] Non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, 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, homologs thereof, or modified versions thereof. These enzymes are known; for example, the amino acid sequence of *S. pyogenes* Cas9 protein may be found in the SwissProt database under accession number Q99ZW2. In some embodiments, the unmodified CRISPR enzyme has DNA cleavage activity, such as Cas9. In some embodiments the CRISPR enzyme is Cas9, and may be Cas9 from *S. pyogenes* or *S. pneumoniae*. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands at the location of a target sequence, such as within the target sequence and/or within the complement of the target sequence. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs from the first or last nucleotide of a target sequence. In some embodiments, a vector encodes a CRISPR enzyme that is mutated to with respect to a corresponding wild-type enzyme such that the mutated CRISPR enzyme lacks the ability to cleave one or both strands of a target polynucleotide containing a target sequence. For example, an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 from S. pyogenes converts Cas9 from a nuclease that cleaves both strands to a nickase (cleaves a single strand). Other examples of mutations that render Cas9 a nickase include, without limitation, H840A, N854A, and N863A. In aspects of the invention, nickases may be used for genome editing via homologous recombination.

[0094] Non-limiting examples of Cas9 amino acid and cDNA sequences are provided below.

[0095] The amino acid sequence of a *Streptococcus pyogenes* Cas9 protein may be found in the SwissProt database under accession number Q99ZW2. This amino acid sequence is:

[0096] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV
DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFK

SNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
TILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM
KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLS
DYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN
AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNI
MNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIV
KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY
SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE
QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA
ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID
LSQLGGD (SEQ ID NO: 1)

[0097] SEQ ID NO: 1 may be encoded by the following nucleotide sequence found in the European Nucleotide Archive under accession number AAK33936.2:

[0098] ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACA
AATAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAA
AAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATC
TTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCGGAAGCGACTCGT
CTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTATTTG
TTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTT
CTTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATG
AACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGA
AATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATA

AAGCGGATTTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTC
GTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGG
ACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAA
AACCCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACG
ATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTG
AGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGA
CCCCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGC
TTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTG
GAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTA
TTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCTCCCC
TATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTC
TTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATC
TTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGAGC
TAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGG
ATGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGC
AAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGT
GAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAA
GACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTAT
GTTGGTCCATTGGCGCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAA
GTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAG
GTGCTTCAGCTCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATC
TTCCAAATGAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGTATTTTA
CGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGA
AAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTC
TTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTT
CAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATA
GATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAAG
ATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATT
GTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACT
TAAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAGCTTAAAC
GTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTA
TTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGAT
GGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACA
TTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAGTTT

ACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTA
TTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCGG
CATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAAC
TCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAA
GGTATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAA
TACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAG
AGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATG
ATGTCGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACA
ATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTT
CCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCT
AAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTG
AACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAA
TTGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAG
TCGCATGAATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTA
AAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCC
AATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGATGCGT
ATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACTTG
AATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGA
TTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTT
TACTCTAATATCATGAACTTCTTCAAAACAGAAATTACACTTGCAAATGGA
GAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAAT
TGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCA
TGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATTC
TCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCG
TAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGATAGTCCAACGG
TAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAG
AAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAG
TTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGG
AAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGT
TAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAA
AGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGC
TAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAA
CAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCA
AATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAA

AGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAG
CAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTG
CTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTCTACAA
AAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATG
AAACACGCATTGATTTGAGTCAGCTAGGAGGTGACTGA (SEQ ID NO: 2)

[0099] The amino acid sequence of a *Streptococcus thermophilus* Cas9 protein may be found in the UniProt database under accession number Q03JI6.1. *See also*, Sapranauskas et al., (2011) *Nucleic Acids Res.* 39:9275-9282. This amino acid sequence is:

[00100] MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSK
KYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLD
DAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKAYHDEFPTIYHLRKYLADSTK
KADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDL
SLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKC
FNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKLYDAILLSGFLTV
TDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNGYA
GYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQ
IHLQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIR
KRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVY
NELTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGY
DGIELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMI
KQRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLID
DGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEVVKSLPGSPAIKKGI
LQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQRLKRLEKSLK
ELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDRLSN
YDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKKRKTFWYQLLK
SKLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNN
KKDENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVV
ASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLA
DGRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGL
DRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSFTVLVK
GTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLF
ELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHR
KYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSS

FIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVT GLYETRIDLAKLGEG (SEQ ID NO: 3)

[00101] SEQ ID NO: 3 may be encoded by the following nucleotide sequence found in the European Nucleotide Archive under accession number ABJ66636.1:

[00102] ATGACTAAGCCATACTCAATTGGACTTGATATTGGAACG AATAGTGTTGGATGGGCTGTAACAACTGATAATTACAAGGTTCCGTCTAA AAAAATGAAAGTCTTAGGAAATACGAGTAAAAAGTATATCAAAAAGAAC CTGTTAGGTGTATTACTCTTTGACTCTGGAATCACAGCAGAAGGAAGAAG ATTGAAGCGTACTGCAAGAAGACGTTATACTAGACGCCGTAATCGTATCC TTTATTTGCAGGAAATTTTTAGCACAGAGATGGCTACATTAGATGATGCTT TCTTTCAAAGACTTGACGATTCGTTTTTAGTTCCTGATGATAAACGTGATA GTAAGTATCCGATATTTGGAAACTTAGTAGAAGAAAAAGCCTATCATGAT GAATTTCCAACTATCTATCATTTAAGGAAATATTTAGCAGATAGTACTAAA AAAGCAGATTTGCGTCTAGTTTATCTTGCATTGGCTCATATGATTAAATAT AGAGGTCACTTCTTAATTGAAGGAGAGTTTAATTCAAAAAATAATGATAT TCAGAAGAATTTTCAAGACTTTTTGGACACTTATAATGCTATTTTTGAATC GGATTTATCACTTGAGAATAGTAAACAACTTGAGGAAATTGTTAAAGATA AGATTAGTAAATTAGAAAAGAAAGATCGTATTTTAAAACTCTTCCCTGGG GAGAAGAATTCGGGGATTTTTTCAGAGTTTCTAAAGTTGATTGTAGGAAA TCAAGCTGATTTTAGGAAATGTTTTAATTTAGACGAAAAAGCCTCCTTACA TTTTTCCAAAGAAAGCTATGATGAAGATTTAGAGACTTTGTTAGGTTATAT TGGAGATGATTACAGTGATGTCTTTCTCAAAGCAAAGAAACTTTATGATG CTATTCTTTTATCGGGTTTTCTGACTGTAACTGATAATGAGACAGAAGCAC CTCTCTCTTCTGCTATGATAAAGCGATATAATGAACACAAAGAAGATTTA GCGTTACTAAAGGAATATATAAGAAATATTTCACTAAAAACGTATAATGA AGTATTTAAAGATGACACCAAAAATGGTTATGCTGGTTATATTGATGGAA AAACAAATCAGGAAGATTTCTACGTATATCTAAAAAAACTATTGGCTGAA TTTGAAGGTGCGGATTATTTTCTTGAAAAAATTGATCGAGAAGATTTTTTG AGAAAGCAACGTACATTTGACAATGGTTCGATACCATATCAGATTCATCTT CAAGAAATGAGAGCAATTCTTGATAAGCAAGCTAAATTTTATCCTTTCTTG GCTAAAAATAAAGAAAGAATCGAGAAGATTTTAACCTTCCGAATTCCTTA TTATGTAGGTCCACTTGCGAGAGGGAATAGTGATTTTGCCTGGTCAATAA GAAAACGAAATGAAAAAATTACACCTTGGAATTTTGAGGACGTTATTGAC

AAAGAATCTTCGGCAGAGGCCTTCATTAATCGAATGACTAGTTTTGATTTG
TATTTGCCAGAAGAGAAGGTACTTCCAAAGCATAGTCTCTTATACGAAAC
TTTTAATGTATATAATGAATTAACAAAAGTTAGATTTATTGCCGAAAGTAT
GAGAGATTATCAATTTTTAGATAGTAAGCAGAAGAAAGATATTGTTAGAC
TTTATTTTAAAGATAAAAGGAAAGTTACTGATAAGGATATTATTGAATATT
TACATGCAATTTATGGGTATGATGGAATTGAATTAAAAGGCATAGAGAAA
CAGTTTAATTCTAGTTTATCTACTTATCACGATCTTTTAAATATTATTAATG
ATAAAGAGTTTTTGGATGATAGTTCAAATGAAGCGATTATCGAAGAAATT
ATCCATACTTTGACAATTTTTGAAGATAGAGAGATGATAAAACAACGTCT
TTCAAAATTTGAGAATATATTCGATAAATCCGTTTTGAAAAAGTTATCTCG
TAGACATTACACTGGCTGGGGTAAGTTATCTGCTAAGCTTATTAATGGTAT
TCGAGATGAAAAATCTGGTAATACTATTCTTGATTACTTAATTGATGATGG
TATTTCTAACCGTAATTTCATGCAACTTATTCACGATGATGCTCTTTCTTTT
AAAAAGAAGATACAGAAAGCACAAATTATTGGTGACGAAGATAAAGGTA
ATATTAAAGAGGTCGTTAAGTCTTTGCCAGGTAGTCCTGCGATTAAAAAA
GGTATTTTACAAAGCATAAAAATTGTAGATGAATTGGTCAAAGTAATGGG
AGGAAGAAAACCCGAGTCAATTGTTGTTGAGATGGCTCGTGAAAATCAAT
ATACCAATCAAGGTAAGTCTAATTCCCAACAACGCTTGAAACGTTTAGAA
AAATCTCTCAAAGAGTTAGGTAGTAAGATACTTAAGGAAAATATTCCTGC
AAAACTTTCTAAAATAGACAATAACGCACTTCAAAATGATCGACTTTACTT
ATACTATCTTCAAAATGGAAAAGATATGTATACCGGAGATGATTTAGATA
TTGATAGATTAAGTAATTATGATATTGATCATATTATTCCTCAAGCTTTTTT
GAAAGATAATTCTATTGACAATAAAGTACTTGTTTCATCTGCTAGTAACCG
TGGTAAATCAGATGATGTTCCAAGTTTAGAGGTTGTCAAAAAAAGAAAGA
CATTTTGGTATCAATTATTGAAATCAAAATTAATTTCTCAACGAAAATTTG
ATAATCTGACAAAAGCTGAACGGGGAGGATTGTCACCTGAGGACAAAGCT
GGTTTTATTCAACGCCAGTTGGTTGAAACACGTCAAATAACAAAACATGT
AGCTCGTTTACTTGATGAGAAATTTAATAATAAAAAAGATGAAAATAATA
GAGCGGTACGAACAGTAAAAATTATTACCTTGAAATCTACCTTAGTTTCTC
AATTTCGTAAGGATTTTGAACTTTATAAAGTTCGTGAAATCAATGATTTTC
ATCATGCTCATGATGCTTACTTGAATGCCGTTGTAGCAAGTGCTTTACTTA
AGAAATACCCTAAACTAGAGCCAGAATTTGTGTACGGTGATTATCCAAAA
TACAATAGTTTTAGAGAAAGAAAGTCCGCTACAGAAAAGGTATATTTCTA
TTCAAATATCATGAATATCTTTAAAAAATCTATTTCTTTAGCTGATGGTAG

AGTTATTGAAAGACCACTTATTGAGGTAAATGAGGAGACCGGCGAATCCG
TTTGGAATAAAGAATCTGATTTAGCAACTGTAAGGAGAGTACTCTCTTATC
CGCAAGTAAATGTTGTGAAAAAAGTTGAGGAACAGAATCACGGATTGGAT
AGAGGAAAACCAAAGGGATTGTTTAATGCAAATCTTTCCTCAAAGCCAAA
ACCAAATAGTAATGAAAATTTAGTAGGTGCTAAAGAGTATCTTGACCCCA
AAAAGTATGGGGGGTATGCTGGAATTTCTAATTCTTTTACTGTTCTTGTTA
AAGGGACAATTGAAAAAGGTGCTAAGAAAAAAATAACAAATGTACTAGA
ATTTCAAGGTATTTCTATTTTAGATAGGATTAATTATAGAAAAGATAAACT
TAATTTTTTACTTGAAAAAGGTTATAAAGATATTGAGTTAATTATTGAACT
ACCTAAATATAGTTTATTTGAACTTTCAGATGGTTCACGTCGTATGTTGGC
TAGTATTTTGTCAACGAATAATAAGAGGGGAGAGATTCACAAAGGAAATC
AGATTTTTCTTTCACAGAAGTTTGTGAAATTACTTTATCATGCTAAGAGAA
TAAGTAACACAATTAATGAGAATCATAGAAAATATGTTGAGAACCATAAA
AAAGAGTTTGAAGAATTATTTTACTACATTCTTGAGTTTAATGAGAATTAT
GTTGGAGCTAAAAAGAATGGTAAACTCTTAAACTCTGCCTTTCAATCTTGG
CAAAATCATAGTATAGATGAACTCTGTAGTAGTTTTATAGGACCTACCGG
AAGTGAAAGAAAGGGGCTATTTGAATTAACCTCTCGTGGAAGTGCTGCTG
ATTTTGAATTTTTAGGTGTTAAAATTCCAAGGTATAGAGACTATACCCCAT
CATCCCTATTAAAAGATGCCACACTTATTCATCAATCTGTTACAGGCCTCT
ATGAAACACGAATAGACCTTGCCAAACTAGGAGAGGGTTAA (SEQ ID NO: 4)

[00103] An example of a Cas9 protein comprising a nuclear localization signal (GGSGPPKKKRKV; SEQ ID NO: 5) at the C-terminus thereof has the following amino acid sequence:

[00104] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV
DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFK
SNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT

VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
TILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM
KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLS
DYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR
MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN
AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNI
MNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIV
KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY
SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE
QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA
ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID
LSQLGGDGGSGPPKKKRKV (SEQ ID NO: 6)

[00105] In some embodiments, a Cas9 nickase may be used in combination with guide sequence(s), e.g., two guide sequences, which target respectively sense and antisense strands of the DNA target. This combination allows both strands to be nicked and used to induce non-homologous end joining (NHEJ).

[00106] As a further example, two or more catalytic domains of Cas9 (RuvC I, RuvC II, and RuvC III) may be mutated to produce a mutated Cas9 substantially lacking all DNA cleavage activity. In some embodiments, a D10A mutation is combined with one or more of H840A, N854A, or N863A mutations to produce a Cas9 enzyme substantially lacking all DNA cleavage activity (where the amino acid numbering is as in SEQ ID NO: 1). In some embodiments, a CRISPR enzyme is considered to substantially lack all DNA cleavage activity when the DNA cleavage activity of the mutated enzyme is less than about 25%, 10%, 5%, 1%, 0.1%, 0.01%, or lower with respect to its non-mutated form. Other mutations may be useful; where the Cas9 or other CRISPR enzyme is from a species other than *S. pyogenes*, mutations in corresponding amino acids may be made to achieve similar effects.

[00107] In some embodiments, an enzyme coding sequence encoding a CRISPR enzyme is codon optimized for expression in particular cells, such as

eukaryotic cells. The eukaryotic cells may be those of or derived from a particular organism, such as a mammal, including but not limited to human, mouse, rat, rabbit, dog, or non-human primate. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g. about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the "Codon Usage Database", and these tables can be adapted in a number of ways. See Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, Pa.), are also available. In some embodiments, one or more codons (e.g. 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a CRISPR enzyme corresponds to the most frequently used codon for a particular amino acid.

[00108] In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of a CRISPR complex to the target sequence. In some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g. the Burrows Wheeler

Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). In some embodiments, a guide sequence is about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length. In some embodiments, a guide sequence is less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length. The ability of a guide sequence to direct sequence-specific binding of a CRISPR complex to a target sequence may be assessed by any suitable assay. For example, the components of a CRISPR system sufficient to form a CRISPR complex, including the guide sequence to be tested, may be provided to a host cell having the corresponding target sequence, such as by transfection with vectors encoding the components of the CRISPR sequence, followed by an assessment of preferential cleavage within the target sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target polynucleotide sequence may be evaluated in a test tube by providing the target sequence, components of a CRISPR complex, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art.

[00109] A guide sequence may be selected to target any target sequence. In some embodiments, the target sequence is a sequence within a genome of a cell. Exemplary target sequences include those that are unique in the target genome. For example, for the S. pyogenes Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNXGG MMMMMMMMNNNNNNNNNNNNXXAGAAW where NNNNNNNNNNNNXGG (N is A, G, T, or C; and X can be a deoxynucleotide) has a single occurrence in the genome. A unique target sequence in a genome may include an S. pyogenes Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNXGG where NNNNNNNNNNNXGG (N is A, G, T, or C; and X can be a deoxynucleotide) has a single occurrence in the genome. For the *S. thermophilus* CRISPR1 Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNXXAGAAW (SEQ ID NO: 7) where NNNNNNNNNNNNXXAGAAW (SEQ ID NO: 8) (N is A, G, T, or C; X can be a

deoxynucleotide; and W is A or T) has a single occurrence in the genome. A unique target sequence in a genome may include an *S. thermophilus* CRISPR1 Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNXXAGAAW (SEQ ID NO: 9) where NNNNNNNNNNNXXAGAAW (SEQ ID NO: 10) (N is A, G, T, or C; X can be a deoxynucleotide; and W is A or T) has a single occurrence in the genome. For the S. pyogenes Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNXGGXG where NNNNNNNNNNNXGGXG (N is A, G, T, or C; and X can be a deoxynucleotide) has a single occurrence in the genome. A unique target sequence in a genome may include an S. pyogenes Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNXGGXG where NNNNNNNNNNNXGGXG (N is A, G, T, or C; and X can be a deoxynucleotide) has a single occurrence in the genome. In each of these sequences "M" may be A, G, T, or C, and need not be considered in identifying a sequence as unique.

[00110] In some embodiments, a guide sequence is selected to reduce the degree of secondary structure within the guide sequence. Secondary structure may be determined by any suitable polynucleotide folding algorithm. Some programs are based on calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example folding algorithm is the online webserver RNAfold, developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (see e.g. A. R. Gruber et al., 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62). Further algorithms may be found in U.S. application Ser. No. 61/836,080; incorporated herein by reference.

[00111] Aspects of the present subject matter relate to delivery of CRISPR/ CRISPR/CRISPR from *Prevotella* and *Francisella* 1 (Cpf1) gene editing complexes or components thereof (*e.g.*, CPf1 proteins). Examples of human codon optimized Cpf1-family proteins are provided below.

Human Codon Optimized Cpf1-family Proteins

[00112] Non-limiting examples of Cpf1-family protein sequences, and aspects of CRISPR/Cpf1 gene-editing, are described in Zetsche et al., Cell 163, 759–771, October 22, 2015, the entire content of which is incorporated herein by reference.

[00113] *Francisella tularensis* subsp. *Novicida U112* (FnCpf1; pY004)), including NLS and HA tag:

[00114] MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEK RAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQ KDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDN GIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIV DDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRV FSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQIN DKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTV EEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLE YITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQC RFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNK IRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVM NKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSE DILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFR FSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFS AYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAK EAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEI NLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNY HDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF EDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQ LTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFS KFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHN WDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTIL QMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKG LMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNNKRPAATKKAGQAKKKKG SYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 11)

[00115] SEQ ID NO: 11 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a Human influenza hemagglutinin (HA) tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00116] SEQ ID NO: 11 may be encoded by the following nucleotide sequence:

[00117] ATGAGCATCTACCAGGAGTTCGTCAACAAGTATTCACTG
AGTAAGACACTGCGGTTCGAGCTGATCCCACAGGGCAAGACACTGGAGA
ACATCAAGGCCCGAGGCCTGATTCTGGACGATGAGAAGCGGGCAAAAGA
CTATAAGAAAGCCAAGCAGATCATTGATAAATACCACCAGTTCTTTATCG
AGGAAATTCTGAGCTCCGTGTGCATCAGTGAGGATCTGCTGCAGAATTAC
TCAGACGTGTACTTCAAGCTGAAGAAGAGCGACGATGACAACCTGCAGAA
GGACTTCAAGTCCGCCAAGGACACCATCAAGAAACAGATTAGCGAGTACA
TCAAGGACTCCGAAAAGTTTAAAAATCTGTTCAACCAGAATCTGATCGAT
GCTAAGAAAGGCCAGGAGTCCGACCTGATCCTGTGGCTGAAACAGTCTAA
GGACAATGGGATTGAACTGTTCAAGGCTAACTCCGATATCACTGATATTG
ACGAGGCACTGGAAATCATCAAGAGCTTCAAGGGATGGACCACATACTTT
AAAGGCTTCCACGAGAACCGCAAGAACGTGTACTCCAGCAACGACATTCC
TACCTCCATCATCTACCGAATCGTCGATGACAATCTGCCAAAGTTCCTGGA
GAACAAGGCCAAATATGAATCTCTGAAGGACAAAGCTCCCGAGGCAATTA
ATTACGAACAGATCAAGAAAGATCTGGCTGAGGAACTGACATTCGATATC
GACTATAAGACTAGCGAGGTGAACCAGAGGGTCTTTTCCCTGGACGAGGT
GTTTGAAATCGCCAATTTCAACAATTACCTGAACCAGTCCGGCATTACTAA
ATTCAATACCATCATTGGCGGGAAGTTTGTGAACGGGGAGAATACCAAGC
GCAAGGGAATTAACGAATACATCAATCTGTATAGCCAGCAGATCAACGAC
AAAACTCTGAAGAAATACAAGATGTCTGTGCTGTTCAAACAGATCCTGAG
TGATACCGAGTCCAAGTCTTTTGTCATTGATAAACTGGAAGATGACTCAG
ACGTGGTCACTACCATGCAGAGCTTTTATGAGCAGATCGCCGCTTTCAAG
ACAGTGGAGGAAAAATCTATTAAGGAAACTCTGAGTCTGCTGTTCGATGA
CCTGAAAGCCCAGAAGCTGGACCTGAGTAAGATCTACTTCAAAAACGATA
AGAGTCTGACAGACCTGTCACAGCAGGTGTTTGATGACTATTCCGTGATTG
GGACCGCCGTCCTGGAGTACATTACACAGCAGATCGCTCCAAAGAACCTG
GATAATCCCTCTAAGAAAGAGCAGGAACTGATCGCTAAGAAAACCGAGA
AGGCAAAATATCTGAGTCTGGAAACAATTAAGCTGGCACTGGAGGAGTTC
AACAAGCACAGGGATATTGACAAACAGTGCCGCTTTGAGGAAATCCTGGC
CAACTTCGCAGCCATCCCCATGATTTTTGATGAGATCGCCCAGAACAAAG
ACAATCTGGCTCAGATCAGTATTAAGTACCAGAACCAGGGCAAGAAAGAC
CTGCTGCAGGCTTCAGCAGAAGATGACGTGAAAGCCATCAAGGATCTGCT
GGACCAGACCAACAATCTGCTGCACAAGCTGAAAATCTTCCATATTAGTC
AGTCAGAGGATAAGGCTAATATCCTGGATAAAGACGAACACTTCTACCTG

GTGTTCGAGGAATGTTACTTCGAGCTGGCAAACATTGTCCCCCTGTATAAC
AAGATTAGGAACTACATCACACAGAAGCCTTACTCTGACGAGAAGTTTAA
ACTGAACTTCGAAAATAGTACCCTGGCCAACGGGTGGGATAAGAACAAG
GAGCCTGACAACACAGCTATCCTGTTCATCAAGGATGACAAGTACTATCT
GGGAGTGATGAATAAGAAAAACAATAAGATCTTCGATGACAAAGCCATT
AAGGAGAACAAAGGGGAAGGATACAAGAAAATCGTGTATAAGCTGCTGC
CCGGCGCAAATAAGATGCTGCCTAAGGTGTTCTTCAGCGCCAAGAGTATC
AAATTCTACAACCCATCCGAGGACATCCTGCGGATTAGAAATCACTCAAC
ACATACTAAGAACGGGAGCCCCCAGAAGGGATATGAGAAATTTGAGTTCA
ACATCGAGGATTGCAGGAAGTTTATTGACTTCTACAAGCAGAGCATCTCC
AAACACCCTGAATGGAAGGATTTTGGCTTCCGGTTTTCCGACACACAGAG
ATATAACTCTATCGACGAGTTCTACCGCGAGGTGGAAAATCAGGGGTATA
AGCTGACTTTTGAGAACATTTCTGAAAGTTACATCGACAGCGTGGTCAATC
AGGGAAAGCTGTACCTGTTCCAGATCTATAACAAAGATTTTTCAGCATAC
AGCAAGGGCAGACCAAACCTGCATACACTGTACTGGAAGGCCCTGTTCGA
TGAGAGGAATCTGCAGGACGTGGTCTATAAACTGAACGGAGAGGCCGAA
CTGTTTTACCGGAAGCAGTCTATTCCTAAGAAAATCACTCACCCAGCTAAG
GAGGCCATCGCTAACAAGAACAAGGACAATCCTAAGAAAGAGAGCGTGT
TCGAATACGATCTGATTAAGGACAAGCGGTTCACCGAAGATAAGTTCTTT
TTCCATTGTCCAATCACCATTAACTTCAAGTCAAGCGGCGCTAACAAGTTC
AACGACGAGATCAATCTGCTGCTGAAGGAAAAAGCAAACGATGTGCACA
TCCTGAGCATTGACCGAGGAGAGCGGCATCTGGCCTACTATACCCTGGTG
GATGGCAAAGGGAATATCATTAAGCAGGATACATTCAACATCATTGGCAA
TGACCGGATGAAAACCAACTACCACGATAAACTGGCTGCAATCGAGAAG
GATAGAGACTCAGCTAGGAAGGACTGGAAGAAAATCAACAACATTAAGG
AGATGAAGGAAGGCTATCTGAGCCAGGTGGTCCATGAGATTGCAAAGCTG
GTCATCGAATACAATGCCATTGTGGTGTTCGAGGATCTGAACTTCGGCTTT
AAGAGGGGGCGCTTTAAGGTGGAAAAACAGGTCTATCAGAAGCTGGAGA
AAATGCTGATCGAAAAGCTGAATTACCTGGTGTTTAAAGATAACGAGTTC
GACAAGACCGGAGGCGTCCTGAGAGCCTACCAGCTGACAGCTCCCTTTGA
AACTTTCAAGAAAATGGGAAAACAGACAGGCATCATCTACTATGTGCCAG
CCGGATTCACTTCCAAGATCTGCCCCGTGACCGGCTTTGTCAACCAGCTGT
ACCCTAAATATGAGTCAGTGAGCAAGTCCCAGGAATTTTTCAGCAAGTTC
GATAAGATCTGTTATAATCTGGACAAGGGGTACTTCGAGTTTTCCTTCGAT

TACAAGAACTTCGGCGACAAGGCCGCTAAGGGGAAATGGACCATTGCCTC
CTTCGGATCTCGCCTGATCAACTTTCGAAATTCCGATAAAAACCACAATTG
GGACACTAGGGAGGTGTACCCAACCAAGGAGCTGGAAAAGCTGCTGAAA
GACTACTCTATCGAGTATGGACATGGCGAATGCATCAAGGCAGCCATCTG
TGGCGAGAGTGATAAGAAATTTTTCGCCAAGCTGACCTCAGTGCTGAATA
CAATCCTGCAGATGCGGAACTCAAAGACCGGGACAGAACTGGACTATCTG
ATTAGCCCCGTGGCTGATGTCAACGGAAACTTCTTCGACAGCAGACAGGC
ACCCAAAAATATGCCTCAGGATGCAGACGCCAACGGGGCCTACCACATCG
GGCTGAAGGGACTGATGCTGCTGGGCCGGATCAAGAACAATCAGGAGGG
GAAGAAGCTGAACCTGGTCATTAAGAACGAGGAATACTTCGAGTTTGTCC
AGAATAGAAATAACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGC
AAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATC
CCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCT
AA (SEQ ID NO: 14)

[00118] *Lachnospiraceae bacterium* MC2017 (Lb3Cpf1; pY005), including NLS and HA tag:

[00119] MDYGNGQFERRAPLTKTITLRLKPIGETRETIREQKLLEQDA
AFRKLVETVTPIVDDCIRKIADNALCHFGTEYDFSCLGNAISKNDSKAIKKETE
KVEKLLAKVLTENLPDGLRKVNDINSAAFIQDTLTSFVQDDADKRVLIQELKG
KTVLMQRFLTTRITALTVWLPDRVFENFNIFIENAEKMRILLDSPLNEKIMKFD
PDAEQYASLEFYGQCLSQKDIDSYNLIISGIYADDEVKNPGINEIVKEYNQQIR
GDKDESPLPKLKKLHKQILMPVEKAFFVRVLSNDSDARSILEKILKDTEMLPS
KIIEAMKEADAGDIAVYGSRLHELSHVIYGDHGKLSQIIYDKESKRISELMETL
SPKERKESKKRLEGLEEHIRKSTYTFDELNRYAEKNVMAAYIAAVEESCAEIM
RKEKDLRTLLSKEDVKIRGNRHNTLIVKNYFNAWTVFRNLIRILRRKSEAEIDS
DFYDVLDDSVEVLSLTYKGENLCRSYITKKIGSDLKPEIATYGSALRPNSRWW
SPGEKFNVKFHTIVRRDGRLYYFILPKGAKPVELEDMDGDIECLQMRKIPNPTI
FLPKLVFKDPEAFFRDNPEADEFVFLSGMKAPVTITRETYEAYRYKLYTVGKL
RDGEVSEEEYKRALLQVLTAYKEFLENRMIYADLNFGFKDLEEYKDSSEFIKQ
VETHNTFMCWAKVSSSQLDDLVKSGNGLLFEIWSERLESYYKYGNEKVLRG
YEGVLLSILKDENLVSMRTLLNSRPMLVYRPKESSKPMVVHRDGSRVVDRFD
KDGKYIPPEVHDELYRFFNNLLIKEKLGEKARKILDNKKVKVKVLESERVKW
SKFYDEQFAVTFSVKKNADCLDTTKDLNAEVMEQYSESNRLILIRNTTDILYY
LVLDKNGKVLKQRSLNIINDGARDVDWKERFRQVTKDRNEGYNEWDYSRTS

NDLKEVYLNYALKEIAEAVIEYNAILIIEKMSNAFKDKYSFLDDVTFKGFETK
LLAKLSDLHFRGIKDGEPCSFTNPLQLCQNDSNKILQDGVIFMVPNSMTRSLD
PDTGFIFAINDHNIRTKKAKLNFLSKFDQLKVSSEGCLIMKYSGDSLPTHNTDN
RVWNCCCNHPITNYDRETKKVEFIEEPVEELSRVLEENGIETDTELNKLNERE
NVPGKVVDAIYSLVLNYLRGTVSGVAGQRAVYYSPVTGKKYDISFIQAMNLN
RKCDYYRIGSKERGEWTDFVAQLINKRPAATKKAGQAKKKKGSYPYDVPDY
AYPYDVPDYAYPYDVPDYA (SEQ ID NO: 15)

[00120] SEQ ID NO: 15 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00121] SEQ ID NO: 15 may be encoded by the following nucleotide sequence:

[00122] ATGGATTACGGCAACGGCCAGTTTGAGCGGAGAGCCCCC
CTGACCAAGACAATCACCCTGCGCCTGAAGCCTATCGGCGAGACACGGGA
GACAATCCGCGAGCAGAAGCTGCTGGAGCAGGACGCCGCCTTCAGAAAG
CTGGTGGAGACAGTGACCCCTATCGTGGACGATTGTATCAGGAAGATCGC
CGATAACGCCCTGTGCCACTTTGGCACCGAGTATGACTTCAGCTGTCTGGG
CAACGCCATCTCTAAGAATGACAGCAAGGCCATCAAGAAGGAGACAGAG
AAGGTGGAGAAGCTGCTGGCCAAGGTGCTGACCGAGAATCTGCCAGATG
GCCTGCGCAAGGTGAACGACATCAATTCCGCCGCCTTTATCCAGGATACA
CTGACCTCTTTCGTGCAGGACGATGCCGACAAGCGGGTGCTGATCCAGGA
GCTGAAGGGCAAGACCGTGCTGATGCAGCGGTTCCTGACCACACGGATCA
CAGCCCTGACCGTGTGGCTGCCCGACAGAGTGTTCGAGAACTTTAATATCT
TCATCGAGAACGCCGAGAAGATGAGAATCCTGCTGGACTCCCCTCTGAAT
GAGAAGATCATGAAGTTTGACCCAGATGCCGAGCAGTACGCCTCTCTGGA
GTTCTATGGCCAGTGCCTGTCTCAGAAGGACATCGATAGCTACAACCTGA
TCATCTCCGGCATCTATGCCGACGATGAGGTGAAGAACCCTGGCATCAAT
GAGATCGTGAAGGAGTACAATCAGCAGATCCGGGGCGACAAGGATGAGT
CCCCACTGCCCAAGCTGAAGAAGCTGCACAAGCAGATCCTGATGCCAGTG
GAGAAGGCCTTCTTTGTGCGCGTGCTGTCTAACGACAGCGATGCCCGGAG
CATCCTGGAGAAGATCCTGAAGGACACAGAGATGCTGCCCTCCAAGATCA
TCGAGGCCATGAAGGAGGCAGATGCAGGCGACATCGCCGTGTACGGCAG
CCGGCTGCACGAGCTGAGCCACGTGATCTACGGCGATCACGGCAAGCTGT

CCCAGATCATCTATGACAAGGAGTCCAAGAGGATCTCTGAGCTGATGGAG
ACACTGTCTCCAAAGGAGCGCAAGGAGAGCAAGAAGCGGCTGGAGGGCC
TGGAGGAGCACATCAGAAAGTCTACATACACCTTCGACGAGCTGAACAGG
TATGCCGAGAAGAATGTGATGGCAGCATACATCGCAGCAGTGGAGGAGTC
TTGTGCCGAGATCATGAGAAAGGAGAAGGATCTGAGGACCCTGCTGAGCA
AGGAGGACGTGAAGATCCGGGGCAACAGACACAATACACTGATCGTGAA
GAACTACTTTAATGCCTGGACCGTGTTCCGGAACCTGATCAGAATCCTGA
GGCGCAAGTCCGAGGCCGAGATCGACTCTGACTTCTACGATGTGCTGGAC
GATTCCGTGGAGGTGCTGTCTCTGACATACAAGGGCGAGAATCTGTGCCG
CAGCTATATCACCAAGAAGATCGGCTCCGACCTGAAGCCCGAGATCGCCA
CATACGGCAGCGCCCTGAGGCCTAACAGCCGCTGGTGGTCCCCAGGAGAG
AAGTTTAATGTGAAGTTCCACACCATCGTGCGGAGAGATGGCCGGCTGTA
CTATTTCATCCTGCCCAAGGGCGCCAAGCCTGTGGAGCTGGAGGACATGG
ATGGCGACATCGAGTGTCTGCAGATGAGAAAGATCCCTAACCCAACAATC
TTTCTGCCCAAGCTGGTGTTCAAGGACCCTGAGGCCTTCTTTAGGGATAAT
CCAGAGGCCGACGAGTTCGTGTTTCTGAGCGGCATGAAGGCCCCCGTGAC
AATCACCAGAGAGACATACGAGGCCTACAGGTATAAGCTGTATACCGTGG
GCAAGCTGCGCGATGGCGAGGTGTCCGAAGAGGAGTACAAGCGGGCCCT
GCTGCAGGTGCTGACCGCCTACAAGGAGTTTCTGGAGAACAGAATGATCT
ATGCCGACCTGAATTTCGGCTTTAAGGATCTGGAGGAGTATAAGGACAGC
TCCGAGTTTATCAAGCAGGTGGAGACACACAACACCTTCATGTGCTGGGC
CAAGGTGTCTAGCTCCCAGCTGGACGATCTGGTGAAGTCTGGCAACGGCC
TGCTGTTCGAGATCTGGAGCGAGCGCCTGGAGTCCTACTATAAGTACGGC
AATGAGAAGGTGCTGCGGGGCTATGAGGGCGTGCTGCTGAGCATCCTGAA
GGATGAGAACCTGGTGTCCATGCGGACCCTGCTGAACAGCCGGCCCATGC
TGGTGTACCGGCCAAAGGAGTCTAGCAAGCCTATGGTGGTGCACCGGGAT
GGCAGCAGAGTGGTGGACAGGTTTGATAAGGACGGCAAGTACATCCCCCC
TGAGGTGCACGACGAGCTGTATCGCTTCTTTAACAATCTGCTGATCAAGG
AGAAGCTGGGCGAGAAGGCCCGGAAGATCCTGGACAACAAGAAGGTGAA
GGTGAAGGTGCTGGAGAGCGAGAGAGTGAAGTGGTCCAAGTTCTACGAT
GAGCAGTTTGCCGTGACCTTCAGCGTGAAGAAGAACGCCGATTGTCTGGA
CACCACAAAGGACCTGAATGCCGAAGTGATGGAGCAGTATAGCGAGTCC
AACAGACTGATCCTGATCAGGAATACCACAGATATCCTGTACTATCTGGT
GCTGGACAAGAATGGCAAGGTGCTGAAGCAGAGATCCCTGAACATCATCA

ATGACGGCGCCAGGGATGTGGACTGGAAGGAGAGGTTCCGCCAGGTGAC
AAAGGATAGAAACGAGGGCTACAATGAGTGGGATTATTCCAGGACCTCTA
ACGACCTGAAGGAGGTGTACCTGAATTATGCCCTGAAGGAGATCGCCGAG
GCCGTGATCGAGTACAACGCCATCCTGATCATCGAGAAGATGTCTAATGC
CTTTAAGGACAAGTATAGCTTCCTGGACGACGTGACCTTCAAGGGCTTCG
AGACAAAGCTGCTGGCCAAGCTGAGCGATCTGCACTTTAGGGGCATCAAG
GACGGCGAGCCATGTTCCTTCACAAACCCCCTGCAGCTGTGCCAGAACGA
TTCTAATAAGATCCTGCAGGACGGCGTGATCTTTATGGTGCCAAATTCTAT
GACACGGAGCCTGGACCCCGACACCGGCTTCATCTTTGCCATCAACGACC
ACAATATCAGGACCAAGAAGGCCAAGCTGAACTTTCTGAGCAAGTTCGAT
CAGCTGAAGGTGTCCTCTGAGGGCTGCCTGATCATGAAGTACAGCGGCGA
TTCCCTGCCTACACACAACACCGACAATCGCGTGTGGAACTGCTGTTGCA
ATCACCCAATCACAAACTATGACCGGGAGACAAAGAAGGTGGAGTTCATC
GAGGAGCCCGTGGAGGAGCTGTCCCGCGTGCTGGAGGAGAATGGCATCG
AGACAGACACCGAGCTGAACAAGCTGAATGAGCGGGAGAACGTGCCTGG
CAAGGTGGTGGATGCCATCTACTCTCTGGTGCTGAATTATCTGCGCGGCAC
AGTGAGCGGAGTGGCAGGACAGAGGGCCGTGTACTATAGCCCTGTGACCG
GCAAGAAGTACGATATCTCCTTTATCCAGGCCATGAACCTGAATAGGAAG
TGTGACTACTATAGGATCGGCTCCAAGGAGAGGGGAGAGTGGACCGATTT
CGTGGCCCAGCTGATCAACAAAAGGCCGGCGGCCACGAAAAAGGCCGGC
CAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGC
TTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTA
TGCCTAA

[00123] *Butyrivibrio proteoclasticus* (BpCpf1; pY006), including NLS and HA tag:

[00124] MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEK
RAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQ
KDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDN
GIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIV
DDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRV
FSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQIN
DKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTV
EEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLE
YITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQC

RFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNK
IRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVM
NKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSE
DILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFR
FSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFS
AYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAK
EAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEI
NLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNY
HDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVF
EDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQ
LTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFS
KFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHN
WDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTIL
QMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKG
LMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNNKRPAATKKAGQAKKKKG
SYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 17)

[00125] SEQ ID NO: 17 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00126] SEQ ID NO: 17 may be encoded by the following nucleotide sequence:

[00127] ATGAGCATCTACCAGGAGTTCGTCAACAAGTATTCACTG
AGTAAGACACTGCGGTTCGAGCTGATCCCACAGGGCAAGACACTGGAGA
ACATCAAGGCCCGAGGCCTGATTCTGGACGATGAGAAGCGGGCAAAAGA
CTATAAGAAAGCCAAGCAGATCATTGATAAATACCACCAGTTCTTTATCG
AGGAAATTCTGAGCTCCGTGTGCATCAGTGAGGATCTGCTGCAGAATTAC
TCAGACGTGTACTTCAAGCTGAAGAAGAGCGACGATGACAACCTGCAGAA
GGACTTCAAGTCCGCCAAGGACACCATCAAGAAACAGATTAGCGAGTACA
TCAAGGACTCCGAAAAGTTTAAAAATCTGTTCAACCAGAATCTGATCGAT
GCTAAGAAAGGCCAGGAGTCCGACCTGATCCTGTGGCTGAAACAGTCTAA
GGACAATGGGATTGAACTGTTCAAGGCTAACTCCGATATCACTGATATTG
ACGAGGCACTGGAAATCATCAAGAGCTTCAAGGGATGGACCACATACTTT

AAAGGCTTCCACGAGAACCGCAAGAACGTGTACTCCAGCAACGACATTCC
TACCTCCATCATCTACCGAATCGTCGATGACAATCTGCCAAAGTTCCTGGA
GAACAAGGCCAAATATGAATCTCTGAAGGACAAAGCTCCCGAGGCAATTA
ATTACGAACAGATCAAGAAAGATCTGGCTGAGGAACTGACATTCGATATC
GACTATAAGACTAGCGAGGTGAACCAGAGGGTCTTTTCCCTGGACGAGGT
GTTTGAAATCGCCAATTTCAACAATTACCTGAACCAGTCCGGCATTACTAA
ATTCAATACCATCATTGGCGGGAAGTTTGTGAACGGGGAGAATACCAAGC
GCAAGGGAATTAACGAATACATCAATCTGTATAGCCAGCAGATCAACGAC
AAAACTCTGAAGAAATACAAGATGTCTGTGCTGTTCAAACAGATCCTGAG
TGATACCGAGTCCAAGTCTTTTGTCATTGATAAACTGGAAGATGACTCAG
ACGTGGTCACTACCATGCAGAGCTTTTATGAGCAGATCGCCGCTTTCAAG
ACAGTGGAGGAAAAATCTATTAAGGAAACTCTGAGTCTGCTGTTCGATGA
CCTGAAAGCCCAGAAGCTGGACCTGAGTAAGATCTACTTCAAAAACGATA
AGAGTCTGACAGACCTGTCACAGCAGGTGTTTGATGACTATTCCGTGATTG
GGACCGCCGTCCTGGAGTACATTACACAGCAGATCGCTCCAAAGAACCTG
GATAATCCCTCTAAGAAAGAGCAGGAACTGATCGCTAAGAAAACCGAGA
AGGCAAAATATCTGAGTCTGGAAACAATTAAGCTGGCACTGGAGGAGTTC
AACAAGCACAGGGATATTGACAAACAGTGCCGCTTTGAGGAAATCCTGGC
CAACTTCGCAGCCATCCCCATGATTTTTGATGAGATCGCCCAGAACAAAG
ACAATCTGGCTCAGATCAGTATTAAGTACCAGAACCAGGGCAAGAAAGAC
CTGCTGCAGGCTTCAGCAGAAGATGACGTGAAAGCCATCAAGGATCTGCT
GGACCAGACCAACAATCTGCTGCACAAGCTGAAAATCTTCCATATTAGTC
AGTCAGAGGATAAGGCTAATATCCTGGATAAAGACGAACACTTCTACCTG
GTGTTCGAGGAATGTTACTTCGAGCTGGCAAACATTGTCCCCCTGTATAAC
AAGATTAGGAACTACATCACACAGAAGCCTTACTCTGACGAGAAGTTTAA
ACTGAACTTCGAAAATAGTACCCTGGCCAACGGGTGGGATAAGAACAAG
GAGCCTGACAACACAGCTATCCTGTTCATCAAGGATGACAAGTACTATCT
GGGAGTGATGAATAAGAAAAACAATAAGATCTTCGATGACAAAGCCATT
AAGGAGAACAAAGGGGAAGGATACAAGAAAATCGTGTATAAGCTGCTGC
CCGGCGCAAATAAGATGCTGCCTAAGGTGTTCTTCAGCGCCAAGAGTATC
AAATTCTACAACCCATCCGAGGACATCCTGCGGATTAGAAATCACTCAAC
ACATACTAAGAACGGGAGCCCCCAGAAGGGATATGAGAAATTTGAGTTCA
ACATCGAGGATTGCAGGAAGTTTATTGACTTCTACAAGCAGAGCATCTCC
AAACACCCTGAATGGAAGGATTTTGGCTTCCGGTTTTCCGACACACAGAG

ATATAACTCTATCGACGAGTTCTACCGCGAGGTGGAAAATCAGGGGTATA
AGCTGACTTTTGAGAACATTTCTGAAAGTTACATCGACAGCGTGGTCAATC
AGGGAAAGCTGTACCTGTTCCAGATCTATAACAAAGATTTTTCAGCATAC
AGCAAGGGCAGACCAAACCTGCATACACTGTACTGGAAGGCCCTGTTCGA
TGAGAGGAATCTGCAGGACGTGGTCTATAAACTGAACGGAGAGGCCGAA
CTGTTTTACCGGAAGCAGTCTATTCCTAAGAAAATCACTCACCCAGCTAAG
GAGGCCATCGCTAACAAGAACAAGGACAATCCTAAGAAAGAGAGCGTGT
TCGAATACGATCTGATTAAGGACAAGCGGTTCACCGAAGATAAGTTCTTT
TTCCATTGTCCAATCACCATTAACTTCAAGTCAAGCGGCGCTAACAAGTTC
AACGACGAGATCAATCTGCTGCTGAAGGAAAAAGCAAACGATGTGCACA
TCCTGAGCATTGACCGAGGAGAGCGGCATCTGGCCTACTATACCCTGGTG
GATGGCAAAGGGAATATCATTAAGCAGGATACATTCAACATCATTGGCAA
TGACCGGATGAAAACCAACTACCACGATAAACTGGCTGCAATCGAGAAG
GATAGAGACTCAGCTAGGAAGGACTGGAAGAAAATCAACAACATTAAGG
AGATGAAGGAAGGCTATCTGAGCCAGGTGGTCCATGAGATTGCAAAGCTG
GTCATCGAATACAATGCCATTGTGGTGTTCGAGGATCTGAACTTCGGCTTT
AAGAGGGGGCGCTTTAAGGTGGAAAAACAGGTCTATCAGAAGCTGGAGA
AAATGCTGATCGAAAAGCTGAATTACCTGGTGTTTAAAGATAACGAGTTC
GACAAGACCGGAGGCGTCCTGAGAGCCTACCAGCTGACAGCTCCCTTTGA
AACTTTCAAGAAAATGGGAAAACAGACAGGCATCATCTACTATGTGCCAG
CCGGATTCACTTCCAAGATCTGCCCCGTGACCGGCTTTGTCAACCAGCTGT
ACCCTAAATATGAGTCAGTGAGCAAGTCCCAGGAATTTTTCAGCAAGTTC
GATAAGATCTGTTATAATCTGGACAAGGGGTACTTCGAGTTTTCCTTCGAT
TACAAGAACTTCGGCGACAAGGCCGCTAAGGGGAAATGGACCATTGCCTC
CTTCGGATCTCGCCTGATCAACTTTCGAAATTCCGATAAAAACCACAATTG
GGACACTAGGGAGGTGTACCCAACCAAGGAGCTGGAAAAGCTGCTGAAA
GACTACTCTATCGAGTATGGACATGGCGAATGCATCAAGGCAGCCATCTG
TGGCGAGAGTGATAAGAAATTTTTCGCCAAGCTGACCTCAGTGCTGAATA
CAATCCTGCAGATGCGGAACTCAAAGACCGGGACAGAACTGGACTATCTG
ATTAGCCCCGTGGCTGATGTCAACGGAAACTTCTTCGACAGCAGACAGGC
ACCCAAAAATATGCCTCAGGATGCAGACGCCAACGGGGCCTACCACATCG
GGCTGAAGGGACTGATGCTGCTGGGCCGGATCAAGAACAATCAGGAGGG
GAAGAAGCTGAACCTGGTCATTAAGAACGAGGAATACTTCGAGTTTGTCC
AGAATAGAAATAACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGC

AAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATC
CCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCT
AA (SEQ ID NO: 18)

[00128] *Peregrinibacteria bacterium* GW2011_GWA_33_10 (PeCpf1; pY007), including NLS and HA tag:

[00129] MSNFFKNFTNLYELSKTLRFELKPVGDTLTNMKDHLEYDE
KLQTFLKDQNIDDAYQALKPQFDEIHEEFITDSLESKKAKEIDFSEYLDLFQEK
KELNDSEKKLRNKIGETFNKAGEKWKKEKYPQYEWKKGSKIANGADILSCQ
DMLQFIKYKNPEDEKIKNYIDDTLKGFFTYFGGFNQNRANYYETKKEASTAV
ATRIVHENLPKFCDNVIQFKHIIKRKKDGTVEKTERKTEYLNAYQYLKNNNKI
TQIKDAETEKMIESTPIAEKIFDVYYFSSCLSQKQIEEYNRIIGHYNLLINLYNQ
AKRSEGKHLSANEKKYKDLPKFKTLYKQIGCGKKKDLFYTIKCDTEEEANKS
RNEGKESHSVEEIINKAQEAINKYFKSNNDCENINTVPDFINYILTKENYEGVY
WSKAAMNTISDKYFANYHDLQDRLKEAKVFQKADKKSEDDIKIPEAIELSGL
FGVLDSLADWQTTLFKSSILSNEDKLKIITDSQTPSEALLKMIFNDIEKNMESFL
KETNDIITLKKYKGNKEGTEKIKQWFDYTLAINRMLKYFLVKENKIKGNSLDT
NISEALKTLIYSDDAEWFKWYDALRNYLTQKPQDEAKENKLKLNFDNPSLAG
GWDVNKECSNFCVILKDKNEKKYLAIMKKGENTLFQKEWTEGRGKNLTKKS
NPLFEINNCEILSKMEYDFWADVSKMIPKCSTQLKAVVNHFKQSDNEFIFPIG
YKVTSGEKFREECKISKQDFELNNKVFNKNELSVTAMRYDLSSTQEKQYIKA
FQKEYWELLFKQEKRDTKLTNNEIFNEWINFCNKKYSELLSWERKYKDALTN
WINFCKYFLSKYPKTTLFNYSFKESENYNSLDEFYRDVDICSYKLNINTTINKS
ILDRLVEEGKLYLFEIKNQDSNDGKSIGHKNNLHTIYWNAIFENFDNRPKLNG
EAEIFYRKAISKDKLGIVKGKKTKNGTEIIKNYRFSKEKFILHVPITLNFCSNNE
YVNDIVNTKFYNFSNLHFLGIDRGEKHLAYYSLVNKNGEIVDQGTLNLPFTD
KDGNQRSIKKEKYFYNKQEDKWEAKEVDCWNYNDLLDAMASNRDMARKN
WQRIGTIKEAKNGYVSLVIRKIADLAVNNERPAFIVLEDLNTGFKRSRQKIDKS
VYQKFELALAKKLNFLVDKNAKRDEIGSPTKALQLTPPVNNYGDIENKKQAG
IMLYTRANYTSQTDPATGWRKTIYLKAGPEETTYKKDGKIKNKSVKDQIIETF
TDIGFDGKDYYFEYDKGEFVDEKTGEIKPKKWRLYSGENGKSLDRFRGEREK
DKYEWKIDKIDIVKILDDLFVNFDKNISLLKQLKEGVELTRNNEHGTGESLRF
AINLIQQIRNTGNNERDNDFILSPVRDENGKHFDSREYWDKETKGEKISMPSS
GDANGAFNIARKGIIMNAHILANSDSKDLSLFVSDEEWDLHLNNKTEWKKQL

NIFSSRKAMAKRKKKRPAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYA YPYDVPDYA (SEQ ID NO: 19)

[00130] SEQ ID NO: 19 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00131] SEQ ID NO: 19 may be encoded by the following nucleotide sequence:

[00132] ATGTCCAACTTCTTTAAGAATTTCACCAACCTGTATGAGC TGTCCAAGACACTGAGGTTTGAGCTGAAGCCCGTGGGCGACACCCTGACA AACATGAAGGACCACCTGGAGTACGATGAGAAGCTGCAGACCTTCCTGAA GGATCAGAATATCGACGATGCCTATCAGGCCCTGAAGCCTCAGTTCGACG AGATCCACGAGGAGTTTATCACAGATTCTCTGGAGAGCAAGAAGGCCAAG GAGATCGACTTCTCCGAGTACCTGGATCTGTTTCAGGAGAAGAAGGAGCT GAACGACTCTGAGAAGAAGCTGCGCAACAAGATCGGCGAGACATTCAAC AAGGCCGGCGAGAAGTGGAAGAAGGAGAAGTACCCTCAGTATGAGTGGA AGAAGGGCTCCAAGATCGCCAATGGCGCCGACATCCTGTCTTGCCAGGAT ATGCTGCAGTTTATCAAGTATAAGAACCCAGAGGATGAGAAGATCAAGAA TTACATCGACGATACACTGAAGGGCTTCTTTACCTATTTCGGCGGCTTTAA TCAGAACAGGGCCAACTACTATGAGACAAAGAAGGAGGCCTCCACCGCA GTGGCAACAAGGATCGTGCACGAGAACCTGCCAAAGTTCTGTGACAATGT GATCCAGTTTAAGCACATCATCAAGCGGAAGAAGGATGGCACCGTGGAG AAAACCGAGAGAAAGACCGAGTACCTGAACGCCTACCAGTATCTGAAGA ACAATAACAAGATCACACAGATCAAGGACGCCGAGACAGAGAAGATGAT CGAGTCTACACCCATCGCCGAGAAGATCTTCGACGTGTACTACTTCAGCA GCTGCCTGAGCCAGAAGCAGATCGAGGAGTACAACCGGATCATCGGCCA CTATAATCTGCTGATCAACCTGTATAACCAGGCCAAGAGATCTGAGGGCA AGCACCTGAGCGCCAACGAGAAGAAGTATAAGGACCTGCCTAAGTTCAA GACCCTGTATAAGCAGATCGGCTGCGGCAAGAAGAAGGACCTGTTTTACA CAATCAAGTGTGATACCGAGGAGGAGGCCAATAAGTCCCGGAACGAGGG CAAGGAGTCCCACTCTGTGGAGGAGATCATCAACAAGGCCCAGGAGGCC ATCAATAAGTACTTCAAGTCTAATAACGACTGTGAGAATATCAACACCGT GCCCGACTTCATCAACTATATCCTGACAAAGGAGAATTACGAGGGCGTGT ATTGGAGCAAGGCCGCCATGAACACCATCTCCGACAAGTACTTCGCCAAT

TATCACGACCTGCAGGATAGACTGAAGGAGGCCAAGGTGTTTCAGAAGGC
CGATAAGAAGTCCGAGGACGATATCAAGATCCCAGAGGCCATCGAGCTGT
CTGGCCTGTTCGGCGTGCTGGACAGCCTGGCCGATTGGCAGACCACACTG
TTTAAGTCTAGCATCCTGAGCAACGAGGACAAGCTGAAGATCATCACAGA
TTCCCAGACCCCCTCTGAGGCCCTGCTGAAGATGATCTTCAATGACATCGA
GAAGAACATGGAGTCCTTTCTGAAGGAGACAAACGATATCATCACCCTGA
AGAAGTATAAGGGCAATAAGGAGGGCACCGAGAAGATCAAGCAGTGGTT
CGACTATACACTGGCCATCAACCGGATGCTGAAGTACTTTCTGGTGAAGG
AGAATAAGATCAAGGGCAACTCCCTGGATACCAATATCTCTGAGGCCCTG
AAAACCCTGATCTACAGCGACGATGCCGAGTGGTTCAAGTGGTACGACGC
CCTGAGAAACTATCTGACCCAGAAGCCTCAGGATGAGGCCAAGGAGAAT
AAGCTGAAGCTGAATTTCGACAACCCATCTCTGGCCGGCGGCTGGGATGT
GAACAAGGAGTGCAGCAATTTTTGCGTGATCCTGAAGGACAAGAACGAG
AAGAAGTACCTGGCCATCATGAAGAAGGGCGAGAATACCCTGTTCCAGAA
GGAGTGGACAGAGGGCCGGGGCAAGAACCTGACAAAGAAGTCTAATCCA
CTGTTCGAGATCAATAACTGCGAGATCCTGAGCAAGATGGAGTATGACTT
TTGGGCCGACGTGAGCAAGATGATCCCCAAGTGTAGCACCCAGCTGAAGG
CCGTGGTGAACCACTTCAAGCAGTCCGACAATGAGTTCATCTTTCCTATCG
GCTACAAGGTGACAAGCGGCGAGAAGTTTAGGGAGGAGTGCAAGATCTC
CAAGCAGGACTTCGAGCTGAATAACAAGGTGTTTAATAAGAACGAGCTGA
GCGTGACCGCCATGCGCTACGATCTGTCCTCTACACAGGAGAAGCAGTAT
ATCAAGGCCTTCCAGAAGGAGTACTGGGAGCTGCTGTTTAAGCAGGAGAA
GCGGGACACCAAGCTGACAAATAACGAGATCTTCAACGAGTGGATCAATT
TTTGCAACAAGAAGTATAGCGAGCTGCTGTCCTGGGAGAGAAAGTACAAG
GATGCCCTGACCAATTGGATCAACTTCTGTAAGTACTTTCTGAGCAAGTAT
CCCAAGACCACACTGTTCAACTACTCTTTTAAGGAGAGCGAGAATTATAA
CTCCCTGGACGAGTTCTACCGGGACGTGGATATCTGTTCTTACAAGCTGAA
TATCAACACCACAATCAATAAGAGCATCCTGGATAGACTGGTGGAGGAGG
GCAAGCTGTACCTGTTTGAGATCAAGAATCAGGACAGCAACGATGGCAAG
TCCATCGGCCACAAGAATAACCTGCACACCATCTACTGGAACGCCATCTT
CGAGAATTTTGACAACAGGCCTAAGCTGAATGGCGAGGCCGAGATCTTCT
ATCGCAAGGCCATCTCCAAGGATAAGCTGGGCATCGTGAAGGGCAAGAA
AACCAAGAACGGCACCGAGATCATCAAGAATTACAGATTCAGCAAGGAG
AAGTTTATCCTGCACGTGCCAATCACCCTGAACTTCTGCTCCAATAACGAG

TATGTGAATGACATCGTGAACACAAAGTTCTACAATTTTTCCAACCTGCAC
TTTCTGGGCATCGATAGGGGCGAGAAGCACCTGGCCTACTATTCTCTGGTG
AATAAGAACGGCGAGATCGTGGACCAGGGCACACTGAACCTGCCTTTCAC
CGACAAGGATGGCAATCAGCGCAGCATCAAGAAGGAGAAGTACTTTTATA
ACAAGCAGGAGGACAAGTGGGAGGCCAAGGAGGTGGATTGTTGGAATTA
TAACGACCTGCTGGATGCCATGGCCTCTAACCGGGACATGGCCAGAAAGA
ATTGGCAGAGGATCGGCACCATCAAGGAGGCCAAGAACGGCTACGTGAG
CCTGGTCATCAGGAAGATCGCCGATCTGGCCGTGAATAACGAGCGCCCCG
CCTTCATCGTGCTGGAGGACCTGAATACAGGCTTTAAGCGGTCCAGACAG
AAGATCGATAAGAGCGTGTACCAGAAGTTCGAGCTGGCCCTGGCCAAGAA
GCTGAACTTTCTGGTGGACAAGAATGCCAAGCGCGATGAGATCGGCTCCC
CTACAAAGGCCCTGCAGCTGACCCCCCCTGTGAATAACTACGGCGACATT
GAGAACAAGAAGCAGGCCGGCATCATGCTGTATACCCGGGCCAATTATAC
CTCTCAGACAGATCCAGCCACAGGCTGGAGAAAGACCATCTATCTGAAGG
CCGGCCCCGAGGAGACAACATACAAGAAGGACGGCAAGATCAAGAACAA
GAGCGTGAAGGACCAGATCATCGAGACATTCACCGATATCGGCTTTGACG
GCAAGGATTACTATTTCGAGTACGACAAGGGCGAGTTTGTGGATGAGAAA
ACCGGCGAGATCAAGCCCAAGAAGTGGCGGCTGTACTCCGGCGAGAATG
GCAAGTCCCTGGACAGGTTCCGCGGAGAGAGGGAGAAGGATAAGTATGA
GTGGAAGATCGACAAGATCGATATCGTGAAGATCCTGGACGATCTGTTCG
TGAATTTTGACAAGAACATCAGCCTGCTGAAGCAGCTGAAGGAGGGCGTG
GAGCTGACCCGGAATAACGAGCACGGCACAGGCGAGTCCCTGAGATTCG
CCATCAACCTGATCCAGCAGATCCGGAATACCGGCAATAACGAGAGAGA
CAACGATTTCATCCTGTCCCCAGTGAGGGACGAGAATGGCAAGCACTTTG
ACTCTCGCGAGTACTGGGATAAGGAGACAAAGGGCGAGAAGATCAGCAT
GCCCAGCTCCGGCGATGCCAATGGCGCCTTCAACATCGCCCGGAAGGGCA
TCATCATGAACGCCCACATCCTGGCCAATAGCGACTCCAAGGATCTGTCC
CTGTTCGTGTCTGACGAGGAGTGGGATCTGCACCTGAATAACAAGACCGA
GTGGAAGAAGCAGCTGAACATCTTTTCTAGCAGGAAGGCCATGGCCAAGC
GCAAGAAGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAA
AGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTAC
GACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA
(SEQ ID NO: 20)

[00133] *Parcubacteria bacterium* GWC2011_GWC2_44_17 (PbCpf1; pY008), including NLS and HA tag:

[00134] MENIFDQFIGKYSLSKTLRFELKPVGKTEDFLKINKVFEKDQ TIDDSYNQAKFYFDSLHQKFIDAALASDKTSELSFQNFADVLEKQNKIILDKK REMGALRKRDKNAVGIDRLQKEINDAEDIIQKEKEKIYKDVRTLFDNEAESW KTYYQEREVDGKKITFSKADLKQKGADFLTAAGILKVLKYEFPEEKEKEFQA KNQPSLFVEEKENPGQKRYIFDSFDKFAGYLTKFQQTKKNLYAADGTSTAVA TRIADNFIIFHQNTKVFRDKYKNNHTDLGFDEENIFEIERYKNCLLQREIEHIKN ENSYNKIIGRINKKIKEYRDQKAKDTKLTKSDFPFFKNLDKQILGEVEKEKQLI EKTREKTEEDVLIERFKEFIENNEERFTAAKKLMNAFCNGEFESEYEGIYLKN KAINTISRRWFVSDRDFELKLPQQKSKNKSEKNEPKVKKFISIAEIKNAVEELD GDIFKAVFYDKKIIAQGGSKLEQFLVIWKYEFEYLFRDIERENGEKLLGYDSCL KIAKQLGIFPQEKEAREKATAVIKNYADAGLGIFQMMKYFSLDDKDRKNTPG QLSTNFYAEYDGYYKDFEFIKYYNEFRNFITKKPFDEDKIKLNFENGALLKGW DENKEYDFMGVILKKEGRLYLGIMHKNHRKLFQSMGNAKGDNANRYQKMI YKQIADASKDVPRLLLTSKKAMEKFKPSQEILRIKKEKTFKRESKNFSLRDLH ALIEYYRNCIPQYSNWSFYDFQFQDTGKYQNIKEFTDDVQKYGYKISFRDIDD EYINQALNEGKMYLFEVVNKDIYNTKNGSKNLHTLYFEHILSAENLNDPVFK LSGMAEIFQRQPSVNEREKITTQKNQCILDKGDRAYKYRRYTEKKIMFHMSL VLNTGKGEIKQVQFNKIINQRISSSDNEMRVNVIGIDRGEKNLLYYSVVKQNG EIIEQASLNEINGVNYRDKLIEREKERLKNRQSWKPVVKIKDLKKGYISHVIHK ICQLIEKYSAIVVLEDLNMRFKQIRGGIERSVYQQFEKALIDKLGYLVFKDNRD LRAPGGVLNGYQLSAPFVSFEKMRKQTGILFYTQAEYTSKTDPITGFRKNVYI SNSASLDKIKEAVKKFDAIGWDGKEQSYFFKYNPYNLADEKYKNSTVSKEW AIFASAPRIRRQKGEDGYWKYDRVKVNEEFEKLLKVWNFVNPKATDIKQEII KKEKAGDLQGEKELDGRLRNFWHSFIYLFNLVLELRNSFSLQIKIKAGEVIAV DEGVDFIASPVKPFFTTPNPYIPSNLCWLAVENADANGAYNIARKGVMILKKI REHAKKDPEFKKLPNLFISNAEWDEAARDWGKYAGTTALNLDHKRPAATKK AGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 21)

[00135] SEQ ID NO: 21 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00136] SEQ ID NO: 21 may be encoded by the following nucleotide sequence:

[00137] ATGGAGAACATCTTCGACCAGTTTATCGGCAAGTACAGC
CTGTCCAAGACCCTGAGATTCGAGCTGAAGCCCGTGGGCAAGACAGAGGA
CTTCCTGAAGATCAACAAGGTGTTTGAGAAGGATCAGACCATCGACGATA
GCTACAATCAGGCCAAGTTCTATTTTGATTCCCTGCACCAGAAGTTTATCG
ACGCCGCCCTGGCCTCCGATAAGACATCCGAGCTGTCTTTCCAGAACTTTG
CCGACGTGCTGGAGAAGCAGAATAAGATCATCCTGGATAAGAAGAGAGA
GATGGGCGCCCTGAGGAAGCGCGACAAGAACGCCGTGGGCATCGATAGG
CTGCAGAAGGAGATCAATGACGCCGAGGATATCATCCAGAAGGAGAAGG
AGAAGATCTACAAGGACGTGCGCACCCTGTTCGATAACGAGGCCGAGTCT
TGGAAAACCTACTATCAGGAGCGGGAGGTGGACGGCAAGAAGATCACCT
TCAGCAAGGCCGACCTGAAGCAGAAGGGCGCCGATTTTCTGACAGCCGCC
GGCATCCTGAAGGTGCTGAAGTATGAGTTCCCCGAGGAGAAGGAGAAGG
AGTTTCAGGCCAAGAACCAGCCCTCCCTGTTCGTGGAGGAGAAGGAGAAT
CCTGGCCAGAAGAGGTACATCTTCGACTCTTTTGATAAGTTCGCCGGCTAT
CTGACCAAGTTTCAGCAGACAAAGAAGAATCTGTACGCAGCAGACGGCAC
CAGCACAGCAGTGGCCACCCGCATCGCCGATAACTTTATCATCTTCCACC
AGAATACCAAGGTGTTCCGGGACAAGTACAAGAACAATCACACAGACCT
GGGCTTCGATGAGGAGAACATCTTTGAGATCGAGAGGTATAAGAATTGCC
TGCTGCAGCGCGAGATCGAGCACATCAAGAATGAGAATAGCTACAACAA
GATCATCGGCCGGATCAATAAGAAGATCAAGGAGTATCGGGACCAGAAG
GCCAAGGATACCAAGCTGACAAAGTCCGACTTCCCTTTCTTTAAGAACCT
GGATAAGCAGATCCTGGGCGAGGTGGAGAAGGAGAAGCAGCTGATCGAG
AAAACCCGGGAGAAAACCGAGGAGGACGTGCTGATCGAGCGGTTCAAGG
AGTTCATCGAGAACAATGAGGAGAGGTTCACCGCCGCCAAGAAGCTGATG
AATGCCTTCTGTAACGGCGAGTTTGAGTCCGAGTACGAGGGCATCTATCT
GAAGAATAAGGCCATCAACACAATCTCCCGGAGATGGTTCGTGTCTGACA
GAGATTTTGAGCTGAAGCTGCCTCAGCAGAAGTCCAAGAACAAGTCTGAG
AAGAATGAGCCAAAGGTGAAGAAGTTCATCTCCATCGCCGAGATCAAGA
ACGCCGTGGAGGAGCTGGACGGCGATATCTTTAAGGCCGTGTTCTACGAC
AAGAAGATCATCGCCCAGGGCGGCTCTAAGCTGGAGCAGTTCCTGGTCAT
CTGGAAGTACGAGTTTGAGTATCTGTTCCGGGACATCGAGAGAGAGAACG
GCGAGAAGCTGCTGGGCTATGATAGCTGCCTGAAGATCGCCAAGCAGCTG

GGCATCTTCCCACAGGAGAAGGAGGCCCGCGAGAAGGCAACCGCCGTGA
TCAAGAATTACGCCGACGCCGGCCTGGGCATCTTCCAGATGATGAAGTAT
TTTTCTCTGGACGATAAGGATCGGAAGAACACCCCCGGCCAGCTGAGCAC
AAATTTCTACGCCGAGTATGACGGCTACTACAAGGATTTCGAGTTTATCAA
GTACTACAACGAGTTTAGGAACTTCATCACCAAGAAGCCTTTCGACGAGG
ATAAGATCAAGCTGAACTTTGAGAATGGCGCCCTGCTGAAGGGCTGGGAC
GAGAACAAGGAGTACGATTTCATGGGCGTGATCCTGAAGAAGGAGGGCC
GCCTGTATCTGGGCATCATGCACAAGAACCACCGGAAGCTGTTTCAGTCC
ATGGGCAATGCCAAGGGCGACAACGCCAATAGATACCAGAAGATGATCT
ATAAGCAGATCGCCGACGCCTCTAAGGATGTGCCCAGGCTGCTGCTGACC
AGCAAGAAGGCCATGGAGAAGTTCAAGCCTTCCCAGGAGATCCTGAGAAT
CAAGAAGGAGAAAACCTTCAAGCGGGAGAGCAAGAACTTTTCCCTGAGA
GATCTGCACGCCCTGATCGAGTACTATAGGAACTGCATCCCTCAGTACAG
CAATTGGTCCTTTTATGACTTCCAGTTTCAGGATACCGGCAAGTACCAGAA
TATCAAGGAGTTCACAGACGATGTGCAGAAGTACGGCTATAAGATCTCCT
TTCGCGACATCGACGATGAGTATATCAATCAGGCCCTGAACGAGGGCAAG
ATGTACCTGTTCGAGGTGGTGAACAAGGATATCTATAACACCAAGAATGG
CTCCAAGAATCTGCACACACTGTACTTTGAGCACATCCTGTCTGCCGAGAA
CCTGAATGACCCAGTGTTCAAGCTGTCTGGCATGGCCGAGATCTTTCAGCG
GCAGCCCAGCGTGAACGAAAGAGAGAAGATCACCACACAGAAGAATCAG
TGTATCCTGGACAAGGGCGATAGAGCCTACAAGTATAGGCGCTACACCGA
GAAGAAGATCATGTTCCACATGAGCCTGGTGCTGAACACAGGCAAGGGCG
AGATCAAGCAGGTGCAGTTTAATAAGATCATCAACCAGAGGATCAGCTCC
TCTGACAACGAGATGAGGGTGAATGTGATCGGCATCGATCGCGGCGAGAA
GAACCTGCTGTACTATAGCGTGGTGAAGCAGAATGGCGAGATCATCGAGC
AGGCCTCCCTGAACGAGATCAATGGCGTGAACTACCGGGACAAGCTGATC
GAGAGGGAGAAGGAGCGCCTGAAGAACCGGCAGAGCTGGAAGCCTGTGG
TGAAGATCAAGGATCTGAAGAAGGGCTACATCTCCCACGTGATCCACAAG
ATCTGCCAGCTGATCGAGAAGTATTCTGCCATCGTGGTGCTGGAGGACCT
GAATATGAGATTCAAGCAGATCAGGGGAGGAATCGAGCGGAGCGTGTAC
CAGCAGTTCGAGAAGGCCCTGATCGATAAGCTGGGCTATCTGGTGTTTAA
GGACAACAGGGATCTGAGGGCACCAGGAGGCGTGCTGAATGGCTACCAG
CTGTCTGCCCCCTTTGTGAGCTTCGAGAAGATGCGCAAGCAGACCGGCAT
CCTGTTCTACACACAGGCCGAGTATACCAGCAAGACAGACCCAATCACCG

GCTTTCGGAAGAACGTGTATATCTCTAATAGCGCCTCCCTGGATAAGATCA
AGGAGGCCGTGAAGAAGTTCGACGCCATCGGCTGGGATGGCAAGGAGCA
GTCTTACTTCTTTAAGTACAACCCTTACAACCTGGCCGACGAGAAGTATAA
GAACTCTACCGTGAGCAAGGAGTGGGCCATCTTTGCCAGCGCCCCAAGAA
TCCGGAGACAGAAGGGCGAGGACGGCTACTGGAAGTATGATAGGGTGAA
AGTGAATGAGGAGTTCGAGAAGCTGCTGAAGGTCTGGAATTTTGTGAACC
CAAAGGCCACAGATATCAAGCAGGAGATCATCAAGAAGGAGAAGGCAGG
CGACCTGCAGGGAGAGAAGGAGCTGGATGGCCGGCTGAGAAACTTTTGG
CACTCTTTCATCTACCTGTTTAACCTGGTGCTGGAGCTGCGCAATTCTTTCA
GCCTGCAGATCAAGATCAAGGCAGGAGAAGTGATCGCAGTGGACGAGGG
CGTGGACTTCATCGCCAGCCCAGTGAAGCCCTTCTTTACCACACCCAACCC
TTACATCCCCTCCAACCTGTGCTGGCTGGCCGTGGAGAATGCAGACGCAA
ACGGAGCCTATAATATCGCCAGGAAGGGCGTGATGATCCTGAAGAAGATC
CGCGAGCACGCCAAGAAGGACCCCGAGTTCAAGAAGCTGCCAAACCTGTT
TATCAGCAATGCAGAGTGGGACGAGGCAGCCCGGGATTGGGGCAAGTAC
GCAGGCACCACAGCCCTGAACCTGGACCACAAAAGGCCGGCGGCCACGA
AAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGT
TCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGA
TGTCCCCGACTATGCCTAA (SEQ ID NO: 22)

[00138] *Smithella* sp. SC_K08D17 (SsCpf1; pY009), including NLS and HA tag:

[00139] MQTLFENFTNQYPVSKTLRFELIPQGKTKDFIEQKGLLKKDE
DRAEKYKKVKNIIDEYHKDFIEKSLNGLKLDGLEKYKTLYLKQEKDDKDKK
AFDKEKENLRKQIANAFRNNEKFKTLFAKELIKNDLMSFACEEDKKNVKEFE
AFTTYFTGFHQNRANMYVADEKRTAIASRLIHENLPKFIDNIKIFEKMKKEAP
ELLSPFNQTLKDMKDVIKGTTLEEIFSLDYFNKTLTQSGIDIYNSVIGGRTPEEG
KTKIKGLNEYINTDFNQKQTDKKKRQPKFKQLYKQILSDRQSLSFIAEAFKND
TEILEAIEKFYVNELLHFSNEGKSTNVLDAIKNAVSNLESFNLTKMYFRSGASL
TDVSRKVFGEWSIINRALDNYYATTYPIKPREKSEKYEERKEKWLKQDFNVS
LIQTAIDEYDNETVKGKNSGKVIADYFAKFCDDKETDLIQKVNEGYIAVKDLL
NTPCPENEKLGSNKDQVKQIKAFMDSIMDIMHFVRPLSLKDTDKEKDETFYS
LFTPLYDHLTQTIALYNKVRNYLTQKPYSTEKIKLNFENSTLLGGWDLNKETD
NTAIILRKDNLYYLGIMDKRHNRIFRNVPKADKKDFCYEKMVYKLLPGANK
MLPKVFFSQSRIQEFTPSAKLLENYANETHKKGDNFNLNHCHKLIDFFKDSIN

KHEDWKNFDFRFSATSTYADLSGFYHEVEHQGYKISFQSVADSFIDDLVNEG
KLYLFQIYNKDFSPFSKGKPNLHTLYWKMLFDENNLKDVVYKLNGEAEVFY
RKKSIAEKNTTIHKANESIINKNPDNPKATSTFNYDIVKDKRYTIDKFQFHIPIT
MNFKAEGIFNMNQRVNQFLKANPDINIIGIDRGERHLLYYALINQKGKILKQD
TLNVIANEKQKVDYHNLLDKKEGDRATARQEWGVIETIKELKEGYLSQVIHK
LTDLMIENNAIIVMEDLNFGFKRGRQKVEKQVYQKFEKMLIDKLNYLVDKN
KKANELGGLLNAFQLANKFESFQKMGKQNGFIFYVPAWNTSKTDPATGFIDF
LKPRYENLNQAKDFFEKFDSIRLNSKADYFEFAFDFKNFTEKADGGRTKWTV
CTTNEDRYAWNRALNNNRGSQEKYDITAELKSLFDGKVDYKSGKDLKQQIA
SQESADFFKALMKNLSITLSLRHNNGEKGDNEQDYILSPVADSKGRFFDSRKA
DDDMPKNADANGAYHIALKGLWCLEQISKTDDLKKVKLAISNKEWLEFVQT
LKGKRPAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
(SEQ ID NO: 23)

[00140] SEQ ID NO: 23 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00141] SEQ ID NO: 23 may be encoded by the following nucleotide sequence:

[00142] ATGCAGACCCTGTTTGAGAACTTCACAAATCAGTACCCA
GTGTCCAAGACCCTGCGCTTTGAGCTGATCCCCCAGGGCAAGACAAAGGA
CTTCATCGAGCAGAAGGGCCTGCTGAAGAAGGATGAGGACCGGGCCGAG
AAGTATAAGAAGGTGAAGAACATCATCGATGAGTACCACAAGGACTTCAT
CGAGAAGTCTCTGAATGGCCTGAAGCTGGACGGCCTGGAGAAGTACAAG
ACCCTGTATCTGAAGCAGGAGAAGGACGATAAGGATAAGAAGGCCTTTG
ACAAGGAGAAGGAGAACCTGCGCAAGCAGATCGCCAATGCCTTCCGGAA
CAATGAGAAGTTTAAGACACTGTTCGCCAAGGAGCTGATCAAGAACGATC
TGATGTCTTTCGCCTGCGAGGAGGACAAGAAGAATGTGAAGGAGTTTGAG
GCCTTCACCACATACTTCACCGGCTTCCACCAGAACCGCGCCAATATGTAC
GTGGCCGATGAGAAGAGAACAGCCATCGCCAGCAGGCTGATCCACGAGA
ACCTGCCAAAGTTTATCGACAATATCAAGATCTTCGAGAAGATGAAGAAG
GAGGCCCCCGAGCTGCTGTCTCCTTTCAACCAGACCCTGAAGGATATGAA
GGACGTGATCAAGGGCACCACACTGGAGGAGATCTTTAGCCTGGATTATT
TCAACAAGACCCTGACACAGAGCGGCATCGACATCTACAATTCCGTGATC

GGCGGCAGAACCCCTGAGGAGGGCAAGACAAAGATCAAGGGCCTGAACG
AGTACATCAATACCGACTTCAACCAGAAGCAGACAGACAAGAAGAAGCG
GCAGCCAAAGTTCAAGCAGCTGTATAAGCAGATCCTGAGCGATAGGCAGA
GCCTGTCCTTTATCGCCGAGGCCTTCAAGAACGACACCGAGATCCTGGAG
GCCATCGAGAAGTTTTACGTGAATGAGCTGCTGCACTTCAGCAATGAGGG
CAAGTCCACAAACGTGCTGGACGCCATCAAGAATGCCGTGTCTAACCTGG
AGAGCTTTAACCTGACCAAGATGTATTTCCGCTCCGGCGCCTCTCTGACAG
ACGTGAGCCGGAAGGTGTTTGGCGAGTGGAGCATCATCAATAGAGCCCTG
GACAACTACTATGCCACCACATATCCAATCAAGCCCAGAGAGAAGTCTGA
GAAGTACGAGGAGAGGAAGGAGAAGTGGCTGAAGCAGGACTTCAACGTG
AGCCTGATCCAGACCGCCATCGATGAGTACGACAACGAGACAGTGAAGG
GCAAGAACAGCGGCAAAGTGATCGCCGATTATTTTGCCAAGTTCTGCGAC
GATAAGGAGACAGACCTGATCCAGAAGGTGAACGAGGGCTACATCGCCG
TGAAGGATCTGCTGAATACACCCTGTCCTGAGAACGAGAAGCTGGGCAGC
AATAAGGACCAGGTGAAGCAGATCAAGGCCTTTATGGATTCTATCATGGA
CATCATGCACTTCGTGCGCCCCCTGAGCCTGAAGGATACCGACAAGGAGA
AGGATGAGACATTCTACTCCCTGTTCACACCTCTGTACGACCACCTGACCC
AGACAATCGCCCTGTATAACAAGGTGCGGAACTATCTGACCCAGAAGCCT
TACAGCACAGAGAAGATCAAGCTGAACTTCGAGAACAGCACCCTGCTGGG
CGGCTGGGATCTGAATAAGGAGACAGACAACACAGCCATCATCCTGAGG
AAGGATAACCTGTACTATCTGGGCATCATGGACAAGAGGCACAATCGCAT
CTTTCGGAACGTGCCCAAGGCCGATAAGAAGGACTTCTGCTACGAGAAGA
TGGTGTATAAGCTGCTGCCTGGCGCCAACAAGATGCTGCCAAAGGTGTTC
TTTTCTCAGAGCAGAATCCAGGAGTTTACCCCTTCCGCCAAGCTGCTGGAG
AACTACGCCAATGAGACACACAAGAAGGGCGATAATTTCAACCTGAATCA
CTGTCACAAGCTGATCGATTTCTTTAAGGACTCTATCAACAAGCACGAGG
ATTGGAAGAATTTCGACTTTAGGTTCAGCGCCACCTCCACCTACGCCGACC
TGAGCGGCTTTTACCACGAGGTGGAGCACCAGGGCTACAAGATCTCTTTT
CAGAGCGTGGCCGATTCCTTCATCGACGATCTGGTGAACGAGGGCAAGCT
GTACCTGTTCCAGATCTATAATAAGGACTTTTCCCCATTCTCTAAGGGCAA
GCCCAACCTGCACACCCTGTACTGGAAGATGCTGTTTGATGAGAACAATC
TGAAGGACGTGGTGTATAAGCTGAATGGCGAGGCCGAGGTGTTCTACCGC
AAGAAGAGCATTGCCGAGAAGAACACCACAATCCACAAGGCCAATGAGT
CCATCATCAACAAGAATCCTGATAACCCAAAGGCCACCAGCACCTTCAAC

TATGATATCGTGAAGGACAAGAGATACACCATCGACAAGTTTCAGTTCCA
CATCCCAATCACAATGAACTTTAAGGCCGAGGGCATCTTCAACATGAATC
AGAGGGTGAATCAGTTCCTGAAGGCCAATCCCGATATCAACATCATCGGC
ATCGACAGAGGCGAGAGGCACCTGCTGTACTATGCCCTGATCAACCAGAA
GGGCAAGATCCTGAAGCAGGATACCCTGAATGTGATCGCCAACGAGAAG
CAGAAGGTGGACTACCACAATCTGCTGGATAAGAAGGAGGGCGACCGCG
CAACCGCAAGGCAGGAGTGGGGCGTGATCGAGACAATCAAGGAGCTGAA
GGAGGGCTATCTGTCCCAGGTCATCCACAAGCTGACCGATCTGATGATCG
AGAACAATGCCATCATCGTGATGGAGGACCTGAACTTTGGCTTCAAGCGG
GGCAGACAGAAGGTGGAGAAGCAGGTGTATCAGAAGTTTGAGAAGATGC
TGATCGATAAGCTGAATTACCTGGTGGACAAGAATAAGAAGGCAAACGA
GCTGGGAGGCCTGCTGAACGCATTCCAGCTGGCCAATAAGTTTGAGTCCT
TCCAGAAGATGGGCAAGCAGAACGGCTTTATCTTCTACGTGCCCGCCTGG
AATACCTCTAAGACAGATCCTGCCACCGGCTTTATCGACTTCCTGAAGCCC
CGCTATGAGAACCTGAATCAGGCCAAGGATTTCTTTGAGAAGTTTGACTCT
ATCCGGCTGAACAGCAAGGCCGATTACTTTGAGTTCGCCTTTGACTTCAAG
AATTTCACCGAGAAGGCCGATGGCGGCAGAACCAAGTGGACAGTGTGCA
CCACAAACGAGGACAGATATGCCTGGAATAGGGCCCTGAACAATAACAG
GGGCAGCCAGGAGAAGTACGACATCACAGCCGAGCTGAAGTCCCTGTTCG
ATGGCAAGGTGGACTATAAGTCTGGCAAGGATCTGAAGCAGCAGATCGCC
AGCCAGGAGTCCGCCGACTTCTTTAAGGCCCTGATGAAGAACCTGTCCAT
CACCCTGTCTCTGAGACACAATAACGGCGAGAAGGGCGATAATGAGCAG
GACTACATCCTGTCCCCTGTGGCCGATTCTAAGGGCCGCTTCTTTGACTCC
CGGAAGGCCGACGATGACATGCCAAAGAATGCCGACGCCAACGGCGCCT
ATCACATCGCCCTGAAGGGCCTGTGGTGTCTGGAGCAGATCAGCAAGACC
GATGACCTGAAGAAGGTGAAGCTGGCCATCTCCAACAAGGAGTGGCTGG
AGTTCGTGCAGACACTGAAGGGCAAAAGGCCGGCGGCCACGAAAAAGGC
CGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATT
ACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCG
ACTATGCCTAA (SEQ ID NO: 24)

[00143] *Acidaminococcus* sp. BV3L6 (AsCpf1; pY010), including NLS and HA tag:

[00144] MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKA
RNDHYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNAL

IEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGT
VTTTEHENALLRSFDKFTTYFSGFYENRKNVFSAEDISTAIPHRIVQDNFPKFK
ENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLY
NQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDR
NTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSIDLTHIFISHK
KLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEII
SAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLY
HLLDWFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEK
FKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFE
PTEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIE
PLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKY
TKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLY
LFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSR
MKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARAL
LPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEH
PETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAA
RQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFGFKSKRTGI
AEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAKMGT
QSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFLHYDVKT
GDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPV
IENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVA
LIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPMDADANGAY
HIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRNKRPAATKKAGQA
KKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 25)

[00145] SEQ ID NO: 25 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00146] SEQ ID NO: 25 may be encoded by the following nucleotide sequence:

[00147] ATGACACAGTTCGAGGGCTTTACCAACCTGTATCAGGTG
AGCAAGACACTGCGGTTTGAGCTGATCCCACAGGGCAAGACCCTGAAGCA
CATCCAGGAGCAGGGCTTCATCGAGGAGGACAAGGCCCGCAATGATCACT
ACAAGGAGCTGAAGCCCATCATCGATCGGATCTACAAGACCTATGCCGAC

CAGTGCCTGCAGCTGGTGCAGCTGGATTGGGAGAACCTGAGCGCCGCCAT
CGACTCCTATAGAAAGGAGAAAACCGAGGAGACAAGGAACGCCCTGATC
GAGGAGCAGGCCACATATCGCAATGCCATCCACGACTACTTCATCGGCCG
GACAGACAACCTGACCGATGCCATCAATAAGAGACACGCCGAGATCTAC
AAGGGCCTGTTCAAGGCCGAGCTGTTTAATGGCAAGGTGCTGAAGCAGCT
GGGCACCGTGACCACAACCGAGCACGAGAACGCCCTGCTGCGGAGCTTCG
ACAAGTTTACAACCTACTTCTCCGGCTTTTATGAGAACAGGAAGAACGTG
TTCAGCGCCGAGGATATCAGCACAGCCATCCCACACCGCATCGTGCAGGA
CAACTTCCCCAAGTTTAAGGAGAATTGTCACATCTTCACACGCCTGATCAC
CGCCGTGCCCAGCCTGCGGGAGCACTTTGAGAACGTGAAGAAGGCCATCG
GCATCTTCGTGAGCACCTCCATCGAGGAGGTGTTTTCCTTCCCTTTTTATA
ACCAGCTGCTGACACAGACCCAGATCGACCTGTATAACCAGCTGCTGGGA
GGAATCTCTCGGGAGGCAGGCACCGAGAAGATCAAGGGCCTGAACGAGG
TGCTGAATCTGGCCATCCAGAAGAATGATGAGACAGCCCACATCATCGCC
TCCCTGCCACACAGATTCATCCCCCTGTTTAAGCAGATCCTGTCCGATAGG
AACACCCTGTCTTTCATCCTGGAGGAGTTTAAGAGCGACGAGGAAGTGAT
CCAGTCCTTCTGCAAGTACAAGACACTGCTGAGAAACGAGAACGTGCTGG
AGACAGCCGAGGCCCTGTTTAACGAGCTGAACAGCATCGACCTGACACAC
ATCTTCATCAGCCACAAGAAGCTGGAGACAATCAGCAGCGCCCTGTGCGA
CCACTGGGATACACTGAGGAATGCCCTGTATGAGCGGAGAATCTCCGAGC
TGACAGGCAAGATCACCAAGTCTGCCAAGGAGAAGGTGCAGCGCAGCCT
GAAGCACGAGGATATCAACCTGCAGGAGATCATCTCTGCCGCAGGCAAGG
AGCTGAGCGAGGCCTTCAAGCAGAAAACCAGCGAGATCCTGTCCCACGCA
CACGCCGCCCTGGATCAGCCACTGCCTACAACCCTGAAGAAGCAGGAGGA
GAAGGAGATCCTGAAGTCTCAGCTGGACAGCCTGCTGGGCCTGTACCACC
TGCTGGACTGGTTTGCCGTGGATGAGTCCAACGAGGTGGACCCCGAGTTC
TCTGCCCGGCTGACCGGCATCAAGCTGGAGATGGAGCCTTCTCTGAGCTTC
TACAACAAGGCCAGAAATTATGCCACCAAGAAGCCCTACTCCGTGGAGAA
GTTCAAGCTGAACTTTCAGATGCCTACACTGGCCTCTGGCTGGGACGTGA
ATAAGGAGAAGAACAATGGCGCCATCCTGTTTGTGAAGAACGGCCTGTAC
TATCTGGGCATCATGCCAAAGCAGAAGGGCAGGTATAAGGCCCTGAGCTT
CGAGCCCACAGAGAAAACCAGCGAGGGCTTTGATAAGATGTACTATGACT
ACTTCCCTGATGCCGCCAAGATGATCCCAAAGTGCAGCACCCAGCTGAAG
GCCGTGACAGCCCACTTTCAGACCCACACAACCCCCATCCTGCTGTCCAA

CAATTTCATCGAGCCTCTGGAGATCACAAAGGAGATCTACGACCTGAACA
ATCCTGAGAAGGAGCCAAAGAAGTTTCAGACAGCCTACGCCAAGAAAAC
CGGCGACCAGAAGGGCTACAGAGAGGCCCTGTGCAAGTGGATCGACTTCA
CAAGGGATTTTCTGTCCAAGTATACCAAGACAACCTCTATCGATCTGTCTA
GCCTGCGGCCATCCTCTCAGTATAAGGACCTGGGCGAGTACTATGCCGAG
CTGAATCCCCTGCTGTACCACATCAGCTTCCAGAGAATCGCCGAGAAGGA
GATCATGGATGCCGTGGAGACAGGCAAGCTGTACCTGTTCCAGATCTATA
ACAAGGACTTTGCCAAGGGCCACCACGGCAAGCCTAATCTGCACACACTG
TATTGGACCGGCCTGTTTTCTCCAGAGAACCTGGCCAAGACAAGCATCAA
GCTGAATGGCCAGGCCGAGCTGTTCTACCGCCCTAAGTCCAGGATGAAGA
GGATGGCACACCGGCTGGGAGAGAAGATGCTGAACAAGAAGCTGAAGGA
TCAGAAAACCCCAATCCCCGACACCCTGTACCAGGAGCTGTACGACTATG
TGAATCACAGACTGTCCCACGACCTGTCTGATGAGGCCAGGGCCCTGCTG
CCCAACGTGATCACCAAGGAGGTGTCTCACGAGATCATCAAGGATAGGCG
CTTTACCAGCGACAAGTTCTTTTTCCACGTGCCTATCACACTGAACTATCA
GGCCGCCAATTCCCCATCTAAGTTCAACCAGAGGGTGAATGCCTACCTGA
AGGAGCACCCCGAGACACCTATCATCGGCATCGATCGGGGCGAGAGAAA
CCTGATCTATATCACAGTGATCGACTCCACCGGCAAGATCCTGGAGCAGC
GGAGCCTGAACACCATCCAGCAGTTTGATTACCAGAAGAAGCTGGACAAC
AGGGAGAAGGAGAGGGTGGCAGCAAGGCAGGCCTGGTCTGTGGTGGGCA
CAATCAAGGATCTGAAGCAGGGCTATCTGAGCCAGGTCATCCACGAGATC
GTGGACCTGATGATCCACTACCAGGCCGTGGTGGTGCTGGAGAACCTGAA
TTTCGGCTTTAAGAGCAAGAGGACCGGCATCGCCGAGAAGGCCGTGTACC
AGCAGTTCGAGAAGATGCTGATCGATAAGCTGAATTGCCTGGTGCTGAAG
GACTATCCAGCAGAGAAAGTGGGAGGCGTGCTGAACCCATACCAGCTGAC
AGACCAGTTCACCTCCTTTGCCAAGATGGGCACCCAGTCTGGCTTCCTGTT
TTACGTGCCTGCCCCATATACATCTAAGATCGATCCCCTGACCGGCTTCGT
GGACCCCTTCGTGTGGAAAACCATCAAGAATCACGAGAGCCGCAAGCACT
TCCTGGAGGGCTTCGACTTTCTGCACTACGACGTGAAAACCGGCGACTTC
ATCCTGCACTTTAAGATGAACAGAAATCTGTCCTTCCAGAGGGGCCTGCC
CGGCTTTATGCCTGCATGGGATATCGTGTTCGAGAAGAACGAGACACAGT
TTGACGCCAAGGGCACCCCTTTCATCGCCGGCAAGAGAATCGTGCCAGTG
ATCGAGAATCACAGATTCACCGGCAGATACCGGGACCTGTATCCTGCCAA
CGAGCTGATCGCCCTGCTGGAGGAGAAGGGCATCGTGTTCAGGGATGGCT

CCAACATCCTGCCAAAGCTGCTGGAGAATGACGATTCTCACGCCATCGAC
ACCATGGTGGCCCTGATCCGCAGCGTGCTGCAGATGCGGAACTCCAATGC
CGCCACAGGCGAGGACTATATCAACAGCCCCGTGCGCGATCTGAATGGCG
TGTGCTTCGACTCCCGGTTTCAGAACCCAGAGTGGCCCATGGACGCCGAT
GCCAATGGCGCCTACCACATCGCCCTGAAGGGCCAGCTGCTGCTGAATCA
CCTGAAGGAGAGCAAGGATCTGAAGCTGCAGAACGGCATCTCCAATCAG
GACTGGCTGGCCTACATCCAGGAGCTGCGCAACAAAAGGCCGGCGGCCA
CGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGA
TGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATA
TGATGTCCCCGACTATGCCTAA (SEQ ID NO: 26)

[00148] *Lachnospiraceae bacterium* MA2020 (Lb2Cpf1; pY011), including NLS and HA tag:

[00149] MYYESLTKQYPVSKTIRNELIPIGKTLDNIRQNNILESDVKR
KQNYEHVKGILDEYHKQLINEALDNCTLPSLKIAAEIYLKNQKEVSDREDFNK
TQDLLRKEVVEKLKAHENFTKIGKKDILDLLEKLPSISEDDYNALESFRNFYT
YFTSYNKVRENLYSDKEKSSTVAYRLINENFPKFLDNVKSYRFVKTAGILADG
LGEEEQDSLFIVETFNKTLTQDGIDTYNSQVGKINSSINLYNQKNQKANGFRKI
PKMKMLYKQILSDREESFIDEFQSDEVLIDNVESYGSVLIESLKSSKVSAFFDA
LRESKGKNVYVKNDLAKTAMSNIVFENWRTFDDLLNQEYDLANENKKKDD
KYFEKRQKELKKNKSYSLEHLCNLSEDSCNLIENYIHQISDDIENIIINNETFLRI
VINEHDRSRKLAKNRKAVKAIKDFLDSIKVLERELKLINSSGQELEKDLIVYSA
HEELLVELKQVDSLYNMTRNYLTKKPFSTEKVKLNFNRSTLLNGWDRNKET
DNLGVLLLKDGKYYLGIMNTSANKAFVNPPVAKTEKVFKKVDYKLLPVPNQ
MLPKVFFAKSNIDFYNPSSEIYSNYKKGTHKKGNMFSLEDCHNLIDFFKESISK
HEDWSKFGFKFSDTASYNDISEFYREVEKQGYKLTYTDIDETYINDLIERNEL
YLFQIYNKDFSMYSKGKLNLHTLYFMMLFDQRNIDDVVYKLNGEAEVFYRP
ASISEDELIIHKAGEEIKNKNPNRARTKETSTFSYDIVKDKRYSKDKFTLHIPIT
MNFGVDEVKRFNDAVNSAIRIDENVNVIGIDRGERNLLYVVVIDSKGNILEQIS
LNSIINKEYDIETDYHALLDEREGGRDKARKDWNTVENIRDLKAGYLSQVVN
VVAKLVLKYNAIICLEDLNFGFKRGRQKVEKQVYQKFEKMLIDKLNYLVIDK
SREQTSPKELGGALNALQLTSKFKSFKELGKQSGVIYYVPAYLTSKIDPTTGF
ANLFYMKCENVEKSKRFFDGFDFIRFNALENVFEFGFDYRSFTQRACGINSK
WTVCTNGERIIKYRNPDKNNMFDEKVVVVTDEMKNLFEQYKIPYEDGRNVK
DMIISNEEAEFYRRLYRLLQQTLQMRNSTSDGTRDYIISPVKNKREAYFNSELS

DGSVPKDADANGAYNIARKGLWVLEQIRQKSEGEKINLAMTNAEWLEYAQT
HLLKRPAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
(SEQ ID NO: 27)

[00150] SEQ ID NO: 27 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00151] SEQ ID NO: 27 may be encoded by the following nucleotide sequence:

[00152] ATGTACTATGAGTCCCTGACCAAGCAGTACCCCGTGTCTA
AGACAATCCGGAATGAGCTGATCCCTATCGGCAAGACACTGGATAACATC
CGCCAGAACAATATCCTGGAGAGCGACGTGAAGCGGAAGCAGAACTACG
AGCACGTGAAGGGCATCCTGGATGAGTATCACAAGCAGCTGATCAACGAG
GCCCTGGACAATTGCACCCTGCCATCCCTGAAGATCGCCGCCGAGATCTA
CCTGAAGAATCAGAAGGAGGTGTCTGACAGAGAGGATTTCAACAAGACA
CAGGACCTGCTGAGGAAGGAGGTGGTGGAGAAGCTGAAGGCCCACGAGA
ACTTTACCAAGATCGGCAAGAAGGACATCCTGGATCTGCTGGAGAAGCTG
CCTTCCATCTCTGAGGACGATTACAATGCCCTGGAGAGCTTCCGCAACTTT
TACACCTATTTCACATCCTACAACAAGGTGCGGGAGAATCTGTATTCTGAT
AAGGAGAAGAGCTCCACAGTGGCCTACAGACTGATCAACGAGAATTTCCC
AAAGTTTCTGGACAATGTGAAGAGCTATAGGTTTGTGAAAACCGCAGGCA
TCCTGGCAGATGGCCTGGGAGAGGAGGAGCAGGACTCCCTGTTCATCGTG
GAGACATTCAACAAGACCCTGACACAGGACGGCATCGATACCTACAATTC
TCAAGTGGGCAAGATCAACTCTAGCATCAATCTGTATAACCAGAAGAATC
AGAAGGCCAATGGCTTCAGAAAGATCCCCAAGATGAAGATGCTGTATAAG
CAGATCCTGTCCGATAGGGAGGAGTCTTTCATCGACGAGTTTCAGAGCGA
TGAGGTGCTGATCGACAACGTGGAGTCTTATGGCAGCGTGCTGATCGAGT
CTCTGAAGTCCTCTAAGGTGAGCGCCTTCTTTGATGCCCTGAGAGAGTCTA
AGGGCAAGAACGTGTACGTGAAGAATGACCTGGCCAAGACAGCCATGAG
CAACATCGTGTTCGAGAATTGGAGGACCTTTGACGATCTGCTGAACCAGG
AGTACGACCTGGCCAACGAGAACAAGAAGAAGGACGATAAGTATTTCGA
GAAGCGCCAGAAGGAGCTGAAGAAGAATAAGAGCTACTCCCTGGAGCAC
CTGTGCAACCTGTCCGAGGATTCTTGTAACCTGATCGAGAATTATATCCAC
CAGATCTCCGACGATATCGAGAATATCATCATCAACAATGAGACATTCCT

GCGCATCGTGATCAATGAGCACGACAGGTCCCGCAAGCTGGCCAAGAACC
GGAAGGCCGTGAAGGCCATCAAGGACTTTCTGGATTCTATCAAGGTGCTG
GAGCGGGAGCTGAAGCTGATCAACAGCTCCGGCCAGGAGCTGGAGAAGG
ATCTGATCGTGTACTCTGCCCACGAGGAGCTGCTGGTGGAGCTGAAGCAG
GTGGACAGCCTGTATAACATGACCAGAAATTATCTGACAAAGAAGCCTTT
CTCTACCGAGAAGGTGAAGCTGAACTTTAATCGCAGCACACTGCTGAACG
GCTGGGATCGGAATAAGGAGACAGACAACCTGGGCGTGCTGCTGCTGAA
GGACGGCAAGTACTATCTGGGCATCATGAACACAAGCGCCAATAAGGCCT
TCGTGAATCCCCCTGTGGCCAAGACCGAGAAGGTGTTTAAGAAGGTGGAT
TACAAGCTGCTGCCAGTGCCCAACCAGATGCTGCCAAAGGTGTTCTTTGCC
AAGAGCAATATCGACTTCTATAACCCCTCTAGCGAGATCTACTCCAATTAT
AAGAAGGGCACCCACAAGAAGGGCAATATGTTTTCCCTGGAGGATTGTCA
CAACCTGATCGACTTCTTTAAGGAGTCTATCAGCAAGCACGAGGACTGGA
GCAAGTTCGGCTTTAAGTTCAGCGATACAGCCTCCTACAACGACATCTCCG
AGTTCTATCGCGAGGTGGAGAAGCAGGGCTACAAGCTGACCTATACAGAC
ATCGATGAGACATACATCAATGATCTGATCGAGCGGAACGAGCTGTACCT
GTTCCAGATCTATAATAAGGACTTTAGCATGTACTCCAAGGGCAAGCTGA
ACCTGCACACACTGTATTTCATGATGCTGTTTGATCAGCGCAATATCGACG
ACGTGGTGTATAAGCTGAACGGAGAGGCAGAGGTGTTCTATAGGCCAGCC
TCCATCTCTGAGGACGAGCTGATCATCCACAAGGCCGGCGAGGAGATCAA
GAACAAGAATCCTAACCGGGCCAGAACCAAGGAGACAAGCACCTTCAGC
TACGACATCGTGAAGGATAAGCGGTATAGCAAGGATAAGTTTACCCTGCA
CATCCCCATCACAATGAACTTCGGCGTGGATGAGGTGAAGCGGTTCAACG
ACGCCGTGAACAGCGCCATCCGGATCGATGAGAATGTGAACGTGATCGGC
ATCGACCGGGGCGAGAGAAATCTGCTGTACGTGGTGGTCATCGACTCTAA
GGGCAACATCCTGGAGCAGATCTCCCTGAACTCTATCATCAATAAGGAGT
ACGACATCGAGACAGATTATCACGCACTGCTGGATGAGAGGGAGGGCGG
CAGAGATAAGGCCCGGAAGGACTGGAACACCGTGGAGAATATCAGGGAC
CTGAAGGCCGGCTACCTGAGCCAGGTGGTGAACGTGGTGGCCAAGCTGGT
GCTGAAGTATAATGCCATCATCTGCCTGGAGGACCTGAACTTTGGCTTCAA
GAGGGGCCGCCAGAAGGTGGAGAAGCAGGTGTACCAGAAGTTCGAGAAG
ATGCTGATCGATAAGCTGAATTACCTGGTCATCGACAAGAGCCGCGAGCA
GACATCCCCTAAGGAGCTGGGAGGCGCCCTGAACGCACTGCAGCTGACCT
CTAAGTTCAAGAGCTTTAAGGAGCTGGGCAAGCAGTCCGGCGTGATCTAC

TATGTGCCTGCCTACCTGACCTCTAAGATCGATCCAACCACAGGCTTCGCC
AATCTGTTTTATATGAAGTGTGAGAACGTGGAGAAGTCCAAGAGATTCTT
TGACGGCTTTGATTTCATCAGGTTCAACGCCCTGGAGAACGTGTTCGAGTT
CGGCTTTGACTACCGGAGCTTCACCCAGAGGGCCTGCGGCATCAATTCCA
AGTGGACCGTGTGCACCAACGGCGAGCGCATCATCAAGTATCGGAATCCA
GATAAGAACAATATGTTCGACGAGAAGGTGGTGGTGGTGACCGATGAGAT
GAAGAACCTGTTTGAGCAGTACAAGATCCCCTATGAGGATGGCAGAAATG
TGAAGGACATGATCATCAGCAACGAGGAGGCCGAGTTCTACCGGAGACTG
TATAGGCTGCTGCAGCAGACCCTGCAGATGAGAAACAGCACCTCCGACGG
CACAAGGGATTACATCATCTCCCCTGTGAAGAATAAGAGAGAGGCCTACT
TCAACAGCGAGCTGTCCGACGGCTCTGTGCCAAAGGACGCCGATGCCAAC
GGCGCCTACAATATCGCCAGAAAGGGCCTGTGGGTGCTGGAGCAGATCAG
GCAGAAGAGCGAGGGCGAGAAGATCAATCTGGCCATGACCAACGCCGAG
TGGCTGGAGTATGCCCAGACACACCTGCTGAAAAGGCCGGCGGCCACGA
AAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGT
TCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGA
TGTCCCCGACTATGCCTAA (SEQ ID NO: 28)

[00153] *Candidatus Methanoplasma termitum* (CMtCpf1; pY012), including NLS and HA tag:

[00154] MNNYDEFTKLYPIQKTIRFELKPQGRTMEHLETFNFFEEDRD
RAEKYKILKEAIDEYHKKFIDEHLTNMSLDWNSLKQISEKYYKSREEKDKKV
FLSEQKRMRQEIVSEFKKDDRFKDLFSKKLFSELLKEEIYKKGNHQEIDALKSF
DKFSGYFIGLHENRKNMYSDGDEITAISNRIVNENFPKFLDNLQKYQEARKKY
PEWIIKAESALVAHNIKMDEVFSLEYFNKVLNQEGIQRYNLALGGYVTKSGE
KMMGLNDALNLAHQSEKSSKGRIHMTPLFKQILSEKESFSYIPDVFTEDSQLL
PSIGGFFAQIENDKDGNIFDRALELISSYAEYDTERIYIRQADINRVSNVIFGEW
GTLGGLMREYKADSINDINLERTCKKVDKWLDSKEFALSDVLEAIKRTGNND
AFNEYISKMRTAREKIDAARKEMKFISEKISGDEESIHIIKTLLDSVQQFLHFFN
LFKARQDIPLDGAFYAEFDEVHSKLFAIVPLYNKVRNYLTKNNLNTKKIKLNF
KNPTLANGWDQNKVYDYASLIFLRDGNYYLGIINPKRKKNIKFEQGSGNGPF
YRKMVYKQIPGPNKNLPRVFLTSTKGKKEYKPSKEIIEGYEADKHIRGDKFDL
DFCHKLIDFFKESIEKHKDWSKFNFYFSPTESYGDISEFYLDVEKQGYRMHFE
NISAETIDEYVEKGDLFLFQIYNKDFVKAATGKKDMHTIYWNAAFSPENLQD
VVVKLNGEAELFYRDKSDIKEIVHREGEILVNRTYNGRTPVPDKIHKKLTDYH

NGRTKDLGEAKEYLDKVRYFKAHYDITKDRRYLNDKIYFHVPLTLNFKANG
KKNLNKMVIEKFLSDEKAHIIGIDRGERNLLYYSIIDRSGKIIDQQSLNVIDGFD
YREKLNQREIEMKDARQSWNAIGKIKDLKEGYLSKAVHEITKMAIQYNAIVV
MEELNYGFKRGRFKVEKQIYQKFENMLIDKMNYLVFKDAPDESPGGVLNAY
QLTNPLESFAKLGKQTGILFYVPAAYTSKIDPTTGFVNLFNTSSKTNAQERKEF
LQKFESISYSAKDGGIFAFAFDYRKFGTSKTDHKNVWTAYTNGERMRYIKEK
KRNELFDPSKEIKEALTSSGIKYDGGQNILPDILRSNNNGLIYTMYSSFIAAIQM
RVYDGKEDYIISPIKNSKGEFFRTDPKRRELPIDADANGAYNIALRGELTMRAI
AEKFDPDSEKMAKLELKHKDWFEFMQTRGDKRPAATKKAGQAKKKKGSYP
YDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 29)

[00155] SEQ ID NO: 29 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00156] SEQ ID NO: 29 may be encoded by the following nucleotide sequence:

[00157] ATGAACAATTACGACGAGTTCACCAAGCTGTATCCTATC
CAGAAAACCATCCGGTTTGAGCTGAAGCCACAGGGCAGAACCATGGAGC
ACCTGGAGACATTCAACTTCTTTGAGGAGGACCGGGATAGAGCCGAGAAG
TATAAGATCCTGAAGGAGGCCATCGACGAGTACCACAAGAAGTTTATCGA
TGAGCACCTGACCAATATGTCCCTGGATTGGAACTCTCTGAAGCAGATCA
GCGAGAAGTACTATAAGAGCAGGGAGGAGAAGGACAAGAAGGTGTTCCT
GTCCGAGCAGAAGAGGATGCGCCAGGAGATCGTGTCTGAGTTTAAGAAG
GACGATCGCTTCAAGGACCTGTTTTCCAAGAAGCTGTTCTCTGAGCTGCTG
AAGGAGGAGATCTACAAGAAGGGCAACCACCAGGAGATCGACGCCCTGA
AGAGCTTCGATAAGTTTTCCGGCTATTTCATCGGCCTGCACGAGAATAGG
AAGAACATGTACTCCGACGGCGATGAGATCACCGCCATCTCCAATCGCAT
CGTGAATGAGAACTTCCCCAAGTTTCTGGATAACCTGCAGAAGTACCAGG
AGGCCAGGAAGAAGTATCCTGAGTGGATCATCAAGGCCGAGAGCGCCCT
GGTGGCCCACAATATCAAGATGGACGAGGTGTTCTCCCTGGAGTACTTTA
ATAAGGTGCTGAACCAGGAGGGCATCCAGCGGTACAACCTGGCCCTGGGC
GGCTATGTGACCAAGAGCGGCGAGAAGATGATGGGCCTGAATGATGCCCT
GAACCTGGCCCACCAGTCCGAGAAGAGCTCCAAGGGCAGAATCCACATG
ACCCCCCTGTTCAAGCAGATCCTGTCCGAGAAGGAGTCCTTCTCTTACATC

CCCGACGTGTTTACAGAGGATTCTCAGCTGCTGCCTAGCATCGGCGGCTTC
TTTGCCCAGATCGAGAATGACAAGGATGGCAACATCTTCGACCGGGCCCT
GGAGCTGATCTCTAGCTACGCCGAGTATGATACCGAGCGGATCTATATCA
GACAGGCCGACATCAATAGAGTGTCCAACGTGATCTTTGGAGAGTGGGGC
ACCCTGGGAGGCCTGATGAGGGAGTACAAGGCCGACTCTATCAATGATAT
CAACCTGGAGCGCACATGCAAGAAGGTGGACAAGTGGCTGGATTCTAAG
GAGTTTGCCCTGAGCGATGTGCTGGAGGCCATCAAGAGGACCGGCAACAA
TGACGCCTTCAACGAGTATATCTCCAAGATGCGGACAGCCAGAGAGAAGA
TCGATGCCGCCCGCAAGGAGATGAAGTTCATCAGCGAGAAGATCTCCGGC
GATGAGGAGTCTATCCACATCATCAAGACCCTGCTGGACAGCGTGCAGCA
GTTCCTGCACTTCTTTAATCTGTTTAAGGCAAGGCAGGACATCCCACTGGA
TGGAGCCTTCTACGCCGAGTTTGACGAGGTGCACAGCAAGCTGTTTGCCA
TCGTGCCCCTGTATAACAAGGTGCGGAACTATCTGACCAAGAACAATCTG
AACACAAAGAAGATCAAGCTGAATTTCAAGAACCCTACACTGGCCAATGG
CTGGGACCAGAACAAGGTGTACGATTATGCCTCCCTGATCTTTCTGCGGG
ACGGCAATTACTATCTGGGCATCATCAATCCTAAGAGAAAGAAGAACATC
AAGTTCGAGCAGGGCTCTGGCAACGGCCCCTTCTACCGGAAGATGGTGTA
TAAGCAGATCCCCGGCCCTAATAAGAACCTGCCAAGAGTGTTCCTGACCT
CCACAAAGGGCAAGAAGGAGTATAAGCCCTCTAAGGAGATCATCGAGGG
CTACGAGGCCGACAAGCACATCAGGGGCGATAAGTTCGACCTGGATTTTT
GTCACAAGCTGATCGATTTCTTTAAGGAGTCCATCGAGAAGCACAAGGAC
TGGTCTAAGTTCAACTTCTACTTCAGCCCAACCGAGAGCTATGGCGACATC
TCTGAGTTCTACCTGGATGTGGAGAAGCAGGGCTATCGCATGCACTTTGA
GAATATCAGCGCCGAGACAATCGACGAGTATGTGGAGAAGGGCGATCTGT
TTCTGTTCCAGATCTACAACAAGGATTTTGTGAAGGCCGCCACCGGCAAG
AAGGACATGCACACAATCTACTGGAATGCCGCCTTCAGCCCCGAGAACCT
GCAGGACGTGGTGGTGAAGCTGAACGGCGAGGCCGAGCTGTTTTATAGGG
ACAAGTCCGATATCAAGGAGATCGTGCACCGCGAGGGCGAGATCCTGGTG
AATAGGACCTACAACGGCCGCACACCAGTGCCCGACAAGATCCACAAGA
AGCTGACCGATTATCACAATGGCCGGACAAAGGACCTGGGCGAGGCCAA
GGAGTACCTGGATAAGGTGAGATACTTCAAGGCCCACTATGACATCACCA
AGGATCGGAGATACCTGAACGACAAGATCTATTTCCACGTGCCTCTGACC
CTGAACTTCAAGGCCAACGGCAAGAAGAATCTGAACAAGATGGTCATCGA
GAAGTTCCTGTCCGATGAGAAGGCCCACATCATCGGCATCGACAGGGGCG

AGCGCAATCTGCTGTACTATTCCATCATCGACAGGTCTGGCAAGATCATCG
ATCAGCAGAGCCTGAATGTGATCGACGGCTTTGATTATCGGGAGAAGCTG
AACCAGAGAGAGATCGAGATGAAGGATGCCCGCCAGTCTTGGAACGCCA
TCGGCAAGATCAAGGACCTGAAGGAGGGCTACCTGAGCAAGGCCGTGCA
CGAGATCACCAAGATGGCCATCCAGTATAATGCCATCGTGGTCATGGAGG
AGCTGAACTACGGCTTCAAGCGGGGCCGGTTCAAGGTGGAGAAGCAGATC
TATCAGAAGTTCGAGAATATGCTGATCGATAAGATGAACTACCTGGTGTT
TAAGGACGCACCTGATGAGTCCCCAGGAGGCGTGCTGAATGCCTACCAGC
TGACAAACCCACTGGAGTCTTTCGCCAAGCTGGGCAAGCAGACCGGCATC
CTGTTTTACGTGCCAGCCGCCTATACATCCAAGATCGACCCCACCACAGG
CTTCGTGAATCTGTTTAACACCTCCTCTAAGACAAACGCCCAGGAGCGGA
AGGAGTTCCTGCAGAAGTTTGAGAGCATCTCCTATTCTGCCAAGGATGGC
GGCATCTTTGCCTTCGCCTTTGACTACAGAAAGTTCGGCACCAGCAAGAC
AGATCACAAGAACGTGTGGACCGCCTATACAAACGGCGAGAGGATGCGC
TACATCAAGGAGAAGAAGCGGAATGAGCTGTTTGACCCTTCTAAGGAGAT
CAAGGAGGCCCTGACCAGCTCCGGCATCAAGTACGATGGCGGCCAGAAC
ATCCTGCCAGACATCCTGAGGAGCAACAATAACGGCCTGATCTACACAAT
GTATTCTAGCTTCATCGCCGCCATCCAGATGCGCGTGTACGACGGCAAGG
AGGATTATATCATCAGCCCCATCAAGAACTCCAAGGGCGAGTTCTTTAGG
ACCGACCCCAAGAGGCGCGAGCTGCCTATCGACGCCGATGCCAATGGCGC
CTACAACATCGCCCTGAGGGGAGAGCTGACAATGAGGGCAATCGCAGAG
AAGTTCGACCCTGATAGCGAGAAGATGGCCAAGCTGGAGCTGAAGCACA
AGGATTGGTTCGAGTTTATGCAGACCAGAGGCGACAAAAGGCCGGCGGCC
ACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACG
ATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCAT
ATGATGTCCCCGACTATGCCTAA (SEQ ID NO: 30)

[00158] *Eubacterium eligens* (EeCpf1; pY013), including NLS and HA tag:

[00159] MNGNRSIVYREFVGVIPVAKTLRNELRPVGHTQEHIIQNGLI
QEDELRQEKSTELKNIMDDYYREYIDKSLSGVTDLDFTLLFELMNLVQSSPSK
DNKKALEKEQSKMREQICTHLQSDSNYKNIFNAKLLKEILPDFIKNYNQYDV
KDKAGKLETLALFNGFSTYFTDFFEKRKNVFTKEAVSTSIAYRIVHENSLIFLA
NMTSYKKISEKALDEIEVIEKNNQDKMGDWELNQIFNPDFYNMVLIQSGIDFY
NEICGVVNAHMNLYCQQTKNNYNLFKMRKLHKQILAYTSTSFEVPKMFEDD
MSVYNAVNAFIDETEKGNIIGKLKDIVNKYDELDEKRIYISKDFYETLSCFMSG

NWNLITGCVENFYDENIHAKGKSKEEKVKKAVKEDKYKSINDVNDLVEKYI
DEKERNEFKNSNAKQYIREISNIITDTETAHLEYDDHISLIESEEKADEMKKRL
DMYMNMYHWAKAFIVDEVLDRDEMFYSDIDDIYNILENIVPLYNRVRNYVT
QKPYNSKKIKLNFQSPTLANGWSQSKEFDNNAIILIRDNKYYLAIFNAKNKPD
KKIIQGNSDKKNDNDYKKMVYNLLPGANKMLPKVFLSKKGIETFKPSDYIISG
YNAHKHIKTSENFDISFCRDLIDYFKNSIEKHAEWRKYEFKFSATDSYSDISEF
YREVEMQGYRIDWTYISEADINKLDEEGKIYLFQIYNKDFAENSTGKENLHT
MYFKNIFSEENLKDIIIKLNGQAELFYRRASVKNPVKHKKDSVLVNKTYKNQL
DNGDVVRIPIPDDIYNEIYKMYNGYIKESDLSEAAKEYLDKVEVRTAQKDIVK
DYRYTVDKYFIHTPITINYKVTARNNVNDMVVKYIAQNDDIHVIGIDRGERNL
IYISVIDSHGNIVKQKSYNILNNYDYKKKLVEKEKTREYARKNWKSIGNIKEL
KEGYISGVVHEIAMLIVEYNAIIAMEDLNYGFKRGRFKVERQVYQKFESMLIN
KLNYFASKEKSVDEPGGLLKGYQLTYVPDNIKNLGKQCGVIFYVPAAFTSKID
PSTGFISAFNFKSISTNASRKQFFMQFDEIRYCAEKDMFSFGFDYNNFDTYNIT
MGKTQWTVYTNGERLQSEFNNARRTGKTKSINLTETIKLLLEDNEINYADGH
DIRIDMEKMDEDKKSEFFAQLLSLYKLTVQMRNSYTEAEEQENGISYDKIISP
VINDEGEFFDSDNYKESDDKECKMPKDADANGAYCIALKGLYEVLKIKSEWT
EDGFDRNCLKLPHAEWLDFIQNKRYEKRPAATKKAGQAKKKKGSYPYDVPD
YAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 31)

[00160] SEQ ID NO: 31 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00161] SEQ ID NO: 31 may be encoded by the following nucleotide sequence:

[00162] ATGAACGGCAATAGGTCCATCGTGTACCGCGAGTTCGTG
GGCGTGATCCCCGTGGCCAAGACCCTGAGGAATGAGCTGCGCCCTGTGGG
CCACACACAGGAGCACATCATCCAGAACGGCCTGATCCAGGAGGACGAG
CTGCGGCAGGAGAAGAGCACCGAGCTGAAGAACATCATGGACGATTACT
ATAGAGAGTACATCGATAAGTCTCTGAGCGGCGTGACCGACCTGGACTTC
ACCCTGCTGTTCGAGCTGATGAACCTGGTGCAGAGCTCCCCCTCCAAGGA
CAATAAGAAGGCCCTGGAGAAGGAGCAGTCTAAGATGAGGGAGCAGATC
TGCACCCACCTGCAGTCCGACTCTAACTACAAGAATATCTTTAACGCCAA
GCTGCTGAAGGAGATCCTGCCTGATTTCATCAAGAACTACAATCAGTATG

ACGTGAAGGATAAGGCCGGCAAGCTGGAGACACTGGCCCTGTTTAATGGC
TTCAGCACATACTTTACCGACTTCTTTGAGAAGAGGAAGAACGTGTTCACC
AAGGAGGCCGTGAGCACATCCATCGCCTACCGCATCGTGCACGAGAACTC
CCTGATCTTCCTGGCCAATATGACCTCTTATAAGAAGATCAGCGAGAAGG
CCCTGGATGAGATCGAAGTGATCGAGAAGAACAATCAGGACAAGATGGG
CGATTGGGAGCTGAATCAGATCTTTAACCCTGACTTCTACAATATGGTGCT
GATCCAGTCCGGCATCGACTTCTACAACGAGATCTGCGGCGTGGTGAATG
CCCACATGAACCTGTACTGTCAGCAGACCAAGAACAATTATAACCTGTTC
AAGATGCGGAAGCTGCACAAGCAGATCCTGGCCTACACCAGCACCAGCTT
CGAGGTGCCCAAGATGTTCGAGGACGATATGAGCGTGTATAACGCCGTGA
ACGCCTTCATCGACGAGACAGAGAAGGGCAACATCATCGGCAAGCTGAA
GGATATCGTGAATAAGTACGACGAGCTGGATGAGAAGAGAATCTATATCA
GCAAGGACTTTTACGAGACACTGAGCTGCTTCATGTCCGGCAACTGGAAT
CTGATCACAGGCTGCGTGGAGAACTTCTACGATGAGAACATCCACGCCAA
GGGCAAGTCCAAGGAGGAGAAGGTGAAGAAGGCCGTGAAGGAGGACAA
GTACAAGTCTATCAATGACGTGAACGATCTGGTGGAGAAGTATATCGATG
AGAAGGAGAGGAATGAGTTCAAGAACAGCAATGCCAAGCAGTACATCCG
CGAGATCTCCAACATCATCACCGACACAGAGACAGCCCACCTGGAGTATG
ACGATCACATCTCTCTGATCGAGAGCGAGGAGAAGGCCGACGAGATGAA
GAAGCGGCTGGATATGTATATGAACATGTACCACTGGGCCAAGGCCTTTA
TCGTGGACGAGGTGCTGGACAGAGATGAGATGTTCTACAGCGATATCGAC
GATATCTATAATATCCTGGAGAACATCGTGCCACTGTATAATCGGGTGAG
AAACTACGTGACCCAGAAGCCCTACAACTCTAAGAAGATCAAGCTGAATT
TCCAGAGCCCTACACTGGCCAATGGCTGGTCCCAGTCTAAGGAGTTCGAC
AACAATGCCATCATCCTGATCAGAGATAACAAGTACTATCTGGCCATCTTC
AATGCCAAGAACAAGCCAGACAAGAAGATCATCCAGGGCAACTCCGATA
AGAAGAACGACAACGATTACAAGAAGATGGTGTATAACCTGCTGCCAGG
CGCCAACAAGATGCTGCCCAAGGTGTTTCTGTCTAAGAAGGGCATCGAGA
CATTCAAGCCCTCCGACTATATCATCTCTGGCTACAACGCCCACAAGCAC
ATCAAGACAAGCGAGAATTTTGATATCTCCTTCTGTCGGGACCTGATCGAT
TACTTCAAGAACAGCATCGAGAAGCACGCCGAGTGGAGAAAGTATGAGTT
CAAGTTTTCCGCCACCGACAGCTACTCCGATATCTCTGAGTTCTATCGGGA
GGTGGAGATGCAGGGCTACAGAATCGACTGGACATATATCAGCGAGGCC
GACATCAACAAGCTGGATGAGGAGGGCAAGATCTATCTGTTTCAGATCTA

CAATAAGGATTTCGCCGAGAACAGCACCGGCAAGGAGAATCTGCACACA
ATGTACTTTAAGAACATCTTCTCCGAGGAGAATCTGAAGGACATCATCAT
CAAGCTGAACGGCCAGGCCGAGCTGTTTTATCGGAGAGCCTCTGTGAAGA
ATCCCGTGAAGCACAAGAAGGATAGCGTGCTGGTGAACAAGACCTACAA
GAATCAGCTGGACAACGGCGACGTGGTGAGAATCCCCATCCCTGACGATA
TCTATAACGAGATCTACAAGATGTATAATGGCTACATCAAGGAGTCCGAC
CTGTCTGAGGCCGCCAAGGAGTACCTGGATAAGGTGGAGGTGAGGACCGC
CCAGAAGGACATCGTGAAGGATTACCGCTATACAGTGGACAAGTACTTCA
TCCACACACCTATCACCATCAACTATAAGGTGACCGCCCGCAACAATGTG
AATGATATGGTGGTGAAGTACATCGCCCAGAACGACGATATCCACGTGAT
CGGCATCGACCGGGGCGAGAGAAACCTGATCTACATCTCCGTGATCGATT
CTCACGGCAACATCGTGAAGCAGAAATCCTACAACATCCTGAACAACTAC
GACTACAAGAAGAAGCTGGTGGAGAAGGAGAAAACCCGGGAGTACGCCA
GAAAGAACTGGAAGAGCATCGGCAATATCAAGGAGCTGAAGGAGGGCTA
TATCTCCGGCGTGGTGCACGAGATCGCCATGCTGATCGTGGAGTACAACG
CCATCATCGCCATGGAGGACCTGAATTATGGCTTTAAGAGGGGCCGCTTC
AAGGTGGAGCGGCAGGTGTACCAGAAGTTTGAGAGCATGCTGATCAATAA
GCTGAACTATTTCGCCAGCAAGGAGAAGTCCGTGGACGAGCCAGGAGGCC
TGCTGAAGGGCTATCAGCTGACCTACGTGCCCGATAATATCAAGAACCTG
GGCAAGCAGTGCGGCGTGATCTTTTACGTGCCTGCCGCCTTCACCAGCAA
GATCGACCCATCCACAGGCTTTATCTCTGCCTTCAACTTTAAGTCTATCAG
CACAAATGCCTCTCGGAAGCAGTTCTTTATGCAGTTTGACGAGATCAGAT
ACTGTGCCGAGAAGGATATGTTCAGCTTTGGCTTCGACTACAACAACTTCG
ATACCTACAACATCACAATGGGCAAGACACAGTGGACCGTGTATACAAAC
GGCGAGAGACTGCAGTCTGAGTTCAACAATGCCAGGCGCACCGGCAAGA
CAAAGAGCATCAATCTGACAGAGACAATCAAGCTGCTGCTGGAGGACAAT
GAGATCAACTACGCCGACGGCCACGATATCAGGATCGATATGGAGAAGAT
GGACGAGGATAAGAAGAGCGAGTTCTTTGCCCAGCTGCTGAGCCTGTATA
AGCTGACCGTGCAGATGCGCAATTCCTATACAGAGGCCGAGGAGCAGGA
GAACGGCATCTCTTACGACAAGATCATCAGCCCTGTGATCAATGATGAGG
GCGAGTTCTTTGACTCCGATAACTATAAGGAGTCTGACGATAAGGAGTGC
AAGATGCCAAAGGACGCCGATGCCAACGGCGCCTACTGTATCGCCCTGAA
GGGCCTGTATGAGGTGCTGAAGATCAAGAGCGAGTGGACCGAGGACGGC
TTTGATAGGAATTGCCTGAAGCTGCCACACGCAGAGTGGCTGGACTTCAT

CCAGAACAAGCGGTACGAGAAAAGGCCGGCGGCCACGAAAAAGGCCGGC
CAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGC
TTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTA
TGCCTAA (SEQ ID NO: 32)

[00163] *Moraxella bovoculi* 237 (MbCpf1; pY014), including NLS and HA tag:

[00164] MLFQDFTHLYPLSKTVRFELKPIDRTLEHIHAKNFLSQDETM
ADMHQKVKVILDDYHRDFIADMMGEVKLTKLAEFYDVYLKFRKNPKDDEL
QKQLKDLQAVLRKEIVKPIGNGGKYKAGYDRLFGAKLFKDGKELGDLAKFVI
AQEGESSPKLAHLAHFEKFSTYFTGFHDNRKNMYSDEDKHTAIAYRLIHENLP
RFIDNLQILTTIKQKHSALYDQIINELTASGLDVSLASHLDGYHKLLTQEGITA
YNTLLGGISGEAGSPKIQGINELINSHHNQHCHKSERIAKLRPLHKQILSDGMS
VSFLPSKFADDSEMCQAVNEFYRHYADVFAKVQSLFDGFDDHQKDGIYVEH
KNLNELSKQAFGDFALLGRVLDGYYVDVVNPEFNERFAKAKTDNAKAKLTK
EKDKFIKGVHSLASLEQAIEHYTARHDDESVQAGKLGQYFKHGLAGVDNPIQ
KIHNNHSTIKGFLERERPAGERALPKIKSGKNPEMTQLRQLKELLDNALNVAH
FAKLLTTKTTLDNQDGNFYGEFGVLYDELAKIPTLYNKVRDYLSQKPFSTEK
YKLNFGNPTLLNGWDLNKEKDNFGVILQKDGCYYLALLDKAHKKVFDNAP
NTGKSIYQKMIYKYLEVRKQFPKVFFSKEAIAINYHPSKELVEIKDKGRQRSD
DERLKLYRFILECLKIHPKYDKKFEGAIGDIQLFKKDKKGREVPISEKDLFDKI
NGIFSSKPKLEMEDFFIGEFKRYNPSQDLVDQYNIYKKIDSNDNRKKENFYNN
HPKFKKDLVRYYYESMCKHEEWEESFEFSKKLQDIGCYVDVNELFTEIETRRL
NYKISFCNINADYIDELVEQGQLYLFQIYNKDFSPKAHGKPNLHTLYFKALFSE
DNLADPIYKLNGEAQIFYRKASLDMNETTIHRAGEVLENKNPDNPKKRQFVY
DIIKDKRYTQDKFMLHVPITMNFGVQGMTIKEFNKKVNQSIQQYDEVNVIGID
RGERHLLYLTVINSKGEILEQCSLNDITTASANGTQMTTPYHKILDKREIERLN
ARVGWGEIETIKELKSGYLSHVVHQISQLMLKYNAIVVLEDLNFGFKRGRFK
VEKQIYQNFENALIKKLNHLVLKDKADDEIGSYKNALQLTNNFTDLKSIGKQT
GFLFYVPAWNTSKIDPETGFVDLLKPRYENIAQSQAFFGKFDKICYNADKDYF
EFHIDYAKFTDKAKNSRQIWTICSHGDKRYVYDKTANQNKGAAKGINVNDE
LKSLFARHHINEKQPNLVMDICQNNDKEFHKSLMYLLKTLLALRYSNASSDE
DFILSPVANDEGVFFNSALADDTQPQNADANGAYHIALKGLWLLNELKNSDD
LNKVKLAIDNQTWLNFAQNRKRPAATKKAGQAKKKKGSYPYDVPDYAYPY
DVPDYAYPYDVPDYA (SEQ ID NO: 33)

[00165] SEQ ID NO: 33 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00166] SEQ ID NO: 33 may be encoded by the following nucleotide sequence:

[00167] ATGCTGTTCCAGGACTTTACCCACCTGTATCCACTGTCCA
AGACAGTGAGATTTGAGCTGAAGCCCATCGATAGGACCCTGGAGCACATC
CACGCCAAGAACTTCCTGTCTCAGGACGAGACAATGGCCGATATGCACCA
GAAGGTGAAAGTGATCCTGGACGATTACCACCGCGACTTCATCGCCGATA
TGATGGGCGAGGTGAAGCTGACCAAGCTGGCCGAGTTCTATGACGTGTAC
CTGAAGTTTCGGAAGAACCCAAAGGACGATGAGCTGCAGAAGCAGCTGA
AGGATCTGCAGGCCGTGCTGAGAAAGGAGATCGTGAAGCCCATCGGCAAT
GGCGGCAAGTATAAGGCCGGCTACGACAGGCTGTTCGGCGCCAAGCTGTT
TAAGGACGGCAAGGAGCTGGGCGATCTGGCCAAGTTCGTGATCGCACAGG
AGGGAGAGAGCTCCCCAAAGCTGGCCCACCTGGCCCACTTCGAGAAGTTT
TCCACCTATTTCACAGGCTTTCACGATAACCGGAAGAATATGTATTCTGAC
GAGGATAAGCACACCGCCATCGCCTACCGCCTGATCCACGAGAACCTGCC
CCGGTTTATCGACAATCTGCAGATCCTGACCACAATCAAGCAGAAGCACT
CTGCCCTGTACGATCAGATCATCAACGAGCTGACCGCCAGCGGCCTGGAC
GTGTCTCTGGCCAGCCACCTGGATGGCTATCACAAGCTGCTGACACAGGA
GGGCATCACCGCCTACAATACACTGCTGGGAGGAATCTCCGGAGAGGCAG
GCTCTCCTAAGATCCAGGGCATCAACGAGCTGATCAATTCTCACCACAAC
CAGCACTGCCACAAGAGCGAGAGAATCGCCAAGCTGAGGCCACTGCACA
AGCAGATCCTGTCCGACGGCATGAGCGTGTCCTTCCTGCCCTCTAAGTTTG
CCGACGATAGCGAGATGTGCCAGGCCGTGAACGAGTTCTATCGCCACTAC
GCCGACGTGTTCGCCAAGGTGCAGAGCCTGTTCGACGGCTTTGACGATCA
CCAGAAGGATGGCATCTACGTGGAGCACAAGAACCTGAATGAGCTGTCCA
AGCAGGCCTTCGGCGACTTTGCACTGCTGGGACGCGTGCTGGACGGATAC
TATGTGGATGTGGTGAATCCAGAGTTCAACGAGCGGTTTGCCAAGGCCAA
GACCGACAATGCCAAGGCCAAGCTGACAAAGGAGAAGGATAAGTTCATC
AAGGGCGTGCACTCCCTGGCCTCTCTGGAGCAGGCCATCGAGCACTATAC
CGCAAGGCACGACGATGAGAGCGTGCAGGCAGGCAAGCTGGGACAGTAC
TTCAAGCACGGCCTGGCCGGAGTGGACAACCCCATCCAGAAGATCCACAA

CAATCACAGCACCATCAAGGGCTTTCTGGAGAGGGAGCGCCCTGCAGGAG
AGAGAGCCCTGCCAAAGATCAAGTCCGGCAAGAATCCTGAGATGACACA
GCTGAGGCAGCTGAAGGAGCTGCTGGATAACGCCCTGAATGTGGCCCACT
TCGCCAAGCTGCTGACCACAAAGACCACACTGGACAATCAGGATGGCAAC
TTCTATGGCGAGTTTGGCGTGCTGTACGACGAGCTGGCCAAGATCCCCAC
CCTGTATAACAAGGTGAGAGATTACCTGAGCCAGAAGCCTTTCTCCACCG
AGAAGTACAAGCTGAACTTTGGCAATCCAACACTGCTGAATGGCTGGGAC
CTGAACAAGGAGAAGGATAATTTCGGCGTGATCCTGCAGAAGGACGGCTG
CTACTATCTGGCCCTGCTGGACAAGGCCCACAAGAAGGTGTTTGATAACG
CCCCTAATACAGGCAAGAGCATCTATCAGAAGATGATCTATAAGTACCTG
GAGGTGAGGAAGCAGTTCCCCAAGGTGTTCTTTTCCAAGGAGGCCATCGC
CATCAACTACCACCCTTCTAAGGAGCTGGTGGAGATCAAGGACAAGGGCC
GGCAGAGATCCGACGATGAGCGCCTGAAGCTGTATCGGTTTATCCTGGAG
TGTCTGAAGATCCACCCTAAGTACGATAAGAAGTTCGAGGGCGCCATCGG
CGACATCCAGCTGTTTAAGAAGGATAAGAAGGGCAGAGAGGTGCCAATC
AGCGAGAAGGACCTGTTCGATAAGATCAACGGCATCTTTTCTAGCAAGCC
TAAGCTGGAGATGGAGGACTTCTTTATCGGCGAGTTCAAGAGGTATAACC
CAAGCCAGGACCTGGTGGATCAGTATAATATCTACAAGAAGATCGACTCC
AACGATAATCGCAAGAAGGAGAATTTCTACAACAATCACCCCAAGTTTAA
GAAGGATCTGGTGCGGTACTATTACGAGTCTATGTGCAAGCACGAGGAGT
GGGAGGAGAGCTTCGAGTTTTCCAAGAAGCTGCAGGACATCGGCTGTTAC
GTGGATGTGAACGAGCTGTTTACCGAGATCGAGACACGGAGACTGAATTA
TAAGATCTCCTTCTGCAACATCAATGCCGACTACATCGATGAGCTGGTGG
AGCAGGGCCAGCTGTATCTGTTCCAGATCTACAACAAGGACTTTTCCCCA
AAGGCCCACGGCAAGCCCAATCTGCACACCCTGTACTTCAAGGCCCTGTT
TTCTGAGGACAACCTGGCCGATCCTATCTATAAGCTGAATGGCGAGGCCC
AGATCTTCTACAGAAAGGCCTCCCTGGACATGAACGAGACAACAATCCAC
AGGGCCGGCGAGGTGCTGGAGAACAAGAATCCCGATAATCCTAAGAAGA
GACAGTTCGTGTACGACATCATCAAGGATAAGAGGTACACACAGGACAA
GTTCATGCTGCACGTGCCAATCACCATGAACTTTGGCGTGCAGGGCATGA
CAATCAAGGAGTTCAATAAGAAGGTGAACCAGTCTATCCAGCAGTATGAC
GAGGTGAACGTGATCGGCATCGATCGGGGCGAGAGACACCTGCTGTACCT
GACCGTGATCAATAGCAAGGGCGAGATCCTGGAGCAGTGTTCCCTGAACG
ACATCACCACAGCCTCTGCCAATGGCACACAGATGACCACACCTTACCAC

AAGATCCTGGATAAGAGGGAGATCGAGCGCCTGAACGCCCGGGTGGGAT
GGGGCGAGATCGAGACAATCAAGGAGCTGAAGTCTGGCTATCTGAGCCAC
GTGGTGCACCAGATCAGCCAGCTGATGCTGAAGTACAACGCCATCGTGGT
GCTGGAGGACCTGAATTTCGGCTTTAAGAGGGGCCGCTTTAAGGTGGAGA
AGCAGATCTATCAGAACTTCGAGAATGCCCTGATCAAGAAGCTGAACCAC
CTGGTGCTGAAGGACAAGGCCGACGATGAGATCGGCTCTTACAAGAATGC
CCTGCAGCTGACCAACAATTTCACAGATCTGAAGAGCATCGGCAAGCAGA
CCGGCTTCCTGTTTTATGTGCCCGCCTGGAACACCTCTAAGATCGACCCTG
AGACAGGCTTTGTGGATCTGCTGAAGCCAAGATACGAGAACATCGCCCAG
AGCCAGGCCTTCTTTGGCAAGTTCGACAAGATCTGCTATAATGCCGACAA
GGATTACTTCGAGTTTCACATCGACTACGCCAAGTTTACCGATAAGGCCA
AGAATAGCCGCCAGATCTGGACAATCTGTTCCCACGGCGACAAGCGGTAC
GTGTACGATAAGACAGCCAACCAGAATAAGGGCGCCGCCAAGGGCATCA
ACGTGAATGATGAGCTGAAGTCCCTGTTCGCCCGCCACCACATCAACGAG
AAGCAGCCCAACCTGGTCATGGACATCTGCCAGAACAATGATAAGGAGTT
TCACAAGTCTCTGATGTACCTGCTGAAAACCCTGCTGGCCCTGCGGTACAG
CAACGCCTCCTCTGACGAGGATTTCATCCTGTCCCCCGTGGCAAACGACG
AGGGCGTGTTCTTTAATAGCGCCCTGGCCGACGATACACAGCCTCAGAAT
GCCGATGCCAACGGCGCCTACCACATCGCCCTGAAGGGCCTGTGGCTGCT
GAATGAGCTGAAGAACTCCGACGATCTGAACAAGGTGAAGCTGGCCATCG
ACAATCAGACCTGGCTGAATTTCGCCCAGAACAGGAAAAGGCCGGCGGC
CACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATAC
GATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCA
TATGATGTCCCCGACTATGCCTAA (SEQ ID NO: 34)

[00168] *Leptospira inadai* (LiCpf1; pY015), including NLS and HA tag:

[00169] MEDYSGFVNIYSIQKTLRFELKPVGKTLEHIEKKGFLKKDKI
RAEDYKAVKKIIDKYHRAYIEEVFDSVLHQKKKKDKTRFSTQFIKEIKEFSELY
YKTEKNIPDKERLEALSEKLRKMLVGAFKGEFSEEVAEKYKNLFSKELIRNEI
EKFCETDEERKQVSNFKSFTTYFTGFHSNRQNIYSDEKKSTAIGYRIIHQNLPK
FLDNLKIIESIQRRFKDFPWSDLKKNLKKIDKNIKLTEYFSIDGFVNVLNQKGID
AYNTILGGKSEESGEKIQGLNEYINLYRQKNNIDRKNLPNVKILFKQILGDRET
KSFIPEAFPDDQSVLNSITEFAKYLKLDKKKKSIIAELKKFLSSFNRYELDGIYL
ANDNSLASISTFLFDDWSFIKKSVSFKYDESVGDPKKKIKSPLKYEKEKEKWL
KQKYYTISFLNDAIESYSKSQDEKRVKIRLEAYFAEFKSKDDAKKQFDLLERIE

EAYAIVEPLLGAEYPRDRNLKADKKEVGKIKDFLDSIKSLQFFLKPLLSAEIFD
EKDLGFYNQLEGYYEEIDSIGHLYNKVRNYLTGKIYSKEKFKLNFENSTLLKG
WDENREVANLCVIFREDQKYYLGVMDKENNTILSDIPKVKPNELFYEKMVY
KLIPTPHMQLPRIIFSSDNLSIYNPSKSILKIREAKSFKEGKNFKLKDCHKFIDFY
KESISKNEDWSRFDFKFSKTSSYENISEFYREVERQGYNLDFKKVSKFYIDSLV
EDGKLYLFQIYNKDFSIFSKGKPNLHTIYFRSLFSKENLKDVCLKLNGEAEMFF
RKKSINYDEKKKREGHHPELFEKLKYPILKDKRYSEDKFQFHLPISLNFKSKER
LNFNLKVNEFLKRNKDINIIGIDRGERNLLYLVMINQKGEILKQTLLDSMQSG
KGRPEINYKEKLQEKEIERDKARKSWGTVENIKELKEGYLSIVIHQISKLMVE
NNAIVVLEDLNIGFKRGRQKVERQVYQKFEKMLIDKLNFLVFKENKPTEPGG
VLKAYQLTDEFQSFEKLSKQTGFLFYVPSWNTSKIDPRTGFIDFLHPAYENIEK
AKQWINKFDSIRFNSKMDWFEFTADTRKFSENLMLGKNRVWVICTTNVERYF
TSKTANSSIQYNSIQITEKLKELFVDIPFSNGQDLKPEILRKNDAVFFKSLLFYIK
TTLSLRQNNGKKGEEEKDFILSPVVDSKGRFFNSLEASDDEPKDADANGAYHI
ALKGLMNLLVLNETKEENLSRPKWKIKNKDWLEFVWERNRKRPAATKKAG
QAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 35)

[00170] SEQ ID NO: 35 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00171] SEQ ID NO: 35 may be encoded by the following nucleotide sequence:

[00172] ATGGAGGACTATTCCGGCTTTGTGAACATCTACTCTATCC
AGAAAACCCTGAGGTTCGAGCTGAAGCCAGTGGGCAAGACACTGGAGCA
CATCGAGAAGAAGGGCTTCCTGAAGAAGGACAAGATCCGGGCCGAGGAT
TACAAGGCCGTGAAGAAGATCATCGATAAGTACCACAGAGCCTATATCGA
GGAGGTGTTTGATTCCGTGCTGCACCAGAAGAAGAAGAAGGACAAGACC
CGCTTTTCTACACAGTTCATCAAGGAGATCAAGGAGTTCAGCGAGCTGTA
CTATAAGACCGAGAAGAACATCCCCGACAAGGAGAGGCTGGAGGCCCTG
AGCGAGAAGCTGCGCAAGATGCTGGTGGGCGCCTTTAAGGGCGAGTTCTC
CGAGGAGGTGGCCGAGAAGTATAAGAACCTGTTTTCTAAGGAGCTGATCA
GGAATGAGATCGAGAAGTTCTGCGAGACAGACGAGGAGCGCAAGCAGGT
GTCTAACTTCAAGAGCTTCACCACATACTTTACCGGCTTCCACTCCAACAG
GCAGAATATCTATTCCGACGAGAAGAAGTCTACAGCCATCGGCTACCGCA

TCATCCACCAGAACCTGCCTAAGTTCCTGGATAATCTGAAGATCATCGAGT
CCATCCAGCGGCGGTTCAAGGACTTCCCATGGTCTGATCTGAAGAAGAAC
CTGAAGAAGATCGATAAGAATATCAAGCTGACCGAGTACTTCAGCATCGA
CGGCTTCGTGAACGTGCTGAATCAGAAGGGCATCGATGCCTACAACACAA
TCCTGGGCGGCAAGTCCGAGGAGTCTGGCGAGAAGATCCAGGGCCTGAAC
GAGTACATCAATCTGTATCGGCAGAAGAACAATATCGACAGAAAGAACCT
GCCCAATGTGAAGATCCTGTTTAAGCAGATCCTGGGCGATAGGGAGACAA
AGAGCTTTATCCCTGAGGCCTTCCCAGACGATCAGTCCGTGCTGAACTCTA
TCACAGAGTTCGCCAAGTACCTGAAGCTGGATAAGAAGAAGAAGAGCAT
CATCGCCGAGCTGAAGAAGTTTCTGAGCTCCTTCAATCGCTACGAGCTGG
ACGGCATCTATCTGGCCAACGATAATAGCCTGGCCTCTATCAGCACCTTCC
TGTTTGACGATTGGTCCTTTATCAAGAAGTCCGTGTCTTTCAAGTATGACG
AGTCCGTGGGCGACCCCAAGAAGAAGATCAAGTCTCCCCTGAAGTACGAG
AAGGAGAAGGAGAAGTGGCTGAAGCAGAAGTACTATACAATCTCTTTCCT
GAACGATGCCATCGAGAGCTATTCCAAGTCTCAGGACGAGAAGAGGGTG
AAGATCCGCCTGGAGGCCTACTTTGCCGAGTTCAAGAGCAAGGACGATGC
CAAGAAGCAGTTCGACCTGCTGGAGAGGATCGAGGAGGCCTATGCCATCG
TGGAGCCTCTGCTGGGAGCAGAGTACCCAAGGGACCGCAACCTGAAGGC
CGATAAGAAGGAAGTGGGCAAGATCAAGGACTTCCTGGATAGCATCAAG
TCCCTGCAGTTCTTTCTGAAGCCTCTGCTGTCCGCCGAGATCTTTGACGAG
AAGGATCTGGGCTTCTACAATCAGCTGGAGGGCTACTATGAGGAGATCGA
TTCTATCGGCCACCTGTATAACAAGGTGCGGAATTATCTGACCGGCAAGA
TCTACAGCAAGGAGAAGTTTAAGCTGAACTTCGAGAACAGCACCCTGCTG
AAGGGCTGGGACGAGAACCGGGAGGTGGCCAATCTGTGCGTGATCTTCAG
AGAGGACCAGAAGTACTATCTGGGCGTGATGGATAAGGAGAACAATACC
ATCCTGTCCGACATCCCCAAGGTGAAGCCTAACGAGCTGTTTTACGAGAA
GATGGTGTATAAGCTGATCCCCACACCTCACATGCAGCTGCCCCGGATCA
TCTTCTCTAGCGACAACCTGTCTATCTATAATCCTAGCAAGTCCATCCTGA
AGATCAGAGAGGCCAAGAGCTTTAAGGAGGGCAAGAACTTCAAGCTGAA
GGACTGTCACAAGTTTATCGATTTCTACAAGGAGTCTATCAGCAAGAATG
AGGACTGGAGCAGATTCGACTTCAAGTTCAGCAAGACCAGCAGCTACGAG
AACATCAGCGAGTTTTACCGGGAGGTGGAGAGACAGGGCTATAACCTGGA
CTTCAAGAAGGTGTCTAAGTTCTACATCGACAGCCTGGTGGAGGATGGCA
AGCTGTACCTGTTCCAGATCTATAACAAGGACTTTTCTATCTTCAGCAAGG

GCAAGCCCAATCTGCACACCATCTATTTTCGGTCCCTGTTCTCTAAGGAGA
ACCTGAAGGACGTGTGCCTGAAGCTGAATGGCGAGGCCGAGATGTTCTTT
CGGAAGAAGTCCATCAACTACGATGAGAAGAAGAAGCGGGAGGGCCACC
ACCCCGAGCTGTTTGAGAAGCTGAAGTATCCTATCCTGAAGGACAAGAGA
TACAGCGAGGATAAGTTTCAGTTCCACCTGCCCATCAGCCTGAACTTCAA
GTCCAAGGAGCGGCTGAACTTTAATCTGAAAGTGAATGAGTTCCTGAAGA
GAAACAAGGACATCAATATCATCGGCATCGATCGGGGCGAGAGAAACCT
GCTGTACCTGGTCATGATCAATCAGAAGGGCGAGATCCTGAAGCAGACCC
TGCTGGACAGCATGCAGTCCGGCAAGGGCCGGCCTGAGATCAACTACAAG
GAGAAGCTGCAGGAGAAGGAGATCGAGAGGGATAAGGCCCGCAAGAGCT
GGGGCACAGTGGAGAATATCAAGGAGCTGAAGGAGGGCTATCTGTCTATC
GTGATCCACCAGATCAGCAAGCTGATGGTGGAGAACAATGCCATCGTGGT
GCTGGAGGACCTGAACATCGGCTTTAAGCGGGGCAGACAGAAGGTGGAG
CGGCAGGTGTACCAGAAGTTCGAGAAGATGCTGATCGATAAGCTGAACTT
TCTGGTGTTCAAGGAGAATAAGCCAACCGAGCCAGGAGGCGTGCTGAAG
GCCTATCAGCTGACAGACGAGTTTCAGTCTTTCGAGAAGCTGAGCAAGCA
GACCGGCTTTCTGTTCTACGTGCCAAGCTGGAACACCTCCAAGATCGACC
CCAGAACAGGCTTTATCGATTTCCTGCACCCTGCCTACGAGAATATCGAG
AAGGCCAAGCAGTGGATCAACAAGTTTGATTCCATCAGGTTCAATTCTAA
GATGGACTGGTTTGAGTTCACCGCCGATACACGCAAGTTTTCCGAGAACC
TGATGCTGGGCAAGAATCGGGTGTGGGTCATCTGCACCACAAATGTGGAG
CGGTACTTCACCAGCAAGACCGCCAACAGCTCCATCCAGTACAATAGCAT
CCAGATCACCGAGAAGCTGAAGGAGCTGTTTGTGGACATCCCTTTCAGCA
ACGGCCAGGATCTGAAGCCAGAGATCCTGAGGAAGAATGACGCCGTGTTC
TTTAAGAGCCTGCTGTTTTACATCAAGACCACACTGTCCCTGCGCCAGAAC
AATGGCAAGAAGGGCGAGGAGGAGAAGGACTTCATCCTGAGCCCAGTGG
TGGATTCCAAGGGCCGGTTCTTTAACTCTCTGGAGGCCAGCGACGATGAG
CCCAAGGACGCCGATGCCAATGGCGCCTACCACATCGCCCTGAAGGGCCT
GATGAACCTGCTGGTGCTGAATGAGACAAAGGAGGAGAACCTGAGCAGA
CCAAAGTGGAAGATCAAGAATAAGGACTGGCTGGAGTTCGTGTGGGAGA
GGAACCGCAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAA
AGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTAC
GACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA
(SEQ ID NO: 36)

[00173] *Lachnospiraceae bacterium* ND2006 (LbCpf1; pY016), including NLS and HA tag:

[00174] MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDE KRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKEL ENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALVNSFNGFTT AFTGFFDNRENMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQ EIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEY INLYNQKTKQKLPKFKPLYKQVLSDRESLSFYGEGYTSDEEVLEVFRNTLNKN SEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIRDKWNAEY DDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEIIIQ KVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKA FFGEGKETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYF QNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVN GNYEKINYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDM FNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEEQGYK VSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDEN NHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTTLSYDV YKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIDRGERN LLYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNW TSIENIKELKAGYISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVY QKFEKMLIDKLNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQNGFIFY IPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYK NFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKY GINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVK NSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKKAEDEKLD KVKIAISNKEWLEYAQTSVKHKRPAATKKAGQAKKKKGSYPYDVPDYAYPY DVPDYAYPYDVPDYA (SEQ ID NO: 37)

[00175] SEQ ID NO: 37 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00176] SEQ ID NO: 37 may be encoded by the following nucleotide sequence:

[00177] ATGAGCAAGCTGGAGAAGTTTACAAACTGCTACTCCCTG
TCTAAGACCCTGAGGTTCAAGGCCATCCCTGTGGGCAAGACCCAGGAGAA
CATCGACAATAAGCGGCTGCTGGTGGAGGACGAGAAGAGAGCCGAGGAT
TATAAGGGCGTGAAGAAGCTGCTGGATCGCTACTATCTGTCTTTTATCAAC
GACGTGCTGCACAGCATCAAGCTGAAGAATCTGAACAATTACATCAGCCT
GTTCCGGAAGAAAACCAGAACCGAGAAGGAGAATAAGGAGCTGGAGAAC
CTGGAGATCAATCTGCGGAAGGAGATCGCCAAGGCCTTCAAGGGCAACG
AGGGCTACAAGTCCCTGTTTAAGAAGGATATCATCGAGACAATCCTGCCA
GAGTTCCTGGACGATAAGGACGAGATCGCCCTGGTGAACAGCTTCAATGG
CTTTACCACAGCCTTCACCGGCTTCTTTGATAACAGAGAGAATATGTTTTC
CGAGGAGGCCAAGAGCACATCCATCGCCTTCAGGTGTATCAACGAGAATC
TGACCCGCTACATCTCTAATATGGACATCTTCGAGAAGGTGGACGCCATCT
TTGATAAGCACGAGGTGCAGGAGATCAAGGAGAAGATCCTGAACAGCGA
CTATGATGTGGAGGATTTCTTTGAGGGCGAGTTCTTTAACTTTGTGCTGAC
ACAGGAGGGCATCGACGTGTATAACGCCATCATCGGCGGCTTCGTGACCG
AGAGCGGCGAGAAGATCAAGGGCCTGAACGAGTACATCAACCTGTATAA
TCAGAAAACCAAGCAGAAGCTGCCTAAGTTTAAGCCACTGTATAAGCAGG
TGCTGAGCGATCGGGAGTCTCTGAGCTTCTACGGCGAGGGCTATACATCC
GATGAGGAGGTGCTGGAGGTGTTTAGAAACACCCTGAACAAGAACAGCG
AGATCTTCAGCTCCATCAAGAAGCTGGAGAAGCTGTTCAAGAATTTTGAC
GAGTACTCTAGCGCCGGCATCTTTGTGAAGAACGGCCCCGCCATCAGCAC
AATCTCCAAGGATATCTTCGGCGAGTGGAACGTGATCCGGGACAAGTGGA
ATGCCGAGTATGACGATATCCACCTGAAGAAGAAGGCCGTGGTGACCGAG
AAGTACGAGGACGATCGGAGAAAGTCCTTCAAGAAGATCGGCTCCTTTTC
TCTGGAGCAGCTGCAGGAGTACGCCGACGCCGATCTGTCTGTGGTGGAGA
AGCTGAAGGAGATCATCATCCAGAAGGTGGATGAGATCTACAAGGTGTAT
GGCTCCTCTGAGAAGCTGTTCGACGCCGATTTTGTGCTGGAGAAGAGCCT
GAAGAAGAACGACGCCGTGGTGGCCATCATGAAGGACCTGCTGGATTCTG
TGAAGAGCTTCGAGAATTACATCAAGGCCTTCTTTGGCGAGGGCAAGGAG
ACAAACAGGGACGAGTCCTTCTATGGCGATTTTGTGCTGGCCTACGACAT
CCTGCTGAAGGTGGACCACATCTACGATGCCATCCGCAATTATGTGACCC
AGAAGCCCTACTCTAAGGATAAGTTCAAGCTGTATTTTCAGAACCCTCAGT
TCATGGGCGGCTGGGACAAGGATAAGGAGACAGACTATCGGGCCACCAT
CCTGAGATACGGCTCCAAGTACTATCTGGCCATCATGGATAAGAAGTACG

CCAAGTGCCTGCAGAAGATCGACAAGGACGATGTGAACGGCAATTACGA
GAAGATCAACTATAAGCTGCTGCCCGGCCCTAATAAGATGCTGCCAAAGG
TGTTCTTTTCTAAGAAGTGGATGGCCTACTATAACCCCAGCGAGGACATCC
AGAAGATCTACAAGAATGGCACATTCAAGAAGGGCGATATGTTTAACCTG
AATGACTGTCACAAGCTGATCGACTTCTTTAAGGATAGCATCTCCCGGTAT
CCAAAGTGGTCCAATGCCTACGATTTCAACTTTTCTGAGACAGAGAAGTA
TAAGGACATCGCCGGCTTTTACAGAGAGGTGGAGGAGCAGGGCTATAAG
GTGAGCTTCGAGTCTGCCAGCAAGAAGGAGGTGGATAAGCTGGTGGAGG
AGGGCAAGCTGTATATGTTCCAGATCTATAACAAGGACTTTTCCGATAAG
TCTCACGGCACACCCAATCTGCACACCATGTACTTCAAGCTGCTGTTTGAC
GAGAACAATCACGGACAGATCAGGCTGAGCGGAGGAGCAGAGCTGTTCA
TGAGGCGCGCCTCCCTGAAGAAGGAGGAGCTGGTGGTGCACCCAGCCAA
CTCCCCTATCGCCAACAAGAATCCAGATAATCCCAAGAAAACCACAACCC
TGTCCTACGACGTGTATAAGGATAAGAGGTTTTCTGAGGACCAGTACGAG
CTGCACATCCCAATCGCCATCAATAAGTGCCCCAAGAACATCTTCAAGAT
CAATACAGAGGTGCGCGTGCTGCTGAAGCACGACGATAACCCCTATGTGA
TCGGCATCGATAGGGGCGAGCGCAATCTGCTGTATATCGTGGTGGTGGAC
GGCAAGGGCAACATCGTGGAGCAGTATTCCCTGAACGAGATCATCAACAA
CTTCAACGGCATCAGGATCAAGACAGATTACCACTCTCTGCTGGACAAGA
AGGAGAAGGAGAGGTTCGAGGCCCGCCAGAACTGGACCTCCATCGAGAA
TATCAAGGAGCTGAAGGCCGGCTATATCTCTCAGGTGGTGCACAAGATCT
GCGAGCTGGTGGAGAAGTACGATGCCGTGATCGCCCTGGAGGACCTGAAC
TCTGGCTTTAAGAATAGCCGCGTGAAGGTGGAGAAGCAGGTGTATCAGAA
GTTCGAGAAGATGCTGATCGATAAGCTGAACTACATGGTGGACAAGAAGT
CTAATCCTTGTGCAACAGGCGGCGCCCTGAAGGGCTATCAGATCACCAAT
AAGTTCGAGAGCTTTAAGTCCATGTCTACCCAGAACGGCTTCATCTTTTAC
ATCCCTGCCTGGCTGACATCCAAGATCGATCCATCTACCGGCTTTGTGAAC
CTGCTGAAAACCAAGTATACCAGCATCGCCGATTCCAAGAAGTTCATCAG
CTCCTTTGACAGGATCATGTACGTGCCCGAGGAGGATCTGTTCGAGTTTGC
CCTGGACTATAAGAACTTCTCTCGCACAGACGCCGATTACATCAAGAAGT
GGAAGCTGTACTCCTACGGCAACCGGATCAGAATCTTCCGGAATCCTAAG
AAGAACAACGTGTTCGACTGGGAGGAGGTGTGCCTGACCAGCGCCTATAA
GGAGCTGTTCAACAAGTACGGCATCAATTATCAGCAGGGCGATATCAGAG
CCCTGCTGTGCGAGCAGTCCGACAAGGCCTTCTACTCTAGCTTTATGGCCC

TGATGAGCCTGATGCTGCAGATGCGGAACAGCATCACAGGCCGCACCGAC
GTGGATTTTCTGATCAGCCCTGTGAAGAACTCCGACGGCATCTTCTACGAT
AGCCGGAACTATGAGGCCCAGGAGAATGCCATCCTGCCAAAGAACGCCG
ACGCCAATGGCGCCTATAACATCGCCAGAAAGGTGCTGTGGGCCATCGGC
CAGTTCAAGAAGGCCGAGGACGAGAAGCTGGATAAGGTGAAGATCGCCA
TCTCTAACAAGGAGTGGCTGGAGTACGCCCAGACCAGCGTGAAGCACAAA
AGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGA
TCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGAT
TATGCATACCCATATGATGTCCCCGACTATGCCTAA (SEQ ID NO: 38)

[00178] *Porphyromonas crevioricanis* (PcCpf1; pY017), including NLS and HA tag:

[00179] MDSLKDFTNLYPVSKTLRFELKPVGKTLENIEKAGILKEDEH
RAESYRRVKKIIDTYHKVFIDSSLENMAKMGIENEIKAMLQSFCELYKKDHRT
EGEDKALDKIRAVLRGLIVGAFTGVCGRRENTVQNEKYESLFKEKLIKEILPD
FVLSTEAESLPFSVEEATRSLKEFDSFTSYFAGFYENRKNIYSTKPQSTAIAYRL
IHENLPKFIDNILVFQKIKEPIAKELEHIRADFSAGGYIKKDERLEDIFSLNYYIH
VLSQAGIEKYNALIGKIVTEGDGEMKGLNEHINLYNQQRGREDRLPLFRPLYK
QILSDREQLSYLPESFEKDEELLRALKEFYDHIAEDILGRTQQLMTSISEYDLSR
IYVRNDSQLTDISKKMLGDWNAIYMARERAYDHEQAPKRITAKYERDRIKAL
KGEESISLANLNSCIAFLDNVRDCRVDTYLSTLGQKEGPHGLSNLVENVFASY
HEAEQLLSFPYPEENNLIQDKDNVVLIKNLLDNISDLQRFLKPLWGMGDEPDK
DERFYGEYNYIRGALDQVIPLYNKVRNYLTRKPYSTRKVKLNFGNSQLLSGW
DRNKEKDNSCVILRKGQNFYLAIMNNRHKRSFENKMLPEYKEGEPYFEKMD
YKFLPDPNKMLPKVFLSKKGIEIYKPSPKLLEQYGHGTHKKGDTFSMDDLHE
LIDFFKHSIEAHEDWKQFGFKFSDTATYENVSSFYREVEDQGYKLSFRKVSES
YVYSLIDQGKLYLFQIYNKDFSPCSKGTPNLHTLYWRMLFDERNLADVIYKL
DGKAEIFFREKSLKNDHPTHPAGKPIKKKSRQKKGEESLFEYDLVKDRRYTM
DKFQFHVPITMNFKCSAGSKVNDMVNAHIREAKDMHVIGIDRGERNLLYICVI
DSRGTILDQISLNTINDIDYHDLLESRDKDRQQEHRNWQTIEGIKELKQGYLSQ
AVHRIAELMVAYKAVVALEDLNMGFKRGRQKVESSVYQQFEKQLIDKLNYL
VDKKKRPEDIGGLLRAYQFTAPFKSFKEMGKQNGFLFYIPAWNTSNIDPTTGF
VNLFHVQYENVDKAKSFFQKFDSISYNPKKDWFEFAFDYKNFTKKAEGSRS
MWILCTHGSRIKNFRNSQKNGQWDSEEFALTEAFKSLFVRYEIDYTADLKTAI
VDEKQKDFFVDLLKLFKLTVQMRNSWKEKDLDYLISPVAGADGRFFDTREG

NKSLPKDADANGAYNIALKGLWALRQIRQTSEGGKLKLAISNKEWLQFVQER
SYEKDKRPAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
(SEQ ID NO: 39)

[00180] SEQ ID NO: 39 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00181] SEQ ID NO: 39 may be encoded by the following nucleotide sequence:

[00182] ATGGACAGCCTGAAGGATTTCACCAACCTGTACCCCGTG
TCCAAGACACTGCGGTTTGAGCTGAAGCCTGTGGGCAAGACCCTGGAGAA
TATCGAGAAGGCCGGCATCCTGAAGGAGGATGAGCACAGAGCCGAGAGC
TACCGGAGAGTGAAGAAGATCATCGATACATATCACAAGGTGTTCATCGA
CAGCTCCCTGGAGAACATGGCCAAGATGGGCATCGAGAATGAGATCAAG
GCCATGCTGCAGTCCTTTTGCGAGCTGTATAAGAAGGACCACAGGACCGA
GGGAGAGGACAAGGCCCTGGATAAGATCAGGGCCGTGCTGAGGGGCCTG
ATCGTGGGAGCCTTCACCGGCGTGTGCGGCCGGCGGGAGAACACAGTGCA
GAATGAGAAGTATGAGAGCCTGTTTAAGGAGAAGCTGATCAAGGAGATC
CTGCCAGATTTCGTGCTGTCTACAGAGGCCGAGTCCCTGCCCTTTTCTGTG
GAGGAGGCCACCAGAAGCCTGAAGGAGTTCGACTCCTTTACATCTTACTT
CGCCGGCTTTTATGAGAACCGGAAGAATATCTACTCTACCAAGCCCCAGA
GCACAGCCATCGCCTATAGACTGATCCACGAGAACCTGCCTAAGTTCATC
GATAATATCCTGGTGTTTCAGAAGATCAAGGAGCCAATCGCCAAGGAGCT
GGAGCACATCAGGGCAGACTTCAGCGCCGGCGGCTACATCAAGAAGGAT
GAGCGCCTGGAGGACATCTTTTCCCTGAACTACTATATCCACGTGCTGTCT
CAGGCCGGCATCGAGAAGTACAATGCCCTGATCGGCAAGATCGTGACCGA
GGGCGATGGCGAGATGAAGGGCCTGAACGAGCACATCAACCTGTATAATC
AGCAGAGGGGCCGCGAGGACCGGCTGCCACTGTTCAGACCCCTGTATAAG
CAGATCCTGTCTGATAGGGAGCAGCTGTCCTATCTGCCAGAGTCTTTCGAG
AAGGACGAGGAGCTGCTGAGGGCCCTGAAGGAGTTTTACGATCACATCGC
AGAGGACATCCTGGGAAGGACCCAGCAGCTGATGACAAGCATCTCCGAGT
ACGATCTGTCCCGGATCTATGTGAGAAACGATAGCCAGCTGACCGACATC
TCCAAGAAGATGCTGGGCGATTGGAATGCCATCTACATGGCCCGGGAGAG
AGCCTATGACCACGAGCAGGCCCCCAAGCGCATCACAGCCAAGTACGAG

AGGGACCGCATCAAGGCCCTGAAGGGCGAGGAGTCTATCAGCCTGGCCA
ACCTGAACAGCTGCATCGCCTTCCTGGACAACGTGAGGGATTGTCGCGTG
GACACCTATCTGTCTACACTGGGACAGAAGGAGGGACCTCACGGCCTGAG
CAACCTGGTGGAGAACGTGTTCGCCTCCTACCACGAGGCCGAGCAGCTGC
TGTCTTTTCCCTATCCTGAGGAGAACAATCTGATCCAGGACAAGGATAAC
GTGGTGCTGATCAAGAACCTGCTGGATAATATCAGCGACCTGCAGAGGTT
CCTGAAGCCACTGTGGGGCATGGGCGATGAGCCCGACAAGGATGAGAGG
TTTTACGGCGAGTACAATTATATCAGGGGCGCCCTGGACCAGGTCATCCCT
CTGTATAACAAGGTGCGGAATTATCTGACCCGCAAGCCATACTCCACACG
CAAGGTGAAGCTGAACTTCGGCAATAGCCAGCTGCTGTCCGGCTGGGATA
GGAACAAGGAGAAGGACAATTCTTGCGTGATCCTGCGCAAGGGCCAGAA
CTTCTACCTGGCCATCATGAACAATCGGCACAAGCGGAGCTTCGAGAATA
AGATGCTGCCCGAGTATAAGGAGGGCGAGCCTTACTTCGAGAAGATGGAT
TATAAGTTTCTGCCAGACCCCAACAAGATGCTGCCCAAGGTGTTCCTGTCT
AAGAAGGGCATCGAGATCTACAAGCCTAGCCCAAAGCTGCTGGAGCAGT
ATGGCCACGGCACCCACAAGAAGGGCGATACCTTCAGCATGGACGATCTG
CACGAGCTGATCGACTTCTTTAAGCACTCCATCGAGGCCCACGAGGATTG
GAAGCAGTTCGGCTTTAAGTTCAGCGACACCGCCACATACGAGAACGTGA
GCAGCTTCTACCGGGAGGTGGAGGACCAGGGCTACAAGCTGTCTTTTAGA
AAGGTGTCCGAGTCTTACGTGTATAGCCTGATCGATCAGGGCAAGCTGTA
CCTGTTCCAGATCTATAACAAGGACTTTAGCCCTTGTTCCAAGGGCACCCC
AAATCTGCACACACTGTACTGGCGGATGCTGTTCGATGAGAGAAACCTGG
CCGACGTGATCTATAAGCTGGATGGCAAGGCCGAGATCTTCTTTCGGGAG
AAGTCCCTGAAGAATGACCACCCAACCCACCCTGCAGGCAAGCCCATCAA
GAAGAAGAGCCGGCAGAAGAAGGGCGAGGAGAGCCTGTTCGAGTACGAT
CTGGTGAAGGACCGGAGATATACCATGGATAAGTTTCAGTTCCACGTGCC
AATCACAATGAACTTTAAGTGCTCTGCCGGCAGCAAGGTGAACGACATGG
TGAATGCCCACATCAGGGAGGCCAAGGACATGCACGTGATCGGCATCGAT
AGGGGCGAGCGCAATCTGCTGTATATCTGCGTGATCGACAGCCGCGGCAC
CATCCTGGATCAGATCTCCCTGAACACAATCAATGACATCGATTATCACG
ATCTGCTGGAGTCCAGGGACAAGGATCGCCAGCAGGAGCACAGGAACTG
GCAGACCATCGAGGGCATCAAGGAGCTGAAGCAGGGCTACCTGTCTCAGG
CCGTGCACCGCATCGCCGAGCTGATGGTGGCCTATAAGGCCGTGGTGGCC
CTGGAGGACCTGAACATGGGCTTCAAGCGGGGCAGACAGAAGGTGGAGA

GCAGCGTGTACCAGCAGTTTGAGAAGCAGCTGATCGACAAGCTGAATTAT
CTGGTGGATAAGAAGAAGCGGCCCGAGGACATCGGAGGCCTGCTGAGAG
CCTACCAGTTCACCGCCCCTTTCAAGAGCTTTAAGGAGATGGGCAAGCAG
AACGGCTTTCTGTTCTATATCCCTGCCTGGAACACATCCAATATCGACCCA
ACCACAGGCTTCGTGAACCTGTTTCACGTGCAGTACGAGAATGTGGATAA
GGCCAAGAGCTTCTTTCAGAAGTTCGACAGCATCTCCTACAACCCTAAGA
AGGATTGGTTTGAGTTCGCCTTTGACTATAAGAACTTCACCAAGAAGGCC
GAGGGCTCTAGGAGCATGTGGATTCTGTGCACCCACGGCTCCCGGATCAA
GAACTTCAGAAATTCTCAGAAGAATGGCCAGTGGGATAGCGAGGAGTTTG
CCCTGACCGAGGCCTTCAAGTCCCTGTTTGTGCGGTACGAGATCGATTATA
CCGCCGACCTGAAAACCGCCATCGTGGACGAGAAGCAGAAGGATTTCTTT
GTGGACCTGCTGAAGCTGTTCAAGCTGACCGTGCAGATGAGAAACTCCTG
GAAGGAGAAGGACCTGGATTACCTGATCTCTCCAGTGGCCGGCGCCGATG
GCAGGTTCTTTGACACACGCGAGGGCAATAAGAGCCTGCCCAAGGACGCA
GATGCAAACGGAGCCTATAATATCGCCCTGAAGGGCCTGTGGGCACTGAG
GCAGATCAGACAGACCTCCGAGGGCGGCAAGCTGAAGCTGGCCATCTCTA
ACAAGGAGTGGCTGCAGTTTGTGCAGGAGAGATCCTACGAGAAGGACAA
AAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGG
ATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGA
TTATGCATACCCATATGATGTCCCCGACTATGCCTAA (SEQ ID NO: 40)

[00183] *Prevotella disiens* (PdCpf1; pY018), including NLS and HA tag:

[00184] MENYQEFTNLFQLNKTLRFELKPIGKTCELLEEGKIFASGSF
LEKDKVRADNVSYVKKEIDKKHKIFIEETLSSFSISNDLLKQYFDCYNELKAFK
KDCKSDEEEVKKTALRNKCTSIQRAMREAISQAFLKSPQKKLLAIKNLIENVF
KADENVQHFSEFTSYFSGFETNRENFYSDEEKSTSIAYRLVHDNLPIFIKNIYIF
EKLKEQFDAKTLSEIFENYKLYVAGSSLDEVFSLEYFNNTLTQKGIDNYNAVI
GKIVKEDKQEIQGLNEHINLYNQKHKDRRLPFFISLKKQILSDREALSWLPDM
FKNDSEVIKALKGFYIEDGFENNVLTPLATLLSSLDKYNLNGIFIRNNEALSSLS
QNVYRNFSIDEAIDANAELQTFNNYELIANALRAKIKKETKQGRKSFEKYEEY
IDKKVKAIDSLSIQEINELVENYVSEFNSNSGNMPRKVEDYFSLMRKGDFGSN
DLIENIKTKLSAAEKLLGTKYQETAKDIFKKDENSKLIKELLDATKQFQHFIKP
LLGTGEEADRDLVFYGDFLPLYEKFEELTLLYNKVRNRLTQKPYSKDKIRLCF
NKPKLMTGWVDSKTEKSDNGTQYGGYLFRKKNEIGEYDYFLGISSKAQLFRK
NEAVIGDYERLDYYQPKANTIYGSAYEGENSYKEDKKRLNKVIIAYIEQIKQT

NIKKSIIESISKYPNISDDDKVTPSSLLEKIKKVSIDSYNGILSFKSFQSVNKEVID
NLLKTISPLKNKAEFLDLINKDYQIFTEVQAVIDEICKQKTFIYFPISNVELEKE
MGDKDKPLCLFQISNKDLSFAKTFSANLRKKRGAENLHTMLFKALMEGNQD
NLDLGSGAIFYRAKSLDGNKPTHPANEAIKCRNVANKDKVSLFTYDIYKNRR
YMENKFLFHLSIVQNYKAANDSAQLNSSATEYIRKADDLHIIGIDRGERNLLY
YSVIDMKGNIVEQDSLNIIRNNDLETDYHDLLDKREKERKANRQNWEAVEGI
KDLKKGYLSQAVHQIAQLMLKYNAIIALEDLGQMFVTRGQKIEKAVYQQFE
KSLVDKLSYLVDKKRPYNELGGILKAYQLASSITKNNSDKQNGFLFYVPAWN
TSKIDPVTGFTDLLRPKAMTIKEAQDFFGAFDNISYNDKGYFEFETNYDKFKIR
MKSAQTRWTICTFGNRIKRKKDKNYWNYEEVELTEEFKKLFKDSNIDYENCN
LKEEIQNKDNRKFFDDLIKLLQLTLQMRNSDDKGNDYIISPVANAEGQFFDSR
NGDKKLPLDADANGAYNIARKGLWNIRQIKQTKNDKKLNLSISSTEWLDFVR
EKPYLKKRPAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDY
A (SEQ ID NO: 41)

[00185] SEQ ID NO: 41 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00186] SEQ ID NO: 41 may be encoded by the following nucleotide sequence:

[00187] ATGGAGAACTATCAGGAGTTCACCAACCTGTTTCAGCTG
AATAAGACACTGAGATTCGAGCTGAAGCCCATCGGCAAGACCTGCGAGCT
GCTGGAGGAGGGCAAGATCTTCGCCAGCGGCTCCTTTCTGGAGAAGGACA
AGGTGAGGGCCGATAACGTGAGCTACGTGAAGAAGGAGATCGACAAGAA
GCACAAGATCTTTATCGAGGAGACACTGAGCTCCTTCTCTATCAGCAACG
ATCTGCTGAAGCAGTACTTTGACTGCTATAATGAGCTGAAGGCCTTCAAG
AAGGACTGTAAGAGCGATGAGGAGGAGGTGAAGAAAACCGCCCTGCGCA
ACAAGTGTACCTCCATCCAGAGGGCCATGCGCGAGGCCATCTCTCAGGCC
TTTCTGAAGAGCCCCCAGAAGAAGCTGCTGGCCATCAAGAACCTGATCGA
GAACGTGTTCAAGGCCGACGAGAATGTGCAGCACTTCTCCGAGTTTACCA
GCTATTTCTCCGGCTTTGAGACAAACAGAGAGAATTTCTACTCTGACGAG
GAGAAGTCCACATCTATCGCCTATAGGCTGGTGCACGATAACCTGCCTAT
CTTCATCAAGAACATCTACATCTTCGAGAAGCTGAAGGAGCAGTTCGACG
CCAAGACCCTGAGCGAGATCTTCGAGAACTACAAGCTGTATGTGGCCGGC

TCTAGCCTGGATGAGGTGTTCTCCCTGGAGTACTTTAACAATACCCTGACA
CAGAAGGGCATCGACAACTATAATGCCGTGATCGGCAAGATCGTGAAGG
AGGATAAGCAGGAGATCCAGGGCCTGAACGAGCACATCAACCTGTATAAT
CAGAAGCACAAGGACCGGAGACTGCCCTTCTTTATCTCCCTGAAGAAGCA
GATCCTGTCCGATCGGGAGGCCCTGTCTTGGCTGCCTGACATGTTCAAGAA
TGATTCTGAAGTGATCAAGGCCCTGAAGGGCTTCTACATCGAGGACGGCT
TTGAGAACAATGTGCTGACACCTCTGGCCACCCTGCTGTCCTCTCTGGATA
AGTACAACCTGAATGGCATCTTTATCCGCAACAATGAGGCCCTGAGCTCC
CTGTCCCAGAACGTGTATCGGAATTTTTCTATCGACGAGGCCATCGATGCC
AACGCCGAGCTGCAGACCTTCAACAATTACGAGCTGATCGCCAATGCCCT
GCGCGCCAAGATCAAGAAGGAGACAAAGCAGGGCCGGAAGTCTTTCGAG
AAGTACGAGGAGTATATCGATAAGAAGGTGAAGGCCATCGACAGCCTGTC
CATCCAGGAGATCAACGAGCTGGTGGAGAATTACGTGAGCGAGTTTAACT
CTAATAGCGGCAACATGCCAAGAAAGGTGGAGGACTACTTCAGCCTGATG
AGGAAGGGCGACTTCGGCTCCAACGATCTGATCGAAAATATCAAGACCAA
GCTGAGCGCCGCAGAGAAGCTGCTGGGCACAAAGTACCAGGAGACAGCC
AAGGACATCTTCAAGAAGGATGAGAACTCCAAGCTGATCAAGGAGCTGCT
GGACGCCACCAAGCAGTTCCAGCACTTTATCAAGCCACTGCTGGGCACAG
GCGAGGAGGCAGATCGGGACCTGGTGTTCTACGGCGATTTTCTGCCCCTG
TATGAGAAGTTTGAGGAGCTGACCCTGCTGTATAACAAGGTGCGGAATAG
ACTGACACAGAAGCCCTATTCCAAGGACAAGATCCGCCTGTGCTTCAACA
AGCCTAAGCTGATGACAGGCTGGGTGGATTCCAAGACCGAGAAGTCTGAC
AACGGCACACAGTACGGCGGCTATCTGTTTCGGAAGAAGAATGAGATCGG
CGAGTACGATTATTTTCTGGGCATCTCTAGCAAGGCCCAGCTGTTCAGAAA
GAACGAGGCCGTGATCGGCGACTACGAGAGGCTGGATTACTATCAGCCAA
AGGCCAATACCATCTACGGCTCTGCCTATGAGGGCGAGAACAGCTACAAG
GAGGACAAGAAGCGGCTGAACAAAGTGATCATCGCCTATATCGAGCAGA
TCAAGCAGACAAACATCAAGAAGTCTATCATCGAGTCCATCTCTAAGTAT
CCTAATATCAGCGACGATGACAAGGTGACCCCATCCTCTCTGCTGGAGAA
GATCAAGAAGGTGTCTATCGACAGCTACAACGGCATCCTGTCCTTCAAGT
CTTTTCAGAGCGTGAACAAGGAAGTGATCGATAACCTGCTGAAAACCATC
AGCCCCCTGAAGAACAAGGCCGAGTTTCTGGACCTGATCAATAAGGATTA
TCAGATCTTCACCGAGGTGCAGGCCGTGATCGACGAGATCTGCAAGCAGA
AAACCTTCATCTACTTTCCAATCTCCAACGTGGAGCTGGAGAAGGAGATG

GGCGATAAGGACAAGCCCCTGTGCCTGTTCCAGATCAGCAATAAGGATCT
GTCCTTCGCCAAGACCTTTAGCGCCAACCTGCGGAAGAAGAGAGGCGCCG
AGAATCTGCACACAATGCTGTTTAAGGCCCTGATGGAGGGCAACCAGGAT
AATCTGGACCTGGGCTCTGGCGCCATCTTCTACAGAGCCAAGAGCCTGGA
CGGCAACAAGCCCACACACCCTGCCAATGAGGCCATCAAGTGTAGGAAC
GTGGCCAATAAGGATAAGGTGTCCCTGTTCACCTACGACATCTATAAGAA
CAGGCGCTACATGGAGAATAAGTTCCTGTTTCACCTGAGCATCGTGCAGA
ACTATAAGGCCGCCAATGACTCCGCCCAGCTGAACAGCTCCGCCACCGAG
TATATCAGAAAGGCCGATGACCTGCACATCATCGGCATCGATAGGGGCGA
GCGCAATCTGCTGTACTATTCCGTGATCGATATGAAGGGCAACATCGTGG
AGCAGGACTCTCTGAATATCATCAGGAACAATGACCTGGAGACAGATTAC
CACGACCTGCTGGATAAGAGGGAGAAGGAGCGCAAGGCCAACCGGCAGA
ATTGGGAGGCCGTGGAGGGCATCAAGGACCTGAAGAAGGGCTACCTGAG
CCAGGCCGTGCACCAGATCGCCCAGCTGATGCTGAAGTATAACGCCATCA
TCGCCCTGGAGGATCTGGGCCAGATGTTTGTGACCCGCGGCCAGAAGATC
GAGAAGGCCGTGTACCAGCAGTTCGAGAAGAGCCTGGTGGATAAGCTGTC
CTACCTGGTGGACAAGAAGCGGCCTTATAATGAGCTGGGCGGCATCCTGA
AGGCCTACCAGCTGGCCTCTAGCATCACCAAGAACAATTCTGACAAGCAG
AACGGCTTCCTGTTTTATGTGCCAGCCTGGAATACAAGCAAGATCGATCCC
GTGACCGGCTTTACAGACCTGCTGCGGCCCAAGGCCATGACCATCAAGGA
GGCCCAGGACTTCTTTGGCGCCTTCGATAACATCTCTTACAATGACAAGGG
CTATTTCGAGTTTGAGACAAACTACGACAAGTTTAAGATCAGAATGAAGA
GCGCCCAGACCAGGTGGACAATCTGCACCTTCGGCAATCGGATCAAGAGA
AAGAAGGATAAGAACTACTGGAATTATGAGGAGGTGGAGCTGACCGAGG
AGTTCAAGAAGCTGTTTAAGGACAGCAACATCGATTACGAGAACTGTAAT
CTGAAGGAGGAGATCCAGAACAAGGACAATCGCAAGTTCTTTGATGACCT
GATCAAGCTGCTGCAGCTGACACTGCAGATGCGGAACTCCGATGACAAGG
GCAATGATTATATCATCTCTCCTGTGGCCAACGCCGAGGGCCAGTTCTTTG
ACTCCCGCAATGGCGATAAGAAGCTGCCACTGGATGCAGACGCAAACGG
AGCCTACAATATCGCCCGCAAGGGCCTGTGGAACATCCGGCAGATCAAGC
AGACCAAGAACGACAAGAAGCTGAATCTGAGCATCTCCTCTACAGAGTGG
CTGGATTTCGTGCGGGAGAAGCCTTACCTGAAGAAAAGGCCGGCGGCCAC
GAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGAT

GTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATAT GATGTCCCCGACTATGCCTAA (SEQ ID NO: 42)

[00188] *Porphyromonas macacae* (PmCpf1; pY09), including NLS and HA tag:

[00189] MKTQHFFEDFTSLYSLSKTIRFELKPIGKTLENIKKNGLIRRD EQRLDDYEKLKKVIDEYHEDFIANILSSFSFSEEILQSYIQNLSESEARAKIEKT MRDTLAKAFSEDERYKSIFKKELVKKDIPVWCPAYKSLCKKFDNFTTSLVPFH ENRKNLYTSNEITASIPYRIVHDNLPKFIQNIEALCELQKKMGADLYLEMMEN LRNVWPSFVKTPDDLCNLKTYNHLMVQSSISEYNRFVGGYSTEDGTKHQGIN EWINIYRQRNKEMRLPGLVFLHKQILAKVDSSSFISDTLENDDQVFCVLRQFR KLFWNTVSSKEDDAASLKDLFCGLSGYDPEAIYVSDAHLATISKNIFDRWNYI SDAIRRKTEVLMPRKKESVERYAEKISKQIKKRQSYSLAELDDLLAHYSEESL PAGFSLLSYFTSLGGQKYLVSDGEVILYEEGSNIWDEVLIAFRDLQVILDKDFT EKKLGKDEEAVSVIKKALDSALRLRKFFDLLSGTGAEIRRDSSFYALYTDRMD KLKGLLKMYDKVRNYLTKKPYSIEKFKLHFDNPSLLSGWDKNKELNNLSVIF RQNGYYYLGIMTPKGKNLFKTLPKLGAEEMFYEKMEYKQIAEPMLMLPKVF FPKKTKPAFAPDQSVVDIYNKKTFKTGQKGFNKKDLYRLIDFYKEALTVHEW KLFNFSFSPTEQYRNIGEFFDEVREQAYKVSMVNVPASYIDEAVENGKLYLFQ IYNKDFSPYSKGIPNLHTLYWKALFSEQNQSRVYKLCGGGELFYRKASLHMQ DTTVHPKGISIHKKNLNKKGETSLFNYDLVKDKRFTEDKFFFHVPISINYKNK KITNVNQMVRDYIAQNDDLQIIGIDRGERNLLYISRIDTRGNLLEQFSLNVIESD KGDLRTDYQKILGDREQERLRRRQEWKSIESIKDLKDGYMSQVVHKICNMV VEHKAIVVLENLNLSFMKGRKKVEKSVYEKFERMLVDKLNYLVVDKKNLSN EPGGLYAAYQLTNPLFSFEELHRYPQSGILFFVDPWNTSLTDPSTGFVNLLGRI NYTNVGDARKFFDRFNAIRYDGKGNILFDLDLSRFDVRVETQRKLWTLTTFG SRIAKSKKSGKWMVERIENLSLCFLELFEQFNIGYRVEKDLKKAILSQDRKEF YVRLIYLFNLMMQIRNSDGEEDYILSPALNEKNLQFDSRLIEAKDLPVDADAN GAYNVARKGLMVVQRIKRGDHESIHRIGRAQWLRYVQEGIVEKRPAATKKA GQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA (SEQ ID NO: 43)

[00190] SEQ ID NO: 43 includes a nuclear localization signal (KRPAATKKAGQAKKKK) (SEQ ID NO: 12), followed by a glycine-serine linker (GS), followed by a HA tag (YPYDVPDYAYPYDVPDYAYPYDVPDYA) (SEQ ID NO: 13).

[00191] SEQ ID NO: 43 may be encoded by the following nucleotide sequence:

[00192] ATGAAAACCCAGCACTTCTTTGAGGACTTCACAAGCCTG
TACTCTCTGAGCAAGACCATCCGGTTTGAGCTGAAGCCAATCGGCAAGAC
CCTGGAGAACATCAAGAAGAATGGCCTGATCCGGAGAGATGAGCAGAGA
CTGGACGATTACGAGAAGCTGAAGAAAGTGATCGACGAGTATCACGAGG
ATTTCATCGCCAACATCCTGAGCTCCTTTTCCTTCTCTGAGGAGATCCTGC
AGTCCTACATCCAGAATCTGAGCGAGTCCGAGGCCAGGGCCAAGATCGAG
AAAACCATGCGCGACACACTGGCCAAGGCCTTCTCTGAGGATGAGAGGTA
CAAGAGCATCTTTAAGAAGGAGCTGGTGAAGAAGGACATCCCCGTGTGGT
GCCCTGCCTATAAGAGCCTGTGCAAGAAGTTCGATAACTTTACCACATCTC
TGGTGCCCTTCCACGAGAACAGGAAGAACCTGTATACCAGCAATGAGATC
ACAGCCTCTATCCCTTATCGCATCGTGCACGTGAACCTGCCAAAGTTTATC
CAGAATATCGAGGCCCTGTGCGAGCTGCAGAAGAAGATGGGCGCCGACCT
GTACCTGGAGATGATGGAGAACCTGCGCAACGTGTGGCCCAGCTTCGTGA
AAACCCCAGACGACCTGTGCAACCTGAAAACCTATAATCACCTGATGGTG
CAGTCTAGCATCAGCGAGTACAACAGGTTTGTGGGCGGCTATTCCACCGA
GGACGGCACAAAGCACCAGGGCATCAACGAGTGGATCAATATCTACAGA
CAGAGGAATAAGGAGATGCGCCTGCCTGGCCTGGTGTTCCTGCACAAGCA
GATCCTGGCCAAGGTGGACTCCTCTAGCTTCATCAGCGATACACTGGAGA
ACGACGATCAGGTGTTTTGCGTGCTGAGACAGTTCAGGAAGCTGTTTTGG
AATACCGTGTCCTCTAAGGAGGACGATGCCGCCTCCCTGAAGGACCTGTT
CTGTGGCCTGTCTGGCTATGACCCTGAGGCCATCTACGTGAGCGATGCCCA
CCTGGCCACAATCTCCAAGAACATCTTTGACAGATGGAATTACATCTCCG
ATGCCATCAGGCGCAAGACCGAGGTGCTGATGCCACGGAAGAAGGAGAG
CGTGGAGAGATATGCCGAGAAGATCTCCAAGCAGATCAAGAAGAGACAG
TCTTACAGCCTGGCCGAGCTGGACGATCTGCTGGCCCACTATAGCGAGGA
GTCCCTGCCCGCAGGCTTCTCTCTGCTGAGCTACTTTACATCTCTGGGCGG
CCAGAAGTATCTGGTGAGCGACGGCGAAGTGATCCTGTACGAGGAGGGC
AGCAACATCTGGGACGAGGTGCTGATCGCCTTCAGGGATCTGCAGGTCAT
CCTGGACAAGGACTTCACCGAGAAGAAGCTGGGCAAGGATGAGGAGGCC
GTGTCTGTGATCAAGAAGGCCCTGGACAGCGCCCTGCGCCTGCGGAAGTT
CTTTGATCTGCTGTCCGGCACAGGCGCAGAGATCAGGAGAGACAGCTCCT
TCTATGCCCTGTATACCGACCGGATGGATAAGCTGAAGGGCCTGCTGAAG

ATGTATGATAAGGTGAGAAACTACCTGACCAAGAAGCCTTATTCCATCGA
GAAGTTCAAGCTGCACTTTGACAACCCATCCCTGCTGTCTGGCTGGGATAA
GAATAAGGAGCTGAACAATCTGTCTGTGATCTTCCGGCAGAACGGCTACT
ATTACCTGGGCATCATGACACCCAAGGGCAAGAATCTGTTCAAGACCCTG
CCTAAGCTGGGCGCCGAGGAGATGTTTTATGAGAAGATGGAGTACAAGCA
GATCGCCGAGCCTATGCTGATGCTGCCAAAGGTGTTCTTTCCCAAGAAAA
CCAAGCCAGCCTTCGCCCCAGACCAGAGCGTGGTGGATATCTACAACAAG
AAAACCTTCAAGACAGGCCAGAAGGGCTTTAATAAGAAGGACCTGTACCG
GCTGATCGACTTCTACAAGGAGGCCCTGACAGTGCACGAGTGGAAGCTGT
TTAACTTCTCCTTTTCTCCAACCGAGCAGTATCGGAATATCGGCGAGTTCT
TTGACGAGGTGAGAGAGCAGGCCTACAAGGTGTCCATGGTGAACGTGCCC
GCCTCTTATATCGACGAGGCCGTGGAGAACGGCAAGCTGTATCTGTTCCA
GATCTACAATAAGGACTTCAGCCCCTACTCCAAGGGCATCCCTAACCTGC
ACACACTGTATTGGAAGGCCCTGTTCAGCGAGCAGAATCAGAGCCGGGTG
TATAAGCTGTGCGGAGGAGGAGAGCTGTTTTATAGAAAGGCCAGCCTGCA
CATGCAGGACACCACAGTGCACCCCAAGGGCATCTCTATCCACAAGAAGA
ACCTGAATAAGAAGGGCGAGACAAGCCTGTTCAACTACGACCTGGTGAAG
GATAAGAGGTTTACCGAGGACAAGTTCTTTTTCCACGTGCCTATCTCTATC
AACTACAAGAATAAGAAGATCACCAACGTGAATCAGATGGTGCGCGATTA
TATCGCCCAGAACGACGATCTGCAGATCATCGGCATCGACCGCGGCGAGC
GGAATCTGCTGTATATCAGCCGGATCGATACAAGGGGCAACCTGCTGGAG
CAGTTCAGCCTGAATGTGATCGAGTCCGACAAGGGCGATCTGAGAACCGA
CTATCAGAAGATCCTGGGCGATCGCGAGCAGGAGCGGCTGAGGCGCCGG
CAGGAGTGGAAGTCTATCGAGAGCATCAAGGACCTGAAGGATGGCTACAT
GAGCCAGGTGGTGCACAAGATCTGTAACATGGTGGTGGAGCACAAGGCC
ATCGTGGTGCTGGAGAACCTGAATCTGAGCTTCATGAAGGGCAGGAAGAA
GGTGGAGAAGTCCGTGTACGAGAAGTTTGAGCGCATGCTGGTGGACAAGC
TGAACTATCTGGTGGTGGATAAGAAGAACCTGTCCAATGAGCCAGGAGGC
CTGTATGCAGCATACCAGCTGACCAATCCACTGTTCTCTTTTGAGGAGCTG
CACAGATACCCCCAGAGCGGCATCCTGTTTTTCGTGGACCCATGGAACAC
CTCTCTGACAGATCCCAGCACAGGCTTCGTGAATCTGCTGGGCAGAATCA
ACTACACCAATGTGGGCGACGCCCGCAAGTTTTTCGATCGGTTTAACGCC
ATCAGATATGACGGCAAGGGCAATATCCTGTTCGACCTGGATCTGTCCAG
ATTTGATGTGAGGGTGGAGACACAGAGGAAGCTGTGGACACTGACCACAT

TCGGCTCTCGCATCGCCAAATCCAAGAAGTCTGGCAAGTGGATGGTGGAG
CGGATCGAGAACCTGAGCCTGTGCTTTCTGGAGCTGTTCGAGCAGTTTAAT
ATCGGCTACAGAGTGGAGAAGGACCTGAAGAAGGCCATCCTGAGCCAGG
ATAGGAAGGAGTTCTATGTGCGCCTGATCTACCTGTTTAACCTGATGATGC
AGATCCGGAACAGCGACGGCGAGGAGGATTATATCCTGTCTCCCGCCCTG
AACGAGAAGAATCTGCAGTTCGACAGCAGGCTGATCGAGGCCAAGGATCT
GCCTGTGGACGCAGATGCAAACGGAGCATACAATGTGGCCCGCAAGGGC
CTGATGGTGGTGCAGAGAATCAAGAGGGGCGACCACGAGTCCATCCACA
GGATCGGAAGGGCACAGTGGCTGAGATATGTGCAGGAGGGCATCGTGGA
GAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAA
GGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCC
TGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA (SEQ ID NO: 44)

[00193] Some of the non-limiting sequences shown above include a sequence such as a nuclear localization signal and/or a tag sequence (such as a HA tags). In various embodiments, a different nuclear localization signal may be present. In some embodiments, no nuclear localization signal is used. In certain embodiments no tag (e.g., no HA tag) is used.

[00194] In various embodiments relating to a protein (such as a protein within a gene-editing complex) the protein may include a nuclear localization signal. For example, the protein (e.g., a Cas protein) may comprise a nuclear localization signal (NLS). Such signals are known in the art, and non-limiting examples are described in Kalderon et al., (1984) *Cell* 39 (3 Pt 2): 499–509; Makkerh et al., (1996) *Curr Biol.* 6 (8): 1025–7; and Dingwall et al., (1991) *Trends in Biochemical Sciences* 16 (12): 478–81, the contents of each of which are hereby incorporated herein by reference. Specific non-limiting examples of nuclear localization signals include GGSGPPKKKRKV (SEQ ID NO: 5), KRPAATKKAGQAKKKK (SEQ ID NO: 12), PKKKRKV (SEQ ID NO: 45), KR[PAATKKAGQA]KKKK (SEQ ID NO: 46), KR[XXXXXXXXXX]KKKK (SEQ ID NO: 47), KKXK (SEQ ID NO: 48), KRXK (SEQ ID NO: 49), KKXR (SEQ ID NO: 50), KRXR (SEQ ID NO: 51), AVKRPAATKKAGQAKKKKLD (SEQ ID NO: 52), MSRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 53), PAAKRVKLD (SEQ ID NO: 54), and KLKIKRPVK (SEQ ID NO: 55).

General Definitions and General Techniques

[00195] Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in cell culture, molecular genetics, and biochemistry).

[00196] As used herein, the term "about" in the context of a numerical value or range means ±10% of the numerical value or range recited or claimed, unless the context requires a more limited range.

[00197] In the descriptions above and in the claims, phrases such as "at least one of" or "one or more of" may occur followed by a conjunctive list of elements or features. The term "and/or" may also occur in a list of two or more elements or features. Unless otherwise implicitly or explicitly contradicted by the context in which it is used, such a phrase is intended to mean any of the listed elements or features individually or any of the recited elements or features in combination with any of the other recited elements or features. For example, the phrases "at least one of A and B;" "one or more of A and B;" and "A and/or B" are each intended to mean "A alone, B alone, or A and B together." A similar interpretation is also intended for lists including three or more items. For example, the phrases "at least one of A, B, and C;" "one or more of A, B, and C;" and "A, B, and/or C" are each intended to mean "A alone, B alone, C alone, A and B together, A and C together, B and C together, or A and B and C together." In addition, use of the term "based on," above and in the claims is intended to mean, "based at least in part on," such that an unrecited feature or element is also permissible.

[00198] The terms "plasma membrane" and "cell membrane" are used interchangeably herein, and refer to the semipermeable membrane that separates the interior of a cell from the environment outside the cell.

[00199] As used herein, an "expression vector" is a DNA or RNA vector that is capable of effecting expression of one or more polynucleotides. Preferably, the expression vector is also capable of replicating within the host cell. Expression vectors can be either prokaryotic or eukaryotic, and are typically plasmids. Expression vectors of the present invention include any vectors that function (*i.e.,* direct gene expression) in host cells of the present invention, including in one of the prokaryotic or eukaryotic cells described herein, *e.g.*, protozoan, algal, fungi, yeast, plant, animal, vertebrate, invertebrate, arthropod, mammalian, rodent, primate, or human cells.

Expression vectors of the present invention contain regulatory sequences such as transcription control sequences, translation control sequences, origins of replication, and other regulatory sequences that are compatible with the host cell and that control the expression of a polynucleotide. In particular, expression vectors of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control sequences are those which control transcription initiation such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art. In preferred embodiments, the methods do not comprise the use of viral vectors such as adenoviruses to deliver nucleic acid molecules or constructs.

[00200] It is understood that where a parameter range is provided, all integers within that range, and tenths thereof, are also provided by the invention. For example, "0.2-5 mg" is a disclosure of 0.2 mg, 0.3 mg, 0.4 mg, 0.5 mg, 0.6 mg etc. up to 5.0 mg.

[00201] Unless otherwise implicitly or explicitly contradicted by the context in which it is used, references to cell "squeeze" "squeezing" "deformation" and the like refer to a process used to deliver macromolecules directly into the cytosol of cells with minimal cytotoxicity. The principle underlying this approach is temporary membrane disruption by rapid mechanical deformation, or squeezing, of the target cell, which permits the uptake by diffusion of macromolecules in the fluid medium and is followed by cell membrane repair (*see, e.g.,* U.S. Patent Application Publication No. 2014/0287509, published September 25, 2014; PCT International Patent Application No. PCT/US2015/058489, filed October 30, 2015; and PCT International Patent Application No. PCT/2015/060689, filed November 13, 2015, the entire contents of each of which are incorporated herein by reference).

[00202] As used herein, "gRNA" refers to a CRISPR-Cas system guide RNA.

[00203] As used herein the term "protein complex" refers to a composite unit arising from the specific binding of a protein with a binding partner, wherein said binding partner can be one or more proteins, one or more nucleic acids, or a combination of one or more proteins and one or more nucleic acids, and the like, to

form said protein complex. Protein complexes may be protein-protein complexes, protein-nucleic acid complexes, and the like. In certain embodiments, a protein complex may comprise protein-protein interactions, e.g. interactions between different proteins, or dimers, trimers, tetramers or higher oligomers of the same protein. Interactions between subunits of protein complexes (e.g., in protein-protein complexes or protein-nucleic acid complexes that comprise more than one protein) or between proteins and nucleic acids (e.g., in protein-nucleic acid complexes) are usually non-binding interactions, such as those interactions caused by hydrogen bridges, pi electron systems such as (optionally conjugated) C--C double bonds or aromatic rings, e.g. phenyl, and heteroaromatic rings, e.g. pyrrole, imidazole, indole, pyrimidine or purine rings, and interactions between metal atoms and oxygen, nitrogen or sulfur atoms, but may also be weak, and in particular reversible, covalent binding interactions, e.g. sulfur-sulfur bridges.

[00204] A "protein-protein complex" means a composite unit that is a combination of two or more proteins formed by interaction between the proteins. Typically but not necessarily, a "protein complex" is formed by the binding of two or more proteins together through specific non-covalent binding affinities. However, covalent bonds may also be present between the interacting partners. For instance, the two interacting partners can be covalently crosslinked so that the protein complex becomes more stable.

[00205] Similarly, a "protein-nucleic acid complex" means a composite unit that is a combination of at least one protein and at least one nucleic acid formed by interactions that include an interaction between a protein and a nucleic acid. Typically but not necessarily, a "protein-nucleic acid complex" is formed by the binding of a protein and a nucleic acid through non-covalent binding affinities.

[00206] In various embodiments, a gene-editing complex is a protein-nucleic acid complex, such as a RNP. A non-limiting example of an RNP is a CRISPR-Cas RNP comprising a Cas protein and a gRNA.

[00207] Methods and devices described herein deliver an intact and functional gene-editing complex into cells. The components of the gene-editing complex do not disassociate during delivery and remain functional after delivery into the cell.

[00208] Various assays are available to determine whether an intact and functional gene-editing complex has been delivered to a cell. For example, the

detection of gene editing by the gene-editing complex may be used to indicate that an intact and functional gene-editing complex was delivered into a cell. Alternatively or in addition, cells to which the gene-editing complex has been delivered may be lysed using non-denaturing conditions (such as a non-denaturing buffer or a French press), and the lysate may be analyzed using a non-denaturing gel to determine whether the gene-editing complex was intact within the cells. Alternatively or in addition, the cells may be lysed using non-denaturing conditions and then immunoprecipitation may be used to isolate the gene-editing complex from the lysate (i.e., to verify that one component of the complex can be co-isolated with another using immunoprecipitation). The isolated gene-editing complex can be assayed before or after delivery to a cell using a non-denaturing gel or a denaturing assay (such as sodium dodecyl sulfate polyacrylamide gel electrophoresis) to determine whether the gene-editing complex was present in a pre-delivery/pre-cell squeeze buffer as well as whether the complex is present after microfluidic/squeeze processing and found intact and/or functional in the treated cells. In some embodiments relating to CRISPR-Cas9 RNPs, a band on a non-denaturing gel of about 145, 150, 155, or 145-160 kDa may indicate that the RNP was delivered as a complete and functional gene-editing complex into the cell.

[00209] As used herein, device dimensions are denoted by a series of numbers indicating length, width, and optionally number of constrictions (e.g., 30μm-6mx5 denotes a device with a 30μm length, 6μm width, and 5 constrictions).

Exemplary Embodiments

[00210] Aspects of the present subject matter provide a method for delivering a protein and a nucleic acid into a cell, the method comprising: providing a cell in a solution; passing the solution through a microfluidic channel that includes a cell-deforming constriction; passing the cell through the constriction such that a pressure is applied to the cell causing perturbations of the cell large enough for the protein and the nucleic acid to pass through; and contacting the cell with the protein and the nucleic acid before, during, and/or after the cell passes through the constriction.

[00211] In some embodiments, said solution comprises the protein and the nucleic acid before, during, and/or after the cell passes through the constriction.

[00212] In some embodiments, the protein and the nucleic acid form a protein-nucleic acid complex.

[00213] In some embodiments, the protein and the nucleic acid are the components of the protein-nucleic acid complex but are not complexed when delivered to the cell.

[00214] In some embodiments, the protein and the nucleic acid form a protein-nucleic acid complex after delivery into the cell.

[00215] In some embodiments, the protein and the nucleic acid form a protein-nucleic acid complex before delivery into the cell.

[00216] In some embodiments, the protein and the nucleic acid comprise gene editing components.

[00217] In some embodiments, said protein-nucleic acid complex comprises a ribonucleoprotein (RNP).

[00218] In some embodiments, (a) the protein is a Cas protein or a Cpf1 protein; and (b) the nucleic acid is a single guide RNA (sgRNA) or a CRISPR RNA (crRNA) and a trans-activating crRNA (tracrRNA).

[00219] In some embodiments, the complex is a RNP comprising a Cas protein or a Cpf1 protein and a sgRNA, wherein the Cas protein or the Cpf1 protein and the sgRNA were complexed using about a 0.5, 2.0, 2.5, or 3.0 molar excess of the Cas protein or Cpf1 protein.

[00220] In some embodiments, the Cas protein comprises a Cas9 protein.

[00221] In some embodiments, said protein-nucleic acid complex comprises a first RNP and a second RNP.

[00222] In some embodiments, the first RNP and the second RNP are both nickases.

[00223] In some embodiments, the first RNP nicks a target sequence different from the target sequence of the second RNP.

[00224] In some embodiments, said protein-nucleic acid complex comprises a TALEN protein, Zinc finger nuclease, mega nuclease, or Cre recombinase.

[00225] In some embodiments, the nucleic acid comprises an mRNA encoding a TALEN protein, a Zinc finger nuclease, a mega nuclease, or a Cre recombinase

[00226] In some embodiments, said protein-nucleic acid complex comprises (a) a nucleic acid molecule that is complexed with a protein via electrostatic attraction; (b) a nucleic acid molecule wrapped around a protein; (c) DNA and a histone; (d) a ribonucleoprotein (RNP); (e) a ribosome, an enzyme telomerase, a vault ribonucleoprotein, RNase P, hnRNP, or a small nuclear RNP (snRNP); or (f) a chromosome comprising a protein.

[00227] In some embodiments, the solution further comprises donor DNA.

[00228] In some embodiments, the solution further comprises donor DNA before, during, and/or after the cell passes through the constriction.

[00229] In some embodiments, said cell comprises a mammalian cell.

[00230] In some embodiments, said cell comprises a human cell.

[00231] In some embodiments, the diameter of the constriction is selected to induce temporary perturbations of the cell membrane large enough for the protein and the nucleic acid to pass through.

[00232] In some embodiments, a diameter of the constriction is about 20-99% of the diameter of the cell.

[00233] In some embodiments, a diameter of the constriction is about 60% of the diameter of the cell.

[00234] In some embodiments, the microfluidic channel is one of a plurality of parallel microfluidic channels in the microfluidic system.

[00235] In some embodiments, the plurality of parallel microfluidic channels comprises at least about 2, 5, 10, 20, 25, 30, 40, 45, 50, 75, 100, 500, 1,000, or 2-1,000 microfluidic channels.

[00236] In some embodiments, the cell is a plurality of cells, and each cell is passed through one of a plurality of parallel microfluidic channels, and wherein each microfluidic channel of the plurality of parallel microfluidic channels includes a cell-deforming constriction.

[00237] In some embodiments, (a) the diameter of the constriction is about 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 2-10μm, or 10-20μm; (b) the length of the constriction is about 10, 15, 20, 24, 30, 40, 50, 60, 70, 80, 90, 100, 10-40, 10-50, 10-60, or 10-100μm; (c) a pressure of about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or 10-100 psi is used to pass the solution through the microfluidic channel; (d) the cell passes through the microfluidic channel at a speed of about 300, 400, 500, 600, 700, 800, 900, 100-300, 200-700, 250-400, 100-1000mm/s, 1-1000mm/s, 1m/s, 2m/s, 3m/s, 4m/s, 5m/s,

6m/s, 7m/s, 8m/s, 9m/s, 10m/s, 0.01-5m/s, 5-10m/s, or 0.01-10m/s; (e) said microfluidic channel comprises multiple cell-deforming constrictions in series; (f) said microfluidic channel comprises a single cell-deforming constriction; (g) the perturbations of the cell membrane include a maximum diameter of about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm; and/or (h) perturbations of the cell membrane having a maximum diameter of about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm persist on the cell membrane for at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 1-10 minutes.

[00238] In some embodiments, (a) the expression of a target gene in the cell is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more; or (b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more, after the protein and the nucleic acid are delivered to the cell.

[00239] In some embodiments, (a) the expression of a target gene in the cell is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more; or (b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more, about 1, 2, 5, 12, 24, 1-12, 6-12, 6-18, 12-24, or 1-24 hours after the protein and the nucleic acid are delivered to the cell.

[00240] In some embodiments, (a) the expression of a target gene in the cell is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more; or (b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more, after the protein and the nucleic acid are delivered to the cell.

[00241] In some embodiments, (a) the expression of a target gene in the cell is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more; or (b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more, about 1, 2, 5, 12, 24, 1-12, 6-12, 6-18, 12-24, or 1-24 hours after the protein and the nucleic acid are delivered to the cell.

[00242] Aspects of the present subject matter provide a device for delivering a protein-nucleic acid complex to a cell, comprising at least one

microfluidic channel, wherein said channel comprises a constriction length of about 30 μm and a constriction width of about 4 μm.

[00243] Examples are provided below to facilitate a more complete understanding of the invention. The following examples illustrate the exemplary modes of making and practicing the invention. However, the scope of the invention is not limited to specific embodiments disclosed in these Examples, which are for purposes of illustration only, since alternative methods can be utilized to obtain similar results.

Example 1: Editing of the B2M locus in primary human T cells using CellSqueeze (SQZ) to deliver CRISPR/Cas9 gene editing complex

[00244] A series of experiments have been undertaken in unstimulated human T cells to demonstrate the ability of the SQZ platform to deliver Cas9 ribonucleoproteins (RNPs; recombinant Cas9 protein complexed with a single-guide RNA) and accomplish efficient genome editing of a model locus, the β_2 microglobulin component of MHC class 1 (B2M).

Delivery of Cas9 RNP to Unstimulated Human T cells

[00245] Fresh PBMCs were isolated from human blood using a standard Ficoll gradient. Next, T cells were negatively selected (Human T cell enrichment kit (StemCell Technologies)) counted, washed and resuspended at 10-20 $x10^6$ cells/mL in OptiMEM for delivery. Ten μg of recombinant CAS9 (PNA Bio) was pre-complexed with a 2.5 molar excess of unmodified gRNA (PNA Bio) designed to specifically target the B2M locus. Recombinant CAS9 is reconstituted to a solution with a final concentration of 20mM Hepes, 150mM KCl, 1% sucrose. gRNA is added directly to the CAS9 solution and incubated on ice for 20 minutes to form the complex. The complex is added directly to resuspended cells. RNP complexes were incubated on ice 20 minutes prior to SQZ-mediated delivery. The RNP (2.2uM) was co-delivered with a 3kD-Cascade Blue Dextran (0.15mg/mL) used as a proxy for delivery efficiency. Two different chips, 10-4 and 30-4 were used to deliver the complex at a pressures of 60 and 90 psi. The chips have constrictions of the same width (4 microns) but have two different constriction lengths (30 vs. 10 microns).

[00246] At 48 hours post-delivery, a FACS based readout was used to determine B2M protein levels. Reduced B2M expression was used as a measure of functional editing. Two controls were used; 1) T cells incubated with the RNP

complex at room temperature for the same time as the delivery process using the Cell Squeeze process (endocytosis control; "endo control"), and 2) T cells squeezed with Cas9 protein but no gRNA. Plots of B2M expression vs. delivered dextran are shown (FIG. 4) for the four different cell populations. B2M expression on the Cas9 control was not significantly different than the endocytosis control. The delivery of the RNP using the 30-4 chip at 90 psi resulted in a 54.4% reduction in B2M expression as compared to the endo control whereas the 10-4 chip at 90 psi resulted in a 25.2% reduction in B2M expression. The longer constriction chip resulted in more delivery of the RNP complex and a larger reduction in B2M expression.

[00247] Dextran delivery was used to define low, mid and high delivered populations. The differences in efficiency of B2M knockdown for these specific populations was then determined using the mean fluorescence intensity (MFI) of B2M staining. For the 10-4 chip, the MFI of the highly delivered population was 18,637 versus 71,173 for the mid delivered populations and 83,676 for the low or non-delivered populations. This nearly 5-fold intensity drop in B2M staining for the high delivered populations demonstrates the degree to which delivery influences RNP activity. Similarly, for the 30-4 chip, the MFI of the highly delivered population was 16,460 versus 44,207 for the mid delivered populations and 54,159 for the low delivered population. These data demonstrated the importance that the cell squeezing delivery system of gene editing complexes to the cytosol of a cell has on editing efficiency, even within a single population.

[00248] To confirm the FACS readout, a second, sequence based analysis, was also employed in which DNA was extracted and amplified using primers flanking the target region thereby generating an amplicon of the edited region for Next Generation Sequencing (NGS). Sequencing results were analyzed using a simple algorithm designed to detect CRISPR variants from NGS reads. As expected, the sequence-based readout showed higher editing efficiencies. Indeed, some of the indels identified in sequencing still resulted in a functional, full length protein (i.e. single base substitutions that did not change the resultant amino acid).

Table: Comparison of FACS- and Sequence-based readouts from 10-4 editing experiment.

	Endo	Cas9	RNP
FACS	0.3	8.15	**20.4**
Sequencing	3.87	3.04	**27.18**

[00249] These data demonstrated successful editing ability of the RNP complex when delivered by the Cell Squeeze platform.

Effect of RNP complex amount on editing efficiency

[00250] RNP complex was delivered to unstimulated human T cells using the 30-4 chip and at two different RNP amounts: 1) the standard 1X RNP complex (10ug Cas9, 2.5 molar excess of gRNA) and, 2) 0.1X the standard RNP complex amount. At 48 hours post-delivery, a FACS based readout was used to determine B2M protein levels. Reduced B2M expression was used as a measure of functional editing. Plots of B2M expression vs. delivered dextran are shown below for the four different cell populations. Two controls were used; 1) T cells incubated in 1X RNP complex at room temperature for the same time as the delivery process using the Cell Squeeze process (endocytosis control), and 2) T cells squeezed with Cas9 protein but no gRNA (FIG. 5).

[00251] B2M expression on the Cas9 control (Cas9 protein with no gRNA) is not significantly different than the endocytosis control. The lower amount of the RNP complex (0.1XRNP) resulted in a 20.7% reduction of B2M positive cells as compared to the 55.4% reduction in B2M positive cells at the higher amount of RNP complex (1X RNP complex (10ug CAS9, 2.5 molar excess of gRNA)). This experiment demonstrates a dose-dependent response directly related to the delivery of the RNP.

OTHER EMBODIMENTS

[00252] Cited references are incorporated herein by reference. To the extent that any of the incorporated material is inconsistent with the present disclosure, the present disclosure shall control. Furthermore, to the extent necessary, material incorporated by reference herein should be disregarded if necessary to preserve the validity of the claims.

[00253] Further, while the description above refers to the invention, the description may include more than one invention.

[00254] The subject matter described herein can be embodied in systems, apparatus, methods, and/or articles depending on the desired configuration. The implementations set forth in the foregoing description do not represent all implementations consistent with the subject matter described herein. Instead, they are merely some examples consistent with aspects related to the described subject matter. Although a few variations have been described in detail above, other modifications or additions are possible. In particular, further features and/or variations can be provided in addition to those set forth herein. For example, the implementations described above can be directed to various combinations and subcombinations of the disclosed features and/or combinations and subcombinations of several further features disclosed above. In addition, the logic flows depicted in the accompanying figures and/or described herein do not necessarily require the particular order shown, or sequential order, to achieve desirable results. Other implementations may be within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A method for delivering a protein and a nucleic acid into a cell, the method comprising:

providing a cell in a solution;

passing the solution through a microfluidic channel that includes a cell-deforming constriction;

passing the cell through the constriction such that a pressure is applied to the cell causing perturbations of the cell large enough for the protein and the nucleic acid to pass through; and

contacting the cell with the protein and the nucleic acid before, during, and/or after the cell passes through the constriction.

2. The method of claim 1, wherein said solution comprises the protein and the nucleic acid before, during, and/or after the cell passes through the constriction.

3. The method of claim 1 or 2, wherein the protein and the nucleic acid form a protein-nucleic acid complex.

4. The method of any one of claims 1-3, wherein the protein and the nucleic acid are the components of the protein-nucleic acid complex but are not complexed when delivered to the cell.

5. The method of claim 4, wherein the protein and the nucleic acid form a protein-nucleic acid complex after delivery into the cell.

6. The method of any one of claims 1-4, wherein the protein and the nucleic acid form a protein-nucleic acid complex before delivery into the cell.

7. The method of any one of claims 1-6, wherein the protein and the nucleic acid comprise gene editing components.

8. The method of any one of claims 3-7, wherein said protein-nucleic acid complex comprises a ribonucleoprotein (RNP).

9. The method of claim 7, wherein

(a) the protein is a Cas protein or a Cpf1 protein; and

(b) the nucleic acid is a single guide RNA (sgRNA) or a CRISPR RNA (crRNA) and a trans-activating crRNA (tracrRNA).

10. The method of claim 9, wherein the complex is a RNP comprising a Cas protein or a Cpf1 protein and a sgRNA, wherein the Cas protein or the Cpf1 protein and the sgRNA were complexed using about a 0.5, 2.0, 2.5, or 3.0 molar excess of the Cas protein or Cpf1 protein.

11. The method of claim 9, wherein the Cas protein comprises a Cas9 protein.

12. The method of any one of claims 3-11, wherein said protein-nucleic acid complex comprises a first RNP and a second RNP.

13. The method of claim 11, wherein the first RNP and the second RNP are both nickases.

14. The method of claim 13, wherein the first RNP nicks a target sequence different from the target sequence of the second RNP.

15. The method of any one of claims 3-8, wherein said protein-nucleic acid complex comprises a TALEN protein, Zinc finger nuclease, mega nuclease, or Cre recombinase.

16. The method of claim 1, wherein the nucleic acid comprises an mRNA encoding a TALEN protein, a Zinc finger nuclease, a mega nuclease, or a Cre recombinase

17. The method of any one of claims 3-8, wherein said protein-nucleic acid complex comprises

(a) a nucleic acid molecule that is complexed with a protein via electrostatic attraction;

(b) a nucleic acid molecule wrapped around a protein;

(c) DNA and a histone;

(d) a ribonucleoprotein (RNP);

(e) a ribosome, an enzyme telomerase, a vault ribonucleoprotein, RNase P, hnRNP, or a small nuclear RNP (snRNP); or

(f) a chromosome comprising a protein.

18. The method of any one of claims 1-17, wherein the solution further comprises donor DNA.

19. The method of claim 18, wherein the solution further comprises donor DNA before, during, and/or after the cell passes through the constriction.

20. The method of any one of claims 1-19, wherein said cell comprises a mammalian cell.

21. The method of any one of claims 1-20, wherein said cell comprises a human cell.

22. The method of any one of claims 1-21, wherein the diameter of the constriction is selected to induce temporary perturbations of the cell membrane large enough for the protein and the nucleic acid to pass through.

23. The method of any one of claims 1-22, wherein a diameter of the constriction is about 20-99% of the diameter of the cell.

24. The method of any one of claims 1-23, wherein a diameter of the constriction is about 60% of the diameter of the cell.

25. The method of any one of claims 1-24, wherein the microfluidic channel is one of a plurality of parallel microfluidic channels in the microfluidic system.

26. The method of claim 25, wherein the plurality of parallel microfluidic channels comprises at least about 2, 5, 10, 20, 25, 30, 40, 45, 50, 75, 100, 500, 1,000, or 2-1,000 microfluidic channels.

27. The method of any one of claims 1-26, wherein the cell is a plurality of cells, and each cell is passed through one of a plurality of parallel microfluidic channels, and wherein each microfluidic channel of the plurality of parallel microfluidic channels includes a cell-deforming constriction.

28. The method of any one of claims 1-27, wherein

(a) the diameter of the constriction is about 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 2-10μm, or 10-20μm;

(b) the length of the constriction is about 10, 15, 20, 24, 30, 40, 50, 60, 70, 80, 90, 100, 10-40, 10-50, 10-60, or 10-100μm;

(c) a pressure of about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or 10-100 psi is used to pass the solution through the microfluidic channel;

(d) the cell passes through the microfluidic channel at a speed of about 300, 400, 500, 600, 700, 800, 900, 100-300, 200-700, 250-400, 100-1000mm/s, 1-1000mm/s, 1m/s, 2m/s, 3m/s, 4m/s, 5m/s, 6m/s, 7m/s, 8m/s, 9m/s, 10m/s, 0.01-5m/s, 5-10m/s, or 0.01-10m/s;

(e) said microfluidic channel comprises multiple cell-deforming constrictions in series;

(f) said microfluidic channel comprises a single cell-deforming constriction;

(g) the perturbations of the cell membrane include a maximum diameter of about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm; and/or

(h) perturbations of the cell membrane having a maximum diameter of about 1-20, 1-600, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, or 600 nm persist on the cell membrane for at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 1-10 minutes.

29. The method of any one of claims 7-11, wherein

(a) the expression of a target gene in the cell is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more; or

(b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more,

after the protein and the nucleic acid are delivered to the cell.

30. The method of claim 29, wherein

(a) the expression of a target gene in the cell is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more; or

(b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is reduced by at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, or 99% or more,

about 1, 2, 5, 12, 24, 1-12, 6-12, 6-18, 12-24, or 1-24 hours after the protein and the nucleic acid are delivered to the cell.

31. The method of any one of claims 7-11, wherein

(a) the expression of a target gene in the cell is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more; or

(b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more,

after the protein and the nucleic acid are delivered to the cell.

32. The method of claim 31, wherein

(a) the expression of a target gene in the cell is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more; or

(b) the cell is a plurality of cells and the expression of a target gene in the plurality of cells is increased by at least about 5, 10, 25, 50, 75, 100, 250, 500% or more,

about 1, 2, 5, 12, 24, 1-12, 6-12, 6-18, 12-24, or 1-24 hours after the protein and the nucleic acid are delivered to the cell.

33. A device for delivering a protein-nucleic acid complex to a cell, comprising at least one microfluidic channel, wherein said channel comprises a constriction length of about 30 μm and a constriction width of about 4 μm.

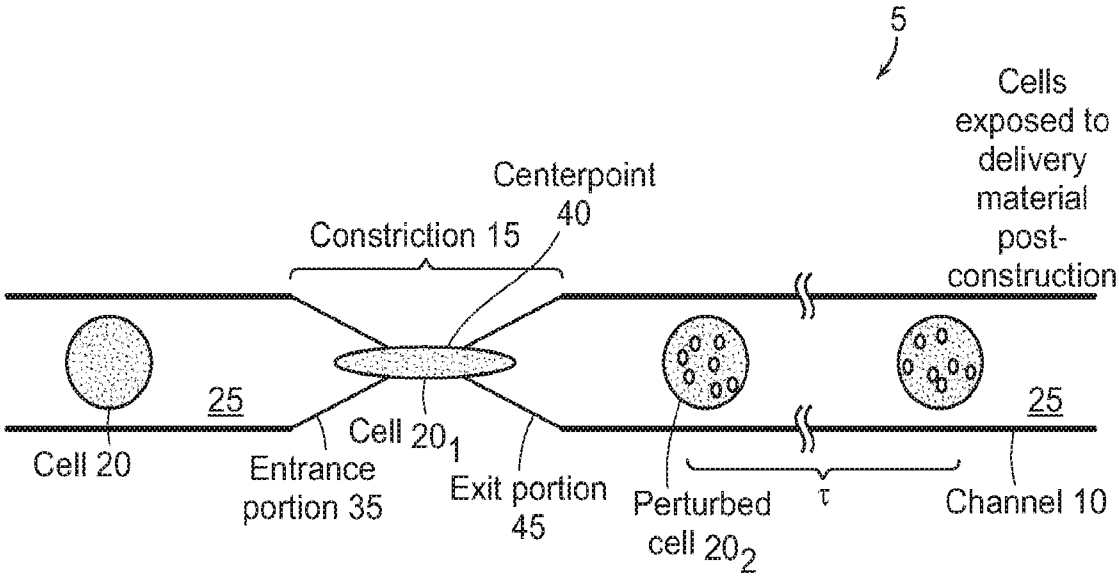

5
Cells exposed to delivery material post-construction
Centerpoint 40
Constriction 15
25
Cell 20
Cell 20_1
Entrance portion 35
Exit portion 45
Perturbed cell 20_2
τ
Channel 10
25

FIG. 1A

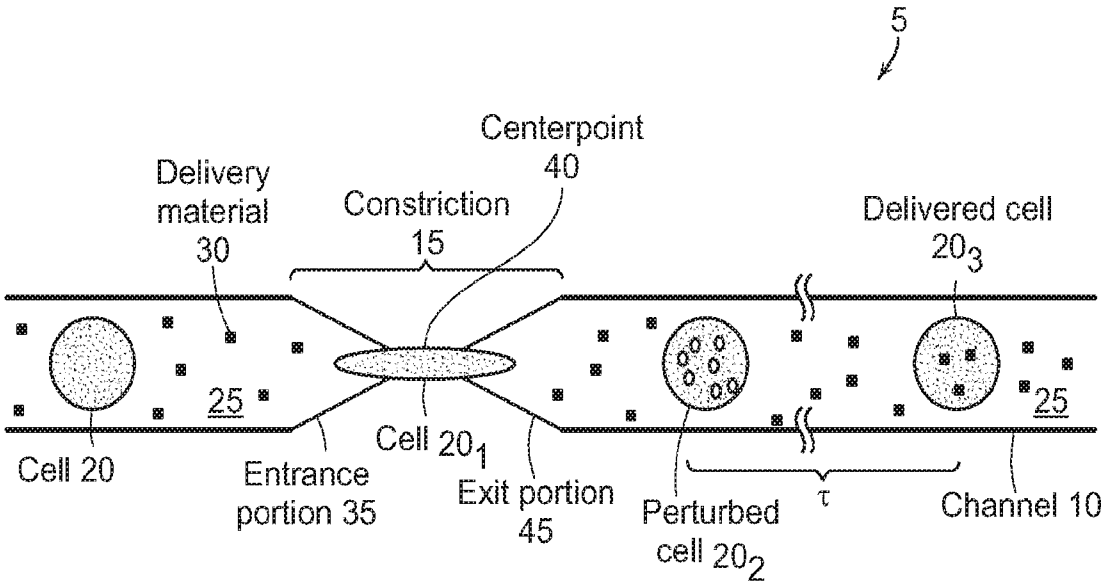

5
Centerpoint 40
Delivery material 30
Constriction 15
Delivered cell 20_3
25
Cell 20
Cell 20_1
Entrance portion 35
Exit portion 45
Perturbed cell 20_2
τ
Channel 10
25

FIG. 1B

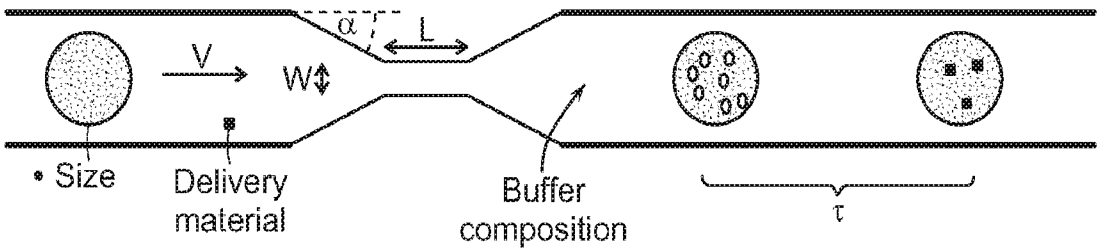

α
L
V
W
• Size
Delivery material
Buffer composition
τ

FIG. 2A

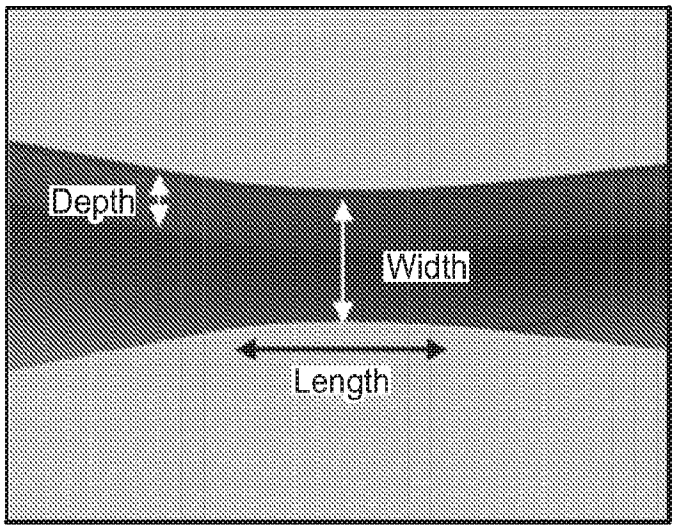
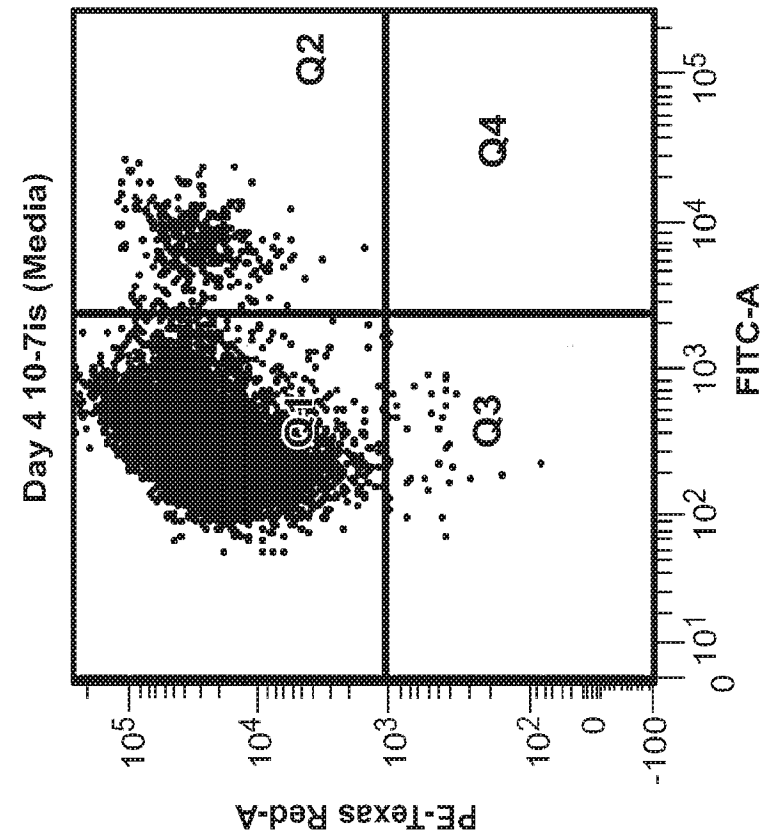

Depth
Width
Length

FIG. 2B


Day 4 10-7is (Media)
Q2
Q4
Q3
@1
PE-Texas Red-A
FITC-A
10^5
10^4
10^3
10^2
0
-100
10^1
0

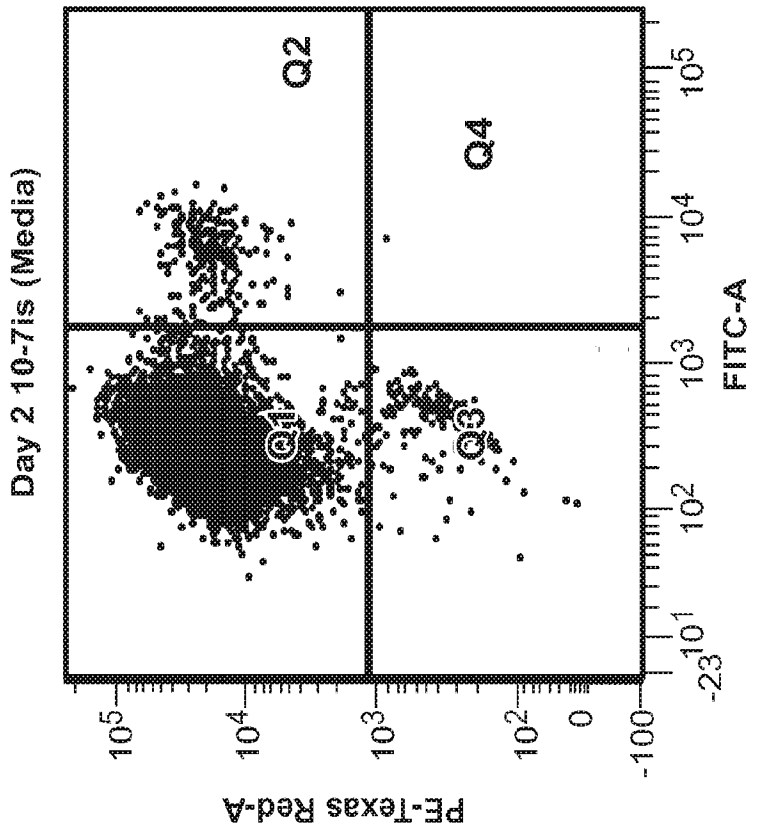

Day 2 10-7is (Media)
Q2
Q4
Q3
@1
PE-Texas Red-A
FITC-A
10^5
10^4
10^3
10^2
0
-100
10^1
-23

FIG. 3A

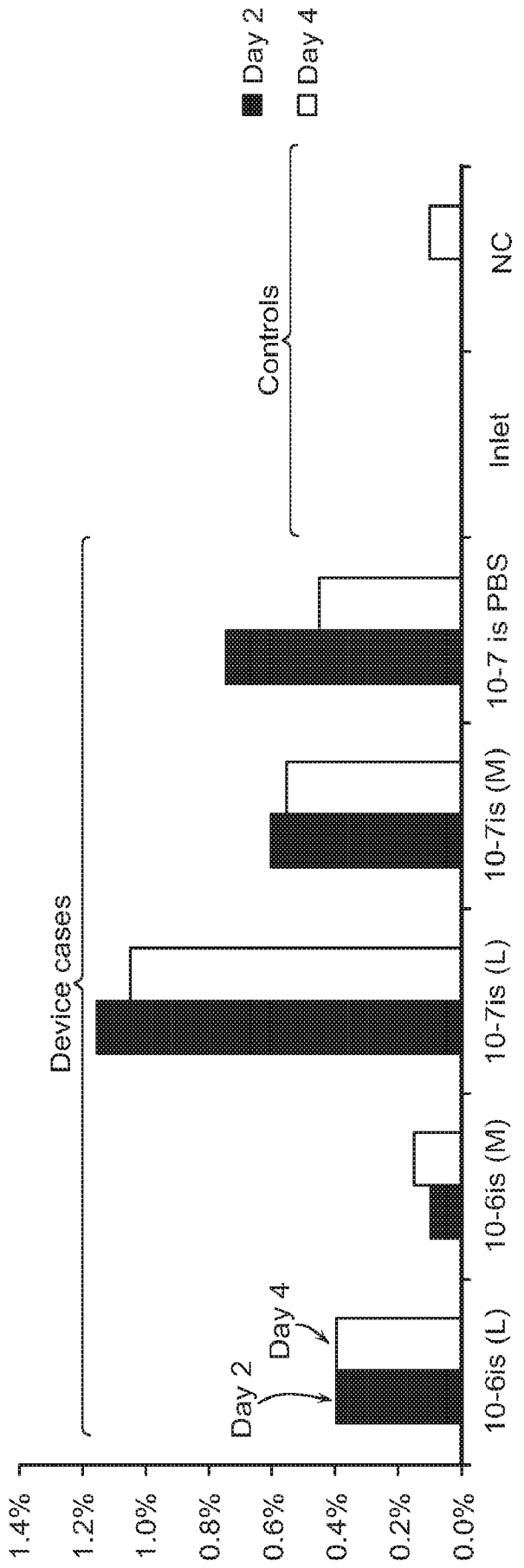

% Recombination
1.4%
1.2%
1.0%
0.8%
0.6%
0.4%
0.2%
0.0%
Device cases
Controls
Day 2
Day 4
Day 2
Day 4
10-6is (L)
10-6is (M)
10-7is (L)
10-7is (M)
10-7 is PBS
Inlet
NC

FIG. 3B

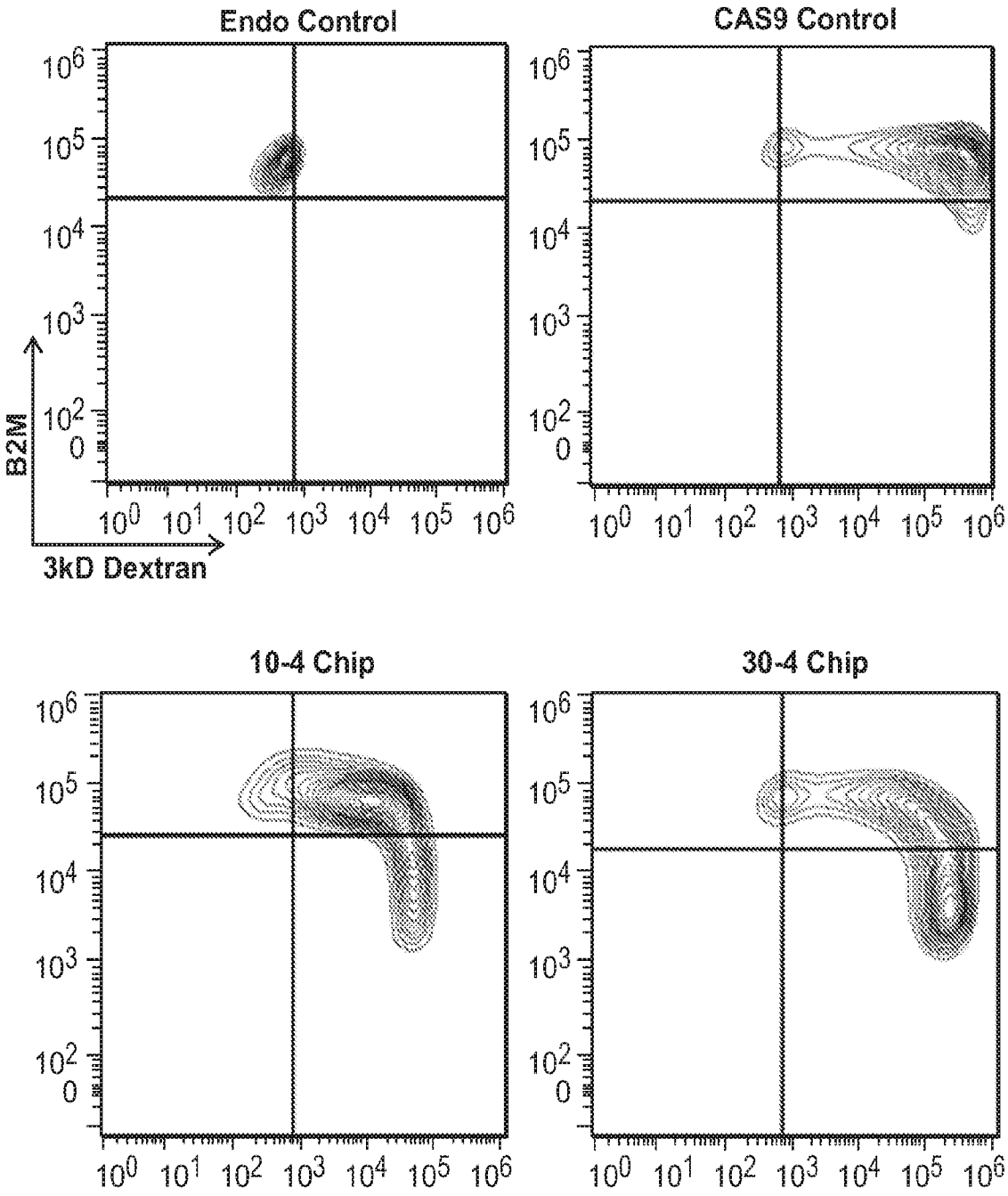

Endo Control
CAS9 Control
10-4 Chip
30-4 Chip
B2M
3kD Dextran
10^0 10^1 10^2 10^3 10^4 10^5 10^6
0

FIG. 4

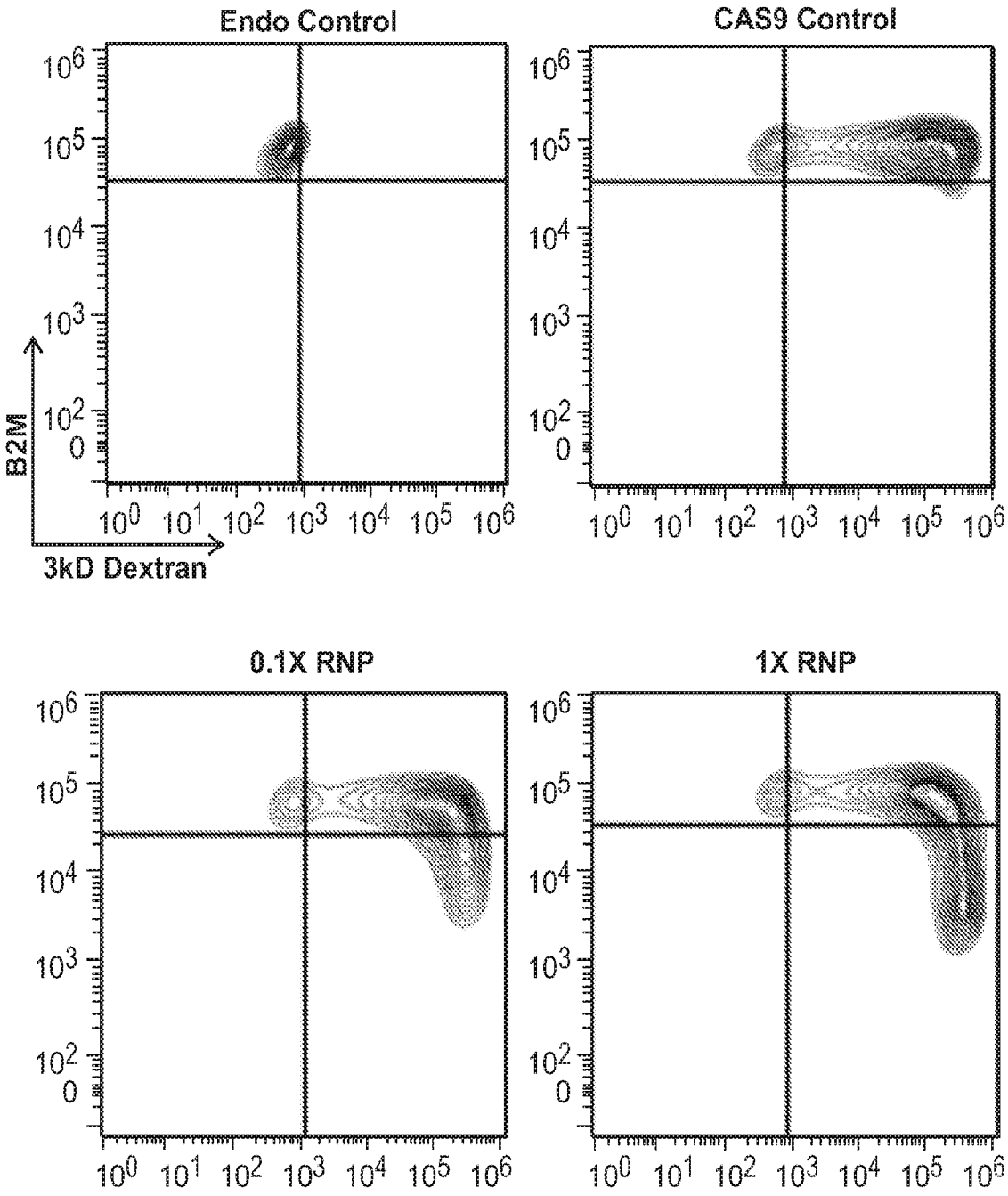

Endo Control
CAS9 Control
0.1X RNP
1X RNP
B2M
3kD Dextran
10^0 10^1 10^2 10^3 10^4 10^5 10^6
0

FIG. 5

TAGGGGATGACAGTGGGCTCTCCGCTTTCTCCTCCATGAAGTAACTTACATGCCCCTCAC
CCTCTGTGTGGGAGGGGTGTTGCAGGGGGGTGCAGAACTCCCCCTCGCCGGGTAGTTCAAGCAA
TGGGGACCATATCAATTCCATCTATAGGGAAACTGAGGCCTGGAGTAGGGCGAGGCCCTCT
GGGAACCCAGCCCTATTCTGTCTCTTTCCCTGGCATTTCCCATCCACACATAGAGCTTCA
GATTCTCTCTTTTCTTTCCCCAGAGACCCCTCAAATATCCCTCTCACTCACAGAATGGTGTCTCT
GCCTGCCTCGGGTTGGCCCTGTGATTTTATTTTAGTTCTTTTCCCTTGTTTTTTTTTTTC
AAACTCTATACACTTTTGTTTTAAAAACTGTGGTTTCTCATGAGCCCTATTATCTCATTG
ATACCTCTCACCTCTGTGGTGAGGGGAAGAAATCATATTTTCAGATGACTCGTAAAGGGCC
AAAGAAAAAAACCCAAAATTTCAAAATTTCCGTTTAAGTCTCATAATCAAGAAAAGGAGA
AACACAGAGAGAGAAAAAAAAACTATGAGAACCCCCTCCCCACCCCGTGATTATCAGCC
GCACACACTCATCGAAAAAAATTTGATTATAGAAGAGAGGTCTGCGGCTTCCACAC
CGTACAGCCGTGGTTTTTCTTCTCGGTATAAAAGCAAAGTTGTTTTTGATACGTGACAGTT
TCCCACAAGCCCAGGCCTGATCCTTTTCTGTCAGTCCACTTCACCAAGGTGAGTGTCCCTGC
TCTCCCCTACCAGATGTGGGCCCCATTGGAGGAGATGGCAGGGAGGGTAGGCACGGCGGGG
GGGTCAGGGGCCCTCTGGTACAGTGGGATGTACCCAGCTACCGTGATTCCAGCCAGGTAA
GGTCTTTAAAAAAATAATAGAATAAATGGCAGAAGACTTAAAGGTGAAGTACTTAATGCA
TGGGCAGGGGCAAAAAATAAATGAATTCATCAATTGATACGTAGATGGATAGGATTTTTTAA
AATATGTCAGCTGCCAGTCTTTATGGATTTCAGAGCTTTGAACAACTTGAAATGCTGTCC
AGGTGCAACTGAAAATTTCAACTGCATCTCTCATCTCTTTGTCTCTTAGCTCTTGATG
CCTCTCATCCTTCTCTGTCTCTGCTCACCCTGTCTTGCTCTTTTACATGTTTCTCTGC
CTCTCTTTTCTCCCGCCATCTGTCTGTCGCTACACTCAGACTCTCTTTCCCCGCATCTCTC
CTGGCCTCTGACCATCTCCCCTATCTCTGTCTCCTCCATCTTGACCACACTCCATCGATC
TCCTTTCCTCCATCTCTCCCGCTCTCTTTTCTTGCCTTCCCCCACCTGCCTTCCCAGCCA
GCATCCAGCCTAGAGCTTTGTACACACAGCTGGCGTTTAATAATAGTAATGTCGCCGAGGAC
CTTTACATTTTATGAGAGTGAGCATCAGTTCTGTTCACCCTAGTAGCAGGAGGATTGCAA
GGTGCCGGGGCAAGCTGAGCACTTAACTGAACTTTGCTTGTATCGCTTTAGAGGAGACCC

FIG. 6

FIG. 6
(continued)

TTGGGGCGGGGGGCAGGAGCACTTCACACCATTGTCAACCGTAGTTTAGCTGTGTGGA
CCTACCCCCAGGACTGCATGTGTGTGACCAGAGTCCATATATTCAGGAAGCAAACAGGTTGTT
CAGGGCTGGGGTGAGCTCGATTCTGCAGGCTTAGCACACAGCGTTTGGCACAAAGTAAGTGC
TCACGAAGTGTTAGCTATTTAGTATTTCTGTAACCAAATTAGAATCATGCTATATATTGGT
TTTGTTCCTGCTTTATTCAATTAACAGTAGTGTTTTTTGAAGCTTACTAAAAATTGTTTGA
ACACATGATCTTTCATGATGGACAAGACTGTCCTCACATCAGCATGAACTTTGATTGATT
TGACCAAGTGCCGATTGTTGGACATTCCGGTTGTTTCTCGTTTTTTTCTATAAACATCCC
TGACTGCTTCCTTGGATTCCTGAAAAGGGAGTTGCTGGGGGCAAAGGGGGCGTGCATTTTT
AAAGTTTATGACCCCTCTCTGCTAATTCTCCCTTCCAAAAGGTTGTCCTATTTTTATGCTCCC
TCTACTTAAGCTAACCCCTAGGTGAGTATTATTATGTTTTTAATTATTTCCCCAACATGACAG
GTAAGAAAAACACATTATCACATTGTTGAAATTTGCATCTCTTGTGATTACTTTTTACTATTC
ATTGGCCATTTCATGCTTTTTTTTTTTTTTGGTTGTCTGGTCATGTCCCTTACTCCCTA
TATTAACCTTTTCATCAGCCATTACATATAGCAAATATTTTTCCCCAATTCATCATTTTGC
TTTTTTTCTATTTGTTTTTTTGTGATGCTTATGATGGTTTTTATACAGTCAAA
TAGGTTAGTCTTTTTTCTGTGGCTTCTCTGTCTCTGGTTTTTGTGCTTAGAAAGTCCTTTTCC
TACTTGAAAATGAGATAAATGTTCACCTATGTTGGCTTCTAGTCTCTTTTTATGGCTTCAT
TTTTTCCATTTACTATAGAGGTTAAGAGTGTGGGTACTGGAGCCAGACTGTCTCTGGGACAA
ACCCAGCGTCACCCCAAGCCCTATGTGTGATTTTTTAGCCAGGCACTTAACCTCTCCCATAC
CTCCATTTTCCTCATATGTACTGCAATGGTTATAATAGTACCTTCCTCAGGAGTCTTTTGTT
TAGATTAAAATTTTTAACCACAGTAAATACTTAGCACACAAGGCCTGACACACAATAAACCC
AAGATCAGCATTAGGTGTTAAAACTTATATCTTGAATGGATCTGGGCATTTTTAGGGTATA
TGATGATGGTGACATTTCAAACTGGGCAGGGAGGGTTGTTGGGATAATGGACTGACTAT
TCACTCAATAACTTTATCTTCTCCCTAATTATCTCAGACACATTTGTTAAACCATACTCC
ATGGTCCTCCAGTTTGAAATGCCACACATTCACATCAAATTCAGTTCTCTCCATATGTGGGT
CCATGTCCAAGCTTTCTATTCTGTTCTCTTCCCCAGTATCTGTTCTTATACCAGGGCCAC
ACTGTTTTTGCTGATTGTTGCTTTGCAATACAATTCAATACCCAGCCATGGGTGTCCCTGG

FIG. 6
(continued)

CACCTTGCCCTTCTTATTCTTGATTATTTCAGGCAGCATCAGTTGAACCAGCCAGAGAC
CAGCAAATGTTCTTATTTGAGGCAATCCTCCTCTCGCACACATGCACTCTGGCTCTCCAT
GCATGTGTCCATTTCTCTCTCACTATCTCTTGTTCTCTCTTGCTCGCTCTTT
GTGTGTGTCTCTCTTTGTCTCTATGTATCCCTATCTCTCAGCCAGTCCCCCTGGATAGAGGGG
CAATTGTCAACTTGAAGCCCTGCAGACACCTAATGACTTAACCAGACAGCGTAGAAGGGC
CCTGGCCCTTGTCTACTCCACGCCTCTCCGTGCTCAGTGTAGAAGGGCAAATTGAAGACC
AGAGATCTCAGGGCCTCACCGAAATGCAGCGGCAGAGTTGAAATCCAAGCCTAGATCTCA
GGACTCTAGGTGGGACCCTGGTTCTGGCTCTCTCCCCAACTGCAGGCCTCAGTTTACCCC
TCAGCACCCAGAAGGGGGAAGGGGAACCTGGGCTACCATTCCCCCTTCTGCCTTTCTCACA
CGTTGGACCCCCAACTTCCCACAGGTTGGACGATCCACGATCACACAGTGTGGGGCCCAGCCT
CACAAGAGCTGGGCTAGGTGAGGCCCCGGACTCCATAGGTCAGGAGGCCTAGTTGGCCAG
AGCGTGGTGATGATGGAGGCATGTCAGTCAGTCAGGCTGTGTGTCCCCCAGAGCTGGTGCT
GGTCCCCCGAAAACCTTGATTGTGGGGCCCCCTCTAGAGAGTCTGATGATGGGCTCTGTATT
GGCGAAGGCTGAGGCTTTTCCAGCTCCCCCCATGAGGCCCAGACCAAAGGCACACCAGCC
TCAACCTCCTCCCTCCCCCTGTTGCCATCTCTGGCGGAGTGGCCATGTATTTGGGAACGTT
GTTCCAGAGTGGACAGGGAGACTGAGGGAGGGCTAGGGAGGCTGGCTCTGTTTCCAGGCCTGG
TAGGCAAGAGGCCCTATGAAGCAGCAAGCTGCCTGACTTTCAGATGGTTCCAAGGAGTTT
GGACACCAGGGACACTGGCCTACACATACTGAGACTTTGGGACCGTAGACCCCCACAGTCT
GTGGTTTTGAGATTCTAGGATCCTTTAAATCTAAGAAATGCTGTTCTATGATTCTGAGGT
CCTGGTGTTATACTATTTGAAGACCCCCAGGGGTCCCAGTATCTGTGGAGCCTGCCTGGCA
CTCTCAGAGCTTCAAACCTGGGTCCTCTCCACAACCCAAGAAGGGCCAGGTCTTCAGAGC
TAGGGGCTTGTCATAGTGGCCAGATGGACATCACCTACCACATCCACCAGCACCCATGTC
ACCCCACCTGGGCCAAGCCTGCTGCAGGACAGGGCAGCCAGTTCTCGGAACGAAACCCTGT
GGGGTGGGGTATCTGCCCTCTTCTCCTTCCGTGGTGTCGATGAAGCCCCGGCGCATCCCG
GCCGCCATGACGTCAATGGCGGAAAAATCTGGGCAAGTCGGGGGCTGTGACAACAGGGCC
CAGATGCAGACCCCCGATATGAAAACATAATCTGTGTCCCAGAAACATCCCCCCATTCAGCT
TCTGAGAAACCCAGTCAGAAAGGGACGTCCCAACAGACAGTGCAGGAAGCCGGGCTGCCCA

GCCCGGCCCTCTAGGTCCTCTACCCCCAGACAGATCATCTCCATGTCCCTGTCTGAGAAT
GTATCTATGCTTTGCTGAGTCAGGCCATCCCACATGTGTTTGGGAGAATTCTTAGCTCT
GGCCAAGTGTCCAGGCAGCTTCAGAAGTGACCACACAGGCCAGCCACATGGGCCAGGCCAGA
GTGGTGGAAAACATCCATTTGCACCGAAATCGGTATTAGTTTGTTCTGGCTGCTATAACA
AAGTACCACAACTGAGTGGCTTCAGCAACAGAAATTGATCATCTCACAGTTCTGGAGGCC
AGAGTCCAAGATCAAGGTGTTGCCAGGGTTGGTTCCTTCTGGGGGCCATAAGGGAGAAGC
TGCTCCAGGCCTCTCCCCCAGCTTCTGGTGGTTGCTGGCAATCTTTGTTATTTCTTGGCT
TGAAGGAGCATCACCCCATCTCTGGCTTCATTCATTCACTCACCTTCTCATCTCACCAGT
GCTCTCCCAGGCCCCTTGTTCGGAGCCTCCACAGACCCACACTCCTGTTTCTCACGCAGTA
ATGTTCTAAGCCCCCAGGAGACTAAGAACTTAATACCTGGATTCTCACACTCATCTCCCCT
CAGCACCACACAGCTCTGACAACCACCCCCCAGGAGGTGACAGCAAAGAAAGGAAGGAGCAAC
TGCTCCCCCCAGAACCCTTCTATACTCCCCCGGGCTGCTCTCAGCAGCTGTGTGCCCAACT
CAGTGCTGTGAGGTGGAACGGGGAGATGGAGACACACACACACACCAGAGAGTTTA
TAAAATGTGCACACGGAAACACTTCCTAAAGAAGAAGTGGAGTTCTCAAGTCACTATGG
GACTACGGAACTCGGCATCTGAGCATCCCCTTCTGTATAATGGGGTTGGCTGGGCAGTCC
CTTCCAGCTCTGGCTTCCTCTCAGACCAGGATTATAGGATCCTGAAAAGCCAGTGCTTCC
CTCAGGGAGTGGAAGGTTCTGGGGGCTCAGCAGGGGGGCCCGGAGCCATTGTGGAGGGCTTT
CAAGGTGAGGACAATAGAAGAGCAGGTCTTGCGTGGCTTGGACAGGGAAAGAGGAGAAGG
AGTGGGGCATTTTGAGGCAGGGGGTGTAGCATGAGGCCAGGGCTTGAGGGGCCAGAAATTGGGGGT
GAACTCTGATGGGGGGCTGGGTAGAGAAGCTTCTACAGGGCCCCAGCTCAAGAGACCCCATC
TCTCCTCCTCCTCTCTGTCACTTGCCATGCTGGATCCGTGCATGATCACACTCCTGGACTCGC
CTCCTTGCCCCTGAGATCCCAGACCCCCCCGTATTTCAGCTGCCCCCCCTCAGCTCCTCCACTCACA
TATTTAATGCCAGACTCTTCATGTCTATCTACACCCTGCACTTTTGCACCCCAATCCCAACTC
CCCGCCATGTCCCCATCTCAGGTAATGTCAGCTCGGTCCCTTCCAGCTGCTCAAGCTAAA
ACCCATGTCACTTTGACTCTCCCTCTTGCCCCACTACATCCAAGCTGCTAGCACTGCTCCCT
GATCCAGCTTCAGATTAAGTCTCAGAATCTACCCCACTTCTCGCCTTCTCCACTGCCACCCA
GCCCATTCTGTGCCAGCATCATCACTTGCCCAGGACTGTTACAATAGCCTCCTCACTAGCCC

FIG. 6
(continued)

CCACTCACAGCAGCCAGATGAATCTTTTGAGTCCATGCCTAGTCACTGGGGCAAAATAGG
ACTCCGAGGAGAAAGTCCGAGACCAGCTCCGGCAAGATGAGCAAACACAGCCTGTGCAGG
GTGCAGGGAGGGCTAGAGGCCTGAGGCTTGAAACAGCTCTCAAGTGGAGGGGGAAACAAC
CATTGCCCTCATAGAGGACACATCCACACCAGGGCTGTGCTAGCGTGGGCAGGCAAGCCA
GGTGCTGGACCTCTGCACGTGGGGCATGTGTGGTATGTACATGTACCTGTGTTCTTGGT
GTGTGTGTGTGTGTGTGTGTCTAGAGCTGGGGTGCAACTATGGGGCCCCT
CGGGACATGTCCCAGCCAATGCCTGCTTTGACCAGAGGAGTGTCCACGTGGCTCAGGTGG
TCGAGTATCTCATACCGCCCTAGCACACGTGTGACTCCTTTCCCCTATTGTCTACGCAGG
CTGCCCTTGGACAAGGACCCGATGCCCAACCCCAGGCCTGGCAAGCCCTCGGCCCTTCC
TTGCCCCTTGGCCCATCCCACGAGCCTCGCCCAGCTGGAGGCTGCACCCAAAGCCTCA
GACCTGCTGGGGGCCCAGGGGGAACCTTCCAGCGCCGAGATCTTCGAGCCGGG
GCCCATGCCTCCTTCTTCTTCCTTGAACCCCATGCCACCATCGCAGCTGCAGGTGAGGCCC
TGGGCCCAGGATGGGGCAGGCAGGGTGGGGTACCTGGACCTACAGGTGCCGACCTTTACT
GTGGCACTGGGCGGGAGGGGGGCTGGCTGGGGCACAGGAAGTGGTTTCTGGGTCCCAGGC
AAGTCTGTGACTTATGCAGATGTTGCAGGGCCAAGAAAATCCCCACCTGCCAGGCCTCAG
AGATTGGAGGCTCTCCCCGACCTCCCAATCCCCTGTCTCAGGAGAGGAGGGCCGTATTG
TAGTCCCATGAGCATAGCTATGTGTCCCCCATCCCCATGTGACAAGAGAAGAGGACTGGGG
CCAAGTAGGTGAGGTGACAGGGCTGAGGCCAGCTCTGCAACTTATTAGCTGTTTGATCTT
TAAAAAGTTACTCGATCTCCATGAGCCTCAGTTTCCATACGTGTAAAAGGGGGATGATCA
TAGCATCTACCATGTGGGCTTGCAGTGCAGAGTATTTGAATTAGACACAGAACAGTGAGG
ATCAGGATGGCCTCTCACCCACCTGCCTTTCTGCCCAGCTGCCCACACTGCCCTAGTCA
TGGTGGCACCCTCCGGGCACGGCTGGGCCCCTTGCCCCACTTACAGGCACTCCTCCAGG
ACAGGCCACATTTCATGCACCAGGTATGGACGGTGAATGGGCAGGGAGGGAGGCAGGT
GGGAGAACTGTGGGGAGGGGCCCCGAGTCAGGCTGAACCACAGCCCACATGTGCCCCCCA
GCTCTCAACGGTGGATGCCCACGCCCGGACCCCCTGTGCTGCAGGTGCACCCCCTGGAGAG
CCCAGCCATGATCAGCCTCACACCCACCGCCCACTGGGTCTTCTCCCTCAAGGC
CCGGCCTGGCCTCCCACCTGGTAACACCTCAGCCCGTACCCCCATGGCTTCACAGAACCCC

FIG. 6
(continued)

FIG. 6
(continued)

CAAGTCCCCAGATCCTTGGCTGTGAGCAGTGTAGGCTATTCTGAATTGCAGTACTCTGGG
GGTCAAAGGTGTCAGGTCTCAGAGGCTTGGAAACTCCACCCTCCAAAAAACGTCAGGTGC
AGAACCTTAAAGATGCAGAATGTCAAAATCACACAAAACCACACAGAGCTTTACAAAGCTAGTC
AAAATGTCAGCACCTGCGAATGGCCCGTCTTTAAGCTTCTCTGCCAGAAGCCTGGGACTTT
GGGGACAGCAGAGCCCCCTGGGAGTCAGGGTTTTCGAGGCTCAGGAGGGTGGGAAGCTCA
AAATGAGAGGCCTTGTGGGCCAAGCTCCAGAGCCCCAGCCTCCATAGGTGCCCCT
GTCCCCACCCACAGGGATCAACGTGGCCAGCCTGGAATGGGTGTCCAGGGAGCCGGCACT
GCTCTGCACCTTCCCAAATCCAGTGCACCCAGGAAGGACAGGTCAGTGGACAGGCCTGG
GAAGGATCCTCGCCCTCCTATCCCCTCCCTGACACCCTCTGTCCCCCAGCACCCTTTC
GGCTGTGCCCCAGAGCTCCTACCCACTGCTGGCAAATGGTGTCTGCAAGTGGCCCGGATG
TGAGAAGGTCTTCGAAGAGGCAGGACTTCCTCAAGTGAGTGGCCCAGGCCTGTGCCAG
TCTACCGGCCCTGGCTTGTGGGGAGAGGGTCTAGGGGTGCAGATCTCAGCTCAGGCTTCATC
CCAATAAGTAGTAAGATGCAAGAGCTAAACTCTGAGACTGTGGGTTCAAATCCCCAGCTCCC
ATCCCCTGACTGGCTGTGTGACCTTGGACAAATTACTTAGCTTCTCTCGACCTCAGTTTTC
TCGTATGTATAATAAGAAACGTTTTGTGCTGTTATTTACCCCCCATTTTTACAGACAAGGAGA
CTGAGGCCAAAGAGGTTAAGTGTCTTGATCAAAATCTCCCCAGTTCCTGAGTGACAGAGCT
GGGATTTGAACCCATGTAGCCAGGCTCTAGAATCTGCACTCCCTGCTTTGCCCTGCTACTC
CTCATGCCAACATGCCAGCCTTATTCCACTGTTCCCAAAGTTCTAGCTCCTCTAGATGGC
TGCTGCTGCCTCCCCTCAGTCCCCCCCATATGTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTCTCTCTCTTTCTCTCTCTCACACACACACACACAAATA
CACACCACACTACACACAACCCTTAGCAGCAACCCCCACATGTCCTGTCCTCCTTGCAGCC
AGGCCTTTGCACAGGCCGTCCCTCCACCCAGAATGCTTTGCCCACTCTCACCTCCCTGGC
CCATTCTCAACTCACCACACTTCCAACATTATCTCCAGCACTCTGGCCAGGCTCAAGTGG
TGAGTTCAGGCCTGACATGCAGTAGGTGCTGAGGGGGCATGTGTTAAGGGAACGAGGGGTG
TGAGAGGGAGACTGAGAGGTAGAGAGGGGAGGGGAGCTGGGGGCTCAGAGGAGAGAACTCCCA
GAGGGTCTGGGCCCCTCCCCATTCAGAGCATTGAGCCCAGACCCAGGCCTGTCGTGGTCACCT
GCATGGAATCTTCTCCCTACTTAGGCACTGCCAGGCCGCCACCATCTTCTGGATGAAACGCG

FIG. 6
(continued)

CAGCGCACAATGTCTCCTCCAGAGAGATGGTACAGTCTCTGGAGCAGCAGGTAATGCC
AGGGGCGGTGGAGGGGTAAGGGATAGTGCGCAAAACCTTCTGTCCACCATGTGCCA
GAAACCAAGTTCACCTGGGACGAGGGCTGGTATAAAGGAAGGAGAGGCGGGCACTCC
CAGGGAAGACCGTAGCCTGGGCAAAGATGTGGCAGAGAAGGGCCAGGCTGAGGCCTCATG
TTTGTGCCATTTCACAGCTGGTGCTGGAGAAGGAGAAGCTGAGTGCCATGCCCCACC
TGGCTGGGAAAATGGCACTGACCAAGGCTTCATCTGTGGTGAGCGACCCCCAGGGACTGGTG
GTGGCAGACTCAGACTGCTGGGGGGCAGGGAGGAGGTCTGCACGGTGCAGGGAACCTAAC
CTCACATTCAGGTCCTGAGAGCTAGGGGCCCCTGTCCCCAGCTCCAAACATGCCCAGACC
TGGAAATCTGTCCCCTTCTACTTACAAACCCTCTGACCCCGGCTGGGCGTGGTGGCTCAC
GCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGTTCAA
GACCAGCCTCGGCCAAGATGGTGAAACCCTGTGTCTACTAAAAGTACAAAAAATTAGCCAG
GCGTGGTGGCAGGTGTCTGTAGTCCCAACTACTTGGGAGGCTGAGGCAGAGAATTGCTTG
AACCTGGGAGGCGGAGTTTGCAGTGAGCTGAGATCATACCACTGCACTCCAGCCTAGGCG
ACAGATCAAGACTCCGTCTCAAGAAAAAAAAAAAAAAAAACCCTCTGACTCTCTAAGATCC
CCAAACACTGTGATCCTGAGTTGTTAAAGCAAATGCAAATAGCCAGACTTGCCAGATGCA
GGCTGTGTGCCAGCAGGACCAGCTATGTAACCTGCAGGGCCCCTTGTTCAGATTTCAAGA
TGGCGACCACAGAGCATTAAACAAAGCACAGGGTGCCACATAACCACACATAGGTCACACGC
CCATGAAGCCAGCCCTGGAGGCCAGGCACTGTTTCTCTGAGCGCTTTGCTGGTGGTAATTTA
TTTCTCATGATGCGAATGTACAGATGAGGAATATTTGAGGCCAGGGGGGGTTTAGGTGACTT
TCCCCAAGGTCACAGTTGGGTGGTAAAGAGCCCCTATTTCAACCCCCAGTTCATGGTCCCAGCA
TCAGTGGCCACATACGACATCCGCACCTGTGCTCTAATAAATACGGCTCATGCTGTTTTG
TGGGATTCCACCTCAGACTGGAATTTAGAAGAGGGCGTCCTTGCTTTGAAAAACACTGAT
TTAAAAATAAAGTGGAGCCAGGCGCAGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAAGCGGGTGGATCACATGAGGTCAGGAGTTCGAGACCAGCCTGGCTAACATGGTGA
AACCCCATCTCTACCAAAAATACAAAAATTAGCTGGGCGTAGTGACGGGCACCCTGTAGTC
CCAGCTACTCGGGAGGCTGAGGCAGAAGACTTGCTTGAATCTGGGAGGTGGGGGTTGCAG
TGAGCCGAGATTGCACCACTGCACTCCAGCCTGGGCGACAGAGCAAGACTCAGTATCAAA

AAAAATAAAAAACATAAAATAGAAAGTAAAAAGTGGGAAGTTTAAGCCTCTGGGTCACCA
GCCTCTCCCCCCTCACCCCAGGCATCATCCGACAAGGGCTCCTGCCTGCATCGTAGCTGCCTGG
CAGCCAAGGCCCTGTCGTCCCAGCCTGGTCTGGCCCCCGGGAGGCCCCCTGACAGCCTGTT
TGCTGTCCGCAGGCACCTGTGGGGTAGCCATGCAAACAGCACATTCCCAGGTAAGAATGG
TCCTTGCACTACACGGTGCCCCCCAAGCTCCTAATCCTGACAGGCTCTGGGTGGAGGGTGC
AAAGGAGCTCCATGCTGCCCCTTCCCACCACCACCACTGCAGCTGCCGCCACTCAGC
CTTTGGGAAAATGCATCCGCTCACAAAAGCTTCCTTTCGGGATGTCCGTGGCCTGAAAGC
CCCCCATATGGTCTCGAGTGTCGGGCCCCTAGCCCCTGACTCCCTTGGGGATTGGGCCAT
GCCTCACCCACTCTGGACTCCAGCTACTATATTCGGCCATCAGAAGGGAGGGACCCTGCT
AAGTAATTCCAGGAGCACCTCCTTTCCTCCCCTGACCAAGGAAAATCGGGGTGGATTCGC
CCGAGCTCACCTATCCACTGCTCTCCACCAGGCCTGGCCCTGTGGGCTTAGCAGGGATCAG
AGACCTTGACTGTCATCCTGGCTCTGCCATTTAACCTCTTTGCATCCTTTGGTGTGCAAGT
TACTCCGCTTCTTTTTCAACCTCGGGGAGAACTATTTTGGCAGAAGTGGTGCAAAGAATAA
ATGATACAACTTATGTCAGGTGCTCAGCAAACAGTACCTGTGCCCCGTGGACACGGGTGTT
GACGGTGAGATCTCAGGCCTGTAGACTCACCTTGTAGGGGGAGGGGACAGGGAGCTAGCT
AGGAGGTCCTGCATGGGGCTTGATTCATCCCCACCCCTCTGACAGAGTTCCTCCACAACAT
CGACTACTTCAAGTTCCACAACATGCCGACCCCCTTTCACCTACGCCACGCTCATCCGGCTG
CGTAAGCAGGGCAGCTCGGCCCCCAAGGAGGAGGAAGACAAAGATGGGGGTGGGGGACCTGC
CTCCCAAACTCCTGTCTCCCTCTGAGTGCCCTCAGAGTGGGTTCCTCCATTCCCAAGCCCC
CAGCCCCCAAGGATCCCCCAAGCCGTGCCTCAAATGTGACCCCCTCATGCTGGCTTCACCCCCA
AACCCTGTCCGCAAATCCAAACCTAAACCACCATCCAGGCCAGAGCATGCCAAATTCTGAC
CCTAAACCCTACCCCCTTTCCAGATGTCCACCTCAGCCCCCATACCTAACCCCTCTCCTGGACC
CATAAAATAGCCTAAAGCTAACCCCCATCTCTGCACCTTGCCCCTAAACGTACCCCCAGCTCT
TACTCTAACTCCCTTCCCCCAGCCTTTATGCCAACCCCAACCCCCATCTCCCTTCCCCTGGCCCCCA
ACGTACCCCCAGTATGTCAATACACCCCCCCAACTGGGGCACCATTCCCCAACCCTTTCCCTTGTAA
CACCCATTTGATCCTTAACTTCATCACCCCCATCCCCAACTCCCTTCATCTCAGCCCTCCCCCAA
TGCCTTCTCAGAACCTTCATCCTAGCGCCACACCCTAACCCCCAAACCCTGAACCCTCACCCCCT

FIG. 6
(continued)

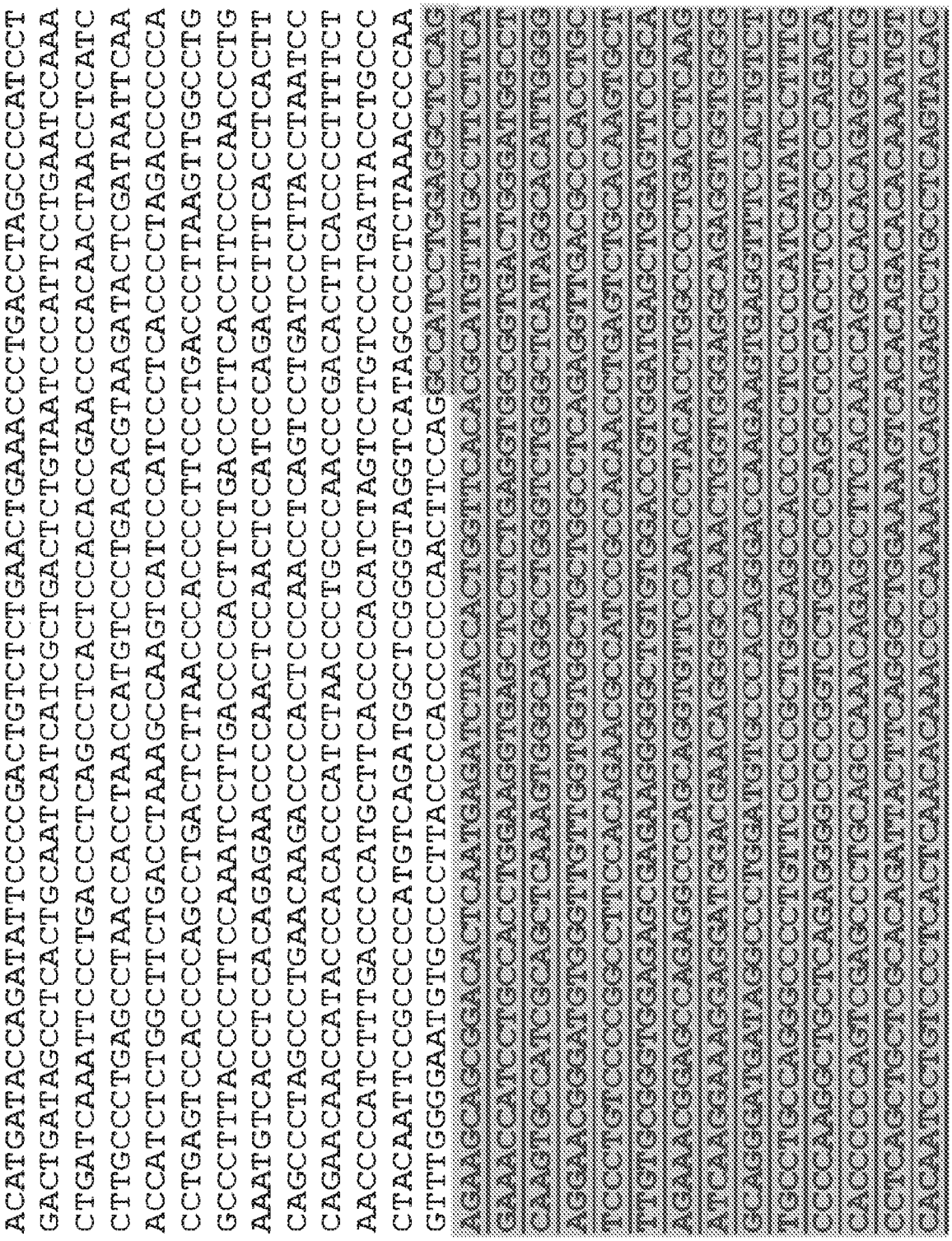


FIG. 6
(continued)

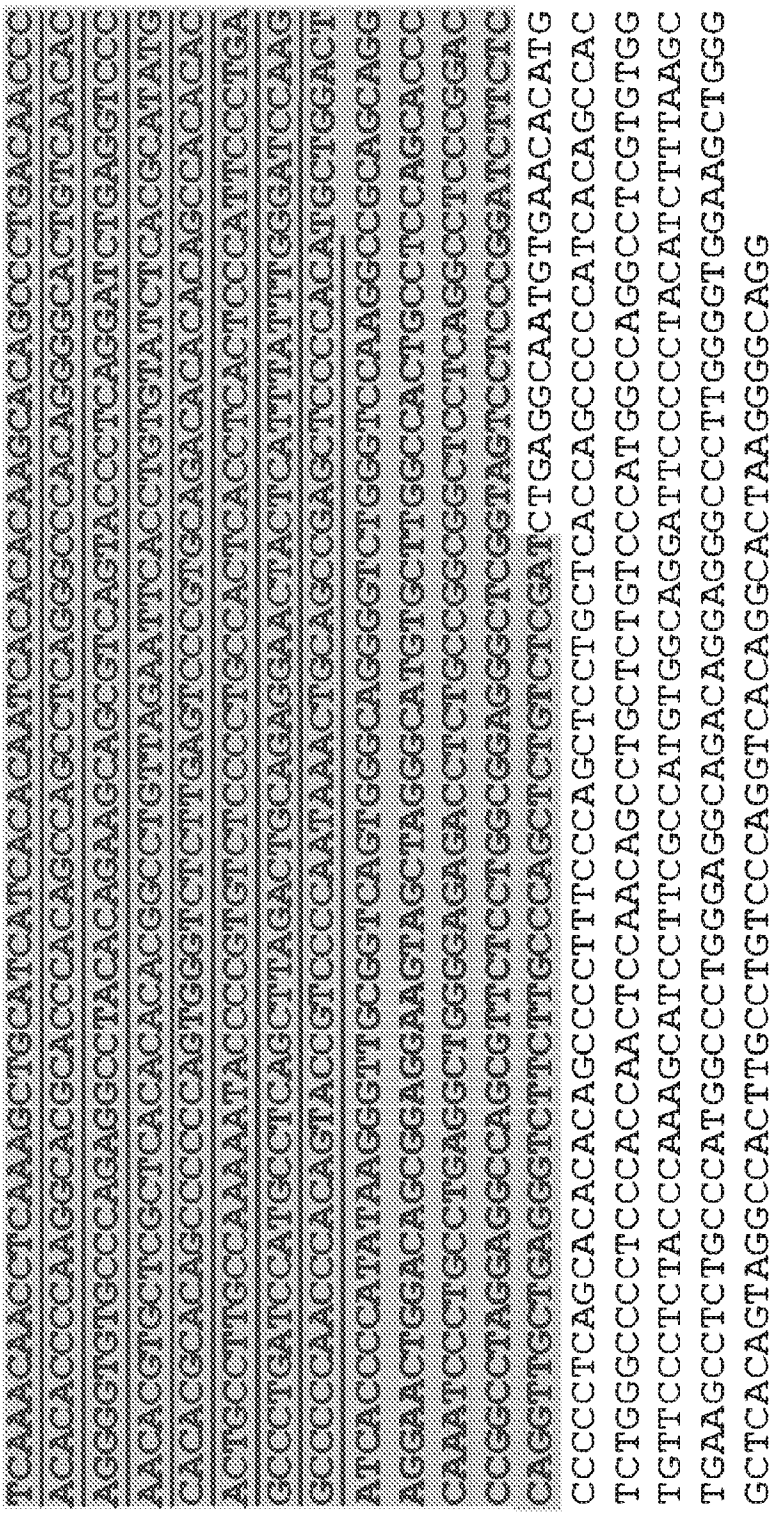


FIG. 6
(continued)

MPNPRPGKPSAPSLALGPSPGASPSWRAAPKASDLLGARGPGGTFQGRDLRGGAHASSSS
LNPMPPSQLQLPTLPLVMVAPSGARLGPLPHLQALLQDRPHFMHQLSTVDAHARTPVLQV
HPLESPAMISLTPPTTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPRKD
STLSAVPQSSYPLLANGVCKWPGCEKVFEEPEDFLKHCQADHLLDEKGRAQCLLQREMVQ
SLEQQLVILEKEKLSAMQAHLAGKMALTKASSVASSDKGSCCIVAAGSQGPVVPAWSGPRE
APDSLFAVRRHLWGSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILEAPEKQRTL
NEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFVRVESEKGAVWTVDELEFRKKRSQR
PSRCSNPTPGP

SEQ ID NO: 57

FIG. 7

CATGACCCTGCCACCCCATGGCCTGCTGGTGGCAGGCGTGGCCCGCCTCCTGAGAGTT
GGCCCTCCCTTGTGCCACTGCCAGGGAAAGGCCCTTGATGTTCCAGACAATAATAAA
TGCCCCCTGTGACTTAGCCTTGGTGTCAGTCTCTTGCGGACCTGACAACCCCCCATCTCTCC
TTCCCTGATTCCCTCTGCCTTTCCAGGCCCCCATCCCCCCTGAACAGCTCCTCCCTATGGTC
CTGGCTGGGCCTAACCCTGCCCCCAGGGCCTAACCCTACCTGAGGCTCCTCCCCCTTCCCCCC
GGGGCAGGTTGAGAGGCTGGAGTGGGTCCCTCAGCGCCCCTGGGTGGGGTGGGCCTGCACAG
GGGGTACCTCCTTCTCTGAGGAACTGGGCTGTTAGGGATTTTCCTTAGGCCCTTTGGTTT
CCGCCTACGGAGAGGTTTCCCCCCATTGGTTGCTCTTCCTCAGCCAGGGTTACTTCCTGGT
CTGTTCCCCTACCCAATACCCCGCCGCTCTGTCAGCTTGAGCTCCAGGTGGAGCTCCAGG
TGGCTCCTCCTCTCCCGGGGGAAGGCGGCCCTGGACCAGCAGGCGGGCCTGCTGTACTCC
CGCTTTGGGGCTGCAGGGAAGCTGGCCGCTGTGGGCGGTCTCGGGCCAGCCCCGCCCCAC
CTGTCCTTTCCCTGGAGACTATTAGTCCAGGGTTTGTCCCTGCAGTGCCATTGGCCCTGGC
AGGCAGGATCGAGGAGGAAGTGGCTGATTACTGAGCGGTTCTTCCTCACCTGGCTTGGGC
CACTGTGCACAGCTGTGCCGCTGGCTCAGCCCCCGCCCCCCTGCGGCCCTCTGCCGTGGCTT
CCCCCTCCCTACAGAGATGCTGTCCCGTGGGTAAGTCCCGGGGCACCATCGGGGTCCCA
GTCTCCTGTTAGTTTTGGAGGGAGGGAGGGCTTTGTTGATGCTCACTCCGACGTGTGTGA
ACGTGAGTGCGATCTGCCGCTGCCCCTGCGCCTGTTTCCGGTCCCCTATGAACTTCCCCCTTC
CCGCAAGGTGTGAGGACCCCCCGGCTCACTCATGCTCCTCTGCCCCCCTCTTTAACATTTTC
CCCTGGACAAGTGTGTATCTGTTCTCTCCATTGCATTTCTACTTCCAGCCTCTGGGCTCC
TGCTTCTGCCTCCTGCTTAGGACCTGTCCCCCTGGGTAGCTCACACAACCTCAAACATAG
CAGTCAGAGGCCACCCGCGAAGGCCCTCCCACGTCCAGCCAACTTCTCCGCACTTCCCAA
CATCAGACTTTGGTCCCATCTTCTTTGTTTCCTTTCACTTCCCTTTCCCCTGCATCATTC
ATTCAACAGGTACGTGTTGAGCATCTATTATGCACCAGGTGCTGTTTAAGATGCTGGTAA
TACTGGAGTGAACAAGACAGACATGGTCTCTGCTCTCACGGAGCTTACATTCCAGTGGGA
GGTTACAGACCGAACAAATAACCCAATAAATTGGATCATTGCAGATTCTCAGAAGTATTA

FIG. 8

FIG. 8
(continued)

CGCAGAAAATAGACAGCCTTGGCCGGGTGTAGTGGTTCACACCTGTGATCCCAGCACTGT
GGGAGGCTGAGGCGAGAGGATTGCTTGAGCCCAGGAGTTTGAGACCAGCCTGGCCAATAT
AGTGAGACCCTGTCTCTACAAAAAATAAGAAATTAGCTGGGTGTGGTGGCACACGTCCTG
TGGTTCCAGCTATGGAGAGGCTAAGGTGAGAGGCTTGCTTGAGCCTGGGAGGTCAAGGCT
GCAGTCAGCGATGATTGCACCACTGCACCAGCCTGGGCGACAGAGTGAGACCTTGTCT
CAAAAAAAAAAAAAAAAAGAAAATGAACCAGCTTCATATGCTAGCAAGTGACTGGGTG
TGCAGGTGACATTACTAGCTGGAGGGATCAGGGAGGCCTTCCCGAGGAGGTGACATTTGA
GCTGAGACCCGGATGAGGAGGAAGAGGAGCTGGCCATGTGACGTAGTGATCAAGAGTCAA
GCATCTCTGGGCAGAGGAGATGGTGAGCACAAAGCCCTAATGTGGGAACAAACAAAAAAA
GGACAGTGTGCCCGTGGCAGAGGACCCTAGTGGAGCGGAGGCAGGGCCACAGCAGGTTAG
ACCATGTTGGAGCTAGGATGTTGAAAGTGAAAACCTGACGAGATGAGGTGGCGCACGTCT
GTGATCCCAGCACTTTGGGAGGCCGAAGGGGGAAGATTGCTTGAGCTCAGGAGTTTAAAA
CCAGCCTGGGCAACATAGAGAGACCCCCATCTCTCTATTAAAAAAAATACTGGGTATGATGG
CCCAAGCATGTGGTAGTCCTAGCAGTTTGGGAGGCTGAGGTGGGAGGATCACTTGAGCCC
AAGAGTTCAAGACCACCCTGGGCAACATAGGGAGAGACCTCATCTCTACTACGACTACGA
CTACTACTACTACTAAATAAATAGCTGGATGTAGTGGCATGCACCTGTGGTCTCAGTTACT
TGGAAGGCTGAGGCAGGAGGATCACCTGAGCCAAGGAGGTCGACGGCTGCAGTGAGTTGGA
TTGTGACACTGCACTTCAGCCTGGGTGATAAAGCAAGATTCTGTGTCAAAAAAAAAAAAA
AAAGAGAGGGAAGGAAGGGAAGGAAAGAAAGAGAAAGAAGGAAGAAGGAAAAAA
AGGAAAGAGCGAGAAAGAAAGGAAGGAAGGAAAGAAAGAAGGAAGAAGGAAG
AAAAAGAAAAAGTGACACCCCAGTCGAAAGAAGGAAAGAAAAAGTGACAAC
CGGTCGAAAGAAAAAAGTGACAACCGGCTGGGCATGGTGGCTCAAGCCTGTAAT
CCCAGCACTTTGGGAGGCCGAGGCAGGTGGATCACGAGGTCAGGAGTTCAAGACCAGCCT
GGCCAACATGGTGAAACCCCTGTCTCAACTAAAGATACAAAAAAAAAATTAGGCTGGCACA
GTGGTGCGCACCTGTGAGTCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATTGCTTGAAC

FIG. 8
(continued)

CCAGGAGGCGGAGGTTGCAGTGAGCCGAGATTGCGTCACTGCACTCCAGCCTGAGTGCAG
CGGGAGAGACTCCATCTCAAAAAAAAAAAAAAGAAAAAGAAAAAGTGACAACCTGCTTA
CAGAGTACTGGCGAGTTTGTGGGTGGGCTCCCTAGCCCTGCTGATTCTTGCTTCTCTCA
CACTCATGTCTGCCCCTGCCCCAGTGCACATCTTGTCACTGTCGGCCCCACCGATGGGGT
TCCTACTGAGTCTTCTGGTCCCTGATCCCGTCTGTGGTCATTTTCCTGCCAGGTAGCTTG
GCCAGGCCTCCCCCTGGTGCAGATTTCATCCTTGGTTTCTCAGCCTGGCCTTGAATGACCC
TCTACAGCAGGGTCCCCACCTCTCAGAACAACTTTGCTCCAGCCACATGGCTTGCTCACG
GCCAGGCACTGCCCATGTGGACTCTGTGCGTGCCACCTCTTTGCCCTGACCCATGTTGCC
TCTGGGGGAGCACTTCTTCCTCCACCTTCCATCATGGGCTGTGGCAGTGCCCATCCCATC
TGCCCCCGACGCTGTCTGCTGCAGTATGGTTGTTGGGGGAAAGGGCACCAGGCTCCGGCG
TCTGACAGCCGTGTTTTACCCACCTTCCTACTCACTAGCTTGTGACCTTGGGCAATTACT
TAACATCTCTGAGTCTTAGTTTCTGTTTCTAAAATTGGGTGAATAACACCTACTAAGTAG
GGTTGGCCTGAGGATTAATAGTATAATGTAAAAGCTGGCAGCACTGAAACCCCTGCCACTT
ACCAGCTTTTTCACATCAGTATTTGGGAAATATTGTTAAGCTCATTTGTCAGGCGGGGATT
CTGAGGCTCAGAGCAGTTCCAGAACTTTCTACAGATTATTTTGCCTTGTTTGCGCTTCCA
GACTGCCTATCTTCTTGTATCACCATTGATCTTGATGGTTTTTAATTTTTTTTTT
TTTTGAGACGGAGTTTCACTCTGTTGCCCAGGCTGGAGTGCGGTGGCATGATCTCGGCTC
ACTGCAACCTCCACCTCCTGAGAAGCTGGGATTACAGGCATGTGCCACCACACCCGGCTA
ATCTTTGTATTTTTAGTAGAGTTGGGGTTTCACTATGTTGGCCAGGCTGGTCTCGATCTC
CTGACCTCGTGATCCGCCAGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA
CTGCGCCCGGCCAACTTCACGTTTATACACACCCATGCAAACAGCATCCAGATAGAGACA
AAGAGCCTTCCCTGTACCCTAAAAGTTTCCCAGAAATTGTTCCCAGTTAGCATATTTATT
TTTATAAAGGTAATGCATGCCCATCATATAACATTCAAAAAGGTATGTAGAGAACCAAGT
GTCTCCCCCCAGCCCTGTCCTCCAGCCACCCAGTTTCCCTCCCTAGGGGAAGCCCACCAATA
TGTGTTTCTTATGTATCCCCTGTTGAGCTGCTTTTCCTCGTTTTGGTTTGGCGGTGTTGA

FIG. 8
(continued)

TGTTTGTATTTGGAATTACAGGTAGGCAGCATCATATACCTTAGTGTTTAGGGCCTCTAA
GATCAACCAGCCCTGAGAAAATCAGCCATGGTGAGGACCTTGTCCCCCCAGCCCCCCAGGA
GATAGGCCCCCCTGGTGGGGCAGTGCTGGGGGCAGGGCCTAGGGACAAGAATTAGAAA
GGACCCATGTTGACAGGGCTGCTCAGGGTCATGTTGTCCATCCCCTCTGCCACAGTGGCAT
GGACAAACTGCATATGTTGGTTAGAGGAGGGCACCCTTCTCTCTTGCAAGCATTGGCAAG
GTCTTAACTATTAGTCTCCTGCTCCCATGGCAGCCCCTTTGGACAAGGAGGCTCTTAATC
TCTGTTCTTTGAAGCCCTGAGGGCTGGTGTATAGGAGTTCAAAGCACTGGCTTTGGAACC
GGACTGTCTGGGTTTGAATCCTGGCACTGCAGCTGACTCACTGATGGACTCAGGCAATGC
CTTAAACTCCCTGAGCCTCAGGTTCCTTGTCTGTAAAATGATAAAGATAGCCCCCTGTTTC
ATAGGGCTGTGGTGAGAAACCAATCAGACAAGGCATGTGAACGCCATTATAGCACAGCGC
CCGGCATCCAGCAGGACTCACTCGATGACAGTTGTCACCGCCATCATTGTTATTAGCGTG
GGCCAGGGAGGGCTGCGTAAAAGCAGCTGGTGGAGGGAGAGATGCCGTGGGACCGTC
TGGGTTCGCATGCGTGAAGTATTATCTGGGCCTGGAGTGTGCAAGGCACACATGTGTCCT
TACTGCATGTGTTGTCACATATGTGCAATGCCATGCTCCTGAGCCTTTGATTGCAGACGT
GTGGGAAGTGGGCCCCGTCCCCACCCCCAGTGCCACCCTGCTCTGCTTCTCTTCCTTGC
TGTGCTCTAAAACGAGAAGTACAAGTGAGTTCCCCCAAGGGTCCGGCCGCCTCTTCCT
GTCCCCGCCCTGCCCGGCTGCCCAGGCCAGTGGAGTGGCAGCCCCAGAACTGGGACCACC
GGGGTGGTGAGGCGCCCCGGCACTGGGAGCCTGCATCTGAGGCTTAGTCCCTGAGCTCTC
TGCCTGCCCAGACTAGCTGCACCTCCTCATTCCCTGCGCCCCCTTCCTCTCCTCCGGAAGCCC
CCAGGATGGTGAGGTAAGGGCCTGCCACCCACGGTAGACAGGAGGCAAGGGTGCCTGGTG
CCCACCGGACCCCCTCCTCACTGCCCTGGGCCGCCCAGGTGGTTTCACCGAGACCCTC
AGTGGGCTGGATGCAGAGACCCTGCTCAAGGGCCGAGGTGTCCACCGGTAGCTTCCTGGCT
CGGCCCAGTCGCAAGAACCAGGGTGACTTCTCGCTCTCCGTCAGGTACGGTGGGCCCCCCG
CAACCCCGGGCATTTTGGCCACTCTCTTGTGCCATCCAGGCCCTGAACCACTCATTCCTG
GTTCCCGGTGCAGTGCTGACTCCCCGTCTGTTCCCTTGCCCCCAACCCCACACTCCCC
ATCCCTGTCTGTGCCCACCCATGCCCATGTGTGCCCCCACCCAGGACCTCAGCCGATCCC

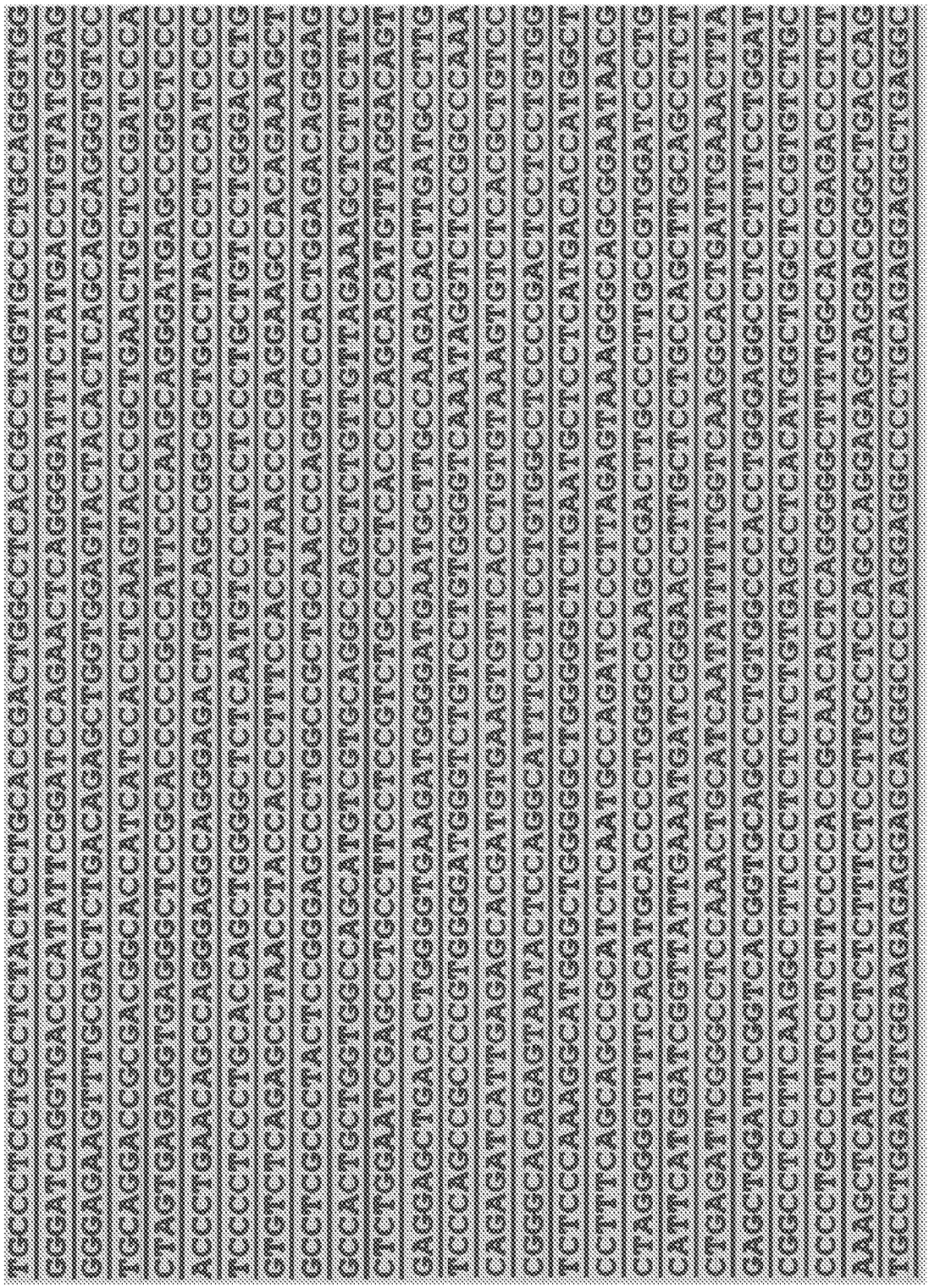


FIG. 8
(continued)

CTGGGTTCAAGGAGAAGAGAGAAGGAAGGGAGGGCAGTGCCCGGGGCGGGAG
GTTAAGACCACGGAAGGCCGCACTGGAGCGGCCCTTTTGCGGTGACCCCTCCCAGGAGCCAGTG
TCACCCCTGAGCCTCGCAGTGTGTGAGAGGCCTCTTTCTCCCACGTTCTGCTGTCCTCT
GCCTTGTCTGTGCGCCTCCTCCTCTGCGAGAATTTGCATCTGTCCCTCGGTGGCTCTGCG
CTTCCTGTGGTCAGCCTGACATTTGCATGGAGACTTCCTCATCCTGGGGCCTGAGGGAAG
GGGCTCAGCCCCCTCCCCGCTACCTGGGGTCCTAGCCTGTCCCCAGGCGGTGGGCTGAAG
TAGCCCAGTGGGGTTAGGAGGCTCTGGGGGTCTCTCGGCTGGAGTCACCTCCGGGCAGGG
GTGAGATGGGTTGGGACAGACTGGTCCTCCCCTCCTTCCCCCCATCCCTGCGGTTGGAAA
ATTTGCCCCGCCCTCCCCTCGTCCCTGGGCTGAGGAAACCTCACACAACCTCACTTCTCACTC
TCTCCCCCAGAAGGAGTTTTGTGTTTTTCCATCACGTGGTTTCCTGTGGGGCTGGGCTTTT
GTGGGGCTACAGTTTCCTCCTGGGAAAGGGGTGTGCTTCGGGGAAAGGGCTTAGTTCTGC
TTTCTGCCCTGACAGCCCCCTTCAAATCCGTTTGAACCCTGGGCTCCCCCTTCAGTGACATC
ATCCAGGGCACCCCCAGAACCCCCTACACCACTCTTTCCCCAGTGGGGTTGTCTTCCCCCGC
CTCCCCTGGCGGAGCGCACCCCCATCCCGCCCTTCCCTTGTGACTTGAGTCTGTGTGTCCATCTC
CCACCACTCCCCTGTGGTGTGGCCTCGGTCTGCGTTTCTCTTTGCCTCTGGTCTCTGCTGG
GGCACAGTCCCCATCCCTTCACGGAGATTCATCCTTAGCTTCTCTCCTCCAAATATTTTGAA
TATTGCCAGCCCTTTCTGCCTTTCAGAGGTGGGCTCTGGGTTCGAAGCCCCGGTTAGAACTC
TGGAGGCTAGGATGGCTTGAACCTGGGAGGTCGAGGCTGCAGAGAGCTGTGTAACCGCGCCA
CTGCACTCCAGCCTGGGCAACAGAGCTCTGGAAGCTTGCCCTAGAGTCAGTCAAGGGCCCC
TAGGCCAGTGAGTAACAGCTCAGCGTCAGTTTCCTCATCTATAAAATGGGGGTAATATCA
TACCTAGCTCTCAGCATGTTTGTGAGAGACCTAAATGAGGTGGTGGATTTGGAAGCATGT
AGCGCAGTGCCTGGCACACAGTAGGTGCTTGATTTCCGGCCCCCTCTCTGTGAATGTCTCT
GCTCAGCGCCTTCCCCCTGTGGCCTGGGTCTTACCTTCCCTGACGCTGCCTTCTCTAGGTG
GTACCATGGCCACATGTCTCGCGGGCAGGCAGACGCCTGCTGCAGGCCAAGGCGGAGCC
CTGGACGTTTCTTGTGCGTGAGAGCCTCAGCCAGCCTGGAGACTTCGTGCTTTCTGTGCT

FIG. 8
(continued)

FIG. 8
(continued)

CAGTGACCAGCCCAAGGCTCTGGCCCAGGCTCCCCCGCTCAGGGTCACCCACATCAAGGTCAT
GTCCGAGGTAAGGCAGCCAGGCGGCGGGGAGGCCTCTGCTGAGGCTCCTGTCTGTGACCA
CAGTGTGGGTGGCAGGGAGGGTCTGCCTGGGCTTGAATTCAAGGCTGGGGACCCAGGGAG
GGAGACTCAAGTCCTGTGAATGGCCTAATTTGGCTCCCCCCCAGGGTGGACGCTACACAGT
GGGTGGTTTGCGACAGCCTCACGGACCTGGTGGAGCATTTCAAGAAGACGGG
GATTGAGGAGGCCTCAGGGCGCCTTTGTCTACCTGCGGCAGGTCAGGGGTGGGCCCAGCTG
CCTCCCCACTTCCCCCTGAGCTGTCCCCCAGATGTGAGGCTTCTCTGGGATCTCTGAGTTGCTG
ACTTCTCGCTCTTCCCCACCCCAGCCGTACTATGCCACCGAGGGTGAATGCGGCCTGACATT
GAGAACCGAGTGTTGGAACTGAACAAGCAGGAGTCCGAGGATACAGCCCAAGGCCTGGC
TTCTCGGAGGAGTTTCAGGTGCATGGTGGGGACCGGCAGGGCTGGGGGCAGCTGAGGTGGT
GGCTGCCCCATTCACCCAGGATACCGCCCCCTGCCCCAGCTGCCTCCCCCTCATCTCACAGG
TCTCCACCCTCCCACGCCAGGAGGGGGCCATCTCCCCCCACACCCCCCCACAGAGCCTCCCCCTT
CTCCAAAAGGCCTCTACTCCCTCCCAGAAGTGCCTCCCCCACCACCAGCAGGCAGGTTGCCC
CCTGCTCCCCAACCTCCTTGTGAACTCCCCTCACTCCCCTCCATACAGATGATCCCCCCACCCCC
TGCTGCCCACAGTCCCCCCGCAAGCCTCATGGCTTCTCTGAGACCAGAATGGCCTGTTAGCTC
AGGAGGGTCTGACCCAGGTGTGGTCAGTCCCTGGCTAACCCAGACCATCTCGCCTCCTCCTCT
CCGCCCACTCCCCAGTTGACCACAGCCGAGTGATCCTGCAGGGACGGGACAGTAACATCCCC
CGGGTCCGACTACATCAATGCCAACTACATCAAGGTCAGCAGTGTGGGCCACGTGGGAGGG
AGAGGCTGGGCCCCTGGGAATTCCCCTGTCTGGTGGGGGGACCCCTAGATCCAGAGGACAGCTG
GGCAAAGCCGAAGCTGGCTTCTTGCATGGGGTGAGGGGTGGCAGTGGTTCAGGGCCTGTGCT
GGGCCAAGGGGGCTCACTGTCTTGGGGTGCGTCTCTCCACGCTTGCGTCCAGAACCCAGCTG

FIG. 8
(continued)

CTAGGCCCTGATGAGAACGCTAAGACCTACATCGCCAGCCAGGGTTGTCTGGAGGCCACG
GTCAATGACTTCTGGCAGATGGCGTGGCAGGAGAACAGCCGTGTCATCGTCATGACCACC
CGAGAGGTGGAGAAAGGCCGGGTAGGGCGCCCCCCCTTCCCCGCATCCGCCCCCGTGCTT
GTGGTCATGCCATTAAGTCGAAGAGAGCAGTCAGATGCCAGGGCAGAAAGGGATCTCAGGGG
TGAGGGGTCCGGCCCTTGTTGGGAAACTGAGGGCTAGTGACAAAGTCTCGACTACACAACG
TGACCCCCCAGATCCCCATGCATCCCCTGGGCTCTTCTGAGCTCCAGACCCAGGTTCCAG
GCTGTCCTCCTTCCTCCTACCCCCTGCCCCCACCCTGTCTGCATCCAGGCCCCCTCCTGTCCTC
CCTGCCCCCATAGATCTCTCTGGAGTCTGCCCCCTTACCCCTGCAGGCTCCCCCCTACACAGCA
CCCTCTGTGCTGCCATTGAAGTGATCCCCATCCGTGACACAAACTGGGTCAAGTTCCTTCC
TTTCTGAAATCTCTTCCATGGCTCCTGGTCACCTTTGGGATAAAGTCGCACTCTAAGGCC
TGGCATTCAAGGTCTGGTGGCTTCCCTCTGACCCCGCACGCTTCTCTTGAAGGCTCACCGC
TAGGAAATGACTCTCTCCACACACTATCTCTGCCTGGCAGATGCCTCGTTTTTGAAGACACA
GCCGGAGCGCTGCCTCCTCTGTGAATCCAGGTCTTGTTTCCTCCAGGACCTAGAGGGAGA
ATTACGTCTTTTCCCAGCCACGCTCCTCAGCGCGGTGTCTCCCCCGGTCACCTGTCTCTGT
GAGCTCCTCGAGGCACAGGGGCACAGACTGGGTGTTATTTTGTGTCTGTGAAGCTGTGTGG
TTTGCACAGCTTCGGGGACAATGCCTGCCCTGGCAACGTTTGTTGAATGACAAACGGATG
TACCGGTGAAGTGGCTGGCCAGGCCTCACCACCTGTTGGTGGTTGATCTGAGACGAGAGC
CCAGGTCTCCTGCCTCTCTGCCAGCCCATCCGTCCATCCAACACAAATGTTTGGGCCGGTGC
CAGGCACTCAGAACATAGAGCAGGACCTGGGATGGGCCACAGTGCCCTGCTCTGTGCCTC
ATCCCCCACCCGACCCCTCCCCTTTCCAGAACAAATGCGTCCCATACTGGCCCGAGGTGGGCA
TGCAGCGCTGCTTATGGCCCCCTACTCTGTGACCAACTGCCGGGAGCATGACACAACCGAAT
ACAAACTCCGTACCTTACAGGTCTCCCCGCTGGACAATGTGAGTGGCCCCCACGCCCCTGC
CCCATTCCGGGAGTCCCTCCCTGGACTTGTTCTCCTCTCTGGTCGGTAGGGTGAGATGG
ATGAGGTGTTCCGAGAGAGGAGGGGCACTGACCCTATGTCCTCGGCTTAGCCGACCCTG

FIG. 8
(continued)

ATTCGGGAGATCTGGCATTACCAGTACCTGAGCTGGCCCGACCATGGGGTCCCCAGTGAG
CCTGGGGGGTGTCCTCAGCTTCCTGGACCAGATCAACCAGCGGCAGGAAAGTCTGCCTCAC
GCAGGGCCCATCATCGTGCACTGCAGGTGAGGATGATAATCCTGATGGTAGTAGTGACAG
CTGAGAAGTAAATACTGCTAAGTGCCATGAGCTGTTATAAGCAATATAAACGTTAGCTCG
CACATTGAGTGCCCTCCGCTCACCCCCGGCTTCTCCTGGGTCCCCTCATGGCTCCAGAAC
CCTGGGTGGATCGTGGCTGGAACCAGCCCCCACTTTGGCCCTCTGCCTGTGGGTATCTTCC
TCAGAGCCCTCTCCGGATGTACCATCTCGCCCCAACCCTGCCAAATACAGAGGAGGAGCCC
GGGACCCAGTTGCTGGCCAGGCCCAAGCTAGTCAGGGCAAGGCCGGGCAGGCACCCACAG
TAGGCCTGTGTCCCGGCTGCTCCGCTTTCTCTCGAGGTCCCATTCTGTTGGTTTCTTCTC
CCAGGAACATCTATGAGGCATGTGCTCCCCCATTCCTCCTCTTTTTCCATCGGTAGCCGCA
GGGCTTCGGCTTCTTCCTGACTCTGCCCTCTCTCCCAGCTTCCCCAGGCAGTGCCCCATC
CTGGCCCCCCAGGGCTGTGTGGGGATGGGTGATGCTTCTTTGGGGCTGCACATAACTCCTC
TGTCTATCTACCCGCATGTTTGTGATCAGGAGACCTCTGGTAAGGTGCAGAGGTGGGGGC
TGCAAGGAGGAGCAGGGGTTCCACAGGTGAGCCCACTGAGCTGGCCTGGGTGGGTGGAT
GAGAGGCAGTGGGTGCAGGGCCCCTCCGCTTACCAGCTGTGTGGTCTTGGACAAATTACT
TAACTTTTCTAACCCTCAGCTTCCTCATCTGTAAAATCAGGATCTCAGGGTTGTCGTGAG
AACTCAATGAGACCCCTATCGTTGTGGCTGGAATTCCGTCAGCCCCTCAAAAACTGGGCGCT
GTTACTAGTTTAGTAACTCACATCAGGCAGAGAATAGGGGGAATGGGAACCTGCCTTGCCC
CGGTCCCCTTCCCCACTCCCCTCCGTGGACCCCCAGGCCCTGCGACGGCCTCTGGCTTCCTCCTC
TTCCCCCCAGCAGCTGTTTGTCCTGGGACAGGGCAAGTCGGCTGAATCTAGAGGGTGCCCCCC
GATGGGCTGTGTCCGGGGACGCGGCTCTGTCCTGTGCTCTCTCAGGGACAGGCCCATCCCCCG
AGAGCTACCCCTCCTGCTCACCCCGCCACACACACATTCACACACTTCTTGAAAGCCCCCATG
GCCTTTATTTAGACGTTACAGGAAGGAAGTGGGTGTGGGGGGTTATTTTTTGACAATCTGG

FIG. 8
(continued)

FIG. 8
(continued)

GTGCTGTCTCCTGACCTGCCAACTGCCTGTACTTGCCCCCCTGCACCCGGCTGCAGAC
ACAAGGAGGATGTGTATGAGAACCTGCCACACTAAGAACAAGCAGGAGGAGAAAGTGAAGA
AGCAGCGGTCAGCAGACAAGGAAGAGCAAGGGTTCCCTCAAGAGGAGTGAGCGGTGCC
TGTCCTCAGGTGGCCATGGTACAGCTCTTCTGCCTGGGTGTCCTCCCTGCCCCTGT
GTCCTTGGCTCCACTGCCTTCCCTGGGTGGATGGGGTGGCCGCAGCCTCATTCTGTGCTT
CCCAGCTGCCCCAGACCCTCTTGTTCCACCTCCAGGTTCCAGCTACCCTCTCACTCCCTC
ACTCCCTTCTCTTGGCAGCCTCAGCCCTGACCCTGTGGAAGCATTTCGCGATGGACAGAC
TCACAACCTGAACCTAGGAGTGCCCCATTCTTTTGTAATTTAAATGGCTGCATCCCCCCC
ACCTCTCCCTGACCCTGTATATAGCCCAGCCAGGCCCCCAGGCAGGGCCAACCCTTCTCCT
CTTGTAAATAAAGCCCTGGGATCACTGTGTCGTCGCCTCTGAGCCCCTTTGCTTGCCCAGTG
AGTGGGCGGCCAGAGGGCAGGGATGGGTAACTGTGTGTGCCTCCGTGCGTGCCTCG
CGTGAAAGCTCCGCCTTCCGTCAGACGGACGTGGGTCGGGACTCCCGCCTCGCACGTGGGA
GGGTGACCGTGGGTGAAGCTCCCCCAGTCTCCTTCTTTAAAATGGAGGGCGATCATAACAG
GGTGGTTGTGAAAAGCACCGAGATGACGGCTGACGATAAGACGGGCACAGTGACTCATCA
CACGCTTGCCATGTGCCCAGGCACTAAAAGACTACACACACGTTAGTTCAGTCTAGGCACTT
CTGTCATTCTCATTTTACCGTGGCGGAAACTGAGGGACAGAAAAACTAAGTAACTTGGTC
ACTTGCCCAAGGTCACAGGGCTATGGAACAGTGAGGCTGGGATTCGAACCCAGGCTGTCT
GACCCCAGAGCCCACACTCCTTACCCTGGAGTTGCAGCTGGGGCCACCCTCAGGGGGGCC
CTGATCACACTCCCCCTGATGCTGAGTTCCAGATCTGAACTAAGAAGAGTAGTTAACAGCC
GGAAGCGCAGACCTGAGGCCAGCCCCGGCT

MVRWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQGDFSLSVRVGDQVTHIRIQNSGDF
YDLYGGEKFATLTELVEYYTQQQGVLQDRDGTIIHLKYPLNCSDPTSERWYHGHMSGGQA
ETLLQAKGEPWTFLVRESLSQPGDFVLSVLSDQPKAGPGSPLRVTHIKVMCEGGRYTVGG
LETFDSLTDLVEHFKKTGIEEASGAFVYLRQPYYATRVNAADIENRVLELNKKQESEDTA
KAGFWEEFESLQKQEVKNLHQRLEGQRPENKGKNRYKNILPFTDHSRVILQGRDSNIPGSD
YINANYIKNQLLGPDENAKTYIASQGCLEATVNDFWQMAWQENSRVIVMTTREVEKGRNK
CVPYWPEVGMQRAYGPYSVTNCGEHDTTEYKLRTLQVSPLDNGDLIREIWHYQYLSWPDH
GVPSEPGGVLSFLDQINQRQESLPHAGPIIVHCSAGIGRTGTIIVIDMLMENISTKGLDC
DIDIQKTIQMVRAQRSGMVQTEAQYKFIYVAIAQFIETTKKKLEVLQSQKGQESEYGNIT
YPPAMKNAHAKASRTSSKSLESSAGTVAASPVRRGGQRGLPVPGPPVLSPDLHQLPVLAP
LHPAADTRRMCMRTCTLRTRGRRK

SEQ ID NO: 59

FIG. 9

AGGCTCAAGCAATCCTCTCACCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCGCGCCA
CCACGCCCGGCTAATTTTTGTATTTTTTGTAGAGATGGGATTTCACTATTTTGCCCGGGC
TGGTTCCCAACTCCTGGACTCAAGCGATTCGCCCGCCTCAGCCTCCCAAAGGGAAGTGCT
GGGATTTCAGGCGTGTGCCACCGCTCCCACCCCAAAGTAGTATTTATTGTAATTATTATT
ATTATTTTGAGACGGAGTCTCGCTCTATTGCCAGGCTGGAGTGCAGTGGCGCGATCTCGG
CTCAATGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCTTCAGACTCCCAAGCAG
CTGGGACTACAGGCGCCCCCCACCACGCCAGGCTAATTCTTGAATTTTTAGTGGAGACGG
GGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCTTGACCTCGTGATCCGCCCACCTCG
GCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCAGCCTATTATTATTTTT
TTAGGCAGTGTCTTGCCCTGTCGCTCAGGGTGTAGTGCAGTGGCGTGATCACGACTCACT
GCAGCCCCGACTTCTCGGGCTTAAGTTATCTTCCCGCCGCAGCCTCCACGCCCGGTTAGT
TTTTTGCATTTTTTGTAGAGATGAGGTCTTGCTTTTTTGCCCAGGCTGGCCTCGAACTCC
TTGGCTTAAGCGAACCTCTTGCCGCAGCCTCCCAAAGTGTTGGGATTACGGGCGTGAACC
ACCGCGCCCAGCCTACTATCTTTATCTTACAGAAAGAAAAGAATGGAGGAAACCGAGGCT
CGGAGACAGTAGGTAATTTCCCCAAGGTTCCACAGCTAATGAGTGGAGCGGCGATTTGTG
CAACGAAATGAATGAAATCGATCTGGCAGCGCGCCCGGACGGTCGGTGGCGTAGACCCG
GAGCCGCAGCTCACACCTGGCGGCCGCGGTTTCCAGGAGGAAGCAAGGATGCTTTGGAC
ACTGTGCGTGGCGCCTCCGCGGAGCCCCCGCGCTGCCATTCCCGGCCGTCGCTCGGTCCT
CCGCTCACGGGAAGCAGGAAGTGGCGCGGCGTCGCGAGCGGTGACATCACGGGGGCGA
CGCCGCGAAGGCCGCGGCCCGAGGAGCAGCGAGCCGGGCCGGCGGGCAGCTGCACAGTC
TCCGGCATCCCAGGCCTCGAGGGGGTCTGTGCGCGGCCGGCTGGCTCTGCCCCGCGTC
CGGTCCCGAGCGGGCCTCCCTCGGGCCAGCCCGATGTGACCGAGCCCAGCGGAGCCTGAG
CAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGAGGAACATGACATCGCGGAGGGTGAG
GAGCCCCGAGGGGCCCGGCGCGGGCCTCGGCCCGGCCACCGCCGCGTTCGGTTAGCCCCG
TCCGGAAGGGGGCGCCCCGGCCGGGCTTCGGGCTCCCGCCCCGGGTCGGGGTTGGGGGCC
GGTTCCCTCCTCGTCCCCTCGCCCTCCAGGGGCCGGGGGCCGGCCCCACCGCGCCCCCAC
CCCTCGGGTCCCCATTCATTTCCTGCCTCCCCGAGTTCCGGCTGCGGCAGCCCCGGGGAT
GCCCGTCAGGCCCGGGGCAGGTAGAGCCGCCGAGGGAACCACGGGTGCCAGCGGCCAGGC
TCAGCGCCGCATTCCTGACCCATTGCCTCATGAGAATTGCCTCATGGTGATTCCGAAATA
ACCCTGCTCACTTGGGGAGGCTCCTTGGGACACGAGAGGGGAGTTGCGCGGGGCCGGGCC
CCCAGTGGTCTAGTCGTTCTGGCTCACTGTGCCACTTTCGTGCATTTGGGGACTTCACGC
AGGACCCCTGACCCTTTTATATGCCTCTTTGTGTCTTCTTTTCCTCCTACCCCTCACGTG
CCAGAAATGGAAAAACTGACTGTATCTGCAGCCACTAGAAGTATTTCCTTCCTCTGCGAT
CTTCGCTTTGGGAGATGGAAAGGAAGGGAGCCGCATCTCGTTATTTAATCCTTCACTGCA
ACCTTAACAGTCAGGTCACTTTACTGGTACCCGTTTTATGGATGAGGAAACCGAGGCCCA

FIG. 10

GAAGCAACATGCTAGTAAATGACAAGATTTGAAACTTAGGAGGATTAGTGAGTTAATGAG
ATCCTTTGAAAGGTCAGGGTAATACTACTACTAATAGCTAACATTTGCTTAGTTCTGACC
ACAGCCCTATCAGATGGCTACTATTATCCCCATTGTAAAGATGAGTAAACCGAGTTTCAG
AGGTTAAGTAAATTGCCTAACCTCACAGCTAGTAGGTGGTGGAGACAGAATCCCTACTTT
TAATCACTATGTTGCTTCTATTATTTTGTAACTATTGCTAACCATTTGTAAGCCTTAATT
TTGTTGTCAAACAGTAGTGTGACCTGTTGTTTTCAGATAGTGATCCTGCTATTTTGTATA
GTCACTCTATATACCACTCACACTTAAGACCCATTGTCTATTCTTTTCCATGATTGTTCA
ATTATGGTCACTGTCTCAGACATTTAAAAAACGATTCAAGCTATTGAGGCTATTTGAATG
AGATTTTCTTTTCTTTTTTTCTTTTTTTTTTTGGAGACGGAGGCTCACTCTGTTGCCCAG
GCTGGAGTGCAGTGGCGCAATCTCGGCTCACCACAATCTCCGCCTCCTAGGTTCAAGCGA
TTCTCCTGCCTCAGCCTCCCAAGTAACTAGGACTACAGGCGCACCACTATGCCCGGCTAA
TTTTTGTATTTTTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGTCTCAAACTCC
TGACCTCGTGATCCGCCCGCCTTGGCCTCCCAAAGTGCTGGAATTACAGGCGTGAGCCAC
CGTACCCAGCCTGAATGAGATTTTTCAAAATATTAGGAATGTCTCCTCCAAACACACCTG
GCATGTTATTCATACATGGATCTGGAATTTAAAAAGGGGAGAAAAAGAAAACTGAGAACT
CGTAGGAAGTGAGTGACTTGGACAGGTCGGTTGGCAAGTGCTTACAGATCTGGGTAATAT
ATAACTGCATTTCAACAGAACAGTGTATAGCCTCAAATGTTCTAATTCTTTAGGGAGCTT
TTAAATAAACAGTTGTCTATTCTTTAATCTGTCAAATAGTCATTGAGCCTTTTGTTCCTG
GTGTCTGCTCTTCCAGACAAGTAAGGATCTGCTGCTTTAGGAGACATCAGACGGGGCTGG
GGGTTGGGAAAAGGTCTGGGTAGTAATAGACCCTACATTGTCCAGTTTGTTCATTTAGAA
GCATAGAAGTGTGGGCATAGTCAAAGTAGCAAGTGGTAAAGATGACAGTTTGAAATGGAG
TAATTCCTTCTCCCCTCCAGCCCTGGTATTATGCACCACCCAAAAAGCCGGGTTATGAAC
ATAATACACATAATTTTGAATGATTCATTATTTTTTGGATTATAAGCCTGTTTTATTTGT
TAACCAGCCTTAATGAGGTATAAATGACATGCAATTAATTGCATATATTTAAATGTACAA
TTTGATCAGTTTTGACATACATATACACTTGGGAAACCACCACCATAGTCAAGATAATGA
ACACATCTATCACCCCTGGTAATTTTGCCTTATGTTCTTTATAATCCTTCCTTTGTTCTT
AGGCAGCCACTATTCTGCTTTCTGTCACTATGTATTAGTTTGCATTTCCTAGAATTTTAT
TTTTAAAAATTTTAAAATTGTTTGAATAGAGATGGGGTCTCACTGTGTTGCCCAGGGCAG
TCTCAAACTCCTGGGTTCAAGTGATCCTCTCACCTTGGCCTCCTGAAGTGTTGGGATTAT
AGGCATGAGACACCCTGCCCAGCCCTAGAATTTTATTATTATTGTTATTATTGTGTTTTT
TTGAGATAGGGTCTCACTTTGTTGCCCAGGCTGGAGTGCAGTGGTGCAATCACTGCAGCC
TTGTTTTCCTAGGCTCAATCCATCCCCCCTCCTCAGCTTTCCGGTTACTGGGGCTACAGG
TGTGCACCACCACACCCGGCTAATTTTTGTATTTTTTTATAGAGACAGGGTTTTGCCATG
TTGGCCAGGCTGGTCTCAAACTCCCGGGCTCAAGCGATCTTCCTGCCTCGGCCTCCCAAA
GTGCTGGGATTACAGGCATGAGCTATTGCGTCCCGCCTTCAAATTACTTTAACCTAGTAT

FIG. 10
(continued)

TAATTCATTCAACAGGAAGTTAATGAGCCAGGCAGGATAAAGCAGTAAGATAGGAAAATA
TTGCTATTTTCATGGCTGAGAGAGAGCAGACAAACACATGACTAAATAGGGCAATTTCAG
GTAGTAATAAATTCTAGGAGGGAAAAAATCCCACAGAAATGTGAGGATGGGAGAATGCAG
TTAGTTTTGATAGGTGGTTTAGAGAAGGTGATCGTGTGAGCTGACACCTGAATGACAATT
AGTAGTCTGAATTTTGTTTTGCTTAATTATCAAAATAACTCCTCTTGGGTTCGGCTTTTA
TATGCATCCAGTAATTAAAATGTAAGTATATTCAATGTACTGATATCTCTCAGCATCATA
GGTAGGAAAACTAAGGCATTCAGCAATTAAGTGACTCCTCCCTTGATCATGTAGCAGTGA
TAGTACTGGATTTAGATTTTGAGGTTGCTTCTCTGCCCTTTTCTGCCTTTGTGAAACCAA
CAAAGCTGCCTGTATTTTCCAACTCTTCCTTCAGCATGTGGTACCTCCTTTACATCTGTT
TTTGTTGCTCTGAAATCCATACGCGACGATGAGCTGAGAGGGGCAGAAAATTGAGCTTGT
TCTGAGACTGGAGGCTTTTGGTTTATCTCTTGCAGGTCAAGTACATTTTGTCCTGGGCTC
TCCCTGGTGGCCACGTTTGTTTATCTCCTGCGGGAGTAAATAAACTTGCCTTGCTGAAAA
ATAACAGTTCTGTGTCTTTGCAGTGGAAACTGGGATGTCTTTATTAACGTTAGGTCCTGA
TGTAAGGCCAAGTTTTTGGTTAGAGTTGCTCAAGTGCAGAGGCCACTGCTAAGATGACTT
ACCCCTCGTGTCCATGGTCAATGTGGAGACTGTTATGAGTGGCACATGATGCTGGAAAAG
CAGAGCCAACTCATGTTTGTAATTGTCCTAGCAGGCCGTGGTGTACTTTGTTAGGCAGCC
ACAGAACAATAGAGAAACTCAGCTTATTCCCCTTCCCTCTGGGAAACACAGACAGTACTT
GCCATCCAACGCCAATGTTTTTAAGGAAGAAAGAGGCAAAAAGTGATGTTGGCAAGGTCT
CTGGGAGTTGTGGACCCCAACCAAGGATTGGAGACCCTGAAATGGATTCAGATGCCCTAA
AATGCAGCCCAGTTCATTACTATGAATTTTGGAGGACTTTGTGCCTTGAGCAAATGTGTA
TATGTGACGCTCTTTGACAACACTGAAATAGGAAAAATACTATCCATGTTCGCGAGGAGC
ACTGAATTTAGAGAGGGAGACAGACTTTTATGCCAGCATCAAATGAATTTGATAAAGCTA
GTACCAAAATGAAATTTGAAATTTTTTTTTTTTGAAATAGAGTCTTACTCAGTCACCCAG
GCTGGAGTGCAGTGATACAATATTGGCTCACTGCAACCTCCACCTCTTGGGTTCAAACAA
TTCTTGTGCCTCAGTCTCCTGAGTAGCTGGGATTACAGGTGCGTGCCACCATGTCTGGCT
AATTTTTATATTTTTAGTAGGGATGGGGTTTCACCATGTTGGCCAGGCCGGTCTTGAACT
CCTGGCCTCAAGTGATCTGCCCACCTTGGCCTTCCAAAGTGCTGGGATTATAGGCATGAG
CTACCACACAAGCCTGAAATTTGAAATGTATTGGTATAGAATATACTGTTTAGAATGTAT
GTGTATATATGTATATTTGTATACTCATATAAACACAAATACACATTGTATGTGTTTCTG
TAATATGTATATCTGTCTACACATACATGTATATACACACATACAATGTCTTTTTTTTTT
TTTTTTTTTTTTGAGACAGGGTCTTACCCTGTTGCCCAGGCTGGAGACTGCAGTGGCATA
ATCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTCAAGTGATCCTCCCATCTCAGCCTCC
TGAGTAGCTGGGACTGACTACAGGCACGTGGCATCAAACTTGTCCAATTTTTCTATTTTT
TTGTAGAGTTAGGGTCTTGCTCTGTTGCCCAGGCTGGTCTCAAATTCCTGGGCTCAAGCT
GTCTGCCTGCCTCGGCCTTCCAAAGTACTAGGATTACAGATGTGAACCACTGTACCTGGC

FIG. 10
(continued)

CTTTACAATGTCTATTTTAAAGATAATGGTTCAAGTTTTTATCATCCCACTGGCCTACTC
TAATGAAACATCTATCCATTCATTGAAGAATTATTTATGGTGGGATAACTCTGTGCCAGG
TACCGTGCTAGGCATTGAGTATTCCAGGTTTTAGGAAACAGCACATGCAAAAGTGCTGAA
GTGGGAGAAGATCTCGGAGTGATTGAAGGCTAGGAGAGAGCAAGTGTGGGAGCTGTGAGG
CTGGGAAGGTGGGAGGTAGGTGGGAGCAGACCACATAGGGATTCTTAATGTCTTTAGTGT
CATGTGGACCATGGAGAGGAGTGTAGATTGTATTTTTAGAGCAATGCAAAATCATAGAAG
GATGTGATCGGGGGGAGTGGCATGAGCTGATCTATTTAAAAATATTTCTCTGGCTGCTGTG
AAGGAAGGATTGTAGGAGGCAGGAGTAGATTCAGGGAGATGAGACAAGTGATGAGAGAGG
CTTTGAACTTGGGTAAAAGTAGTTTGTGGAAAGTCTTTTTTGGAGGTAGTTTTTGTTTAT
TGCCTTGTCATCAAAGCAGAGATGCTGACCAATGAAACTCCATGAGAAAATAGTGATTTA
TAAAGACATATCTATGCACTGCCATTAAAAAGCTGCTTGGAAAAAAAGGATAAAAAGCTG
CTTTAACAACTTTTTTTTTTGAGATGGGGTCTTACTCTGTCACCCAGGCTCACGACCTCA
GCTCACTGCAACCTCTGCCTCCCAGGCTCAAGCATTCTCCCACCTCAGCCTCCCGAGTGG
CTGGGACTGCAGGCACACGCCACCATGTCAGGCTAATTGTGTGTGTGTGTGTGTGTGTGT
ATGTGTGTGTGTGTGTGTGTGTGTGTGCTGGGACTGCAGGCACACACCACCATGTCAG
GCTAATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA
TGTAGAGATGGGGTTTTGCCATGTTGCCCAGGCTGGTCTCAAAATGTTGCCCAGGCTGGT
CTCAAACTCCTGAGCTCAGGTGATCCACCCGCCTCGGCCTCCAAAGTGCTGGAGATTACA
GACGTGAGCCACTGTGCCCACCTAACAACTTTAAAAAAATTTTGACATTTAGTAGGATAT
TTATTGCATTATTGTTGAGATGGCAAAATATTGGAGACAACTGAAATGTTCATCAGTGGG
GGGGGCTAGTTAAATGAAATACAGTGTAGCATGCATTAGAACACTTTTCAAGAATTTAAC
TTTTTTTGTAGCCTTTTACTTATAATGCTTGTCCCTATTGATGCCTTTTTTTTCAGCATG
ACTTACTCTTTTACTATAGGATATTAAAATTTAATTAGATTAGAAATGAGGAATATTCTT
GTAATCTGTAGAAAGTAACAAACTATAAACTTATTCCCCAAGAACAAATATAATAATTTT
TCTGGAGTAGCAGGTAAGAAAGATATAAATTTATATGTATACAAGAAACTGAAATTAGAC
TTTATACATTTAAAGGTTACAAGTGCAGTTTTATTACATGAATGTATTATCCAGCATTGA
AGTCTGGGCTTTTAGTGTAACCAGCACCTGAATAACATACATTGTACCCATTAAGTAATT
TCTCATCCCTCAAACCCCTCCCACCCTGAAATTAGACTTTGGATCCCTAGTTTAAATTCC
ACCCCTCTCTTTTTTTTGAGACAAGGTCTCACTCTGTCACCCAGGCTGGAGGGCAATGTTG
CAATGATAGCTTACTGTAGCCTCAACCTCCTGGGCTCAAGGGATACACCCTCCTCAGCCT
CCTGAGTAGCTGGAACTGCAGGCGTGCACCACCACATTCAGCTAATTTTTTGATTTTTTT
ATAGAGATGAGGTCGGAACTCCTGGGCTCAAGCGATTCTCCCCAAGTGCTGGGGTTACAC
ACATGGGCCACTGCCCCCAGCCTAAACCTCCTTTCTCAGTATAGCAGCCTTGAGATGAAG
TTCCTGAAATTACTGGCCAGCTTGACTGTTTCCCCACATCACTGGAGGAGGGGGATGCAT
AGATAAAACAAAATATTCAGCATCATTGTATTTTCTTTTTGTTTCATCAGCATCTTTTTT

FIG. 10
(continued)

TAAAACTCACTTGACATAAGTCCCTAGCCTCAAAGAGTAAAGCCTTTGCAGAATCTGCAT
TCAGATTTCGGGTGTGATTTCCTGACAGATAGTTCAGGTTTGTAAACTCTTTTTTTTTTC
TTTGAGACAGAGTTTCACTCTTGTAGCGCAGGCTGGAGTGCAGTGGCACCATCTTGCCTC
ACTGCAACTTCTGCCCCCTTGATTCACGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCATGCGCCACCACACCTGGGTAATTTTTGTATTTTTAGTAGAGATGGGGT
TTCACCATGTTGGCCAGGCTGGTTTTGAACTCCTGACTTCAGGTGATCTACCTGCCTCAG
CCTCCCAAAGTGATGGGATTACAGGTGTGAGCCACCGCAGCCGGCCAAAACTTTGTTTTT
TTTCCTCTTTTTGTTGCTGAGAAATGTAAACTCTTACAGACACAAATTATGTCTCCCATT
TTTTAAAACCCACTCAACACAGGGGTCATGTGTAATAGGCCCTGGAGCTTATTTTAGACA
TTGATTTGAGGCTCTTTTCCCCAAGTGCTGGTTTGTGTGTGTGTGTATGTGTGTGTAAGT
CTTTCTATGAGATGAGTGGTACCTACCTGGGCTGTGTGATCTTTTTTATTTTATTTATTT
TATTTTTGTAGATACGAGGTCTCACTATGTTGCTCAGGCTGGTCTTGAACTCTGGGGCTC
AACCTATCCTCCCTCCTTGGCCTCCTAGAGTGCTGAGATTACAGGTGTGAGCCACTGCAC
CTGGCCAGCGATCCTTAATAAATATAGATAATGGCCGGGCGTGGTGGCTCACACCTATAA
TACCAGTACTTTGAGGGGCCGAGGCTGGCAGGTCACCTGAGCTGAGGAGTTTGAGACCAG
CCTGGGTAACGTGGGTGAAACCCTGTCTCTACAGAAAATAGAAAAATTAGCCAGGTGTGG
TGGTGCATGCCTGTAGTCACAGCTACTTGGGAGGTTGAGACAGGAGAATTGCTTGAACCT
GGAAGGTGGAGGTTGCAGTGAGCCGAGATCGTGTCTTTGAACTCCAGCCTGGGTGACAGA
GTGAGACCTTGTCTCAAAAAAAAATATAGATATAGGCTGGGCGTGGTGGCTCACACCTGT
AATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCAGGAGGTCAGGAGATCGAGACCAT
CCTAGCTAACATGGTGAAACCCTGTCTCTACTAAAAATACAAACAATTAGCCAGGCCTGG
TGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCC
GGGAGGTGGAGGTTGCAGTGAGCCGAGACTGTGCCACTGCCCTCCAGCCTGGGCGACAGA
GCGAGACTCTGTCTCAAAAAAAAAAAATCTATATATCTATATATCTATATCTATATAGAT
ATAGATATAGATAATGCCAGATGATGGCTGGTTAGAAGGGATTGTCAGGGGCTGGCAGGT
TTTGCAGGTGTTAGAATGAGCAAGATGAGGAGAAGGATGCTTACTTCCCTCTCCTTGTAA
CTCTCTACCCCCTCCCCTCAGTGTTTTTTTATTTTTATTTTTATTTATTTATTTTTTTTG
AGACAAGGTCTTGCTCTGTCACCCACACTGGATTGCAGTGATGCAATCATAGCTCATTGA
AGCCCAAACTCCTGGGCTCAAGTGATCCTCTTGCCTCAGCCTCCCAAGTAACTGGGACCA
CAGGTGCGTACAACTATGCCCAGTTAAGTTTTTCATTTTTTATACAGACGGGGTCTTGCT
ATGCTGTCCAGGCTGGACTTGCACTTCTGGCTTCAAGTGATTCTCTTGCCTCAGTTTCCC
AAAGTGCTGGCATTATGGGCATAAGCCACTGTGCCTAGCCCATCAGTGTCTTTTTATCCT
TTACTCCTATCAAAATTCATTCACTCAGCAGCCATTGATCAAGTGCCTACTATATACATG
TTGAGGACTGGAAATTTATTTGTCTCTTCTCATCTTATCTGGACCCTCTGTGTTAATTGT
AATTAACTGTAATCATTCTGTATTAATTGTAATAAACTTGTTGATAAACTCAAATGAGGC

FIG. 10
(continued)

CATACCGTTTTGCCACTTCCCCTCCTTCCAGGTTATATGGATGTACTTACATTGCAGGTT
TCATTTGTTGGTTCAGTTTTTAAACTAAGCCCTATTGTGTCAAATTATGCTAGGTGTGAG
ATGGGGAGTTCAAGCTGTGTGTTGTCTTTTTTTTTTTTTTTTTTTTGCCTCACTTACTA
ATATACAAGCGCTTATAACCTTTGAGGCTGGCCCTATACATTAAGATTTTTATTAATTCC
ACTGTTCTTTATCTTCTCTTACTAAGTTCTCAGGGTCGAATGAACTCTAACTGCTCCTTG
CTAGTGATAAGCAAGTTGCAAATTACAGAATTGTCAGTGATTGAATACACGTATTAAACC
TGTAACTGGGAAGCATTTTTGGTAATTATGAATACTTTTGGAAAAAAAAAAGCTATGGAA
GGAAAGTTTAAAATCTACGAAAGCTCAAGTAGATGGTCATGGAATAGCTATTTCAATTTC
TAACTATATATTACTTATTTATTTATTTATTTTTGAGACGGAGTTTAGCTCTTGTTGCCC
AGGCTGGAGTGTAATGGCGTGATCTCAGCTCACTGCAACCTCCACCTCCCGGGTTCAAGC
TATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTATAGACATGTGCCACCACGCCAGG
CTAATTTTGTATTTTTAGTAGAGACGGGGTTTCTCCACATTGGTCAGGCTGGTCTCGAAC
TCCCAACCTCAGCTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGA
GCCACCGCGTCCGGCCTCTTAACTATTGTTTGAAATAATGTAGAGACAGCTCCAGAGCCA
TGAAGAAGTGTATGAAGAAGCAGTGTTAGCTTAAATGACATACATGTCACAATTGCCTAT
GTGAAACTATCATAATTATGCATGAGAAGTATCTATCCTGCATAACCTCCACCAATAATA
ATAATGTTAATAATAGTGAAAACTAATGTTTATTAAGTCCTTACTGTCTCCAGCCTCTGT
GCTAAATACTGGTTACTAAGTTTCCCTGAAAATACTATTCTCATCTGTTTGTTCTTAATA
ACAGGATAGCATAATTGTAAGTTGTAAATGAAATAATACAGTTTATGTAATAAAAGGGTA
AAAGAGAAGACCACCTACCTTATCTTCTGTTGCTGATCTGGATGGATGTAGGTGGTGTTT
ACCTAGTTTCACCTTTGGCAGTTGAAACTACTTTTTTTTTTTTTTTTTTTTTTTTAAGA
GACAGGGTGGGCCAGGCGCAGTGGCTCACGCCTGTAATCCCCGCACTTTGGGAGGCTGAG
GCGGACAGATCACTTGAGGTCAGAAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCT
GTCTCTACTAAAAATACAGAAAAATTAACTGGGTGTGGTGGTACACACCTGTAATTCCAG
CTACGTGGGAGGCTGAAGCAGGAGAATCGCTTGAACCCGGGAGTGGAGGTTGCAGTGAGC
TGAGATTGTGCCACTGCACTCCAGCCTGGGTGACAGAGCAGGACTCCGTCTCAAAAAAAA
AAACAACAACAAAAAAAGAAATTTTTAGAAATATGAGATGACAGCAAGAATGAGGGTATT
AAAAAGAAATTTTTAGAACTAAATAGCAGAATGTAATGGTGAAAAGTTTGATTTCTCAAG
TCTGCTTTGCACACAGGCATGTGGCAAACATTCAGTAAGTATAGCTGTAATTTTAACCAG
CTGTAATGTATAATAGCCAACATATCACATTTTTCTTTTTTCTTTTTTGAGACAGAGTCT
TGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCACCATCTCGGCTCACTGCAACCTCTGCCT
CCTGAGTTCAAGTGATTCTTGTGCCTCAGCCTCTCAAGTAGCTGGGATTACAGGTGTGTG
CCACCACACTCGGCTATTTTTTGCATTTTTAGTAGAGATGGGGCTGGTCTTGAACTCCCA
GCCTCAGGTGATCTGCCTGCCTCAGCCTCCCAAAGTGCTGAGATTACAGGTGTGAGCCAC
AGCGCCTGGCCATATATTGCTTTTTTCTTATTATCAGAGCCAGTTCATAATTGTGGAAAA

FIG. 10
(continued)

ATAGTGTTTGTAACAATGTAAGTATGGATAAATCATCTTTTTAATTTTGTGATTCATATA
GGTTTGTTGTTGTTGTTGTTTTGTTTTTATCTTGAGACAGAGTCTTGGTCTGTCACC
CAGGCTGGAGTGTAATGGCACAACCATGGCTCACTGCAGCCTCAGATGCCTGGGTTCAAG
CAATCCTCCCGTCTCAGCCTCTAGAGTAGATGGGACCACAGGTGTGGGCCACCATGCCTG
GGTAATTACAAAACTTTTTTTTTTTTTTCTAGAGATGAGGTCTCACTATGTTGCCCAGGC
TGGTCTCAAACCTTTGACCTCGCTTCAGCCTTTAGAGTAGCTATGACTATAGGCATGTGC
CATCACCCAGCTAATTAAAATTTTTTTTCTTTTTTTTTTTTGGTGGAGATGCGGTCTTACT
TTGTTACCCAGACTGCAAGTTAGTTTCAGATATCAACATTTGGTGTTTCCAAATGCACGG
GGAGGCTTTGGAGCAAGTTTTTGGCTCATATGCATAGGTGTCCTAGACATTCACTTTGCA
AATTCTTATTAAAATGACTACAGTAGCATACAGATAGGGAAAAATATCCTTGTCAGTACC
ACCGATTGGGTGAGAAGAGACTGTATATTAAAAACAATGACCATCTTTTTGCCACATAAA
TTGCTGGTGGGGCCAGTTTGAAGAGGGCTTTGTCAGCTGCCTTCTGCCTCTTCCTCTTGA
GTACGTGGAGTTGGAGTCATCCTTGACAGCCTCCTGTTGACACCACCCGGGTCACAGATG
TGAAACTGTGTGGATGTAGGAGAGAGCAGTGATGGGGCTTACCCCAAGGTTGCTCTTCCT
TCCCTCTGGCCACAAATGTTTAGTAAGGAACTGCTCTGTATTAACCATTTGCTAGGGGCT
GCAGATACGGTGGTGAAGAAATAGACATGTTCCTACTCGGGATGCTGAGGTGGGAGGATT
GCTTGAGCCCAGGAGTTGGAGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCTG
GGGGACAGAGCGAGACCCTATCTCTAAAAAACAATAAAAGAAATAGATGTGTCCTTCACC
CTCATGGAACTGCCAGTCTAGCCTTCAACCTGGTGACTGTAGAAATGTGTGATTAGATGC
TATATTGCCATGTTGAGTGTCACCCCTGAGAAGCAGGGTTTTTTTTGAGAAGGTAGGATG
GGGGATCTGACTGTGGGACCACCAGAGGGAAAAGCACATGTAAAAGCTGCGTGTACCAAC
TGGAGGAAATCGGAGACGTGATCAGAGAACCAGAGTCAACCAGGGGCCATGCCGTACAGG
GTCCTGTTAAGATCTGTGACTTTTTTCTAAACGTTTTCTTCTGGATAACATCTAAATTTC
TAGTTCCAAATGTGAAACTCCAAGGGCGTTCTGTGCTAAACATTTTGCATGTATTAATTA
ATTTCCACCACACAACATTGCTGTGAATTAAGACAGTTTCTAAGCATGGCAAGAAACCCA
GAAATCATAATGGAAAAATCTGATAAATTTAACAATGCCAACATGAACCTCTGTAGGAAA
AAAAATACCACAGACTAAAAAGGGGGGAAAAAAACCAGAGACAAATATTTGCAACACATA
CAGTAAAGGGTAATTTTCTGGTTATATCAAGAGCTCCTACAAATCAGTAAGAAAAAAAAT
CTAATAGGAAATGAGCAACGACAAACTGACAACTCATAGAAAAGGAAACACAAGTGGTCT
GAAAACATGAAAAAGTGCTCAGTCTCACAAAGAAATGCAAACTAACATGGTACCATTTTC
CATTAATCAGATAGACAAAGATGAAAGAGTTTGGTAATGTATGTAGTATTGGCACAAGTG
AGGGAAAACAGGGGATTTCACACTCTATGCCCGTCCAAACCAGTACCTTATTTTGAGGGT
GGTTTGACAATATTTGTCAAAATAAAAAAATTATATATAGTCATTTGCCACATAATGATG
GTTCAGTTGATGATGGACGGCATACATAATGGTGGTCCCATAAGAATATAATGGGCTGGG
TGCAGTGGCTCTCACCTGCAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATTGCCTG

FIG. 10
(continued)

AGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTGCTAAAAACA
TACAAACAATTAGCCAGGCATGGTGGCGGGTGCCTGTAATCCCAGCTACTCAGGAGGCAG
AGGCAGGAGAATCGCTTGAACCCGGAAGGCGGAGGTTGCAGTGAGGTGAGATTGGGCCAC
TGCACTCCCATCTAGATGACAAGGCAAAACTCCATCTCAAAAAAAAAAAAAAAAAAGAAT
ATTATGGGCCCAGCCACAGTGGCTCACACCTGTAATCCCAGTACTTTGGTAGGCCAAGGC
AGGAGAATCATTTGAACTCAGGAGTTTGAGACTAGTGGGGACAACATAGCAAGACCCCAT
CTCAAAAAAAAAAAGATTATGGTGGAGCTGTCCTGTATAGACATACCATTTTTAACTTTTT
TTTTTTTTGAGATGGAGTCTTGCTGTGTCACCCAGGCTGATGTGTAGTGGCGTGATCTGG
GCTTACTGAAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCTTCCTGAGTA
GCTGGGACTGCAGGCGCAGGACACCATATCTGGCTAATTTTTATATATTTAGTAGAGATG
GGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCCGCCTGCC
TCAGCCTCCCAAAGTGCTGGGATTACAGGCATTAGCCACCATTTACAGGCACCTGGCCAC
CATTTTTAATCTTTTATATTGTATTTAAACTGTACCTTTTCTATGTATGGATGTGTTTAG
ATACACAAATACCATTGTGTTACAGTTACTTACAGTATTCAGTACAGTAGCATGCTGTAC
AGGTGTGTAGCCTAGGAGCAATAGGTTATACCATATAGCCCAGGTGTGTAGTAGGCTCTG
CCATCTAGGTTTGTGTAAGTACGCTCCATGATGTTACCACAGTGACGAAATCGCCTAATG
ATGCATTTCTCAGAACATATTCCTGTTGTTAAGCAATGCATGACCGTATCTTGACAAAGC
CATTTTATTTCTAAAACTTTAATTTTACAGATTTATTTGTAAAAGTATGTAAAAATGATT
GTAAAGGATATGTTCTGCTGCATTATTTGTAATAACAAAAAACCAGAGGATAACATAAAT
GTCCTATAAGAAGGGTTAGATTATGGATGGCACATTCATACAATGGGGTATTATGTAGCC
ATTGAATAAAAGGGTACTGGCTGGGCGCAGTGGCTCATGCCTATAATCTCAACACTTTGG
GTGGCCAAAGAAGGAGGATTGCTTGAAGCCAGGAGCTTGGGGCCAGCCTGGGCAACATAG
CAAGACCCTATCTCTACAAAGGAAAAATAAAACAATTAGCCAGGTTTGGTATTGGACACC
TTCATGGTCCCAGCTACTGAGGAGGCTGAGATTGGAGGGATCGCTTGTGCCTGGCAGGTT
GAGGCTGTAGTGAGCCATGATTGTGCCACTGCACTCCAGGCTGGGAGATAGAGTGGGACC
CTATCTCAAAAAAACAAAAACAAAAACAAAACCTCCTGTAAAATGTCAAGAAGTCCTAGA
TGTGGGCCAGGTGTGGTGGCTCACACTTGTAATCCCTGCACTTTGGGAGGCTGAGGCCAG
GAGTTTGAGACCAGGCAGAGCAAGATAGCAAGACTCCATTTCTACAAAAAATAAAAAAAA
TTAGTTGGGCATAGTGGTGCATTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGTGGGAG
GATTGCTTGAGCCTGGGAGGTTGAGGCTGCAGTGAGCCATGATCACACCTCTGCACTCCA
ACCTGCGCAACAGAGTGAGACCCTGTCTCTAAAAACAACAACCAAAAAAACCCAGCAAAG
TACTGATAAAGATCTTTGGCTGGGCGCAGTGGCTCACACCTGTAATCCCAACACTTCAGG
AGGCTGAGGCGGGCAGGTCACAAGATCAAGAGATCAAGACCATCCTGGCCAACATGGTGA
AACCCGGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCGTGCACCTGTAGTC
TCTGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGTGGCAGAGGTTGCAG

FIG. 10
(continued)

TGAGCCGAGATCACGCCACTGCATTCCAGCCTGGCGACAGAGCAAGACTCCGTCTCAAAA
AAAAAAAAAAGAGAGAAAGATCTTCAAGTTGTAGTATGTGAAAAAATCAGGGTGTAAAAC
AAGAGAATCCCATTTGTGTGTGTGTCGAGTGTGTTTCACACAGGCTCAGAGGGAGTAGTG
TGTATATGCACATGAACATACGTGTCAGTGTATATATGTATATATACAAGGTTGTGGGTT
TGTTTGTTTTTTTTGAGACAGAGTCTTACTCTGTTGCCCAGGCTGGGGTGCAGTGGTGCA
ATCTTGACCCACTGCAACCTTCACCTCCCAGGTTCAAGTGATTCTTGTGCCTCAGCCTCC
CAAGTAGCTGAGACTACAGGCACGCACCACCATGCCCAGTTAATTTTTGTATTTTTAGTA
GAGATGGGGTTTCATCATGTTGCCCAGGCTGGTCTGGAACTCCTGGCCTCAAGTGCTCTG
CCCGCCTTGGCCTCCGAAAGTGCTGTTGCCCAGGCTGGAGCTCAGTGGCACAATCGCAGC
TCACTGCAACCCCGACGTCCCAGGCTCAGGCAATCTTTCCGTCTTAGCTTCCCAAGTAAC
TGGGACTACAGGTGTGTGCCATCAATGCCCCACCAATTTTTTAATTTTTTGTAGAGATGG
GGTTTCCCTACGTTGCCCAGGCTGATCTTGAACTCCTGGTCTCAAGCAATCCTCCCACCT
CAGCCTCCCAAAGTGCTGCGATTACAGGTGTGAGCCACCTTGCCCTGCCCTGTACAAAGA
TCTGCATAAAAGCAGTTAATAATACTATGTTTGAGGCTGCCATCACAGGGGTGAGGTCAA
GGACAAGTGTGAGAAATTCTTTTAGAATCTATTTTAAAAAAAGAAGAGATGACAGTGGTG
ACAGTCAGGGAACAGATAAGCAGGTAGATTGTGGGGGTCTAGGCTGTCTAACTGGTGTTT
AAAATGAAGCAACCGCTGAGCCTGCTGTATTTCATTTAATGGAGACTAGTAAAACAACAG
CCAGAAATTCTTCACTTTCCATCTAAGAGAGGCAAAAGTTATTTTCCCTTCAATAACCTG
GGACTGTAGGATTAAGGTTTTTTTTTTTTTTTTTTAAATACTACAATATGACTACCAGT
ATAATTTAAAAATGATTAGAATTCTATTTGAGTAAGAAATAGGTGTCTGCCTGAAGTAGA
CAGTCACTGAAGTCACTAAGTGGCAAAAGACAGAAAAAAAATTGAAAGTAGGAAACAATC
AGCAGATATGATACCAAACATGAGCTGTCAGTGATAATGGATTAAGTCCTTCAATAATGG
CTGAGCCAGATGGAATTAAAAGAAAAAATCCAGGCCGGGCATGGTGGCTCACACCTGTAA
TCCCAGCACTTTGGGAGGCTGAGGTGGGAGGATCACTTGAGTCCAGGAGTTTGAGACCAG
CCTGAACAACATAGTGGGACCCCATCTCTATTTTATAAAAATATTTTGAAAAAAGAAAAA
AAAATTCAGTTGTGTTCTGCTTTAAAAAGACAAATTGGCACAGAATGTCAAAGAATAAAT
AAAACAAACATGGGCAAAAGAGATTCAGGTGGTACCAATATCGGGCTAAGTAGCATTCAA
GATAAAGATTATTAAATAATAAGTTAGTTAATACTAGAGTAATTGCATATTAATGAAACA
TAATCTATGGTAGAGATATTATAGTCAATAATTGTTTTATGTATTCATTAAGGTAACAAC
AAGCAAACAAGCTTTAATAGTTTTAAATGCTTTATATGCTTTATAGTTCTTTTATGTGCA
TTAATTCATTAATTCTCATTTCCTATGAGGTAAACACTATTATTATCCACATTTTACAGA
TGTAAAAACCGAAGCAGAGAGATTAATTAGCTTGCCCAGGAGATGTGGCATTCTGGGATT
TGAGACAGTGGTTTGGCTCTGTAGGTTGCTTCAATAACCAAGAGATGCTTCAAATCAGAT
TTTTAAAATATGTTTTTCAGAAGCATTTTCCTGATACTTCTCCCCTTACATGGGTGTTAG
TCTTTTGGGTTGAAAAACATGAGTAAGTGCTAGAAGAGCAAAATATGCATCCAGATTTAA

FIG. 10
(continued)

TAGTATGTCTGTTTTTCTGAGCCTTGGCATTTCATTGCTTTTATAATAGAAATGAAGGCT
TTTTTTTTTTTTTGGCTGAGAATAGCACTGAACTCAGTGGGAGGGACTGTGGGTTGTAAG
TTGTCCGCCTCTGAATGGAGTTGAATTTAAGTTTCTTGGTTTCCAAAGAATGATTGATTT
AAAGACCCTCAAATTGCAAGTTAGAACTGACTTCAGTCCTTGAGGTTTTTTACCATTTAA
TGAATAATTAAATTTATGGTAATAAATGGTAATAAATGGTAAAAATGGTAATAAATTTTA
CCATTTAATGAATTTTTCTTAAAAAGCAATTGAATTGTTGATGAAAGGTGATGTTAAAAT
TATCCCAGATTTATCAATCTTTTTTTTATTGCCCCTGGATTTTGAGTCATAGAAAGCCTT
TCCTTATTCTAAGGTTAACAAGACATTCACCCATGTTTTCCTCTAGTATTGCATTGTTTC
ATCTTTTACGTTTATTATTTATTTTATTTTATTTTTTTGAGACAGGGTCTCACTGTGTCA
CTCAGGCTGGAGTGCAGTGGAATGATCTTGGCTCACTGCAGCCTCTGCCTCCCGCCTCCC
GGGTTCAAGCGATTCTGCTGCCTCGGCCTCCCAAGTAGCTGGGATTACAGGCACCTGCCA
CCGCGCCTGGCTAATTTTTGTATTTTTTTTTTAGTACAGATGGGGTTTTGCTGTTGGCCA
GGCTGGTCTCGAACTCCTGACCTTAAGTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGG
GATTACAGGCATGAGCCACCGTGCCCGGCCTAAAATTTATTCTGATATGTGATATGATGT
ATGGTTCTAACTACTTTGTTACGGTGCATTATTTTCTAAATGTGGTATTGGATTCTTTTA
TATTTTGTTTAGAAGTTCTGCATCAATATTCATGAGTACCATTGGTCTCTGTTGTTTTTC
TTGTGCCATCTTTATTGGTATAGGTATCAGTGTTATATTTAGTTTGTAAAAGGAAGTTGG
AAGTTTTCCTTTCTTTTTAGTACTCAGGAATGATTTTAAGAATTGAGACTATTTGGTCTT
TGAAGGTTTGGTAGAAGTCCATTGGGAATCCATCTGGGCCTGGTGATTTTCTGTGCGGTA
GTTCCTTAATTGTTTTCCCTATTTTTTCTTATTTTTAATCAGGTAGCCTCTGAACCAGAA
TAGGTTCAGAGAGGCTCCCTCTATTTTTTTTAATACAAGTTGGTCTGCCTAAGTTTTCTT
ACTCTAATGGGTTAATTTTTGTAGACTGCATTTCCCTGAAAAATTACACGTTTGTTCTAG
GTTTTCTGACTTATTTCCACAACTTTTTAGTCTTTCCCCCTGGAATCATGCCCCTTTCCA
TAAACAGGACTCTGATGTACCTGAAGTATTTTCACACTTCGGGTGGACTTTCTGTTTCTG
GGGGTGGTTTTAGAGCAATTTTAGGCCTGCCACTAGCTACCCTGTTCTCTACACCATGCT
GTTTTTCTCAGAATGCTCTTCTTTTGCACAAAGGCTTGGAGTAGGAGGTTGAGCAGTCAC
TCACTGACGTTTGGTATATTTTCTTTTTTTTGCTTACAGGTAATCTGGAAGTTTGGGCAT
TCTCTTTAAGTTGAGGGTGTGGTTTTCATGTCATTTTATTTGTTTATTGTTTTCTTGTGT
GTGTTTCTTAGAGACAGGGTCCCACTCTTGCCCTGGCTGGAGTGCAGTGGCGTCTTGATC
ATAGCTTACTGCATCCTCAAGCTGCTGGGCTTAGATGAACCTCCCACCTCAGCCTCCTGA
GTAGCTGGGACTACAGGAGCACACCACCATACCTAATTTTTTTTTTTTTGAGACGAAGTC
TTGCTCTGTCCCCCAGATTGGAGTGTAGTGGTGCAATCTCGGCTCACTGCAACCTCTGCC
TCCCGGGTTCAAGCGATTCTCTCACCTCAGCCTCCCGAGTAGCTGAGACTGCAGGTGCAT
GCCACCATACCCGGCTAATTTTTGTATTTTTTAGTAGAAACAGGGTTTCACCATGTTGGC
TAGGCTGGTCTCAAACTCTTGACCTCAAGTGATCCACCCACCTTGGCCTCCCAAAGTGCT

FIG. 10
(continued)

GGGATTACAGGCTTGAGCCACTGTGCCTGGTCCCTGGCTAATTTTTAATTTTTTTGTAGA
GATGGGATCTTGCTATGTTGCCCAGGCTGGTCTTGAACACCTGGCCTTAAGCAATCCTCC
CACCCTAGCCTGCCAAAACACTGGGATTTACAGGCATGAACCATTGTGCCTGGCTTGTTT
TGTTTTTAATTCTATGTTGTTTTTGAAGGATGTATGGGGAGAGATGGATTTAGGCAATCA
TCGTTGTCCTTGGCTACCTGAAAGTCCAGGCACTCTTCTAGATACTTTATAAATATTAAC
TCATTTTATCCTCTCAACAACACTATGACATGGGTACTGTTACACCTTCCATTTTATAGG
ACTTAACAGAGAGGTTAAATATGTAGCCCAGGGTCACAGAGAGCTGGGCTTCAGACCAAG
ACAATCTGGCACCAGAGTCTATGTGGCTACCCCTAAGGCTTTGCCACCATGTGTTAGTGA
TTCTCAGCCTGTCATTTGGGGAGGGGATTGCCCTTTTTTTTAAACTTTTTAAAAAATTTA
TTCTTATTTTATTATATTTTTGAGACAGAGTCTCCCTCTTTTGCCGAGGCTGGAGTGGAG
TGGTGTGATTTCAGCTCACTGTAACCTCTGCCTCTGGGGTTCAAGTGATTCTCATGCCTC
AGCCTCCCAAGTAGCTGGGATTACAGTTGCCAGCCACCATGCCCAGCTAATTTTTGTATT
ATTATTATTATTATTTGAGACGGAGTCTCGCTCTTTTGTTCAGGCTGGAGTGCAGTGCTG
TGATCTCGGCTCTCTGTAACCTTCGTCTCCTGGGTTCAGGTGATTCTCCTGCCTCAGCCT
CCGGAGTAGCTGGGACTATAGGCGCGCACCACCATACTTGGCTAATTTTTTGTATTTTTA
GTAGAGACGGGGTTTCACTATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGAT
CTACCTGCCTTGGCCTTCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATGCATGGCTG
GATTGTCCTTTTTTAAAAAAAAAAACAAAAACAAAAAAAAAAACCCAAACCATAAACCCA
ATATTCTGAAAGATTTGGTCTCCACACCTGTGTTATATAATAATTAGTTTTTCCATTTTT
TTCCTCTTGGTAGAAGGCACATATGCCACTCAGTTTCCAGTTGCCACACCCAATTAACAT
AATTGTTTTGCAGCCAAAAGCAAAAGAGAGTTGACATTTTAATTAGCTTATGTAGGTAGA
CAAATTGAGGCCTAATGTAAGAGTTTCATTATACCTTTTTGAAAAACTATAAATAGCTAG
AAGCCAGTTGTCATTACTTTTTGATTCCTTAGAATTCTGGGCATCTTTCATCTGGAACCA
CAGATGAAAGAAGCTGCAAGGAAGGATTTTTTTTCTTAACGGAATAGTTTAACCATTCTG
AATGCAAAAGTATTGGATGCTAGAATAATAGGTATCACATAAATTGAGGTTGACGTTTTC
CCGGGTGAAATTCTATTCTGTCTCAATTTTCCTTTTTTTTTTGAGACGGAATCTTGCTCTG
TCGCCCAGGCTGGAGTGCAGTGGCATGATCTCGGCTCACTGCAAGCTCCACCTCCTGGGT
TCATGCCATTTTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGGGCCTGCCACAAC
ACCCAGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCCCAGGATGGTCTCAATCT
CCTGACCTCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCCGGGATTACAGGCGTGAGCC
ACTGTGCCTGGCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGACAGAGTCTCGCTTTG
TTGCCTAGGCTGGAGCGCAGTGGCATGATCTCAGCTTATTGCAACCTCCGCCTCCCGGGT
TCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTATCTGAGATTACAGATGTGTGCCACCAT
GCCTGGCTAATTTTTGTATTTTTAGTACAGATGAGGTTTTGCCATGTTGCCCAGGCTGGC
CTCAAACTCCTGACCTCAGGTAATCCTCCTGCCTCAGCTCTTCCCAAAGTGCTGGGATTA

FIG. 10
(continued)

TAGGCATGAGTCACCGGGCCCAGACTCAATCTTCTGACAAGCTCTCAGAGAGAGTAAAAA
GCAAATGAATATTTCATTATTTTGATCTGAGCTTTACGATTTTTCTTTTCTTTTCTTTTT
TTTTTTTTTTTGAGATGGAGTTTTGCGTTGTTGCCCAGGCTAGAGTGCAGTGGTGGCGAT
CTTGGCTCACCGCACCCTCCGCTTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCTG
AGTAACTGGGATTACAGGCATGCGCCACCATGCCCGGCTGATTTTGTATTTTTAGTAGGG
ACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTTAAGCTCCCGACCTCAGGTGATCCACCT
GCCTCGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCACCTTGCCCAGCCTTTTTTT
TTTAAATCTGAGAAGAGGTCTTGCTCGATTGCCTAGGCTGGAGTGCAGTGGTGCGATCTC
TGCTCACTGCATTCTCTGCCTCCCAGACTCAAGCAATCCTCCCACCTTAGCCTCCTGAGT
AGCTGGGACTACAGGCATATGCCACCACACCTGGCTAATGTTCGTATTTTTTTGTAGAGA
CAGGGTTTTGCCATTTTGCCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCTCCCA
CCTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACTGTGCCTGGTCTCCTTCAC
TGTTGTAAGATACTTGAATTGGGTCAATATTTGTGGAGAAGTCTCTTAAAAGTTCACTTG
ATTGTCAGTACTAGAACTCTACATTTAATATTGACATATTCCTGGGAGCATTTCAGAGCA
TTCTATTAGCTTAGAAAGGTCCAGGATAATTTGACTTTAGAAGTTACTGTTACCATGAAT
CTCAATGACTTTTGAAATCCATGAAGAATATCTTTTTTTTTTTTTTGAGACGGAGTCTCA
CTCTGTCGCCCAGGCTGGAGTGCAGTGGTGATCTGGGCTCACTGCAAGCTCCGCCTACTG
GGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACATGCCAC
CACGCCTGGCTAATTTTTTTGCATTTTTAGTAGAGAGGGGGTTTCACTGTGTTAGCCAGG
ATGGTCTCGATCTCCTGACCTTGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGCGCCTGCCCAAGAATATCTTTTTGCTGGTAACTAGAGAGGACT
CCTCTGAAGCAGATGCCATTCATGATGGATTTCATCATTTATGGGTTTTAAAAAACATTT
TATTTTGAAATAATTTCAAATTTAAATAAGAGTTGCAAAATAGTACAAATAATTCGTGTT
AACTTTTCATCCAGATTTACAAGTCAACCTTATACAGGTTGAGTATCCCTTATCCAAAAT
GCTTGGGACCAGAAGTGTTTTGGATTTCAGATTTTTTCGAATTTTGGAATATTTTTATTA
TATACTTAAGCATCTCTAATCCCCAAATCTCAAATCTGAAATATCTGAAATGCTATGATG
AGCATTTCCTTTGAGTGTTATGTGGGCACTTTTTAAATTTATTTAATTAATTTATTTTTT
GAGATGGAGTATTGCTCCATCACCCAGGCTGGAGTGCAGTGAGCGATCTTGGCTTATTGC
AAACTTCACCTTCTGGGTTCAAGTGATTCTCCTGCCTCAGCCCCCTGAGTAGTTGGGACT
ATAGGCGCTTGCCACCACGGCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCAC
CGTGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGGTCCACCTGCCTCCGCCTCC
CAAAGTGCTGGGATTACAGGAGTGAACCACCGCGCCTGGCCATGGATTTTGCAGCATTTT
AGATTTGGGATACTCAACCTGTACCATGTTTACTCTCTCTCCTCTCTCTCTCTCTCTTTT
TATATATATATATATATATATATATATATATATATATATATATATATATAAATTATATATAC
ACTACACATATATGTATGTATATGTATGTATTTTATATATAAAATACATATCTACATATA

FIG. 10
(continued)

AAATACACATGTATATATACATGTGTACATATATGTGTCTCTATATTTAAGTTTTGTTGG
AACCACTTGAGGGTAAGTTGCAGACATGGCGTCTCATTGCTCCAAAATACTTCAGTGTGT
ATTTCTTAAATACAAGGACACTTGGTTACATAACCACAGTATATCACCAAATGTATATTA
TAACAAGACTACCATCAAATCCTTATATCTCTTTCAAATTGTTTTAGTAATATCCTTATA
GCAAAAGACAAAACAACAACAAAAACTGTTCCCTTTTATTTTGTTTGTTTTGGTCCATTA
TATGTCCAGGTTATGCATTAATGCATTGTGTTACTTGCTAAGTCTTGTTACTGGCCTTTA
ATTAGGATATTTCTTTGCATCCCGCCAAACTCCTCTTCATGGTTGTATCTTTTTTTTTTT
TTTTGGAGATGGAATTTTGCTTATGTTGCCCAGGCTGGAGTATAATGATGCGATCTTGGC
TCACTGCAACCTCCGTCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAAC
TGGGATTGCAGGCCTGCGCCACCTTGCCCAGCTAATTTTGGAATTTTGTGAGACGGGGTT
TTGCCATGTTGGTCAGACTAGTCTCGAACTCCTGACCTCATGATCCGCCCGCCTTGGCCT
CCCAAACTGTTGGGATTACAGGTGTGAGCCACTGTGCCCGGTCTTTTTTTTTTTTTTTTT
GAGACAGGGTCTTATTCTGTTGCCTGGCCTGGAGTGCAGTGGTATGATCTTGGCTCACTG
CAACCTGGACCTCCTGGGCTCAGGCGATCCTCCCACCTCAGCCTCCTTAGTAGCTGGGAC
TATAGGCACACACCACCATGCATGGCTAATTTTTATATTTTTTTGTAGAGACTGGGTTTC
GCCATGTTGCCCAAGCTGGTCTTGAACTCCTGGGCTCAAGTGATCCACCTGCCTTGGCCT
CCCAAAATGCTAGGATTACAGGTGTAAGCCACTGCGCCTGGCCCTAATTTTTGCATTTTT
TGTAGAGATGGGGTTTCACTATATTGCCCAGGCTGGTCTTGAACTCCTGGGCTCAAGTGA
TCTTCCCATCACAGCCCCCTAAAGTGCTGGGATTATAGGCGTGAACCACTGTGCCTGGCT
GAGGATTAAGTTTCAACCTCAGGGGAGCGGCATTCAAACTATAGCATTGTCCTTTAGTGA
CTGGCTTAGTTCACTTAGAATGTTTGTCTATTCATCCATCTATAGACACTGTTTTCTTTC
ACCTTTTGGCTTTGCAAATAATGCTGCTGTGAATATGAGTTATAGAAAAATACCAATTTG
AATCCGTGTTTTCAATTACTTTGAGTATATACCTGGAAGTGGAATTTCTGGATCATATGG
TACTTCCAAGTTTTTTTTTTTCTTTTTTGAGACAAGGTCTCACTCTGTCACCCAGGCTG
GAGTGTAGTGGCACGATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCT
CCTGCCTCAGCCTCTCAAGTAGCTGGGATTACAGGCACGCGCCACCACGCCCAACTAATT
TTGTATTTTTAGTAGAGATGGGTTTCTCCATGTTGGTCAGGCTGCTCCCGAACTCCCGAC
CTCAGGTGATCTGCCTGCCTCAGCCTCCCAAAATTCTGGGATTACAGGTGTGAGCCACCG
CACCTGGCCTCCATGTTTCAATTTTTAAACAAACAATTAGTTAAAAAAATAGGAAACTAA
GAGAATGAACTATTTCCTGTTTTATTCAGTGGGTTATAATCTGTTACTATCATTGTTTAT
TTTGAGGTACAAATTGTCCCTACTTTGGCCAGCAGAGGATCCTGCAGTTTGTCTCCTGTG
TCCTTTTCATAGCTCCTTGTTGGAACTCTTACTGGCCCACAATAGGATGTTCCAAGTTCA
TCTTCTTACTTTTACTGCCCCAACGCTGGGATCAGCCATTTCTTCAAGGAGGCCAGTTCC
TTTCATTGGAGAATGGAAAACCCAATATGTAGAAACCAAGATAGAGGTGTTAGGTGTGAT
TGCTACTGGAGTGTCATTGCTTCCAAACCCTTTCAGAAGAGACCTAGGAAATGTGTGTGT

FIG. 10
(continued)

GTGTGTATATATATATGTGTGTGTGTGTGTATTCATAAAAGCACATACACATACACAT

ACCCCGAAGCATGTATTTCTGTATTATTATTATTTTTTTGAGATGGAGTCTTGCTCTGTC

GCCCAGGCTGGAGTACAGTGGCACGATCATGGCTCACTGCAACCTCTGCCTCCTGGATTC

AAGCAATTCTCCTGTCTCAGCCTCCTGAGTAGCTGGGATTACAGGTGTCCACCACCACGC

CCACCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCACATTGGCCAGGATGGTCT

TGAACTCCTGACGTCAAGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTATAGG

CGTGAGCCACTGTTCCCATCCAGAAGCATACATATCTATTTCTATATCTACATTTCTGTC

TTTACATGTATATATTAAAAATTACAGTTTGCACTAATACCTCCAATTACAATCTAACAT

CATGGGATTTATTCTGGCTTTCTCCCTTCTCATATTTGTGTCTCCCCAACAGTGAGAAAC

CTGGCTTGCTATCCTCAACATGGTAACTTATTTATTAAGAAACTTATTCTTTTTTTTTTT

TTTTTTCTGAGATTGAGTTTCGCTCTTGTTGCCCAAGCTGGAGTGCAGTGGTGTGATCTT

GGCTCACCGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCTTCTCAAGT

AGCTGGGATTACAGGCATGCACCACCATGCCCAGCTAATTTCGTATTTTTAGTAGAGATG

GGTTTCTCCATGTTGGTCAGGCTGCTCTGGAACTCCCGACCCCAGCTGATCTGCCTGCCT

CGGCCTCCCAAAGTCCTGGGATTACAGGCGTGAGCCACCGTGCCCTGCCTCTAGTTTATT

TATTTTTATTCCATGTGCTCAGTCTTGCGAGCACGTGGTCTGTTTTCTTGGGCCTGGCCC

CCTCAGTGCACTGTCTTAATACCCTAGCCCCCAGTCCCTCTGATCATATCCCCAGACACC

CCTACTGAATCCCAGGTCTCTACCAAGGGAAAGGCAGGGAGGAGGCATTGACCAAGGAGA

AGAGGGGGAAGGGACAGGGAAGGTCTTGATTTGTATTTTCTAAAATTTTCTACTCTGCTC

ATAATGCGTCTTAGCTGTGTTGTTGTGGAAAGTAGTGCTGACAGTGTCTTGTTTTTTTAT

TACTTACTTTGTCTTTCTTTTTAAGATGGTTTCACCCAAATATCACTGGTGTGGAGGCAG

AAAACCTACTGTTGACAAGAGGAGTTGATGGCAGTTTTTTGGCAAGGCCTAGTAAAAGTA

ACCCTGGAGACTTCACACTTTCCGTTAGGTAAGTTGGAATGAAAAGAGAGGATCCTGAGA

GTGTTTTCTAGGTAGGAAGTGGTAAAACCATGCTTGGATAGCTTGCTGCCTGCATTTCGA

GTTTGAAGGCCTTATCTGAGCCCTGGGCTGCCTTCAGGGTTTGGGGAGTGGCCTCCTGGA

CATTTAGCAGAAGAGGAGTAAGGAGGGCCCTTCTTCTCCCTCTGAGACCTCATGGAAGGT

GAGTTGGAGCAGGTCATAGAAGTTCTTAAGCCCTCCAGTGCTTGAGACTTGTTCCACACA

TCTTGAACCTGGTTTCTGCATTTTTCTTTTCCTTCCTGTTGATTTATTTAAAAATTTTAT

TTCTTTTCAATTTTTTTTTTTTTTTAAATAGAGGTGGGATCTTCCAATGTTGGCCAGGTT

GGCCTTGAACTTCTGGCCTCAAGCAATCCTGCCTCGGCCTCCCAAAGTGTTAGGATTACA

GGCGTGAGCCACTATGCCTGGCCTTCTTTTTTTGAGACAAGCTGTTGCTCTGTTGCCCAG

GCTGGAGTGCAGTGGTACGATCACAGCTTACAGCAGCCTTGAACTCCTGGGCTTAAGTGA

TCCTCCCGCCTCAGCCTCCCGGGTAGCTGGGACTCCAGGCTTGTGCCACCATGCTCAGCA

TTTTTAAAAAATATTTTTTGTAGAGATGAGGTCTCACTGTATTACCAAGGCTGATCTTTA

ACTCTTAGCCTCAAGTGATCCTCCTGCCTCAGCCTCCCAAAGTGTTGGGATTACAGGCAT

FIG. 10
(continued)

GAGCCACCACACTCAGACTTTGTTGACTTCTTAATAAGAAAAATACTTGTTAAGAGTTTC
TTCAGATCACTTTCCTTTATCAACAAGTAAAACATGACTGAGGAAGTTGTGGTCCCCTTT
GCTTCCCTGCCCAGGCCCGTTTCCCTCCCTCTTTCCCCAGAGGAAACCACCAAGAGGTTG
GCATATATTCTTCCTGAACGTGTTTTTATAGTTGTACTGCACTTGTACTGTGTATGAACA
ATATAAAGTTGGTTTGTGTGTTTAAAAAATTCACATACATGGATTTATAATGTATGTATC
ATTTTGCAACTTAAAAATTTTTTTTTGAGCTCCATGCTGATTGATAACGATCTATTTTTT
TTTTTTGAGATGGAGTTTCAGTCTTATTGCCCAGGCTGAAGTGCAATGGCGTGATCTCAG
CTCACTGCAACCTCAGCCTCCTGGGTTCAAGCTATTCTCCTGTCTCAGCCTCCGGAGTGG
CTGGGATTACAGGTGCATGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAGATGG
GGTTTCACCATGTCGACCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCTGCCTGCCT
TGGCCTCCCAAAGTGCTGGAATTACAGGCATGAGCTACCATGCCTGGCCTTTTTTTTTTT
TTTTTTTTGAGACAAAGTCTTGCTCTTTTTCCCAGGCTGGAGTGCAGTGGCCACAATCTT
GGCTCACTGCAACCTCTGCCTCCTGAGTTCAAGCAGTTCTCCTGCCTCAGCCTCCTGAGT
AGCTGGGATTACAGACATGTACCACCATGCCAAGTTAATTTTTGTATTTTTTGTAGAGAC
TAGGTTTTACCATGTTGGCCAGGCTGGTCCTGAACTCCTGACTTAAAGTGATCCATCTGC
CTTGGCTTCCCAAAGTGCTGGGGTTACAGGCATGAGCTATCGCGCCTGGCCTGAGAAATC
TCATTCTTACTCCTACTCCCTTGCACACTATCTCCATTCTGTAGGTAGCCATTTCTATTA
ATTTCTTGTTTACCCTTCTGTGTTTCTTTCATTCTTTTTCTTTTTTTCTTTTTTTTTTTT
GAGACAATCTTGCTCTGTTGCCCAGACTGGAGTGCAGTGGTGTGATCTTGGCTCACCGCA
ACCTCCACCTCCTGGGTTCAAGTGATTTTCATGACTCAGCCACCTAAGTAGTTGGGATTA
CAGCGCCTGGTGTACACTACCACACCCAGCTAATTTGTGTATTTTTAGTAGAGATGGGGT
TTCACCATGTTGTCCAGGCTAATCTCCAACTCTTGGCCTCAAGGGATCTGCCTGTCTCAG
CCTCCCAAAGTGCTGGGATTATAGGCATGAGCCACCATGCCTGGCCCTATGTTTCTTTTT
ATAAAAATAAGCAAATTAATATTTTTATTACTATTTTCCTTTTATTTTTACACATCAAGT
AGAACATTAAATATATTTCTCTGTAATTTTTTTCAGTTACCTAAATCTTTTAGTGATCTC
TCTCATCTTTTTAATCAGCTGGATCGCATTCTATCATGTGAATATTTTATAACTTCTATA
TACTGTCACCAGCAGGTAGCGATTTAGTTGTGTCTAATATTTTAAAATGATATATAATGC
CTCAATGAATATAGTAACCTTTTGCATATATTGTTTTGTGCTTTGGGATAACACTACCTC
GTATTGGAAACTGTGTCATTACATGTGTCTTTAAAATTACATGTGTCTTTTTATTTTTAT
TTTTATTTTTTTTGAGTGGGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAGTGGTGAG
ATCTCGGCCGACTGCAACTTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
CCAGTAGGTGAGATTACAGGTGCCTGCCACCACGCCCAGCTAATTTTTGTATTTTTAGTA
GGGACGGGGTTTCACCATGTTGGCCAGGCTGGTATCGGTCTGCTGACCTCAGGTGATCCT
CCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGACGTGAGCCACCATGCCTGGCCATCA
CTTTTTTTTTTTTCTTAATTGCTGCATAGTGGCCGGGCACAGTGGCTCACGCCTGTAATC

FIG. 10
(continued)

CCAGCACTTTGGGAGGCCAAGGCAGGCGGCGGATCATGAGGTCAGGAGACCAATACCATC
CTGGCTAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAAATTTAGCTGGGCGTC
GTGGCGGGCGCCTGTAGTCCCAGCTACTTGGGAGGTTGAGGCAGGAGAATGGTGTGAACC
CGGGACGTGGAGCTTGCAGTGAGCCAAGATTGCACCACTGCACTCCAGCCTGGGTGATGG
AGTGAGACTCTGTCTCAAAAACAAACAAACAAACAAAAAAATTGCTGCATAGTATTCCAT
TGTATGAGTAGTAACACAACAATTTTTATAATGCATAGTATTCCATTGTATGAATAGTAA
TGTAGCACTATTTGTTTATACATTTTTATGATTAAAAAACAAAATGTTTTTCTATTATGA
ATAAAGTGGCAATGAATATTTTTGTACAAGTGTTTTGGTAGCTATACAGTTATTGTCACT
TAATATATGCAATTCGATAGGCCAGTCATTCAAAATAGAAGATATACAAGGTAGGCCGGG
CGTGGTGGCTCACGCCTGTAATCTCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTG
TGGTTAGGAGTTTCAGACCAGCCTGACCAACATGGAGAAACCTCATCTCTACTAAAAATA
CAAAAGTAGCTGAGCGTGGTGGCGCATTCCTGTAATCCCAGCTTCTTGGGAGGCTGAGGT
AGGAGAATCACTTGAACCTGGATTTATAATGTATGTAAATCCACCGCGAAGGTTGCGGTG
AACCGAGATCACGTCATTGCACTCCAGCCTGGGCAATAAGAGCGAAACTCCATCTCAAAA
AAAAAAAAAAAAGATATGCAAGGTAAAGATACTAATAAAGACCTTTGTGTTGAGTTGGTT
GACATGTGGTTATTTCACCCATCGTATTTCTTATAGGGAATAGGTAAATTCGTTCCTTGG
GTTTCTTTCAACACTTAGGTAAAATCCGACGTGGAAGATGAGATCTGATTTTACTGGTGT
AACTCTTTATTTGTCCCCTTGCCTCCCTTTCCAATGGACTATTTTAGAAGAAATGGAGCT
GTCACCCACATGAAGATTCAGAACACTGGTGATTACTATGACCTGTATGGAGGGAGAAA
TTTGCCACTTTGCCTGAGTTGGTCCAGTATTACATGGAACATCACGGGCAATTAAAAGAG
AAGAATGGAGATGTCATTCAGCTTAAATATCCTCTGAACTGTGCAGATCCTACCTCTGAA
AGGTCAGTAACATTTTAGTGACCACAAAGTCTGCTGCTCCCTTGTGCCCTGAGTGTCAGA
AATGCATGACGGTCTGTGTATGACTCTCTGACTCCAAAGGCTTGTGACTGTTTTTTGAGC
TGTAATCTTTAAAGAATTACTAAAGTGAGACTAATAGCATCAAATTATTTTCAGAGTACC
TTTTTCCTGCAAAAGTTTTAATCAGTGTTACTTACACTCATCCTATAGGGGTTGCATACC
ATTCCTGCATATACTTGGTACGTGTATTAGTTTTAAGACTTATTGAACTTCAGCAGATAA
TCTTTGAGAGTTATTAGAGGAAAACAAATGATAATGGAGACACCAAAATAGCAGCAGTTT
TCTATGGTGGCTCTCGACCAGTTATTCAGCAATGTCACCAACAGATGTCAGTTTAAGCTC
AGAAGTGGAAAAGCAGAGAGCTCAGAGGGTCAGCTTTTTCATCAGTTCTTTTAATGTTAT
CACCACAATTATGTGAGAATGACCTTGCTTAGAGAAAATTATGTTATTTTCGAGATCTTT
CCCCCTGTGTTGGAACTAGGCTGATGAAAGCATGGGCTTGACTTATTTATTGATTGTATT
CGTTTTGTACATTCCCAATCTCCTCTCTGACTTGGTGCAAATTCAGGATCTCTTAGTTAG
TTTGTATATTTTGTGTCTTCAGGTATGATTTTTTCAGCTTATACCTTTATGTCAGTGCTA
TTATGTGCTGATAATTTGTTTCTCTAGCTACCACCGTAGCTTCAGGCAAAAGGCTGTCAG
CCAACTCTGTACAGTTTATTTCTAAATTTTACTGTTTTCAGTTGAGTATGGATGAAGAAT

FIG. 10
(continued)

AACTCAAAGTTTATTCTTTTGATGATGAGCCCTTAACACCACCTGCCATGATAGTACTTG
CTTTCTGACCAAGATCCTGAGGGAAAAAGCCACTTTATTATTAGAACTATGTTAAGATGC
TTCCCAAAAAACATGGAGCAGTATTGTCTCAAAGTCTGTCCTTGGATGGCTTTGGATGCC
TACATCAGGACTGTCTGATGTGCTGGTTAAAATGCAGATTCCTGGGCCTCATTCAGACTT
ACATGTATTGATATTGCTGGTTGTGGAGCCTGGGAATTCATATTTTTAGCAAAATCCCTC
ATTTTTACTCCAAGTCTTATGTGCATTATACAGTTTGAGATGATCACCCAGGATATAGTC
CAAAGACACTGGAGGCTGTTGAAGTATAGGTTGTATATATGGAAAAGGTTGGAATGTTTG
AATTAATTTATAATGAAGATCCTTTTTAATTGAGTGTTCACATGCCAAGGCAAGGACAAA
CATTCAAAATGATTTTCTGTCTCTGTTACAACTTTTTCTTTCTTTTTTTTAATTTATTTA
TTTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTCAAGTGACGCGATCTCGGCTCA
CTACAACCTCCGCCTCCCAGATTCAAGTAATTCTCTTGCCTCAGCCTCCCGAGTAGCTGG
GACTACAGGCATGTGCCACCATGCCCAGTTAATTTTTGTATTTTTAGTAGAGACAGGGTT
TTGTCATGTTTGCCAGGCTGGTCTCAAACTCCTGAACTCAGGTGATCCGCCCACCTTGAC
CTCTCAAAGTGCTGGGATTATAGGCGTGAGCCACCGTGCCTGTCTCTATTACAACTTTTT
ATTACAACTTCTTTATTTTGACTTTATTTTTACAAATTATTTATTTATTTTTTTTGAGAT
GGAGTTTCGCTCGTCACCCAGGCTGGAGTGCAATGGTGCGATCTCAGCTCACTGCAACCT
CCGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGG
CACTTGCCACCACACCCGGCCAATTTTGTATTTTTAGCAGAGACAGGGTTTCACCATGTT
GGTCAGGCTGGTCTCGAATTCTTGACCTCAGGTGATCCACCTGCCTCGGCCTCCCAAAGT
GTTGGGATTACAGGCATGAGCCACCACGTCCGGCCGACTTTTATTTTTTTTTCTTGAGAC
AGGGTCTTGCTCTGTCACCCAAGCTGGAGTGCGGTGGCATGATCATAGCGCACTGCAGCC
TCGACCTCCTGGACTCAAGTGATCCTCCTGCCTCGGCCTTGTGTATAGCTGGGATTACAG
GCAGTTGCCACCATGCCAGGCTAATTTTTAATTGTTTTGTGAAGATGGGGATTTCACTGT
GTTGCCCAGACTGGTCTTGAACTCCTGGCCTCAAGTGATCTTCCTGCCTTGGCCTTCCAA
AGTGTTGGGATTACAGGCATAAGCCACTATGCATGGCCTGTAACTTCTTTAAATGGCTAT
AATTAAACAGTTGGTCCTTTTAAGATTGGGCAATGGACGAATGGCAAATTGCATTTTTAA
AAGAGGAGGGATTTAAAAAAAAACAGGAAAGATTGGGGCATTTGTCTCTAAAGGACTGTG
GACTCATTTAAGAAGTTTAGTGGTCATTCTTACCATCTTTGTGGTTTTTCCTGCCTGCAT
GGGATGCAGATTTTCTGTCTCAGGTGGGATTGATCAATCCCTTGGAGGAATGTGTCTACT
TTTTAATTGTGTTTAGGAGAGCTGACTGTATACAGTAGTTTTGTGAAAGAACAACATGAA
CCCATAGTAGAGCTAAATTCTTTTTTATTTTTTAAAAACTTTAGGTGGTTTCATGGACAT
CTCTCTGGGAAAGAAGCACAGAAATTATTAACTGAAAAAGCAAAACATGGTAGTTTTCTT
GTACCAGAGAGCCAGAGCCACCCTGGAGATTTTGTTCTTTCTGTGCGCACTGGTGATGAC
AAAGGGGAGAGCAATGACGGCAAGTCTAAAGTGACCCATGTTATGATTCGCTGTCAGGTA
AATCTCCAGTTGAAAAATGGGTCTGGCAAGATGTTACCTTTGGGTGATTTTTCTGCTGAC

FIG. 10
(continued)

AGAAGACAGACACCATTACATTCAAAGTCAGATTGTCTTTTATTTATTTATTTATTTATT
TATTTATTTGAGACAGGGTCTTGCTCTATCACCTACAGATGGGGTTTCACCACGTTGGGT
CTGGTGACCCAAATCTTTGGGTGATTTTTCTGCTGGAAGAGGACAAACACCATTACATTC
AAAGTCAGATTTTCTGTTTTTTTTTTTTTTTTGTTTTTGTTTTTTTAATATTCATTTGTT
TATTCATTTGAGACTGGGTCTTGCTCTGTCACGCAGGCTGGAGTGCAACCTCCCTGGGCT
CAGTTGATCTTCCCTCAGCCTCTTGAGTAGCTGGGACTACAGGTGTGTGCCACCATGCCC
AGCTAGTGTTTGTATTTTTTGTGGAGATGGTGTTTTGCCGCATTGCCCAGTGTGGTCTTG
AACTAGTGCTCAAGAGGCCTGCCTCCTTCAACCTCTCAAAGTGTTAGGATTACAGATGTG
AACTACTGTGCCTGATCCAAAGTCAGATTTTCTTTGCTTACTTAGTCAAGTTCGTCTATG
CTTTTATTATACTTAATATATTAGTATAGTTACTGTATTAGTATATTAGCATATTTAATA
TATTATTATACTTATCATACTTGAGTATATTGAGTATATTTACACTTTTAGTATATTTGT
ATACACACACCACATTTTTATTATTTATCTTTTTTTTGAGACAGAGTCTCCCTCTGTCTC
CCAGGCTGAAGCACAGTTGGCTCACTGCAACCTCTGCCTCTTGGGCTCAAGTGATTCTCG
TGCCTCACCCTCCTGAGTAGCAGGGATTACAGGTGTCCACCACCAAGCCTGGCTAATTTT
TGTATTTTTAGTGGATATGGGGTTTTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGAC
CTCAAATGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGAATTACTGGCGTGAGCCACTG
CACCCAGCCTATTATCTGTCTTTTGATGGACATTTAAGTTGTCTCTATATACTAGCTATT
GTGAATAATGCTGCAGTGAACATGAGAGTGCTTGAAAACACTAATGTAACATAAAGGTAA
CAAATAATAAATGTCATGTGTTTATCTTGAAAGCAACTGAAATACGACGTTGGTGGAGGA
CAACGGTTTGATTCTTTGACAGATCTTGTGCAACATTATAAGAAGAATCCTATGGTGGAA
ACATTGGGTACAGTACTACAACTCAAGCAGGTGAGCAGATTGGAAAGCTCAAGCTTTCTC
CTTAAAAACTTAAAACAAATCCTAATAGAGAATTTTGCAAACATACAGAGGTAGACAGAA
TAGTATCATCAGCCTCCATGTACCCATTGCAGCTTCAACTATCAAATCTTTTTTTTTTTT
TTTTTTTTTGAGACAGTCTTACTCTGTCACCCAGTCTGGAGTACAGTGTTGCAATCTTGG
CTCACTACAACCTCTGCTTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAG
CTGGGACTACAGGTGCCCACCACCATGCCCGGCTAGTTTTTGTGTTTTTAATAGAGATGG
GGTTTCACCATGTTGGCCTGGCTGGTCTTGAATTCCCGACCTCAGGTTTTCTGCCCGCCT
TGGCCTCCCGAAGTTTTGGGATTACAGGCGTGAGCTACCACGCCCGGCCCTAAATCTTTT
CTTATTATGATTCCACTCACTGACTGCCGCTATAGTACTTGGAAACATATTCCAGATTTA
TATTATTCCCATATTTATCTGTAAAAGGCATTACAGAGGTTCTTTTTTTTTTTTTTTTTT
TTTGAGATGGAGTTTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGTGTTCTTGGCTCA
CTGCAACCTCTGCGTCCCGGGTTCAAGAGCTTCTCCTGCCTCAGCCTCCTGAGTAGCTGG
GATTATAGGTGGTGCCACTACACCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTT
CACCATGTTAGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCTGCCTGCCTCAGCC
TCTCAAAGTGCTGGGATTATAGGCATGAGCCACTGCATCTGGCCTAAGGCTGTACAGAGT

FIG. 10
(continued)

TTTAAAGCAAGTTTTCATTATAGATCCACTTCTGGTTACCTTTAGGTAACCTCACTTATT
CACTTTGGCATTGTTGCTATTTCAAATTTCACCTTTATGATAGTGGAAAATGATATAATC
TCTCTAAATAATGTGGTCTATTCATAAAGAAAAATAGGCTTGAATTTATATCAGCAGAGT
AAAGTGTATGTGAAGACTGAAGAAAGATACATTTTCTGGCTGAACAGAAAACACGGTGAA
ACGATTTGAAAACTTTTATTGTGAATTACAGGGTCCTATGAACCCTCTGTCCGTGCCTTT
ATGAATATCAACATAGACATGTTTTTTTTTTTTTTTTGCATTAACACCGTTTTCTGTAA
TATTTTCTTTATTTTACATCAACTGCTGTACTCGATCAGCCCCTTAACACGACTCGTATA
AATGCTGCTGAAATAGAAAGCAGAGTTCGAGAACTAAGCAAATTAGCTGAGACCACAGAT
AAAGTCAAACAAGGCTTTTGGGAAGAATTTGAGGTAAGTTATTAAAAAACTGTTTTTACG
TGAGTTGTTATATCCTATTTTTAGTGGAGGAGAAGTTGCTCTTGTGTTTGGAATTGGACC
TGAGAGACTTGAAACTGACGTCCTTTTTTAATTCGGCCATTGATTGACACGGAGCAAGTT
GCTGAGAGGGCTTCTTCGAAACAGAAGAGCATTGTGTTCTGAGGGAAGGGAGTTGGCAGT
GAGTAGTCAATGGATGTGCTAGCCGCTCCATTTGGCTCTTTTGGTTTGGACTGGTGGCAA
AATCTCAGAGAAACAAAAGGATCTAATTTCTTCGAAAGATTTCCAGCATGCACTGGGGTC
TTTAGAAACAATCTATAGCCTTAGTGCAGCAAATGAGTATGAGTAAAAGAGAAACACCTT
GTGGTGGCTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCGCTCTGTCGCCGAAGCTGGAG
TGTAGTGGCGTGATCTCGGTTTACTGCAGCCCCGTCCTCCCTGGGCTCAAGTGATCTTCC
CATCTCAGCCTACTGAGTAGCTGGGACTACAGGCACATGCCCCTATGCCTGGCTAATTTT
TGTATTTTTGGTAGAGATGAGGTTTTGCAGTGTTGCCCAGGCTGGTCTTGAACTCTTGGG
CTCAAGTGATCCTCCTACTTAAGCTTCCCGAGTAGCTGGGACTACAGGCACACGATACCA
TGCCCATCTAATTTTTGTATTTTTTTGTAGAGATGGGGTTTTGCAGTGTTGCCCAGGCTG
GTCTTGAACTCTTGGGCTCAAGTGATCCTCCAGCTTTGACGTGCCAAATGTGGTGGCTTT
AATTTCAGAGTTCAAATTGATAACTCTGGTAAGTTAAGTGAACTGATTTCTTTTTTTTTT
AAATTATTTTTGTTGATTATACTTTAAGTTCTGGGATATATGTGCAGAACGTGCAGGTTT
GTACATAGGTATACATGTGCCATCATGGTTTGCTGCACACATTAACCCATCATTTAGGTT
TTAAGTCCTGCATGCATTAGGTGTTTGTCCTAATGCTCTCCCTCCCCTTTAATGCATCAG
TGAAAAAGTGATGATAGGCTGGGCGTGGTGGCTCACTCCTGTAATCTCAGCACTTTGAGA
GGGTGAGGCAGGTGGACCACTTGAATCCAGGAGTTTGCCCCCATCCCCAGACAGTGTGTG
TGATGTTCCCCTCCCTGTGTCCATGTGTTCTCATTGTTTGGTTTTCTGTTCCTGTGTTAG
TTTGCTGAGAATGATGGTTTCCAGCTTCATCCATGACCCTGCAAAGGACATGAACTCATT
CTTTTTTTATGGCTGCATAGTATTCCATGGTGTGTATGTGCCACATTTTCTTTATCCGGT
CTATCATTGATGGCCATTTGGGTTGGTTCCAAGTCTTTGCTATTGTAAATAGTGCTGCAA
TAAACATATGTGTGCATATGTCTTTATAGTAGAATGTTTTATAATCCTTTGGGTATATAC
CCAGTAATGGGATTGCTGGGTCAAATGGTATTTCTGGTTCTAGATCCTTGAGGAGTCACC
ACACTGTCTTCCACAATGGTTCAACTAATTTACACTCCCACCAACAGTGTAAAAGCATTC

FIG. 10
(continued)

CTATTTCTCCACATCTTCTCCAGCATCTGTTGTTTCCTGACTTTAAGTGAACTGATCTCT
TTCCTGAAACTAACTTGGGTTGGAGAATGTCCCTGATGGGAATGTGCTGTGTTCCCATTG
CACTCTTCTATATCACTTACCCATTGACAATGTGATCTCTTTCATTTTCTCCTCATCCAT
TTGACAGAAAACTTCAAAAACAAGGATTCTGGCATATTTACCTTTGCAGTTGTCCCCAGC
ATGTAGCACGGTGCCTAGTACACAGAAGAAACTCCATAAATGTTTGTTGAATGAGATTTA
CATTTAACTCATGTTTACATCATTTTATTTTCCTGTTCTGTTTTATGGGAATGATTATTC
TATGCTTTTTGAGGACTACAATTTATAAATATTTGTGGATTGAATGAATAAGTGAATACT
GGGCAAATAAAGTCCTTTTAGCCAGAGTATGTCTGAACAACTTGCTGAGATAGATATGAT
TTCCCATTTTCCAGCTGAGGGGCCTAAGGGAGGTTAAGTAAATTATTCAATCTTCATACC
ACAGTTTTTGTTTTGTTTTGTTTTGTTTTTTTTCCTCCTGAGACAGAGTCTCACTTTGCT
GCCATACTGGAGTACAGTGGTGCAATCATAGCTCACTGCAGCGTCCAACTTCTGGGCTCA
CGCCATCCTCCCACCTCAGCCTCCTGAGTAGCTGGTACTACAGGTGTGCACCACCATAGC
CGGCTAATTTTTCATTTTTTGTAGATATGGGGTCTCACTGTGTTACTCAGGTTGGTCTTG
AACTTCTGAGCTCAAACAATTCTCCTGTCTTGGCCTCTCAAAGTGTTGGGATTACAGGTG
TGAGCCACTGTGCCCGGCCCATACCACAGATATTGATTGAATTCCAGCAGTGGGGAGGAG
TGTGGAATAGAACATTCTCAGTCCTTGCTCAACATTACTGAACAGAGACTTGAATTTGAG
TTTATTCTCTCATCCCAGGCTTCGCGTTAGGCTCTGAAGACACTAGTGAACAAGACAGAC
AGGGTTACTGCCTTTAAAGGGAGCTTTTAGTTGAGAGAAGGAAAACAGTGATGAAAAGCA
TCAGTGAAAAAGTGATGATAGGCTGGGGCGTAGTGGCTACTCCTGTAATCTCAGCACTTT
TAGAGGGTGAGGCAGGCAGCTCACTTGATTCCAGGAGTTTGAGACCAGGCTGGGCAACAT
GGTAAAACCCCGTCTCTACAAAAAATACAAAAAGTAGCTGGGTGTGGGGGTGCGCACCCA
CAGTCCCAGCTACTCTGGGGGTTGAGGTGGGAGGATTGCTCGAGCCTGGGAGATTGAGGC
TGCAGTGAGCTGAGATCACGTCACTGCTCTCCAGCCTGAGCAACAGAGCCAGAACCTGTC
CCAAAAAAAAAAAAAATTGATGATAAACATAGTGAGACAGAATTTTGAAATCTCAGCCTC
ACTGTTGCCTTCCTTGTCCCCTGCCTGCCTAAATAATAAAAGGCAGCATTTCAGCAGTCA
TTCATTTCATTACTTTCACTTCATTTCACCTTCATAAAGCCTCATGAGGTAAGATGGGAA
GATACAGAAGTTTTAGAAACCGCTCATCAAAATTGAATGGAAAGCCGATTGTTCCAAAAC
TTTTTAGTGTGGAAAATTTCTATTATATGCAAAAGTAGAGAGAATGGGATAGTTATAGCA
GTATACCTGACACCCAGCATTAACAACTGTTGATAATATGGCCAATCTTTTTCGACTCTG
CCCCACTCACTTCCCCAGCCCTGACTTGTCTTGAAGCAAATACTTTTTTTTTTTTTTTGA
GATAGAGTTTTGTTTTGTTTTGTTTTTTGTTTTTGAGATGGAGTCTCACTCTGTCCCCCA
AGCTGGAGTGCTGTGGCTTGATCTTGGCTCACTACAACCTCCGCCTCCTGGGTTCAAGTG
ATTCTTGTGCCTCAGCCTCCTGAGTAACTGGGATTACAGGTGTGTACCACCATGCCCAGC
TAATTTTTGTATTTTTAGTAGGGACAGGGTTTTCACTATGTTGGCCACGCTGGTCTCAAA
CTCCTGACCTCAGGTGATCCGCCTGACTTGGCCTCCGAAAGTGCTGGGATTGTAGGTGTG

FIG. 10
(continued)

AGCCACTGCTCCCGGCCTTGAAGCAAATCTTAACACATCATTTCGTCTGTAACTATTTTA
TTTCAAAAAATTATAACCTGAATAGCATTATCATATCTAAAACTATTAACAGTATTTCCT
TAATATTAACACATATCAGTCACATTTTCCTGATTGCTACACACACACACACACACACAC
ACACACACACACTTGCAATTTGTGTTTTTTTCTTTTTAGATGGATCTCACTCTGTTGCCC
AGGCTGGAGTGCAATGGTGCATTCTCAGCTCACTGCAACCTCCACCTCCTGGGCTCAACT
GATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACCTCACCTGG
CTAGTTTTTGTATTTTTAGTAGAGGTGGGGTTTCACCATGTTGGCCAGGTTGGTCTCAAA
CTTCCGACCTCAGGTGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATG
AGCCACTGTGCCCAGCAGCAATTTGTTTGAATTGGGAGTGCTTTCTTCCACCTTGATTAT
GAAAAAATTTCAAATGTGTATAAAACAGATTCATATAAAGGATCCTGATATGCCATTATC
AGCTTTATCAATTATCCCTGTCATCATATTTTTTATTTATAAATATTTCAATATTTGTGG
AATCCTTAAAAATGCATCACATAACCCAACATTGTTCATATTATACCAATTGTCTTATAA
TTTAAAAATATTTTGTTCAATCATTTTTCAGATAAGCTTCACACACTGTGGTTGGCTAAG
TCTCATAATATTTCTGTTGTAAAAATCTTAAGTCTGGGCGTGGTGGCACACGGCTGTCAT
TCCAGCACTTTGGGAGGCTGAGGTGGGCGGATCACGAGGTCAAGAGATCGAGACCATCCT
GGCCAACATGGTGAAACCCGGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTAGT
GCGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGAA
GGTGGCAGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTAGAGACAGAGTGCG
GCTTCATCTCAAAACGAAACAAAACAAAACAATCTTAAGTCTCTTAGAATACTTTGATGC
CCCTTCCATCTCTCTTTTTCTGTCTTCCTTCCCCCTCTCCCTGTCTTTTCTGCTGTTGAA
GAAAGCAGATCATTTGTCCTGAGAGTTACTTATAGTCTGAATTTTGCTGAGTGCCTCTCT
GTGGTGGACTTAAGCATGTATCCATCCCTTATATTTCTTGTAAGTTGATATATCTAGAGA
CTTCATTGGATACAAGTTTTCTTTGGCAAGATAGCATGTATGGTGGTGTATCAGGAGGTG
TTTATGTCCTGTTGTTTCTTCTCTGATTTTCTTAGCAGCTCCTGATCATTATTACTTAGA
TCCATTAATTCATAAGGGACTATATGGTAGTGATATTGTAATTTTATCATTCTTCTTCAT
TTGTTAGGTTGGCATATTTCTATAAAAAGCTTTTCATCGCCGAGGGTTGATTTTTTCCTT
CTTACTAAGCAGTTTTCTTTTCTTTTTCTTTTTTTTTTTTTTGAGGTAGGTCTCACTGTG
TTGCTCAGGCTGGTGTGCAGTGGCGCAAACACACAGTTGCGAACTCTTGGGCTGAGGTGA
TCCTCCTGCCTCAGTTTCCTGTGTAGTTGGGACCACAGGTGCATGCCACCATGCCTGGCT
AATTTTTTGATTCTTTTGTAGAGATGAGGTCTCACTTTATTTCCCAGGCTGGTCTTGAAT
GTCTGGGCTCAAGCAATCTTTCTACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACAT
ACCACCATGCCCAGCTAATTTTTTAATTTTTATTTTTAGTAGAGATGTGGTCGTATTATG
TTGCTCAGGATGGTCTCGAACTGCAGAGCTCAAGTGATCCTCCTGCCTCAGCCTCCCAGT
GTGCTGGGATTATAGGTGTACTACAGGCAAGAGCCAATGAGCCTGGTCAGATTTTTTTTT
CCTGATTTGAAATCTGTTATGGGTTCAATTGATACTTCCAAATCAAACTCAGGGTTTCAG

FIG. 10
(continued)

GATTTTTACTAACCTCATTGATCTTACCCATGTATCTCCTTTCTCTAATGCCAAAAATCC
TACTTCTTGAAGCCATAATAAGATTATTCATTTGTTTTATCCCACATTACACACAACAAT
CTTAGAATAATGACTTCCCAATAATATGATTACTGAAAACAGTTTAATTTTTTTTGCGCT
TTTCAAAAAAATCCTTCAGAGATGTGTAGTCAAGTTACTGTATTCTGCTGGGCACAGTGG
CTCACGCCTATAATCCCAGTACTTTGGGAGGACAAGAAGGGAGGATCGCTGGACCTCAGG
AGTTTGAGACCAGCCGGGGCAATATAGTGAGACCCTGTCTCTACAAAAGAAAATTAAAAA
TTAACCAGACATGGTGGCATGTCCCTATAGTCCCAGCTATTGAGAGGCTGTGGCGAGAGT
AGGCTTAAGCCCAGGAGTTTGAAGCTGCAGTGAGATACGATTGTGACACTGTACTCTAGG
GTGACAGAGCAGGGACCCTGTTTTTAAAAAAAAAAAATGAAAAAACTTCCTGTGCCTTAG
ACTCATTTGTAATCGTCCTTCTCTCTGTGTGGCTATATGCTAACTGGGTATATGGTTAGT
TTATTTGTTTCATTTAAAAAATCTCTTTCTGTTAAGTTTTATTTATAATTACACAAATAC
TGGCTTTGATAGTCAAATTGAAAAAACAAAGTGTATTCAAAGAAGTCTACCTTCTATCCT
TGTCCTTTCCTATGTTTTAGCCATAGTATAAAAAGTTATGGTTTATCATTATATTTCAAA
AATATAAGAAGATATTCCCATATCCCACTTTTTCTTAAACAGTAGCATAACTTTACATAC
TTTTTTCTAACCTTGCTTTTTTAAATATCCTGGACATCCTGGATATCCATAATAGTGTCT
AGAGATAGTCTTCATTCTTTTTTTACTGTATAGTAATCCACTGTGTACTTGTACCATAGT
TTATTCAACCTATTGATGGGCATTTGGGTAGTTTCCAAATGTATCACAGAGAGGATTACA
GTGAATAGCCTTGTGTATGCATCCTGCTTTACTTTTGCTGACTACTGGTAATATTAACAT
TTTTTATGTTCTGTATTTAAAAAATGGTGGTTATTATTCATCTATAACTTTTATTATACA
TGACTTTGGTTAGCATGCTTTAACCTTTTAGCATAACATTTGCAAGCTACTTGTTTTAAT
TAAAATTTTGGTTAAATGTAAAAAATAGTGAGCTATTTTGTAATCTAGATTCAATAGAAT
CTTATACTTCCTTTACAAATGATAGCTGAGTTGATCATTTGTGTAAATGACTGTGAACTT
AAAAATTACAGCATTTTTTAAAATAAATTTTTTTAACATTTTAAAATTATTTAAAATAAT
AGACACACAAAGTAAAAAGAGAAGAAAAAAAAAAGAGACAGGGTCTTGCTATGTTGCCCA
GGCTGGTCTCAAACTCCCAGGCTCAAATGATCCTCCTGCCTTGGCCTCCTAAAGTGTAAG
CCACCACACTTGGCAAAAATTAGTTTCTTTAAAACAAAAACATTACAGGTTATCTGGTAC
CATGGTAGCTTCTTTAACACTAGGTTCACTTAGAACAAAGCTTAGGAACAAAGTCAGACT
TTCACAAAGAGCTTGTGTGGCAATGGGGTATTTTTTGCAAATTCCATTGGTGGGGTCAAG
ATGTGAGTTTAGAAGGAACTCTTAGCCTGACTCTTCTGGCCATGGAAAAAGATGGTTGCT
TCTAAATGCTGACCTGGTGATTTTACACTGTCACATCTCAAATTGTGGTCATCTTTTATA
CATTATTAACAACAAAAGGGAAAAATTGAGTTGACTTTAAGAGGAAGTGGAAAATAACGA
GATCACATCTGTACTCTACAGGCTCTCCACAGAGGTCAGACTGAGGTGGTAAAATTGTTG
TGCACTAAATTAGGGCATTAACGTTTCATGGAAACTGAAGCTATATCTAAATAGCTGATG
GCCTGCTTTCTAGATCTCCTATATACCTGCTTCTCAAATTCAGTCTGTTTTAAAAAATTG
CCCTTTGAGGTTGGAACCAGCGAAATAAGGCTGAAAACAGAATAAGCCATTATTGAAAAA

FIG. 10
(continued)

ATTAGGAACTTGGAAGCAGATACTCATAATCTAAATCCTCTGAAGCTAAAGTTTGATCCA
CAATAGCAAAGCATTATCATTTTAGTGATTGTACCTTAGTTGTTTCCTGGCAGGTGATAA
ATTTGGGATCACTTTCTTCTTACAGTGTGCTCTGATAGTCTTTAAAACAAACCAGAGCTC
TAAATTGTAATGCCATTGGTAATTTAACTCTGATTTGTCTCTATGCCTGTCTCCTGGTGT
TCTGTAAAATTCTACACGTCATTTCAGGTATCACTATCCAGAAGACGTTACTTTTGCCTT
TGATGCACTTTAAAATGTGAAGTCTCTTGTGAAGCTCTTTGGTTATTTTCTCCTTTGCTG
CTGAAATAAATTCAGGTTGATGATTTTCTTGTAGGATATGTTGTGTGATCTAGACATTGC
AAACCCAAGTCTTTGATTTTTTTTTCCCTACAGATTGCCTGTTTCTTTTTTATTTTAATT
TTTATTAGTTATTATTATTTTTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCA
GAGGTGTGATAGCTCACTGCAACCTCCACCTCCCGGGTTCTTGTGCCTCAGCCACCCAGG
TAGCTGGGATTACAGGCACGTACCACCACTCTCAGCTAATTTTTTTGTATTTTTAGTAGG
GATGGGATTTCTCCATGTTGGCCAGGCTGATCTCAAACTCCTGACCTTAAGTGATCTTCC
TGCCTTGGTCTCTGAAAGTGTTGGGATTACAGGTGTGAGCCACTGTGCCTGGCCAGTTAT
TAATTTTTTTAAAGAGATGGGGTCTCACTATCTTGCCCAGGCTGGAGTGCAGTGGCTCTT
TACAGGCACTGTTGTAGTGCACTGCAGCCTTGAACTCCTGGGCTCAAGTGATCCTCCTGA
GAGGCTGGAATTACAGGCACACACCACTGTGTCCAACAGATTGCCCATTTGTGATCTGTG
TAAATATCTCTCACTTCCTGCAGTATCTCTGCTCAAGAATGTAAAGAGATGGATAATATT
TTTAGATTTGTTGAAACAAAGTAAAGTTCTGCTCAAATGAGAATGACACTAACTAAATGA
AAAGGCCGGTTATAATTCTGTAATTTTGTGCCTGCAATGTGTGTGTTATTGTACACTTGA
ATCGGCCCTGTGCATTGTGGCGAGGTGCATATTGCATGGTTGTATTGAAAAGGTGCTTGG
GCCGGGCGTGGTGGCTCACACCTGTAATCCCAGCAATTTGGGAGGCTGAGGCAGCTGGAT
TACCTGAGGTTAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCTGTTTCTAGTA
AAAAATACAAAAAATTAGCTGGGTGTGGTGGTGGGTGCCTGTAATACCAGCTACTAGGGA
GGCTAAGGCAGGGAGAATTGCTTAAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTG
TGCCACTGCACTCCAGCCTGAGTGTATCACAAAAAAAAAAAAAAAAGGTTTTTGCCCTCT
CTCTGTGCCTGCTGCTCCCTGTTGAGTCCTATAGGCCTGAGCTGCCAGGGGGTACTGTGG
GCTGAGACTGGACATTGCAACCGACTGCAAGGCACCGTGGGACCCAGGTTGTGGATGGAC
TGTCTCTCGGGCTTTCTTCTTTCCATTCATCTTCCTCCTCTAACTCCCCTCTGTATCCAG
TATCCTTGCTCTCCATACACCTGCTTCATTCTTTTTCCTTCAGTAGATTTTTCTGCTTCT
TGACTTACAAACCCTACTTCTAGCCCCTTTCAGATATTGAAACTAGCAACTTTCAGGCTT
TGTACCAAAGTCTCAGAGATTCTCATTGACTCGGATGCCATCCATCTCTAGTCCAAAGAA
CAATGTCAAGGACATGAACATGTGGAACAAAAGTGTCTGCTGTGGACACCTTTGGGGAGA
AATAGTTTTCAGTGATGAGGGTTGTAGTGAGTTGGGCAGATATCCCAAAAATATCTGCCA
AAAACTATAGACACTTCTGGTTGCAGTGACTTATTCCTTCCTTCATTCAGCAAATACTGA
TTGAACACCGACTGTATGTCTGGATCTATTCTAGGTTTTGGGGGTGGAGCAGTGAACAAA

FIG. 10
(continued)

TCAGTCTTTATCTTTATAGAGTGTACAGTCAAGTGGGAGAGACAGGCAGTAAACAAAGAA
ACAGTTCAATATTCAATCTGTGAGATGGTGATAAGTGCTACAGAGAAAACAAACTAGTGT
AAGATAAAAAGGGTGTTTTGATAGGCCTTTACTATTTAGGTCTCTTTGATAAGGTGGCAT
TTGAACAAAGCTCTGAAGGAAATAATGGAGCCAACCATGCATATAACCTCAGGGAGAACA
TTCTAGGTAGAGGGAACAGCAAGTGCAAAGGCCCTGAAGTGGGGGTTTGTTTACCTTGTT
GCACAATCTGCACACAGGCCAGTACAATTGGAATGGATGGGAAATGTAAAAGAGAGAAGT
TGAAAAGGCCAGGTGCAGTGGCTCATGCCTACAATCCCAGCATTTTGGGAGGCTGAAGTG
GGAGGAATTTGAGATCAGCCTGGGCAACAGAACCAGACCTCGGGCTAATTTTTGTATTTT
TAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGATCTCAAACTCCTGACCTCAGGTG
ATCCTCCTGCCTCAGCCTCCCAAAGTGCTAGGATTACAGGTGTGAGCCATGGCCCCCAGC
CGTATCTTTGTCTTAAAAAGTAATCTCTGTGCTTGGTAGGCCAAGAATTTAAAATATAAA
AAATTTAAGAAAGAAAAAAAATAAGTAAAGTAACTATACAGGTTGGTCTGGCCGTAATGG
TGAGTGTCATTATTTTTCTTCCCTAGGTATTTTGGCTCTGTTGCTCAGAGCAGTGCAGGC
GAAATGGTCATTAGGGCATCGTCATGGTGCCTGGGGATGCCTGGCTCAGCCAGTTTATTT
TCTGTCTGCCTCTCTCCTTGGTCCTTTTCCTCCACTTTCATTCATGAAATTCTAGTCAAG
AGCTGGGTCCAGTGGTTTTCAATCCAAGGGCTTTGGAAGCCTCTGGGGTCTATTTTGGTC
ATTGCAGTCACTGGGCTGCTGCTCCTGGCATTTAGGTTGGCAGGGGTCTGGGCTGGGAAG
CAGGAATGTTCAGTGGCCATAAATGTAAGGGTTGGTCTTACATTTACATAAGGGAGACAA
TGAAAACTTAACTCCTCCACAGTAGTGGAGTAGTGCCGTTGGGTACTCACAGTCAGTAGT
GCCGTTGGGTACTCACATGTACAACATGGATCAGGACATTGACTTTCTGTGGATACCTTT
TAATAGTTTATTAGATGTGTTAGGCTGTTTTGCACTGCTCTAAAGGAATATCTGAGTCTA
GGTAATTTATAAAGACAAGAGGTTTAATTGGCTCATGGTTCTGAAGGCTGTACAAGCATG
GCTCCAGCATCTGCTTCTGGTGAGGGCCTCAGGAAGCTTCCGGTCATAGTGGAAGGCAAA
AGGAGGGCAGACGATCACATGGCCGGAGTGGTGGCAAGGGTGGGGTGGGAGCCACGCTCT
TTTTTTAATTTTATTTTAATTTGAGACAGTGTCTCACTCTTTTGCCCAGCCTGGAGTGCA
GTGGCGTGATCTCAGCTCACTGCAGCCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCT
CAGCCTCCTGAGTAGTTGGGACTACAGGCGCGCATCACAATGCCCAGCTGATTTTTGTAT
TTTTAGCAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGGACTCCTGATCTCAA
GTAATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCAC
GGCCACCACACTGTTTTAAACAACCAGATTGCACGTGAACTTAGAGTGAGAACTCACTGT
GAGGATGGCACCAAAACATTCATGAAGGATCCACCACCTTCCTTTAGGCCCCACCTCCAA
CACTGGAGGTCATATTTCAACTTGAGATTTGGAGGGGACAGACATCCAAACCGTATCATT
AAATTTAATAGTTTTATGCAGTTTTTTTGGCTCTAGATCTGTTTAGACTCCTGCAGTCAG
GTGTCTGTAACTAGCCTCTGGTCCTTTTTGAGAGTTCACAGTTTGGTGCAAACCCTTTGG
ATGTATTATTTGGGAAAATGGGATATCTGGCAGCCTGTGTCCCTGCTTTACATTATCCTT

FIG. 10
(continued)

TTTGCTGCCTGCCCCAAGCCTCCTCATTAGCATCCCTGCCAAGGCCAGTGGAGAAGGATG
GAGATGCGGTGACATTCAGCTTGACAGGTCATTAGCAGCTTTTGTGCCCTAGGGACTGCT
GGTGGGAGGGAGGTTGTGGAAGATAAACCCTGACAGGAATGTATTCTCCTCGAGGGCAGG
GTTTATTTGATATTTTTCTGGAGCTTAGAACCATAAGCCTGGTGCTGGGGAGGAAGCGCC
CTTAGCATTTGGTAGCCTCTGTGGGCAGAGCATGGAAAGTCACAACTTCTGAATTGTTTG
TATTTTCAGTCTCACTCTAGATGGATGGCATCTTCTGCTATGGGAAATGAAATATGTTTA
GGCAACTTGAGTCCCAGGTGCAGATGAGGCTGGGCTAATTGGTGCACTAGGGAAGGAGCC
GGGGGAGAGATGTGCTGTTAGCTATTATCAATCTGTGACAACTGTCAGCTGCTGGCAGTT
AGCACCCACCTGAGCCTGGGATGCAGGGGTGCCTCTCCTGTCCTCTGTGGAAGCCTCTGG
ACCCAGCAGCCATCTTGACTGTGCACTGTTCAAGCCCCAAGTCCGCCTGGAAGAGGTGAT
TGAGAACTTACTGCAGGATAAGGAAAGCGCAGGACAGGTGCAGTGGCTCACGCCTGTAAT
CTCAGTGCTTTGGGAGGCTGAGGCCGGAGGAGGGCTGGAGTCCTTGAGTGCGAGACCAGC
CTGGGCAACATAGTGAGACCCTGTCTTTACAAAAAGGAAAAGAATTAGCCAGATGTGGTG
GTGCGTGCCTGTAGTCCCAGCCACTCAAGAGGCTGAGGTGCGAGGATCACTTGAGCCCAG
GAGTTTGAGGTTACAGTGAGCTATGATCATACCACTGCATTCCAGCCTGGGTGAGAGAGC
ATGACTCTGTCCCAACAACAAAAAAAAAGATTAAGGGAAGCCTCTGGCAGACCTGATGAT
GGGTGGCCCAGCCAAAATGAGTATTGATGAGGATTTCCCTGGTCTGGAACTCTGAATTTA
GTCTGGCAAAGTATTCCCTTTGTGTTGTGAGATGATTCTTGGTGTTACCCCATCACGGTA
GGTAAGATGAATTAGCAAATGAGAAAGGCTTTCTCTTTTTCATCCTTATCTAGTCCGTAG
ATGAAGCCTGAAGAAGGTCTCCATATGGTAGTAGTAAGTGTTTAACATCTACCTCTAACA
CTTGCCTGTGTCTTTTTTTTTTTGCAAAGCCTCAGGAATGCCCCAGTATCTAGGTAGAAT
TTGATAATATTTCATTTTTGTTATATTCCCTTTTCTGTTTACCTTCTATATACAGCAAAA
TGAAAAAATTTTTAAAATTTGTGCAAGTAAGGGCAATTTCTTTTTTCTTTTTCTTTTTTT
TTGAGACAGGGTCTTGCTCTGGCACCCAGGCTGGAGTGCAGTGACACAATCTCGGCTCAC
TGCAACCTCTGCTTCCTGGGTTTAAGCGATTCTCCTGCCTCAGGCTTCCAAGTAGCTGGG
ATTACAGGTGCCTGCCACCACTCCCAGCTAATTTTCATATTTTTAGTAGAGACCAGGTTT
TGCCATGTTGACTGGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCATCCACCTTGGCC
TCCCAAAGTGCTGGGATTATAGGCTTGAGCCACTGGGCCTGGCTGAGGCAGTTTCTTTTT
GAAATATATTTTGTGAAGGAGAAAAAGAGGAGTTCAGTTTAAAGAAACAAATGACATAAG
AGGTGGTATGCAGAGATGCCAAAGCATCTTGAAGGTGCTTTTTTTTTTTGGAAACAGAGTC
TTGCTTCATTGCCCAGTCTGGTCTGCAGTGGTGCAATCATGGTTCCCTGCAGCCTTGACC
TTCTGGGCTCAAGTAATCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGATGCAT
GCCACTATGTCTGGCTAATCTTTAAATTTTTTGTAGAAGCCAGCTCTCACCATATTGCCC
AGGCTGGTCTTGACCTCCTGTCCTCGAGCAAAAATACCGATTTTGATTAAGTCTGGGGTA
GGACCTGGGGCTGGGATTCTAACCAGCTCCCAGGTGGTGCTAATGCTGCTGGTCTACAGA

FIG. 10
(continued)

CCACACGTGGAGTAGCCAGTGTAGAGTTCATGTAGCAATAGTGATGTCATAGAAATAGCC
AGTATCTGTATACTTGCTTTGTTGTATGTCACGCACTGTATAGTGATGTACATGCATCTC
ATTTGACCCTCACCCCGCCCCTTTGGGGGTAGAAAGGATTGTGCTCATTTCACACTCAAG
GAAACTGAGGCACAGACAGGCAAAGTAGCTTGGCGAAACAGAAAGGAACTTAGAGGCAGG
CCCTGATTAGCTCAGAGACTAGAAGGCCTTGTGCGTCATCCTGAACAGCTTGGACTTGAT
CTTGAAGGTGGAGGGAGAAATTGAAGGGTAATTAAACAGGAACTGTAGGAAATTCACCTT
GCATAGTGATTGCTTTGGCCACGTGTGCCCTGCCACCGCCCCCCCACCTCAGTGAAGTGT
CATGCGAAGTTGGGTTCGTAAATGAAGGCCCGAATGCTTTCCTGACAAGTTTGTTTTAAA
TCAAGCTGCTAATTAGTCCCAGTCCCCCTCCCCCGGTATGTATTTTTTTGTTGATGTCGT
TTCACTTCATTTAGTTGAAGTGATTGATTCAGTTCAGTGTTTGAACTTCTTTTTGAACCT
CACCTTAATAACCTGTCTAAACATCAAGGTTAAACCTTCTTGCTAACACAGCAGTATTGC
TTGGTAAGACTGGCTCACAGTCCAAGGAAATGCTTGCCCAGAGAGGGCAAACTGCCTTAA
CTCCTTAACCTGAGCTCATTAAAAAAAATTCAAATGACTGATTCCTTGTCACAGTTCTAC
CTACATTGTTTTTATTTTTGTCCAGGTTTCAGCTAGTTAAATGCTTTTGTGATGAGCTTA
TGTCCAGGCTGAAGGTTGCATTTTGAAACTGAGCGTCAAATACCAATTTAAAGTCCAGAC
CTTTACACTTGTGAAATTCAGATAAATGAAATGGAAATAAAACAGGGCTGCTGTGTTGTG
AAATATGACTGTGTTTTTCCTTGTAGGACTCTTTGAGGGTAGCCATTTTGGCATTTTATA
TATAAATTTTCTTTTCTTAGCCTACCTTTTACTTTCTTGATTTGCCTATTTGTGATTTCC
CATTAAACACTAGGCTTTTTGTAAACCAATTATCCCTTGAAATTGACTTTTTTTTTTTTT
GAGACAGGATCTTGTTTTGCCACACAGGCTGGAGTGCCGTGGCTCCATCATATGATAAAC
AGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACCCTGTCTTATTTAAAACA
AAAAAAGAAGAAGAAAAAAAGAATATAGATCACAGCTGTTATTTGTATATGCTACGCCAA
TCCTTGTTGGGTTTCATTCTTTATAATTGTTATTTTTAAAGATTTTTCTTATGAATATTC
TATTGTTTCATTGTAGAAAATTTAAGGGAGAACACAGTGGGAAAAAAAAAAACAAGAAAAG
GACTTCATAATCCTGCTACCCTGGGAGAAAAAAAAAATCACCATTACCTATTTGGTTCTT
CTCCCACTTTTTTTTTTTTCGAGATGGAGTCTCCCTTTGTTACCCAGGCTGGAGGGCAGG
GACGTGATCTTGGCTCTCTGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCGTGCCTCA
GCCTCCCGAGTATCTGGGATTACAGGGGTGTGCCATCACACCTGGCTAATTTTTGTATTT
TTAGTAGAGACGGGGTTTTGTCATGTTGGCCAGGCTGGTTTGTTGGCCATGTCTGGTTTT
TTGTCATATTGGCCAGTCTGTTTGTCATGTCAGGCTGACATGTTTTGTCATGTTGGCCAG
GCTGGTCTTTAACTCCTGACTTCAGGTAATCCTGAAGTGCTAGGATTATAGGCGTGAGCC
ATTGCACCTGGCCTTCTGCCTTTTTTTTAAAGAAAAAAAATTAAAACATTTTTTTCTTTT
TAAGATAGCGTCTCATTTTGTTGCCCAGGCTGGTCTTGAACTCCTGGGCTCAAGTGATCC
TCCAGCCTCAGCCTCTGGAGTAGCTGGGACTACAGATGCACATCATGGTGTCCTTATGCC
ATTTCTTTTGTACGTAGGTGAATGCAAGTGTATGATTACATCATATGCTATTTTGGAGGT

FIG. 10
(continued)

TTGACTTTCTTTTCACTTTCATCATCTTTCCAAGGTGTTATTTTCCTAGTACATCTTTTT
AAATGGACATAGAACATTCTTTTGTATGAACAAACAATAGTTTTATTTAGGCGGTCCTTT
CCTGTTGGACATTTATATTATTTTCAGCATTTCTCCACAGTTGTTGCAGCATTCAGATGA
ACCTTCTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTTTCGCCCAGGCTGGAGTGCAG
TGGCACAATCTCTCCTCAAGTGATTCCTGTGTCACCCTCCCACGTAGCTGGGATTACAGG
TGCCCATGTCTGGCTAATTTTTGTGTTTTTGGTAGAGCTGTGGTTTTACCATGTTGGCCA
GGCTGGTTTCGAACTCCTGCCCTGAAGTGATCTGCCCACCTCAGCCTCCCAAAGTGTGGG
GATTACAGGTGTAAGCCATCACGCCTGACCCAGATGAACATTCTTGTAGCTATCGCACAC
AATTCTGAACATTTCCTAGGATGAATTCCTTAAAGAAGTAATGCTGATCCAGGCTTTTTT
CTTTTTCTGTGACTCTTTGACACGTAATAATATTGACTTTTCTTTCTTTCCAGACACTAC
AACAACAGGAGTGCAAACTTCTCTACAGCCGAAAGAGGGTCAAAGGCAAGAAAACAAAA
ACAAAAAATAGATATAAAAACATCCTGCCCTGTAAGTATCAATATTCCGCTCAGTAATAGT
CACTCTTGGAGATTTTGATTCCTAGCACCTCTGTACCTTTCCTCAGGGTCGTGTGCTCTT
GTTAGCACATCGGAGGCCTTAGCTTCTTTAATTGCAAGCAGTTTCCAAAATAATCAACCA
TGGTGGGTGTTGATGACTTCATTCACTGAGCTCCCGTGATGCTGATTACTGAGTAAAGTT
GCCACTAGGTGGCTTTGTCTGTGGTTGGTTCCTTCTGTTAATTAATTTTCTGTCTGCCCA
AGATAGATCATCTCAAGGCTTGGGATCTCTCAGTGTCAGGGACCTTAGGGTGCCAGATTT
GTGTCTTGACTCCTCCTCACTGGGCCTGTGAGTCCTGGGTAAGGCCTGCCTCCTTTCTGG
GACTCAGTTCCCTTAAGTGGGAAACAGACAAACACCTCCTGAGGGCTCCTAGAACTGTTC
TGCTTGCTGATCCCCTGAGCTCAAGTTACTGGAGAAAGGGTATATACCTAAACTGCTCAG
AAGAAGACTTTGTGGGCCGGGCGCAGTGGCTCACACCTGTAATCCCAGCACTTTCGGAGG
CCGAGGCAAGCGGATCACCTCTGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAA
ACCCCATCTCTACTAAAAATACAAAAATTAGCCATATGTGGTGGTGTGCGCCTGTAATCC
CAGCTACTCGGGAGGCTGAGGCGGGAAATTGGTTGAACCCAGGAGATGGAGGTTGCAGTG
AGCCGAGATGTGCCATTGCACTCCAGCCTGGGTGACAAGAGCAAAACTCCGTCTCAAAAA
AAAAAAAGGAAGACTTTGTGAATATTCGCAAAGCTGTAAAGCTGTACCTTTCAATTTTTT
TTTGAGACATAGTCTCACTCTGTTGCTCAGGGTGCAGTCACAGCTCACTGTAGCCTCAAC
CTCCTGGGCTCAAGCGATTCTCCCACCTCAGCCTCCTGATTAGCTGGGACAATAGGCAGG
CACCAGTACACCTGGTTGATTTTACAGTTTTTCTGTAGGCCGGCGCAGTGGCTTACGCCT
GTAATCCCAGCACCCTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTTAGGAGTTCGAGA
GTAGCCTGGCCAACATGGTGAAACCCCATCTCTATTAAAAATTACAAAAATTAGCTGGGC
GTGGTGGTGGATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCTGAGGCAGGAGAATC
GCTTGAACCTGGGAGGCGGAGGTTGCAATGAGCCGGAGGTGCTATGTGCACCACTGCACT
CCAGGCTGGGCGACAGAGTGAGACTCTGTCTCAAAACAAAAAACGATTTAAAAAATAATA
AAATTTTTTCTAGGGCGGGGTCTCCCTATGTTGCCCAGGCTGGTCTTGAACTCCTGGGCT

FIG. 10
(continued)

CAAGTAGTCCTCCTGCCTCAGCCTCCCAAACTGTTGGGATTACCAGTGCAAGCCATTGTG
CCTGGCTGTACCTTCTGTAACACCCAAATGCCACCTGGCAAAGCCCAAGTTGAATCATGA
GGAAAAAAGGCCTGGAAGGATGTAGACCTTCCTTTTTTCTACTTATTTATTTATTTATTT
TTGAGATAGGGTCTTACTCTGTTGCCCAGGCTGGAGTGCAGTGGCATGATCATGGGTCAC
TGCAGCCTCAACCTCCCGGGCTCAAGTGGTCCTTCCCACCCCAGCCTGCAATGTAGCTGG
GACTACAGGCATGTGCTACCATGCCCAGCTAATTTTTGTATTTTTTGTAATTATTTTTTT
TGTAGAGACAGGGTTTCGTCATGTTGCCTAGGCTGGTCTCGAATTCCTGGGCTCAAACGA
TCTGCCTGCATCGGCCTCCCAAAGTGTTGGGATTACAGGTGTGAACCACTGTGTCTGGCT
ATATCTTCTGTAACACCCAAATGCCACCAGGCAAAGCCCAAGTTGAACCAGGAGGGAAAA
AGGCCTGGCAGGATGTAGGCCTTGCATGAGGATCTCAGAAACTGCACTAAACCAGTCACA
GTTCCTCTCTCCCGAGGTCTAACTCTATGCTGAACTCTTTGCATTTTTATCTCACTTAAT
CCATATCACATGCACAGGAAGGAAGCATTCGTAGTATCCTGGTTTCCTAGACCATTTTAG
CAAGGTTATAAGTGAAGGGGAGTGGGTGGGAGAACTGGCACTAGAGCCCCCAAAGTCACT
GTTCTTAGCACCACTCTAATGCATGGGGTTCTCCATTGATGTGCTATGCAAGGCAGTGCA
CTGAGGAGAAAGGAAGGAACATTTACAACTTCTCTTTATTTATATCCTGTCCCTAAAAAA
AAAAGAAAAAGAAAAATTTGTCTGAGGCCTAGATTGATTGCAGGGAGTGCATAATGTTTT
ATTGATTGATTGATTGATTGTATATAGAGATGGGGGGTCTCACTATATTGCCCAGGCTGA
TCTCGAACTCCTAGGCTCAAGCAATCCTCCTGCTTTGGCTTCCCAAAGTGCTGGGATTAC
AGGCATGAGCGACTGCACCTGGCTATGCATACTATATTTATCCAACTTACAAATAAGGCT
TGCTTGCCTGTAGTGCATATGTGTATACATTTCAGCATAGAAAAACTGTGTGATTGGGGG
TTGTGATCAAATTTGGAGAGCATTGCTCTCATGTCTTATCAGGTCAGAGTCATTTTGTCA
AATCTTGTAAACCATTCTTTGTGTGTGTCTATGCATGAAACATAGTCTTTCTCTTTCTGC
ATGCATATGTACATATACATGGTATATATGTATATCATATCTACATGGATATTGTAATGT
ATATGTATGAGGATGGGGGAAAGTGGAGACATTTGTAATACTGAGAAAAGGCAGTGAGGA
ATTTGCAGAGAAGCAGTTTGAGCTGTAGCATGGTACTAGTGACCTTGAGGAAGCCTTATC
CTTTTTTTTTGGAATTTATTTTTTCAATTTTTAGAAATAGACAAGAGTTTCTCTATGTTG
CCCAGGCTGGTCTTGACCTCCTGGGCCCAAACTATCCTCCTGCCTTGGCTTCCCAAAGTG
CCAGGATTACAGGTGTGGACCACCATGCCTGGCCACCTTGTCCTTTCTATGTCTAAGTTG
TGACATCTGCTCAGGGGTCAGGTGGTATTAAATGGTATAAAATGTATGGGAAAGTGAAGG
GATCAATGGTATGCAGTATCTAAATAGAATATCGCTTTTTCCTCCCTTAAAGGTCTCATT
CAGATGTTTCCTCTGATGAACATCTCATTTCCTTAAAGATGAGGAGTCTGAAGCAAAAAA
GACATTATTCTTTTAAGACACATGGCTGTCTTACTAATTCCCATTGCAAAATATGTTGTT
TAGGTAGAGCACTCAGATTTTTATACGAATAATAGACTTTTGTACAGAATTTGGACAGTT
GATACTATCAGAGCCTTGTGATATTCCACTGCATTATGCTTCACTAAAAAATACCTGGCT
GGGTGCGGTGGCTCACAACTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGCAGATCAC

FIG. 10
(continued)

CTGAGGTCAGGAGTTCAAGATCAGCCTGGCTAACATGGCAAAACCCCATCTCTACTAAAA
ATACAAAAATTAGCCAGATGTGGTGGCACGCTCCTGTAATCCCAGTTACTCAGGAGGCTG
AGGTATGAGAATTGCTTGAGCCCAGGAGGCAGAGGTTGCAGAGAGCCGAGATAGTGCTAT
TGCACTCCAACCTGGGTGACAGAGGAAAACCCTGTCTCAAAAAATAAATTTAAAACAACA
ACAACAACAACAACAAAAACCCCTCTTTATTATGGAAATTTTCAAATATATTCAAGAGCA
TAAAGAACCCACATGTACCCATCACCCAGCTTCAACAATTATCAACTCATGCCCAGTCTT
GGTTTCATCTATACTCTGATCCACATCTCCTCTCTCCTTGAATTATTTTGAAGCCCATCT
CAGACATCATGTCATATATGTATACTTCAATCTTCTTTTTTTTTAAAACTCCCCCTCCCC
TTTTCTTTTTTCTTGAGACTGTGTCTCACTCTGTCATCCAGGCTGGAGTGATCTTGGCTC
ACTGCAATGTCCGCCTCTCGGGTTCAAGCGATTTTTGTACCTCAGCCTCCCTAGTAGCTA
GGATTACAGATGTGGACCAACATGCCTGGCTAATTTTTGTATTTTTAATAGAGACAGGGT
TTTGTCATGTTGGCCAGGCTGGTCTTGACCTCCTGACCTCATATGATCCACCTGCCTTGG
CCTCCCAAAGTGCTGAAATTATAGGCCACTGCGCCCAGCCCAAAATTTCTTGGTTTGAAA
TAATTTTGGAACTCATAAGAAGTTACACATATAGTAGAGAGAATTTTCTTGTACCTTCTC
TGAGCTTCCTATATACCCAATGATAACATCCTATATACCCATAGTATATGATCAAAACTA
GGAAATTGTGAAGATGGCATTTTGAGACATCAGGCAGTGTTCACGTTACTGTTTTGCTTA
CCTGGGCTTTAATTTTTATGTGTTTTTTTTTCAATCATTGAATGAACAAAACTTGGACTA
GGCTGGGGAGTAACTGATTTGAACTGTTTTTTCCTGAAGCAGTCCAGGACTTATGTGACC
GTGGTCTCTTTTTCTTCTAGTTGATCATACCAGGGTTGTCCTACACGATGGTGATCCCAA
TGAGCCTGTTTCAGATTACATCAATGCAAATATCATCATGGTAAGCTTTGCTTTTCACAG
TGTTTTCTGACCATACATTTCTAGCCTATTTTTGTATTTTAAATCCTTCCTCATGTCCTG
AAAGTAACTTTAAGGTGTTTGAAGGATTTTCTTCCTAAATTTCTAGCCTGAATTTGAAAC
CAAGTGCAACAATTCAAAGCCCAAAAAGAGTTACATTGCCACACAAGGCTGCCTGCAAAA
CACGGTGAATGACTTTTGGCGGATGGTGTTCCAAGAAAACTCCCGAGTGATTGTCATGAC
AACCAAAGAAGTGGAGAGAGGAAAGGTAAATCACAGAAACTTCTTTTCTGCTAAACTGTT
TTTAAAGTATCAGACATGTCAGATTGGCCATGTTTAGGAATTGAATAAATGAATTAAGCT
TACTGTAACTGATTCTCTGGAAAAAAGGGACTAGGAGAAATTTGATTATGTTATTCCTTG
GTGTAGTTTTCTTTATGTTTCTTCTGCTTGGGATTTGTTGAGCTTCTTGGCTCCATGGAT
TTGTAGTTTTCCTTAAATTTGGATAATGTTCAGTCTTAGTTTCTTCAGATACATATCCTG
GGCTGGGCATGGTGGCTCATGCCTGTAGTCCCAGCACTGTGGGGTGTTGAGGTGGGCGGA
TCACTTGAGGTCAGGAGTTTGAGACCAGCCTGGGCAATGTAGTAAGACCCCATCTCTTAA
AAAAAAAAAATGTACCCTGCACAACCTTGTCCTAGGACAGCAGTCATACGTGTATTAGAC
TACTTGAAGTTGTCTCATAGCCCACTGATACTTGGTTTATTTTATTCAGTTTTTTCTCCC
CGTGTTTCATTTCGAATAGCTTCTTTTGCTATGTCTCCAAGTTAATCTTCTGCAATATGT
CATCCGCTCTTAATCCTATCCAGAGTATTTTTCATCACAGACATTGTATTTTTCATCTCT

FIG. 10
(continued)

AGAAGTGTTAATGTCATCTATAGCTTTCCTTTTAACATGTGTAGCATTTTCCTTACCTTT
TGAATGTATGGAGTATTTCTGTTGTTGTTTTTTGTTTTGTAGAGACAGGGTCTCGGTCTG
TTGCCCAGGCCGGAGTGCAGTGGCATGATCTCAGCTCACTGCAGCCTCTGCCTCCCGGTT
CAAATGATTCTCATGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGTGCGTGCCACCACG
CCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGTTGGCCAGGCTGGTT
TTGGAACCCCTGAGCTTAGGTGATCCACCTTCCTTGACCTCCCAAAGTGTTGGGATTATA
GGTGTGAGCCACCATGCCTGGCCATGTTGTCTGTTTTAATTAACTCTGCCTAACTGTCCT
CCCAAATGGTTGCTGCAGTGCTCACTCCCACCAGCAGCACCTGCCTAGGACTCATTACTC
CATACTCTTCAAGACACTTCAGATTAAAAAAATAAATTGTAACACCCCACACCTACAGAA
GAGCGGACAGATCTTATTGAGTGACAGCCCTCTGTGTTATCTCAAAGTGAGCCCACCATG
GTGGTTTTTTTTTTAAATATGGAAAAGTTCTGTGTTTTTGTTTGTGTTCTAGTGAAAGTT
CTTTTTTAGATATCCTTTAATTGGTTTATATAAGATTTTATGTGGAATGTAGCAGTCATA
CCTATAAATTAAACCTAAGGCAGATGGAGAACTTTGGAGTTGAGCCTTCCTACTGTAATT
TTCATATTGGATGTGAAGGGCAGTGTGATTTTCATAAGACTTTCATTGTTGTACTCCTAG
TTGGTATACTTCTGAATACCTTTGAGGCCAGTTCTGGTCATCGTGAAACAAAGGTTTCCT
TCAGCAAATGCCTGTGGTAACATTAGGTGTTCTTGAATTAATGGACCAATGAAAACATCT
TTGTAGTTTCTGCTTCAGGCAAGGGTTTTTTGCCCTAAATGTGGATAGGAAGAATGAAGC
CCTTCATCCTCCTTTTTGCCTGATTATAGCTATAGGAGGTTCACCTGTTCTCAGAAGACA
TGAGGATTGTGAAGAGAGGGGTCTTGTGTTGCTTCAGAGGAATCAGTATCAGTCCCTTTC
AGAAGCTCTCCTGGATAGACAGGCATTAGGGCCAAATCACTCTGCCCCACCCCTCACCAC
CATGTCCTACTCTCTGCTCCCTGTCTCATTCTTCCTCTTTACTTTGGTGGTGCCGAGAGG
ATGACATGATGGGTATTGATTCTCTCCACAGACCTTTCTGACATCCTACTTTCAGTATCC
CCCCAGTGCACAGAAGACAAGCCAGACTGTGGACTGTGTTTGATTCCTGGGCTCTATTTT
AAAAGACAGTGTATTAGTTCTCACATTTTAGAATTTGTTTGCCAAGGTTTCCACGGGAGT
TTAGAAACTAGGGGGAGGGCTGATGTTTAAAGTTAGCTAAAATGTTCTTTTCAGGGTCAT
GATTTAATTTTATATTCTCTGGTGAGTTCCCTATAGTGACTGGGAGCAGTCCTCAGTCTT
GATTGGCCAGTGACAGCATAGAGTACAATTAATATTAGGAGTGCTCATTTGGGGAAACTA
AAATTTGCATCAAATCTGTCAGAGGTGTTTGGATCTACAAAATACCGGAGGGAAAGCTGA
ATTGAGAATCATAATAAATAAAAGACCACATCGTTCTTTTTTTTTTTTTTTTTGGGACT
GTATCTTGCTCTGTCACTCAGGCTGCAGTGCAGTGGCACTATCTTGGATCACTGCAGGCT
CCGCCTCCCGGATTCAAGCGATTTTCCTGCCTCAGTGCCTGAGTAGCTGGGATTACAGGC
GTGTGCCACTACACCTGGCTAATTTTTGTAATTTTAGTAGAGACAGGTTTCACCATGTTG
GCCAGGCTGGTCTCAAACTCCTGGCCTCAAGTGATCCACCCGGCTTCCCAAAGTGCTGGG
ATTACAGGCGTGAGCCACTGCGCCCAACCAAGACCACATCCTTTTATTGAACGTTCCTCC
TACCATGTTTTCTTTTTTCTTTCAATTAATCATTGACTCATTGACTCTCACTGTTGATGT

FIG. 10
(continued)

CTGTAGCTGCTCTCTTATTTCCAGTTTTATAGCTGTAAATTTCTCTGTCTTCCTAAGATA
CAAGGTAAATTTCTCTTGCTGATATTGGTGGTTTTGGAAAGTGAGTGGTGTGGATGACTG
CCCAGAAAACAACAGAACACAAAAGCATTCTCTGCCCAGAACACATCACCAAATAGATAC
AAACTCATCTCTTACTGAGTGAAATAGCTTCCTTTTTGGCAGCAAGAATGATTTTCTTGG
TGCCATATTTTTCAATCCGCCTGCTCTTGAAGCCAGCAGCTATTGCAGACTTGGCATTCC
CAGGCACCCAGTTAAGGGAAAGTGACGTGTAGAGGAGGTATCAGATGGGTCTGGATATAG
AAAAAGCAGCTGGTTCAAAACCCCATGGGCTGCCTTTCTGTGATAGAGTTATTCACACTT
GGGTTAGATAAGGCACAGAGTCCTCCTACACTGGTGCGGAAATGAAACAGACAGTCTGGC
TCGTTGGGCAGCCTAGCCTCCTCCAGAATCTGTGCTTGCCTTCCCTATGGAGTGACTGGT
AGATCTTAGAATTCAGACCTCAGTGGTTGCTAGCCAGCACTCTCACATTGGTTGGTCCTT
CTCTCTGCATCTTTGATTCTTTAGAGATAGATAAACCAAGCACCGACTCTCCTTTGACAT
GTGCTTGGAACAGACACCTGCACGAGCTGCCTTTCTCCTCCCACTTCTGCCTGGTCTTCC
AAACACCTGCTTTTCTTGTTTGAACTCTTCCTTTTTTTTTGAGACAGAACCTCTCTCTGT
CACCCAGGCTGGAGTGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCAGGTT
CAAATAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCTGCTATCACG
CCTGGCTAATTTTTGTATTTTTAGTAGAGACACGGTTTCACCATTTGGCCAGGTTGGTCT
CAAACCTCTGGTCTCAAGTGATCTGCCCGCCTCGGCCACCCGAACTGCTGGGATTACAGG
CATGAGCCACTGCGCCCCAGCTGATTCTTTACAGATAAACAAACATTGACTCTGCTTTGA
CATGTGCTTGGATCAGGTAACTGCACCAGCTGCCTTTCTCCTCCCACTTCTGCCTGGTCC
TCCGAATGCCTGCTTTTCTTATTTGAACTCTTCTGTCCTTTTCTGAAAACCTAACAGATG
CGAAACAGGCCATTTTCCATGTTGGTGGTTATTAAGCAAGACTTGAACATTTGTTTGTTG
CTTGTTTAGGCTTTTATTTCAGAGTTCACAGAATTAACTTTCTTTTTTTCTGATCTCTTC
CAGAGTAAATGTGTCAAATACTGGCCTGATGAGTATGCTCTAAAAGAATATGGCGTCATG
CGTGTTAGGAACGTCAAACAAAGCGCCGCTCATGACTATACGCTAAGAGAACTTAAACTT
TCAAAGGTTGGACAAGTAAGTATATTGTCGTATTCTAGAGACTTTGGGAACTGTTGATGG
TGTGTAGGAATTCAGGGTCTTGCCGTTACTCATGTTTGCATACATGCATGCATTCGCTCA
CTCATTGATTCAGTAGCCATTTATTAGCTTCCTTCTATGTGCCAGGTACAGTTTAAGCAG
TACTGGTACATTGTGAACAAGGCAGGTAGTGTTCCTGCCCTCATCGAGCCTAGGGAGATA
GACAATTTAAAAACAAATAACTGGCCAGGCGCCGTGGCTCAGGCCTGTAATCCCAGCACT
TTGGGAGGCTGAGGTGGGTGGATCGCTTGAGCCGGGGAGTTCGAGACCAGCCCTGGGTGG
GAGACTGGGATAGGGTGACCTGAGTGGCTACAAGGTCTGTTAGGAGGCCTCCGCAGGGGC
CTATGTTGATGGCCTCCTCTCCAAGTATCCACAGACTTCAGCAGTTGTTCTTTTTTGTTC
CTTCCTTTGGAATGGAATATTATATAAAATGGCAGAATAAACTGGAAGAGAAGCAGTAGA
TGTGAGAGGTGCCGGGGGGTGAAGTCTGCAGGATGTGGGGATTGTTTGGCTTTTGGAGGA
GGAAGGAGGGATTCAAGACACATTGTAGAGGTTTGAGTCTGAGCGGACAGTGGTGCTGTG

FIG. 10
(continued)

GCAGACACCACAAAAGCTGGAAGGAGAACTGATGTGGGCAGTGATTTGTTTTCTTCTGGA
TGTGTTCAGCTGGGCATCTGAACAGTCATGTGGACATTCATCTATTCATTCAGAGATATT
TGTTCAATGACCTCTTGGTTCCTGGCACCATGCTGCTTGCTGGAGATAGAGCTGGGGAAC
AAAACAGATGGAATCCCTGCACTCCCAAGTGTACACTATACTGGCCAGTAATCTACCAGC
CCAGTAATTGCACATATAAATATATCATTATAAACTGTAATCAGGGCTAGAAAGAAAAAA
TGCAGGAGTTTAGGGTTCATTTGGAGGGGGAAGGGACTTTTTTTTTTTTTTTTTTGAAAC
AGAATCTTGTTCTGTCACCCAGACTGGAGTGCACTGGTGCATTCACGGCTCACTGCAGCC
ACAACCTCCTAAGCTCAAGTGATCCTCTCACCTCAGCCTCCCATGTAGCTGGGGGCTACA
GGTGTGTGCCACCATGCCCACCCAATTGTTAAATTTTTTATAGAGACGGTTGTCTCATTA
TGTTGCCCAGGCTGGTCTTGAACTCCTGGGCTTAAGCGATCCTGCTGCCACATGCAGCCT
CCCAAGGTGCTGGAATTACAGGCGTGAGCCAGCGCACCCGGCCAAGGGAGGGGAGGTTCT
TAAGGCATAGGGAACAATGTGTTTGAGTCAGCAAAGGAGGTTGTGGGGGTTTGTCCTAAG
TGTGGTAAGCAGCCAGAGTTGGATTTAAGTTTTTAAGAGATTCCCCTCCACCCTGTAGAG
ACTGGAGGGGGCAGGAGTTGTTCTAGGGATTAGGACCAATTTGGAGGTAGTGCAGCCGTC
AGAGTAAAAAATAATAGGGATTGAACTAGGCCAGTGCCCAGGGTGCCTGAAAGAAGAGGA
CCCAGTAGAGCTGACTGGAGGCAGACATGCAGGGATTCAGTGAAGGAGTGTACCAAGGGC
GAGGGTGGTGTGCAGGGTGACTGGCAATTTTCTAGCTTGAGAAAGGTCCGGGGGGATGGC
AGTGGAGTTGAGGAAGCTGGGAGGATCAAGGACCTTTTTGTGAACACACAAAGTTTGAGA
TGCCTTGGACACATTGAAGTGGAGCGGTCAGGGAGGCAAGGGTGGAGGTGGGATGCGGAG
GGGAGGTGGGATGCAGAGCGTCGTGGATGGATCAGTTTTGCTCGATAGAGGGACATGTTT
TTCTGTGGCAACAGGAGGGCAAAAGGAGAAGGTGGCCACAGATGCCGGTAGATGAGCTGA
GAGTGATTGTATTCCCTATCCTCTCGGAAGCTTGAGGCAAGGCCATCAACAGACAATCAG
AGGGAATAAGAAGAGATAGAATATATGAAGAAAGGGAGAAAAGATGAAATCGTAATTGTG
TAGCAGGGCAAGAAGTCCAGAAATTTCTGTGCTGTGCCAAGTTCCCAGTTGAGGCGGTGA
ACATGAAAATATACTGATACCCATTGCCTGGTTTTTCTCCAAGGACACTTGGCTCCTAGG
GCACAAAACAGAAAGTACGTGGTTTGTCCAGGCCGAGGGCTTTGCATAGTTGCAGTGGAT
GGAGAGGAGGTCAAGGAATGGAGGCACATGGTAGAGAGAGACTGTCCCCAGAGCACGGGG
ACTCCTGGCCGGATGAGGGGGACAGGGGCAGGAGGAGGCAGGTGGAAAGTAGAGGGAGGG
CTCAGTGGTCTGGAGGCTACAGGAAGTGACGGGGGGACCAGAAGGAGCTGGAAACCAGTG
TGGTTGTGGCCCAGGGTGGGATGTTTGGATTTCTGATGTCAGAGAGGGTCCAGTCCTTCT
GATGATGGGGAGGGGTGGAGGCTGAATCTATGGTAGAGATAGTGAGAGGAACTGGAACAA
TGTAGCTGTCAAGTGGAAATGGGAGAAAGGGCTGGGCGTGGTGGCTCACGCCTGTAATCC
CAGCATATTGGGAGGCTGAGGCAAGAGGATCGTGTTAGCTCAGGAGTTCTGGGCTGCATT
GAGCTGTGATTGTGCCACTGCACTCCAGCCTTGGCAACAGAGTGCCCAGTTAAAAATAAA
AATAAAATAAAATAAAAAAATTAAAAAAAAAAGAAGAAGAAAAAAGAGAAAAGTGTCCTT

FIG. 10
(continued)

TTACATCCCTTTTAAAAATGTCACTTAAGGCTGGGCAAAGTGGCTCATGCCTGTAATCCC
TGCACTTTGGGAGGCTGAAGTGGGTGGATTACTTGAGGTCAGGAGTACAAGACCAGCCTG
GCCAACATGGCGAAACTCCTTCTCTACTAAAATTAGCTGGATGTGGTACATGCCTGTAGT
CCCAGCTACTCGGGAGTCGAGTCTGAGGCCCAAGAATTGCTTGAATCGGGGAGGCGTAGG
TTGCAGTGAGCTGTGATCAGGTCACTGTGCACCAGCCTGGATGACAGAGTGAGACTCTGT
CTCAAAAAAAAAAGTCACTTAGCTTAGATTGTCTCTACATATATAGGAAGAAGATGTAGG
AATGAATGGTGCTGCTACAATTACGTCATCTGGATAGACCCAGAAACATGATACTTTTTG
GTTTTCTGTAGCCTTGGTGCCATTGTTGATCTTTATTAATTATCATTATCCTCAAAATAG
CCATAATGTGCTGAGTCTCTTCCTATTTGCTGGGCAGAGGCTGAGTATTTCAGCGAGCTC
ACTGAGTCCTTAAAATTGCATTATGATAGAGAGAAAGAGATTATTATTTGCATTTTGCAA
AATGAAGAAATTGAGGTTTAGAGATACCCAAGGGCCACGTGAGTGTGAGTGCCTGGAATT
GGAGCCTAAATCTAGTCATCTGATAGCAAAGCCTGTTTTCTTATCTGCTTTGCATTAAAT
ATAAGTTTAAAATAGAACAATACTGGCCAGGCTGGGTGGCTCACGCCTGTAATCCCAGCA
CTTTGGGAGGTCGAGGCAGGCAGATCACCTGAGGTCAGGAGTTTGCAACCAGCCTGGCCA
ATATGGCGAAAGAAACCCCATCGCTACTAAAAATACAAAAATTAGCCAGGCATGGTGATG
TGTGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCTTGAACCCGGGAG
GCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGTAAG
ACTCTGTCTTGGAAAAAAAAAAAAAAAAGAATGATACTATAGTCTGTGTTTATATGGTGG
GGAAGGTTGAGTATCAAAAAAATAACAAAGAGGAATGAATGTCTTAAGTGAATGCCTGTT
TCCCCATCTGCTTCCTCTTCTGCTGGGAGGAGAGACCTGGATCCCTAGAGGTTTCAGTTG
CCTCCAGAGCTGAGTGCCACAGGGATGCAGGGGAATAGGGATGTTACCTGTCGCTGGTAA
TTCAGAGAGATGATTCAGGGTATAGTTACCTGAAAGAACAAATTGCCATGCCAGACGTCT
TGGTTCTTATGACAGAGGCAAAGAGTTGCCTCCAGGATTGCCCAAAAGGAGACGAGTTCT
GGGAACCTCACGAAGAGGACCTTTCAGTGGAACCTGGGGAGATTCTCTTCCTCTCCATTG
GATTTAGGAAAGCTTAGAACCGGGTGATTCCTCAACCTCTTGATTTATTTAATTCTTTTC
TGGTTTTTTCTTGGCTCTACTCCAGGGGAATACGGAGGAGAACGGTCTGGCAATACCACTTT
CGGACCTGGCCGGACCACGGCGTGCCCAGCGACCCTGGGGGCGTGCTGGACTTCCTGGAG
GAGGTGCACCATAAGCAGGAGAGCATCATGGATGCAGGGCCGGTCGTGGTGCACTGCAGG
TGACAGCTCCTGCTGCCCCTCTAGGCCACAGCCTGTCCCTGTCTCCTAGCGCCCAGGGCT
TGCTTTTACCTACCCACTCCTAGCTCTTTAACTGTAGGAAGAATTTAATATCTGTTTGAG
GCATAGAGCAACTGCATTGAGGGACATTTTGATCCCAAGGCATATTTCTCCTAGACCCTA
CAGCACTGCCATTGGCCATGGCCATGGCAACATGCTCAGTTAAAACAGCAAAGACTAAGT
CAGCATTATCTCTGAGTCCACCACAAGTTGTGCATTAAACAACTTCATCCTGGCTCTGCA
GTTTCTCCTTATTCTTCATGATGTTTGCTTTGTAGCTGTTGACTGCTTTGTAGGTATTGA
GGTGGTGGGGGTGTGGTGGAAATAGGCCTGACTCTTGAGGATCCCTTAAGTCATTTTTGC

FIG. 10
(continued)

TTGGTTCTCTTTTTCCTTCTTTTCTTCTACTCTTCTATGATTCATCTCTTTGATTGTGAT
TCTGTTCTCTCTCTCTCTCTCTTTTTTTTTTTTCGTTTTTGAGACAGAGTCTTGTTTT
GTTGCCCAGGCTAGAGTGCAGTGGTGCCATCTTGGCTCACTGCAACCTCCGCCTCCCGGG
TTCAGGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATCTGACACTA
CGCCCGGCTAATTTTTGTATTTTAATAGAGACAAGGTTTTGTCATGTTGGCCAGGCTGGT
CTCGAACCCTTGACCTCAGGTGATCCACCTGCCTTGTCCTTCCAAAGTGCTGGGATTACA
GGTATGAGCTACCATGCCCGGCCCATTCTGTTCTCTTCTACCATAAATATATTTCTCCCC
TAACACTATATTTGTTTGCTTCACAAGATTCCAGCTGCTTTTCCACCAAGGCCTTTGATG
GAAGCTGTGCTGTGACCTCTGTAATGAGTCTGTGGGCTGCTGATTCTCCAGTTTGGGCTT
CATGATTATACTGGGGAATATTGGGTTTCCTAAATCTCATTCATTTCTTGGGCAAGTAGA
TATATGTGAAAGTGTTTATTTGTCCAGTTGTTAAAGAAGCTACCATTTATTGAGCCAGCC
TCTGAGCACAATGTTTTTTGTTTTGTTTTGTTTTTAATTTTTAAAATTATTTACTTCTTC
TATTTCAATAACTTTATTATTATTATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGC
TAGAGTGCAATTGAGCGATCTTAGCTCACTGCAACCTCTGCTTTCTGGGTTCAAGCAATT
CTCATGTCTCAGCCTCCCGAGTAGCTGGGATTACTGGTACGTGACAACATGCCTGGCTAA
TTTTTGTGTTTTTAGTAGAGACGAGGTTTTGCTATGTTGGCCAGGCTGGTCTGGAACTCC
TGGCCCCAAGTGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGTATTATAGGTGAGAGCC
ACTGCGCCCGGCCCTCTTTCAGTAATTTTGATGTATTTTTTTGTATATGATTCCTGTTTC
ATTCTGTCCAACCAGCACTCTGTATGGTATGTGCTGTTGTCCCCATTTCACAGATGCAGA
AATTAAGGGTCAGAGAGGTTAAGGGACTTACCTCAGGCACGTTGTACTGGAGAAGCTGAA
CTCCAAGAGCAGGTTTGGGCTGACTCCAAAGCCCTATGCTTTTTGCCAACATATTTTCAA
ACATAAATAGACAATTTTATAAATAGCTCCAAAGAGTAGACATTGTTTCTGTTGATATTA
ATGGCTTGGTTTTGAGTCTGAAACCCCCATGAATGATTCTGTTGTCCCTGCTTTTTGTCC
TTCTGCCCGCAGTGCTGGAATTGGCCGGACAGGGACGTTCATTGTGATTGATATTCTTAT
TGACATCATCAGAGAGAAAGGTGGGTCATCTGGTGGGCAAGAAGCGACAGTTTCTGTTTT
TAGTTTATGGAAGGAAAGTGCTCACGAAAACAGTCTGGGGAAGAGAGGTTGAATGGGAAA
ATTCTTTCACAAAAATCTGGGCTGAAGACTTCAGTGTGTCTGCCTGAGAACAGAAGTGAC
ACTATTTGAGCTTTTGGCATAAAATGAAGTCTAGGAGCTGCAGAACCCACTGCCATGGCC
TTTTGTTGCATACACAGTGGTGGTCTCTATCCAGCCACCTGACCTTGTTTACAGTATGGG
GTGATTTGTTGGCAAGTGAGGGAATCCTGACTTCTGCCACTTCGTTATTTATGTAGTCTT
CTGGGATCATTGGTATTGGTCAGAAGTTCAACACTGTAGCCATTGCAACATGCTCAGTTA
AAACAGCAAAGACTAAATTAGCATTGTCTCTGAGTCCACTAAAAGTTGTGCATTAAACAA
CTTCATCCTGGCTCTGCAGTTTCTCTTTATTCTTCATGATGTTTCCTTCGTAGGTGTTGA
CTGCGATATTGACGTTCCCAAAACCATCCAGATGGTGCGGTCTCAGAGGTCAGGGATGGT
CCAGACAGAAGCACAGTACCGATTTATCTATATGGCGGTCCAGCATTATATTGAAACACT

FIG. 10
(continued)

ACAGCCAGGATTGAAGAAGAGGAGGTACCAGCCTGAGGGCTGGCATGCGGATTCTCATT
CTCTTGCTAGGCCTCTTGGATACGCTCTCCTTTTGAGCAGGAGGACAGGCTCTGATAGAC
AACTGTTTGATTTCGGAATGGGAAACAAACTCCCAACTAAAAGGGCCTCTGGAAACTGTC
AATTATTCTCCACTTCTCAGCTCTGATTTTTCACTGCAGAGGAGCTTAGGGAAGGGCACC
ATCCTATCAGCCTGGCCTGCCAGATTGAAGAACTGCCATGCAGAAAGGTTCTGATGTTCT
CAGGCTCATGTGGCAAGCGTAAAACTCAAAGCCTTGAAGTTTCTAGCCTGTTCCAGCCTT
GATCCAGGCCATGTTTATCCTGATTCCATCCTTTAAAACGAATGCCTCACTCTTAATAGC
GCACGGCAGTTTGAACCACTAATTTGGTCGAGTTGGAAACAGTGAAATTTCAATTTTAAT
AAGCTGTGCATAATGAAGAGGAATGTGGAATTGGAGCCTTTCCATCTGAAGCTATTCATA
ACAGGCACAAAGCTGAGTTAATTAGGAATATGCTGAGATGAAGGAAATGAGGAGAGCTGC
TCTTTTGGGGGCTGTGCTTCTCTCCCCAACCCCTCAACCCCATTGCCATGCTGCAGATGG
GGTGGTGTCTAAACATCAGTGGCGAGTGCCTGCATTACTCTGCTCGTTGCCTTCCAGAGA
ACTCAGCTTCTCCAAATGCTGAGCTCTTTTCAGAATGGGACCTGCCACCAGTATTTGAAA
GATTTCTAGCCTAGCAGAACAGCAGCCACGTTATCAAAGTTTGGTTGGCCAAAGGAAGGT
ACTTGCTAATTAGTTTAGTAGGTTTTCAGTCCGCACAGACATACGGGATTGTTTTATTGT
ACATAGACATCTTCAGAAACAGTGTATGTATAGAAATGTAAGGTCAAAATTTGAACCTCA
GTGCTTTAAATCTGAATTTGTATTAACTGATATGAAATATTTAGACGGTTACTTTATTTT
ATATCTGTCTTCCATTATACTTAATTTGGCTCAAGAATAGTTAGGCAAAAAGTTGCCCAA
AGAGAAGGATCTCCTAGTAAATACAAAGAGAATGTAACATAGTTGCTACAAGTTGGAGCA
TGTTCAGGGATGTCTTTTTTTTTTTTTTTTTTTGAGAGAGAGGTCTCTCTCTGTTGCCCA
GGCTGGAGTGCAGTGGTGTAATCATGGCTCACTGCAGCCTCAATCTCCCAGGCTTAAGCG
ATCCTCCCACCTCAGCCTCCCAAGTAGCTGGGACTATAGGCATGCGCCACCACACCTAGC
TAATTTTCGCATTTTTTGTAGTGTCACAGTTTCGCCATGTTGCCCAGGCTAGTCTCGAAT
TCCTAGGCTCAAGCAGTGCTTCTGCCTCAGCCTCTCTGAGTAGTTAGGACTACAAATTTG
TGGCTCCATGCCCGGCTAATTTTTTTATCTTTATTTTGTAGAGACAAGGTCTCACTGTGT
TGCCCAGGCTAGTCTTGAACTCCTGGGCTCAAACAACCCTCCCACTTTGGGTTTCCAAAG
TGCTGGGATTACAAGTGTGAGCCACTGAGCCCAGTGACCTCTGGGTTTTAAAAATGTGTA
GGCTTCAATTATTTATTTTAAAAAATGAAATCCTGCAATATATAGTTTTCTGCGTTGTGT
GGTTTGAATCAATCTGGGAACTGGCTTGCTGGCTGATTGTGGTAAAGTAAGAAGTACTTA
ATTTAGTAGAAAGTTTAAATGGCAGACATAACATTAAACCCAGCTGATTTATAAATGAAG
CAAAAGAACAAAACTCATTCAGGATAATTGGTTATTCTAAAATACAGTCATTTCTAAAAT
TATGAAGTGTTCAGGACCTTTGGGAGTGAAAGAATTTGCTAAAGAAGGATCAGTGAAAAA
AAGGAATGATGGGTGAAGAGCTGTGGAGAAGGAAGAGAAGAAACAGCACAAGGAAGGAAG
AATATAAAATCAGATGTGGGAATCCAGGGGAAAGTGCAAACGAAGCAAGATTGAGAAAAT
TCTCAAGTTTTTATAAACAGTTCTCACACTCTGCCAGTTCCTTGGAGGTAGACTTTTTTG

FIG. 10
(continued)

TTAACTTCCAACTACAGTAGTGAAAAAAAAAAAAAAACCCTCAAATTTGCAAAAGCAGTC
TGTGGAATTTTCTTTACCCAGCTTTCCTGACTGTTAACTTTTTAGCACACTTAACTTTAT
CATTCGTTTATTCTCTCTGTTTAAAATTAAAAATGTAAATTTTAAAAAGTAAAATGTTTG
TTGGTTACAAACATTTATACCCCTTTGTCTCTAAATATCATTTCATTTTAAAAAATGAAT
AATCTAAGCCTACACATTCTAAAATGTGTATATTTTCTAAAAATAAGGGCATTCTCTTAC
ATAACCAATGTCACAATTATTTGATACAGTGATCAAAATCAGGAAACTAACATTGATATA
ACACTATTATCTAACCTACAGACCATCTTCAAATTTTGTCCTGCTAGTATCTTTTATGGG
TCCAGGGTCACACAGTGCATTTGGCTATAATGTATCTTTTTTCTCTTTTTTTGAGACAGG
GTCTCACTTTGTTGCCCAGGTTGGAGTGCAGTGGTGCAATTATGGCTCACGGCAGCCTTG
ACCTCCTTGGGCTCAGGTGATCCTCCCACCTCAGCCTCTCGAGTAGCTGGAGACCACAGG
TGTGCACCACCATGCCTGGCTAAGTTTTGTATTTTTTGTAGAGATGGAGCTTCGCCGTGT
TGCCCCGGCTGGCCTTGAACTCCTGGGCTCAAGTGACCCTCCCGCCTTGGCCTCCCAAAG
TGCTGGGATTACAGGCGTGAGTCACCACACCTGGCCAGTTATTAGTATGTTTAGTCTCTT
TAATCTGGAACAGTTTCTCAGTCATTCTTTATTTTTCATGACCTGGATGTTTTTGAAGAG
TTTAGGCCAGCTATTTAGCAGAATGCCTTTCAGTTTGGATTTGTCCAGTGTTTTCTCTTG
ACTATATTCTAGTCATGCATTTTTGGCAGGACTGTCACAGAAATGTTGTTGTAGTCTTCT
TAGTACATCACATCAGGTACACACTGTTGATCTGATTCATTACTAGTGGTGTTAACTTTG
ATCACTTGAATAAGGTGGTGTCTGTCAAATTTGTCCACCGTAAAGTTACTTGAGCAAAAC
GTAGCTGGGACTACAGGCGTAGCAAAAAATGTAGCAAAAAGTAGTATTTTTGCTACATTT
TTTTTTTAGGAACAAAGTATTTTTCCCTTTTAAGTTAATCTCTTGTCCATAAAGTTATTA
TTTTTCCCTTTTAAGTTAATATCTTGTGGGTAGATACTGGAGACTGCGTAAATTACCTAT
TTCTCATAATACTTTTTTTTTTTTTGAGATGGAGTCTCGCACCGTCTCCCAGGCTGGAGT
GCAGTGGTGCAATCTCGGGTCACTGCAAGCTCCACCTCCCGGGTTGACGCCATTCTCCTG
CCTCAGCCTCCCAAGTAGTTGGGACTACAGGCGCCCGCCATCACACCTGGCTAATTTTTT
GTATTTTTAGTAGAGACGGGGTCTCACCGTGTTAGCCAGGATGGTCTTGATCTCCTGACC
TTGTGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGTCACTGCGC
CCGGCTCTCATAATACTTTTTGCCTACTAATTTTATATTCATTGATTAAATTCTTGCCTG
AAAAAATTATTACTGTGGTATTTGCCAAATGGCAATTTTCTGTTTCCATCATTGCCTTTC
CCCCGCTTTTAAAAGTATAAGTGACAAAGAAAAACTGTATATAAAGTGTACACCATGATA
TTTTGATATATGTATACTTTGTGAAATGATTATCAAAATTGAGTTAAATAATGCATCCAA
CATCTCAGTTACTTTTTTTTTTTTTTGAGACAGAGTCTTGGTTTGTCACTAAGGCTGGAG
TGCAGTGCCACAATCTCGGCTCATTACAACCTCCACCTCCCAGGTTCAAGTGATTCTCCT
GCCTTGGCCTCCCCAGTAGCTGGGATTACAGGTGCCCACCATCACACCCGGCTAATTTTT
GTATTTTTAGTAGAGGTGGGGTTTCACTACGTTGGCCAGGCTGGTCTCGAACTCCTGACC
TCAAATGATCCTCCCGTCTCAGCTTTCCAAAGTGGTGGGATTACAGGCGTGAGCCACTGT

FIG. 10
(continued)

GCCCGGCCACTCTTAGTAAATTTTAAGTGTACATTTTTTTTTTTTTTTTTTTGAGATGGA
GTCTCACTTTGTCACCCTGGCTGGAGTGCAGTGGCATGATCTTGCCACACTGGAACCTCT
GCCTCCTGGGTTCATTCAGGTGCTTCTCCCACCTCAGCCTCCCAAGTAGCTGAGACTACA
GGTACCCGCCACCATGCCTGGCTAATTATTGTATTTTTAGTAGAGATGGGGGTTCACCAT
GTTAGCCAGGCTGGCCTCAAACTCCTGACCTCAGGTGATCTACCCACCTCGGCCTCCCAA
AGTACTGAGATTACAGGCATGAGCCACCACACCCAGCCACATTACGTTAGTATTAACTAT
AATCACCATGCTGTACATTAGATCTCCAAAATGTATTCATCTTATGTAACTTCAAGTTTG
TACCCTTTGACCAAAGTCTCCTTGTTTTCCCTACCCCCAACCCCTGGTAATCACTGCTTT
AATCTCAGTTTTTATGAGTTTGACTGGTTTAGATTCCACATACAAATGAGATCAGGCAGT
GATGGTTTATTTCACTTAGCATAATGTCATCCATGTTCTTGCAAATGACAGGATTTTCTT
CTTTTTAAAACTAATATCCATGCTGGACACGGTGGCTCATGCCTGTAATCCCAGCACTTT
GGAAGGCTGAGGAGGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACAT
GGTGAAACCCCATCTCTACCAAAAATATAAAAAATTAGCTGGATGTGGTGGCGCACACCT
GTGATCCCAGCTACTTGGGACACTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCGGAGG
TTGCAGTGAGCCAAGATGGTGCCACTGCACTTTAGCCTGGATGTTGATGTTGTTCCACTT
GTTTATTTTTATTTTGTTCCCTGTGCTTTTGGTATCAAATCCTAAAAACCATTGCCATGA
CCATTGTCATGTTACTTTCCCCATATGCTTTCTTCTAGAACTTTTAAGGTTCATCATTCC
CTTTTCTGTTTTTAGTTGCAAGCCTACTATAAGGAAGGGCTTTTCTTTCTTCCTTATTTA
TTTATTCATGTCTATCAGAATGGGCACCTTACTACTATTTTTGTTGTTATTGCTTGAATT
GACTTGAATTTGGCTAGTGGAAACCTTTTCAGATCGGGTACTCTGTCCTTTTGATCTCTT
TCCATTTTCAAGCACTTCTTTAGACTTAAGATGGTCTAGGCTCATCTTCTCCTTTCCCAG
CCATTTTTCAAAGGAACCTGATTCCTTTTAGTGAAGAGCAGTATTTTGAAACCAAGATCT
GGGCACTGGGTCTACTTGTTTGTACTGGTACAGTGTTCTTTGAATTGCTAATTAGCTGAT
CAATTACTGCTCTATTTGAGTTCCCTCTTTCTAAAACCTCACATATGTGTACAGACGGTC
CCTGACTTATGATGGTTCGACTTATGATTTTTGATTTTATGATGGTTTGAGAGCAATACA
TCCATTCTGTTTTTCACTTTTCATTCAACACTTTATTTTAAAATAGGGATTGTGAGATGA
TATTGCCCACGTGTAGGCTAATGTAAGTGTTCTGAGCACGTTTAAAGTAGGCTAGGCTAA
GCTGTGGTGTTTGGTAGGTTAGATATGTTAAATGCATTTTCGACTAGTGATATTTTCAAC
TTATGATGAGTTTATTGGGATGTATCCCCATAAAGTCGAGGAGCATTATACATATCTCTG
TATAACAGAGTGAGTTCCTTATACCTTTCATCCACTTTCCCCTGAAGTTAACATTTTACC
TAACCATGATACATTTATCAAAACTAAAACATTAACATCAATACATTGCTATTAACTAAA
CTAGAGTTTAATTGGATTTTGCCAGTTTTCCAATGAATATCCTTTTTCTGTTCCTTGATC
CAATTCATGGTCACACACTGAGTTTGGTCACTTGTCACTGTAGTCTTCTCCAATCTGCGA
CAGCTTCTTAGGCTTTCCTTGTTTTTCATGTACTCTTGACGATTTTTAAGAGTACTGGTC
AGATATCTTGTAGGATATCCCACAACTTGTGTTTAATCTTATGTTTTCTCATGATTAGAC

FIG. 10
(continued)

TTGAGTAATGGATTTTTGGGAAGAATACCACAGAGGTATATTGTTAAGTGTTCTCATCAC
TTGGAGGTAAATGTTATCAACATGGCCTGGTGATGTTAAACTTGTCAGTTTGTTTAGTTA
GTATCTGCCAGATTTTTCTCACTGCATAATTACAAATCCTCCTTAACTTATGATGGGGTT
ACAGCCTGATAAGCCCATCATAAATTGAAAATATCATAAGTCAAAAATGCATTTAATGCA
TCTAAACTACTAAACATCACAGCTTAGCCTAGCCTGCCTTGAACGTATTCAGGACACTTA
CATTAGCCTACAGTTGGGCAAAATCATCTCATGGGAAGCCTGTTTTATAATGTGTTGCAT
ATCTTATGTAATGTGTTGAGTACTGTACTCAGAATGAAAAACAGAAGGGTTGTATTGCTT
TTGCACCATCATAAAATCAAAAAAACCATAAGGCAAACCATCATGAAGTTGGGGACTGCC
TGTACTTTTTTCCTCTTTCCCTGTTCAATTCCTTGGAAGAAAGTCATTTAGTTCAGACCA
TACTCAAGAAAAGGGAAATAAAGCTCCATCTCTTGGAGCTTAATTGAAACTGGAATGACT
AGTTTCTATATACATTATTTAGAATCCTTTTGTAAGAAAGATTTGTTCCTTCTCTCCATT
TATTTATTCCATTATTTATATTGATAGAGACGCATGTACATTTATTTTATACTTTGGGTT
ATAATCTATTTTTCTTGCTCAAATTGTTACAGCTTTGGTCACTGGGAGGTTCTTCAGATT
GGCTCCTGTGTCATTTGACATGTCCCCACCCTCTCGTTTCTGAGTACTTCTCTACTTTGG
CATTACAAAAGATGTTCCAGGCTCCTCTTATATTTTTCCCTGCCGCAGCCCTAGAATCAT
CCATTTTTCTATGGTGCCCTGGTTCCTTTTACTTTAGATGGGGGTTTAGAAACCAATCTG
GGTGTTGGGTGTGCTCATTGCTACTGGAATCACTGCTTCTAGGCCCTCTCAGCAGATAGA
GCTAGAAAACATATGGCTGTATATGAATCCATGGATTCATATATATCTATAATTGTTTTC
TGTATCTGGCCATCTATATATATATTAAGCTAAACATGAATTCATACTGATGTCTCAGAC
TCGAATCCATTGCCGCAGGGCTCATTCTTGCCTTCCTCTTGCTTATTTGTGACTTCTTTC
TCTAACAGGGAGAAACCCCAGTCTCATTATCACCAACCTATCTACTCATTTGTTCAACCC
TGGTATAGGTGTAAAGTAGTTTCAGAATTACTAACCTATACCCATGTGAGAATTGTATTT
GCACTTCTTGTTTGAAGGAAATACATACAACACAGGTAGCGTCTCTACACTTCAGTATAC
AGAGATCTGAACAGTGTTCTCTCTGAGTGAATCATATTGCAGGACAGAAATTACTTTTAA
AAATTCTGTAATGGGTCAGGCCTATAATCCTAGCACTTTGGGAGGCTGAGGTGGGCAGAT
CACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAAAATGGTAAAACCCCATCTCTACAA
AAAATACAAAAATTAGCCAGGCGTAGTGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGG
CTGAGGCACGAGAATCACTTGAACCTGGGAGGCAGAGCTTGCAGTGAGCTGAGATTGAGC
CACTGCACTCCAGTCTGGGCGACAGAGCGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAA
AATTCCATAATGATAGCAGAGCTGGAATAGAAATGGGATTGCACAGGCTGAATCTGAGTT
GTTGCAACAGTAAACGAGCAAGATTTAAACTGGCCTTGTGTAGCACTTGCTATTTGGCTC
CTCATATTTTATTAGACGCTTATTCTTTTTTGTTTGGTGTCATTCCTTTGAGAAATATTT
GAGTGCCTTTTCTGTTGCAGACATTGATTAGATGCTGAGGTTGTAACAATGAAGAAGATA
GCCATCGCTGTTGCCTCATGGAACTGAAGTTTTACTAGATGTAAAATTTGAGTTAACATG
AGGCCGTGCCCCTATGTGCCCTATTGTTTCTTCACACAGCTCCCTTCATCTCCTTGGTCC

FIG. 10
(continued)

AATGAAAAGGTTTTTTCATACTTGTTCATTCATTCCTGCATTAATTAAAGTAGGTTGTAC
TGTGCCAGGCACTGGGAATATTTAAGTAGTTGTGTTCCTGAATTGGAAATGAATCCAGCA
TGGTTGGAGTAGAAGGAGCTGGGGGGCAATGTGGAGTGTGATGGGGAGATTGGAAAAGTA
AGCTGAGACCAGATTTTTCAGTTTGGAGGGAGAGGTGGGCCTTGTAGGCCATATTACAGA
TTGTAGACTTTATTTGGAGGGACATGGAAGTCATTGAGGAGTCTGAAGCAGGGGAATGAC
ATAAAAAGATCCTCATTTTAGGCCGGATGTGGTGGCTCACGCCTGTAATCCCAGCACTTT
GGGAGGTTGAAGTGGGTGGATTGCTTGAGGCCAAGAGTTTGAGACTAGCCTGGGCAACAT
GGTGAAACCCTGTCTCTATCAAAAATACAAAAATTAGCTGGGCATGGTGGCTCACACCTG
TAGTCCCAGCTACTTGGGAGGCTGAGGCATGAGAATCGCTTGAACCCGGGAGGCAGAGAT
TGCAGTGAGCCGAGATTGTGCCACTGCATTCCAGCCTGGGTGACAGAGTGAGACTTCGTG
TCAAAAAAAAAACAAAAAACCCCTCATTTTGAAAGGGAACCCTGGCTTGAGGGTGAAGAA
TGGGTGGGCACTAGGCTAGAGCAGCTGCAGGGTCAGTGAGGAGCTGCCGCAGTGCTGCAC
GTGAGAACCCGTCATGGTTTGGTCAGGGTGGGCAGGACTGACAGTGAGCACAGAGCGAAG
TAAAACCAGCAAAATTTCATGATTGGATAGTGGAAGGAATCATGGTGTTTGTAGTCTTCA
AATGTGAACCCAGAGTGCACTGGACAAGTAGTCTAGGCTGCTCTGTAACCAAGGCAAGTG
TTTTCATTTTACCCTCTCTTCCTGCTCTTGGCCTTTGGATTTTTTGTAATTTAAGGTTTA
TGAATGTAATCAGTTACTTAACATGGAAAGATACTTAATACCAGATGATTTTGGAGTCTT
GTGATCAATACCTTCTCTCAATCTTGGGTGTGTGTCAGTTGGCAAGGCCATAAAATTTGT
TATAAACATTGCAGAAGGCTTGGTTACTGTGCTGTGACGTTGAATTTGGGTGGAGATAGA
TCAATTTCAGTTGATTTTCTAGGCTTCAGAAACACATTACCCTCTACTCCACAAACACAA
ATCAAAACAAAACAATCCCTATTCCCTGAGCATTTCTCTTGATCTATAACACAGCCTGGG
CTGTCACAGTACTAAGACAAGCCCATCTGATTTGTGAGTCAGTTTTATTTCTTGGTCTTC
TACATAAGCTAAAAAGTTTCAACATTTTAATGCTTTTCCTTGGATTCCTTTGAGTCATTG
AAGTAATTCCTGTTTCATTTGTACTAATTATTCCACACTAGAAAATTCTGTTGTAATCAC
TTTATGTATTAATAGAAATACTGATTTTTATTTTCAAGGAAGTATTGAGTAGGGAGGGGG
AAATAGGGATTTGCTGTTCAATGGGTATAGAGTTTCAGTAATACAAGACAAAAAACTTCA
GAGATCTTCTATACAGCAGTGGGTATATAGTTAACAATACTGCACATCTAACAGTTTGTT
AAGAGGGTAGATCTCATGTCATGTGTTTTTAAAAATTGCTTTTAAAAAAAGTATCGAGTA
AAAAAGCAGTTTTACTCCTCAGTTTCTATTTATATTTAAAATTTTTATTTAAAAAGTGAG
TTGAGATTTTTAAACCTCAGGATAAGTTTTATTTTTTAAAAAATTTATTTTTTTATTATTT
TTTGAGATGGAGTCTCACTCCATCTCAAGTCACCCAGGCTGGAGTGCAGTGGTGTCTTGG
CTCACTGCGACCTCTATCTCCCAGGTTCAAGTGTTTCTGCTGCTTCAGCCTCCTGAGTAG
CTGGGATTACAGGTCTGCACCACCACGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGG
GGTGTCACCATGTTGGCCAGGTTTGTCTTGAACTCCTAACCTCAAGTGACCACCTGCCTT
GGCCTCTCAAAGTGCTGGGATTACAGGTATGAGCCACAGTGCCCGGCGGGATAAGTTTTA

FIG. 10
(continued)

AAATAATATTCTCTGCTGGCTGGGCATGGTGGCTCATGCCTGTAAACCCAGCACTTTGGG
AGGCTGAGGCAGGAGCATCACTCGAGGCCAAGAGTTTGAGACCAGTCTGGGCAACATAAT
GAGACCCCCTCTCTACAAAAAATAAAAAAAATTTGGCTGAGTGTGGCATGTTCCTGTAGC
TATCGGGAGGCTGAGATGGGAGGATTGCTTGAGCCCAGGAGTTTGAGGCTGCAGTGAGCT
ATGATTGCACCACTGCGCTCTAGTCTGGGTGACAGTGTGAGACCCTGTCTCTTAAAAAAA
AAAAAAAAAAAGGCCAGGCACAGTGGCTCAGGCCTGTAACCCCAGCACTTTGGGAGGCCG
AGGCGGGTGGATCACTTGAGGCCAGGAATTTGAGACCAGGCTGGCCAACATGATGAAACC
CCGTCTCTACTAAAAATACAAAAATAAGCTGGGTGTTGTGGTGCACACCTGTAATCCCAG
CTACTTGGGAGGCTGAGGGAGAGAATTGCTTGAACCTGGGAGGCAGAGGCTACAGTGAGC
CGAGATCACACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAACAA
CAACAACAAAAAAACCAAATGTTCTTGCCAATTCTTCCATTTAATATTTAATTTTGAATT
ATATTGTATCTTTCTAAGGATTGTTTCTTATATAAGCAAAGATTTTTCAGTGCTAAACAT
TTACGACTGCTATTCAGAAATGGTTATTTACAAGTCTTTTTGTTTTAAGAAAATGGCTGT
TCAAAAAATTAAAATAGTATATAAACCAAACAAAATATTTTTGCTTTGGATGTCTGTTTT
GCAGCTTCTTCCCTACACTATAAGTTCTTACTGACTGCTTTATCACTTAATAAATTGGTT
TGGCTACTTTAACAGAGGCAAATAGTATCAGGCAAAAAATTATTTTTTATTTTTATTTTT
TGAGACAGTCTCACTCCATCACCCAGGCTGCAGTGCAGTGGCCTGATCTTGGCTCACTGC
AACCTCCACCTCCCAGGTTCAAGCGATTCTCATGCCTCAGCCTCCTGAGTAGCTGGAATT
ATAGGCATGCACCACCACACTCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTGC
CATGTTGACCAGGCTAGTCTTGAACTCCTGACCTCAAGTGATCCATCTGCTTTGGCCTCC
CAAAGTGCTGGGATAACAGGCATGAGCCACCATGCCCAGCCCTATTTTTTATTTTTTAGA
GATGGGTCTCGCTTTTTTAGAGATGGGTCTTGTTGCCCAGGCCAGAGTGCAGTGGTGCGAT
CATAGCTTACTGCAGCCTTGAATTCCTGGGCTCAAGCAATTCTCCTGCCTCAGCCTCCCG
AGTAGCTGGGACTACAGGCCTGTGCCACCAGGCCTGGCTTGTACATTAGTATTTGATATG
GCTACCCTAAGGGCAATCCTATAGTGAAGTCAACATTAGATAATGATGCTCATCTGATGG
ATTAGATTTTCAGAGTTGGCTGTTTCCAGGTGCCTATAGGAGTAGAAAAGGGTGACAAAC
CTCCTAACTAGATGTCCTACCAAATATAGTTCACTCCACATCTGAGATGAGACTGCATGA
CTGCTGGTTTTCTTTGCCTTTTCCCCCCCAGGGTATCATCAGAACCAAAAATAAAGTTTT
AAAGGTGGGTCAGGTGTGTGTTGGCTCATGCCTGTAATCCTAGCACTTTGGGAGGCTGAG
GCAGGTGGATCATCTGAGCTCAGGAGTTCAAGACCAGCCTGGCTAATAACATGGTTAAGC
CCCATCTCTACTAAAAATACAAAAAGTTAGCTGGGCATGGTGGTGGGCACCTGTAATCCCA
GCTACTCAGGAGGCTGAGGCATGAAAATCGCTTGAACCCCAGAGGCGGGGGTTGCAGTGA
GCCGAGATCATGCCACTGCACACTAGCCTGAACAACAGAGCAAGGCTCTGTCTCCAAACA
AACAAAAATGGTGCCAGAGTCTTTTCCAGGGCTGAGGGGAGATACAATGAAGTGTGTTAT
TTTTTCTGATAAGAGTGCTACCATCTTTCATTCTTGTGTGCCATTTCTAGTTGGGGTGAA

FIG. 10
(continued)

TTTGTTTTCGGAGTTCCTTTCCCAGCTGTTTGCCTGAAAAACCATGAAATGTGTTCCACA
TGAACTATGAAATGATTAGATGCTAATGTGGCAAAGAAAGTGTGAATTCTCTTGTAGAAA
CAGGGACATTTGGTTCGGTACAGTAAGTTGTTAATGCGTGACTCTGTGCTTTCAAATTCT
GTGGTTCAAAAGTACTTTTCACTCCTACTGTGTATTTACCTTGAGAAGGTGAATCCCCTA
ACAATTTGGTCAATGTATCAGTATTCTCAACCCGTCTATCAATTTTTTTTTCTTTCTCCC
TCTTTTTTCTTTTTTTGGGCAAAATACCTTTTTTGCTTTTTATCCCCTTAAAATAACCAT
TGTCCCTCACATGTGCACTCTTCCAAATTTCAGAAAAGCAAGAGGAAAGGGCACGAATAT
ACAAATATTAAGTATTCTCTAGCGGACCAGACGAGTGGAGATCAGAGCCCTCTCCCGCCT
TGTACTCCAACGCCACCCTGTGCAGAGTAAGTAGTGCTGAAGGAAATTCTTTTTACCTGG
TCATGGTGGTTTAAAAAGGTTTAAAAAACAAAAACAAAAACAAAACACAAGTTTGTAGCA
CATGCCTTTCACTGGTGCACGTTCCTGTTGCCCTACTGTTAGTGTATCTGTGACTGGTGA
TATCTATTGATTGTGTTAATGCTATCTCAACCACGTTTTAATTTTCCTAAGCTGGCCAGG
CACGGTGGCTAACGCCTGTAATCCCAGTGCTTTGGGAGGCCGAGGTTCATGGATTACTTT
GAAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAAT
ACAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAATCCCAGCTACTCAGGAGGCTGAG
GCAGGAGAATCGCTTGAACCCAGGAAACGGATGTTGCAGTGAGCCGAGATCATGCCACTG
CACTCCAGCCTGGGCGATAGAGTGAGCCTCTGTCTAAAAATAAAATAAAATAAAATAAAT
TCCTAAACTGAAGGCTGACTGCTATGCTAGCTAGGATTATATGGGATTTTAAGTATATCA
AGTGGTGGTTCTCCAAGAAGAATCTAATTTTTCTTTTGATGGGCTGGGGATTGTAACAAA
GGAAGGTCATATGTCTTAATGATGTGTTAAGGCTCTTTGCAAAATCAAAGTAAATAAATT
GACCACTAATGTGTCAGCCCAGCCATGTTCTGCTCATTTGCCACCAGTCAACAGAAATCT
ACTTTGGGTGTTTAAACCAGGAGTCAGCAAACTACAGCTCACAAGGCCAGATGTGGGCCA
TGGCCTGTTACTGTATGGCCTGTTAATGGTTTTAAAGGGTTGTAAAACAAAAGAACACAA
AACAAAGACCCAATAACAAAACAAAGCCCGAAGAATAATATGCGACAGAGACCATGTATG
GCATATAGAGCCTAAAATACTGACTCTCAAGCCCTTCCCAGAAATCCTTCCCGACTCCTT
GTTGAAAACACGGTAGGAAAGCATTTGTCAAATTGAGGATATGAATAGCAATTGTAAGTT
ATTATTTTTCTATATATTCGAAAGTCACTTGCTAGTATAACATTTACCTTTTATTTTTCC
CTAAGAATCTTCTCTCTGTTTGCTTTCGACATGGATTTTTAAACCCCTGCAGATTTTAAT
ATTCTATATAAATGTTTTAGGTGGCATATATGAGGTTTGTATTAACATTTGCTTTCTATT
TAACATTGAAATGAAATTATACAGCAGAGGTATTTTCTCGTCCAAGTTGCCACTTCTTTC
TATCTTTTTTCTTTTCTTTCCCAGTGGACTGCCTGGGAAAATTGATATTTTAAATTGCTC
TCTGCAATAATTTGCAATGGAACTGGAATGCCAGGGTTCTGAGTCCTTGCCAGACAGCTC
GTCCCTCCTGTTGGCATGACTGAGTCAGCTGTCATGATTCCCTCAGTACCAGTGGCATGC
CTGTGACAGACAGCCTGTCTGCCTTTCATTCCCGTCGTCTCCCTTGTAGGGTTCAGATCC
AGGATACACTGGTCCTGGAGCCCCTCTCAGCCTGGCACCCACAGCTGCTGGGTTCCTTAC

FIG. 10
(continued)

TCTCCTGGACTGCTCTGATGTCATCTCCCTGCTCAGCAGAAAGAAGTCTGGGATCTTGAT
GCTTTGGCCCTCTGTCCTAGGCCCTAAACCACCCATTGCCCTTCACATAACCTGAGCTGG
GGCTAAATAGATCTCTCATCACTGCCTGCCTGCTCCTGTATTTTCCCTTCTTGGAGCTTT
TGCCTGTTCAGATCCCTCTACTGGAAATTAATAGGATTTCATTCTATGTGTGCATTTCCA
ACCTTTCTTCACAGTGCGATCCAAATGCCTCATCCTACAGGCCTCCTTAAAACAACCTGC
TTTCTGCCAGACCCCAGGGAGCACCAGGACTTGAGGCTTTTATTGCACTTCTGTTGTTTT
TTTGAGATGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGATCTCTGCTCA
CTGCAACCTCCATCTCCCGAGTTCAAGAGATTCTTCTGCCTCAGCCTCTCAAGCAGCTGG
GACTACAGGCATGTGCCATGACACCCGGATAATTTTTGTATTTTTAGTAGAGACGGGGTT
CACCATATTGGCCAGGCTGGTCTCAAACTCCTGACCTCGTGATCCACCCACCTGGGCCTC
CCAAAGTTCTGGGATTACAGGCGTGAGCCACCATGCCCAGCGTTATTTCACTTCTGCCTC
TGTAATTATATTGCTGTATGGCTATCTCTTCTCTCCCTGGGAATGTCAGGTCCTAGGCAC
AGGAACTGTGTCTGTACCATATCTGGTGCCCAAAGAATGTAGTATGTGTTTTATAGATAT
CATGTAAGCTTAAACAGCGTGGTCTACATTTTTGTAAATGTCTTTCTTTTTCTTTTCTCT
CCAGAATGAGAGAAGACAGTGCTAGAGTCTATGAAAACGTGGGCCTGATGCAACAGCAGA
AAAGTTTCAGATGACAAAACCTGCCAAAACTTCAGCACAGAAATAGGTATTTAAATGCAA
GTGCTCTATTGGTTAATTGTTTATATAATTGGCAGTATTTTTAAGCAGGCAAGCAATTTG
GGAATGTTTTAGCAAAGTGTACCATAATTGAGTTTTACAAACCAGGCTCCTTTTTCCTCT
CCCTGTACTTCTTTTTCCAAGATGGTTTTAGTTTAGAGTTCATTAAACATTAAAATCAAA
CACAGAATTAATTCTGCATGAGGCAAGGCTAGCACTTATTCCAGAGAAATGGCTGATACT
GGTGGTAGAGTGCAGGTATCACTGTTCCTGCAATTTTTATTAGAGTTGGTTAGCCCAGGC
TGTGCTGGGGGATGATCTGTAGGGATCTGGGAAGCATCGGGACTCAGCACTGGGTGGTTG
GGAGTCAGGAAGCCTGAGTTCTCATTTCAGTCAGTCTCTGACCAACTGTGTGGCATGGGG
TGCTAGACCACTTGGCTGCCGACTGGGTCACCGACATCCCTTCCAGCTCTGCTGCTGGAA
ATTCATCTCTCCCATATGTTGCCTCCCCATCAATTACGTTTTTTAAGTGTGACCCAAGTA
TATGATGTATGTTTTCATGATAAATTAGAAACTTATCTGGGCATGGTGGCTCATACCTGT
AATCCCAGCACTTTGGGAGGCTGAGGTGGGCGGATCACCTGAGGTCAGGAGTTCGAGACC
AGCCTGACCAACTAAAATAGTAGAGACCAACCCGTCTCTACTAAAAATAGAAAATTAGCT
GAGCATGGTGGTGCATGCCTATAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAGGCAG
CGGTTGCAGTGTGCCAAGATCGCGCCATTGCACTCCACCTGGGCCACAAGAGTGAAACTC
CATCTCAAAAAAAAAAAAAAAAAAAAAAAAACTCAGTGTCAGTATTTCATGTCGAAATTC
CACTTCAATGGGTAGTGTAGTTAAAAGCTCTAAGTCTACCTTAAAATCACCTAATGCTTT
GTTAAGCTTTTAGATATATGTTCCTTAAAAACTCTTAACTTATTTCTTCCCCAGATGTGG
ACTTTCACCCTCTCCCTAAAAAGATCAAGAACAGACGCAAGAAAGTTTATGTGAAGACAG
AATTTGGATTTGCAAGGCTTGCAATGTGGTTGACTACCTTTTGATAAGCAAAATTTGAAA

FIG. 10
(continued)

CCATTTAAAGACCACTGTATTTTAACTCAACAATACCTGCTTCCCAATTACTCATTTCCT
CAGATAAGAAGAAATCATCTCTACAATGTAGACAACATTATATTTTATAGAATTTGTTTG
AAATTCAGGAAGCAGTTAAATTGTGCCCTGTATTTTGCAGATTATGGGATTCAAATTCT
AGTAATAGGCTTTTTTTATTTTTATTTTTATACCCTTAACCAGTTTAATTTTTTTTTTCCT
CATTGTTGGGGATGATGAGAAGAAATGATTTGGGAAAATTAAGTAACAACGACCTAGAAA
AGTGAGAACAATCTCATTTACCATCATCTATCCAGTAGTGGATAATTCATTTTGATGGCT
TCTATTTTTGGCCAAATGAGAATTAAGCCAGTGCCTGAGACTGTCAGAAGTTGACCTTTG
CACTGGCATTAAAGAGTCATAGAAAAAGAATCATGGATATTTATGAATTAAGGTAAGAGG
TGTGGCTTTTTTTTTTTTTCTTTTTTCCAGCCGTTGACCAATTATAGTTGGCTGTTGAGT
GAGAAGTTTGTGGTGGCAAAACGTTTGCCATATTTTCTTTGCATTTGAATAATTGTCTTG
TACTTAGAAAAAAGGCGTCTATGAATGACCACTGTTTTTTGGTCGCCAAATGTTCCTGACA
AACTTATCCCAAAACTTTAGTGCCTTAAAAAAACCTGCCCCCAACTGTTAGTCAATCTGA
GCTGGGCTCAGCTGGGCTGTTCTTCTGCCAGCCTGCAGGTGGCCACTCATGTGGTCAGCA
GGTCGGCGGAGAGACTGGCATGGCTGGCCTTCTCTCTGCCTGCAGTCCTGAGTCTCTTC
CTTCTTCGTGTAGTCTCTTTCAGTGGCCTGGCTGGCAGGGTAGCTAGACCTCTCACATGC
AGCTCAGAGCTCCCAAGAGCTCAAAAGCAGAAATGCCCAGCCCTTCTGAAAACTTAAGTC
CAGAATTGTCACAGTGTCCCTTCTACTTCCCTCTATTGATGATGATGATGATGATGATGA
TGATGATGATGATGATGATGGTTTTTTTCTAATCAGAAGAAAGCTGGGTATGCCCTC
TACTTACTAAACAAGTCACAAGCCCAGCTCAGATTCAAGAAAAGGGTGTGAAGTAGAGGT
GCAGTTAAGTGGCGGCCCACTAGTCTAACAGACGGTCACAACCAGTGCCATGGAAAACCA
AGGATATTAGCAAAAGCAGAAGTTGCTAGTGACCTTGGGAAGCCGAAGCTGCTTACAGTA
GCTGGGACAAGCTGAAAGTCAGACTAAGAAATAAAGAGAGGGCCTTCAAGAAGCTTCCTG
AATGATTTCTGCTAGCCCTGAGCCTATTTTTGGAACCAGCACTTGGGGAAACTGATCTTG
TGAGGATGGATGTGTTTAGGGACACAGGGCTTTTGAGAGCAGCACCACCCCACTGGGGCA
TCCCCAGACTTGGGAAACCTGACTCTTTCTTAATGCCACTGGGTTTTAGTCAGGCCACAG
TGAGAAGGAACAGCCCTAACAGGCCTCCAGCCAGGTTGAATGAGCTCATTTTTGTTGTAG
CCAACCAGTRAGATTTGCTAATGTTCTACATTAAGTGCCTTCTCCAAAGACATCCCTCTT
TGCCTCATATGTTGAATCATCCAGTGCGGATATTTCAATGAAAATATCATTGGTTGACTT
TTGTGATGGTAATAATGCTATGGCATCTTTGCCATGAAGTTGTGGCCTCCTTGCATTCTT
CTGACTTTGGCTTCTGAAAGGAAGGCCTAGATCCAGCCCTGGTGGTAGTTCCTTTCTGAG
GTCTCTCAGTCCCTTGAGACTTTGGGCTAGTTTGGCTGCCATTCTCACTGACAAAATGTA
TATCAGCCCCCACCTCCACCCCCCAATATTCCTTGAACTTTGAATTGCTTCAGAACACAG
GTGTGGCCTGAAGGTATTCCCTTATTAGGGAAGTGTCACTGCTGTCTTCTAGTCAAACTT
GTAAAGAAAAAGATTCCAGTTCAGTATTTGCAGCAAGAAGCTTGAATGCTGTTCTTTTTA
TCGCATTGTTACATCGACTCATTCTCCATTTTGCTTTGGTTTTGTCTTGACTTGACTTGA

FIG. 10
(continued)

CTTTGGGGGTAAAGTCTTTCACCAGCACACAAGAGTTTGATTGTACAAATATATCTTCTG
CATTAACATCTCTGCCTGTTGCTTAAGATCAGTTGCTTTTATACTCAGAATGGAAATACC
TGATCTTGGCTAGTTTTGTTATAAGATATTGATTTCATTTAGATTTCCCTCCACCAGGTC
AGCAAACTATCATGTTCTTATGTAAACTTAGGCCAAGGCCAGAGTTATCATAGTCCCTAG
GTTGCTACGGCTTATCATGTGCTTGGTAAAAGGTGATCGCAGGTTCTCAGACGAGTTTAC
TTTACATGACATGGAATCAGGCACAGAGGCTGGGATGATGGAGAAAGCTCGAGGTGAAGT
TTTAAAAAAAAAACTTGTGGAAAGCAAAGTTCCAAAGAGGTGGTTTCTGAGGAAGTCAGAG
CGCCCAGGGCCAGAGCAGTCAGTAATGGGTGAATGAGGTTGTTTGGAAAGTGGCTGTGAC
AGACACATGGATGCCATCTACTTCTAGGTTGCTGGTGGGTATTAAATATGCACAATATTC
CATAGCTCACTGAGGATTTTAAAATTATAAGCATAGGATTTTATATTTTGGGGTGAAAGA
ATTATCTGGCACATTAGGTATTGGAGTTTAAAAAAAAAGCCAAATTTCACAGTCTTAATA
ACTTTTTTTTAAAAAAAACTAAAAGGCCCTTCATGTCCAGTGTGTGGCCCTTCTGAAACTT
ATGGTCATCTCTCCCACTGAAACCAAGGTCTTTTCAAATGTGGCTAAATGGGGATGAGGA
CACACGGGTAGGACTTTCTTGGTGTGTCTGCATTCTTTAAAGAGCCAAGTTGCTTCGGGG
AAACAGCCAGGAAAATGGTCAAGATTATTTTAGAGGTTATTTTATTGGGGATTTTAAGA
ACTAATAACATCTTCAGTTATTTTTAATTCACGGGGATGTCGAAAGGTTTGCAATTGTCA
AGTGTTTTGTTGTAGCTTAGTATCCATAAGGCAAACTTAGACTATAGACATAACTACAAA
GCCAGTGCAGCTTTTGTTTTCTGTATGTTGTTGGGGATCAACTTTCACACATAGCAAGC
ACATGGCCTCCCTGATGTCAGGATGCCTTTGTTAGGATCTGTATTTGCCCTTAATTTTGT
TGAAATCTTTTTTCCTTCTTCCTCTTGAAAAGTTCCAAAATATAGTTTATTGTATCTTTC
ATCACTAAAAATTGTTCCTTTTTCACTATGGCAGTTCACACAAGGCAAAAACTATTGA
ACAGTTGGTTTTAGTGTGTTGTATAACTTTGCTGTATATCAAACTAATTTTGACAAGTTT
TCATGCTAAGCCTCAAATCATGTAATTAATAATTTGCCTGTTTATTTATGACCTAATTGT
GATTCTTTTATTAATAAAAGCTAATGGCAAAAGGATCCCTGATTAAGCTGATGACTAGAC
CTACAATTAATTTTCCTGCAGTATATGAAGTATTGTACCAGAGTATTAAAAGATATGTAA
TATTTTATTGATAAATCTATCCTTTAAAAGGAATACGTTTTAGGATGTCATCATTTTGAT
GTGAATCATGTAAATGTTGATAATATGCTGTTTATTATACATTTAGTGTTTCAAGAGATT
CACTTAATTGCCTTTTTGCCCACGTATATTATGTAGTCTATTTGCAACTGTTCTTAAAAA
AATGACATTAAAAGAATAGTTTATGTAGAGAAACATTAGTGGATGTTAATTGTCTCCCCA
CCTATATTTATGGGTGTTAGCGCAACTGCTTTGCTAGTTGCAAAGCTGTATTATCAGAGT
AAAAGTGTATTTGTAAACTGTATGGGAACTAAAAATTAGGAATAAAACCATTTTCTTATA
TGATGGCATTTGTCGTTTGCTTCATCAGAAATGTCCAGGAAAAAAATGGGATTATTGGTC
ACTCCACCTCTCACACTGGCAAAATACTGACATTTAGCAGCTCTTATCTAGAAGTGACTT
GGAACATAGAATAAAGGCATGAGTTCCTGAAGAATTCATTGAGTGTTTCCTGTAGAAATA
GCTTTAGGAGATAGGGAGTTCTATCTGGGAGAACATATGAGTAACTCAAGAGTAAAAAGT

FIG. 10
(continued)

ATAGTCTGTGTAAACTATAGAAGAAATGCTGGGCATGGTGGCGCGCCCCTGTAATCTCAG

CTACTTGGAGGCTGAGACGGGAGGATTCCTTGAACCCAGGAGCCCAGGAGTTTTAGACCA

GTCTGGGTAACATAGTGAGACCCTTTCTCACCTACTCTCACTGCATGCCCCCCAAAAATA

TATATGTGCGCGCACGCGCGCGCACACACACATACACACACACACACACACACACACACA

CAGAGGAAATTGTTAGAAAACACACAGAACTGAATGTAAATAGTATTAGGTGGGAATAAG

AAGTAAAGGGATGGTAAGGAGGCTTGGAGGAGGAGTAAATTATCTGCTATGGGACATCAG

CTC

FIG. 10
(continued)

MTSRRWFHPNITGVEAENLLLTRGVDGSFLARPSKSNPGDFTLSVRRNGAVTHIKIQNTG
DYYDLYGGEKFATLAELVQYYMEHHGQLKEKNGDVIELKYPLNCADPTSERRWFHGHLSGK
EAEKLLTEKGKHGSFLVRESQSHPGDFVLSVRTGDDKGESNDGKSKVTHVMIRCQELKYD
VGGGERFDSLTDLVEHYKKNPMVETLGTVLQLKQPLNTTRINAAEIESRVRELSKLAETT
DKVKQGFWEEFETLQQQECKLLYSRKEGQRQENKNKNRYKNILPFDHTRVVLHDGDPNEP
VSDYINANIIMPEFETKCNNSKPKKSYIATQGCLQNTVNDFWRMVFQENSRVIVMTTKEV
ERGKSKCVKYWPDEYALKEYGVMRVRNVKESAAHDYTLRELKLSKVGQGNTERTVWQYHF
RTWPDHGVPSDPGGVLDFLEEVHHKQESIMDAGPVVVHCSAGIGRTGTFIVIDILIDIIR
EKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIETLQRRIEEEQKSKRKGH
EYTNIKYSLADQTSGDQSPLPPCTPTPPCAEMREDSARVYENVGLMQQQKSFR

SEQ ID NO: 61

FIG. 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/013113

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/00 (2016.01)
CPC - C12M 23/16 (2016.02)
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)
IPC(8) - B07B 1/00; C12M 3/00; C12Q 1/00 (2016.01)
CPC - C12M 23/16, 35/00, 35/04; C12Q 1/00 (2016.02)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 324/92; 435/287.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, Google Patents, PubMed.

Search terms used: microfluidic channel constriction pressure protein nucleic

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 2014/0287509 A1 (SHAREI et al) 25 September 2014 (25.09.2014) entire document	1-3, 33 ---------- 16
Y	US 6,410,329 B1 (HANSEN et al) 25 June 2002 (25.06.2002) entire document	16
A	US 2004/0176282 A1 (DALBY et al) 09 September 2004 (09.09.2004) entire document	1-3, 16, 33
A	US 8,844,570 B2 (GLICK) 30 September 2014 (30.09.2014) entire document	1-3, 16, 33
A	SHAREI et al. "A vector-free microfluidic platform for intracellular delivery," Proc Natl Acad Sci USA 22 January 2013 (22.01.2013), Vol. 110, Pgs.2082-7 entire document	1-3, 16, 33

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 March 2016	21 MAR 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/013113

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

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

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

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

3. ☒ Claims Nos.: 4-15, 17-32
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

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

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.