ENHANCED NUCLEAR DELIVERY OF DNA AND COMPOSITIONS FOR USE IN PRACTICING THE SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

Pursuant to 35 U.S.C. § 119 (e), this application claims priority to the filing date of the United States Provisional Application Serial No. 63/563,878 filed on March 11, 2024; United States Provisional Application Serial No. 63/633,236 filed on April 12, 2024; and United States Provisional Application Serial No. 63/702,825 filed on October 3, 2024, the disclosures of which are herein incorporated by reference.

INTRODUCTION

Field of the Invention

This invention pertains to the delivery of nucleic acids to cells and tissues.

Background

There are many instances in which nuclear delivery of a nucleic acid is desired, where such instances include research, diagnostic and therapeutic applications. An example of such a therapeutic application is gene therapy. In the field of gene therapy, viral vectors, such as vectors based on the virus AAV, are commonly employed to deliver genes to the nucleus. However, AAV's genome is limited in size, so any gene greater than 4.7kB will not be suitable for use with AAV vectors, which limits the utility of such vectors for many indications. In addition, viral vectors, such as AAV, induce an adaptive immune response, such that they can only be delivered once, which is not suitable for some indications, such as indications in the liver where the cells are slowly dividing and will lose the transduced genome, thereby requiring redosing. Furthermore, viral vectors such as AAV are toxic at the doses that are required for a therapeutic benefit in some indications. Accordingly, what is needed is a new delivery vehicle for delivering nucleic acids, such as DNA, to cells, particularly in patients in need of gene therapy but also *in vitro* during research.

That next generation delivery vehicle is a nanoparticle. Of particular interest are lipid nanoparticles, given how LNPs have been de-risked by being used in Onpattro and the Covid vaccines. The problem is that, in contrast to AAV's ability to carry its DNA cargo all the way into the nucleus, nanoparticles only deliver their cargo to the cytoplasm. In the case of dividing cells, delivery to the cytosol may be sufficient as the nuclear envelope dissolves in the course of replication and new DNA gets captured as the nucleus re-forms in the daughter cells. However, it is a problem for nondividing or slowly dividing cells. What is needed to make LNPs into a successful delivery vehicle for DNA is a mechanism

to enhance the transport of the DNA from the cytoplasm into the nucleus. The inventors have identified methods and compositions to satisfy the above need.

SUMMARY

Methods and compositions for enhancing nuclear DNA delivery are provided. Aspects of the methods include contacting a cell with a DNA to be delivered to the nucleus of the cell and an agent that modulates the activity of the protein Barrier-to-Autointegration Factor (BAF) in a cell. In addition, compositions, including reagents, devices and kits thereof, that find use in practicing embodiments of the subject methods are provided.

Aspects of the disclosure include compositions for the delivery of DNA to the nucleus of a cell are provided. In embodiments, the compositions include a cytosolic delivery vehicle, a DNA, and an agent that modulates the protein Barrier to Autointegration Factor (BAF).

In some embodiments, the agent is a protein having an amino acid sequence having a sequence identity of 90% or more to wild type BAF, or an mRNA encoding the protein having an amino acid sequence having a sequence identity of 90% or more to wild type BAF. In some instances, the protein includes a mutation such as a substitution at a residue relative to the wild type human BAF protein sequence selected from the group consisting of the threonine at residue 2, the threonine at residue 3, the serine at residue 4, the alanine at residue 12, the glutamic acid at residue 28, the glycine at residue 31, the lysine at residue 33, the glutamic acid at residue 36, the arginine at residue 37, the leucine at residue 58, the alanine at reside 71, or the glycine at residue 79of the wild type human BAF protein or a corresponding residue in a BAF ortholog. In some instances, the protein includes a mutation relative to wild type human BAF selected from the group consisting of T2D, T3D, S4D, A12T, E28D, G31S, K33R, E36D, R37K, L58R, A71K, and G79R. In some instances, the protein includes a heterologous UTR. In some embodiments, the agent is a BAF kinase or an mRNA encoding a BAF kinase. In some instances, the BAF kinase includes an amino acid sequence having a sequence identity of 90% or more to vaccinia virus Ankara strain B1 protein, where in some instances the BAF kinase includes an amino acid sequence having a sequence identity of 90% or more to human VRK1, VRK1-X1, human VRK2A, human VRK2B, or human VRK3. In some embodiments, the agent is a siRNA that is specific for BAF. In some embodiments, the agent is a small molecule, e.g., a small molecule is selected from the group consisting of obtusilactone B, Mahubanolide, kotomolide B, epilitsenolide D2, and rabeprazole. In some embodiments, the agent includes a LEM domain or an mRNA encoding a LEM domain. In some instances, the LEM domain includes an animo acid sequence having a sequence identity of 90% or more to a sequence selected from the group consisting of:

a. an emerin LEM domain
MDNYADLSDTELTTLLRRYNIPHGPVVGSTRRLYEKKIFEYETQ (SEQ ID NO:283);

b. a LAP2-beta LEM domain:
DLDVTELTNEDLLDQLVKYGVNPGPIVGTTRKLYEKKLLKLREQ (SEQ ID NO:284); and

c. a MAN1 LEM domain:
ASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKKLKKLREE(SEQ ID NO:285).

In some embodiments, the agent does not include a LEM domain or a nucleic acid encoding a LEM domain.

In some embodiments, the cytosolic delivery vehicle includes a non-viral delivery vehicle. In some instances, the non-viral delivery vehicle is a nanoparticle, such as a lipid nanoparticle (LNP). In some instances, the non-viral delivery vehicle is a vesicle, such as a vesicle selected from the group consisting of an exosome, an extracellular vesicle, and a micelle. In some embodiments, the composition includes a single cytosolic delivery vehicle having both the DNA and the agent. In some instances, the composition includes a first cytosolic delivery vehicle that includes the DNA and a second cytosolic delivery vehicle that includes the agent.

In some embodiments, the composition further includes an mRNA encoding a nuclear targeting factor (NTF), and the DNA includes a DNA nuclear targeting sequence (DTS). In some instances,

a. the NTF is a Gal4 NTF and the DTS is a Gal4 binding sequence; or

b. the NTF is a TetR NTF and the DTS is a TetR binding sequence.

In some embodiments, the composition includes a nuclease, an integrase or a transposase.IN some embodiments, the composition encodes a nuclease and the composition includes a guide RNA (gRNA).

In some embodiments, the DNA is 100 – 15,000 bp. In some embodiments, the DNA includes a coding sequence. In some embodiments, the DNA incldues an expression cassette comprising a promoter operably linked to the coding sequence. In some instances, the DNA includes sequences flanking the coding sequence or the expression cassette that are homologous to genomic sequences of the cell for which the DNA is configured to be used.

Aspects of the disclosure include methods, where the methods include contacting a cell with a DNA and an agent that modulates the activity of the protein Barrier to Autointegration Factor (BAF). In some embodiments, the agent is a protein that includes an amino acid sequence having a sequence identity of 90% or more to wild type BAF, or an mRNA encoding the protein that includes an amino acid sequence having a sequence identity of 90% or more to wild type BAF. In some instances, the protein includes a mutation such as a substitution at a residue relative to the wild type human BAF protein sequence selected from the group consisting of the threonine at residue 2, the threonine at residue 3, the

serine at residue 4, the alanine at residue 12, the glutamic acid at residue 28, the glycine at residue 31, the lysine at residue 33, the glutamic acid at residue 36, the arginine at residue 37, the leucine at residue 58, the alanine at reside 71, or the glycine at residue 79of the wild type human BAF protein or a corresponding residue in a BAF ortholog. In some instances, the protein includes a mutation relative to wild type human BAF selected from the group consisting of T2D, T3D, S4D, A12T, E28D, G31S, K33R, E36D, R37K, L58R, A71K, and G79R. In some instances, the protein includes a heterologous UTR. In some embodiments, the agent is a BAF kinase or an mRNA encoding a BAF kinase. In some instances, the BAF kinase includes an amino acid sequence having a sequence identity of 90% or more to vaccinia virus Ankara strain B1 protein, where in some instances the BAF kinase includes an amino acid sequence having a sequence identity of 90% or more to human VRK1, VRK1-X1, human VRK2A, human VRK2B, or human VRK3. In some embodiments, the agent is a siRNA that is specific for BAF. In some embodiments, the agent is a small molecule, e.g., a small molecule is selected from the group consisting of obtusilactone B, Mahubanolide, kotomolide B, epilitsenolide D2, and rabeprazole. In some embodiments, the agent includes a LEM domain or an mRNA encoding a LEM domain. In some instances, the LEM domain includes an animo acid sequence having a sequence identity of 90% or more to a sequence selected from the group consisting of:

a. an emerin LEM domain
MDNYADLSDTELTTLLRRYNIPHGPVVGSTRRLYEKKIFEYETQ (SEQ ID NO:283);

b. a LAP2-beta LEM domain:
DLDVTELTNEDLLDQLVKYGVNPGPIVGTTRKLYEKKLLKLREQ (SEQ ID NO:284); and

c. a MAN1 LEM domain:
ASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKKLKKLREE(SEQ ID NO:285).

In some embodiments, the agent does not include a LEM domain or a nucleic acid encoding a LEM domain.

In some embodiments of the methods, the cell is simultaneously or sequentially contacted with the DNA and the agent. In some instances, the DNA and the agent are present in a cytosolic delivery vehicle. In some instances, the DNA and the agent are present in the same cytosolic delivery vehicle. In some instances, the DNA and the agent are present in different cytosolic delivery vehicles. In some instances, the cytosolic delivery vehicle includes a nanoparticle, such as a lipid nanoparticle. In some instances, the cytosolic delivery vehicle includes a vesicle, such as a vesicle selected from the group consisting of an exosome, an extracellular vesicle, and a micelle.

In some embodiments, the methods include electroporating the cell. In some embodiments, the cell is *in vivo*. In some embodiments, the methods include orally, parenterally, subretinally, intravitreally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal

administration, intrapleurally, intravenously, intra-arterially, intraperitoneally, intracranially, subcutaneously, intramuscularly, intranasally, intrathecally and/or intraarticularly delivering the DNA and the agent. In some embodiments, the cell is *ex vivo*. In some embodiments, the cell is selected from the group consisting of hepatocytes, stellate cells, T lymphocytes, B lymphocytes, NK cells, hematopoietic stem cells, skeletal muscle cells, cardiomyocytes, neurons, astrocytes, oligodendrocytes, dendritic cells and skin cells.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A, 1A', 1B, 1B', 1C, and 1C' illustrate the formation of cytoplasmic puncta of DNA following transfection into the cell, and the subcellular localization of Barrier to Autointegration Factor (BAF) to these puncta. (A) and (A'): low and high magnification images of labeled DNA, following transfection in HepG2 cells. (B) and (B'): low and high magnification images of BAF protein in DNA-transfected HepG2 cells. (C) and (C'): low and high magnification images of BAF protein in untransfected HepG2 cells. Images were taken 16 hours after transfection.

Figures 2A-2F illustrate the subcellular localization of the nuclear envelope protein Emerin before and after transfection of cells with DNA or RNA. In these panels, Emerin protein is visualized with an Emerin-specific antibody (red) and the nuclei are stained with DAPI (blue). (A) Untreated HepG2 cells stained with an Emerin-specific antibody; (B) untreated HepG2 cells co-stained with DAPI; (C) mock transfected HepG2 cells (transfection reagents only) (D) HepG2 cells transfected with an mRNA encoding a nuclear translocating factor (NTF); (E) HepG2 cells transfected with DNA; (F) HepG2 cells transfected with DNA and the mRNA encoding an NTF.

Figures 3A-3C illustrate by confocal microscopy the envelope-type structure that forms from Emerin that is colocalized with transfected DNA. (A) Emerin protein, as visualized with an Emerin-specific antibody (red): (B) DNA, as visualized with an EdU (5-ethynyl-2'-deoxyuridine) label that was incorporated into the DNA prior to transfection (yellow); (C) merge of A and B. Nucleus is stained with DAPI (blue).

Figure 4 shows the amount of human Factor IX protein that has accumulated in mouse plasma seven days after delivery of an LNP co-formulated with a DNA comprising an hFIX expression cassette and an mRNA encoding the BAF modulator BAF-L58R (groups 1-4), or an mRNA encoding BAF-L58R protein and an mRNA encoding a Gal4 nuclear translocating factor (group 5), or no additional mRNA agents (group 6). Groups 1-4: DNA (0.3mg/kg) co-formulated with decreasing amounts of BAF-L58R mRNA (0.6 mg/kg, 0.3 mg/kg, 0.1 mg/kg, 0.03 mg/kg). Group 5: DNA (0.3mg/kg) co-formulated with BAF-L58R mRNA (0.3 mg/kg) and Gal4 NTF (0.6 mg/kg). Group 6: DNA alone (0.3 mg/kg). Group 7: vehicle (PBS).

Figure 5 shows the amount of human Factor IX protein that has accumulated in mouse plasma seven days after delivery of an LNP co-formulated with a DNA comprising an hFIX expression cassette and either an mRNA encoding the BAF-L58R protein based on mouse BAF ("mBAF-L58R") and an mRNA encoding a Gal4 nuclear translocating factor (group 1), an mRNA encoding BAF-L58R protein based on human BAF ("hBAF-L58R") and an mRNA encoding a Gal4 nuclear translocating factor (group 2), an mRNA encoding BAF-L58R protein based on mouse BAF but utilizing alternative UTRs to alter mRNA translation kinetics and an mRNA encoding a Gal4 nuclear translocating factor (group 3), or simply an mRNA encoding the BAF-L58R protein based on mouse BAF, i.e. without an mRNA encoding a Gal4 nuclear translocating factor (group 4).

Figure 6 shows the amount of human Factor IX protein that has accumulated in mouse plasma seven days after delivery of an LNP co-formulated with a DNA comprising an hFIX expression cassette, an mRNA encoding a GAL4 nuclear translocating factor, and either an mRNA encoding mBAF-L58R (group 1) or an mRNA encoding mBAF-L58R with a C-terminal NLS (group 2).

Figure 7 shows the amount of human Factor IX protein that has accumulated in mouse plasma seven days after delivery of an LNP formulated with a DNA comprising an hFIX expression cassette, an mRNA encoding a GAL4 nuclear translocating factor, and an mRNA encoding mouse wild type BAF (group 1).

Figure 8 shows the amount of human Factor IX protein that has accumulated in mouse plasma three and seven days after delivery of an LNP formulated with a DNA comprising an hFIX expression cassette, an mRNA encoding a Gal4 nuclear translocating factor and either: an mRNA encoding vaccinia virus Ankara strain B1 protein (group 1); an mRNA encoding a vaccinia virus Western Region strain B1 protein (group 2); or an mRNA encoding BAF-L58R protein based on mouse BAF. Group 4: vehicle (PBS).

Figures 9A-9D provide images of sections of mouse liver labeled with an antibody specific for human Factor IX at 15 or 27 days following i.v. injection of: (A) PBS; (B) an LNP formulated with a DNA comprising an hFIX expression cassette; (C) an LNP co-formulated with a DNA comprising an hFIX expression cassette and an mRNA encoding a GAL4 nuclear translocating factor; or (D) an LNP co-formulated with a DNA comprising an hFIX expression cassette, an mRNA encoding BAF-L58R protein based on mouse BAF, and an mRNA encoding a GAL4 nuclear translocating factor. Tissue sections in Figures 9A-C are from 27 days post-dose; tissue section in Figure 9D is from 15 days post-dose.

Figures 10A-10C illustrate the amount of human Factor IX protein that has accumulated in plasma between 3 and 14 days after delivery of an LNP formulated with a DNA comprising an hFIX expression cassette and an mRNA encoding a mutant human BAF protein fused to an NLS into (A) mice, (B) rats, or (C) NHPs.

Figure 11 demonstrates the increase in GFP expression following delivery of LNPs formulated with a DNA comprising a GFP expression cassette and an mRNA encoding BAF variants of the present disclosure into primary human hepatocytes (PHHs).

INCORPORATION BY REFERENCE

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

DEFINITIONS

The terms "polypeptide," "polypeptide sequence," "peptide," "peptide sequence," "protein," "protein sequence" and "amino acid sequence" are used interchangeably herein to designate a linear series of amino acid residues connected one to the other by peptide bonds, which series may include proteins, polypeptides, oligopeptides, peptides, and fragments thereof. The protein may be made up of naturally occurring amino acids and/or synthetic (e.g., modified or non-naturally occurring) amino acids. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. The terms "polypeptide", "peptide", and "protein" includes fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence, fusions with heterologous and homologous leader sequences, with or without N-terminal methionine residues; immunologically tagged proteins; fusion proteins with detectable fusion partners, e.g., fusion proteins including as a fusion partner a fluorescent protein, beta-galactosidase, luciferase, and the like. Furthermore, it should be noted that a dash at the beginning or end of an amino acid sequence indicates either a peptide bond to a further sequence of one or more amino acid residues or a covalent bond to a carboxyl or hydroxyl end group. However, the absence of a dash should not be taken to mean that such peptide bond or covalent bond to a carboxyl or hydroxyl end group is not present, as it is conventional in representation of amino acid sequences to omit such.

The term "polynucleotide," "polynucleotide sequence," "oligonucleotide," "oligonucleotide sequence," "oligomer," "oligo," "nucleic acid sequence" or "nucleotide sequence" used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer having purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.

The terms "derivative" and "variant" refer without limitation to any compound such as nucleic acid or protein that has a structure or sequence derived from the compounds disclosed herein and whose

structure or sequence is sufficiently similar to those disclosed herein such that it has the same or similar activities and utilities or, based upon such similarity, would be expected by one skilled in the art to exhibit the same or similar activities and utilities as the referenced compounds, thereby also interchangeably referred to "functionally equivalent" or as "functional equivalents." Modifications to obtain "derivatives" or "variants" may include, for example, addition, deletion and/or substitution of one or more of the nucleic acids or amino acid residues.

The functional equivalent or fragment of the functional equivalent, in the context of a protein, may have one or more conservative amino acid substitutions. The term "conservative amino acid substitution" refers to substitution of an amino acid for another amino acid that has similar properties as the original amino acid. The groups of conservative amino acids are as follows:

Group	Name of the amino acids
Aliphatic	Gly, Ala, Val, Leu, Ile
Hydroxyl or Sulfhydryl/Selenium-containing	Ser, Cys, Thr, Met
Cyclic	Pro
Aromatic	Phe, Tyr, Trp
Basic	His, Lys, Arg
Acidic and their Amide	Asp, Glu, Asn, Gln

Conservative substitutions may be introduced in any position of a preferred predetermined peptide or fragment thereof. It may however also be desirable to introduce non-conservative substitutions, particularly, but not limited to, a non-conservative substitution in any one or more positions. A non-conservative substitution leading to the formation of a functionally equivalent fragment of the peptide would for example differ substantially in polarity, in electric charge, and/or in steric bulk while maintaining the functionality of the derivative or variant fragment.

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

The terms "identical" or percent "identity" in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity over a specified region, e.g., the entire polypeptide

sequences or individual domains of the polypeptides), when compared and aligned for maximum correspondence over a comparison window or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be "substantially identical." This definition also refers to the complement of a test sequence.

The term "complementary" or "substantially complementary," interchangeably used herein, means that a nucleic acid (e.g., DNA or RNA) has a sequence of nucleotides that enables it to non-covalently bind, i.e. form Watson-Crick base pairs and/or G/U base pairs to another nucleic acid in a sequence-specific, antiparallel, manner (i.e., a nucleic acid specifically binds to a complementary nucleic acid). As is known in the art, standard Watson-Crick base-pairing includes: adenine (A) pairing with thymidine (T), adenine (A) pairing with uracil (U), and guanine (G) pairing with cytosine (C).

A DNA sequence that "encodes" a particular RNA is a DNA nucleic acid sequence that is transcribed into RNA when placed under the control of appropriate regulatory sequences. A DNA polynucleotide may encode an RNA (mRNA) that is translated into protein, or a DNA polynucleotide may encode an RNA that is not translated into protein (e.g., tRNA, rRNA, or a guide RNA; also called "non-coding" RNA or "ncRNA"). A protein coding sequence or a sequence that encodes a particular protein or polypeptide, is a nucleic acid sequence that is transcribed into mRNA (in the case of DNA) and is translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences.

As used herein, "codon" refers to a sequence of three nucleotides that together form a unit of genetic code in a DNA or RNA molecule. As used herein the term "codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. Nucleotides are typically referred to as the following: A (adenine), G (guanine), C (cytosine), T (thymine), U (uracil), N (= A or C or G or T/U), B (= C or G or T/U), D (= A or G or T/U), H (= A or C or T/U), K (= G or T/U), M (= A or C), R (= A or G), S (= C or G), V(= A or C or G), W (= A or T/U), (= C or T/U). When describing the RNA herein, it will be appreciated by the ordinarily skilled artisan that in the course of transcription, a DNA sequence will be transcribed into an mRNA sequence having the same nucleotide sequence but for thymine, which will be encoded as uracil in the RNA. Accordingly, the ordinarily skilled artisan will readily be able to convert the DNA sequences disclosed herein or known in the art to mRNA sequences.

The term "codon-optimized" or "codon optimization" refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA. Such optimization includes replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the

genes of that organism. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at www.kazusa.or.jp/codon/ (visited Mar. 20, 2008). By utilizing the knowledge on codon usage or codon preference in each organism, one of ordinary skill in the art can apply the frequencies to any given polypeptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypeptide, but which uses codons optimal for a given species. Codon-optimized coding regions can be designed by various methods known to those skilled in the art.

The term "recombinant" or "engineered" when used with reference, for example, to a cell, a nucleic acid, a protein, or a vector, indicates that the cell, nucleic acid, protein or vector has been modified by or is the result of laboratory methods. Thus, for example, recombinant or engineered proteins include proteins produced by laboratory methods. Recombinant or engineered proteins can include amino acid residues not found within the native (non-recombinant or wild-type) form of the protein or can include amino acid residues that have been modified, e.g., labeled. The term can include any modifications to the peptide, protein, or nucleic acid sequence. Such modifications may include the following: any chemical modifications of the peptide, protein or nucleic acid sequence, including of one or more amino acids, deoxyribonucleotides, or ribonucleotides; addition, deletion, and/or substitution of one or more of amino acids in the peptide or protein; and addition, deletion, and/or substitution of one or more of nucleic acids in the nucleic acid sequence.

The term "genomic DNA" or "genomic sequence" refers to the DNA of a genome of an organism including, but not limited to, the DNA of the genome of a bacterium, fungus, archea, plant or animal.

As used herein, "transgene," "exogenous gene" or "exogenous sequence," in the context of nucleic acid, refers to a nucleic acid sequence or gene that was not present in the genome of a cell but artificially introduced into the genome, e.g., via genome-edition.

As used herein, "endogenous gene" or "endogenous sequence," in the context of nucleic acid, refers to a nucleic acid sequence or gene that is naturally present in the genome of a cell, without being introduced via any artificial means.

The term "expression cassette" refers to a DNA coding sequence operably linked to a promoter. "Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression. The terms "recombinant expression vector," or "DNA construct" are used interchangeably herein to refer to a DNA molecule having a vector and at least one insert. Recombinant expression vectors are usually generated for the purpose of expressing and/or propagating the insert(s), or for the construction of other recombinant nucleotide sequences. The nucleic acid(s) may or may not be operably linked to a promoter sequence and may or may not be operably linked to DNA regulatory sequences.

The term "operably linked" means that the nucleotide sequence of interest is linked to regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence. The term "regulatory sequence" is intended to include, for example, promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are well known in the art and are described, for example, in Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cells, and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the target cell, the level of expression desired, and the like.

A cell has been "genetically modified" or "transformed" or "transfected" by exogenous DNA, e.g., a recombinant expression vector, when such DNA has been introduced inside the cell. The presence of the exogenous DNA results in permanent or transient genetic change. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. The genetically modified (or transformed or transfected) cells that have therapeutic activity, e.g., treating hemophilia A, can be used and referred to as therapeutic cells.

The term "concentration" used in the context of a molecule such as peptide fragment refers to an amount of molecule, e.g., the number of moles of the molecule, present in a given volume of solution.

The terms "individual," "subject" and "host" are used interchangeably herein and refer to any subject for whom diagnosis, treatment or therapy is desired. In some aspects, the subject is a mammal. In some aspects, the subject is a human being. In some aspects, the subject is a patient. In some aspects, the subject is a human patient. In some aspects, the subject can have or is suspected of having a disorder or health condition associated with a gene-of-interest (GOI). In some aspects, the subject is a human who is diagnosed with a risk of disorder or health condition associated with a GOI at the time of diagnosis or later. In some cases, the diagnosis with a risk of disorder or health condition associated with a GOI can be determined based on the presence of one or more mutations in the endogenous GOI or genomic sequence near the GOI in the genome that may affect the expression of GOI.

The term "treatment" referring to a disease or condition means that at least an attenuation (amelioration) of the symptoms associated with the condition afflicting an individual is achieved, where attenuation is used in a broad sense to refer to at least a reduction in the magnitude of a parameter, e.g., a symptom, associated with the condition (e.g., hemophilia) being treated. As such, treatment also includes situations where the pathological condition, or at least symptoms associated therewith, are completely inhibited, e.g., prevented from happening, or eliminated entirely such that the host no longer suffers from the condition, or at least the symptoms that characterize the condition. Thus, treatment includes: (i)

prevention, that is, reducing the risk of development of clinical symptoms, including causing the clinical symptoms not to develop, e.g., preventing disease progression; (ii) inhibition, that is, arresting the development or further development of clinical symptoms, e.g., reducing the rate of progression of disease; (iii) mitigating, that is, reducing the symptoms of, or completely inhibiting, an active disease.

The terms "effective amount," "pharmaceutically effective amount," or "therapeutically effective amount" as used herein mean a sufficient amount of the composition to provide the desired utility when administered to a subject having a particular condition. The term "therapeutically effective amount" therefore refers to an amount of therapeutic cells or a composition having therapeutic cells that is sufficient to promote a particular effect when administered to a subject in need of treatment. An effective amount would also include an amount sufficient to prevent or delay the development of a symptom of the disease, alter the course of a symptom of the disease (for example but not limited to, slow the progression of a symptom of the disease), or reverse a symptom of the disease. It is understood that for any given case, an appropriate "effective amount" can be determined by one of ordinary skill in the art using routine experimentation.

The term "pharmaceutically acceptable excipient" as used herein refers to any suitable substance that provides a pharmaceutically acceptable carrier, additive or diluent for administration of a compound(s) of interest to a subject. "Pharmaceutically acceptable excipient" can encompass substances referred to as pharmaceutically acceptable diluents, pharmaceutically acceptable additives, and pharmaceutically acceptable carriers.

DETAILED DESCRIPTION

Methods and compositions for enhancing nuclear DNA delivery are provided. Aspects of the methods include contacting a cell with a DNA to be delivered to the nucleus of the cell and an agent that modulates the activity of the protein Barrier-to-Autointegration Factor (BAF) in a cell. In addition, compositions, including reagents, devices and kits thereof, that find use in practicing embodiments of the subject methods are provided.

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

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit

of that range and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Certain ranges are presented herein with numerical values being preceded by the term "about." The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, representative illustrative methods and materials are now described.

All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

It is noted that, as used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the context clearly dictates otherwise. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation.

As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present invention. Any recited method can be carried out in the order of events recited or in any other order which is logically possible.

While the apparatus and method has or will be described for the sake of grammatical fluidity with functional explanations, it is to be expressly understood that the claims, unless expressly formulated under 35 U.S.C. §112, are not to be construed as necessarily limited in any way by the construction of "means" or "steps" limitations, but are to be accorded the full scope of the meaning and equivalents of the definition provided by the claims under the judicial doctrine of equivalents, and in the case where the claims are expressly formulated under 35 U.S.C. §112 are to be accorded full statutory equivalents under 35 U.S.C. §112.

METHODS

As summarized above, methods for enhancing nuclear DNA delivery are provided. Aspects of the methods include contacting a cell with a DNA to be delivered and an agent that modulates the activity of the protein Barrier-to-Autointegration Factor (BAF) in a cell. By "enhancing" nuclear DNA delivery is meant that embodiments of the invention provide for increased DNA delivery to the nucleus, e.g., compared to a suitable control. The magnitude of such increase may vary, where in some instances the magnitude of enhancement is 2-fold or greater, such as 5-fold or greater, including 10-fold or greater, 25-fold or greater, 50-fold or greater, 100-fold greater, or 1000-fold greater. Embodiments of the methods provide for transportation of a DNA, which may be a cargo nucleic acid, from the cytosol of a cell into the nucleus of cell. As such, embodiments of methods of invention may be characterized as methods of facilitating or promoting transport of a DNA into the nucleus of a cell. Where the DNA includes a coding sequence, practicing embodiments of the methods of the invention may result in enhancement of expression of the coding sequence of the DNA. By "enhancing" expression of the coding sequence is meant expression of the coding sequence of the DNA is increased, e.g., compared to a suitable control. The magnitude of such increase may vary, where in some instances the magnitude of enhancement is 2-fold or greater, such as 5-fold or greater, including 10-fold or greater, 25-fold or greater, 50-fold or greater, or 100-fold or greater. Embodiments of methods of the invention may provide for an improvement, or an enhancement, in delivery and/or expression of DNA to the nucleus over DNA delivery methods that depend on the delivery of DNA that is not provided in conjunction with an agent that modulates BAF. In some instances, the improvement is a 2-fold increase or more in the amount of DNA in the nucleus relative to the amount of DNA that would be found in the nucleus if that DNA was not provided in conjunction with an agent that modulates BAF, for example, a 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 10-fold, 15-fold, 20-fold, 30-fold, 40-fold, or 50-fold increase in the amount of DNA in the nucleus. This may be detected as an increase in the expression of a coding sequence of the DNA, for example, as a 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 10-fold, 15-fold, 20-fold, 30-fold, 40-fold, 50-fold, 100-fold, 200-fold, 500-fold or 1000-fold increase or more in RNA or protein

encoded by the DNA over that expressed in a cell that has not been contacted with DNA or a cell that has been contacted with a DNA not in conjunction with the agent that modulates BAF .

Various aspects of the methods are now reviewed in greater detail.

Modulation of the protein Barrier to Autointegration Factor (BAF)

As reviewed above, aspects of embodiments of the invention include contacting the cell with a DNA to be delivered and an agent that modulates the activity of the protein Barrier to Autointegration Factor (BAF). Barrier-to-Autointegration Factor (BAF, also referred to as BAF nuclear assembly factor, or BANF, encoded by the BANF1 gene, the human ortholog having the protein sequence MTTSQKHRDFVAEPMGEKPVGSLAGIGEVLGKKLEERGFDKAYVVLGQFLVLKKDEDLFREWLKDTCGANAKQSRDC FGCLREWCDAFL (SEQ ID NO:257) and the mouse ortholog having the protein sequence MTTSQKHRDFVAEPMGEKPVGSLAGIGDVLSKRLEERGFDKAYVVLGQFLVLKKDEDLFREWLKDTCGANAKQSRDC FGCLREWCDAFL(SEQ ID NO:258) is known in the art to be a conserved nuclear envelope (NE) component that binds chromatin in a 3d-structure dependent, sequence-independent fashion and helps it to anchor to the nuclear envelope. During mitosis, phosphorylation by vaccinia-related kinase 1 (VRK1) or 2 (VRK2) causes BAF to release chromosomal DNA, allowing the nuclear envelope to dissemble. BAF dephosphorylation by protein phosphatase 2A (PP2A) at the end of mitosis allows BAF to recapture the DNA and, working in concert with LEM-domain containing proteins, mediate its packaging within a new nuclear envelope. BAF has also been observed in the art to recognize and bind exogenous DNA entering the cell and nucleate the formation of a nuclear envelope-like structure called an exclusome around the DNA. For the sake of clarity, "BAF" does *not* refer to the "BRG1- or BRM-associated factor" chromatin remodeling complex, also known in the art as the mammalian SWI/SNF complex.

By an agent that modulates BAF (referred to herein as a "BAF modulator"), it is meant an agent that suppresses, reduces, attenuates, inhibits, promotes, agonizes, supplements or otherwise modifies one or more activities of BAF in a manner that provides for a desired result, e.g., an enhancement of nuclear delivery and/or expression of a DNA, such as described above. In some embodiments, BAF modulator agents are agents that reduce the affinity of BAF for DNA. In some embodiments, BAF modulator agents are agents that reduce the ability of BAF to dimerize. In some embodiments, BAF modulator agents are agents that reduce the affinity of BAF for LEM domain-containing proteins such as LAP2, Emerin, and MAN1. In some embodiments, BAF modulator agents are agents that compete with endogenous BAF to inhibit functions of endogenous BAF. In some embodiments, BAF modulator agents are agents that interact with a BAF-interacting protein to modulate the interaction between BAF and that BAF-interacting protein, thereby altering the activity of BAF on DNA, e.g. as it relates to BAF's ability to condense DNA, localize DNA to a region of the cell, or otherwise modify the ability of DNA to be re-

localized to the nucleus of a cell. In some embodiments, BAF modulator agents are agents that interact directly with a DNA, so as to modify the ability of DNA to be re-localized to the nucleus of a cell, e.g. by altering the condensation of DNA, the association of DNA with other proteins, the subcellular localization of the DNA to a region of the cell, and the like.

In some instances, the BAF modulator is a protein or a nucleic acid, e.g. a DNA or an mRNA, encoding a protein. In some such instances, the BAF modulator is a wild type BAF protein or a nucleic acid, e.g., a DNA or an mRNA, encoding a wild type BAF protein, e.g. a human BAF protein having the protein sequence

MTTSQKHRDFVAEPMGEKPVGSLAGIGEVLGKKLEERGFDKAYVVLGQFLVLKKDEDLFREWLKDTCGANAKQSRDC
FGCLREWCDAFL(SEQ ID NO:260) or the mouse ortholog having the protein sequence
MTTSQKHRDFVAEPMGEKPVGSLAGIGDVLSKRLEERGFDKAYVVLGQFLVLKKDEDLFREWLKDTCGANAKQSRDC
FGCLREWCDAFL (SEQ ID NO:260). Thus, for example, the BAF modulator may be a DNA encoding a polypeptide comprising the human BAF protein, e.g. a deoxyribopolynucleotide having a sequence identity of 65% or more to the sequence

atgacaacctcccaaaagcaccgagacttcgtggcagagcccatgggggagaagccagtggggagcctggctgggat
tggtgaagtcctgggcaagaagctggaggaaaggggttttgacaaggcctatgttgtccttggccagtttctggtgc
taaagaaagatgaagacctcttccgggaatggctgaaagacacttgtggcgccaacgccaagcagtcccgggactgc
ttcggatgccttcgagagtggtgcgacgccttcttgtga(SEQ ID NO:261) or to the sequence:
atgaccaccagccagaagcacagggacttcgtggccgagcccatgggcgagaagcccgtgggcagcctggccggcat
cggcgaggtgctgggcaagaagctggaggagaggggcttcgacaaggcctacgtggtgctgggccagttcctggtgc
tgaagaaggacgaggacttgttcagggagtggctgaaggacacctgcggcgccaacgccaagcagagcagggactgc
ttcggctgcctgagggagtggtgcgacgccttcctgtaa(SEQ ID NO:262); or an mRNA encoding a polypeptide comprising human BAF protein sequence, e.g. a ribopolynucleotide having a sequence identity of 65% or more to the sequence:

atgaccaccagccagaagcacagggacttcgtggccgagcccatgggcgagaagcccgtgggcagcctggccggcat
cggcgaggtgctgggcaagaagctggaggagaggggcttcgacaaggcctacgtggtgctgggccagttcctggtgc
tgaagaaggacgaggacttgttcagggagtggctgaaggacacctgcggcgccaacgccaagcagagcagggactgc
ttcggctgcctgagggagtggtgcgacgccttcctg(SEQ ID NO:263).

As another example, the BAF modulator may be a DNA encoding a polypeptide comprising the mouse BAF protein, e.g. a deoxyribopolynucleotide having a sequence identity of 65% or more to the sequence;

atgacaacctcccaaaagcaccgagacttcgtggcagagcccatgggggaaaagccagtggggagcctggccgggat
tggtgacgtcctgagcaagaggctggaggaaaggggctttgacaaggcttatgtggtccttggccagtttctggtgc
taaagaaagatgaagacctcttccgagaatggctgaaggatacatgtggtgccaatgccaagcagtcccgggactgc
tttgggtgccttcgagaatggtgtgatgccttcttgtag(SEQ ID NO:264) or to the sequence:

atgaccaccagccagaagcacagggacttcgtggccgagcccatgggcgagaagcccgtgggcagcctggccggcat
cggcgacgtgctgagcaagaggctggaggagaggggcttcgacaaggcctacgtggtgctgggccagttcctggtgc

tgaagaaggacgaggacttgttcagggagtggctgaaggacacctgcggcgccaacgccaagcagagcagggactgc ttcggctgcctgagggagtggtgcgacgccttcctg(SEQ ID NO:265); or an mRNA encoding a polypeptide comprising mouse BAF protein sequence, e.g. a ribopolynucleotide having a sequence identity of 65% or more to the sequence:

aggggataataggaaaaatctatagcagaagtcagcggtagacgcacggcatagcatccaacgccaccatgaccacc agccagaagcacagggacttcgtggccgagcccatgggcgagaagcccgtgggcagcctggccggcatcggcgacgt gctgagcaagaggctggaggagaggggcttcgacaaggcctacgtggtgctgggccagttcctggtgctgaagaagg acgaggacttgttcagggagtggctgaaggacacctgcggcgccaacgccaagcagagcagggactgcttcggctgc ctgagggagtggtgcgacgccttcctg (SEQ ID NO:266).

In other instances, the BAF modulator is a BAF mutant protein or a nucleic acid, e.g., a DNA or an mRNA, encoding a BAF mutant protein. BAF mutant proteins finding use in embodiments of the invention may vary, where examples of BAF mutant proteins include proteins comprising an amino acid sequence having a sequence identity of 75% or more, such as 80% or more, including 90% or more, e.g., 95% or more, to a wild type BAF. In some instances, the BAF mutant comprises a mutation, e.g. a substitution or deletion, relative to wild type BAF that mimics hyperphosphorylation of BAF, that reduces the ability of BAF to be dephosphorylated, that reduces the ability of BAF to bind DNA and/or that reduces the ability of BAF to interact with LEM-domain containing proteins. In some embodiments, the mutation includes a substitution at a residue relative to the wild type human BAF protein sequence selected from the group consisting of the threonine at residue 2, the threonine at residue 3, the serine at residue 4, the alanine at residue 12, the glutamic acid at residue 28, the glycine at residue 31, the lysine at residue 33, the glutamic acid at residue 36, the arginine at residue 37, the leucine at residue 58, the alanine at reside 71, or the glycine at residue 79of the wild type human BAF protein or a corresponding residue in a BAF ortholog. For example, the mutation may include a heterologous amino acid (relative to wild type human BAF) selected from the group consisting of an aspartic acid at residue 2, an aspartic acid at residue 3, an aspartic acid at residue 4, a threonine at residue 12, an aspartic acid at residue 28, a serine at reside 31, an arginine at residue 33, an aspartic acid at residue 36, a lysine at residue 37, an arginine at residue 58, a lysine at reside 71, and an arginine at residue 79; or in a corresponding residue in a BAF ortholog, for example, T2D, T3D, S4D, A12T, E28D, G31S, K33R, E36D, R37K, L58R, A71K, G79R, and the like. In other embodiments, the BAF modulator is a barrier to autointegration factor-like (BAFL) protein, also referred to as BAF2 or BANF2, or a nucleic acid, e.g., a DNA or an mRNA, encoding a BAFL protein, including proteins comprising an amino acid sequence having a sequence identity of 75% or more, such as 80% or more, including 90% or more, e.g., 95% or more, to wild type BAFL (MDNMSPRLRAFLSEPIGEKDVCWVDGISHELAINLVTKGINKAYILLGQFLLMHKNEAEFQRWLICCFGATECEAQ QTSHCLKEWCACFL(SEQ ID NO:267)).

In yet other embodiments, the BAF modulator is a kinase that targets BAF, i.e. a kinase that phosphorylates BAF or nucleic acid, e.g., a DNA or an mRNA, encoding a kinase that phosphorylates BAF, referred to herein as a "BAF kinase" or a "BAF-targeting kinase". Examples of BAF kinases finding use in embodiments of the invention include any polypeptide that phosphorylates BAF, including, but not limited to, B1 proteins of *Poxviridae*, and mutants thereof having sequence identity of 75% or more, such as 80% or more, including 90% or more, e.g., 95% or more, 96% or more, 97% or more to a B1 ortholog that retain kinase activity on BAF, and polypeptides comprising an amino acid sequence having a sequence identity of 75% or more, such as 80% or more, including 90% or more, e.g., 95% or more, 96% or more, or 97% or more to a B1 ortholog. B1 proteins of *Poxviridae* are serine/threonine kinases that phosphorylate BAF at the N terminus, thereby inhibiting BAF binding to DNA. B1 orthologs are highly conserved within the members of the *Poxviridae* family that infect mammals, and include but are not limited to the B1 protein from vaccinia virus Ankara strain (GenBank Accession No. O57252):

MNFQGLVLTDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHNIKH
VGLITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNNRLPKRSVMLIGIEILNTIQFMHEQGYSHGDIKASNIV
LDQIDKNKLYLVDYGLVSKFMSNGEHVPFIRNPNKMDNGTLEFTPIDSHKGYVVSRRGDLETLGYCMIRWLGGILPW
TKISETKNCALVSATKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEEPNYDKFRHILMQGVYY(SEQ ID NO:268), as encoded by, for example, an mRNA having the sequence:

gaagagcagcctaatacgactcactataaggggataatagtcaacacaacatatacaaaacaaacgaatctcaagca
atcaagcattctacttctattgcagcaatttaaatcatttcttttaaagcaaaagcaatttttctgaaaatttttcacc
atttacgaacgataggccaccatgaacttccagggcctggtgctgaccgacaactgcaagaaccagtgggtggtggg
ccccctgatcggcaagggcggcttcggcagcatctacaccaccaacgacaacaactacgtggtgaagatcgagccca
aggccaacggcagcctgttcaccgagcaggccttctacaccagggtgctgaagcccagcgtgatcgaggagtggaag
aagagccacaacatcaagcacgtgggcctgatcacctgcaaggccttcggcctgtacaagagcatcaacgtggagta
caggttcctggtgatcaacaggctgggcgccgacctggacgccgtgatcagggccaacaacaacaggctgcccaaga
ggagcgtgatgctgatcggcatcgagatcctgaacaccatccagttcatgcacgagcagggctacagccacggcgac
atcaaggccagcaacatcgtgctggaccagatcgacaagaacaagctgtacctggtggactacggcctggtgagcaa
gttcatgagcaacggcgagcacgtgccccttcatcaggaaccccaacaagatggacaacggcaccctggagttcaccc
ccatcgacagccacaagggctacgtggtgagcaggaggggcgacctggagaccctgggctactgcatgatcaggtgg
ctgggcggcatcctgccctggaccaagatcagcgagaccaagaactgcgccctggtgagcgccaccaagcagaagta
cgtgaacaacaccgccaccctgctgatgaccagcctgcagtacgcccccagggagctgctgcagtacatcaccatgg
tgaacagcctgacctacttcgaggagcccaactacgacaagttcaggcacatcctgatgcagggcgtgtactactaa
ctagtgactgactaggatctggttaccactaaaccagcctcaagaacacccgaatggagtctctaagctacataata
ccaacttacacttacaaaatgttgtcccccaaaatgtagccattcgtatctgctcctaataaaaagaaagtttcttc
acatttcttctctcccttgcacctaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa(SEQ ID NO:269);

The B1 protein from vaccinia virus Western Reserve strain (GenBank Accession No. YP_233065):

MNFQGLVLTDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHNIKH
VGLITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNNRLPKRSVMLIGIEILNTIQFMHEQGYSHGDIKASNIV
LDQIDKNKLYLVDYGLVSKFMSNGEHVPFIRNPNKMDNGTLEFTPIDSHKGYVVSRRGDLETLGYCMIRWLGGILPW
TKISETKNCALVSATKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEEPNYDEFRHILMQGVYY(SEQ ID
NO:270), as encoded by, for example, an mRNA having the sequence:
cgcatcatagactagatctactatctagcggtagagcagcctaatacgactcactataaggggataataggaaaaat
ctatagcagaagtcagcggtagacgcacggcatagcatccaacgccaccatgaacttccagggcctggtgctgaccg
acaactgcaagaaccagtgggtggtgggccccctgatcggcaagggcggcttcggcagcatctacaccaccaacgac
aacaactacgtggtgaagatcgagcccaaggccaacggcagcctgttcaccgagcaggccttctacaccagggtgct
gaagcccagcgtgatcgaggagtggaagaagagccacaacatcaagcacgtgggcctgatcacctgcaaggccttcg
gcctgtacaagagcatcaacgtggagtacaggttcctggtgatcaacaggctgggcgccgacctggacgccgtgatc
agggccaacaacaacaggctgcccaagaggagcgtgatgctgatcggcatcgagatcctgaacaccatccagttcat
gcacgagcagggctacagccacggcgacatcaaggccagcaacatcgtgctggaccagatcgacaagaacaagctgt
acctggtggactacggcctggtgagcaagttcatgagcaacggcgagcacgtgcccttcatcaggaaccccaacaag
atggacaacggcaccctggagttcacccccatcgacagccacaagggctacgtggtgagcaggaggggcgacctgga
gaccctgggctactgcatgatcaggtggctgggcggcatcctgccctggaccaagatcagcgagaccaagaactgcg
ccctggtgagcgccaccaagcagaagtacgtgaacaacaccgccaccctgctgatgaccagcctgcagtacgccccc
agggagctgctgcagtacatcaccatggtgaacagcctgacctacttcgaggagcccaactacgacgagttcaggca
catcctgatgcagggcgtgtactactaaagaccagacagctggagcctcggtggccatgcttcttgccccttgggcc
tccccccagccctcctcccctcctgcacccgtacccccgtggtctttgaataaagtctgagtgggcggcaagctt
ttcttctctcccttgcacctaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa(SEQ ID NO:271) ;

The horsepox virus OPG187 protein (GenBank Accession No. YP_010509385):
MNFQGLVLTDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHNIKH
VGLITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNNRLPKRSVMLIGIE
ILNTIQFMHEQGYSHGDIKASNIVLDQIDKNKLYLVDYGLVSKFMSNGEHVPFIRNPNKMDNGTLEFTPIDSHKGYV
VSRRGDLETLGYCMIRWLGGILPWTKISETKNCALVSATKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEE
PNYDKFRHILMQGVYY(SEQ ID NO:272);

The cowpox CPXV196 protein (GenBank Accession No. ADZ29301):
MNFQGLVLTDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHNIKH
VGLITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNNRLPKRSVMLIGIE
ILNTIQFMHEQGYSHGDIKASNIVLDQIDKNKLYLVDYGLVSKFMSNGEHIPFIRNPNKMDNGTLEFTPIDSHKGYV
VSRRGDLETLGYCMIRWLGGILPWTKISETKNCALVSATKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEE
PNYDEFRHILMQGVYY(SEQ ID NO:273);

the variola virus protein (GenBank Accession No. ABF22933):
MNFQGLVLTDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHHISH
VGVITCKAFGLYKSINTEYRFLVINRLGADLDAVIRANNNRLPKRSVMLVGIEILNTIQFMHEQGYSHGDIKASNIV

LDQMDKNKLYLVDYGLVSKFMSNGEHVPFIRNPNKMDNGTLEFTPIDSHKGYVVSRRGDLETLGYCMIRWLGGILPW
TKIAETKNCALVSATKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEEPNYDKFRHILMQGAYY(SEQ ID
NO:274);

and the Monkeypox OPG187 protein (GenBank Accession No. UZV11619.1):

MKFQGLVLIDNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRVLKPSVIEEWKKSHNIKH
VGITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNNRLPERSVMLIGIEILNTIQFMHEQGYSHGDIKASNIVL
DQIDKNKLYLVDYGLVSKFMSNGEHVPFIRNPNKMDNGTLEFTPIDSHKGYVVSRRGDLETLGYCMIRWLGGILPWT
KISETKNSALVSAAKQKYVNNTATLLMTSLQYAPRELLQYITMVNSLTYFEEPNYDKFRRVLMNGVMKNFC(SEQ
ID NO:275).

Examples of BAF kinases finding use in embodiments of the invention also include, but are not limited to, mammalian VRK proteins, as well as mutants thereof that retain kinase activity on BAF. VRK proteins are a group of eukaryotic serine/threonine kinases having homology (~40% identity) to the *Poxviridae* B1-like proteins, including, e.g. the vaccinia-related kinase 1 (VRK1) protein (GenBank Accession no. NM_003384.3) having a sequence:

MPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAKKEWKVGLPIGQGGFGCIYLADMNSSESVGSDAPCVVKVEPSDN
GPLFTELKFYQRAAKPEQIQKWIRTRKLKYLGVPKYWGSGLHDKNGKSYRFMIMDRFGSDLQKIYEANAKRFSRKTV
LQLSLRILDILEYIHEHEYVHGDIKASNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKEYKEDPKRCHDGTIEFTSID
AHNGVAPSRRGDLEILGYCMIQWLTGHLPWEDNLKDPKYVRDSKIRYRENIASLMDKCFPEKNKPGEIAKYMETVKL
LDYTEKPLYENLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRKKEIEESKEPGVEDTEWSNTQTEEAIQ
TRSRTRKRVQK (SEQ ID NO:276);

the vaccinia-related kinase 1-X1 (VRK1-X1) protein (GenBank Accession no. XM_047431752), having a sequence:

MKMPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAKKEWKVGLPIGQGGFGCIYLADMNSSESVGSDAPCVVKVEPS
DNGPLFTELKFYQRAAKPEQIQKWIRTRKLKYLGVPKYWGSGLHDKNGKSYRFMIMDRFGSDLQKIYEANAKRFSRK
TVLQLSLRILDILEYIHEHEYVHGDIKASNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKEYKEDPKRCHDGTIEFTS
IDAHNGVAPSRRGDLEILGYCMIQWLTGHLPWEDNLKDPKYVRDSKIRYRENIASLMDKCFPEKNKPGEIAKYMETV
KLLDYTEKPLYENLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRKKEIEESKEPGVEDTEWSNTQTEEA
IQTRSEESQGAIHRSMSQPEASSSSYDSSRTRKRVQK(SEQ ID NO:277),

the vaccinia-related kinase 2A (VRK2A) protein (GenBank Accession no. NM_006296.7) having a sequence:

MPPKRNEKYKLPIPFPEGKVLDDMEGNQWVLGKKIGSGGFGLIYLAFPTNKPEKDARHVVKVEYQENGPLFSELKFY
QRVAKKDCIKKWIERKQLDYLGIPLFYGSGLTEFKGRSYRFMVMERLGIDLQKISGQNGTFKKSTVLQLGIRMLDVL
EYIHENEYVHGDIKAANLLIGYKNPDQVYLADYGLSYRYCPNGNHKQYQENPRKGHNGTIEFTSLDAHKGVALSRRS
DVEILGYCMLRWLCGKLPWEQNLKDPVAVQTAKTNLLDELPQSVLKWAPSGSSCCEIAQFLVCAHSLAYDEKPNYQA
LKKILNPHGIPLGPLDFSTKGQSINVHTPNSQKVDSQKAATKQVNKAHNRLIEKKVHSERSAESCATWKVQKEEKLI
GLMNNEAAQESTRRRQKYQESQEPLNEVNSFPQKISYTQFPNSFYEPHQDFTSPDIFKKSRSPSWYKYTSTVSTGIT
DLESSTGLWPTISQFTLSEETNADVYYYRIIIPVLLMLVFLALFFL(SEQ ID NO:278);

the vaccinia-related kinase 2B (VRK2B) protein, having a sequence:
MPPKRNEKYKLPIPFPEGKVLDDMEGNQWVLGKKIGSGGFGLIYLAFPTNKPEKDARHVVKVEYQENGPLFSELKFY
QRVAKKDCIKKWIERKQLDYLGIPLFYGSGLTEFKGRSYRFMVMERLGIDLQKISGQNGTFKKSTVLQLGIRMLDVL
EYIHENEYVHGDIKAANLLLGYKNPDQVYLADYGLSYRYCPNGNHKQYQENPRKGHNGTIEFTSLDAHKGVALSRRS
DVEILGYCMLRWLCGKLPWEQNLKDPVAVQTAKTNLLDELPQSVLKWAPSGSSCCEIAQFLVCAHSLAYDEKPNYQA
LKKILNPHGIPLGPLDFSTKGQSINVHTPNSQKVDSQKAATKQVNKAHNRLIEKKVHSERSAESCATWKVQKEEKLI
GLMNNEAAQVEA(SEQ ID NO:279); and the vaccinia related kinase 3 (VRK3) protein, encoding:
MISFCPDCGKSIQAAFKFCPYCGNSLPVEEHVGSQTFVNPHVSSFQGSKRGLNSSFETSPKKVKWSSTVTSPRLSLF
SDGDSSESEDTLSSSERSKGSGSRPPTPKSSPQKTRKSPQVTRGSPQKTSCSPQKTRQSPQTLKRSRVTTSLEALPT
GTVLTDKSGRQWKLKSFQTRDNQGILYEAAPTSTLTCDSGPQKQKFSLKLDAKDGRLFNEQNFFQRAAKPLQVNKWK
KLYSTPLLAIPTCMGFGVHQDKYRFLVLPSLGRSLQSALDVSPKHVLSERSVLQVACRLLDALEFLHENEYVHGNVT
AENIFVDPEDQSQVTLAGYGFAFRYCPSGKHVAYVEGSRSPHEGDLEFISMDLHKGCGPSRRSDLQSLGYCMLKWLY
GFLPWTNCLPNTEDIMKQKQKFVDKPGPFVGPCGHWIRPSETLQKYLKVVMALTYEEKPPYAMLRNNLEALLQDLRV
SPYDPIGLPMVP(SEQ ID NO:280).

In some embodiments, the VRK protein is a mutant that does not comprise a functional NLS, e.g. the sequence KTITKKRKKE (SEQ ID NO:281) and/or RQSSAKRHLAEQFAVGEIITDMAKKEWKVGLPI (SEQ ID NO:282) of VRK1, for example a deletion of the NLS or a substitution in the NLS that renders it inactive. In some embodiments, the VRK protein is a mutant comprising a modification that reduces the ability of an modulator to inhibit the kinase, e.g. the binding of an modulator of the kinase, for example, the binding of RAN-GDP. Kinase-active mutants of B1 proteins and their orthologs, including the VRK1, VRK2, and VRK3 proteins, can be readily determined by one of ordinary skill in the art by using well developed methods for detecting the presence of a phosphate group on a target protein to assess the ability of the subject kinase to phosphorylate BAF protein.

In some embodiments, the BAF modulator comprises a LEM domain or a nucleic acid, e.g., a DNA or an mRNA, encoding a LEM domain. As is known in the art, a LEM domain is a ~40-residue helix-loop-helix fold structure that directly binds to BAF. While the LEM domain may vary, LEM domains of interest include, but are not limited to: an Emerin LEM domain (MDNYADLSDTELTTLLRRYNIPHGPVVGSTRRLYEKKIFEYETQ(SEQ ID NO:283)); a LAP2-beta LEM domain: (DLDVTELTNEDLLDQLVKYGVNPGPIVGTTRKLYEKKLLKLREQ(SEQ ID NO:284)); a MAN1 LEM domain (ASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKKLKKLREE(SEQ ID NO:285)); and the like. In other embodiments, the BAF modulator does not comprise a LEM domain or a nucleic acid, e.g., a DNA or an mRNA, encoding a LEM domain. In some such embodiments, the BAF modulator does not comprise a LEM-domain containing protein 2 (LEMD2) polypeptide. In some such embodiments, the BAF modulator does not comprise a LEM-domain containing protein 3 (LEMD3) polypeptide.

In some embodiments, the protein agent that is the BAF modulator, e.g. as described herein, comprises a domain that reduces the half-life of the protein, e.g. a degradation domain. Examples of degradation domains include, but are not limited to NSGDSGYPSE(SEQ ID NO:286), SDNDSGICMS(SEQ ID NO:287), PSEDSGYGME(SEQ ID NO:288), DSGIHS(SEQ ID NO:289), SYLDSGIHSG(SEQ ID NO:290), SYLDSGIHSGATTTAPSLS(SEQ ID NO:291), KAAVSHWQQQSYLDSGIHSGATTTAPSLS(SEQ ID NO:292), PLPSGLLTPPQSPKKQS(SEQ ID NO:293), AARGLLTPPASPLELLEA(SEQ ID NO:294), GQSQGPPTPPTTPKTDV(SEQ ID NO:295), THGMMLTPPSSPLEL(SEQ ID NO:296), SGFKPLTPPTTPLSP(SEQ ID NO:295), LPTPP(SEQ ID NO:298), QFERQ(SEQ ID NO:299), KFERQ(SEQ ID NO:300), PEST(SEQ ID NO:301), and PESTEPTSPTSE(SEQ ID NO:302).

In some instances, it is advantageous to engineer into the BAF modulator the functionalities of a nuclear targeting factor (NTF) as known in the art and as described herein. Thus, while in some embodiments, the protein agent that is the BAF modulator, e.g. as described herein, does not comprise a heterologous nuclear localization sequence (NLS) domain, in other instances, the protein agent that is the BAF modulator is engineered to comprise an NLS domain. In some instances, the NLS is engineered to be within the center of the protein, i.e. the NLS is flanked on both N and C termini by 20 or more amino acids endogenous to the BAF modulator. In some instances, the NLS is engineered to be at the N terminus. In some instances,, the NLS is engineered to be at the C terminus. In some instances, the NLS is engineered to be at both the N terminus and the C terminus. When engineered to reside at a terminus, the NLS may be engineered to be within 0-20 amino acids of the terminus, in some instances, within 0-10 amino acids of the terminus, in certain instances within 0-5 amino acids of the terminus, in some such cases, at the terminus. In some instances, the NLS may be flanked by one or more amino acids, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids. In some instances, the BAF modulator will comprise one NLS, in other instances multiple NLSs. In some instances, in which multiple NLSs are employed, the same NLS is employed multiple times. In other instances, in which multiple NLSs are employed, different NLSs are used. In embodiments, when multiple NLSs are employed, they will be separated by a linker, e.g., a 6xK linker. Exemplary NLSs and their origins are provided in Table 5. Surprisingly, it has been observed that unlike other nuclear translocating DNA systems as known in the art and described herein, a BAF protein variant that is engineered to comprise an NLS does not require a specific nucleotide sequence, i.e. a DTS, on the DNA that it is translocating in order to be able to exert its additional function as an NTF. Thus, in such instance, the DNA does not comprise a BAF-modulator-specific DTS.

In some embodiments, the BAF modulator is a nucleic acid. BAF modulator nucleic acids may include DNA or RNA molecules, for example a DNA or RNA molecule encoding a protein that modulates BAF as described herein, some certain embodiments of which are described in Table 10

herein. In certain embodiments, the nucleic acids modulate, e.g., inhibit or reduce, the activity of a gene or protein, e.., by reducing or downregulating the expression of the gene. The nucleic acid may be single stranded or double-stranded and may include modified or unmodified nucleotides or non-nucleotides or various mixtures and combinations thereof. In some cases, the BAF modulator includes intracellular gene silencing molecules by way of RNA splicing and molecules that provide an antisense oligonucleotide effect or an RNA interference (RNAi) effect useful for inhibiting gene function. In some cases, gene silencing molecules, such as, e.g., antisense RNA, short temporary RNA (stRNA), double-stranded RNA (dsRNA), small interfering RNA (siRNA), short hairpin RNA (shRNA), microRNA (miRNA), tiny non-coding RNA (tncRNA), snRNA, snoRNA, and other RNAi-like small RNA constructs, may be used to target a protein-coding as well as non-protein-coding genes. In some cases, the nucleic acids include aptamers (e.g., spiegelmers). In some cases, the nucleic acids include antisense compounds. In some cases, the nucleic acids include molecules which may be utilized in RNA interference (RNAi) such as double stranded RNA including small interfering RNA (siRNA), locked nucleic acid (LNA) modulators, peptide nucleic acid (PNA) modulators, etc. In some cases, the nucleic acids include aptamers.

In some instances, the BAF modulator is a small molecule agent that exhibits the desired activity, e.g., inhibiting BAF expression and/or activity. Naturally occurring or synthetic small molecule compounds of interest include numerous chemical classes, such as organic molecules, e.g., small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents may include cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Such molecules may be identified, among other ways, by employing screening protocols. Small molecule BAF modulators include, but are not limited to, obtusilactone B, Mahubanolide, kotomolide B, epilitsenolide D2, rabeprazole, and the like.

Deoxyribonucleic Acids (DNAs)

As reviewed above, embodiments of the methods include contacting a cell with a DNA. DNAs are double stranded deoxyribonucleic acids that may vary in length, ranging in some instances from 15nt to 15,000 nt, such as 100 to 10,000 nt and including 100 to 5000nt. Where the DNA is double-stranded, the DNA may vary in length, ranging in some instances from 15 to 15,000 bp, such as 100 to 15,000 bp, e.g., 100 to 10,000 bp, including 100 to 5000 bp. The DNA may vary as desired. In some instances, the

DNA is 500 nt or more, for example 1 kb, 2 kb, 3 kb, 4 kb, or 5 kb or more, e.g., 6 kb, 7 kb, 8 kb, 9 kb, 10 kb or more, in some cases, 15 kb or more. The DNA may have any desired sequence. In some instances, the DNA may include one or more of coding sequences, promoters, sequences homologous to the genomic DNA of the targeted nucleus (e.g., to provide for genomic integration of the cargo nucleic acid), untranslated sequences (5' UTR, 3' UTR), polyadenylation sequences, and the like.

A given DNA to be delivered to the nucleus may be configured to be maintained episomally or integrated into the genome, as desired. As such, in some instances a DNA is configured to be maintained episomally in the nucleus of a target cell, such that it is not genomically integrated. In other instances, the DNA may be configured to be integrated into the genome of a target cell. In such instances, integration may be accomplished using any convenient protocol, such as by using a gene editing system, e.g., as described in greater detail below.

In some instances, the DNA includes a coding sequence. By a coding sequence it is meant a nucleic acid sequence that encodes for any gene product, e.g., micro RNA (miRNA), small hairpin RNA (shRNA), circular RNA (circRNA), long noncoding RNA (lncRNA), mRNA, peptide, polypeptide, protein. Where desired, a given coding sequence can encode multiple gene products, e.g., separated by IRES sequence, 2A sequence, or the like. For example, a coding sequence to be delivered to the nucleus by a DNA can be configured so that it can be integrated into the genome in operable linkage with its native promoter, for example to replace a mutant coding sequence or to be in operable linkage with an active promoter at a safe harbor (in which instances, the cargo nucleic acid need not include, and in some instances does not include, a promoter).

In some instances, a DNA may include a promoter. As used herein, the term "promoter" refers to any nucleic acid sequence that regulates the expression of another nucleic acid sequence by driving transcription of the nucleic acid sequence, which can be a heterologous target gene encoding a protein or an RNA. Promoters can be constitutive, inducible, repressible, tissue-specific, or any combination thereof. A promoter is a control region of a nucleic acid sequence at which initiation and rate of transcription of the remainder of a nucleic acid sequence are controlled. A promoter can also contain genetic elements at which regulatory proteins and molecules can bind, such as RNA polymerase and other transcription factors. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Various promoters, including inducible promoters, may be used to drive the expression of transgenes. A promoter sequence may be bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Promoters useful in DNAs of embodiments of the invention include, for example constitutively active

promoters, such as the CMV promoter, CAG promoter (which combines the CMV enhancer with the chicken β-actin (CBA) promoter), β-actin promoter, SV-40 promoter, 4xGRM6-SV40, hTTR, hAAT, 3x-Serpina, ubiquitin B/C, EF1-Alpha, EFS, and HBV promoter, etc. Promoters useful in DNAs of embodiments of the invention also include promoters having more cell-type specific expression patterns, for example for hepatocytes, may include, without limitation the TTR, hAAT (and derivatives), 3xSerpina-TTR, HBV, UbiC, and P3-hybrid promoter. A cargo nucleic acid may include a promoter sequence to be delivered to the nucleus so that it can be integrated into the genome, for example to replace a mutant promoter.

In some instances, the DNA includes an expression cassette. By an expression cassette it is meant a nucleic acid sequence comprising a promoter, e.g., as described above, operably linked to a coding sequence, e.g., as described above (also referred to herein as a transgene). In some instances, the expression cassette may also comprise one or more nucleic acid sequences including a 5' untranslated region (5' UTR), 3' untranslated region (3' UTR), polyA tail, miRNA regulatory element. In embodiments, the expression cassette may include a transgene and one or more regulatory sequences that allows and/or controls the expression of the transgene, e.g., where the expression cassette can include one or more of, e.g., in this order: an enhancer/promoter, an ORF reporter (transgene), a post-transcription regulatory element (e.g., WPRE), and a polyadenylation and termination signal (e.g., BGH polyA). The expression cassette can also comprise an internal ribosome entry site (IRES) and/or a 2A element. The cis-regulatory elements include, but are not limited to, a promoter, a riboswitch, an insulator, a mir-regulatable element, a post-transcriptional regulatory element, a tissue- and cell type-specific promoter and an enhancer. As desired, the expression cassette can comprise 4000 or more nucleotides, 5000 or more nucleotides, 10,000 or more nucleotides or 20,000 or more nucleotides, or 30,000 or more nucleotides, or 40,000 or more nucleotides or 50,000 or more nucleotides, and in some instances may range between 4000-10,000 nucleotides or 10,000-50,000 nucleotides, or more than 50,000 nucleotides. The expression cassette to be delivered to the nucleus may be configured so that it can be maintained episomally. Alternatively, an expression cassette to be delivered to the nucleus may be configured so that it can be integrated into the genome.

The coding sequence e.g., transgene, of the expression cassette may vary. In some embodiments, the expression cassette can include a transgene in the range of 500 to 50,000 nucleotides in length. In some embodiments, the expression cassette can include a transgene in the range of 500 to 75,000 nucleotides in length. In some embodiments, the expression cassette can include a transgene which is in the range of 500 to 10,000 nucleotides in length. In some embodiments, the expression cassette can include a transgene which is in the range of 1000 to 10,000 nucleotides in length. In some embodiments, the expression cassette can include a transgene which is in the range of 500 to 5,000 nucleotides in length.

The DNA constructs of embodiments of the invention do not have the size limitations of encapsidated AAV vectors, and thus enable nuclear delivery of large-size expression cassettes to provide efficient transgene.

A given expression cassette can include, for example, an expressible exogenous sequence (e.g., open reading frame) or transgene that encodes a protein that is either absent, inactive, or insufficient activity in the recipient subject or a gene that encodes a protein having a desired biological or a therapeutic effect. The transgene can encode a gene product that can function to correct the expression of a defective gene or transcript. In principle, the expression cassette can include any gene that encodes a protein, polypeptide or RNA that is either reduced or absent due to a mutation or which conveys a therapeutic benefit when overexpressed is considered to be within the scope of the disclosure. The expression cassette can include any transgene useful for treating a disease or disorder in a subject.

DNAs, e.g., as described herein, can be used to deliver and express any gene of interest in a subject, where such genes include, but are not not limited to, nucleic acids encoding polypeptides, or non-coding nucleic acids (e.g., RNAi, miRs etc.), as well as exogenous genes and nucleotide sequences, including virus sequences in a subjects' genome, e.g., HIV virus sequences and the like. In some instances, a DNA (e.g., as disclosed herein) is used for therapeutic purposes (e.g., for medical, diagnostic, or veterinary uses) or immunogenic polypeptides. In certain embodiments, a DNA is useful to express any gene of interest in the subject, which includes one or more polypeptides, peptides, ribozymes, peptide nucleic acids, siRNAs, RNAis, antisense oligonucleotides, antisense polynucleotides, or RNAs (coding or non-coding; e.g., siRNAs, shRNAs, micro-RNAs, and their antisense counterparts (e.g., antagoMiR)), antibodies, antigen binding fragments, or any combination thereof. As such, expression cassettes can encode polypeptides, sense or antisense oligonucleotides, or RNAs (coding or non-coding; e.g., siRNAs, shRNAs, micro-RNAs, and their antisense counterparts (e.g., antagoMiR)). Expression cassettes can include an exogenous sequence that encodes a reporter protein to be used for experimental or diagnostic purposes, such as β-lactamase, β-galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, EPO, and others well known in the art.

Sequences provided in a given expression cassette, expression construct of a DNA, e.g., as described herein, can be codon optimized for the target host cell. As used herein, the term "codon optimized" or "codon optimization" refers to the process of modifying a nucleic acid sequence for enhanced expression in the cells of the vertebrate of interest, e.g., mouse or human, by replacing at least one, more than one, or a significant number of codons of the native sequence (e.g., a prokaryotic sequence) with codons that are more frequently or most frequently used in the genes of that vertebrate. Various species exhibit particular bias for certain codons of a particular amino acid. Typically, codon

optimization does not alter the amino acid sequence of the original translated protein. In some embodiments, a transgene expressed by the DNA is a therapeutic gene. In some embodiments, a therapeutic gene is an antibody, or antibody fragment, or antigen-binding fragment thereof, e.g., a neutralizing antibody or antibody fragment and the like. In some instances, a therapeutic gene is one or more therapeutic agent(s), including, but not limited to, for example, protein(s), polypeptide(s), peptide(s), enzyme(s), antibodies, antigen binding fragments, as well as variants, and/or active fragments thereof, for use in the treatment, prophylaxis, and/or amelioration of one or more symptoms of a disease, dysfunction, injury, and/or disorder. A coding sequence to be delivered to the nucleus so that it can be integrated into the genome in operable linkage with its native promoter, for example to replace a mutant coding sequence or to be in operable linkage with an active promoter at a safe harbor (in which instances, the polynucleotide does not comprise a promoter); or the coding sequence of an expression cassette to be delivered to the nucleus to be either maintained episomally or integrated into the genome. By a coding sequence it is meant a nucleic acid sequence that encodes for any gene product, e.g., micro RNA (miRNA), small hairpin RNA (shRNA), circular RNA (circRNA), long noncoding RNA (lncRNA), mRNA, peptide, polypeptide, protein. Can encode multiple gene products, e.g., separated by IRES sequence, 2A sequence, or the like.

Where desired, a given DNA may include flanking sequences that are homologous to genomic regions of the cell, e.g., to promote genomic integration of the cargo nucleic acid. When present, such sequences may vary in length, ranging in some instances from 30 to 5,000 nt, such as 50 to 1000 nt. While the sequences of such regions may vary depending on the genomic integration location of interest, examples of such sequences include, but are not limited to: actin, ADA, albumin, α-globin, β-globin, CD2, CD3, CD5, CD7, CCR5, E1α, IL2RG, Ins1, Ins2, NCF1, p50, p65, PF4, PGC-γ, PTEN, TERT, TRAC, UBC, and VWF, and the like.

In embodiments where the DNA is to be integrated into the genome of a target cell in process mediated by a gene editing system, e.g., as described below, the DNA may be serve as a donor DNA or donor template in such a system. Site-directed polypeptides, such as a DNA endonuclease, can introduce double-strand breaks or single-strand breaks in nucleic acids, e.g., genomic DNA. The double-strand break can stimulate a cell's endogenous DNA-repair pathways (e.g., homology-dependent repair (HDR) or non-homologous end joining or alternative non-homologous end joining (A-NHEJ) or microhomology-mediated end joining (MMEJ). NHEJ can repair cleaved target nucleic acid without the need for a homologous template. This can sometimes result in small deletions or insertions (indels) in the target nucleic acid at the site of cleavage, and can lead to disruption or alteration of gene expression. HDR, which is also known as homologous recombination (HR) can occur when a homologous repair template, or donor, is available.

The homologous donor template has sequences that are homologous to sequences flanking the target nucleic acid cleavage site. The sister chromatid is generally used by the cell as the repair template. However, for the purposes of genome editing, the repair template is often supplied as an exogenous nucleic acid, such as a plasmid, duplex oligonucleotide, single-strand oligonucleotide, double-stranded oligonucleotide, or viral nucleic acid. With exogenous donor templates, it is common to introduce an additional nucleic acid sequence (such as a transgene) or modification (such as a single or multiple base change or a deletion) between the flanking regions of homology so that the additional or altered nucleic acid sequence also becomes incorporated into the target locus. MMEJ results in a genetic outcome that is similar to NHEJ in that small deletions and insertions can occur at the cleavage site. MMEJ makes use of homologous sequences of a few base pairs flanking the cleavage site to drive a favored end-joining DNA repair outcome. In some instances, it can be possible to predict likely repair outcomes based on analysis of potential microhomologies in the nuclease target regions.

Thus, in some cases, homologous recombination is used to insert an exogenous polynucleotide sequence into the target nucleic acid cleavage site. An exogenous polynucleotide sequence is termed a donor polynucleotide (or donor DNA or donor sequence or polynucleotide donor template) herein, and may in embodiments of the present invention be a DNA or component thereof, e.g., cargo nucleic acid component of a DNA. In some embodiments, the donor polynucleotide, a portion of the donor polynucleotide, a copy of the donor polynucleotide, or a portion of a copy of the donor polynucleotide is inserted into the target nucleic acid cleavage site. In some embodiments, the donor polynucleotide is an exogenous polynucleotide sequence, i.e., a sequence that does not naturally occur at the target nucleic acid cleavage site.

When an exogenous DNA molecule is supplied in sufficient concentration inside the nucleus of a cell in which the double strand break occurs, the exogenous DNA can be inserted at the double strand break during the NHEJ repair process and thus become a permanent addition to the genome. These exogenous DNA molecules are referred to as donor templates in some embodiments. If the donor template contains a coding sequence for a gene-of-interest optionally together with relevant regulatory sequences such as promoters, enhancers, polyA sequences and/or splice acceptor sequences (also referred to herein as a "donor cassette"), the gene of interest can be expressed from the integrated copy in the genome resulting in permanent expression for the life of the cell. Moreover, the integrated copy of the donor DNA template can be transmitted to the daughter cells when the cell divides.

In the presence of sufficient concentrations of a donor DNA template that contains flanking DNA sequences with homology to the DNA sequence either side of the double strand break (referred to as homology arms), the donor DNA template can be integrated via the HDR pathway. The homology arms act as substrates for homologous recombination between the donor template and the sequences either side

of the double strand break. This can result in an error free insertion of the donor template in which the sequences either side of the double strand break are not altered from that in the un-modified genome.

Supplied donors for editing by HDR vary markedly but generally contain the intended sequence with small or large flanking homology arms to allow annealing to the genomic DNA. The homology regions flanking the introduced genetic changes can be 30 bp or smaller, or as large as a multi-kilobase cassette that can contain promoters, cDNAs, etc. Both single-stranded and double-stranded oligonucleotide donors can be used. These oligonucleotides range in size from less than 100 nt to over many kb, though longer ssDNA can also be generated and used. Double-stranded donors are often used, including PCR amplicons, plasmids, and mini-circles.

In some embodiments, an exogenous sequence, e.g., cargo nucleic acid, that is intended to be inserted into a genome is a gene-of-interest (GOI) or functional derivative thereof. The exogenous gene can include a nucleotide sequence encoding a GOI product, e.g., GOI protein, or functional derivative thereof. The functional derivative of a GOI can include a nucleic acid sequence encoding a functional derivative of a GOI protein that has a substantial activity of a wildtype GOI protein such as. the wildtype human GOI protein, e.g., at least about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95% or about 100% of the activity that the wildtype GOI protein exhibits. In some embodiments, the functional derivative of a GOI protein can have at least about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98% or about 99% amino acid sequence identity to the GOI protein, e.g., the wildtype GOI protein. In some embodiments, one having ordinary skill in the art can use a number of methods known in the field to test the functionality or activity of a compound, e.g., peptide or protein. The functional derivative of the GOI protein can also include any fragment of the wildtype GOI protein or fragment of a modified GOI protein that has conservative modification on one or more of amino acid residues in the full length, wildtype GOI protein. Thus, in some embodiments, the functional derivative of a nucleic acid sequence of a GOI can have at least about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98% or about 99% nucleic acid sequence identity to the GOI, e g. the wildtype GOI.

In some embodiments where the insertion of a GOI or functional derivative thereof is concerned, a cDNA of GOI or functional derivative thereof can be inserted into a genome of a patient having defective GOI or its regulatory sequences. In such a case, a donor DNA or donor template can be an expression cassette or vector construct having the sequence encoding GOI or functional derivative thereof, e.g., cDNA sequence. In some embodiments, the expression vector contains a sequence encoding a modified GOI protein, which is described elsewhere in the disclosures, can be used.

In some embodiments, according to any of the donor templates described herein comprising a donor cassette, the donor cassette is flanked on one or both sides by a gRNA target site. For example, such a donor template may comprise a donor cassette with a gRNA target site 5' of the donor cassette and/or a gRNA target site 3' of the donor cassette. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 3' of the donor cassette. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette and a gRNA target site 3' of the donor cassette. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette and a gRNA target site 3' of the donor cassette, and the two gRNA target sites comprise the same sequence. In some embodiments, the donor template comprises at least one gRNA target site, and the at least one gRNA target site in the donor template comprises the same sequence as a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated. In some embodiments, the donor template comprises at least one gRNA target site, and the at least one gRNA target site in the donor template comprises the reverse complement of a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated. In some embodiments, the donor template comprises at least one gRNA target site, and the at least one gRNA target site in the donor template does not comprise the same sequence as a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette and a gRNA target site 3' of the donor cassette, and the two gRNA target sites in the donor template comprises the same sequence as a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette and a gRNA target site 3' of the donor cassette, and the two gRNA target sites in the donor template comprises the reverse complement of a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated. In some embodiments, the donor template comprises a donor cassette with a gRNA target site 5' of the donor cassette and a gRNA target site 3' of the donor cassette, and the two gRNA target sites in the donor template do not comprise the same sequence as a gRNA target site in a target locus into which the donor cassette of the donor template is to be integrated.

A given DNA may or may not be configured to be maintained in a bacterium, as desired. For example, in some instances a DNA is configured to be maintained in a bacterium. In such instances, the DNA may include a bacterial origin of replication and a selection element, e.g., a plasmid or a nanoplasmid. In other instances, a DNA may be configured to not be maintained in a bacterium. In such instances, the structure of the DNA may vary, wherein examples of such structures include minicircle

DNA (mcDNA), a linear DNA, a covalently closed DNA (doggybone DNA, or "dbDNA", ministring DNA, etc.), double stranded linear DNA comprising terminal repeat sequences at both ends, a 3DNA structure, etc., and the like.

Nuclear Targeted Deoxyribonucleic Acids (NTDNAs) and Nuclear Targeting Factors

Nuclear Targeted Deoxyribonucleic Acids (NTDNAs). In some embodiments, the DNA, e.g., as described above, is a component of a Nuclear Targeted Deoxyribonucleic Acid (NTDNA). NTDNAs finding use in embodiments of the invention are described in PCT Application Serial No. PCT/US2023/035932 filed on October 25, 2023, the disclosure of which is herein incorporated by reference. NTDNAs employed in methods of the invention comprise both a DNA nuclear targeting sequence (DTS) and a DNA, e.g., as described above, which may be referred to as a "cargo nucleic acid sequence" in the context of a NTDNA, where the cargo nucleic acid is heterologous to the DTS. As such, the NTDNAs include a DTS domain and a cargo nucleic acid domain that is heterologous to the DTS. By "heterologous" to the DTS it is meant that the cargo nucleic acid is not naturally associated with the DTS, e.g., is not part of the same gene as the DTS in nature. NTDNAs employed in methods of the invention comprise both a DNA nuclear targeting sequence (DTS) and a cargo nucleic acid sequence that is heterologous to the DTS. As such, the NTDNAs include a DTS domain and a cargo nucleic acid domain. Each of these domains is now described further in greater detail.

A DTS refers to a nucleotide sequence that mediates the translocation of a polynucleotide that comprises it into the nucleus of a cell. Without wishing to be bound by theory, it is believed that DTSs leverage the movement of nuclear-acting DNA binding proteins as they move from the cytoplasm into the nucleus. These nuclear-acting DNA binding proteins act like nuclear targeting factors for DNA, binding to sequences on the DNA and dragging the DNA into the nucleus as the DNA binding protein translocates into the nucleus. Accordingly, the nuclear-acting DNA binding proteins that are leveraged are referred to herein as nuclear targeting factors (NTFs), and the DNA sequences to which they bind are referred to herein as nuclear targeting factor binding sites (NTFBSs, or more simply, TFBSs).

In some embodiments, the DTS comprises a TFBS that is bound by a nuclear targeting factor that is provided to the cell as a protein or an mRNA encoding a protein. As will be appreciated by one of ordinary skill in the art, any protein that binds to DNA and that traffics to the nucleus when delivered to the cytoplasm can be provided to the cell to mediate nuclear translocation. Thus, for example, any of the naturally occurring proteins described in Tables 1 and 2 may be exogenously provided. As another example, a protein that is not native to the cell, i.e., that is heterologous to the cell, may be provided. Such a protein may be naturally occurring or engineered. Examples include any of the proteins listed in

Table 1. Exemplary proteins of each class and the DNA sequences to which they bind are well known in the art and include those described in greater detail below.

Table 1. Classes of nuclear targeting factors that may be co-delivered with NTDNA to achieve DNA nuclear transport.

Protein/DNA binding system	**Example NTF protein**
Bacterial regulatory proteins	Tetracycline repressor (TetR), Purine synthesis repressor (PurR); Lac repressor (LacR)
Yeast regulatory proteins	Gal4; GCN4
Bacteriophage regulatory proteins	Arc; Mnt; Bac434
Mammalian regulatory proteins	YY1; HEY1; MyoD; GR; ER; PR
Zing finger proteins (ZFP)	ZFP-CCR5
Transcription Activator-Like Effectors (TALE)	TALE-CCR5
Inactivated nucleases	Inactivated I-SceI
Inactivated Site-Specific Recombinases	Cre; Piggybac
Transcription factor DNA-binding domain	STAT1; ESR1; Gal4
Inactivated RNA-guided endonuclease	dCas9; HypaCas9; miniCas9; Cas12a
Single strand DNA binding protein	hSSB1; POT1

In some embodiments, the TFBS is a binding sequence for a Tet Repressor (TetR) protein, for example, a TetO sequence such as YCTATCANTGATAGA (SEQ ID NO:303), for example, TCCCTATCAGTGATAGAGA(SEQ ID NO:304) or TCGAGTTTACTCCCTATCAGTGATAGAGAACG(SEQ ID NO:305). In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the TetR TFBS. In certain embodiments, the DTS comprises 7 copies of the TetR TFBS. In certain embodiments, the DTS comprises 10 or more copies of the TetR TFBS. In some embodiments, the DTS comprises a sequence having 80% identity or more to a sequence listed in Table 2, for example, 85% or 90% identity or more, in some cases 95% identity or more, e.g., 96%, 97%, 98%, or 99% identity to a sequence in Table 2. In some instances, the DTS binding sequence is identical to a sequence in Table 2. In some instances, the DTS consists essential of a sequence in Table 2. In certain embodiments, the DTS comprises a tetracycline response element (TRE), as known in the art. In certain embodiments, the DTS consists essentially of a TRE.

Table 2. Examples of DTSs that comprise a TetR binding sequence (e.g., TetO) in varying numbers and sequences.

DTS(SEQ ID)	**Sequence (5' to 3')**
TetO Motif	YCTATCANTGATAGA (SEQ ID NO:306)
TetO-1 (Core)	TCCCTATCAGTGATAGAGA (SEQ ID NO:307)

TetO-2	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTAC(SEQ ID NO:308)
TetO-3 (309)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTAC
TetO-4 (310)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT AGAGAACGTATGTCGAGTTTAC
TetO-5 (311)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT AGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTA
TetO-6 (312)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT AGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCTATCA GTCATAGAGAACGTATGTCGAGTTTAC
TetO-7 (313)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT AGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCTATCA GTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGGTAGGCGT GTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGAATTCA
TetO-7-m (314)	CTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGA ACGATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGAT AGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCTATCA GTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGGTA

In some embodiments, the TFBS is a binding sequence for the DNA binding domain of a gene editing system, e.g., as described herein or as known in the art, e.g., the guide RNA of a Cas nuclease, the zinc finger domain of a zinc finger nuclease, the TALE DNA binding domain of a TALEN.

In some embodiments, the TFBS is a binding sequence for a zinc-finger containing protein ("ZF protein"). As will be appreciated by the ordinarily skilled artisan, any ZF protein and its cognate ZF binding sequence may be used as a nuclear targeting factor (NTF) and cognate TFBS in the compositions and methods of the present disclosure. In some embodiments the DTS comprises a ZF-responsive TFBS having a sequence identity of 90% or more to AAACTGCAAAAG(SEQ ID NO:315).

In some embodiments, the TFBS is a binding sequence for a TAL effector DNA-binding domain-containing protein ("TALE protein"). A TALE is a protein of 32 fixed amino acids and 2 variable residues, the 2 residues being engineered to recognize specific nucleotides (e.g., NN for G, NI for A, HD for C, etc.). As will be appreciated by the ordinarily skilled artisan, any TALE protein and its cognate TALE binding sequence may be used in the compositions and methods of the present disclosure, such examples being found in, e.g., Li et al. 2011 (*Modularly assembled designer TAL effector nucleases for*

targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Research, Volume 39, Issue 14, pp 6315–6325) and Kim et. al. 2013 (*A library of TAL effector nucleases spanning the human genome.* Nature Biotechnology volume 31, pages 251–258 (2013). In some embodiments, the TALE TFBS has 90% identity or more to a sequence is selected from the group consisting of TTCATTACACCTGCAGCT(SEQ ID NO:316), ATAAACCCCCTCCAA(SEQ ID NO:317), and TCGAGTTTACTCCCTATCAGTGATAGAGAACG(SEQ ID NO:318).

In some embodiments, the TFBS is a binding sequence for a GAL4 protein. In some such embodiments, the GAL4 TFBS comprises the sequence CGG-N_{11}-CCG, for example, CGGAGGACTGTCCTCCG(SEQ ID NO:319). In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the GAL4 TFBS. In certain embodiments, the DTS comprises 5 copies of the GAL4 TFBS. In certain embodiments, the DTS comprises 7 copies of the GAL4 TFBS. In certain embodiments, the DTS comprises 10 or more copies of the GAL4 TFBS. In certain embodiments, the DTS comprises an upstream activation sequence (UAS) for the native GAL4 protein, as known in the art. In certain embodiments, the DTS consists essentially of a UAS. In some embodiments, the DTS comprises a sequence having 80% identity or more to a sequence listed in Table 3, for example, 85% or 90% identity or more, in some cases 95% identity or more, e.g., 96%, 97%, 98%, or 99% identity to a sequence in Table 3. In some instances, the DTS is identical to a sequence in Table 3.

Table 3. Examples of DTSs comprising a GAL4 TFBS (e.g., UAS)

DTS	Sequence (5' to 3') (SEQ ID)
UAS Motif	CGG-N_{11}-CCG(SEQ ID NO:319)
UAS Core A	CGGTGGCTTCTAATCCG (SEQ ID NO:320)
UAS Core B	CGGGTGACAGCCCTCCG (SEQ ID NO:321)
UAS Core C	CGGAGGAGAGTCTTCCG (SEQ ID NO:322)
UAS Core D	CGGAGTACTGTCCTCCG (SEQ ID NO:323)
UAS Core E	CGGAGGACTGTCCTCCG (SEQ ID NO:324)
UAS-1	ACACTAGAGTCGGAGGACTGTCCTCCGTGAGTCCTAG (SEQ ID NO:325)
UAS-2	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCG (SEQ ID NO:326)
UAS-3	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCG (SEQ ID NO:327)
UAS-4	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCG (SEQ ID NO:328)
UAS-5	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCG (SEQ ID NO:329)
UAS-5-NR	CGGTGGCTTCTAATCCGTGAGTCCTAGCGGGTGACAGCCCTCCGTCTTCACAGGCGGAGGAGAGTC TTCCGTAGGGTTCCTCGGAGTACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCG (SEQ ID NO:330)

UAS-6	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCGGATAGTA TCACGGAGGACTGTCCTCCG (SEQ ID NO:331)
UAS-7	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCGGATAGTA TCACGGAGGACTGTCCTCCGAGCAAACGAACGGAGGACTGTCCTCCG(SEQ ID NO:332)
UAS-10	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCGGATAGTA TCACGGAGGACTGTCCTCCGAATGGACTAACGGAGGACTGTCCTCCGAGCAAACGAACGGAGGACT GTCCTCCGTACACCGATCCGGAGGACTGTCCTCCGGCGCTCGGCACGGAGGACTGTCCTCCG(SEQ ID NO:333)
UAS-20	CGGAGGACTGTCCTCCGTGAGTCCTAGCGGAGGACTGTCCTCCGTCTTCACAGGCGGAGGACTGTC CTCCGTAGGGTTCCTCGGAGGACTGTCCTCCGACACTAGAGTCGGAGGACTGTCCTCCGGATAGTA TCACGGAGGACTGTCCTCCGAATGGACTAACGGAGGACTGTCCTCCGACTTTAGCGCCGGAGGACT GTCCTCCGTACACCGATCCGGAGGACTGTCCTCCGGCGCTCGGCACGGAGGACTGTCCTCCGCGAA AACGGGCGGAGGACTGTCCTCCGTGCGTGGATTCGGAGGACTGTCCTCCGAGCCAGGAGCCGGAGG ACTGTCCTCCGAAGACCACGGCGGAGGACTGTCCTCCGACCAGTACCCCGGAGGACTGTCCTCCGA TGTCTTGTACGGAGGACTGTCCTCCGCTGACTATCGCGGAGGACTGTCCTCCGCCTACGTCATCGG AGGACTGTCCTCCGATAGAAACCTCGGAGGACTGTCCTCCGCATGTCTAACCGGAGGACTGTCCTC CG (SEQ ID NO:334)

In some embodiments, the TFBS is a binding sequence for an Arc protein, where Arc is a bacteriophage regulatory protein. In some such embodiments, the Arc TFBS comprises the sequence RYRVTAGANNNNNTCTABYRY(SEQ ID NO:335), for example, ATGATAGAAGCACTCTACTAT(SEQ ID NO:336). In some embodiments, the Arc TFBS comprises a sequence having 80%, 85%, 90% identity or more to ATGATAGAAGCACTCTACTAT(SEQ ID NO:337). In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the Arc TFBS. In certain embodiments, the DTS comprises 5 copies of the Arc TFBS. In certain embodiments, the DTS comprises 7 copies of the Arc TFBS. In certain embodiments, the DTS comprises 10 or more copies of the Arc TFBS. In some embodiments, the DTS comprises a sequence having 80%, 85%, or 90% identity or more to the sequence ATGATAGAAGCACTCTACTATTGAGTCCTAGATGATAGAAGCACTCTACTATTCTTCACAGGATGATAGAAGCACTC TACTATTAGGGTTCCTATGATAGAAGCACTCTACTATACACTAGAGTATGATAGAAGCACTCTACTATGATAGTATC AATGATAGAAGCACTCTACTATAGCAAACGAAATGATAGAAGCACTCTACTAT(SEQ ID NO:338), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence.

In some embodiments, the TFBS is a binding sequence for a Mnt protein, where Mnt is a bacteriophage regulatory protein. In some such embodiments, the Mnt TFBS comprises the sequence GGNCCACNGTGGNCC(SEQ ID NO:339), for example, ATAGGTCCACGGTGGACCATA(SEQ ID NO:340). In some embodiments, the Mnt TFBS comprises a sequence having 80%, 85%, 90% identity or more to ATAGGTCCACGGTGGACCATA. In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10

copies of the Mnt TFBS. In certain embodiments, the DTS comprises 5 copies of the Mnt TFBS. In certain embodiments, the DTS comprises 7 copies of the Mnt TFBS. In certain embodiments, the DTS comprises 10 or more copies of the Mnt TFBS. In some embodiments, the DTS comprises a sequence having 80%, 85%, or 90% identity or more to the sequence ATAGGTCCACGGTGGACCATATGAGTCCTAGATAGGTCCACGGTGGACCATATCTTCACAGGATAGGTCCACGGTGG ACCATATAGGGTTCCTATAGGTCCACGGTGGACCATAACACTAGAGTATAGGTCCACGGTGGACCATAGATAGTATC AATAGGTCCACGGTGGACCATAAGCAAACGAAATAGGTCCACGGTGGACCATA (SEQ ID NO:341), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence.

In some embodiments, the TFBS is a binding sequence for a purine synthesis repressor (PurR) protein. In some such embodiments, the PurR TFBS comprises a sequence having 80%, 85%, 90% identity or more to ACGCAAACGTTTTCGT (SEQ ID NO:342), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence. In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the PurR TFBS. In certain embodiments, the DTS comprises 5 copies of the PurR TFBS. In certain embodiments, the DTS comprises 7 copies of the PurR TFBS. In certain embodiments, the DTS comprises 10 or more copies of the PurR TFBS. I In some embodiments, the DTS comprises a sequence having 80%, 85%, or 90% identity or more to the sequence ACGCAAACGTTTTCGTTGAGTCCTAGACGCAAACGTTTTCGTTCTTCACAGGACGCAAACGTTTTCGTTAGGGTTCC TACGCAAACGTTTTCGTACACTAGAGTACGCAAACGTTTTCGTGATAGTATCAACGCAAACGTTTTCGTAGCAAACG AAACGCAAACGTTTTCGT (SEQ ID NO:343), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence.

In some embodiments, the TFBS is a binding sequence for a Bac434 protein, where Bac434 is a bacteriophage regulatory protein. In some such embodiments, the Bac434 TFBS comprises a sequence that has 80%, 85%, 90% identity or more to ACAAGAAAGTTTGT (SEQ ID NO:344), ACAAGATACATTGT (SEQ ID NO:345), or ACAAGAAAAACTGT (SEQ ID NO:346), in some cases 95% identity or more to ACAAGAAAGTTTGT (SEQ ID NO:347), ACAAGATACATTGT (SEQ ID NO:348), or ACAAGAAAAACTGT (SEQ ID NO:349), in certain cases sharing 100% identity with ACAAGAAAGTTTGT (SEQ ID NO:350), ACAAGATACATTGT (SEQ ID NO:351), or ACAAGAAAAACTGT (SEQ ID NO:352). In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the Bac434 TFBS. In certain embodiments, the DTS comprises 5 copies of the Bac434 TFBS. In certain embodiments, the DTS comprises 7 copies of the Bac434 TFBS. In certain embodiments, the DTS comprises 10 or more copies of the Bac434 TFBS. In certain embodiments, the DTS comprises a sequence having 80%, 85%, or 90% identity or more to a sequence listed in Table 4, for example, 85% or 90% identity or more, in some cases 95% identity or more, e.g., 96%, 97%, 98%, or 99% identity to a sequence in Table 4. In some instances, the DTS is identical to a sequence in Table 4.

Table 4. Examples of DTSs comprising a Bac434 TFBS

DTS	Sequence (5' to 3')
Bac434-OR1-Core	ACAAGAAAGTTTGT (SEQ ID NO:353)
Bac434-OR2-Core	ACAAGATACATTGT (SEQ ID NO:354)
Bac434-OR3-Core	ACAAGAAAAACTGT (SEQ ID NO:355)
Bac434-OR321-Core	ACAAGAAAAACTGTATTTGACAAACAAGATACATTGTATGAAAATACAAGAAAGTTTGT (SEQ ID NO:356)
Bac434-OR1-7	ACAAGAAAGTTTGTTGAGTCCTAGACAAGAAAGTTTGTTCTTCACAGGACAAGAAAGTTTG TTAGGGTTCCTACAAGAAAGTTTGTACACTAGAGTACAAGAAAGTTTGTGATAGTATCAAC AAGAAAGTTTGTAGCAAACGAAACAAGAAAGTTTGT (SEQ ID NO:357)
Bac434-OR3-7	ACAAGAAAAACTGTTGAGTCCTAGACAAGAAAAACTGTTCTTCACAGGACAAGAAAAACTG TTAGGGTTCCTACAAGAAAAACTGTACACTAGAGTACAAGAAAAACTGTGATAGTATCAAC AAGAAAAACTGTAGCAAACGAAACAAGAAAAACTGT (SEQ ID NO:358)
Bac434-OR321-3	CCCAATCTTCACAAGAAAAACTGTATTTGACAAACAAGATACATTGTATGAAAATACAAGA AAGTTTGTTGATGGAGGCACAAGAAAAACTGTATTTGACAAACAAGATACATTGTATGAAA ATACAAGAAAGTTTGTTGATGGAGGCACAAGAAAAACTGTATTTGACAAACAAGATACATT GTATGAAAATACAAGAAAGTTTGTTGATGGAGGC (SEQ ID NO:359)

In some embodiments, the TFBS is a binding sequence for a GCN4 protein. In some such embodiments, the GCN4 TFBS comprises a sequence having 80%, 85%, 90% identity or more to TGACTC, in some cases 95% identity or more to TGACTC, in certain cases 100% identity with TGACTC, for example, AGTGACTCATT. In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the GCN4 TFBS. In certain embodiments, the DTS comprises 5 copies of the GCN4 TFBS. In certain embodiments, the DTS comprises 7 copies of the GCN4 TFBS. In certain embodiments, the DTS comprises 10 or more copies of the GCN4 TFBS. In some embodiments, the DTS comprises a sequence having 80%, 85%, or 90% identity or more to the sequence AGTGACTCATTTGAGTCCTAGAGTGACTCATTTCTTCACAGGAGTGACTCATTTAGGGTTCCTAGTGACTCATTACA CTAGAGTAGTGACTCATTGATAGTATCAAGTGACTCATTAGCAAACGAAAGTGACTCATT(SEQ ID NO:360), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence.

In some embodiments, the TFBS is a binding sequence for a Lactose Modulator (LacI) protein, also referred to herein as a Lactose Repressor (LacR) protein, for example, a LacO sequence, e.g., TTGTTATCCGCTCACAA(SEQ ID NO:361). In some such embodiments, the LacR TFBS comprises a sequence having 80%, 85%, 90% identity or more to TTGTTATCCGCTCACAA(SEQ ID NO:362). In certain embodiments, the DTS comprises a lactose operon (LacO), as is known in the art. In certain embodiments, the DTS consists essentially of a LacO. In some embodiments, the DTS comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the LacR TFBS. In certain embodiments, the DTS comprises 5 copies of the LacR TFBS. In certain embodiments, the DTS comprises 7 copies of the LacR TFBS. In certain

embodiments, the DTS comprises 10 or more copies of the LacR TFBS. In some embodiments, the DTS comprises a sequence that has 80%, 85%, or 90% identity or more to the sequence TTGTTATCCGCTCACAATTCCACATGTGGCCACAAATTGTTATCCGCTCACAATTCCACATGTGGCCACAAATTGTTATCCGCTCACAATTCCACATGTGGCCACAAATTGTTATCCGCTCACAATTCCACATGTGGCCACAAATTGTTATCCGCTCACAATTCCACATGTGGCCACAAATTGTTATCCGCTCAC AA(SEQ ID NO:363), in some cases 95% identity or more to this sequence, in certain cases sharing 100% identity with this sequence.

In some embodiments, the TFBS is a binding sequence for an endonuclease, e.g., an I-SceI D44A protein. In some embodiments the DTS comprises a TFBS having 90% identity or more to the sequence TAGGGATAACAGGGTAAT(SEQ ID NO:364).

It will be appreciated by one of ordinary skill in the art that DNA binding proteins can tolerate some degree of nucleotide substitution in their binding sequences and that the TFBSs provided herein are but examples of sequences that may be used. In some embodiments, the TFBS shares 80% identity or more with a sequence disclosed herein, for example, 85%, 90%, 95% identity or more, e.g., 96%, 97%, 98%, or 99% identity or more, in certain instances 100% identity to the sequence. Publicly available databases such as Uniprot, Cis-BP, Transfac, and GrassiusX can be consulted to identify which nucleotides can be varied and which should be conserved in designing DTSs for use in the inventions of the present disclosure.

A given DTS may include 2 or more different TFBSs (i.e., TFBSs that differ from each other by nucleotide sequence and are therefore distinct), such as 2, 3, 4, 5, 6, 7, 8, 9 or 10 different TFBSs. A given DTS may include 1 or more copies of the same TFBS, such as 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, copies of the same TFBS, where in some instances the number of identical TFBS copies does not exceed 10.

The length of a given DTS may vary depending on the number of TFBSs comprised by it, the lengths of the TFBS sequence to which each nuclear targeting factor binds, and the number nucleotides between TFBSs (the spacer sequence). Where two or more TFBSs (either the same or different) are present in a given DTS, the distance between any two TFBSs may vary as desired, ranging in some instances from 5 to 100 bp, such as 10 to 75 bp, including 15 to 50 bp. For example, the TFBSs may be separated from one another by 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,16, 17, 18, 19, or 20 or more nts, for example, 0-5 nts, 6-10 nts, 11-15 nts, 16-20 nts, 21-25 nts, 26-30 nts, 31-35 nts, 36-40 nts, in some instances 40 – 50 nts. In some instances, the length of a given DTS ranges from 10 to 500 nt, such as 100 to 300 nt and including 150 to 200 nt.

A given NTDNA may include a single DTS or a plurality of DTSs, as desired. Where a given NTDNA includes a plurality of DTSs, the disparate DTSs may be the same or different. As such, a given

NTDNA may include 2 or more different DTSs (i.e., DTSs that differ from each other by nucleotide sequence and are therefore distinct).

The DTS may be incorporated in the NTDNA in any of a variety of positions. For example, the DTS may be placed 5' of the promoter, 3' of the promoter and 5' of the expression cassette, within an intron of the expression cassette, or 3' of the expression cassette. In some instances, the DTSs is placed 5' of the promoter. In some embodiments, the DTS is placed 3' of the expression cassette.

Nuclear targeting factors. As summarized above, in some embodiments, in addition to the DNA and the agent that modulates BAF, a cell may also be contacted with a nuclear targeting factor or an mRNA that encodes the nuclear targeting factor to which the DTS binds to mediate nuclear entry of the DNA. Without wishing to be bound by theory, it is believed that the exogenously provided nuclear targeting factor is translated from the mRNA and translocates from the cytosol to the nucleus, and in doing so brings the NTDNA associated with it (via the binding interaction between the DTS and the nuclear targeting factor) into the nucleus. In embodiments, the nuclear targeting factor will comprise a nuclear localization sequence (NLS) or a fragment thereof, as described herein or known in the art. In some instances, the NLS will be native, or endogenous, to the nuclear targeting factor. In other instances, a DNA binding protein will be engineered to comprise an NLS so as to create the nuclear targeting factor. In such instances, the NLS will be heterologous to the protein. If engineered into the NTF, the heterologous NLS may be engineered to be anywhere within the NTF. In some instances, the heterologous NLS is engineered to be within the center of the protein, i.e. the NLS is flanked on both N and C termini by 20 or more amino acids endogenous to the DNA binding protein. In some instances, the heterologous NLS is engineered to be at the N terminus. In some instances, the heterologous NLS is engineered to be at the C terminus. In some instances, the heterologous NLS is engineered to be at both the N terminus and the C terminus. When engineered to reside at a terminus, the NLS is typically engineered to be within 0-20 amino acids of the terminus, in some instances, within 0-10 amino acids of the terminus, in certain instances within 0-5 amino acids of the terminus, in some such cases, at the terminus. Typically, the NLS is engineered to be at a site that is distinct from the protein domain that mediates binding to the DNA. In some instances, the NLS may be flanked by one or more amino acids, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids. In some instances, the nuclear targeting factor will comprise one NLS, in other instances multiple NLSs. In some instances, in which multiple NLSs are employed, the same NLS is employed multiple times. In other instances, in which multiple NLSs are employed, different NLSs are used. Often, when multiple NLSs are employed, they will be separated by a linker, e.g., a 6xK linker. Exemplary NLSs and their origins are provided in Table 5.

Table 5. Exemplary NLS sequences

NLS origin	Sequence
SV40 Large T antigen SMALL	PKKKRKVED (SEQ ID NO:365)
SV40 Large T antigen	KKKSSSDDEATADSQHSTPPKKKRKVEDPKDFPSELLS (SEQ ID NO:366)
Optimized SV40 Large T antigen	PSSDDEATADSQHAAPPKKKRKVEDPKDFPSELLS (SEQ ID NO:367)
Heterogeneous nuclear RNP A1.1 M9 domain	GNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGG (SEQ ID NO:368)
Heterogeneous nuclear RNP A1.2 M9 domain	FGNYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGG (SEQ ID NO:369)
Heterogeneous nuclear RNP A1.3 M9 domain	FGNYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGGSSSSSSSYGSGRRF(SEQ ID NO:370)
HIV viral protein 1 (VP1)	DTWTGVEALIRILQQLLFIHFRIGCRHSRIGIIQQRRTRNGA (SEQ ID NO:371)
HIV viral protein 2 (VP2)	GDTWAGVEAIIRILQQLLFIHFRIGCRHSRIGVTRQRRARNGASRS (SEQ ID NO:372)
HIV viral protein 3 (VP3)	EQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQHIYETYGDTWAGVEAIIRILQQLLFIHFRIGCRHSRIGVTRQRRARNGASRS (SEQ ID NO:373)
Influenza virus A nuclear protein (INF-A)	ASQGTKRSYEQMETDGERQ (SEQ ID NO:374)
SV40 short fused to Influenza A NLS	PKKKRKVASQGTKRSYEQMETDGERQ(SEQ ID NO:375)
Influenza A NLS fused to SV40 short	MASQGTKRSYEQMETDGERQPKKKRKV (SEQ ID NO:376)
Cellular myc transcription factor (c-myc)	PAAKRVKLD (SEQ ID NO:377)
c-myc fused to INF-A	PAAKRVKLDASQGTKRSYEQMETDGERQ(SEQID NO:378)
c-myc NLS fused to Influenza A NLS, fused to SV40 long version	PAAKRVKLDASQGTKRSYEQMETDGERQKKKSSSDDEATADSQHSTPPKKKRKVEDPKDFPSELLS (SEQ ID NO:379)
c-myc NLS fused to Influenza A NLS, fused to a second c-myc NLS	PAAKRVKLDASQGTKRSYEQMETDGERQPAAKRVKLD (SEQ ID NO:380)
Influenza A NLS fused to c-myc NLS	MASQGTKRSYEQMETDGERQPAAKRVKLD (SEQ ID NO:381)
Biparticle nucleus localization signal	KRTADGSEFESPKKKRKV (SEQ ID NO:382)

In some embodiments, the nuclear targeting factor comprises a nuclear export signal. In some embodiments, the NES is NMD3 ribosome export adaptor. In some such embodiments, the NES comprises a sequence having 85% identity or more to NELALKLAGLDINKT(SEQ ID NO:383). In some such embodiments, the NES comprises a sequence having 85% identity or more to EHVNKMNSDRVPDVVLIKKSYDRTKRQRRRNWKLKELA(SEQ ID NO:384).

In some embodiments, the nuclear targeting factor comprises a tag, a linker or other sequence that may be used, e.g., to preserve the structure of an appended NLS or NES, to purify protein in recombinant protein production, to visualize the protein in cells, etc. Examples of tags or linkers include but are not

limited to CS3, HIS, Flag, GGS, RIGID, SpyTag, Strep tags, etc. Exemplary linkers are provided in Table 6.

Table 6. Exemplary linkers.

NAME	AMINO ACID
S1 (StrepTag 1X)	WSHPQFEK (SEQ ID NO:385)
S2 (StrepTag 2X)	WSHPQFEKWSHPQFEK (SEQ ID NO:386)
S3 (StrepTag 3X)	WSHPQFEKWSHPQFEKWSHPQFEK(SEQ ID NO:387)
GGS 1X	GGS (SEQ ID NO:388)
GGS 2X	GGSGGS (SEQ ID NO:389)
GGS 3X	GGSGGSGGS (SEQ ID NO:390)
GGGGS 1X	GGGGS (SEQ ID NO:391)
GGGGS 2X	GGGGSGGGGS (SEQ ID NO:392)
GGGGS 3X	GGGGSGGGGSGGGGS (SEQ ID NO:393)
GGGGS 4X	GGGGSGGGGSGGGGSGGGGS (SEQ ID NO:394)
QSGSGSTATGSG linker	QSGSGSTATGSG(SEQ ID NO:395)
Rigid	GTPTPTPTPTG (SEQ ID NO:396)
M_Flex	SGGGSGGSGSS (SEQ ID NO:397)
W_Flex	GSAGSAAGSGEF (SEQ ID NO:398)

Any method for making mRNA may be employed to generate the subject mRNA encoding the NTF. As one example, mRNA may be chemically synthesized using solid-phase methods. As another example, mRNA may be synthesized from a DNA template by an *in vitro* transcription (IVT) reaction, in which a DNA template comprising a promoter sequence for an RNA polymerase operably linked to a sequence encoding the target mRNA (in this instance, the nuclear targeting factor) is contacted with the RNA polymerase, which binds the template at the promoter region and starts the RNA synthesis. As yet another example, the mRNA may be synthesized *in vivo* and purified from the biological source. The mRNA may comprise naturally occurring ribonucleotides and/or chemically modified ribonucleotides. Typically, some chemically modified nucleotides may be included to reduce the immunogenicity of the mRNA, for example, N^1-methylpseudouridine, 2-thiouridine (s^2U), 5-methylcytidine (m^5C), N^6-methyladenosine (m^6A), 2'-*O*-methyluridine (Um), 2'-*O*-methylcytidine (Cm), 2'-*O*-methyladenosine (Am), and 2'-*O*-methylguanosine (Gm), N^4-Acetyl-cytidine 5-triphosphate (AC4C). The mRNA may comprise a 5' cap or analog thereof, e.g., an m^7GpppG cap, anti-reverse cap analog (ARCA), a two-headed cap, an S cap, a 2S cap, and the like. The mRNA will typically comprise a 5' untranslated region (UTR). The mRNA will typically comprise a 3' UTR, e.g., as described in Table 7. The mRNA may comprise a tail modification, e.g., a ribose modified adenosine, 8-azaadenosine, cordycepin, and the like.

Table 7. Exemplary 3' UTRs

UTR	Sequence
T7 promoter	TAATACGACTCACTATAGG (SEQ ID NO: 399)
T7 promoter v2	TAATACGACTCACTATAAG (SEQ ID NO: 400)
T7 promoter m6Am	TAATACGACTCACTATAAGTCA (SEQ ID NO: 401)
T7 enhancer	GGGATAAT (SEQ ID NO: 402)
T7 modified enhancer	GGGATAATAG (SEQ ID NO: 403)
Kozak	GCCACCATG (SEQ ID NO: 404)
modified Kozak	GCCGCCACCATG (SEQ ID NO: 405)
m6a modification site	GGACA (SEQ ID NO: 406)
m6a modification site 2	TGACT (SEQ ID NO: 407)
m6a modification site 3	AGACCAGACA (SEQ ID NO: 408)
m6a modification site 4	GGACTGGACTGGACT (SEQ ID NO: 409)
CAP binding site	TAATGTGAGTTAGCTCACTCAT (SEQ ID NO: 410)
2' *O* methyl site	AGATCGAGGA (SEQ ID NO: 411)
5-methylcytidine sute	CGGGGAACTCTCCAATTCCCC (SEQ ID NO: 412)
b-globin	ACTCTTCTGGTCCCCACAGACTCAGAGAGAA (SEQ ID NO: 413)
b-globin v2	AAGAGAGACTCAGACACCCCTGGTCTTCTCAACTCTTCTGG TCCCCACAGACTCAGAGAGAA (SEQ ID NO: 414)
Translation Initiator of Short 5' UTR (TISU) sequence	AAGATCG (SEQ ID NO: 415)
Synthetic UTR	CCGTGACGCAGGGATAAAGCAATCAAGGTTGTGTCAGTGA CGTGCTATAA (SEQ ID NO: 416)
Complement Factor 3 (C3)	GAGCCAGATAAAAAGCCAGCTCCAGCAGGCGCTGCTCAC TCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCC TGCACTGTCCC (SEQ ID NO: 417)
CYP2e1	ACAGGATTGTCTCCCGGGCTGGCAGCAGGGCCCCAGC (SEQ ID NO: 418)
Sv40 PolyA	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA TTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC TCATCAATGTATCTTA (SEQ ID NO: 419)
Amino terminal enhancer of *Split*	CTGGTACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGTACCCCG AGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCACTC ACCACCTCTGCTAGTTCCAGACACCTCC (SEQ ID NO: 420)
mtRNR1	CAAGCACGCAGCAATGCAGCTCAAAACCTTAGCCTAGCACACCCCCACGGGAA ACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAA CCCCAGGGTTGGTCAATTTCGTCCAGCCACACC (SEQ ID NO: 421)
a-globin	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGG CGGCA (SEQ ID NO: 422)
UTR30 (Sample PJ et al. Nature Biotechnol 2019	CCGTAAATCGGATAGGGCGGAAACGATGGACACGAAGTCTATAGTCGGGT(SE Q ID NO: 423)
UTR31 (Sample PJ et al., *supra*)	CTCACGGGATGACCCACATTGCGATGATGTAGTTCCGGGCCGCCAAATAG (SEQ ID NO: 424)
UTR32 (Sample PJ et al., *supra*)	AAGTACTGCAGGCCCGAATCACTCGGCCAAATGGCCTGCATAATGCTGAT (SEQ ID NO: 425)

UTR33 (Sample PJ et al., *supra*)	TAAACTTCCAAGTCGCCTCGCAGGCTCCCACCAATGCCGCCGCCTAAGGG (SEQ ID NO: 426)
UTR34 (Sample PJ et al., *supra*)	CTCAGGAGGGTGGTAAGCCACCGCCCACGCGTGGGGCACTCGTGTATTCT (SEQ ID NO: 427)
UTR35 (Sample PJ et al., *supra*)	AGCGGACATCAGCCATTGAATACATGAAACAAATCGCAAAAGCAATAGCT (SEQ ID NO: 428)
UTR36 (Sample PJ et al., *supra*)	CGTCGCTCCGGGGAGCTGTGTGTACGAGTAAAAAATTTAGATCACCAAAA (SEQ ID NO: 429)
UTR37 (Sample PJ et al., *supra*)	ACAATCGCCAGGGACAGATAGGAAGTTATAGACAGCAGGCCGACACAAGT (SEQ ID NO: 430)
UTR38 (Sample PJ et al., *supra*)	GAAAAATCTATAGCAGAAGTCAGCGGTAGACGCACGGCATAGCATCCAAC (SEQ ID NO: 431)
UTR39 (Sample PJ et al., *supra*)	GGGCGCTCGAGCAGGTTCAGAAGGAGATCAAAAACCCCCAAGGATCAAAC (SEQ ID NO: 432)

For example, when the DTS comprises a binding sequence for a Tet Repressor (TetR) protein, e.g., a TCCCTATCAGTGATAGAGA(SEQ ID NO: 433) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a TetR protein. In some embodiments, the TetR protein has been engineered to comprise an NLS. In some embodiments, the NLS is an NLS listed in Table 5. In some certain embodiments, the NLS is the NLS of SV40 Large T antigen or an optimized variant thereof, the NLS of the influenza A nuclear protein INF-A, the NLS of c-myc, or a fusion of the c-myc NLS to the influenza A nuclear protein INF-A NLS. In some embodiments, the NLS is located proximal to the N terminus. In other embodiments, the NLS is located proximal to the C-terminus. In some embodiments, the NLS is fused to the TetR protein with a linker.

In some embodiments, the TetR protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to the wild type *Escherichia coli* transcriptional regulator TetR protein sequence:

MSRLDKSKVINSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEMLDRHHTHFCPLEGESWQD FLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFLCQQGFSLENALYALSAVGHFTLGCVLEDQEHQVA KEERETPTTDSMPPLLRQAIELFDHQGAEPAFLFGLELIICGLEKQLKCESGS(SEQ ID NO: 434).

In some embodiments, the TetR protein is encoded by a polynucleotide comprising a sequence having a sequence identity of 80% or more to a sequence listed in Table 8, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity across the length of a sequence in Table 8. In some embodiments, the polynucleotide sequence has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human cells or mouse cells.

Table 8. RNA polynucleotide sequences for expression of TetR protein

TetR variant	TetR RNA Sequence (5′ to 3′)
Wild Type	AUGAGCAGACUGGACAAGAGCAAAGUGAUCAACAGCGCCCUGGAACUGCUGAACGAAGUGGGCAU CGAGGGCCUGACCACAAGAAAGCUGGCCCAGAAGCUGGGCGUCGAGCAGCCUACACUGUACUGGC ACGUGAAGAACAAGCGGGCCCUGCUGGAUGCCCUGGCCAUUGAGAUGCUGGACCGGCACCACACA CACUUUUGCCCUCUGGAAGGCGAGAGCUGGCAGGACUUCCUGAGAAACAACGCCAAGAGCUUCAG AUGCGCCCUGCUGAGCCAUAGAGAUGGCGCCAAAGUGCACCUGGGCACCAGACCUACAGAGAAGC AGUACGAGACACUGGAAAACCAGCUGGCCUUCCUGUGCCAGCAGGGAUUCAGCCUGGAAAACGCC CUGUAUGCCCUGUCUGCCGUGGGCCACUUUACACUGGGAUGCGUGCUGGAAGAUCAAGAGCACCA GGUGGCCAAAGAGGAAAGAGAGACACCCACCACCGACAGCAUGCCUCCACUGCUGAGACAGGCCA UCCGAGCUGUUCGAUCACCAAGGCGCCGAACCUGCCUUUCUGUUCGGACUGGAACUGAUCAUCUGC GGCCUCGAAAAGCAGCUGAAGUGCGAGAGCGGCUCC (SEQ ID NO: 435)
Human (Jcat)	AUGAGCAGGCUGGACAAGAGCAAGGUGAUCAACAGCGCCCUGGAGCUGCUGAACGAGGUGGGCAU CGAGGGCCUGACCACCAGGAAGCUGGCCCAGAAGCUGGGCGUGGAGCAGCCCACCCUGUACUGGC ACCUGAAGAACAAGAGGGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGACAGGCACCACACC CACUUCUGCCCCCUGGAGGGCGAGAGCUGGCAGGACUUCCUGAGGAACAACGCCAAGAGCUUCAG GUGCGCCCUGCUGAGCCACAGGGACGGCGCCAAGGUGCACCUGGGCACCAGGCCCACCGAGAAGC AGUACGAGACgCUGGAGAACCAGCUGGCCUUCCUGUGCCAGCAGGGCUUCAGCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGCCACUUCACCCUGGGCUGCGUGCUGGAGGACCAGGAGCACCA GGUGGCCAAGGAGGAGAGGGAGACgCCCACCACCGACAGCAUGCCCCCCCUGCUGAGGCAGGCCA UCGAGCUGUUCGACCACCAGGGCGCCGAGCCCGCCUUCCUGUUCGGCCUGGAGCUGAUCAUCUGC GGCCUGGAGAAGCAGCUGAAGUGCGAGAGCGGCAGC(SEQ ID NO: 436)
Human (Vector Builder)	AUGAGCAGGCUGGAUAAGAGCAAGGUGAUUAACUCCGCCCUGGAGCUGCUGAACGAAGUGGGAAU CGAGGGCCUGACUACAAGGAAGUUGGCCCAGAAGCUGGGCGUGGAGCAGCCUACCCUGUACUGGC ACGUGAAGAACAAGAGAGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGACAGGCAUCACACU CACUUUUGCCCUCUCGAGGGCGAGAGCUGGCAGGACUUCCUGAGGAACAACGCCAAGUCAUUUCG GUGUGCCCUGCUGAGCCAUCGAGAUGGGGCCAAGGUGCACCUGGGAACCAGACCCACAGAGAAGC AGUACGAGACACUGGAAAACCAGCUGGCUUUCCUGUGCCAGCAGGGCUUCUCCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGCCAUUUCACCCUGGGUUGUGUGCUGGAGGAUCAGGAGCACCA GGUGGCCAAGGAGGAAAGGGAAACCCCUACCACCGACUCAAUGCCACCUCUGCUGCGCCAGGCCA UCGAGCUGUUUGACCACCAGGGAGCCGAGCCCGCCUUCCUGUUCGGGCUGGAGCUGAUUAUUUGC GGCCUGGAGAAGCAGCUGAAGUGUGAGAGUGGCAGC (SEQ ID NO: 437)
Human (Novopro)	AUGUCCCGACUGGACAAAUCCAAGGUUAUCAAUAGCGCCCUGGAGCUGCUCAAUGAAGUGGGUAU CGAGGGCCUGACGACACGCAAGCUUGCACAGAAGUUGGGCGUCGAGCAGCCAACUCUGUAUUGGC AUGUGAAGAACAAGAGGGCCCUGCUCGACGCCCUCGCGAUUGAAAUGCUGGACAGACACCACACA CAUUUUUGUCCUUUGGAGGGCGAGUCCUGGCAGGACUUUCUUAGAAAUAACGCCAAGUCCUUCCG CUGCGCACUGCUCAGCCACCGAGAUGGGGCUAAAGUGCAUCUUGGAACACGCCCCACCGAGAAAC AGUACGAAACUCUCGAGAAUCAGCUUGCAUUUCUGUGUCAGCAAGGGUUCAGUCUGGAAAACGCC CUGUACGCUCUGUCCGCCGUGGGGCAUUUUACACUGGGAUGUGUGCUGGAAGACCAAGAGCAUCA GGUCGCCAAGGAAGAAAGAGAGACUCCGACAACCGACUCCAUGCCUCCCCUUGCUCCGGCAGGCCA UUGAGCUGUUCGACCACCAGGGCGCCGAACCUGCCUUCUUGUUCGGUCUUGAGCUGAUCAUCUGC GGACUGGAAAAGCAACUGAAAUGUGAGAGCGGGUCC (SEQ ID NO: 438)
Human (IDT)	AUGAGUCGACUUGAUAAGAGCAAGGUUAUAAAUUCCGCUUUGGAACUGUUGAAUGAGGUGGGCAU CGAGGGUCUUACGACCCGCAAAUUGGCACAAAAGUUGGGACUUGAACAGCCAACGCUUUAUUGGC ACGUUAAAAAUAAGCGGGCUCUGCUGGACGCUCUGGCUAUAGAAAUGCUUGACCGGCAUCAUACC CAUUUUUGUCCUCUGGAAGGCGAAUCUUGGCAGGACUUUCUCAGGAAUAACGCCAAAUCUUUUAG AUGCGCUCUUUUGUCACACCGCGAUGGAGCCAAGGUCCAUCUUGGCACGCGGCCAACGGAGAAAC AAUAUGAGACCCUCGAAAACCAGCUCGCCUUCCUCUGUCAACAGGGAUUCAGUCUUGAGAAUGCC CUGUACGCUCUUUCUGCGGUCGGGCAUUUUACACUCGGAUGUGUUUUGGAGGACCAAGAACAUCA AGUAGCGAAAGAGGAACGGGAAACUCCCACGACCGAUUCAAUGCCGCCACUCCUCCGACAAGCUA UUGAAUUGUUCGAUCAUCAAGGGGCUGAACCUGCCUUCCUCUUCGGUCUGGAACUUAUAAUCUGC GGGCUCGAGAAACAGUUGAAAUGCGAGUCAGGAUCA (SEQ ID NO: 439)

Human (Genewiz)	AUGAGCAGACUGGACAAGAGCAAGGUGAUCAACAGCGCCCUGGAGCUGCUGAACGAGGUGGGCAU CGAGGGCCUGACCACAAGAAAGCUGGCUCAGAAGCUGGGCGUGGAGCAGCCCACCCUGUACUGGC ACGUGAAGAACAAGAGAGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGACAGACACCACACC CACUUCUGCCCCCUGGAGGGCGAGAGCUGGCAAGACUUCCUGAGAAACAACGCCAAGAGCUUCAG AUGCGCCCUGCUGAGCCACAGAGACGGCGCCAAGGUGCACCUGGGCACAAGACCCACCGAGAAGC AGUACGAAACCCUGGAGAAUCAGCUGGCCUUCCUGUGUCAGCAAGGCUUCAGCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGCCACUUCACCCUGGGCUGCGUGCUGGAGGACCAAGAGCACCA AGUGGCCAAGGAGGAGAGAGAAACCCCCACCACCGACAGCAUGCCCCCCCUGCUGAGACAAGCCA UCGAGCUGUUCGACCACCAAGGCGCCGAGCCCGCCUUCCUGUUCGGCCUGGAGCUGAUCAUCUGC GGCCUGGAGAAGCAGCUGAAGUGCGAAAGCGGCAGC (SEQ ID NO: 440)
Human (Genesmart)	AUGAGCAGACUGGAUAAGAGCAAGGUCAUCAACAGCGCCCUCGAACUGCUGAACGAGGUGGGCAU CGAAGGCCUUACAACAAGAAAGCUGGCUCAGAAGCUGGGCGUGGAACAACCUACCCUGUAUUGGC AUGUGAAGAACAAGCGGGCUCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGACAGACACCACACC CACUUUUGCCCACUGGAGGGCGAGAGCUGGCAGGACUUCCUGAGAAAUAACGCCAAGUCCUUCCG GUGCGCCCUCCUGAGCCACCGGGACGGCGCCAAGGUGCACCUGGGAACCAGACCUACCGAGAAGC AGUACGAGACACUGGAAAACCAGCUGGCCUUCCUGUGCCAGCAGGGCUUCAGCCUGGAGAAUGCC CUGUACGCCCUGUCUGCCGUGGGCCACUUCACCCUGGGUUGUGUGCUGGAAGAUCAGGAGCACCA AGUGGCCAAAGAGGAAAGAGAGACACCUACAACCGACAGCAUGCCCCCCCUGCUGCGGCAGGCCA UUGAGCUGUUCGACCACCAGGGAGCCGAGCCUGCUUUUCUGUUCGGCCUGGAGCUGAUCAUCUGC GGCCUGGAAAAACAGCUGAAGUGUGAAUCUGGCUCC (SEQ ID NO: 441)
Human (Geneart)	AUGAGCAGACUGGACAAGAGCAAAGUGAUCAACAGCGCCCUGGAACUGCUGAACGAAGUGGGCAU CGAGGGCCUGACCACAAGAAAGCUGGCCCAGAAGCUGGGCGUCGAGCAGCCUACACUGUACUGGC ACGUGAAGAACAAGCGGGCCCUGCUGGAUGCCCUGGCCAUUGAGAUGCUGGACCGGCACCACACA CACUUUUGCCCUCUGGAAGGCGAGAGCUGGCAGGACUUCCUGAGAAACAACGCCAAGAGCUUCAG AUGCGCCCUGCUGAGCCAUAGAGAUGGCGCCAAAGUGCACCUGGGCACCAGACCUACAGAGAAGC AGUACGAGACACUGGAAAACCAGCUGGCCUUCCUGUGCCAGCAGGGAUUCAGCCUGGAAAACGCC CUGUAUGCCCUGUCUGCCGUGGGCCACUUUACACUGGGAUGCGUGCUGGAAGAUCAAGAGCACCA GGUGGCCAAAGAGGAAAGAGAGACACCCACCACCGACAGCAUGCCUCCACUGCUGAGACAGGCCA UCGAGCUGUUCGAUCACCAAGGCGCCGAACCUGCCUUUCUGUUCGGACUGGAACUGAUCAUCUGC GGCCUCGAGAAGCAGCUGAAGUGCGAGUCUGGAUCU (SEQ ID NO: 442)
Mouse (Jcat)	AUGAGCAGGCUGGACAAGAGCAAGGUGAUCAACAGCGCCCUGGAGCUGCUGAACGAGGUGGGCAU CGAGGGCCUGACCACCAGGAAGCUGGCCCAGAAGCUGGGCGUGGAGCAGCCCACCCUGUACUGGC ACGUGAAGAACAAGAGGGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGACAGGCACCACACC CACUUCUGCCCCCUGGAGGGCGAGAGCUGGCAGGACUUCCUGAGGAACAACGCCAAGAGCUUCAG GUGCGCCCUGCUGAGCCACAGGGACGGCGCCAAGGUGCACCUGGGCACCAGGCCCACCGAGAAGC AGUACGAGACCCUGGAGAACCAGCUGGCCUUCCUGUGCCAGCAGGGCUUCAGCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGCCACUUCACCCUGGGCUGCGUGCUGGAGGACCAGGAGCACCA GGUGGCCAAGGAGGAGAGGGAGACCCCCACCACCGACAGCAUGCCCCCCCUGCUGAGGCAGGCCA UCGAGCUGUUCGACCACCAGGGCGCCGAGCCCGCCUUCCUGUUCGGCCUGGAGCUGAUCAUCUGC GGCCUGGAGAAGCAGCUGAAGUGCGAGAGCGGCAGC (SEQ ID NO: 443)
Mouse (Vector Builder)	AUGUCCCGCCUGGACAAGUCUAAGGUGAUCAAUUCCGCCCUGGAACUGCUGAAUGAGGUGGGAAU CGAGGGCCUGACCACAAGGAAGCUGGCCCAGAAGCUGGGCGUGGAGCAGCCUACACUGUACUGGC AUGUGAAGAACAAGCGGGCCCUCCUGGACGCCCUGGCCAUCGAAAUGCUGGACCGCCACCACACC CACUUCUGUCCCCUGGAGGGCGAGAGCUGGCAGGAUUUCCUGCGGAAUAAUGCCAAGAGCUUCAG AUGUGCUCUGCUGAGCCACAGAGAUGGCGCAAAGGUGCACCUGGGCACCAGACCCACCGAGAAGC AGUAUGAGACUCUGGAGAACCAGCUGGCAUUCCUGUGCCAACAGGGGUUUUCCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGACACUUCACCCUGGGCUGUGUGCUGGAGGACCAGGAGCACCA GGUGGCCAAGGAGGAAAGAGAGACACCAACUACCGACUCCAUGCCUCCUCUCCUCCGGCAGGCCA UUGAACUGUUUGAUCAUCAGGGCGCCGAGCCAGCCUUUCUGUUUGGCCUGGAACUGAUCAUCUGU GGCCUGGAGAAGCAGCUGAAGUGCGAAUCUGGCAGC (SEQ ID NO: 444)
Mouse (Novopro)	AUGUCUAGGUUGGAUAAAUCAAAGGUCAUCAAUUCUGCUCUGGAGCUCCUCAAUGAAGUGGGAAU CGAAGGUCUCACUACUCGGAAGCUGGCACAAAAACUCGGGGUGGAACAACCUACCCUUUAUUGGC ACGUGAAAAACAAAAGGGCCUUGCUGGACGCUCUGGCCAUAGAGAUGUUGGACCGACAUCACACG CAUUUCUGCCCUCUGGAGGGGAAUCCUGGCAAGAUUUCCUGCGCAAUAACGCAAAGUCUUUCCG AUGCGCUCUGCUCUCUCAUAGGGACGGCGCUAAGGUCCACCUGGGAACCAGACCUACGGAGAAGC AGUAUGAAACCCUGGAGAACCAGCUGGCAUUCCUCUGCCAGCAGGGCUUCAGUCUCGAGAACGCA

	CUGUACGCACUGUCUGCUGUGGGUCACUUCACCCUGGGCUGUGUGUUGGAGGACCAGGAGCACCA GGUGGCCAAGGAAGAGCGCGAAACCCCUACAACGGAUUCUAUGCCCCCCUUGCUUAGGCAGGCCA UCGAGCUCUUCGAUCACCAGGGGGCCGAACCUGCCUUUUUGUUUGGCCUGGAACUCAUUAUAUGU GGGCUGGAGAAGCAGCUGAAAUGUGAGAGUGGAAGC (SEQ ID NO: 445)
Mouse (IDT)	AUGUCACGAUUGGACAAAAGUAAGGUUAUAAACUCUGCUCUCGAACUUCUGAACGAAGUUGGGAU CGAAGGUUUGACCACCCGCAAGUUGGCUCAAAAACUUGGUGUGGAGCAGCCUACCCUGUAUUGGC ACGUAAAGAACAAGAGAGCCUUGCUGGACGCCCUUGCAAUCGAGAUGCUGGACAGACACCACACC CAUUUUUGUCCCUUGGAAGGGAGAGUUGGCAAGAUUUUCUGCGGAACAAUGCUAAGAGCUUCCG CUGUGCCCUGCUUUCCCAUCGAGAUGGCGCAAAAGUGCACUUGGGAACUCGGCCCACUGAGAAGC AAUACGAGACUCUUGAGAACCAACUCGCAUUUCUUUGUCAACAAGGAUUUUCCCUGGAGAAUGCU CUGUAUGCUUUGAGUGCAGUAGGCCAUUUCACCCUCGGCUGCGUACUGGAGGACCAAGAGCAUCA AGUUGCCAAAGAAGAACGGGAAACCCCAACUACAGAUAGCAUGCCUCCUUUGCUCCGACAGGCCA UAGAACUCUUUGAUCAUCAAGGUGCAGAGCCAGCAUUCCUCUUCGGCCUUGAGCUCAUUAUUUGC GGACUUGAGAAGCAACUGAAGUGUGAAAGCGGAAGU (SEQ ID NO: 446)
Mouse (Genewiz)	AUGUCUAGGCUGGACAAGAGCAAGCUGAUCAACAGCGCCCUGGAGCUGCUGAACGAGGUGGGCAU CGAGGGCCUGACCACAAGGAAGCUGGCUCAGAAGCUGGGCGUGGAGCAGCCUACCCUGUACUGGC ACGUGAAGAACAAGAGGGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGAUAGGCACCACACC CACUUCUGCCCUCUGGAGGGCGAGAGCUGGCAAGACUUCCUGAGGAACAACGCCAAGAGCUUUAG GUGCGCCCUGCUGAGCCAUAGGGACGGCGCCAAGGUGCACCUGGGCACAAGGCCUACCGAGAAGC AGUACGAAACCCUGGAGAAUCAGCUGGCCUUCCUGUGUCAGCAAGGCUUCAGCCUGGAGAACGCC CUGUACGCCCUGAGCGCCGUGGGCCACUUCACCCUGGGCUGCGUGCUGGAGGACCAAGAGCACCA AGUGGCCAAGGAGGAGAGGGAAACCCCUACCACCGACAGCAUGCCUCCUCUGCUGAGGCAAGCCA UCGAGCUGUUCGACCACCAAGGCGCCGAGCCUGCCUUCCUGUUCGGCCUGGAGCUGAUCAUCUGC GGCCUGGAGAAGCAGCUGAAGUGCGAAAGCGGCAG (SEQ ID NO: 447)
Mouse (Genesmart)	AUGUCUCGUCUCGACAAGUCAAAAGUCAUCAAUAGUGCUCUGGAGCUGCUUAAUGAAGUAGGGAU AGAAGGUCUGACCACCAGGAAGCUGGCCCAGAAGCUGGGAGUUGAACAACCUACCCUCUACUGGC ACGUGAAAAACAAGCGCGCACUUUUAGACGCCCUGGCGAUUGAGAUGCUGGAUCGGCACCAUACA CAUUUCUGCCCCCUAGAAGGAGAAUCCUGGCAGGACUUUCUCCGAAACAACGCCAAAUCCUUCCG CUGUGCACUGCUGAGCCAUCGAGAUGGAGCGAAAGUGCACCUGGGGACGCGGCCUACUGAGAAAC AGUACGAAACUCUAGAGAACCAGUUGGCCUUCCUCUGCCAGCAGGGAUUCAGUUUAGAGAAUGCA CUCUAUGCUCUCUCUGCAGUGGGCCACUUCACAUUGGGCUGCGUUUUGGAAGACCAGGAGCAUCA GGUGGCCAAGGAGGAAAGAGAGACACCUACUACAGAUAGCAUGCCUCCACUGCUGAGACAAGCCA UUGAACUGUUUGACCACCAGGGUGCUGAGCCAGCCUUUUUGUUUGGUUUAGAACUUAUCAUCUGU GGCCUAGAGAAGCAGCUGAAAUGUGAGUCUGGCAGC (SEQ ID NO: 448)
Mouse (Geneart)	AUGAGCAGACUGGACAAGAGCAAAGUGAUCAACAGCGCCCUGGAACUGCUGAACGAAGUGGGCAU CGAGGGCCUGACCACAAGAAAGCUGGCUCAGAAGCUGGGCGUCGAGCAGCCUACACUGUACUGGC ACGUGAAGAACAAGAGAGCCCUGCUGGACGCCCUGGCCAUCGAGAUGCUGGAUAGACACCACACA CACUUUUGCCCUCUGGAAGGCGAGAGCUGGCAGGACUUCCUGAGAAACAACGCCAAGAGCUUCAG AUGCGCCCUGCUGAGCCAUAGAGAUGGCGCCAAAGUGCACCUGGGCACCAGACCUACAGAGAAGC AGUACGAGACACUGGAAAACCAGCUGGCCUUCCUGUGCCAGCAGGGAUUCUCUCUGGAAAACGCC CUGUACGCCCUGAGCGCUGUGGGCCACUUUACACUGGGAUGCGUGCUGGAAGAUCAAGAGCACCA GGUGGCCAAAGAGGAAAGAGAGACACCCACCACCGACAGCAUGCCUCCACUGCUGAGACAGGCUA UCGAGCUGUUCGACCACCAAGGCGCCGAGCCUGCUUUUCUGUUCGGACUGGAACUGAUCAUCUGC GGCCUCGAGAAGCAGCUGAAGUGCGAGUCUGGAUCU (SEQ ID NO: 449)

As another example, when the DTS comprises a binding sequence for a ZF protein, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a ZF protein. For example, when the DTS comprises the sequence AAACTGCAAAAG or a variant thereof, the cell may also be contacted with an mRNA encoding the ZF protein ZF-CCR5. In some embodiments, the ZF protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In

some embodiments, the ZF protein comprises an NLS at both the N terminus and the C terminus. In some embodiments, the NLS is fused to the ZF protein with a linker.

In some embodiments, the ZF protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g. is 100% identical to the ZF-CCR5 fusion protein: MRPFQCRICMRNFSRSDNLSVHIRTHTGEKPFACDICGRKFAQKINLQVHTKIHTGEKPFQCRICMRNFSRSDVLSE HIRTHTGEKPFACDICGRKFAQRNHRTTHTKIHLR(SEQ ID NO: 450), for example, the variant MAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSRSDNLSVHIRTHTGEKPFACDICGRKFAQKINLQVHTKIHTGEK PFQCRICMRNFSRSDVLSEHIRTHTGEKPFACDICGRKFAQRNHRTTHTKIHLR(SEQ ID NO: 451).

As another example, when the DTS comprises a binding sequence for a TALE protein, e.g., TTCATTACACCTGCAGCT(SEQ ID NO: 452), ATAAACCCCCTCCAA(SEQ ID NO: 453), or TCGAGTTTACTCCCTATCAGTGATAGAGAACG(SEQ ID NO: 454), or variants thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a TALE protein. In some embodiments, the TALE protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the TALE protein comprises an NLS at both the N terminus and the C terminus. In some embodiments, the NLS is fused to the TALE protein with a linker.

In some embodiments, the TALE protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to TMVLAQNRKKSLHCFEGLFTAVVTSNSDHLVRSQ(SEQ ID NO: 455), for example MLTPDQVVAIASN_GGKQA_ETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVA IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPV_CQDHGLTPDQVVAIASNIGGKQ ALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLL PVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIA SHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQAL ETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG(SEQ ID NO: 456). Other exemplary TALEs and methods for engineering them may be found in Li et a. 2011 (*supra*) and Kim et al. 2011 (*supra*).

As another example, when the DTS comprises the binding sequence for a GAL4 protein, for example CGG-N_{11}-CCG or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a GAL4 protein. In some embodiments, the GAL4 protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In

some embodiments, the GAL4 protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the GAL4 protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5. In some certain embodiments, the NLS is the NLS of SV40 Large T antigen or an optimized variant thereof, the NLS of the influenza A nuclear protein INF-A, the NLS of c-myc, or a fusion of the c-myc NLS to the influenza A nuclear protein INF-A NLS.

In some embodiments, the GAL4 protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to wild type *Saccharomyces cerevisiae* GAL4 protein MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESRLERLE(SEQ ID NO: 457). In some cases, the GAL4 protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to: MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLTEVESRLERLEQLFLLIFPREDL DMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSEESSNKGQRQLTVS(SEQ ID NO: 458).

In some embodiments, the GAL4 protein is encoded by a polynucleotide comprises a sequence having a sequence identity of 80% or more to a sequence listed in Table 9, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity across the length of a sequence in Table 9. In some embodiments, the GAL4 polynucleotide has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

Table 9. RNA polynucleotide sequences encoding GAL4 protein.

GAL4 variant	GAL4 mRNA Sequence (5′ to 3′)
Wild Type	AUGAAACUGCUGUCAUCCAUCGAGCAAGCUUGUGACAUAUGCCGGCUGAAGAAAUUGAAAUGUAGC AAGGAGAAGCCAAAAUGUGCGAAAUGCCUCAAGAAUAAUUGGGAAUGCAGAUAUUCCCCUAAGACU AAGCGUAGCCCCUUGACCCGGGCCCACCUUACUGAGGUCGAGAGCAGAUUAGAGCGCCUCGAGCAA UUGUUCCUGCUCAUCUUCCCCAGAGAGGAUCUGGAUAUGAUCUUAAAGAUGGACAGCCUUCAAGAC AUUAAGGCUCUGCUCACUGGGCUGUUCGUGCAGGACAACGUCAACAAGGACGCUGUGACGGACCGC CUGGCUUCCGUAGAAACGGACAUGCCCCUGACUCUGAGGCAACACCGCAUUUCUGCUACGAGUUCU UCCGAGGAAUCCUCAAAUAAGGGACAGCGUCAACUGACCGUCUCU (SEQ ID NO: 459)
Human (Jcat)	AUGAAGCUGCUGAGCAGCAUCGAGCAGGCCUGCGACAUCUGCAGGCUGAAGAAGCUGAAGUGCAGC AAGGAGAAGCCCAAGUGCGCCAAGUGCCUGAAGAACAACUGGGAGUGCAGGUACAGCCCCAAGACC AAGAGGAGCCCCCUGACCAGGGCCCACCUGACCGAGGUGGAGAGCAGGCUGGAGAGGCUGGAGCAG CUGUUCCUGCUGAUCUUCCCCAGGGAGGACCUGGACAUGAUCCUGAAGAUGGACAGCCUGCAGGAC AUCAAGGCCCUGCUGACCGGCCUGUUCGUGCAGGACAACGUGAACAAGGACGCCGUGACCGACAGG CUGGCCAGCGUGGAGACgGACAUGCCCCUGACCCUGAGGCAGCACAGGAUCAGCGCCACCAGCAGC AGCGAGGAGAGCAGCAACAAGGGCCAGAGGCAGCUGACCGUGAGC (SEQ ID NO: 460)

Human (Vector Builder)	AUGAAGCUGCUGUCUAGCAUCGAGCAGGCGUGUGACAUCUGCAGACUGAAGAAACUGAAAUGCAGU AAAGAGAAGCCAAAGUGUGCUAAGUGUCUGAAAAACAAUUGGGAGUGUAGAUACUCCCCCAAGACA AAGAGAAGCCCACUGACAAGGGCCCACCUGACCGAGGUGGAGAGUAGACUGGAGAGACUGGAGCAG CUGUUUCUGCUGAUCUUCCCAAGGGAGGACCUGGACAUGAUUCUGAAGAUGGACUCCCUGCAGGAU AUCAAAGCCCUGCUGACCGGACUGUUCGUGCAGGACAACGUGAACAAGGACGCCGUGACCGACAGG CUGGCCAGCGUGGAGACgGACAUGCCCCUGACACUGAGACAGCACCGGAUCAGCGCCACCAGCUCU UCCGAGGAGAGUUCCAACAAGGGACAGCGGCAGUUGACAGUGUCC (SEQ ID NO: 461)
Human (Genewiz)	AUGAAGCUGCUGAGCAGCAUCGAGCAAGCCUGCGACAUCUGUAGGCUGAAGAAGCUGAAGUGCAGC AAGGAGAAGCCUAAGUGCGCCAAGUGCCUGAAGAACAACUGGGAGUGUAGGUACAGCCCUAAGACG AAGAGGAGCCCUCUGACAAGGGCCCACCUGACCGAGGUGGAGUCUAGGCUGGAGAGGCUGGAGCAG CUGUUCCUGCUGAUCUUCCCUAGGGAGGACCUGGACAUGAUCCUGAAGAUGGACAGCCUGCAAGAC AUCAAGGCCCUGCUGACCGGCCUGUUCGUGCAAGACAACGUGAACAAGGACGCCGUGACCGAUAGG CUGGCUAGCGUGGAAACCGACAUGCCUCUGACCCUGAGGCAGCAUAGGAUCAGCGCCACAAGCAGC AGCGAGGAGAGCAGCAACAAGGGACAGAGGCAGCUGACCGUGAGC (SEQ ID NO: 462)
Human (Genesmart)	AUGAAGUUACUAUCUUCGAUUGAGCAAGCCUGUGACAUCUCCAGACUUAAGAAACUGAAAUGCAGC AAAGAGAAGCCCAAAUGUGCUAAGUGCCUGAAGAACAACUGGGAGUGUCGUUACAGCCCAAAGACC AAGCGCUCACCUCUUACUCGAGCUCACCUGACAGAGGUGGAAAGUCGCCUGGAGCGGUUAGAGCAG CUGUUCCUGCUCAUCUUCCCCAGAGAAGACCUGGACAUGAUCCUCAAGAUGGACAGCCUGCAGGAU AUUAAAGCCCUGCUCACAGGCCUGUUCGUGCAGGACAAUGUCAAUAAAGAUGCUGUGACUGAUAGA CUUGCAUCUGUAGAAACGGACAUGCCUCUGACUCUGCGGCAGCAUCGGAUAUCUGCCACCUCCUCC UCAGAAGAAAGUAGCAACAAGGGCCAGAGGCAGUUGACCGUUUCU (SEQ ID NO: 463)

As another example, when the DTS comprises a binding sequence for an Arc protein, e.g., RYRVTAGANNNNNNCTABYRY(SEQ ID NO: 464) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding an Arc protein. In some embodiments, the Arc protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the Arc protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the Arc protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5. In some certain embodiments, the NLS is the NLS of SV40 Large T antigen or an optimized variant thereof, the NLS of the influenza A nuclear protein INF-A, the NLS of c-myc, or a fusion of the c-myc NLS to the influenza A nuclear protein INF-A NLS.

In some embodiments, the Arc protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to the sequence encoding the *Salmonella* phage Arc-like repressor: MKGMSKMPQFNLRWPREVLDLVRKVAEENGRSVNSEIYQRVMESFKKEGRIGA(SEQ ID NO: 465). For example, the Arc protein may be the st11 variant MKGMSKMPQFNLRWPREVLDLVRKVAEENGRSVNSEIYQRVMASFAKEGRIAAKNQH E(SEQ ID NO: 466) or a protein having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more to the st11 variant.

In some embodiments, the Arc protein is encoded by a polynucleotide comprising a sequence having a sequence identity of 80% or more, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity to the Salmonella phage Arc-like repressor wild type mRNA sequence:

AUGAAGGGCAUGAGCAAGAUGCCCCAGUUCAACCUGCGCUGGCCCCGCGAGGUGCUGGACCUGGUGCGCAAGGUGGC CGAGGAGAACGGCCGCAGCGUGAACAGCGAGAUCUACCAGCGCGUGAUGGAGAGCUUCAAGAAGGAGGGCCGCAUCG GCGCC(SEQ ID NO: 467). In some embodiments, the Arc protein is encoded by a polynucleotide comprises a sequence having a sequence identity of 80% or more, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity to the sequence encoding the st11 Arc variant:

AUGAAGGGCAUGAGCAAGAUGCCCCAGUUCAACCUGCGCUGGCCCCGCGAGGUGCUGGACCUGGUGCGCAAGGUGGC CGAGGAGAACGGCCGCAGCGUGAACAGCGAGAUCUACCAGCGCGUGAUGGCCAGCUUCGCCAAGGAGGGCCGCAUCG CCGCCAAGAACCAGCACGAG(SEQ ID NO: 468). In some instances, the polynucleotide has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human cells.

As another example, when the DTS comprises the binding sequence for a Mnt protein, e.g., GGNCCACNGTGGNCC(SEQ ID NO: 469) or a variant thereof, for example ATAGGTCCACGGTGGACCATA(SEQ ID NO: 470), in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a Mnt protein. In some embodiments, the Mnt protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the Mnt protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the Mnt protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5. In some certain embodiments, the NLS is the NLS of SV40 Large T antigen or an optimized variant thereof, the NLS of the influenza A nuclear protein INF-A, the NLS of c-myc, or a fusion of the c-myc NLS to the influenza A nuclear protein INF-A NLS.

In some embodiments, the Mnt protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to native wild-type Mnt:

MARDDPHFNFRMPMEVREKLKFRAEANGRSMNSELLQIVQDALSKPSPVTGYRNDAERLADEQSELVKKMVFDTLKD LYKKTT(SEQ ID NO: 471).

In some embodiments, the Mnt protein is encoded by a polynucleotide comprising a sequence having a sequence identity of 80% or more, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity across

the sequence:

AUGGCCCGCGACGACCCCCACUUCAACUUCCGCAUGCCCAUGGAGGUGCGCGAGAAGCUGAAGUUCCGCGCCGAGGC
CAACGGCCGCAGCAUGAACAGCGAGCUGCUGCAGAUCGUGCAGGACGCCCUGAGCAAGCCCAGCCCCGUGACCGGCU
ACCGCAACGACGCCGAGCGCCUGGCCGACGAGCAGAGCGAGCUGGUGAAGAAGAUGGUGUUCGACACCCUGAAGGAC
CUGUACAAGAAGACCACC(SEQ ID NO: 472). In some embodiments, the Mnt polynucleotide has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

As another example, when the DTS comprises the binding sequence for a PurR protein, e.g., ACGCAAACGTTTTCGT(SEQ ID NO: 473) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a PurR protein. In some embodiments, the PurR protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the PurR protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the PurR protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5.

In some embodiments, the PurR protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g. is 100% identical to wild-type *Escherichia coli* HTH-type transcriptional repressor PurR:

MATIKDVAKRANVSTTTVSHVINKTRFVAEETRNAVWAAIKELHYSPSAVARSLKVNHTKSIGLLATSSEAAYFAEI
IEAVEKNCFQKGYTLILGNAWNNLEKQRAYLSMMAQKRVDGLLVMCSEYPEPLLAMLEEYRHIPMVVMDWGEAKADF
TDAVIDNAFEGGYMAGRYLIERGHREIGVIPGPLERNTGAGRLAGFMKAMEEAMIKVPESWIVQGDFEPESGYRAMQ
QILSQPHRPTAVFCGGDIMAMGALCAADEMGLRVPQDVSLIGYDNVRNARYFTPALTTIHQPKDSLGETAFNMLLDR
IVNKREEPQSIEVHPRLIERRSVADGPFRDYRR(SEQ ID NO: 474).

In some embodiments, the PurR protein is encoded by a polynucleotide comprising a sequence having a sequence identity of 80% or more, for example 85%, 90%, or 95% or more, in some instances 96%, 97%, 98%, or 99% identity, in some cases 100% sequence identity across the length of the sequence:

AUGGCCACCAUCAAGGACGUGGCCAAGCGCGCCAACGUGAGCACCACCACCGUGAGCCACGUGAUCAACAAGACCCG
CUUCGUGGCCGAGGAGACGCGCAACGCCGUGUGGGCCGCCAUCAAGGAGCUGCACUACAGCCCCAGCGCCGUGGCCC
GCAGCCUGAAGGUGAACCACACCAAGAGCAUCGGCCUGCUGGCCACCAGCAGCGAGGCCGCCUACUUCGCCGAGAUC
AUCGAGGCCGUGGAGAAGAACUGCUUCCAGAAGGGCUACACCCUGAUCCUGGGCAACGCCUGGAACAACCUGGAGAA
GCAGCGCGCCUACCUGAGCAUGAUGGCCCAGAAGCGCGUGGACGGCCUGCUGGUGAUGUGCAGCGAGUACCCCGAGC
CCCUGCUGGCCAUGCUGGAGGAGUACCGCCACAUCCCCAUGGUGGUGAUGGACUGGGGCGAGGCCAAGGCCGACUUC
ACCGACGCCGUGAUCGACAACGCCUUCGAGGGCGGCUACAUGGCCGGCCGCUACCUGAUCGAGCGCGGCCACCGCGA

GAUCGGCGUGAUCCCCGGCCCCCUGGAGCGCAACACCGGCGCCGGCCGCCUGGCCGGCUUCAUGAAGGCCAUGGAGG
AGGCCAUGAUCAAGGUGCCCGAGAGCUGGAUCGUGCAGGGCGACUUCGAGCCCGAGAGCGGCUACCGCGCCAUGCAG
CAGAUCCUGAGCCAGCCCCACCGCCCCACCGCCGUGUUCUGCGGCGGCGACAUCAUGGCCAUGGGCGCCCUGUGCGC
CGCCGACGAGAUGGGCCUGCGCGUGCCCCAGGACGUGAGCCUGAUCGGCUACGACAACGUGCGCAACGCCCGCUACU
UCACCCCCGCCCUGACCACCAUCCACCAGCCCAAGGACAGCCUGGGCGAGACGGCCUUCAACAUGCUGCUGGACCGC
AUCGUGAACAAGCGCGAGGAGCCCCAGAGCAUCGAGGUGCACCCCCGCCUGAUCGAGCGCCGCAGCGUGGCCGACGG
CCCCUUCCGCGACUACCGCCGC(SEQ ID NO: 475). In some embodiments, the polynucleotide encoding the PurR protein has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

As another example, when the DTS comprises the binding sequence for a Bac434 protein, e.g., ACAAGAAAGTTTGT(SEQ ID NO: 476), ACAAGATACATTGT(SEQ ID NO: 477), or ACAAGAAAAACTGT(SEQ ID NO: 478) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a Bac434 protein. In some embodiments, the Bac434 protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the Bac434 protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the Bac434 protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5.

In some embodiments, the Bac434 protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g., is 100% identical to the wild type *Escherichia coli* Phage 434 Repressor protein CI: MSISSRVKSKRIQLGLNQAELAQKVGTTQQSIEQLENGKTKRPRFLPELASALGVSVDWLLNGTSDSNVRFVGHVEP KGKYPLISMVRAGSWCEA(SEQ ID NO: 479).

In some embodiments, the Bac434 protein is encoded by a polynucleotide comprising a sequence having a sequences identity of 80% or more, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity to: AUGAGCAUCAGCAGCCGCGUGAAGAGCAAGCGCAUCCAGCUGGGCCUGAACCAGGCCGAGCUGGCCCAGAAGGUGGG CACCACCCAGCAGAGCAUCGAGCAGCUGGAGAACGGCAAGACCAAGCGCCCCCGCUUCCUGCCCGAGCUGGCCAGCG CCCUGGGCGUGAGCGUGGACUGGCUGCUGAACGGCACCAGCGACAGCAACGUGCGCUUCGUGGGCCACGUGGAGCCC AAGGGCAAGUACCCCCUGAUCAGCAUGGUGCGCGCCGGCAGCUGGUGCGAGGCC(SEQ ID NO: 480). In some embodiments, the polynucleotide sequence encoding the Bac434 protein has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

As another example, when the DTS comprises the binding sequence for a GCN4 protein, e.g., AGTGACTCATT or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a GCN4 protein. In some

embodiments, the GCN4 protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the GCN4 protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the GCN4 protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5.

In some embodiments, the GCN4 protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g. is 100% identical to the wild type sequence for *Saccharomyces cerevisiae* GCN4:

MSEYQPSLFALNPMGFSPLDGSKSTNENVSASTSTAKPMVGQLIFDKFIKTEEDPIIKQDTPSNLDFDFALPQTATA
PDAKTVLPIPELDDAVVESFFSSSTDSTPMFEYENLEDNSKEWTSLFDNDIPVTTDDVSLADKAIESTEEVSLVPSN
LEVSTTSFLPTPVLEDAKLTQTRKVKKPNSVVKKSHHVGKDDESRLDHLGVVAYNRKQRSIPLSPIVPESSDPAALK
RARNTEAARRSRARKLQRMKQLEDKVEELLSKNYHLENEVARLKKLVGER(SEQ ID NO: 481).

In some embodiments, the GCN4 protein is encoded by a polynucleotide comprising a sequence having a sequences identity of 80% or more, for example 85%, 90%, or 95% sequence identity or more, in some instances 96%, 97%, 98%, or 99% sequence identity, in some cases 100% sequence identity to

AUGUCCGAAUAUCAACCAUCUUUAUUUGCAUUGAAUCCAAUGGGGUUUUCCCCACUUGAUGGAUCCAAAUCCACAAA
UGAAAAUGUUUCUGCAUCAACAAGUACUGCAAAGCCAAUGGUCGGGCAAUUGAUCUUUGAUAAAUUUAUUAAGACAG
AAGAAGAUCCGAUUAUAAAACAAGAUACUCCCAGUAAUUUGGAUUUUGAUUUUGCGUUGCCUCAAACAGCUACAGCU
CCAGAUGCUAAAACAGUCUUACCAAUUCCAGAACUCGAUGAUGCUGUCGUCGAAAGUUUCUUUUCAAGUUCCACGGA
UAGUACACCAAUGUUUGAGUAUGAAAAUUUGGAAGAUAAUUCAAAAGAAUGGACUUCUCUCUUUGACAAUGACAUUC
CAGUCACUACUGAUGAUGUUUCCCUUGCAGAUAAAGCUAUUGAAUCUACAGAAGAAGUAUCUUUAGUCCCAUCCAAU
CUUGAAGUUUCCACUACGAGUUUUCUCCCUACGCCUGUACUUGAAGAUGCAAAAUUGACGCAAACGCGGAAAGUCAA
CAAACCAAAUUCUGUUGUCAAGAAAUCUCAUCAUGUUGGAAAAGAUGAUGAAUCUAGGCUCGAUCAUCUUGCAGUUG
UCGCUUAUAAUAGGAAACAAAGAUCUAUACCACUCUCCCCUAUAGUCCCAGAAAGUAGUGAUCCUGCUGCUUUAAAA
CGGGCUAGGAAUACAGAAGCGGCUCGAAGGUCCAGGGCCCGGAAAUUACAACGGAUGAAACAACUCGAAGAUAAAGU
AGAAGAAUUGCUUUCUAAGAAUUAUCAUCUUGAGAAUGAAGUUGCACGGUUGAAGAAACUCGUUGGGGAACGG(SEQ
ID NO: 482). In some embodiments, the GCN polynucleotide has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

As another example, when the DTS comprises the binding sequence for a Lac Repressor protein ("LacR"), e.g., TTGTTATCCGCTCACAA(SEQ ID NO: 483) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a LacR protein. In some embodiments, the LacR protein has been engineered to comprise an NLS. In some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the LacR protein is engineered to comprise an NLS at both the N-

terminus and the C-terminus. In some embodiments, the NLS is fused to the LacR protein with a linker. In some embodiments, the NLS is selected from those listed in Table 5.

In some embodiments, the LacR protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g. is 100% identical to the wild type sequence for the *Escherichia coli* DNA-binding transcriptional repressor LacI:

MKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYIPNRVAQQLAGKQSLLIGVATSSLALHAPS
QIVAAIKSRADQLGASVVVSMVERSGVEACKAAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQT
PINSIIFSHEDGTRLGVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQTMQ
MLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPLTTIKQDFRLLGQTSVDRLLQLS
QGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSLMQLARQVSRLESGQ(SEQ ID NO: 484).

In some embodiments, the LacR protein is encoded by a polynucleotide comprising a sequence having a sequence identity of 80% or more, for example 85%, 90%, or 95% or more, in some instances 96%, 97%, 98%, or 99% identity, in some cases 100% sequence identity across the length of the sequence:

AUGAAACCAGUCACACUCUAUGAUGUAGCGGAAUAUGCUGGGGUAUCCUAUCAAACAGUUAGUAGAGUUGUUAAUCA
AGCUUCACAUGUUUCCGCCAAAACGAGGGAAAAGGUUGAAGCAGCUAUGGCUGAACUUAAUUAUAUACCAAAUCGAG
UAGCGCAACAACUCGCAGGGAAACAAUCUUUACUCAUUGGAGUCGCAACAUCCUCACUCGCUCUCCAUGCUCCUUCU
CAAAUAGUCGCUGCUAUUAAAAGUAGAGCUGAUCAAUUAGGGGCAUCUGUCCUAGUCUCUAUGGUCGAAAGAUCUGG
UGUAGAAGCUUGUAAAGCUGCAGUCCAUAAUCUCCUCGCUCAACGGGUCAGCGGGCUUAUAAUUAAUUAUCCUCUUG
AUGAUCAAGAUGCAAUAGCUGUCGAAGCAGCUUGUACUAAUGUCCCUGCGUUGUUUCUUGAUGUCUCAGAUCAAACA
CCAAUAAAUUCAAUUAUAUUUAGUCAUGAAGAUGGUACAAGGCUCGGGGUUGAACAUCUCGUAGCACUCGGUCAUCA
ACAAAUAGCACUCCUUGCUGGGCCACUCUCUUCCGUUUCUGCACGGCUCAGGCUCGCAGGAUGGCAUAAAUAUUUGA
CACGUAAUCAAAUUCAACCUAUAGCUGAAAGGGAAGGUGAUUGGUCCGCUAUGAGUGGGUUUCAACAAACAAUGCAA
AUGUUGAAUGAAGGGAUAGUUCCAACUGCAAUGCUCGUUGCUAAUGAUCAAAUGGCGCUUGGUGCGAUGCGAGCUAU
UACAGAAUCCGGACUUCGGGUUGGUGCUGAUAUAUCCGUAGUCGGAUAUGAUGAUACAGAAGAUUCCUCAUGUUAUA
UUCCACCUCUUACAACUAUUAAACAAGAUUUUAGGCUUCUCGGGCAAACAUCCGUCGAUAGACUUUUACAACUCUCC
CAAGGACAAGCAGUAAAAGGAAAUCAAUUAUUGCCGGUUAGUCUUGUCAAACGUAAAACAACGCUUGCGCCAAAUAC
ACAAACGGCUUCUCCUAGAGCACUUGCUGAUAGUCUUAUGCAACUCGCACGACAAGUCUCUCGAUUGGAAUCUGGAC
AA(SEQ ID NO: 485). In some embodiments, the polynucleotide encoding the LacR protein has been codon optimized for expression in a particular host cell, e.g., codon optimized for expression in human or mouse cells.

As another example, when the DTS comprises the binding sequence for a I-SceI D44A protein, e.g., TAGGGATAACAGGGTAAT(SEQ ID NO: 486) or a variant thereof, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding a I-SceI D44A protein. In some embodiments, the I-SceI protein has been engineered to comprise an NLS. In

some embodiments, the NLS is proximal to the N terminus. In other embodiments, the NLS is proximal to the C-terminus. In some embodiments, the I-SceI protein is engineered to comprise an NLS at both the N-terminus and the C-terminus. In some embodiments, the NLS is fused to the I-SceI protein with a linker.

In some embodiments, the I-SceI protein comprises a sequence having a sequence identity of 90% or more, in some instances 92%, 93%, 94%, 95% or more, in certain instances 96%, 97%, 98%, 99% identity or more, e.g. is 100% identical to

MKNIKKNQVMNLGPNSKLLKEYKSQLIELNIEQFEAGIGLILGAAYIRSRDEGKTYCMQFEWKNKAYMDHVCLLYDQ
WVLSPPHKKERVNHLGNLVITWGAQTFKHQAFNKLANLFIVNNKKTIPNNLVENYLTPMSLAYWFMDDGGKWDYNKN
STNKSIVLNTQSFTFEEVEYLVKGLRNKFQLNCYVKINKNKPIIYIDSMSYLIFYNLIKPYLIPQMMYKLPNTISSE
TFLK(SEQ ID NO: 487), e.g. a sequence having 85% identity more to:

MPKKKRKVPKKHAAPPKKKRKVEDPRFMYPYDVPDYAGMKNIKKNQVMNLGPNSKLLKEYKSQLIELNIEQFEAGIG
LILGAAYIRSRDEGKTYCMQFEWKNKAYMDHVCLLYDQWVLSPPHKKERVNHLGNLVITWGAQTFKHQAFNKLANLF
IVNNKKTIPNNLVENYLTPMSLAYWFMDDGGKWDYNKNSTNKSIVLNTQSFTFEEVEYLVKGLRNKFQLNCYVKINK
NKPIIYIDSMSYLIFYNLIKPYLIPQMMYKLPNTISSETFLK(SEQ ID NO: 488).

As another example, when the TFBS is the binding sequence for a gene editing system, e.g. the guide RNA for a Cas nuclease, the zinc finger of a zinc finger nuclease, the TALE of a TALEN, e.g. as described below and as known in the art, in addition to contacting the cell with the DNA and the BAF modulator, the cell may also be contacted with an mRNA encoding the Cas nuclease, the ZF nuclease, or the TALEN. In some certain embodiments, the nuclease of the gene editing system is catalytically active. Put another way, the nuclease is able to modify DNA, for example to nick it, to create a double-stranded break, to replace one or more nucleotides within it. In other certain embodiments, the gene editing system is catalytically attenuated, e.g., its activity is reduced 50% or more, in some instances 60%, 70%, 80%, 90% or more, in some cases 95% or more, e.g., the catalytic activity has been abrogated.

Delivery Compositions and Methods of Delivery

Where desired, the DNA and/or BAF modulator and other agents as disclosed herein may be present in a delivery composition that includes the DNA and/or BAF modulator and a cytosolic delivery vehicle component, e.g., a delivery vehicle component that mediates entry of the DNA and/or BAF modulator into the cytosol from an extracellular location. Typically, the cytosolic delivery vehicle is a non-naturally occurring vehicle, i.e. a non-viral delivery vehicle. The DNA and BAF modulator may be present in the same cytosolic delivery vehicle or present in different cytosolic delivery vehicles, i.e., a first cytosolic delivery vehicle for the DNA and a second cytosolic delivery vehicle for the BAF modulator. When present in separate cytosolic deliver vehicles, the DNA and BAF modulator may be

contacted with the cell simultaneously or sequentially, where when the DNA and BAF modulator are contacted sequentially with the cell, the DNA may be contacted with the cell prior to the BAF modulator or the BAF modulator may be contacted with the cell prior to the DNA. In some aspects, the methods provided herein comprise delivering a DNA and BAF modulator to a target cell.

Methods of delivery of nucleic acids can include lipofection, nucleofection, microinjection, biolistics, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355 and lipofection reagents are sold commercially (e.g., Transfectam™ and Lipofectin™). Delivery can be to cells (e.g., *in vitro* or *ex vivo* administration) or target tissues (e.g., in vivo administration), as indicated above.

Various techniques and methods are known in the art for delivering nucleic acids to cells. For example, a DNA and/or BAF modulator can be delivered to a cell by conjugating the nucleic acid with a ligand that is internalized by the cell. For example, the ligand can bind a receptor on the cell surface and internalized via endocytosis. The ligand can be covalently linked to a nucleotide in the nucleic acid. Exemplary conjugates for delivering nucleic acids into a cell are described, example, in WO2015/006740, WO2014/025805, WO2012/037254, WO2009/082606, WO2009/073809, WO2009/018332, WO2006/112872, WO2004/090108, WO2004/091515 and WO2017/177326.

DNAs and/or BAF modulators, e.g., as described herein, can also be delivered to a cell by transfection. Useful transfection methods include, but are not limited to, lipid-mediated transfection, cationic polymer-mediated transfection, or calcium phosphate precipitation. Transfection reagents are well known in the art and include, but are not limited to, TurboFect Transfection Reagent (Thermo Fisher Scientific), Pro-Ject Reagent (Thermo Fisher Scientific), TRANSPASS™ P Protein Transfection Reagent (New England Biolabs), CHARIOT™ Protein Delivery Reagent (Active Motif), PROTEOJUICE™ Protein Transfection Reagent (EMD Millipore), 293fectin, LIPOFECTAMINE™ 2000, LIPOFECTAMINE™ 3000 (Thermo Fisher Scientific), LIPOFECTAMINE™ (Thermo Fisher Scientific), LIPOFECTIN™ (Thermo Fisher Scientific), DMRIE-C, CELLFECTIN™ (Thermo Fisher Scientific), OLIGOFECTAMINE™ (Thermo Fisher Scientific), LIPOFECTACE™, FUGENE™ (Roche, Basel, Switzerland), FUGENE™ HD (Roche), TRANSFECTAM™ (Transfectam, Promega, Madison, Wis.), TFX-10™ (Promega), TFX-20™ (Promega), TFX-50™ (Promega), TRANSFECTIN™ (BioRad, Hercules, Calif.), SILENTFECT™ (Bio-Rad), Effectene™ (Qiagen, Valencia, Calif.), DC-chol (Avanti Polar Lipids), GENEPORTER™ (Gene Therapy Systems, San Diego, Calif.), DHARMAFECT 1™ (Dharmacon, Lafayette, Colo.), DHARMAFECT 2™ (Dharmacon), DHARMAFECT 3™ (Dharmacon), DHARMAFECT 4™ (Dharmacon), ESCORT™ III (Sigma, St. Louis, Mo.), and ESCORT™ IV (Sigma

Chemical Co.). Nucleic acids, such as DNA, can also be delivered to a cell via microfluidics methods, e.g., such as those known to those of skill in the art.

Methods of non-viral delivery of nucleic acids in vivo or ex vivo include electroporation, lipofection (see, U.S. Pat. Nos. 5,049,386; 4,946,787 and commercially available reagents such as Transfectam™ and Lipofectin™), microinjection, biolistics, LNPs, virosomes, liposomes (see, e.g., Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787), immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, and agent-enhanced uptake of DNA. Sonoporation using, e.g., the Sonitron 2000 system (Rich-Mar) can also be used for delivery of nucleic acids.

For example, DNAs and/or BAF modulators can be formulated into lipid nanoparticles (LNPs), lipidoids, liposomes, lipoplexes, or core-shell nanoparticles. Delivery reagents such as liposomes, nanocapsules, microparticles, microspheres, lipid nanoparticles, vesicles, and the like, can be used for the introduction of the compositions of the present disclosure into suitable host cells. In particular, the nucleic acids can be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, a nanoparticle, a gold particle, or the like. Such formulations can be preferred for the introduction of pharmaceutically acceptable formulations of the nucleic acids disclosed herein.

Various delivery methods known in the art or modifications thereof can be used to deliver DNAs and/or BAF modulators as described herein *in vitro* or *in vivo*. For example, in some embodiments, DNAs and/or BAF modulators are delivered by making transient penetration in cell membrane by mechanical, electrical, ultrasonic, hydrodynamic, or laser-based energy so that DNA entrance into the targeted cells is facilitated. For example, a DNA and/or BAF modulator can be delivered by transiently disrupting cell membrane by squeezing the cell through a size-restricted channel or by other means known in the art. In some cases, a DNA and/or BAF modulator alone is directly injected as naked DNA into skin, thymus, cardiac muscle, skeletal muscle, or liver cells.

In some cases, a DNA and/or BAF modulator is delivered by gene gun. Gold or tungsten spherical particles (1-3 μm diameter) coated with DNA can be accelerated to high speed by pressurized gas to penetrate into target tissue cells.

In some embodiments, electroporation is used to deliver DNA and/or an BAF modulator to a target cell. Electroporation causes temporary destabilization of the cell membrane target cell tissue by insertion of a pair of electrodes into the tissue so that DNA molecules in the surrounding media of the

destabilized membrane would be able to penetrate into cytoplasm and nucleoplasm of the cell. Electroporation has been used *in vivo* for many types of tissues, such as skin, lung, and muscle.

In some cases, DNA and/or an BAF modulator is delivered by hydrodynamic injection, which is a simple and highly efficient method for direct intracellular delivery of any water-soluble compounds and particles into internal organs and skeletal muscle in an entire limb.

In some cases, DNAs and/or BAF modulators are delivered by ultrasound by making nanoscopic pores in membrane to facilitate intracellular delivery of DNA particles into cells of internal organs or tumors, so the size and concentration of plasmid DNA have great role in efficiency of the system. In some cases, DNAs and/or BAF modulators are delivered by magnetofection by using magnetic fields to concentrate particles containing nucleic acid into the target cells.

In some cases, chemical delivery systems can be used, for example, by using nanomeric complexes, which include compaction of negatively charged nucleic acid by polycationic nanomeric particles, belonging to cationic liposome/micelle or cationic polymers. Cationic lipids used for the delivery method includes, but not limited to monovalent cationic lipids, polyvalent cationic lipids, guanidine containing compounds, cholesterol derivative compounds, cationic polymers, (e.g., poly(ethylenimine), poly-L-lysine, protamine, other cationic polymers), and lipid-polymer hybrid.

DNAs and/or BAF modulators, e.g., as described herein, can also be administered directly to an organism for transduction of cells *in vivo*. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells including, but not limited to, injection, infusion, topical application and electroporation. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

Compositions comprising a DNA and/or BAF modulator, e.g., as described herein, and a cytosolic delivery vehicle are specifically contemplated herein. In some embodiments, a DNA and/or BAF modulator is formulated with a lipid delivery system, for example, LNPs as described herein. In some embodiments, such compositions are administered by any route desired by a skilled practitioner. The compositions may be administered to a subject by different routes including orally, parenterally, subretinally, intravitreally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal administration, intrapleurally, intravenous, intra-arterial, intraperitoneal, intracranially, subcutaneous, intramuscular, intranasal, intrathecal, and intraarticular or combinations thereof. For veterinary use, the composition may be administered as a suitably acceptable formulation in accordance with normal veterinary practice. The veterinarian may readily determine the dosing regimen and route of administration that is most appropriate for a particular animal. The compositions may be administered by

traditional syringes, needleless injection devices, "microprojectile bombardment gene guns", or other physical methods such as electroporation ("EP"), hydrodynamic methods or ultrasound.

Microparticle/Nanoparticles. In some embodiments, a DNA and/or BAF modulator as described herein is delivered by a nanoparticle. One example of a nanoparticle that finds use is a lipid nanoparticle (LNP). Generally, LNPs of the present disclosure may be composed of nucleic acid (e.g., DNA and/or BAF modulator) molecules, one or more ionizable or cationic lipids (or salts thereof), one or more non-ionic or neutral lipids (e.g., a phospholipid), a molecule that prevents aggregation (e.g., PEG or a PEG-lipid conjugate), and optionally a sterol (e.g., cholesterol). For example, lipid nanoparticles may comprise an ionizable amino lipid (e.g., heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate, DLin-MC3-DMA, a phosphatidylcholine (1,2-distearoyl-sn-glycero-3-phosphocholine, DSPC), cholesterol and a coat lipid (polyethylene glycol-dimyristolglycerol, PEG-DMG), for example as disclosed by Tam et al. (2013). Advances in Lipid Nanoparticles for siRNA delivery. Pharmaceuticals 5(3): 498-507.

In some embodiments, a lipid nanoparticle has a mean diameter between about 10 and about 1000 nm. In some embodiments, a lipid nanoparticle has a diameter that is less than 300 nm. In some embodiments, a lipid nanoparticle has a diameter between about 10 and about 300 nm. In some embodiments, a lipid nanoparticle has a diameter that is less than 200 nm. In some embodiments, a lipid nanoparticle has a diameter between about 25 and about 200 nm. In some embodiments, a lipid nanoparticle preparation (e.g., composition comprising a plurality of lipid nanoparticles) has a size distribution in which the mean size (e.g., diameter) is about 70 nm to about 200 nm, and more typically the mean size is about 100 nm or less.

In some aspects, the disclosure provides for a lipid nanoparticle comprising a DNA and/or BAF modulator as described herein and an ionizable lipid. The ionizable lipid is typically employed to condense the nucleic acid cargo, e.g., DNA at low pH and to drive membrane association and fusogenicity. Generally, ionizable lipids are lipids comprising at least one amino group that is positively charged or becomes protonated under acidic conditions, for example at pH of 6.5 or lower. Ionizable lipids are also referred to as cationic lipids herein. Exemplary ionizable lipids are described in International PCT patent applications PCT/US2023/079923 and PCT/US2023/76457 and International PCT publications WO2015/095340, WO2015/199952, WO2018/011633, WO2017/049245, WO2015/061467, WO2012/040184, WO2012/000104, WO2015/074085, WO2016/081029, WO2017/004143, WO2017/075531, WO2017/117528, WO2011/022460, WO2013/148541, WO2013/116126, WO2011/153120, WO2012/044638, WO2012/054365, WO2011/090965, WO2013/016058, WO2012/162210, WO2008/042973, WO2010/129709, WO2010/144740, WO2012/099755, WO2013/049328, WO2013/086322, WO2013/086373, WO2011/071860, WO2009/132131, WO2010/048536, WO2010/088537, WO2010/054401, WO2010/054406,

WO2010/054405, WO2010/054384, WO2012/016184, WO2009/086558, WO2010/042877, WO2011/000106, WO2011/000107, WO2005/120152, WO2011/141705, WO2013/126803, WO2006/007712, WO2011/038160, WO2005/121348, WO2011/066651, WO2009/127060, WO2011/141704, WO2006/069782, WO2012/031043, WO2013/006825, WO2013/033563, WO2013/089151, WO2017/099823, WO2015/095346, and WO2013/086354, and US patent publications US2016/0311759, US2015/0376115, US2016/0151284, US2017/0210697, US2015/0140070, US2013/0178541, US2013/0303587, US2015/0141678, US2015/0239926, US2016/0376224, US2017/0119904, US2012/0149894, US2015/0057373, US2013/0090372, US2013/0274523, US2013/0274504, US2013/0274504, US2009/0023673, US2012/0128760, US2010/0324120, US2014/0200257, US2015/0203446, US2018/0005363, US2014/0308304, US2013/0338210, US2012/0101148, US2012/0027796, US2012/0058144, US2013/0323269, US2011/0117125, US2011/0256175, US2012/0202871, US2011/0076335, US2006/0083780, US2013/0123338, US2015/0064242, US2006/0051405, US2013/0065939, US2006/0008910, US2003/0022649, US2010/0130588, U52013/0116307, US2010/0062967, US2013/0202684, US2014/0141070, US2014/0255472, US2014/0039032, US2018/0028664, U52016/0317458, and US2013/0195920.

Various LNP formulations known in the art can be used to deliver DNAs and/or BAF modulators as described herein. For example, various LNP formulations and delivery methods using lipid nanoparticles are described in U.S. Pat. Nos. 9,404,127; 9,006,417; 9,518,272; and US Patent Application No. 63/415,229. Such particles can be prepared by high energy mixing of ethanolic lipids with aqueous DNA and/or BAF modulator at low pH which protonates the ionizable lipid and provides favorable energetics for nucleic acid/lipid association and nucleation of particles. The particles can be further stabilized through aqueous dilution and removal of the organic solvent. The particles can be concentrated to the desired level.

Another example of a nanoparticle that finds use in the delivery of the subject DNAs and/or BAF modulators is a metal nanoparticle. In some embodiments, a DNA and/or BAF modulator as described herein is delivered by a gold nanoparticle. Generally, a nucleic acid can be covalently bound to a gold nanoparticle or non-covalently bound to a gold nanoparticle (e.g., bound by a charge-charge interaction), for example as described by Ding et al. (2014). Gold Nanoparticles for Nucleic Acid Delivery. Mol. Ther. 22(6); 1075-1083. In some embodiments, gold nanoparticle-nucleic acid conjugates are produced using methods described, for example, in U.S. Pat. No. 6,812,334.

In some embodiments, DNAs and/or BAF modulators described herein can be readily formulated in high concentrations of chitosan-nucleic acid polyplex compositions and administered orally in DNA enteric coated pills described in U.S. Pat. Nos. 8,846,102; 9,404,088; and 9,850,323, each of which is incorporated herein by its entirety.

Exosomes. In some embodiments, a DNA and/or BAF modulator as described herein is delivered by being packaged in an exosome. Exosomes are small membrane vesicles of endocytic origin that are released into the extracellular environment following fusion of multivesicular bodies with the plasma membrane. Their surface consists of a lipid bilayer from the donor cell's cell membrane, they contain cytosol from the cell that produced the exosome, and exhibit membrane proteins from the parental cell on the surface. Exosomes are produced by various cell types including epithelial cells, B and T lymphocytes, mast cells (MC) as well as dendritic cells (DC). Some embodiments, exosomes with a diameter between 10 nm and 1 μm, between 20 nm and 500 nm, between 30 nm and 250 nm, between 50 nm and 100 nm are envisioned for use. Exosomes can be isolated for delivery to target cells using either their donor cells or by introducing specific nucleic acids into them. Various approaches known in the art can be used to produce exosomes containing capsid-free AAV vectors of the present invention.

Conjugates. In some embodiments, a DNA and/or BAF modulator as described herein as disclosed herein is conjugated (e.g., covalently bound to an agent that increases cellular uptake. An "agent that increases cellular uptake" is a molecule that facilitates transport of a nucleic acid across a lipid membrane. For example, a nucleic acid can be conjugated to a lipophilic compound (e.g., cholesterol, tocopherol, etc.), a cell penetrating peptide (CPP) (e.g., penetratin, TAT, Syn1B, etc.), and polyamines (e.g., spermine). Further examples of agents that increase cellular uptake are disclosed, for example, in Winkler (2013). Oligonucleotide conjugates for therapeutic applications. Ther. Deliv. 4(7); 791-809.

In some embodiments, a DNA and/or BAF modulator as described herein as disclosed herein is conjugated to a polymer (e.g., a polymeric molecule) or a folate molecule (e.g., folic acid molecule). Generally, delivery of nucleic acids conjugated to polymers is known in the art, for example as described in WO2000/34343 and WO2008/022309. In some embodiments, a DNA and/or BAF modulator as disclosed herein is conjugated to a poly(amide) polymer, for example as described by U.S. Pat. No. 8,987,377. In some embodiments, a nucleic acid described by the disclosure is conjugated to a folic acid molecule as described in U.S. Pat. No. 8,507,455. In some embodiments, a DNA and/or BAF modulator as described herein as disclosed herein is conjugated to a carbohydrate, for example as described in U.S. Pat. No. 8,450,467.

Nanocapsules. Alternatively, nanocapsule formulations of a DNA and/or BAF modulator can be used. Nanocapsules can generally entrap substances in a stable and reproducible way. To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use.

Liposomes. A DNA and/or BAF modulator as described herein can be added to liposomes for delivery to a cell or target organ in a subject. Liposomes are vesicles that possess at least one lipid bilayer

and an aqueous core. Liposomes are typically used as carriers for drug/therapeutic delivery in the context of pharmaceutical development. They work by fusing with a cellular membrane and repositioning its lipid structure to deliver a drug or active pharmaceutical ingredient (API). Liposome compositions for such delivery are composed of phospholipids, especially compounds having a phosphatidylcholine group, however these compositions may also include other lipids.

The formation and use of liposomes is generally known to those of skill in the art. Liposomes have been developed with improved serum stability and circulation half-times (U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5,795,587).

Also provided herein are cells produced by such methods, and organisms (such as animals, plants, or fungi) comprising or produced from such cells.

Target Cells

Cells employed in embodiments of the invention, referred to herein as "target cells" may vary. It should be appreciated that the target cell may be of any origin, for example from an organism. In some embodiments, the target cell is a mammalian cell. Some non-limiting examples of a mammalian cell include, without limitation, a mouse cell, a rat cell, hamster cell, a rodent cell, and a nonhuman primate cell. In some embodiments, the target cell is a human cell. It should also be appreciated that the target cell may be of any cell type. For example, the target cell may be a stem cell, which may include embryonic stem cells, induced pluripotent stem cells (iPS cells), fetal stem cells, cord blood stem cells, or adult stem cells (i.e., tissue specific stem cells). In other cases, the target cell may be any differentiated cell type found in a subject. Cells of interest include both dividing cells and non-dividing cells. Examples of specific target cells of interest include, but are not limited to: hepatocytes, stellate cells, T lymphocytes, B lymphocytes, NK cells, skeletal muscle cells, cardiomyocytes, neurons, astrocytes, oligodendrocytes, dendritic cells, skin cells, photoreceptors, RPE cells, radial glia, etc.

In some embodiments, the target cell is a cell *in vitro*, and the method includes contacting the cell *in vitro*. In some embodiments, the target cell is a cell in a subject, and the methods include administering the DNA and the agent that modulates BAF activity (where either or both may be present in the same or different suitable delivery vehicle, as desired) to the subject. In some embodiments, the subject is a mammalian subject, for example, a rodent, a mouse, a rat, a hamster, or a non-human primate. In some embodiments, the subject is a human subject.

The target cell may be contacting simultaneously or sequentially with the DNA and the agent that modulates BAF, as desired. In some instances, the target cell is contacted simultaneously with the DNA and the agent that modulates BAF. As such, the DNA and the agent that modulates BAF are contacted at

the same time with the target cells. In other instances, the DNA and the agent that modulates BAF are sequentially contacted with the cell. For example, the target cell may be contacted with the DNA before being contacted with the agent that modulates BAF. Alternatively, the target cell may be contacted with the DNA after being contacted with the agent that modulates BAF.

In embodiments of the invention in which a nuclear targeting factor is provided to the cell, the target cell may be contacted with the DNA and the BAF modulator simultaneously or sequentially with an mRNA encoding the NTF. In some instances, the target cell is contacted simultaneously with the DNA, the agent that modulates BAF, and the NTF. In certain such embodiments, the DNA, the agent that modulates BAF, and the NTF are provided as a single formulation in a delivery vehicle (as described in greater detail below) that is administered to the target cell. In other embodiments, the DNA and agent that modulates BAF are provided as a separate formulation from the NTF that are administered simultaneously to the cell, i.e., as an admixture, or sequentially. In other embodiments, the DNA and the NTF are provided as a separate formulation from the BAF modulator that are administered simultaneously to the cell, i.e., as an admixture, or sequentially. In yet other embodiments, all of the elements are formulated individually, i.e. the DNA is provided as a first formulation, the BAF modulator is provided as a second formulation, and the NTF is provided as a third formulation, where all may be administered simultaneously to the cell, i.e., as an admixture, or sequentially. Preferably, the DNA, the agent that modulates BAF, and the NTF are prepared as a single formulation in the delivery vehicle for contacting with the cell.

Genomic Integration

In some instances, the DNA and agent that modulates BAF are contacted with a target cell in conjunction with a gene editing system, e.g., that is configured to provide genomic integration of the cargo nucleic acid component of the DNA. In such embodiments, as the DNA is contacted with the target cell in conjunction with a gene editing system, it may be contacted with the target cell simultaneously or sequentially with the gene editing system. In some instances, one or more components of a gene editing system may be present with the DNA in a delivery composition, such as a cytosolic delivery composition (e.g., LNP), such as described in greater detail below. Gene editing systems that may be employed in such embodiments may vary, as desired. In some embodiments, the employed gene editing system is configured to genomically integrate the DNA (or desired component thereof, e.g., coding sequence, expression cassette, etc., such as described above) into a specific safe harbor location in the genome, for example to a genomic location within or near an endogenous albumin locus. Generally, one skilled in the art will understand that a safe harbor locus is a location within a genome that can be used for integrating exogenous nucleic acids, where the addition of exogenous nucleic acids into the safe harbor locus does

not cause significant effect on the growth of the host cell by the addition of the nucleic acids alone. In some embodiments, the cargo nucleic acid may be inserted into the specific safe harbor location in the genome that may either utilize the promoter found at that safe harbor locus, or allow the expressional regulation of a coding sequence of the cargo nucleic acid by an exogenous promoter that is fused to the cargo nucleic acid coding sequence prior to insertion.

Gene editing can be conducted using nucleases engineered to target specific sequences. To date there are four major types of nucleases: meganucleases and their derivatives, zinc finger nucleases (ZFNs), transcription activator like effector nucleases (TALENs), and CRISPR-Cas9 nuclease systems. The nuclease platforms vary in difficulty of design, targeting density and mode of action, particularly as the specificity of ZFNs and TALENs is through protein-DNA interactions, while RNA-DNA interactions primarily guide Cas9. Cas9 cleavage also requires an adjacent motif, the PAM, which differs between different CRISPR systems. Cas9 from Streptococcus pyogenes cleaves using a NRG PAM, CRISPR from Neisseria meningitidis can cleave at sites with PAMs including NNNNGATT, NNNNNGTTT and NNNNGCTT. A number of other Cas9 orthologs target protospacer adjacent to alternative PAMs. CRISPR endonucleases, such as Cas9, can be used in various embodiments of the methods of the disclosure. However, the teachings described herein, such as therapeutic target sites, could be applied to other forms of endonucleases, such as ZFNs, TALENs, HEs, or MegaTALs, or using combinations of nucleases. These different systems are now described in further detail.

A given DNA delivery cytosolic delivery composition may include one or more elements of a given gene editing system. Gene editing system employed in embodiments of the invention may include a number of different elements, such as nucleic acid elements, e.g., genome-targeting nucleic acids or Guide RNAs, nucleic acids, e.g., mRNAs encoding endonucleases; polypeptide components, e.g., endonucleases, etc. These various components are now reviewed in greater detail in conjunction with the description of representative endonuclease based genomic integration systems.

In some embodiments, the methods of genome edition and compositions therefore use a nucleic acid sequence (or oligonucleotide) encoding a site-directed polypeptide or DNA endonuclease. The nucleic acid sequence encoding the site-directed polypeptide can be DNA or RNA. If the nucleic acid sequence encoding the site-directed polypeptide is RNA, it can be covalently linked to a gRNA sequence or exist as a separate sequence. In some embodiments, a peptide sequence of the site-directed polypeptide or DNA endonuclease can be used instead of the nucleic acid sequence thereof.

The modifications of the target DNA due to NHEJ and/or HDR can lead to, for example, mutations, deletions, alterations, integrations, gene correction, gene replacement, gene tagging, transgene insertion, nucleotide deletion, gene disruption, translocations and/or gene mutation. The process of integrating non-native nucleic acid, e.g., an exogenously provided DNA, into genomic DNA is an

example of genome editing. A site-directed polypeptide is a nuclease used in genome editing to cleave DNA. The site-directed can be administered to a cell or a patient as either: one or more polypeptides, or one or more mRNAs encoding the polypeptide. In some embodiments, a site-directed polypeptide has a plurality of nucleic acid-cleaving (i.e., nuclease) domains. Two or more nucleic acid-cleaving domains can be linked together via a linker. In some embodiments, the linker has a flexible linker. Linkers can have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40 or more amino acids in length.

CRISPR. In the context of a CRISPR/Cas or CRISPR/Cpf1 system, the site-directed polypeptide can bind to a guide RNA that, in turn, specifies the site in the target DNA to which the polypeptide is directed. In some embodiments of CRISPR/Cas or CRISPR/Cpf1 systems herein, the site-directed polypeptide is an endonuclease, such as a DNA endonuclease.

Naturally-occurring wild-type Cas9 enzymes have two nuclease domains, a HNH nuclease domain and a RuvC domain. Herein, the "Cas9" refers to both naturally-occurring and recombinant Cas9s. Cas9 enzymes contemplated herein have a HNH or HNH-like nuclease domain, and/or a RuvC or RuvC-like nuclease domain.

HNH or HNH-like domains have a McrA-like fold. HNH or HNH-like domains has two antiparallel β-strands and an α-helix. HNH or HNH-like domains has a metal binding site (e.g., a divalent cation binding site). HNH or HNH-like domains can cleave one strand of a target nucleic acid (e.g., the complementary strand of the crRNA targeted strand).

RuvC or RuvC-like domains have an RNaseH or RNaseH-like fold. RuvC/RNaseH domains are involved in a diverse set of nucleic acid-based functions including acting on both RNA and DNA. The RNaseH domain has 5 β-strands surrounded by a plurality of a-helices. RuvC/RNaseH or RuvC/RNaseH-like domains have a metal binding site (e.g., a divalent cation binding site). RuvC/RNaseH or RuvC/RNaseH-like domains can cleave one strand of a target nucleic acid (e.g., the non-complementary strand of a double-stranded target DNA).

In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 10%, at least 15%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% amino acid sequence identity to a wild-type exemplary site-directed polypeptide [e.g., Cas9 from S. pyogenes, US2014/0068797 Sequence ID No. 8 or Sapranauskas et al., Nucleic Acids Res, 39(21): 9275-9282 (2011)], and various other site-directed polypeptides). In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 10%, at least 15%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% amino acid sequence identity to the nuclease domain of a wild-type exemplary site-

directed polypeptide (e.g., Cas9 from S. pyogenes, supra). In some embodiments, a site-directed polypeptide has at least 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids. In some embodiments, a site-directed polypeptide has at most: 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids. In some embodiments, a site-directed polypeptide has at least: 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids in a HNH nuclease domain of the site-directed polypeptide. In some embodiments, a site-directed polypeptide has at most: 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids in a HNH nuclease domain of the site-directed polypeptide. In some embodiments, a site-directed polypeptide has at least: 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids in a RuvC nuclease domain of the site-directed polypeptide. In some embodiments, a site-directed polypeptide has at most: 70, 75, 80, 85, 90, 95, 97, 99, or 100% identity to a wild-type site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra) over 10 contiguous amino acids in a RuvC nuclease domain of the site-directed polypeptide.

In some embodiments, the site-directed polypeptide has a modified form of a wild-type exemplary site-directed polypeptide. The modified form of the wild-type exemplary site-directed polypeptide has a mutation that reduces the nucleic acid-cleaving activity of the site-directed polypeptide. In some embodiments, the modified form of the wild-type exemplary site-directed polypeptide has less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nucleic acid-cleaving activity of the wild-type exemplary site-directed polypeptide (e.g., Cas9 from S. pyogenes, supra). The modified form of the site-directed polypeptide can have no substantial nucleic acid-cleaving activity. When a site-directed polypeptide is a modified form that has no substantial nucleic acid-cleaving activity, it is referred to herein as "enzymatically inactive."

In some embodiments, the modified form of the site-directed polypeptide has a mutation such that it can induce a single-strand break (SSB) on a target nucleic acid (e.g., by cutting only one of the sugar-phosphate backbones of a double-strand target nucleic acid). In some embodiments, the mutation results in less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nucleic acid-cleaving activity in one or more of the plurality of nucleic acid-cleaving domains of the wild-type site directed polypeptide (e.g., Cas9 from S. pyogenes, supra). In some embodiments, the mutation results in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the complementary strand of the

target nucleic acid, but reducing its ability to cleave the non-complementary strand of the target nucleic acid. In some embodiments, the mutation results in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the non-complementary strand of the target nucleic acid, but reducing its ability to cleave the complementary strand of the target nucleic acid. For example, residues in the wild-type exemplary S. pyogenes Cas9 polypeptide, such as Asp10, His840, Asn854 and Asn856, are mutated to inactivate one or more of the plurality of nucleic acid-cleaving domains (e.g., nuclease domains). In some embodiments, the residues to be mutated correspond to residues Asp10, His840, Asn854 and Asn856 in the wild-type exemplary S. pyogenes Cas9 polypeptide (e.g., as determined by sequence and/or structural alignment). Non-limiting examples of mutations include D10A, H840A, N854A or N856A. One skilled in the art will recognize that mutations other than alanine substitutions are suitable.

In some embodiments, a D10A mutation is combined with one or more of H840A, N854A, or N856A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. In some embodiments, a H840A mutation is combined with one or more of D10A, N854A, or N856A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. In some embodiments, a N854A mutation is combined with one or more of H840A, D10A, or N856A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. In some embodiments, a N856A mutation is combined with one or more of H840A, N854A, or D10A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. Site-directed polypeptides that have one substantially inactive nuclease domain are referred to as "nickases".

In some embodiments, variants of RNA-guided endonucleases, for example Cas9, can be used to increase the specificity of CRISPR-mediated genome editing. Wild type Cas9 is generally guided by a single guide RNA designed to hybridize with a specified about.20 nucleotide sequence in the target sequence (such as an endogenous genomic locus). However, several mismatches can be tolerated between the guide RNA and the target locus, effectively reducing the length of required homology in the target site to, for example, as little as 13 nt of homology, and thereby resulting in elevated potential for binding and double-strand nucleic acid cleavage by the CRISPR/Cas9 complex elsewhere in the target genome--also known as off-target cleavage. Because nickase variants of Cas9 each only cut one strand, in order to create a double-strand break it is necessary for a pair of nickases to bind in close proximity and on opposite strands of the target nucleic acid, thereby creating a pair of nicks, which is the equivalent of a double-strand break. This requires that two separate guide RNAs--one for each nickase--must bind in close proximity and on opposite strands of the target nucleic acid. This requirement essentially doubles the minimum length of homology needed for the double-strand break to occur, thereby reducing the likelihood that a double-strand cleavage event will occur elsewhere in the genome, where the two guide

RNA sites--if they exist--are unlikely to be sufficiently close to each other to enable the double-strand break to form. As described in the art, nickases can also be used to promote HDR versus NHEJ. HDR can be used to introduce selected changes into target sites in the genome through the use of specific donor sequences that effectively mediate the desired changes. Descriptions of various CRISPR/Cas systems for use in gene editing can be found, e.g., in international patent application publication number WO2013/176772, and in Nature Biotechnology 32, 347-355 (2014), and references cited therein.

In some embodiments, the site-directed polypeptide (e.g., variant, mutated, enzymatically inactive and/or conditionally enzymatically inactive site-directed polypeptide) targets nucleic acid. In some embodiments, the site-directed polypeptide (e.g., variant, mutated, enzymatically inactive and/or conditionally enzymatically inactive endoribonuclease) targets DNA. In some embodiments, the site-directed polypeptide (e.g., variant, mutated, enzymatically inactive and/or conditionally enzymatically inactive endoribonuclease) targets RNA.

In some embodiments, the site-directed polypeptide has one or more non-native sequences (e.g., the site-directed polypeptide is a fusion protein). In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), a nucleic acid binding domain, and two nucleic acid cleaving domains (i.e., a HNH domain and a RuvC domain). In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), and two nucleic acid cleaving domains (i.e., a HNH domain and a RuvC domain). In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), and two nucleic acid cleaving domains, wherein one or both of the nucleic acid cleaving domains have at least 50% amino acid identity to a nuclease domain from Cas9 from a bacterium (e.g., S. pyogenes). In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), two nucleic acid cleaving domains (i.e., a HNH domain and a RuvC domain), and non-native sequence (for example, a nuclear localization signal) or a linker linking the site-directed polypeptide to a non-native sequence. In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), two nucleic acid cleaving domains (i.e., a HNH domain and a RuvC domain), wherein the site-directed polypeptide has a mutation in one or both of the nucleic acid cleaving domains that reduces the cleaving activity of the nuclease domains by at least 50%. In some embodiments, the site-directed polypeptide has an amino acid sequence having at least 15% amino acid identity to a Cas9 from a bacterium (e.g., S. pyogenes), and two nucleic acid cleaving domains (i.e., a HNH domain and a RuvC domain), wherein one of the nuclease domains

has mutation of aspartic acid 10, and/or wherein one of the nuclease domains has mutation of histidine 840, and wherein the mutation reduces the cleaving activity of the nuclease domain(s) by at least 50%.

In some embodiments, the one or more site-directed polypeptides, e.g., DNA endonucleases, include two nickases that together effect one double-strand break at a specific locus in the genome, or four nickases that together effect two double-strand breaks at specific loci in the genome. Alternatively, one site-directed polypeptide, e.g., DNA endonuclease, affects one double-strand break at a specific locus in the genome.

In some embodiments, a polynucleotide encoding a site-directed polypeptide can be used to edit genome. In some of such embodiments, the polynucleotide encoding a site-directed polypeptide is codon-optimized according to methods standard in the art for expression in the cell containing the target DNA of interest. For example, if the intended target nucleic acid is in a human cell, a human codon-optimized polynucleotide encoding Cas9 is contemplated for use for producing the Cas9 polypeptide.

A CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) genomic locus can be found in the genomes of many prokaryotes (e.g., bacteria and archaea). In prokaryotes, the CRISPR locus encodes products that function as a type of immune system to help defend the prokaryotes against foreign invaders, such as virus and phage. There are three stages of CRISPR locus function: integration of new sequences into the CRISPR locus, expression of CRISPR RNA (crRNA), and silencing of foreign invader nucleic acid. Five types of CRISPR systems (e.g., Type I, Type II, Type III, Type U, and Type V) have been identified.

A CRISPR locus includes a number of short repeating sequences referred to as "repeats." When expressed, the repeats can form secondary hairpin structures (e.g., hairpins) and/or have unstructured single-stranded sequences. The repeats usually occur in clusters and frequently diverge between species. The repeats are regularly interspaced with unique intervening sequences referred to as "spacers," resulting in a repeat-spacer-repeat locus architecture. The spacers are identical to or have high homology with known foreign invader sequences. A spacer-repeat unit encodes a crisprRNA (crRNA), which is processed into a mature form of the spacer-repeat unit. A crRNA has a "seed" or spacer sequence that is involved in targeting a target nucleic acid (in the naturally occurring form in prokaryotes, the spacer sequence targets the foreign invader nucleic acid). A spacer sequence is located at the 5' or 3' end of the crRNA.

A CRISPR locus also has polynucleotide sequences encoding CRISPR Associated (Cas) genes. Cas genes encode endonucleases involved in the biogenesis and the interference stages of crRNA function in prokaryotes. Some Cas genes have homologous secondary and/or tertiary structures.

crRNA biogenesis in a Type II CRISPR system in nature requires a trans-activating CRISPR RNA (tracrRNA). The tracrRNA is modified by endogenous RNaseIII, and then hybridizes to a crRNA

repeat in the pre-crRNA array. Endogenous RNaseIII is recruited to cleave the pre-crRNA. Cleaved crRNAs are subjected to exoribonuclease trimming to produce the mature crRNA form (e.g., 5' trimming). The tracrRNA remains hybridized to the crRNA, and the tracrRNA and the crRNA associate with a site-directed polypeptide (e.g., Cas9). The crRNA of the crRNA-tracrRNA-Cas9 complex guides the complex to a target nucleic acid to which the crRNA can hybridize. Hybridization of the crRNA to the target nucleic acid activates Cas9 for targeted nucleic acid cleavage. The target nucleic acid in a Type II CRISPR system is referred to as a protospacer adjacent motif (PAM). In nature, the PAM is essential to facilitate binding of a site-directed polypeptide (e.g., Cas9) to the target nucleic acid. Type II systems (also referred to as Nmeni or CASS4) are further subdivided into Type II-A (CASS4) and II-B (CASS4a). Jinek et al., Science, 337(6096):816-821 (2012) showed that the CRISPR/Cas9 system is useful for RNA-programmable genome editing, and international patent application publication number WO 2013/176772 provides numerous examples and applications of the CRISPR/Cas endonuclease system for site-specific gene editing.

Type V CRISPR systems have several important differences from Type II systems. For example, Cpf1 is a single RNA-guided endonuclease that, in contrast to Type II systems, lacks tracrRNA. In fact, Cpf1-associated CRISPR arrays are processed into mature crRNAS without the requirement of an additional trans-activating tracrRNA. The Type V CRISPR array is processed into short mature crRNAs of 42-44 nucleotides in length, with each mature crRNA beginning with 19 nucleotides of direct repeat followed by 23-25 nucleotides of spacer sequence. In contrast, mature crRNAs in Type II systems start with 20-24 nucleotides of spacer sequence followed by about 22 nucleotides of direct repeat. Also, Cpf1 utilizes a T-rich protospacer-adjacent motif such that Cpf1-crRNA complexes efficiently cleave target DNA preceded by a short T-rich PAM, which is in contrast to the G-rich PAM following the target DNA for Type II systems. Thus, Type V systems cleave at a point that is distant from the PAM, while Type II systems cleave at a point that is adjacent to the PAM. In addition, in contrast to Type II systems, Cpf1 cleaves DNA via a staggered DNA double-stranded break with a 4 or 5 nucleotide 5' overhang. Type II systems cleave via a blunt double-stranded break. Similar to Type II systems, Cpf1 contains a predicted RuvC-like endonuclease domain, but lacks a second HNH endonuclease domain, which is in contrast to Type II systems.

Exemplary CRISPR/Cas polypeptides include the Cas9 polypeptides in FIG. 1 of Fonfara et al., Nucleic Acids Research, 42: 2577-2590 (2014). The CRISPR/Cas gene naming system has undergone extensive rewriting since the Cas genes were discovered.

A genome-targeting nucleic acid interacts with a site-directed polypeptide (e.g., a nucleic acid-guided nuclease such as Cas9), thereby forming a complex. The genome-targeting nucleic acid (e.g., gRNA, such as described in greater detail below) guides the site-directed polypeptide to a target nucleic

acid.

In some embodiments the site-directed polypeptide and genome-targeting nucleic acid can each be administered separately to a cell or a patient. On the other hand, in some other embodiments the site-directed polypeptide can be pre-complexed with one or more guide RNAs, or one or more crRNA together with a tracrRNA. The pre-complexed material can then be administered to a cell or a patient. Such pre-complexed material is known as a ribonucleoprotein particle (RNP).

Genome-Targeting Nucleic Acid or Guide RNA. Genomic editing components may include a genome-targeting nucleic acid that can direct the activities of an associated polypeptide (e.g., a site-directed polypeptide or DNA endonuclease) to a specific target sequence within a target nucleic acid. In some embodiments, the genome-targeting nucleic acid is an RNA. A genome-targeting RNA is referred to as a "guide RNA" or "gRNA" herein. A guide RNA has at least a spacer sequence that hybridizes to a target nucleic acid sequence of interest and a CRISPR repeat sequence. In Type II systems, the gRNA also has a second RNA called the tracrRNA sequence. In the Type II guide RNA (gRNA), the CRISPR repeat sequence and tracrRNA sequence hybridize to each other to form a duplex. In the Type V guide RNA (gRNA), the crRNA forms a duplex. In both systems, the duplex binds a site-directed polypeptide such that the guide RNA and site-direct polypeptide form a complex. The genome-targeting nucleic acid provides target specificity to the complex by virtue of its association with the site-directed polypeptide. The genome-targeting nucleic acid thus directs the activity of the site-directed polypeptide.

In some embodiments, the genome-targeting nucleic acid is a double-molecule guide RNA. In some embodiments, the genome-targeting nucleic acid is a single-molecule guide RNA. A double-molecule guide RNA has two strands of RNA. The first strand has in the 5' to 3' direction, an optional spacer extension sequence, a spacer sequence and a minimum CRISPR repeat sequence. The second strand has a minimum tracrRNA sequence (complementary to the minimum CRISPR repeat sequence), a 3' tracrRNA sequence and an optional tracrRNA extension sequence. A single-molecule guide RNA (sgRNA) in a Type II system has, in the 5' to 3' direction, an optional spacer extension sequence, a spacer sequence, a minimum CRISPR repeat sequence, a single-molecule guide linker, a minimum tracrRNA sequence, a 3' tracrRNA sequence and an optional tracrRNA extension sequence. The optional tracrRNA extension may have elements that contribute additional functionality (e.g., stability) to the guide RNA. The single-molecule guide linker links the minimum CRISPR repeat and the minimum tracrRNA sequence to form a hairpin structure. The optional tracrRNA extension has one or more hairpins. A single-molecule guide RNA (sgRNA) in a Type V system has, in the 5' to 3' direction, a minimum CRISPR repeat sequence and a spacer sequence.

By way of illustration, guide RNAs used in the CRISPR/Cas/Cpf1 system, or other smaller RNAs can be readily synthesized by chemical means as illustrated below and described in the art. While

chemical synthetic procedures are continually expanding, purifications of such RNAs by procedures such as high performance liquid chromatography (HPLC), which avoids the use of gels such as PAGE) tends to become more challenging as polynucleotide lengths increase significantly beyond a hundred or so nucleotides. One approach used for generating RNAs of greater length is to produce two or more molecules that are ligated together. Much longer RNAs, such as those encoding a Cas9 or Cpf1 endonuclease, are more readily generated enzymatically. Various types of RNA modifications can be introduced during or after chemical synthesis and/or enzymatic generation of RNAs, e.g., modifications that enhance stability, reduce the likelihood or degree of innate immune response, and/or enhance other attributes, as described in the art.

In some embodiments of genome-targeting nucleic acids, a spacer extension sequence can modify activity, provide stability and/or provide a location for modifications of a genome-targeting nucleic acid. A spacer extension sequence can modify on- or off-target activity or specificity. In some embodiments, a spacer extension sequence is provided. A spacer extension sequence can have a length of more than 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 1000, 2000, 3000, 4000, 5000, 6000, or 7000 or more nucleotides. A spacer extension sequence can have a length of about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 1000, 2000, 3000, 4000, 5000, 6000, or 7000 or more nucleotides. A spacer extension sequence can have a length of less than 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 1000, 2000, 3000, 4000, 5000, 6000, 7000 or more nucleotides. In some embodiments, a spacer extension sequence is less than 10 nucleotides in length. In some embodiments, a spacer extension sequence is between 10-30 nucleotides in length. In some embodiments, a spacer extension sequence is between 30-70 nucleotides in length.

In some embodiments, the spacer extension sequence has another moiety (e.g., a stability control sequence, an endoribonuclease binding sequence, a ribozyme). In some embodiments, the moiety decreases or increases the stability of a nucleic acid targeting nucleic acid. In some embodiments, the moiety is a transcriptional terminator segment (i.e., a transcription termination sequence). In some embodiments, the moiety functions in a eukaryotic cell. In some embodiments, the moiety functions in a prokaryotic cell. In some embodiments, the moiety functions in both eukaryotic and prokaryotic cells. Non-limiting examples of suitable moieties include: a 5' cap (e.g., a 7-methylguanylate cap (m7 G)), a riboswitch sequence (e.g., to allow for regulated stability and/or regulated accessibility by proteins and protein complexes), a sequence that forms a dsRNA duplex (i.e., a hairpin), a sequence that targets the RNA to a subcellular location (e.g., nucleus, mitochondria, chloroplasts, and the like), a modification or sequence that provides for tracking (e.g., direct conjugation to a fluorescent molecule, conjugation to a

moiety that facilitates fluorescent detection, a sequence that allows for fluorescent detection, etc.), and/or a modification or sequence that provides a binding site for proteins (e.g., proteins that act on DNA, including transcriptional activators, transcriptional repressors, DNA methyltransferases, DNA demethylases, histone acetyltransferases, histone deacetylases, and the like).

The spacer sequence hybridizes to a sequence in a target nucleic acid of interest. The spacer of a genome-targeting nucleic acid interacts with a target nucleic acid in a sequence-specific manner via hybridization (i.e., base pairing). The nucleotide sequence of the spacer thus varies depending on the sequence of the target nucleic acid of interest.

In a CRISPR/Cas system herein, the spacer sequence is designed to hybridize to a target nucleic acid that is located 5' of a PAM of the Cas9 enzyme used in the system. The spacer can perfectly match the target sequence or can have mismatches. Each Cas9 enzyme has a particular PAM sequence that it recognizes in a target DNA. For example, S. pyogenes recognizes in a target nucleic acid a PAM that has the sequence 5'-NRG-3', where R has either A or G, where N is any nucleotide and N is immediately 3' of the target nucleic acid sequence targeted by the spacer sequence.

In some embodiments, the target nucleic acid sequence has 20 nucleotides. In some embodiments, the target nucleic acid has less than 20 nucleotides. In some embodiments, the target nucleic acid has more than 20 nucleotides. In some embodiments, the target nucleic acid has at least: 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. In some embodiments, the target nucleic acid has at most: 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. In some embodiments, the target nucleic acid sequence has 20 bases immediately 5' of the first nucleotide of the PAM. For example, in a sequence having 5'-NNNNNNNNNNNNNNNNNNNNRG-3', the target nucleic acid has the sequence that corresponds to the Ns, wherein N is any nucleotide, and the underlined NRG sequence (R is G or A) is the Streptococcus pyogenes Cas9 PAM. In some embodiments, the PAM sequence used in the compositions and methods of the present disclosure as a sequence recognized by S.p. Cas9 is NGG.

In some embodiments, the spacer sequence that hybridizes to the target nucleic acid has a length of at least about 6 nucleotides (nt). The spacer sequence can be at least about 6 nt, about 10 nt, about 15 nt, about 18 nt, about 19 nt, about 20 nt, about 25 nt, about 30 nt, about 35 nt or about 40 nt, from about 6 nt to about 80 nt, from about 6 nt to about 50 nt, from about 6 nt to about 45 nt, from about 6 nt to about 40 nt, from about 6 nt to about 35 nt, from about 6 nt to about 30 nt, from about 6 nt to about 25 nt, from about 6 nt to about 20 nt, from about 6 nt to about 19 nt, from about 10 nt to about 50 nt, from about 10 nt to about 45 nt, from about 10 nt to about 40 nt, from about 10 nt to about 35 nt, from about 10 nt to about 30 nt, from about 10 nt to about 25 nt, from about 10 nt to about 20 nt, from about 10 nt to about 19 nt, from about 19 nt to about 25 nt, from about 19 nt to about 30 nt, from about 19 nt to about 35 nt, from about 19 nt to about 40 nt, from about 19 nt to about 45 nt, from about 19 nt to about 50 nt, from about 19

nt to about 60 nt, from about 20 nt to about 25 nt, from about 20 nt to about 30 nt, from about 20 nt to about 35 nt, from about 20 nt to about 40 nt, from about 20 nt to about 45 nt, from about 20 nt to about 50 nt, or from about 20 nt to about 60 nt. In some embodiments, the spacer sequence has 20 nucleotides. In some embodiments, the spacer has 19 nucleotides. In some embodiments, the spacer has 18 nucleotides. In some embodiments, the spacer has 17 nucleotides. In some embodiments, the spacer has 16 nucleotides. In some embodiments, the spacer has 15 nucleotides.

In some embodiments, the percent complementarity between the spacer sequence and the target nucleic acid is at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97%, at least about 98%, at least about 99%, or 100%. In some embodiments, the percent complementarity between the spacer sequence and the target nucleic acid is at most about 30%, at most about 40%, at most about 50%, at most about 60%, at most about 65%, at most about 70%, at most about 75%, at most about 80%, at most about 85%, at most about 90%, at most about 95%, at most about 97%, at most about 98%, at most about 99%, or 100%. In some embodiments, the percent complementarity between the spacer sequence and the target nucleic acid is 100% over the six contiguous 5'-most nucleotides of the target sequence of the complementary strand of the target nucleic acid. In some embodiments, the percent complementarity between the spacer sequence and the target nucleic acid is at least 60% over about 20 contiguous nucleotides. In some embodiments, the length of the spacer sequence and the target nucleic acid can differ by 1 to 6 nucleotides, which can be thought of as a bulge or bulges.

In some embodiments, the spacer sequence is designed or chosen using a computer program. The computer program can use variables, such as predicted melting temperature, secondary structure formation, predicted annealing temperature, sequence identity, genomic context, chromatin accessibility, % GC, frequency of genomic occurrence (e.g., of sequences that are identical or are similar but vary in one or more spots as a result of mismatch, insertion or deletion), methylation status, presence of SNPs, and the like.

Zinc Finger Nucleases. Zinc finger nucleases (ZFNs) are modular proteins having an engineered zinc finger DNA binding domain linked to the catalytic domain of the type II endonuclease FokI. Because FokI functions only as a dimer, a pair of ZFNs must be engineered to bind to cognate target "half-site" sequences on opposite DNA strands and with precise spacing between them to enable the catalytically active FokI dimer to form. Upon dimerization of the FokI domain, which itself has no sequence specificity per se, a DNA double-strand break is generated between the ZFN half-sites as the initiating step in genome editing.

The DNA binding domain of each ZFN generally has 3-6 zinc fingers of the abundant Cys2-His2 architecture, with each finger primarily recognizing a triplet of nucleotides on one strand of the target DNA sequence, although cross-strand interaction with a fourth nucleotide also can be important. Alteration of the amino acids of a finger in positions that make key contacts with the DNA alters the sequence specificity of a given finger. Thus, a four-finger zinc finger protein will selectively recognize a 12 bp target sequence, where the target sequence is a composite of the triplet preferences contributed by each finger, although triplet preference can be influenced to varying degrees by neighboring fingers. An important aspect of ZFNs is that they can be readily re-targeted to almost any genomic address simply by modifying individual fingers, although considerable expertise is required to do this well. In most applications of ZFNs, proteins of 4-6 fingers are used, recognizing 12-18 bp respectively. Hence, a pair of ZFNs will generally recognize a combined target sequence of 24-36 bp, not including the 5-7 bp spacer between half-sites. The binding sites can be separated further with larger spacers, including 15-17 bp. A target sequence of this length is likely to be unique in the human genome, assuming repetitive sequences or gene homologs are excluded during the design process. Nevertheless, the ZFN protein-DNA interactions are not absolute in their specificity so off-target binding and cleavage events do occur, either as a heterodimer between the two ZFNs, or as a homodimer of one or the other of the ZFNs. The latter possibility has been effectively eliminated by engineering the dimerization interface of the FokI domain to create "plus" and "minus" variants, also known as obligate heterodimer variants, which can only dimerize with each other, and not with themselves. Forcing the obligate heterodimer prevents formation of the homodimer. This has greatly enhanced specificity of ZFNs, as well as any other nuclease that adopts these FokI variants.

A variety of ZFN-based systems have been described in the art, modifications thereof are regularly reported, and numerous references describe rules and parameters that are used to guide the design of ZFNs; see, e.g., Segal et al., Proc Natl Acad Sci USA 96(6):2758-63 (1999); Dreier B et al., J Mol Biol. 303(4):489-502 (2000); Liu Q et al., J Biol Chem. 277(6):3850-6 (2002); Dreier et al., J Biol Chem 280(42):35588-97 (2005); and Dreier et al., J Biol Chem. 276(31):29466-78 (2001).

Transcription Activator-Like Effector Nucleases (TALENs). TALENs represent another format of modular nucleases whereby, as with ZFNs, an engineered DNA binding domain is linked to the FokI nuclease domain, and a pair of TALENs operate in tandem to achieve targeted DNA cleavage. The major difference from ZFNs is the nature of the DNA binding domain and the associated target DNA sequence recognition properties. The TALEN DNA binding domain derives from TALE proteins, which were originally described in the plant bacterial pathogen Xanthomonas sp. TALEs have tandem arrays of 33-35 amino acid repeats, with each repeat recognizing a single base pair in the target DNA sequence that is generally up to 20 bp in length, giving a total target sequence length of up to 40 bp. Nucleotide specificity

of each repeat is determined by the repeat variable diresidue (RVD), which includes just two amino acids at positions 12 and 13. The bases guanine, adenine, cytosine and thymine are predominantly recognized by the four RVDs: Asn-Asn, Asn-Ile, His-Asp and Asn-Gly, respectively. This constitutes a much simpler recognition code than for zinc fingers, and thus represents an advantage over the latter for nuclease design. Nevertheless, as with ZFNs, the protein-DNA interactions of TALENs are not absolute in their specificity, and TALENs have also benefitted from the use of obligate heterodimer variants of the FokI domain to reduce off-target activity.

Additional variants of the FokI domain have been created that are deactivated in their catalytic function. If one half of either a TALEN or a ZFN pair contains an inactive FokI domain, then only single-strand DNA cleavage (nicking) will occur at the target site, rather than a DSB. The outcome is comparable to the use of CRISPR/Cas9/Cpf1 "nickase" mutants in which one of the Cas9 cleavage domains has been deactivated. DNA nicks can be used to drive genome editing by HDR, but at lower efficiency than with a DSB. The main benefit is that off-target nicks are quickly and accurately repaired, unlike the DSB, which is prone to NHEJ-mediated mis-repair.

A variety of TALEN-based systems have been described in the art, and modifications thereof are regularly reported; see, e.g., Boch, Science 326(5959):1509-12 (2009); Mak et al., Science 335(6069):716-9 (2012); and Moscou et al., Science 326(5959):1501 (2009). The use of TALENs based on the "Golden Gate" platform, or cloning scheme, has been described by multiple groups; see, e.g., Cermak et al., Nucleic Acids Res. 39(12):e82 (2011); Li et al., Nucleic Acids Res. 39(14):6315-25 (2011); Weber et al., PLoS One. 6(2):e16765 (2011); Wang et al., J Genet Genomics 41(6):339-47, Epub 2014 Can 17 (2014); and Cermak T et al., Methods Mol Biol. 1239:133-59 (2015).

Homing Endonucleases. Homing endonucleases (HEs) are sequence-specific endonucleases that have long recognition sequences (14-44 base pairs) and cleave DNA with high specificity--often at sites unique in the genome. There are at least six known families of HEs as classified by their structure, including LAGLIDADG, GIY-YIG, His-Cis box, H-N-H, PD-(D/E)xK, and Vsr-like that are derived from a broad range of hosts, including eukarya, protists, bacteria, archaea, cyanobacteria and phage. As with ZFNs and TALENs, HEs can be used to create a DSB at a target locus as the initial step in genome editing. In addition, some natural and engineered HEs cut only a single strand of DNA, thereby functioning as site-specific nickases. The large target sequence of HEs and the specificity that they offer have made them attractive candidates to create site-specific DSBs.

A variety of HE-based systems have been described in the art, and modifications thereof are regularly reported; see, e.g., the reviews by Steentoft et al., Glycobiology 24(8):663-80 (2014); Belfort and Bonocora, Methods Mol Biol. 1123:1-26 (2014); Hafez and Hausner, Genome 55(8):553-69 (2012); and references cited therein.

MegaTAL/Tev-mTALEN/MegaTev. As further examples of hybrid nucleases, the MegaTAL platform and Tev-mTALEN platform use a fusion of TALE DNA binding domains and catalytically active HEs, taking advantage of both the tunable DNA binding and specificity of the TALE, as well as the cleavage sequence specificity of the HE; see, e.g., Boissel et al., NAR 42: 2591-2601 (2014); Kleinstiver et al., G3 4:1155-65 (2014); and Boissel and Scharenberg, Methods Mol. Biol. 1239: 171-96 (2015).

In a further variation, the MegaTev architecture is the fusion of a meganuclease (Mega) with the nuclease domain derived from the GIY-YIG homing endonuclease I-TevI (Tev). The two active sites are positioned about 30 bp apart on a DNA substrate and generate two DSBs with non-compatible cohesive ends; see, e.g., Wolfs et al., NAR 42, 8816-29 (2014). It is anticipated that other combinations of existing nuclease-based approaches will evolve and be useful in achieving the targeted genome modifications described herein.

dCas9-FokI or dCpf1-FokI and Other Nucleases. Combining the structural and functional properties of the nuclease platforms described above offers a further approach to genome editing that can potentially overcome some of the inherent deficiencies. As an example, the CRISPR genome editing system generally uses a single Cas9 endonuclease to create a DSB. The specificity of targeting is driven by a 20 or 22 nucleotide sequence in the guide RNA that undergoes Watson-Crick base-pairing with the target DNA (plus an additional 2 bases in the adjacent NAG or NGG PAM sequence in the case of Cas9 from S. pyogenes). Such a sequence is long enough to be unique in the human genome, however, the specificity of the RNA/DNA interaction is not absolute, with significant promiscuity sometimes tolerated, particularly in the 5' half of the target sequence, effectively reducing the number of bases that drive specificity. One solution to this has been to completely deactivate the Cas9 or Cpf1 catalytic function--retaining only the RNA-guided DNA binding function--and instead fusing a FokI domain to the deactivated Cas9; see, e.g., Tsai et al., Nature Biotech 32: 569-76 (2014); and Guilinger et al., Nature Biotech. 32: 577-82 (2014). Because FokI must dimerize to become catalytically active, two guide RNAs are required to tether two FokI fusions in close proximity to form the dimer and cleave DNA. This essentially doubles the number of bases in the combined target sites, thereby increasing the stringency of targeting by CRISPR-based systems.

As further example, fusion of the TALE DNA binding domain to a catalytically active HE, such as I-TevI, takes advantage of both the tunable DNA binding and specificity of the TALE, as well as the cleavage sequence specificity of I-TevI, with the expectation that off-target cleavage can be further reduced.

Additional details regarding gene editing systems that find use in embodiments of the invention may be found in United States Published Patent Application Publication No. 20210348159.

Additional components

In some embodiments, a delivery composition may include a DNA and/or BAF modulator, as well as one or more additional components. In some instances, the one or more additional components may be one or more additional components of a gene editing system, e.g., as described above, which one or more additional components mediate genomic integration of the DNA. As such, a DNA lipid nanoparticle may further include one or more of: an agent that modulates BAF, a guide RNA, an endonuclease, or a nucleic acid coding sequence therefore, e.g., RNA (such as mRNA) or DNA, etc.

The one or more additional compounds can be a therapeutic agent. The therapeutic agent can be selected from any class suitable for the therapeutic objective. In other words, the therapeutic agent can be selected from any class suitable for the therapeutic objective. In other words, the therapeutic agent can be selected according to the treatment objective and biological action desired. For example, if the DNA within the LNP is useful for treating cancer, the additional compound can be an anti-cancer agent (e.g., a chemotherapeutic agent, a targeted cancer therapy (including, but not limited to, a small molecule, an antibody, or an antibody-drug conjugate). In another example, if the LNP containing the DNA is useful for treating an infection, the additional compound can be an antimicrobial agent (e.g., an antibiotic or antiviral compound). In yet another example, if the LNP containing the DNA is useful for treating an immune disease or disorder, the additional compound can be a compound that modulates an immune response (e.g., an immunosuppressant, immunostimulatory compound, or compound modulating one or more specific immune pathways). In some embodiments, different cocktails of different lipid nanoparticles containing different compounds, such as a DNA encoding a different protein or a different compound, such as a therapeutic may be used in the compositions and methods of the invention. In some embodiments, the additional compound is an immune modulating agent. For example, the additional compound is an immunosuppressant. In some embodiments, the additional compound is immune stimulatory agent.

COMPOSITIONS

Also provided are compositions that find use in practicing embodiments of the invention. Compositions of the invention include those having a DNA and/or BAF modulator, e.g., as described, where the DNA and/or BAF modulator may be present in combination with one or more additional components, such as but not limited to, components of a gene editing system, e.g., as described above, such as a gRNA, endonuclease, or nucleic acid encoding the same, etc.

In some embodiments, a composition can have a DNA and/or BAF modulator cytosolic delivery vehicle, e.g., as described above, such as liposome or a lipid nanoparticle or other non-naturally occurring delivery vehicle. Therefore, in some embodiments, any compounds (e.g., a DNA endonuclease or a

nucleic acid encoding thereof, gRNA and DNA and/or BAF modulator) of the composition can be formulated in a liposome or lipid nanoparticle. In some embodiments, one or more such compounds are associated with a liposome or lipid nanoparticle via a covalent bond or non-covalent bond. In some embodiments, any of the compounds can be separately or together contained in a liposome or lipid nanoparticle. Therefore, in some embodiments, each of a DNA endonuclease or a nucleic acid encoding thereof, gRNA and DNA and/or BAF modulator (donor template) is separately formulated in a liposome or lipid nanoparticle. In some embodiments, a DNA endonuclease is formulated in a liposome or lipid nanoparticle with gRNA. In some embodiments, a DNA endonuclease or a nucleic acid encoding thereof, gRNA and donor template are formulated in a liposome or lipid nanoparticle together.

In some embodiments, a composition described above further has one or more additional reagents, where such additional reagents are selected from a buffer, a buffer for introducing a polypeptide or polynucleotide into a cell, a wash buffer, a control reagent, a control vector, a control RNA polynucleotide, a reagent for *in vitro* production of the polypeptide from DNA, adaptors for sequencing and the like. A buffer can be a stabilization buffer, a reconstituting buffer, a diluting buffer, or the like. In some embodiments, a composition can also include one or more components that can be used to facilitate or enhance the on-target binding or the cleavage of DNA by the endonuclease, or improve the specificity of targeting.

Also provided herein is a pharmaceutical composition comprising the delivery vehicle - encapsulated DNA and/or BAF modulator and a pharmaceutically acceptable carrier or excipient. In some aspects, the disclosure provides for a lipid nanoparticle formulation further comprising one or more pharmaceutical excipients. In some embodiments, the lipid nanoparticle formulation further comprises sucrose, tris, trehalose and/or glycine.

In some embodiments, any components of a composition are formulated with pharmaceutically acceptable excipients such as carriers, solvents, stabilizers, adjuvants, diluents, etc., depending upon the particular mode of administration and dosage form. In some embodiments, guide RNA compositions are generally formulated to achieve a physiologically compatible pH, and range from a pH of about 3 to a pH of about 11, about pH 3 to about pH 7, depending on the formulation and route of administration. In some embodiments, the pH is adjusted to a range from about pH 5.0 to about pH 8. In some embodiments, the composition has a therapeutically effective amount of at least one compound as described herein, together with one or more pharmaceutically acceptable excipients. Optionally, the composition can have a combination of the compounds described herein, or can include a second active ingredient useful in the treatment or prevention of bacterial growth (for example and without limitation, anti-bacterial or anti-microbial agents), or can include a combination of reagents of the disclosure. In some embodiments, gRNAs are formulated with other one or more oligonucleotides, e.g., a nucleic acid encoding DNA

endonuclease and/or a donor template. Alternatively, a nucleic acid encoding DNA endonuclease and a donor template, separately or in combination with other oligonucleotides, are formulated with the method described above for gRNA formulation.

Suitable excipients can include, for example, carrier molecules that include large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Other exemplary excipients include antioxidants (for example and without limitation, ascorbic acid), chelating agents (for example and without limitation, EDTA), carbohydrates (for example and without limitation, dextrin, hydroxyalkylcellulose, and hydroxyalkylmethylcellulose), stearic acid, liquids (for example and without limitation, oils, water, saline, glycerol and ethanol), wetting or emulsifying agents, pH buffering substances, and the like.

In some embodiments, a composition refers to a therapeutic composition having therapeutic cells, e.g., modified via methods of the invention such as described herein, that are used in an *ex vivo* treatment method. In some embodiments, therapeutic compositions contain a physiologically tolerable carrier together with the cell composition, and optionally at least one additional bioactive agent as described herein, dissolved or dispersed therein as an active ingredient. In some embodiments, the therapeutic composition is not substantially immunogenic when administered to a mammal or human patient for therapeutic purposes, unless so desired. In general, the genetically-modified, therapeutic cells described herein are administered as a suspension with a pharmaceutically acceptable carrier. One of skill in the art will recognize that a pharmaceutically acceptable carrier to be used in a cell composition will not include buffers, compounds, cryopreservation agents, preservatives, or other agents in amounts that substantially interfere with the viability of the cells to be delivered to the subject. A formulation having cells can include, e.g., osmotic buffers that permit cell membrane integrity to be maintained, and optionally, nutrients to maintain cell viability or enhance engraftment upon administration. Such formulations and suspensions are known to those of skill in the art and/or can be adapted for use with the progenitor cells, as described herein, using routine experimentation. In some embodiments, a cell composition can also be emulsified or presented as a liposome composition, provided that the emulsification procedure does not adversely affect cell viability. The cells and any other active ingredient can be mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient, and in amounts suitable for use in the therapeutic methods described herein. Additional agents included in a cell composition can include pharmaceutically acceptable salts of the components therein. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) that are formed with inorganic acids, such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, tartaric, mandelic and the like. Salts formed with the free carboxyl groups can also be derived from

inorganic bases, such as, for example, sodium, potassium, ammonium, calcium or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine and the like. Physiologically tolerable carriers are well known in the art. Exemplary liquid carriers are sterile aqueous solutions that contain no materials in addition to the active ingredients and water, or contain a buffer such as sodium phosphate at physiological pH value, physiological saline or both, such as phosphate-buffered saline. Still further, aqueous carriers can contain more than one buffer salt, as well as salts such as sodium and potassium chlorides, dextrose, polyethylene glycol and other solutes. Liquid compositions can also contain liquid phases in addition to and to the exclusion of water. Exemplary of such additional liquid phases are glycerin, vegetable oils such as cottonseed oil, and water-oil emulsions. The amount of an active compound used in the cell compositions that is effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques.

KITS

Aspects of the present disclosure also include kits. Some aspects of this disclosure provide kits that include a DNA and/or BAF modulator, e.g., as described above. Kits of the invention may further include one or more additional components, e.g., an inducing agent (e.g., as described above), a gene editing system or components thereof, e.g., as described above, etc. In some instances, the various components of a given kit may be combined into a single composition, e.g., with a cytosolic delivery vehicle, as a pharmaceutical composition, etc., such as described above.

In some instances, kits may include an article of manufacture containing materials useful for the treatment of the diseases described above is included. In some embodiments, the article of manufacture comprises a container and a label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. In some embodiments, the container holds a composition that is effective for treating a disease described herein and may have a sterile access port. For example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle. The active agent in the composition is a compound of the invention. In some embodiments, the label on or associated with the container indicates that the composition is used for treating the disease of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, or dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.

Components of the kits may be present in separate containers, or multiple components may be present in a single container. In addition to the above-mentioned components, a subject kit may further include instructions for using the components of the kit, e.g., to practice the subject methods. The instructions are generally recorded on a suitable recording medium. For example, the instructions may be printed on a substrate, such as paper or plastic, etc. As such, the instructions may be present in the kits as a package insert, in the labeling of the container of the kit or components thereof (i.e., associated with the packaging or sub-packaging) etc. In other embodiments, the instructions are present as an electronic storage data file present on a suitable computer readable storage medium, e.g., CD-ROM, diskette, Hard Disk Drive (HDD), portable flash drive, etc. In yet other embodiments, the actual instructions are not present in the kit, but means for obtaining the instructions from a remote source, e.g., via the internet, are provided. An example of this embodiment is a kit that includes a web address where the instructions can be viewed and/or from which the instructions can be downloaded. As with the instructions, this means for obtaining the instructions is recorded on a suitable substrate.

UTILITY

The subject methods and compositions, e.g., as described above, can be used in any application where nuclear delivery of a DNA is desired. Applications of interest include both research and therapeutic applications. Applications of interest include, but are not limited to: research applications, diagnostic applications and therapeutic applications. In some instances, nucleic acids that may be introduced into a nucleus via methods of the invention include those encoding research proteins, diagnostic proteins and therapeutic proteins.

Research proteins are proteins whose activity finds use in a research protocol. As such, research proteins are proteins that are employed in an experimental procedure. The research protein may be any protein that has such utility, where in some instances the research protein is a protein domain that is also provided in research protocols by expressing it in a cell from an encoding vector. Examples of specific types of research proteins include, but are not limited to: transcription modulators of inducible expression systems, members of signal production systems, e.g., enzymes and substrates thereof, hormones, prohormones, proteases, enzyme activity modulators, perturbimers and peptide aptamers, antibodies, modulators of protein-protein interactions, genomic modification proteins, such as CRE recombinase, meganucleases, Zinc-finger nucleases, CRISPR/Cas-9 nuclease, TAL effector nucleases, etc., cellular reprogramming proteins, such as Oct 3/4, Sox2, Klf4, c-Myc, Nanog, Lin-28, etc., and the like.

Diagnostic proteins are proteins whose activity finds use in a diagnostic protocol. As such, diagnostic proteins are proteins that are employed in a diagnostic procedure. The diagnostic protein may be any protein that has such utility. Examples of specific types of diagnostic proteins include, but are not

limited to: members of signal production systems, e.g., enzymes and substrates thereof, labeled binding members, e.g., labeled antibodies and binding fragments thereof, peptide aptamers and the like.

Proteins of interest further include therapeutic proteins. Therapeutic proteins are proteins that provide a therapeutic benefit to a patient, and include secreted proteins, transmembrane proteins, and intracellularly acting proteins. As will be appreciated by one of ordinary skill in the art, cargo nucleic acids that encode any protein that is associated with a liver disease or that, upon secretion from the liver, find use in treating another organ in the body, may be delivered using the subject compositions and methods.

Target cells to which nucleic acids may be delivered in accordance with the invention may vary widely. Target cells of interest include, but are not limited to: cell lines, HeLa, HEK, CHO, 293 and the like, Mouse embryonic stem cells, human stem cells, mesenchymal stem cells, primary cells, tissue samples and the like. Some non-limiting examples of a mammalian cell include, without limitation, a mouse cell, a rat cell, hamster cell, a rodent cell, and a nonhuman primate cell. In some embodiments, the target cell is a human cell. It should also be appreciated that the target cell may be of any cell type. For example, the target cell may be a stem cell, which may include embryonic stem cells, induced pluripotent stem cells (iPS cells), fetal stem cells, cord blood stem cells, or adult stem cells (i.e., tissue specific stem cells). In other cases, the target cell may be any differentiated cell type found in a subject. Cells of interest include both dividing cells and non-dividing cells. Examples of specific target cells of interest include, but are not limited to: hepatocytes, stellate cells, T lymphocytes, B lymphocytes, NK cells, skeletal muscle cells, cardiomyocytes, neurons, astrocytes, oligodendrocytes, dendritic cells, skin cells, etc.

In some instances, the application of interest is a therapeutic application, for example, in the treatment of a disease. For example, the compositions and methods of the present application may be used to deliver a nucleic acid sequence to the nucleus of a cell to complement a genetic deficiency. As one nonlimiting example, compositions of the present application may be used in the treatment of a genetic deficiency that impacts the function of hepatocytes, or in the treatment of a genetic deficiency elsewhere in the body that can be remedied by leveraging hepatocytes as a biofactory to secrete the deficient protein.

The following example(s) is/are offered by way of illustration and not by way of limitation.

EXAMPLES

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the

experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

Exogenous DNA becomes associated with nuclear envelope proteins upon entry into the cytoplasm of cells. To assess the extent to which DNA is able to access the nucleus in nondividing cells, growth-arrested HepG2 cells (Fig. 1-3) and primary hepatocytes (data not shown) were transfected with EdU-labeled doggybone DNA (i.e., doggybone DNA with 5-ethynyl-2'-deoxyuridine (EdU) incorporated to enable conjugation of a fluorophore during the staining protocol) and analyzed by confocal microscopy 16 hours later. As shown in Fig. 1A and at higher magnification in Fig. 1A', the EdU-labeled DNA was observed to form puncta throughout the cytoplasm of the cell. To better understand what was driving the formation of these puncta, the cells were incubated with a panel of antibodies that are specific for proteins previously reported to associate with DNA and assessed by immunofluorescence. Two proteins were observed to colocalize to the DNA puncta: Barrier-to-autointegration factor (BAF) and emerin. BAF was observed to reside primarily in the nucleus of untransfected HepG2 controls (Fig.1C and 1C'), and colocalized with DNA in the cytoplasm of EdU-DNA transfected cells (Fig.1B and 1B'). Similarly, emerin was observed to reside primarily at the nuclear membrane in untransfected (Fig.2A, B), mock transfected (Fig.2C), and mRNA-transfected (Fig. 2D) cells, but colocalized with the EdU-labeled DNA in DNA-transfected cells (Fig.1E, 1F). Confocal microscopy revealed emerin (Fig. 3A and 3C, red) enveloping the EdU-DNA (Fig.3B and 3C, yellow).

Exogenous mRNAs encoding BAF mutant proteins enhance expression of co-delivered DNA. BAF and emerin are two proteins that have been reported to be involved in the formation of the nuclear envelope around chromosomal DNA following mitosis. BAF is known in the art to bind chromatin and help it to anchor to the nuclear envelope. During mitosis, phosphorylation by vaccinia-related kinase 1 (VRK1) or 2 (VRK2) causes BAF to release chromosomal DNA, allowing the nuclear envelope to disassemble. BAF dephosphorylation by protein phosphatase 2A (PP2A) at the end of mitosis allows BAF to recapture the DNA and, working in concert with LEM-domain containing proteins LAP2, emerin and MAN1, mediate its packaging within a new nuclear envelope. BAF has also been reported to recognize and bind foreign (viral and plasmid) DNA immediately following entry into the cytoplasm, nucleating the formation of a nuclear envelope-like structure around the DNA so as to sequester the DNA for degradation.

To determine if inhibiting BAF could render more DNA available to translocate to the nucleus, we attempted to create a BAF-deficient cell line by CRISPR-directed knockout; however, this deficiency

was lethal, indicating that at least some amount of BAF is required for cell viability. In an alternative strategy, preparations of lipid nanoparticles (LNPs) were co-formulated with a nanoplasmid DNA (npDNA) encoding a human Factor IX (hFIX) expression cassette and an mRNA encoding a BAF mutant comprising a L58R mutation, which is predicted to disrupt BAF's binding to LEM-domain containing proteins and the formation/stability of the DNA-BAF-LEM complex. Four different LNP formulations were generated following methods as described in PCT Application Serial No. PCT/US2023/079923 filed on November 15, 2023, the disclosure of which is herein incorporated by reference. Each formulation comprised a different weight ratio of DNA to BAF – 1:2, 1:1, 1:0.33, and 1:0. The LNP formulations were administered to cohorts of wild type mice by i.v. injection at a dose of 0.3 mg/kg DNA (resulting in doses of 0.3mg/kg DNA/0.6 mg/kg mRNA [Fig.4, group 1]; 0.3 mg/kg DNA/0.3 mg/kg mRNA [Fig.4, group 2]; 0.3 mg/kg DNA /0.1 mg/kg mRNA [Fig.4, group 3]; and 0.3 mg/kg DNA/0.03 mg/kg mRNA [Fig.4, group 4, respectively), and the amount of hFIX in the plasma of the treated mice was detected by ELISA 7 days later. In these experiments and below, the concentration of hFIX in plasma is reported relative to the amount of hFIX in plasma of normal human samples, i.e. "% normal hFIX levels". Given that the average concentration of hFIX in pooled normal human samples is 5 ug/ml, 100% normal human Factor IX levels is approximately 5 ug/ml.

Expecting that the mRNA encoding the BAF modulator would take time to be translated while BAF and emerin would colocalize with DNA and begin to sequester it within a nuclear envelope-type structure immediately upon entry of the DNA into the cell, we were surprised to detect levels of hFIX in the plasma of mice injected with LNPs comprising DNA+BAF modulator that were 4-to 150-fold greater than in mice injected with LNPs comprising only the npDNA (Fig.4, groups 1-4 versus DNA only group 6). Expression levels appeared to be dependent on the dose of BAF modulator delivered.

To determine if these expression levels could be increased further by actively driving nuclear translocation of the available DNA, an LNP was formulated comprising a npDNA encoding the human Factor IX (hFIX) expression cassette and comprising a UAS DNA targeting sequence (DTS) as disclosed in PCT/US2023/035932; an mRNA encoding the BAF-L58R protein; and an mRNA encoding a GAL4 nuclear translocating factor (NTF), i.e. a GAL4 protein fused to an NLS at its C terminus, as disclosed in PCT/US2023/035932, in a weight ratio of 1:1:2 respectively. Administration of this LNP at a dose of 0.3 mg/kg DNA / 0.3 mg/kg BAF modulator mRNA / 0.6 mg/kg GAL4 mRNA yielded amounts of hFIX protein in plasma at day 7 that were 200-fold greater than the amounts achieved by the LNP comprising the npDNA alone (Fig.4, group 5 versus DNA only group 6), and 4-fold greater than the LNP comprising the npDNA + BAF modulator (Fig.4, group 5 versus group 2).

The aforementioned study was performed with an mRNA encoding a mouse BAF-L58R protein. To determine if a human BAF protein comprising the L58R substitution would behave similarly, the

UAS-hFIX-npDNA, the GAL4 mRNA, and either the mRNA encoding the mouse BAF-L58R protein used previously or an mRNA encoding a human BAF-L58R protein were coformulated into LNPs in a 1:1:2 weight ratio and administered to mice by i.v. bolus injection at a dose of 0.3 mg/kg DNA / 0.3 mg/kg BAF modulator mRNA / 0.6 mg/kg GAL4 mRNA. The human BAF-L58R appeared to be approximately 50% as efficient as the mouse BAF-L58R at facilitating DNA uptake and translocation by the nuclear translocating factor in mouse tissue (Fig.5, group 2 versus group 1), suggesting some species selectivity.

To determine how the kinetics of the BAF modulator impacted DNA delivery and expression, an mRNA was designed encoding a mouse BAF-L58R protein modified to comprise an alternative UTR known to reduce the half-life of the mRNA (UTR38, disclosed above). The UAS-hFIX-npDNA, the GAL4 mRNA, and the mRNA encoding the BAF-L58R with UTR38 were co-formulated into LNPs in a 1:1:2 weight ratio and administered to mice by i.v. bolus injection at a dose of 0.3 mg/kg DNA / 0.3 mg/kg BAF modulator mRNA / 0.6 mg/kg GAL4 mRNA. Decreasing the stability of the BAF modulator mRNA yielded a 2-fold increase in expression from the npDNA expression cassette (Fig.5, group 3) over the unmodified mouse BAF-L58R protein (Fig.5, group 1).

To determine if forced nuclear localization of the BAF-L58R protein could further enhance delivery of DNA to the nucleus, e.g. as a nuclear translocating factor, an mRNA was designed to encode a protein that comprised the mouse BAF-L58R protein with an NLS at the C-terminus. The UAS-hFIX-npDNA, the mRNA encoding the BAF modulator (i.e. mBAF-L58R or mBAF-L58R-C'NLS), and an mRNA encoding GAL4 were co-formulated into LNPs in a 1:0.3:1 weight ratio and administered to mice by i.v. bolus injection at a dose of 0.3 mg/kg DNA / 0.1 mg/kg BAF modulator mRNA / 0.3 mg/kg GAL4 mRNA. Inclusion of an NLS enhanced DNA delivery 80-fold (Fig.6, group 2).

Exogenous mRNAs encoding wild type BAF enhance expression of co-delivered DNA. To determine the effect that an excess amount of wild type BAF protein might have on the sequestration of DNA in the cytoplasm, mRNA encoding wild type mouse BAF protein was co-formulated in LNPs with UAS-hFIX-npDNA and GAL4 mRNA. The nucleic acid species were formulated in a weight ratio of 1:1:1 [DNA : wt BAF mRNA : GAL4 mRNA], and administered to wild type mice by i.v. injection at a dose of 0.3 mg/kg DNA / 0.3 mg/kg wt BAF mRNA / 0.3 mg/kg GAL4 mRNA. Levels of hFIX in plasma were detected at day 7 post-dosing by ELISA (Fig. 7). hFIX levels were augmented by exogenously added wild type BAF 160-fold or more over DNA alone (81% normal human levels with Fig. 7 group 1 versus 0.5% normal human levels for DNA alone historical controls).

Exogenous mRNAs encoding BAF kinases enhance expression of co-delivered DNA. BAF engagement with DNA is modulated by phosphorylation at BAF's N-terminus, with phosphorylation of residues Threonine 2, Threonine 3, and Serine 4 triggering BAF release of DNA. To determine the effect

that a co-delivered BAF kinase would have on BAF's ability to mediate DNA sequestration, mRNAs encoding the vaccinia B1 protein Ankara strain and the vaccinia B1 protein WR strain were co-formulated in LNPs with UAS-hFIX-npDNA and GAL4 mRNA. An LNP co-formulated with UAS-hFIX-npDNA and mRNAs encoding GAL4 and mouse BAF-L58R was also prepared as a control. As before, the nucleic acid species were formulated in a weight ratio of 1:1:2 [DNA : BAF modulator (kinase or BAF-L58R mutant) mRNA : GAL4 mRNA], and administered to wild type mice at a dose of 0.3 mg/kg DNA / 0.3 mg/kg BAF modulator mRNA / 0.6 mg/kg GAL4 mRNA. Levels of hFIX in plasma were detected at day 3 and day 7 post-dosing by ELISA. Surprisingly, hFIX levels were augmented 3000-fold or more over DNA alone by the mRNA encoding the Ankara strain B1 protein (1661% normal human levels with Fig. 8 group 1 versus 0.5% normal human levels for DNA alone historical controls) and 600-fold or more over DNA alone by the mRNA encoding the Western Reserve strain B1 protein (310% normal human level with Fig. 8, group 2 versus 0.5% normal human levels for DNA alone historical controls). Interestingly, relative to the BAF-L58R control (Fig. 8, group 3), a 40-fold boost and 8-fold boost in expression were observed with the Ankara and Western Reserve B1 proteins, respectively, suggesting that BAF kinases may be acting on other proteins in addition to BAF to facilitate DNA delivery to the nucleus or that BAF kinases are more efficient modulators of endogenous BAF than BAF mutant proteins.

BAF inhibition increases the number of cells expressing protein from the co-delivered DNA. Increases in hFIX levels in plasma following exposure of a tissue to an agent that modulates BAF could be due to the few cells that expressed protein in the absence of agent expressing more protein in the presence of the agent, more cells in the tissue expressing protein, or both. To determine whether the increase in hFIX expression in the mice receiving LNPs comprising BAF modulator was due to an increase in the total number of cells expressing protein, an increase in the amount of protein expressed per cell, or both, immunohistochemistry for hFIX protein was performed on sections of liver from mice dosed with PBS (Fig.9A, 27 days post-dosing) or with LNPs formulated with UAS-hFIX-npDNA (Fig.9B, 27 days post-dosing), with UAS-hFIX-npDNA + GAL4 mRNA (Fig.9C, 27 days post-dosing), or with UAS-hFIX-npDNA + GAL4 mRNA + BAF-L58R mRNA (Fig.9D, 15 days post dosing). As illustrated in Figure 8, it appears that both the number of cells and the amount of protein within each cell increases when mRNA encoding a BAF modulator is included in the formulation, with 15% of cells expressing hFIX in mice receiving the BAF modulator versus 0.5% of cells receiving DNA alone and 4% of cells receiving DNA and the mRNA encoding the nuclear translocating protein GAL4.Additional exemplary embodiments of proteins that modulate BAF and the mRNAs that encode them are provided in Table 10.

BAF inhibition results in better FIX expression across species, from mice to rats to NHPs. LNP formulations were generated following methods as described in PCT Application Serial No.

PCT/US2023/079923 filed on November 15, 2023, the disclosure of which is herein incorporated by reference. LNP formulations comprising 0.3 mg/kg FIX npDNA (comprising the hFIX transgene operably linked to a promoter but no BAF modulator DTS sequence) and 0.2 mg/kg of mRNA encoding a mutant human BAF protein of the present disclosure fused to an NLS were intravenously administered to cohorts of wild type mice, rats, or NHPs. Levels of hFIX protein in plasma were detected from day 3 to day 14 post-dosing by ELISA. In mice, delivery of the npDNA + BAF mRNA LNP resulted in hFIX levels over 100% of normal human levels, which persisted over 14 days. In rats, delivery of the npDNA + BAF mRNA LNP resulted in hFIX levels of approximately 100% of normal human levels, which persisted over 14 days. In NHPs, delivery of the npDNA + BAF mRNA LNP resulted in hFIX levels of approximately 5% of normal human levels, which persisted over 14 days.

BAF inhibition results in better FIX expression in human cells in vitro. Primary human hepatocytes (PHHs) were plated in 384-well collagen-coated tissue culture plates. LNP formulations comprising 40 ng GFP npDNA and 20 ng mRNA encoding BAF variants of the present disclosure were added to the PHHs in OptiMEM with 45 ug/mL ApoE. GFP fluorescence intensity was measured in each well 72 h post-transfection. BAF variant 3 produced the highest increase in GFP expression in human cells, approximately 5-fold higher than the GFP signal observed in untreated cells (Fig. 11).

Table 10. Exemplary embodiments of agents that modulate BAF

SEQ NO.	In vitro transcription template Sequence	mRNA Sequence	Protein Sequence
1	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGA	MKFERQGGSTTSQKHR
	CGACTCACTATAAGGGGATAATA	AGUCAGCGGUAGACGCACGGCAUAGCA	DFVAEPMGEKPVGSLA
	GGAAAAATCTATAGCAGAAGTCA	UCCAACGCCACCAUGAAAUUCGAAAGG	GIGDVLSKRLEERGFD
	GCGGTAGACGCACGGCATAGCAT	CAAGGCGGCAGCACUACGAGUCAGAAG	KAYVVLGQFLVLKKDE
	CCAACGCCACCATGAAATTCGAA	CAUCGCGAUUUUGUAGCUGAGCCAAUG	DRFREWLKDTCGANAK
	AGGCAAGGCGGCAGCACTACGAG	GGCGAAAAGCCUGUGGGAAGCUUGGCU	QSRDCFGCLREWCDAF
	TCAGAAGCATCGCGATTTTGTAG	GGAAUAGGGGACGUACUGAGCAAAAGG	LWSHPQFEKWSHPQFE
	CTGAGCCAATGGGCGAAAAGCCT	CUGGAGGAAAGAGGCUUUGAUAAGGCU	KWSHPQFEKPAAKRVK
	GTGGGAAGCTTGGCTGGAATAGG	UAUGUGGUACUUGGGCAGUUUCUGGUA	LDASQGTKRSYEQMET
	GGACGTACTGAGCAAAAGGCTGG	UUGAAAAAGGACGAGGACAGGUUCAGG	DGERQ
	AGGAAAGAGGCTTTGATAAGGCT	GAAUGGCUCAAAGACACAUGUGGUGCA	
	TATGTGGTACTTGGGCAGTTTCT	AACGCCAAACAAAGCCGUGACUGCUUC	
	GGTATTGAAAAAGGACGAGGACA	GGAUGCCUGCGGGAGUGGUGUGACGCA	
	GGTTCAGGGAATGGCTCAAAGAC	UUCCUGUGGAGCCACCCUCAGUUCGAA	
	ACATGTGGTGCAAACGCCAAACA	AAAUGGUCUCAUCCUCAGUUCGAAAAA	
	AAGCCGTGACTGCTTCGGATGCC	UGGUCUCAUCCCCAGUUUGAGAAGCCC	
	TGCGGGAGTGGTGTGACGCATTC	GCCGCAAAGCGUGUAAAACUGGACGCC	
	CTGTGGAGCCACCCTCAGTTCGA	UCCCAGGGAACAAAGAGGUCUUACGAA	
	AAAATGGTCTCATCCTCAGTTCG	CAGAUGGAAACAGAUGGCGAACGACAA	
	AAAAATGGTCTCATCCCCAGTTT	UAAAGACCAGACAGCUGGAGCCUCGGU	
	GAGAAGCCCGCCGCAAAGCGTGT	GGCCAUGCUUCUUGCCCCUUGGGCCUC	
	AAAACTGGACGCCTCCCAGGGAA	CCCCCAGCCCCUCCUCCCCUUCCUGCA	

	CAAAGAGGTCTTACGAACAGATG GAAACAGATGGCGAACGACAATA AAGACCAGACAGCTGGAGCCTCG GTGGCCATGCTTCTTGCCCCTTG GGCCTCCCCCCAGCCCCTCCTCC CCTTCCTGCACCCGTACCCCCGT GGTCTTTGAATAAAGTCTGAGTG GGCGGCAAGCTTTTCTTCTCTCC CTTGCACCT	CCCGUACCCCCGUGGUCUUUGAAUAAA GUCUGAGUGGGCGGCAAGCUUUUCUUC UCUCCCUUGCACCUAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
2	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGGAGGTTCA GGAGGTAGCGGAGGAAGCACTAC AAGCCAAAAACATAGGGATTTTG TGGCAGAGCCAATGGGGGAGAAA CCTGTCGGATCCCTTGCAGGTAT CGGGGACGTTTTGAGCAAGAGGT TGGAAGAAAGGGGATTCGATAAG GCATACGTCGTGCTCGGACAGTT TCTTGTGTTAAAAAAGGATGAAG ACCGTTTCCGGGAGTGGTTGAAA GATACCTGTGGCGCCAACGCTAA ACAGAGCAGAGATTGTTTGGGT GTTTGCGAGAGTGGTGCGACGCT TTCCTTTGGAGCCACCCTCAGTT TGAAAAGTGGTCCCACCCCCAGT TCGAAAAATGGAGCCACCCACAG TTTGAAAAACCCGCCGCCAAGCG CGTGAAGTTGGACGCTAGTCAAG GGACTAAGAGATCTTACGAACAG ATGGAAACAGATGGCGAGAGACA GTAAAGACCAGACAGCTGGAGCC TCGGTGGCCATGCTTCTTGCCCC TTGGGCCTCCCCCCAGCCCCTCC TCCCCTTCCTGCACCCGTACCCC CGTGGTCTTTGAATAAAGTCTGA GTGGGCGGCAAGCTTTTCTTCTC TCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGGAGGUUCAGGAGGU AGCGGAGGAAGCACUACAAGCCAAAAAC AUAGGGAUUUUGUGGCAGAGCCAAUGGG GGAGAAACCUGUCGGAUCCCUUGCAGGU AUCGGGGACGUUUUGAGCAAGAGGUUGG AAGAAAGGGGAUUCGAUAAGGCAUACGU CGUGCUCGGACAGUUUCUUGUGUUAAAA AAGGAUGAAGACCGUUUCCGGGAGUGGU UGAAAGAUACCUGUGGCGCCAACGCUAA ACAGAGCAGAGAUUGUUUUGGGUGUUUG CGAGAGUGGUGCGACGCUUUCCUUUGGA GCCACCCUCAGUUUGAAAAGUGGUCCCA CCCCCAGUUCGAAAAAUGGAGCCACCCA CAGUUUGAAAAACCCGCCGCCAAGCGCG UGAAGUUGGACGCUAGUCAAGGGACUAA GAGAUCUUACGAACAGAUGGAAACAGAU GGCGAGAGACAGUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCAAGCU UUUCUUCUCUCCCUUGCACCUAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	MGGSGGSGGSTTSQKHR DFVAEPMGEKPVGSLAG IGDVLSKRLEERGFDKA YVVLGQFLVLKKDEDRF REWLKDTCGANAKQSRD CFGCLREWCDAFLWSHP QFEKWSHPQFEKWSHPQ FEKPAAKRVKLDASQGT KRSYEQMETDGERQ
3	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGAAATTTGAA AGACAGTGGAGCCACCCACAGTT CGAAAAGTGGAGCCACCCCCAGT TTGAGAAATGGAGCCATCCTCAA TTTGAAAAGACCACCAGCCAGAA GCACCGGGATTTTGTTGCCGAGC CAATGGGGGAAAAACCAGTAGGA TCACTGGCAGGGATTGGTGATGT TCTCTCAAAGAGACTGGAAGAAC GCGGATTCGATAAAGCATACGTC GTCTTGGGACAGTTCCTTGTTTT GAAGAAAGATGAGGATAGATTCA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAAAUUUGAAAGACAG UGGAGCCACCCACAGUUCGAAAAGUGGA GCCACCCCCAGUUUGAGAAAUGGAGCCA UCCUCAAUUUGAAAAGACCACCAGCCAG AAGCACCGGGAUUUUGUUGCCGAGCCAA UGGGGGAAAAACCAGUAGGAUCACUGGC AGGGAUUGGUGAUGUUCUCUCAAAGAGA CUGGAAGAACGCGGAUUCGAUAAAGCAU ACGUCGUCUUGGGACAGUUCCUUGUUUU GAAGAAAGAUGAGGAUAGAUUCAGGGAA UGGUUAAAAGACACCUGCGGUGCAAAUG CGAAGCAGAGUAGAGACUGUUUCGGGUG UCUUCGAGAAUGGUGUGAUGCCUUCUUA UGGUCCCACCCCCAGUUUGAGAAAUGGU	MKFERQWSHPQFEKWSH PQFEKWSHPQFEKTTSQ KHRDFVAEPMGEKPVGS LAGIGDVLSKRLEERGF DKAYVVLGQFLVLKKDE DRFREWLKDTCGANAKQ SRDCFGCLREWCDAFLW SHPQFEKWSHPQFEKWS HPQFEKPAAKRVKLDAS QGTKRSYEQMETDGERQ

	GGGAATGGTTAAAAGACACCTGC GGTGCAAATGCGAAGCAGAGTAG AGACTGTTTCGGGTGTCTTCGAG AATGGTGTGATGCCTTCTTATGG TCCCACCCCCAGTTTGAGAAATG GTCTCATCCACAGTTTGAGAAAT GGTCCCATCCACAATTCGAGAAA CCCGCTGCAAAAAGGGTCAAGCT GGATGCAAGCCAGGGGACCAAAA GATCCTACGAACAAATGGAGACA GACGGAGAGAGACAGTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC AAGCTTTTCTTCTCTCCCTTGCA CCT	CUCAUCCACAGUUUGAGAAAUGGUCCCA UCCACAAUUCGAGAAACCCGCUGCAAAA AGGGUCAAGCUGGAUGCAAGCCAGGGGA CCAAAAGAUCCUACGAACAAAUGGAGAC AGACGGAGAGAGACAGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
4	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGAGCCATTGG CAACAGCAGAGTTACTTGGACAG TGGAATACACAGTGGGGCAACAA CCACAGCTCCTAGCCTTCCGGA AAGGGTAATCCGTGGAGCCATCC ACAGTTCGAAAAGTGGTCTCACC CTCAGTTTGAGAAGTGGTCCCAT CCTCAGTTTGAAAAAACCACGAG CCAAAAGCATAGGGACTTTGTAG CAGAACCCATGGGGGAAAAACCC GTTGGATCTTTAGCCGGAATAGG AGATGTGCTTTCAAAGAGATTAG AAGAACGCGGATTTGATAAAGCA TACGTCGTTCTGGGCCAATTTCT GGTTTTGAAGAAGGATGAAGATA GATTCAGAGAATGGTTGAAGGAC ACTTGCGGTGCCAACGCCAAACA GAGCAGGGATTGTTTCGGCTGCC TTAGAGAATGGTGTGATGCTTTC CTGTGGAGCCATCCCCAGTTCGA GAAATGGTCCCACCCCCAATTCG AAAAGTGGTCACATCCGCAGTTC GAGAAGCCCGCTGCAAAACGCGT TAAACTCGATGCCTCCCAGGGAA CAAAAAGATCTTATGAGCAGATG GAAACGGATGGAGAGCGGCAGTA AAGACCAGACAGCTGGAGCCTCG GTGGCCATGCTTCTTGCCCCTTG GGCCTCCCCCCAGCCCCTCCTCC CCTTCCTGCACCCGTACCCCCGT GGTCTTTGAATAAAGTCTGAGTG GGCGGCAAGCTTTTCTTCTCTCC CTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAGCCAUUGGCAACAG CAGAGUUACUUGGACAGUGGAAUACACA GUGGGGCAACAACCACAGCUCCUAGCCU UUCCGGAAAGGGUAAUCCGUGGAGCCAU CCACAGUUCGAAAAGUGGUCUCACCCUC AGUUUGAGAAGUGGUCCCAUCCUCAGUU UGAAAAAACCACGAGCCAAAAGCAUAGG GACUUUGUAGCAGAACCCAUGGGGGAAA AACCCGUUGGAUCUUUAGCCGGAAUAGG AGAUGUGCUUUCAAAGAGAUUAGAAGAA CGCGGAUUUGAUAAAGCAUACGUCGUUC UGGGCCAAUUUCUGGUUUUGAAGAAGGA UGAAGAUAGAUUCAGAGAAUGGUUGAAG GACACUUGCGGUGCCAACGCCAAACAGA GCAGGGAUUGUUUCGGCUGCCUUAGAGA AUGGUGUGAUGCUUUCCUGUGGAGCCAU CCCCAGUUCGAGAAAUGGUCCCACCCCC AAUUCGAAAAGUGGUCACAUCCGCAGUU CGAGAAGCCCGCUGCAAAACGCGUUAAA CUCGAUGCCUCCCAGGGAACAAAAAGAU CUUAUGAGCAGAUGGAAACGGAUGGAGA GCGGCAGUAAAGACCAGACAGCUGGAGC CUCGGUGGCCAUGCUUCUUGCCCCUUGG GCCUCCCCCCAGCCCCUCCUCCCCUUCC UGCACCCGUACCCCCGUGGUCUUUGAAU AAAGUCUGAGUGGGCGGCAAGCUUUUCU UCUCUCCCUUGCACCUAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAA	MSHWQQQSYLDSGIHSG ATTTAPSLSGKGNPWSH PQFEKWSHPQFEKWSHP QFEKTTSQKHRDFVAEP MGEKPVGSLAGIGDVLS KRLEERGFDKAYVVLGQ FLVLKKDEDRFREWLKD TCGANAKQSRDCFGCLR EWCDAFLWSHPQFEKWS HPQFEKWSHPQFEKPAA KRVKLDASQGTKRSYEQ METDGERQ

5	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MSENSGDSGYPSEKRGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	WSHPQFEKWSHPQFEKW
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGUCCGAGAACAGUGGU	SHPQFEKTTSQKHRDFV
	GCGGTAGACGCACGGCATAGCAT	GAUAGCGGUUACCCCAGUGAGAAAAGAG	AEPMGEKPVGSLAGIGD
	CCAACGCCACCATGTCCGAGAAC	GAGAGUGGAGUCACCCCCAGUUUGAAAA	VLSKRLEERGFDKAYVV
	AGTGGTGATAGCGGTTACCCCAG	AUGGAGUCACCCUCAGUUCGAAAAAUGG	LGQFLVLKKDEDRFREW
	TGAGAAAAGAGGAGAGTGGAGTC	UCACAUCCACAGUUUGAAAAGACAACCA	LKDTCGANAKQSRDCFG
	ACCCCCAGTTTGAAAAATGGAGT	GCCAGAAGCAUCGAGACUUUGUUGCCGA	CLREWCDAFLWSHPQFE
	CACCCTCAGTTCGAAAAATGGTC	GCCAAUGGGGGAAAAGCCGGUGGGAUCU	KWSHPQFEKWSHPQFEK
	ACATCCACAGTTTGAAAAGACAA	UUAGCAGGUAUAGGCGACGUGCUCAGUA	PAAKRVKLDASQGTKRS
	CCAGCCAGAAGCATCGAGACTTT	AACGACUCGAAGAGCGUGGAUUUGACAA	YEQMETDGERQ
	GTTGCCGAGCCAATGGGGGAAAA	AGCCUAUGUGGUCCUGGGCCAAUUCCUC	
	GCCGGTGGGATCTTTAGCAGGTA	GUUCUGAAGAAAGAUGAAGAUAGAUUCC	
	TAGGCGACGTGCTCAGTAAACGA	GGGAAUGGCUGAAAGAUACAUGCGGAGC	
	CTCGAAGAGCGTGGATTTGACAA	UAACGCUAAGCAGAGUAGGGAUUGCUUC	
	AGCCTATGTGGTCCTGGGCCAAT	GGAUGCCUUAGAGAGUGGUGCGACGCUU	
	TCCTCGTTCTGAAGAAAGATGAA	UUCUUUGGUCACAUCCCCAAUUUGAGAA	
	GATAGATTCCGGGAATGGCTGAA	AUGGUCCCACCCCCAAUUUGAGAAAUGG	
	AGATACATGCGGAGCTAACGCTA	AGCCAUCCUCAGUUCGAAAAGCCAGCUG	
	AGCAGAGTAGGGATTGCTTCGGA	CAAAGCGCGUUAAGUUGGAUGCUUCCCA	
	TGCCTTAGAGAGTGGTGCGACGC	AGGGACCAAGCGAUCCUACGAGCAGAUG	
	TTTTCTTTGGTCACATCCCCAAT	GAAACAGACGGAGAACGCCAGUAAAGAC	
	TTGAGAAATGGTCCCACCCCCAA	CAGACAGCUGGAGCCUCGGUGGCCAUGC	
	TTTGAGAAATGGAGCCATCCTCA	UUCUUGCCCCUUGGGCCUCCCCCCAGCC	
	GTTCGAAAAGCCAGCTGCAAAGC	CCUCCUCCCCUUCCUGCACCCGUACCCC	
	GCGTTAAGTTGGATGCTTCCCAA	CGUGGUCUUUGAAUAAAGUCUGAGUGGG	
	GGGACCAAGCGATCCTACGAGCA	CGGCAAGCUUUUCUUCUCUCCCUUGCAC	
	GATGGAAACAGACGGAGAACGCC	CUAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCT	AAAAAA	
6	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MAYISKIPHCTKEEDAP
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	SDNDSGICWSHPQFEKW
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGGCUUAUAUAUCCAAA	SHPQFEKWSHPQFEKTT
	GCGGTAGACGCACGGCATAGCAT	AUUCCACAUUGUACAAAGGAGGAAGAUG	SQKHRDFVAEPMGEKPV
	CCAACGCCACCATGGCTTATATA	CCCCUAGUGAUAAUGACUCUGGGAUCUG	GSLAGIGDVLSKRLEER
	TCCAAAATTCCACATTGTACAAA	UUGGUCACACCCACAGUUCGAAAAAUGG	GFDKAYVVLGQFLVLKK
	GGAGGAAGATGCCCCTAGTGATA	UCUCACCCCCAGUUUGAAAAAUGGUCAC	DEDRFREWLKDTCGANA
	ATGACTCTGGGATCTGTTGGTCA	ACCCACAAUUUGAAAAAACCACCUCCCA	KQSRDCFGCLREWCDAF
	CACCCACAGTTCGAAAAATGGTC	GAAACACAGAGACUUCGUAGCCGAGCCG	LWSHPQFEKWSHPQFEK
	TCACCCCCAGTTTGAAAAATGGT	AUGGGGGAGAAACCAGUUGGAAGUUUGG	WSHPQFEKPAAKRVKLD
	CACACCCACAATTTGAAAAAACC	CAGGCAUUGGAGAUGUGCUCUCUAAGAG	ASQGTKRSYEQMETDGE
	ACCTCCCAGAAACACAGAGACTT	GCUGGAGGAACGGGGAUUUGACAAAGCC	RQ
	CGTAGCCGAGCCGATGGGGGAGA	UACGUGGUAUUAGGGCAGUUCCUCGUAU	
	AACCAGTTGGAAGTTTGGCAGGC	UGAAGAAGGAUGAGGAUAGAUUUAGAGA	
	ATTGGAGATGTGCTCTCTAAGAG	GUGGCUGAAAGAUACAUGUGGAGCAAAC	
	GCTGGAGGAACGGGGATTTGACA	GCCAAACAGUCUAGAGAUUGCUUCGGGU	
	AAGCCTACGTGGTATTAGGGCAG	GCUUGCGAGAGUGGUGCGACGCAUUCCU	
	TTCCTCGTATTGAAGAAGGATGA	GUGGUCCCAUCCGCAAUUCGAGAAGUGG	
	GGATAGATTTAGAGAGTGGCTGA	AGCCAUCCACAGUUCGAGAAGUGGAGUC	
	AAGATACATGTGGAGCAAACGCC	ACCCACAAUUUGAAAAGCCCGCCGCAAA	
	AAACAGTCTAGAGATTGCTTCGG	AAGGGUUAAAUUAGACGCCAGUCAGGGG	
	GTGCTTGCGAGAGTGGTGCGACG	ACCAAGAGAAGCUACGAACAGAUGGAGA	

	CATTCCTGTGGTCCCATCCGCAA TTCGAGAAGTGGAGCCATCCACA GTTCGAGAAGTGGAGTCACCCAC AATTTGAAAAGCCCGCCGCAAAA AGGGTTAAATTAGACGCCAGTCA GGGGACCAAGAGAAGCTACGAAC AGATGGAGACAGATGGTGAAAGA CAATAAAGACCAGACAGCTGGAG CCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACC CCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCAAGCTTTTCTTC TCTCCCTTGCACCT	CAGAUGGUGAAAGACAAUAAAGACCAGA CAGCUGGAGCCUCGGUGGCCAUGCUUCU UGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGUACCCCCGUG GUCUUUGAAUAAAGUCUGAGUGGGCGGC AAGCUUUUCUUCUCUCCCUUGCACCUAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA	
7	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGCCAGGACAT CTTCCAAGCGAAGACTCTGGGTA TGGTATGGAAATGCCTGCGGACA AGGGTCCTTCCTGGTCACATCCA CAGTTTGAAAAATGGAGTCACCC TCAGTTTGAGAAATGGTCACATC CTCAATTTGAGAAGACGACTTCA CAAAAGCATCGGGATTTCGTGGC AGAACCAATGGGGGAGAAGCCCG TAGGCTCCTTAGCTGGAATTGGA GACGTACTTAGTAAGAGACTTGA AGAGAGAGGATTCGATAAAGCAT ATGTCGTACTTGGACAATTCTTA GTCCTCAAAAAAGATGAAGATAG ATTTCGCGAGTGGCTGAAAGATA CATGTGGGGCCAATGCCAAACAG AGCAGAGATTGTTTTGGGTGCCT GAGGGAATGGTGCGACGCCTTCC TTTGGAGCCATCCTCAATTTGAG AAGTGGAGCCACCCTCAGTTCGA AAAATGGTCCCACCCCCAGTTTG AGAAACCAGCTGCCAAACGCGTT AAACTGGATGCATCCCAGGGTAC TAAGCGCAGCTATGAACAGATGG AAACGGACGGGGAGCGCCAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCCAGGACAUCUUCCA AGCGAAGACUCUGGGUAUGGUAUGGAAA UGCCUGCGGACAAGGGUCCUUCCUGGUC ACAUCCACAGUUUGAAAAAUGGAGUCAC CCUCAGUUUGAGAAAUGGUCACAUCCUC AAUUUGAGAAGACGACUUCACAAAAGCA UCGGGAUUUCGUGGCAGAACCAAUGGGG GAGAAGCCCGUAGGCUCCUUAGCUGGAA UUGGAGACGUACUUAGUAAGAGACUUGA AGAGAGAGGAUUCGAUAAAGCAUAUGUC GUACUUGGACAAUUCUUAGUCCUCAAAA AAGAUGAAGAUAGAUUUCGCGAGUGGCU GAAAGAUACAUGUGGGGCCAAUGCCAAA CAGAGCAGAGAUUGUUUUGGGUGCCUGA GGGAAUGGUGCGACGCCUUCCUUUGGAG CCAUCCUCAAUUUGAGAAGUGGAGCCAC CCUCAGUUCGAAAAAUGGUCCCACCCCC AGUUUGAGAAACCAGCUGCCAAACGCGU UAAACUGGAUGCAUCCCAGGGUACUAAG CGCAGCUAUGAACAGAUGGAAACGGACG GGGAGCGCCAAUAAAGACCAGACAGCUG GAGCCUCGGUGGCCAUGCUUCUUGCCCC UUGGGCCUCCCCCCAGCCCCUCCUCCCC UUCCUGCACCCGUACCCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGCAAGCUU UUCUUCUCUCCCUUGCACCUAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	MPGHLPSEDSGYGMEMP ADKGPSWSHPQFEKWSH PQFEKWSHPQFEKTTSQ KHRDFVAEPMGEKPVGS LAGIGDVLSKRLEERGF DKAYVVLGQFLVLKKDE DRFREWLKDTCGANAKQ SRDCFGCLREWCDAFLW SHPQFEKWSHPQFEKWS HPQFEKPAAKRVKLDAS QGTKRSYEQMETDGERQ

8	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGCCAGACCGG AAAGCCGCAGTGTCTCACTGGCA GCAACAAAGCTACTTAGATTCTG GGATTCACTCCGGGGCCTGGTCA CACCCACAGTTCGAAAAATGGTC CCATCCGCAGTTTGAGAAATGGA GTCACCCACAGTTTGAAAAAACC ACTTCACAGAAGCATCGCGACTT TGTGGCAGAACCAATGGGAGAGA AGCCCGTCGGCTCCTTGGCAGGT ATAGGCGACGTGCTGTCCAAAAG ACTCGAGGAGAGAGGATTCGATA AGGCGTATGTTGTTCTGGGACAG TTTTTAGTCTTAAAGAAAGATGA GGATCGCTTCAGGGAATGGCTTA AGGACACCTGTGGCGCAAACGCG AAGCAGAGCAGGGATTGCTTTGG CTGCCTTCGGGAATGGTGTGATG CGTTTCTGTGGTCCCACCCACAG TTCGAGAAGTGGTCTCACCCACA ATTCGAAAAGTGGAGTCACCCTC AGTTCGAAAAGCCTGCCGCTAAG AGGGTTAAACTCGACGCCAGCCA AGGCACTAAGAGGTCTTACGAGC AGATGGAAACCGATGGTGAGCGC CAATAAAGACCAGACAGCTGGAG CCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACC CCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCAAGCTTTTCTTC TCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCCAGACCGGAAAGCC GCAGUGUCUCACUGGCAGCAACAAAGCU ACUUAGAUUCUGGGAUUCACUCCGGGGC CUGGUCACACCCACAGUUCGAAAAAUGG UCCCAUCCGCAGUUUGAGAAAUGGAGUC ACCCACAGUUUGAAAAAACCACUUCACA GAAGCAUCGCGACUUUGUGGCAGAACCA AUGGGAGAGAAGCCCGUCGGCUCCUUGG CAGGUAUAGGCGACGUGCUGUCCAAAAG ACUCGAGGAGAGAGGAUUCGAUAAGGCG UAUGUUGUUCUGGGACAGUUUUUAGUCU UAAAGAAAGAUGAGGAUCGCUUCAGGGA AUGGCUUAAGGACACCUGUGGCGCAAAC GCGAAGCAGAGCAGGGAUUGCUUUGGCU GCCUUCGGGAAUGGUGUGAUGCGUUUCU GUGGUCCCACCCACAGUUCGAGAAGUGG UCUCACCCACAAUUCGAAAAGUGGAGUC ACCCUCAGUUCGAAAAGCCUGCCGCUAA GAGGGUUAAACUCGACGCCAGCCAAGGC ACUAAGAGGUCUUACGAGCAGAUGGAAA CCGAUGGUGAGCGCCAAUAAAGACCAGA CAGCUGGAGCCUCGGUGGCCAUGCUUCU UGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGUACCCCCGUG GUCUUUGAAUAAAGUCUGAGUGGGCGGC AAGCUUUUCUUCUCUCCCUUGCACCUAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA	
9	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGTCTGGGGAC AGTGGTTACCCCAGCGAGAAGCG AGGCGAGTTAGATGATTGGAGCC ATCCCCAATTTGAAAAATGGAGT CATCCACAGTTTGAGAAATGGTC ACATCCACAGTTCGAAAAAACAA CTTCCCAAAAGCATCGCGACTTC GTTGCCGAACCCATGGGCGAAAA GCCTGTAGGCTCTTTGGCAGGAA TCGGCGACGTATTAAGCAAAAGG TTAGAAGAAAGAGGGTTTGACAA GGCATATGTAGTCCTGGGGCAAT TTTTGGTGCTGAAAAAAGACGAA GATAGATTTCGAGAGTGGCTTAA GGATACGTGCGGCGCCAATGCCA AGCAGTCACGCGACTGTTTTGGG TGCCTGCGTGAGTGGTGCGATGC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUCUGGGGACAGUGGU UACCCCAGCGAGAAGCGAGGCGAGUUAG AUGAUUGGAGCCAUCCCCAAUUUGAAAA AUGGAGUCAUCCACAGUUUGAGAAAUGG UCACAUCCACAGUUCGAAAAAACAACUU CCCAAAAGCAUCGCGACUUCGUUGCCGA ACCCAUGGGCGAAAAGCCUGUAGGCUCU UUGGCAGGAAUCGGCGACGUAUUAAGCA AAAGGUUAGAAGAAAGAGGGUUUGACAA GGCAUAUGUAGUCCUGGGGCAAUUUUUG GUGCUGAAAAAAGACGAAGAUAGAUUUC GAGAGUGGCUUAAGGAUACGUGCGGCGC CAAUGCCAAGCAGUCACGCGACUGUUUU GGGUGCCUGCGUGAGUGGUGCGAUGCCU UUCUGUGGUCCCACCCUCAGUUUGAAAA GUGGUCUCACCCCCAAUUCGAGAAGUGG AGUCAUCCCCAGUUCGAGAAGCCCGCCG CCAAAAGAGUGAAAUUGGACGCCUCCCA GGGGACGAAGCGCAGUUACGAACAGAUG	MSGDSGYPSEKRGELDD WSHPQFEKWSHPQFEKW SHPQFEKTTSQKHRDFV AEPMGEKPVGSLAGIGD VLSKRLEERGFDKAYVV LGQFLVLKKDEDRFREW LKDTCGANAKQSRDCFG CLREWCDAFLWSHPQFE KWSHPQFEKWSHPQFEK PAAKRVKLDASQGTKRS YEQMETDGERQ

	CTTTCTGTGGTCCCACCCTCAGT	GAGACUGACGGAGAGCGGCAGUAAAGAC	
	TTGAAAAGTGGTCTCACCCCCAA	CAGACAGCUGGAGCCUCGGUGGCCAUGC	
	TTCGAGAAGTGGAGTCATCCCCA	UUCUUGCCCCUUGGGCCUCCCCCCAGCC	
	GTTCGAGAAGCCCGCCGCCAAAA	CCUCCUCCCCUUCCUGCACCCGUACCCC	
	GAGTGAAATTGGACGCCTCCCAG	CGUGGUCUUUGAAUAAAGUCUGAGUGGG	
	GGGACGAAGCGCAGTTACGAACA	CGGCAAGCUUUUCUUCUCUCCCUUGCAC	
	GATGGAGACTGACGGAGAGCGGC	CUAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCT	AAAAAA	
10	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MPDRKAAVSHWQQQSYL
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	DSGIHSGATTTAPSLSG
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGCCCGAUCGCAAAGCG	KGNPEEWSHPQFEKWSH
	GCGGTAGACGCACGGCATAGCAT	GCCGUGAGUCAUUGGCAGCAACAAUCUU	PQFEKWSHPQFEKTTSQ
	CCAACGCCACCATGCCCGATCGC	AUCUUGACAGCGGCAUUCAUUCCGGAGC	KHRDFVAEPMGEKPVGS
	AAAGCGGCCGTGAGTCATTGGCA	UACAACCACCGCACCCAGUUUAUCAGGG	LAGIGDVLSKRLEERGF
	GCAACAATCTTATCTTGACAGCG	AAAGGCAACCCGGAGGAGUGGUCACAUC	DKAYVVLGQFLVLKKDE
	GCATTCATTCCGGAGCTACAACC	CCCAGUUUGAGAAGUGGUCUCACCCUCA	DRFREWLKDTCGANAKQ
	ACCGCACCCAGTTTATCAGGGAA	GUUUGAAAAAUGGUCCCAUCCCCAAUUU	SRDCFGCLREWCDAFLW
	AGGCAACCCGGAGGAGTGGTCAC	GAGAAAACUACUAGUCAGAAGCAUAGGG	SHPQFEKWSHPQFEKWS
	ATCCCCAGTTTGAGAAGTGGTCT	AUUUUGUCGCCGAGCCUAUGGGAGAAAA	HPQFEKPAAKRVKLDAS
	CACCCTCAGTTTGAAAAATGGTC	ACCCGUUGGAAGCCUCGCAGGCAUAGGC	QGTKRSYEQMETDGERQ
	CCATCCCCAATTTGAGAAAACTA	GAUGUGCUCUCUAAAAGAUUGGAAGAGC	
	CTAGTCAGAAGCATAGGGATTTT	GCGGGUUCGACAAAGCGUACGUAGUCUU	
	GTCGCCGAGCCTATGGGAGAAAA	GGGGCAGUUCUUGGUACUCAAAAAAGAC	
	ACCCGTTGGAAGCCTCGCAGGCA	GAGGAUAGGUUUCGCGAAUGGUUGAAGG	
	TAGGCGATGTGCTCTCTAAAAGA	AUACUUGUGGUGCCAACGCCAAACAAUC	
	TTGGAAGAGCGCGGGTTCGACAA	CCGUGAUUGCUUUGGAUGUUUACGGGAA	
	AGCGTACGTAGTCTTGGGGCAGT	UGGUGCGACGCAUUUCUCUGGUCUCAUC	
	TCTTGGTACTCAAAAAAGACGAG	CGCAAUUUGAGAAAUGGAGUCACCCUCA	
	GATAGGTTTCGCGAATGGTTGAA	GUUCGAGAAGUGGAGUCAUCCUCAGUUC	
	GGATACTTGTGGTGCCAACGCCA	GAAAAGCCCGCAGCUAAGCGGGUCAAAC	
	AACAATCCCGTGATTGCTTTGGA	UCGAUGCAUCCCAGGGAACUAAGCGCAG	
	TGTTTACGGGAATGGTGCGACGC	CUAUGAACAGAUGGAGACGGACGGCGAA	
	ATTTCTCTGGTCTCATCCGCAAT	CGACAGUAAAGACCAGACAGCUGGAGCC	
	TTGAGAAATGGAGTCACCCTCAG	UCGGUGGCCAUGCUUCUUGCCCCUUGGG	
	TTCGAGAAGTGGAGTCATCCTCA	CCUCCCCCCAGCCCCUCCUCCCCUUCCU	
	GTTCGAAAAGCCCGCAGCTAAGC	GCACCCGUACCCCCGUGGUCUUUGAAUA	
	GGGTCAAACTCGATGCATCCCAG	AAGUCUGAGUGGGCGGCAAGCUUUUCUU	
	GGAACTAAGCGCAGCTATGAACA	CUCUCCCUUGCACCUAAAAAAAAAAAAA	
	GATGGAGACGGACGGCGAACGAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT	AAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCT		

11	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGATAAAGCT AGGGCAAAAAAGGCCATGCTTAG CGAACAGAATAGGGCGTCACCAT TACCTTCTGGGTTGCTTACCCCT CCTCAAAGCGGAAAGAAGCAAAG CTCCGGCTGGTCTCACCCGCAGT TTGAAAAGTGGTCCCATCCACAG TTTGAGAAATGGTCCCATCCTCA GTTTGAGAAAACCACGTCACAAA AACACCGGGATTTTGTGGCCGAA CCAATGGGTGAAAAGCCTGTGGG CTCCCTTGCTGGTATCGGAGATG TCCTTTCTAAGCGGCTGGAAGAG CGAGGATTTGATAAAGCTTATGT GGTTCTCGGACAGTTTCTGGTTC TTAAAAAAGACGAGGACAGGTTT CGTGAGTGGCTGAAGGATACTTG TGGTGCTAACGCCAAACAGAGTA GAGACTGTTTCGGATGCTTAAGA GAATGGTGCGACGCATTTTTATG GTCACATCCACAATTCGAGAAAT GGAGTCACCCTCAATTTGAGAAA TGGTCTCATCCCCAATTTGAAAA ACCCGCTGCCAAACGAGTGAAAC TGGATGCCTCCCAGGGGACAAAG CGGAGTTATGAACAGATGGAGAC AGACGGAGAGCGGCAATAAAGAC CAGACAGCTGGAGCCTCGGTGGC CATGCTTCTTGCCCCCTGGGCCT CCCCCCAGCCCCTCCTCCCCTTC CTGCACCCGTACCCCCGTGGTCT TTGAATAAAGTCTGAGTGGGCGG CAAGCTTTTCTTCTCTCCCTTGC ACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGAUAAAGCUAGGGCA AAAAAGGCCAUGCUUAGCGAACAGAAUA GGGCGUCACCAUUACCUUCUGGGUUGCU UACCCCUCCUCAAAGCGGAAAGAAGCAA AGCUCCGGCUGGUCUCACCCGCAGUUUG AAAAGUGGUCCCAUCCACAGUUUGAGAA AUGGUCCCAUCCUCAGUUUGAGAAAACC ACGUCACAAAAACACCGGGAUUUUGUGG CCGAACCAAUGGGUGAAAAGCCUGUGGG CUCCCUUGCUGGUAUCGGAGAUGUCCUU UCUAAGCGGCUGGAAGAGCGAGGAUUUG AUAAAGCUUAUGUGGUUCUCGGACAGUU UCUGGUUCUUAAAAAAGACGAGGACAGG UUUCGUGAGUGGCUGAAGGAUACUUGUG GUGCUAACGCCAAACAGAGUAGAGACUG UUUCGGAUGCUUAAGAGAAUGGUGCGAC GCAUUUUUAUGGUCACAUCCACAAUUCG AGAAAUGGAGUCACCCUCAAUUUGAGAA AUGGUCUCAUCCCCAAUUUGAAAAACCC GCUGCCAAACGAGUGAAACUGGAUGCCU CCCAGGGGACAAAGCGGAGUUAUGAACA GAUGGAGACAGACGGAGAGCGGCAAUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MDKARAKKAMLSEQNRA SPLPSGLLTPPQSGKKQ SSGWSHPQFEKWSHPQF EKWSHPQFEKTTSQKHR DFVAEPMGEKPVGSLAG IGDVLSKRLEERGFDKA YVVLGQFLVLKKDEDRF REWLKDTCGANAKQSRD CFGCLREWCDAFLWSHP QFEKWSHPQFEKWSHPQ FEKPAAKRVKLDASQGT KRSYEQMETDGERQ
12	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGGGCTCCTG ACTCCCCCAGCCTCACCACTGGA GCTCCTTGAAGCTAAACCCAAGC GCGGGAGAAGGTGGTCACACCCA CAGTTCGAAAAATGGTCTCACCC ACAGTTTGAGAAATGGTCTCACC CCCAGTTCGAGAAAACCACCTCT CAGAAGCACAGGGACTTCGTGGC TGAACCCATGGGCGAAAAACCCG TCGGAAGCCTCGCAGGAATCGGC GACGTGCTGTCAAAACGCTTGGA GGAACGAGGATTTGACAAGGCTT ATGTCGTACTTGGTCAGTTTTTG GTACTTAAGAAGGATGAAGATAG ATTTCGCGAATGGCTCAAAGATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGGGCUCCUGACUCCC CCAGCCUCACCACUGGAGCUCCUUGAAG CUAAACCCAAGCGCGGGAGAAGGUGGUC ACACCCACAGUUCGAAAAAUGGUCUCAC CCACAGUUUGAGAAAUGGUCUCACCCCC AGUUCGAGAAAACCACCUCUCAGAAGCA CAGGGACUUCGUGGCUGAACCCAUGGGC GAAAAACCCGUCGGAAGCCUCGCAGGAA UCGGCGACGUGCUGUCAAAACGCUUGGA GGAACGAGGAUUUGACAAGGCUUAUGUC GUACUUGGUCAGUUUUUGGUACUUAAGA AGGAUGAAGAUAGAUUUCGCGAAUGGCU CAAAGAUACUUGUGGAGCAAACGCUAAA CAAUCCCGGGAUUGUUUUGGAUGCCUGA GAGAAUGGUGUGACGCAUUUCUGUGGUC ACAUCCCCAAUUUGAGAAGUGGUCUCAC CCUCAAUUUGAAAAGUGGAGUCACCCUC	MGLLTPPASPLELLEAK PKRGRRWSHPQFEKWSH PQFEKWSHPQFEKTTSQ KHRDFVAEPMGEKPVGS LAGIGDVLSKRLEERGF DKAYVVLGQFLVLKKDE DRFREWLKDTCGANAKQ SRDCFGCLREWCDAFLW SHPQFEKWSHPQFEKWS HPQFEKPAAKRVKLDAS QGTKRSYEQMETDGERQ

	CTTGTGGAGCAAACGCTAAACAA TCCCGGGATTGTTTTGGATGCCT GAGAGAATGGTGTGACGCATTTC TGTGGTCACATCCCCAATTTGAG AAGTGGTCTCACCCTCAATTTGA AAAGTGGAGTCACCCTCAGTTTG AAAAGCCAGCAGCAAAGCGAGTC AAGTTGGACGCATCACAGGGGAC TAAAAGGTCATATGAGCAGATGG AAACCGACGGGGAAAGACAGTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AGUUUGAAAAGCCAGCAGCAAAGCGAGU CAAGUUGGACGCAUCACAGGGGACUAAA AGGUCAUAUGAGCAGAUGGAAACCGACG GGGAAAGACAGUAAAGACCAGACAGCUG GAGCCUCGGUGGCCAUGCUUCUUGCCCC UUGGGCCUCCCCCCAGCCCCUCCUCCCC UUCCUGCACCCGUACCCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGCAAGCUU UUCUUCUCUCCCUUGCACCUAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	
13	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGCAGAGCCAG GGCCCACCAACCCCTCCAACAAC ACCAAAAACAGATGTGCAGCCAG GGAAACCCGATCTGAAGCGCGAG GGATGGTCCCACCCACAGTTTGA AAAATGGAGTCACCCCCAGTTCG AGAAATGGTCCCACCCCCAATTC GAAAAGACCACCAGTCAGAAACA CCGGGACTTCGTGGCAGAGCCTA TGGGGGAAAAACCAGTGGGCTCA TTAGCTGGCATCGGCGATGTTCT GAGTAAAAGGCTTGAAGAGCGCG GATTTGACAAGGCCTACGTAGTG CTCGGACAATTTCTTGTGCTCAA AAAGGACGAAGACAGATTTAGAG AATGGTTGAAAGATACTTGTGGT GCGAATGCCAAGCAGTCAAGGGA TTGTTTCGGCTGCCTCCGGGAGT GGTGCGACGCCTTCTTGTGGAGT CACCCCCAATTCGAGAAGTGGAG TCATCCCCAGTTTGAAAAGTGGA GCCATCCACAGTTCGAAAAACCT GCCGCCAAGAGAGTGAAGCTGGA CGCAAGCCAAGGCACAAAGAGGA GCTACGAGCAAATGGAAACAGAT GGTGAGAGACAGTAAAGACCAGA CAGCTGGAGCCTCGGTGGCCATG CTTCTTGCCCCTTGGGCCTCCCC CCAGCCCCTCCTCCCCTTCCTGC ACCCGTACCCCCGTGGTCTTTGA ATAAAGTCTGAGTGGGCGGCAAG CTTTTCTTCTCTCCCTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGAGCCAGGGCCCA CCAACCCCUCCAACAACACCAAAAACAG AUGUGCAGCCAGGGAAACCCGAUCUGAA GCGCGAGGGAUGGUCCCACCCACAGUUU GAAAAAUGGAGUCACCCCCAGUUCGAGA AAUGGUCCCACCCCCAAUUCGAAAAGAC CACCAGUCAGAAACACCGGGACUUCGUG GCAGAGCCUAUGGGGGAAAAACCAGUGG GCUCAUUAGCUGGCAUCGGCGAUGUUCU GAGUAAAAGGCUUGAAGAGCGCGGAUUU GACAAGGCCUACGUAGUGCUCGGACAAU UUCUUGUGCUCAAAAAGGACGAAGACAG AUUUAGAGAAUGGUUGAAAGAUACUUGU GGUGCGAAUGCCAAGCAGUCAAGGGAUU GUUUCGGCUGCCUCCGGGAGUGGUGCGA CGCCUUCUUGUGGAGUCACCCCCAAUUC GAGAAGUGGAGUCAUCCCCAGUUUGAAA AGUGGAGCCAUCCACAGUUCGAAAAACC UGCCGCCAAGAGAGUGAAGCUGGACGCA AGCCAAGGCACAAAGAGGAGCUACGAGC AAAUGGAAACAGAUGGUGAGAGACAGUA AAGACCAGACAGCUGGAGCCUCGGUGGC CAUGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUGA GUGGGCGGCAAGCUUUUCUUCUCUCCCU UGCACCUAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MQSQGPPTPPTTPKTDV QPGKPDLKREGWSHPQF EKWSHPQFEKWSHPQFE KTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

14	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTHGMMLTPPSSPLELM
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PPGSCMPEEPKPKRGRW
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCCAUGGAAUGAUG	SHPQFEKWSHPQFEKWS
	GCGGTAGACGCACGGCATAGCAT	UUAACCCCACCAUCAAGCCCCCUCGAAU	HPQFEKTTSQKHRDFVA
	CCAACGCCACCATGACCCATGGA	UAAUGCCACCUGGCUCCUGUAUGCCUGA	EPMGEKPVGSLAGIGDV
	ATGATGTTAACCCCACCATCAAG	AGAACCCAAACCUAAAAGGGGCAGGUGG	LSKRLEERGFDKAYVVL
	CCCCCTCGAATTAATGCCACCTG	AGUCACCCUCAGUUUGAGAAAUGGUCUC	GQFLVLKKDEDRFREWL
	GCTCCTGTATGCCTGAAGAACCC	AUCCCCAAUUUGAGAAGUGGUCCCAUCC	KDTCGANAKQSRDCFGC
	AAACCTAAAAGGGGCAGGTGGAG	CCAGUUUGAGAAGACAACGAGUCAGAAG	LREWCDAFLWSHPQFEK
	TCACCCTCAGTTTGAGAAATGGT	CAUCGGGACUUUGUAGCAGAACCAAUGG	WSHPQFEKWSHPQFEKP
	CTCATCCCCAATTTGAGAAGTGG	GCGAAAAACCCGUCGGCUCUCUGGCCGG	AAKRVKLDASQGTKRSY
	TCCCATCCCCAGTTTGAGAAGAC	AAUAGGCGAUGUCCUGUCUAAGCGUUUG	EQMETDGERQ
	AACGAGTCAGAAGCACCGGGACT	GAAGAACGGGGCUUCGAUAAAGCUUAUG	
	TTGTAGCAGAACCAACGGGCGAA	UCGUGCUGGGCCAGUUUUUGGUGCUCAA	
	AAACCCGTCGGCTCTCTGGCCGG	AAAAGACGAGGACCGUUUCCGGGAAUGG	
	AATAGGCGATGTCCTGTCTAAGC	UUGAAAGAUACUUGUGGAGCCAAUGCAA	
	GTTTGGAAGAACGGGGCTTCGAT	AGCAAAGCAGGGACUGCUUUGGUUGUCU	
	AAAGCTTATGTCGTGCTGGGCCA	GCGCGAGUGGUGCGAUGCCUUUUUAUGG	
	GTTTTTGGTGCTCAAAAAAGACG	AGCCACCCUCAGUUUGAGAAAUGGUCUC	
	AGGACCGTTTCCGGGAATGGTTG	AUCCCCAGUUCGAGAAAUGGUCACACCC	
	AAAGATACTTGTGGAGCCAATGC	UCAAUUUGAAAAACCAGCUGCCAAACGA	
	AAAGCAAAGCAGGGACTGCTTTG	GUGAAGUUGGACGCUUCUCAGGGCACAA	
	GTTGTCTGCGCGAGTGGTGCGAT	AACGAUCUUACGAACAAAUGGAGACUGA	
	GCCTTTTTATGGAGCCACCCTCA	UGGUGAAAGACAGUAAAGACCAGACAGC	
	GTTTGAGAAATGGTCTCATCCCC	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGTTCGAGAAATGGTCACACCCT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	CAATTTGAAAAACCAGCTGCCAA	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	ACGAGTGAAGTTGGACGCTTCTC	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGGGCACAAAACGATCTTACGAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	CAAATGGAGACTGATGGTGAAAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACAGTAAAGACCAGACAGCTGGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCGGTGGCCATGCTTCTTGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCTTGGGCCTCCCCCCAGCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCTCCCCTTCCTGCACCCGTAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCGTGGTCTTTGAATAAAGTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGAGTGGGCGGCAAGCTTTTCTT	AAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCTCCCTTGCACCT		
15	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MYDRKPPSGFKPLTPPT
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	WSHPQFEKWSHPQFEKW
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGUAUGAUCGCAAACCU	SHPQFEKTTSQKHRDFV
	GCGGTAGACGCACGGCATAGCAT	CCUAGUGGGUUCAAGCCUCUCACUCCUC	AEPMGEKPVGSLAGIGD
	CCAACGCCACCATGTATGATCGC	CUACAUGGAGUCACCCUCAGUUUGAGAA	VLSKRLEERGFDKAYVV
	AAACCTCCTAGTGGGTTCAAGCC	GUGGAGUCACCCGCAGUUCGAAAAGUGG	LGQFLVLKKDEDRFREW
	TCTCACTCCTCCTACATGGAGTC	UCACACCCUCAGUUCGAGAAAACUACAA	LKDTCGANAKQSRDCFG
	ACCCTCAGTTTGAGAAGTGGAGT	GUCAGAAGCAUAGGGAUUUUGUUGCCGA	CLREWCDAFLWSHPQFE
	CACCCGCAGTTCGAAAAGTGGTC	ACCGAUGGGGGAAAAGCCUGUCGGGAGC	KWSHPQFEKWSHPQFEK
	ACACCCTCAGTTCGAGAAAACTA	CUCGCAGGCAUAGGGGACGUAUUAAGUA	PAAKRVKLDASQGTKRS
	CAAGTCAGAAGCATAGGGATTTT	AGCGCCUCGAAGAGAGGGGUUUCGACAA	YEQMETDGERQ
	GTTGCCGAACCGATGGGGGAAAA	GGCAUAUGUGGUACUUGGACAGUUUUUG	
	GCCTGTCGGGAGCCTCGCAGGCA	GUGCUGAAAAAAGACGAGGAUCGUUUUC	
	TAGGGGACGTATTAAGTAAGCGC	GUGAAUGGUUGAAAGAUACAUGUGGGGC	
	CTCGAAGAGAGGGGTTTCGACAA	AAAUGCUAAACAGAGCAGGGAUUGUUUU	
	GGCATATGTGGTACTTGGACAGT	GGUUGCCUCCGGGAGUGGUGCGACGCCU	
	TTTTGGTGCTGAAAAAAGACGAG	UCCUUUGGUCUCACCCACAGUUCGAGAA	
	GATCGTTTTCGTGAATGGTTGAA	GUGGUCACAUCCCCAGUUUGAGAAAUGG	
	AGATACATGTGGGGCAAATGCTA	AGCCAUCCACAGUUCGAGAAACCAGCAG	
	AACAGAGCAGGGATTGTTTTGGT	CCAAAAGGGUUAAAUUGGACGCCUCUCA	

	TGCCTCCGGGAGTGGTGCGACGC	GGGAACCAAAAGAAGCUAUGAGCAGAUG	
	CTTCCTTTGGTCTCACCCACAGT	GAAACCGAUGGGGAGCGUCAAUAAAGAC	
	TCGAGAAGTGGTCACATCCCCAG	CAGACAGCUGGAGCCUCGGUGGCCAUGC	
	TTTGAGAAATGGAGCCATCCACA	UUCUUGCCCCUUGGGCCUCCCCCCAGCC	
	GTTCGAGAAACCAGCAGCCAAAA	CCUCCUCCCCUUCCUGCACCCGUACCCC	
	GGGTTAAATTGGACGCCTCTCAG	CGUGGUCUUUGAAUAAAGUCUGAGUGGG	
	GGAACCAAAAGAAGCCATGAGCA	CGGCAAGCUUUUCUUCUCUCCCUUGCAC	
	GATGGAAACCGATGGGGAGCGTC	CUAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCT	AAAAAA	
16	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACCUCCCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGAUUUUGUUGCAGAACCAAUGG	LRKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACCTCC	GAGAAAAACCUGUGGGCUCUCUCGCGGG	ANAKQSRDCFGCLREWC
	CAGAAGCACAGAGATTTGTTGC	CAUCGGAGACGUCUUGUCCAAAAGGCUC	DAFLWSHPQFEKWSHPQ
	AGAACCAATGGGAGAAAAACCTG	GAAGAAAGGGGAUUUGAUAAAGCAUACG	FEKWSHPQFEKPAAKRV
	TGGGCTCTCTCGCGGGCATCGGA	UGGUUCUGGGCCAGUUUCUUGUGUUGAG	KLDASQGTKRSYEQMET
	GACGTCTTGTCCAAAAGGCTCGA	AAAAGAUGAAGAUCGCUUUAGGGAGUGG	DGERQ
	AGAAAGGGGATTTGATAAAGCAT	CUGAAGGACACUUGCGGGGCCAAUGCCA	
	ACGTGGTTCTGGGCCAGTTTCTT	AGCAAUCUAGAGACUGCUUCGGGUGUUU	
	GTGTTGAGAAAAGATGAAGATCG	AAGGGAGUGGUGCGAUGCCUUUUUGUGG	
	CTTTAGGGAGTGGCTGAAGGACA	UCUCAUCCACAGUUUGAAAAAUGGUCAC	
	CTTGCGGGGCCAATGCCAAGCAA	ACCCCCAAUUCGAAAAGUGGAGUCACCC	
	TCTAGAGACTGCTTCGGGTGTTT	UCAGUUUGAAAAGCCAGCCGCCAAGAGA	
	AAGGGAGTGGTGCGATGCCTTTT	GUCAAAUUGGACGCAUCCCAAGGAACAA	
	TGTGGTCTCATCCACAGTTTGAA	AACGUUCUUAUGAACAGAUGGAGACUGA	
	AAATGGTCACACCCCCAATTCGA	CGGAGAAAGACAGUAAAGACCAGACAGC	
	AAAGTGGAGTCACCCTCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCAGCCGCCAAGAGAGTC	CCUUGGGCCUCCCCCCCAGCCCCUCCUCC	
	AAATTGGACGCATCCCAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAACGTTCTTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACTGACGGAGAAAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAA	
17	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDILSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAUCUCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGAUUUCGUAGCAGAGCCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACATCT	GCGAAAAACCAGUGGGUUCACUGGCUGG	ANAKQSRDCFGCLREWC
	CAAAAACACAGGGATTTCGTAGC	UAUUGGGGAUAUCCUCAGCAAAAGGUUA	DAFLWSHPQFEKWSHPQ
	AGAGCCAATGGGCGAAAAACCAG	GAAGAACGCGGAUUCGAUAAGGCAUACG	FEKWSHPQFEKPAAKRV
	TGGGTTCACTGGCTGGTATTGGG	UGGUGCUGGGCCAGUUCCUGGUCCUCAA	KLDASQGTKRSYEQMET
	GATATCCTCAGCAAAAGGTTAGA	AAAGGAUGAAGAUAGAUUCCGGGAAUGG	DGERQ
	AGAACGCGGATTCGATAAGGCAT	CUGAAGGAUACAUGCGGAGCCAAUGCCA	
	ACGTGGTGCTGGGCCAGTTCCTG	AACAGAGCCGGGACUGCUUCGGUUGCCU	
	GTCCTCAAAAAGGATGAAGATAG	UAGGGAGUGGUGCGACGCAUUCCUCUGG	

	ATTCCGGGAATGGCTGAAGGATA	AGCCACCCUCAAUUUGAAAAGUGGUCUC	
	CATGCGGAGCCAATGCCAAACAG	AUCCCCAGUUCGAAAAGUGGAGUCACCC	
	AGCCGGGACTGCTTCGGTTGCCT	CCAGUUUGAAAAACCUGCUGCUAAGCGA	
	TAGGGAGTGGTGCGACGCATTCC	GUGAAGUUAGAUGCAUCACAGGGCACCA	
	TCTGGAGCCACCCTCAATTTGAA	AGAGAUCUUACGAACAGAUGGAGACCGA	
	AAGTGGTCTCATCCCCAGTTCGA	UGGAGAGCGCCAGUAAAGACCAGACAGC	
	AAAGTGGAGTCACCCCCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCTGCTGCTAAGCGAGTG	CCUUGGGCCUCCCCCCCAGCCCCUCCUCC	
	AAGTTAGATGCATCACAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGAGATCTTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACCGATGGAGAGCGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
18	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDALSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACUAGCCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGGGACUUCGUCGCCGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACTAGC	GAGAGAAACCCGUGGGUUCAUUGGCUGG	ANAKQSRDCFGCLREWC
	CAGAAGCACCGGGACTTCGTCGC	UAUAGGUGACGCUUUAAGUAAACGGCUC	DAFLWSHPQFEKWSHPQ
	CGAACCAATGGGAGAGAAACCCG	GAAGAAAGGGGCUUCGAUAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TGGGTTCATTGGCTGGTATAGGT	UAGUGCUGGGCCAGUUCCUCGUACUUAA	KLDASQGTKRSYEQMET
	GACGCTTTAAGTAAACGGCTCGA	AAAAGACGAAGAUCGCUUCAGAGAAUGG	DGERQ
	AGAAAGGGGCTTCGATAAGGCCT	CUUAAAGAUACAUGUGGAGCUAACGCAA	
	ACGTAGTGCTGGGCCAGTTCCTC	AACAAAGUCGAGACUGCUUCGGCUGCUU	
	GTACTTAAAAAAGACGAAGATCG	AAGAGAAUGGUGUGAUGCCUUUUUGUGG	
	CTTCAGAGAATGGCTTAAAGATA	UCUCAUCCCCAAUUUGAGAAGUGGUCCC	
	CATGTGGAGCTAACGCAAAACAA	ACCCUCAGUUUGAAAAGUGGUCUCAUCC	
	AGTCGAGACTGCTTCGGCTGCTT	CCAAUUUGAAAAGCCUGCCGCGAAACGC	
	AAGAGAATGGTGTGATGCCTTTT	GUGAAAUUAGACGCAAGUCAGGGUACCA	
	TGTGGTCTCATCCCCAATTTGAG	AGCGAUCUUAUGAGCAGAUGGAAACCGA	
	AAGTGGTCCCACCCTCAGTTTGA	UGGGGAAAGACAAUAAAGACCAGACAGC	
	AAAGTGGTCTCATCCCCAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCTGCCGCGAAACGCGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTAGACGCAAGTCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGCGATCTTATGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGGGAAAGACAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
19	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACAAGUCAGAAG	EDRGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGAUUUUGUCGCCGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACAAGT	GAGAAAAACCAGUGGGCUCCUUGGCCGG	ANAKQSRDCFGCLREWC
	CAGAAGCATAGAGATTTGTCGC	AAUCGGUGACGUUCUGAGCAAAAGGUUA	DAFLWSHPQFEKWSHPQ
	CGAACCAATGGGAGAAAAACCAG	GAGGACAGAGGUUUCGACAAGGCUUAUG	FEKWSHPQFEKPAAKRV
	TGGGCTCCTTGGCCGGAATCGGT	UCGUACUUGGCCAGUUUCUGGUUUUGAA	KLDASQGTKRSYEQMET
	GACGTTCTGAGCAAAAGGTTAGA	GAAAGAUGAAGAUCGUUUUAGGGAAUGG	DGERQ

	GGACAGAGGTTTCGACAAGGCTT	CUGAAAGAUACAUGCGGGGCUAACGCCA	
	ATGTCGTACTTGGCCAGTTTCTG	AACAGUCCCGCGAUUGCUUCGGAUGUCU	
	GTTTTGAAGAAAGATGAAGATCG	UCGCGAAUGGUGCGACGCGUUUCUUUGG	
	TTTTAGGGAATGGCTGAAAGATA	AGCCACCCUCAAUUUGAAAAGUGGAGUC	
	CATGCGGGGCTAACGCCAAACAG	ACCCACAGUUUGAGAAAUGGAGUCACCC	
	TCCCGCGATTGCTTCGGATGTCT	CCAGUUUGAAAAACCUGCAGCCAAAAGG	
	TCGCGAATGGTGCGACGCGTTTC	GUGAAAUUGGACGCAUCCCAGGGCACCA	
	TTTGGAGCCACCCTCAATTTGAA	AAAGGUCAUACGAGCAAAUGGAAACAGA	
	AAGTGGAGTCACCCACAGTTTGA	CGGGGAACGCCAAUAAAGACCAGACAGC	
	GAAATGGAGTCACCCCCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCTGCAGCCAAAAGGGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTGGACGCATCCCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAAAGGTCATACGAGCAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGACGGGGAACGCCAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
20	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDTLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACAAGUCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGGGACUUUGUCGCAGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACAAGT	GAGAAAAACCAGUUGGCAGCUUAGCGGG	ANAKQSRDCFGCLREWC
	CAGAAGCATAGGGACTTTGTCGC	CAUUGGGGACACCCUGUCAAAGAGACUC	DAFLWSHPQFEKWSHPQ
	AGAACCTATGGGAGAAAAACCAG	GAGGAAAGGGGCUUUGAUAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TTGGCAGCTTAGCGGGCATTGGG	UGGUUCUCGGGCAGUUUCUCGUCUUAAA	KLDASQGTKRSYEQMET
	GACACCCTGTCAAAGAGACTCGA	GAAGGAUGAAGACCGGUUUAGGGAAUGG	DGERQ
	GGAAAGGGGCTTTGATAAGGCCT	CUGAAAGAUACAUGCGGAGCUAAUGCCA	
	ACGTGGTTCTCGGGCAGTTTCTC	AACAGAGUAGGGACUGCUUCGGUUGUCU	
	GTCTTAAAGAAGGATGAAGACCG	GCGUGAAUGGUGCGACGCAUUCCUCUGG	
	GTTTAGGGAATGGCTGAAAGATA	AGUCACCCCCAAUUUGAGAAAUGGAGCC	
	CATGCGGAGCTAATGCCAAACAG	AUCCACAGUUUGAAAAAUGGAGUCACCC	
	AGTAGGGACTGCTTCGGTTGTCT	CCAAUUCGAGAAACCCGCAGCAAAGAGG	
	GCGTGAATGGTGCGACGCATTCC	GUCAAAUUGGAUGCUAGCCAGGGUACCA	
	TCTGGAGTCACCCCCAATTTGAG	AGCGCAGUUACGAGCAGAUGGAAACCGA	
	AAATGGAGCCATCCACAGTTTGA	UGGAGAAAGACAGUAAAGACCAGACAGC	
	AAAATGGAGTCACCCCCAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCCGCAGCAAAGAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTGGATGCTAGCCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGCGCAGTTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGAGAAAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
21	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACAAGUCAGAAG	EEKGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGAUUUUGUCGCAGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACAAGT	GCGAAAAGCCUGUCGGCUCCCUGGCAGG	ANAKQSRDCFGCLREWC
	CAGAAGCACAGAGATTTTGTCGC	AAUUGGAGACGUAUUGAGUAAAAGACUU	DAFLWSHPQFEKWSHPQ

	AGAACCTATGGGCGAAAAGCCTG	GAGGAAAAAGGAUUCGACAAAGCAUAUG	FEKWSHPQFEKPAAKRV
	TCGGCTCCCTGGCAGGAATTGGA	UUGUACUGGGGCAGUUCCUGGUUCUGAA	KLDASQGTKRSYEQMET
	GACGTATTGAGTAAAAGACTTGA	AAAGGAUGAAGAUAGGUUCAGAGAAUGG	DGERQ
	GGAAAAAGGATTCGACAAAGCAT	CUCAAAGAUACCUGUGGUGCCAAUGCCA	
	ATGTTGTACTGGGGCAGTTCCTG	AACAAAGUCGAGACUGUUUUGGCUGUCU	
	GTTCTGAAAAAGGATGAAGATAG	GAGAGAGUGGUGCGAUGCUUUCCUGUGG	
	GTTCAGAGAATGGCTCAAAGATA	UCUCACCCUCAAUUCGAGAAAUGGAGUC	
	CCTGTGGTGCCAATGCCAAACAA	AUCCACAGUUCGAAAAGUGGUCACACCC	
	AGTCGAGACTGTTTTGGCTGTCT	ACAGUUCGAGAAACCAGCCGCCAAGAGG	
	GAGAGAGTGGTGCGACGCTTTCC	GUCAAGCUGGAUGCCAGCCAGGGCACAA	
	TGTGGTCTCACCCTCAATTCGAG	AGAGAUCCUACGAGCAGAUGGAGACAGA	
	AAATGGAGTCATCCACAGTTCGA	UGGAGAACGCCAGUAAAGACCAGACAGC	
	AAAGTGGTCACACCCACAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCAGCCGCCAAGAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTGGATGCCAGCCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGAGATCCTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACAGATGGAGAACGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
22	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGTLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAUCCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGCGAUUUUGUGGCCGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACATCC	GAGAGAAGCCAGUGGGGACUCUGGCCGG	ANAKQSRDCFGCLREWC
	CAAAAGCATCGCGATTTTGTGGC	CAUAGGAGAUGUUCUGAGUAAGAGACUG	DAFLWSHPQFEKWSHPQ
	CGAACCCATGGGAGAGAAGCCAG	GAAGAGAGAGGCUUCGAUAAGGCUUACG	FEKWSHPQFEKPAAKRV
	TGGGGACTCTGGCCGGCATAGGA	UAGUCCUGGGACAGUUCUUGGUACUUAA	KLDASQGTKRSYEQMET
	GATGTTCTGAGTAAGAGACTGGA	AAAAGACGAGGAUCGAUUCCGCGAGUGG	DGERQ
	AGAGAGAGGCTTCGATAAGGCTT	CUCAAAGAUACCUGCGGUGCCAAUGCUA	
	ACGTAGTCCTGGGACAGTTCTTG	AGCAGUCAAGGGAUUGCUUCGGCUGCUU	
	GTACTTAAAAAAGACGAGGATCG	GAGAGAAUGGUGCGACGCAUUUUUAUGG	
	ATTCCGCGAGTGGCTCAAAGATA	UCCCACCCACAGUUCGAGAAGUGGUCCC	
	CCTGCGGTGCCAATGCTAAGCAG	AUCCCCAAUUCGAGAAAUGGAGCCACCC	
	TCAAGGGATTGCTTCGGCTGCTT	CCAGUUUGAAAAACCCGCGGCCAAGAGA	
	GAGAGAATGGTGCGACGCATTTT	GUGAAGCUCGAUGCUUCUCAAGGGACAA	
	TATGGTCCCACCCACAGTTCGAG	AACGAUCCUACGAGCAAAUGGAGACCGA	
	AAGTGGTCCCATCCCCAATTCGA	CGGCGAACGACAGUAAAGACCAGACAGC	
	GAAATGGAGCCACCCCCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCCGCGGCCAAGAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTCGATGCTTCTCAAGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAACGATCCTACGAGCAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACCGACGGCGAACGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	

23	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACCAGUCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGGGAUUUUGUAGCUGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACCAGT	GAGAAAAGCCUGUAGGGUCACUGGCGGG	ANAKQSRDCYGCLREWC
	CAAAAACATCGGGATTTTGTAGC	AAUUGGCGACGUUCUCAGCAAGCGUUUG	DAFLWSHPQFEKWSHPQ
	TGAACCTATGGGAGAAAAGCCTG	GAGGAGAGAGGCUUUGAUAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TAGGGTCACTGGCGGGAATTGGC	UUGUGCUGGGACAGUUCUUGGUGUUAAA	KLDASQGTKRSYEQMET
	GACGTTCTCAGCAAGCGTTTGGA	GAAAGAUGAGGAUAGGUUUAGAGAAUGG	DGERQ
	GGAGAGAGGCTTTGATAAGGCCT	CUGAAAGAUACAUGCGGGGCAAACGCUA	
	ACGTTGTGCTGGGACAGTTCTTG	AGCAAAGUCGCGAUUGUUAUGGAUGUCU	
	GTGTTAAAGAAAGATGAGGATAG	UAGAGAGUGGUGUGAUGCCUUUCUUUGG	
	GTTTAGAGAATGGCTGAAAGATA	AGCCAUCCACAAUUCGAGAAAUGGUCUC	
	CATGCGGGGCAAACGCTAAGCAA	ACCCACAAUUCGAGAAGUGGUCACACCC	
	AGTCGCGATTGTTATGGATGTCT	ACAGUUUGAGAAGCCAGCAGCUAAGAGG	
	TAGAGAGTGGTGTGATGCCTTTC	GUCAAACUCGAUGCCUCCCAAGGAACAA	
	TTTGGAGCCATCCACAATTCGAG	AGAGAAGUUAUGAACAAAUGGAAACCGA	
	AAATGGTCTCACCCACAATTCGA	CGGCGAAAGGCAGUAAAGACCAGACAGC	
	GAAGTGGTCACACCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCAGCAGCTAAGAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTCGATGCCTCCCAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGAGAAGTTATGAACAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGACGGCGAAAGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
24	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGEVLGKKL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAUCACAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGAUUUUGUGGCCGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACATCA	GGGAGAAGCCAGUCGGGUCACUCGCUGG	ANAKQSRDCFGCLREWC
	CAAAAGCACAGGGATTTTGTGGC	GAUUGGCGAAGUAUUGGGUAAGAAACUG	DAFLWSHPQFEKWSHPQ
	CGAACCAATGGGGGAGAAGCCAG	GAGGAACGCGGAUUCGAUAAGGCGUAUG	FEKWSHPQFEKPAAKRV
	TCGGGTCACTCGCTGGGATTGGC	UGGUUCUGGGUCAGUUCUUAGUGUUGAA	KLDASQGTKRSYEQMET
	GAAGTATTGGGTAAGAAACTGGA	AAAGGACGAAGAUCGCUUCAGAGAAUGG	DGERQ
	GGAACGCGGATTCGATAAGGCGT	UUGAAAGAUACAUGUGGAGCUAAUGCUA	
	ATGTGGTTCTGGGTCAGTTCTTA	AACAGUCACGGGAUUGUUUCGGCUGUUU	
	GTGTTGAAAAAGGACGAAGATCG	AAGAGAAUGGUGUGAUGCCUUCUUGUGG	
	CTTCAGAGAATGGTTGAAAGATA	UCUCACCCCCAAUUUGAGAAAUGGAGUC	
	CATGTGGAGCTAATGCTAAACAG	ACCCUCAGUUCGAGAAGUGGUCCCAUCC	
	TCACGGGATTGTTTCGGCTGTTT	UCAAUUUGAAAAGCCAGCAGCCAAACGU	
	AAGAGAATGGTGTGATGCCTTCT	GUUAAACUGGAUGCCAGCCAGGGGACAA	
	TGTGGTCTCACCCCCAATTTGAG	AGCGUUCUUAUGAACAAAUGGAAACCGA	
	AAATGGAGTCACCCTCAGTTCGA	UGGGGAAAGACAAUAAAGACCAGACAGC	
	GAAGTGGTCCCATCCTCAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCAGCAGCCAAACGTGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGATGCCAGCCAGGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGCGTTCTTATGAACAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGGGAAAGACAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
25	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACCTCC CAGAAACATCGCGATTTTGTCGC CGAGCCTATGGGAGAAAAACCAG TCGGATCTCTGGCCGGCATCGGG GATGTCTTAGGGAAGAAATTGGA GGAGAGAGGATTTGACAAGGCTT ACGTTGTTCTCGGTCAATTCTTA GTTCTGAAAAAGGATGAGGATAG ATTCCGCGAATGGCTGAAAGATA CGTGTGGGGCGAACGCCAAACAG TCCAGAGATTGTTTTGGGTGCCT GCGCGAGTGGTGTGATGCTTTTT TGTGGAGCCATCCCCAATTTGAA AAGTGGAGCCACCCCCAGTTCGA AAAATGGAGCCACCCGCAATTTG AGAAGCCCGCAGCAAAAAGGGTG AAACTGGACGCGTCCCAAGGAAC CAAAAGATCCTACGAACAAATGG AAACCGATGGGGAGAGACAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACCUCCCAGAAA CAUCGCGAUUUUGUCGCCGAGCCUAUGG GAGAAAAACCAGUCGGAUCUCUGGCCGG CAUCGGGGAUGUCUUAGGGAAGAAAUUG GAGGAGAGAGGAUUUGACAAGGCUUACG UUGUUCUCGGUCAAUUCUUAGUUCUGAA AAAGGAUGAGGAUAGAUUCCGCGAAUGG CUGAAAGAUACGUGUGGGGCGAACGCCA AACAGUCCAGAGAUUGUUUUGGGUGCCU GCGCGAGUGGUGUGAUGCUUUUUUGUGG AGCCAUCCCCAAUUUGAAAAGUGGAGCC ACCCCCAGUUCGAAAAAUGGAGCCACCC GCAAUUUGAGAAGCCCGCAGCAAAAAGG GUGAAACUGGACGCGUCCCAAGGAACCA AAAGAUCCUACGAACAAAUGGAAACCGA UGGGGAGAGACAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
26	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACCACATCT CAAAAGCACCGGGATTTTGTGGC TGAACCCATGGGGGAAAAGCCAG TGGGATCTTTAGCTGGCATTGGA GAGGTGCTTTCCAAGAAACTTGA GGAACGTGGTTTTGACAAAGCTT ACGTGGTCCTGGGACAATTCTTG GTACTGAAGAAAGATGAAGACAG ATTTAGGGAGTGGCTGAAGGACA CATGCGGTGCCAACGCAAAACAA TCACGCGACTGTTTTGGATGCCT CAGGGAATGGTGCGACGCCTTTT TATGGTCTCACCCACAATTCGAA AAGTGGTCCCATCCTCAGTTTGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACAUCUCAAAAG CACCGGGAUUUUGUGGCUGAACCCAUGG GGGAAAAGCCAGUGGGAUCUUUAGCUGG CAUUGGAGAGGUGCUUUCCAAGAAACUU GAGGAACGUGGUUUUGACAAAGCUUACG UGGUCCUGGGACAAUUCUUGGUACUGAA GAAAGAUGAAGACAGAUUUAGGGAGUGG CUGAAGGACACAUGCGGUGCCAACGCAA AACAAUCACGCGACUGUUUUGGAUGCCU CAGGGAAUGGUGCGACGCCUUUUUAUGG UCUCACCCACAAUUCGAAAAGUGGUCCC AUCCUCAGUUUGAGAAAUGGAGCCAUCC ACAGUUCGAAAAACCUGCUGCCAAAAGA GUGAAGUUAGACGCUUCCCAAGGAACCA AAAGAUCCUAUGAGCAGAUGGAAACUGA CGGGGAAAGGCAGUAAAGACCAGACAGC	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLSKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	GAAATGGAGCCATCCACAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCTGCTGCCAAAAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGTTAGACGCTTCCCAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAAAGATCCTATGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACTGACGGGGAAAGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
27	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGEVLGKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACUUCCCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGAUUUUGUUGCAGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACTTCC	GGGAAAAACCCGUUGGAAGCCUGGCGGG	ANAKQSRDCFGCLREWC
	CAAAAACACAGGGATTTTGTTGC	UAUCGGUGAAGUGUUGGGAAAAAGGCUC	DAFLWSHPQFEKWSHPQ
	AGAGCCTATGGGGGAAAAACCCG	GAGGAGAGAGGGUUUGAUAAAGCCUACG	FEKWSHPQFEKPAAKRV
	TTGGAAGCCTGGCGGGTATCGGT	UGGUGUUAGGUCAAUUUUUGGUGCUCAA	KLDASQGTKRSYEQMET
	GAAGTGTTGGGAAAAAGGCTCGA	AAAGGACGAAGAUAGAUUCAGAGAAUGG	DGERQ
	GGAGAGAGGGTTTGATAAAGCCT	UUAAAGGACACUUGUGGAGCUAACGCAA	
	ACGTGGTGTTAGGTCAATTTTTG	AACAGUCUCGCGAUUGCUUCGGGUGCUU	
	GTGCTCAAAAAGGACGAAGATAG	GAGGGAGUGGUGCGAUGCCUUCCUGUGG	
	ATTCAGAGAATGGTTAAAGGACA	AGCCACCCACAGUUCGAGAAAUGGAGUC	
	CTTGTGGAGCTAACGCAAAACAG	ACCCUCAAUUCGAAAAAUGGUCCCAUCC	
	TCTCGCGATTGCTTCGGGTGCTT	ACAGUUUGAGAAGCCAGCUGCAAAGCGA	
	GAGGGAGTGGTGCGATGCCTTCC	GUGAAACUCGAUGCUUCUCAGGGCACAA	
	TGTGGAGCCACCCACAGTTCGAG	AGAGAAGUUAUGAGCAGAUGGAAACGGA	
	AAATGGAGTCACCCTCAATTCGA	CGGAGAACGUCAGUAAAGACCAGACAGC	
	AAAATGGTCCCATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCAGCTGCAAAGCGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTCGATGCTTCTCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGAGAAGTTATGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACGGACGGAGAACGTCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
28	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKKL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAUCACAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGGGAUUUCGUCGCUGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACATCA	GGGAAAAGCCUGUGGGAAGUCUGGCUGG	ANAKQSRDCFGCLREWC
	CAGAAGCATAGGGATTTCGTCGC	UAUAGGGGAUGUCCUGAGCAAGAAACUG	DAFLWSHPQFEKWSHPQ
	TGAACCCATGGGGGAAAAGCCTG	GAAGAACGAGGAUUCGAUAAAGCCUACG	FEKWSHPQFEKPAAKRV
	TGGGAAGTCTGGCTGGTATAGGG	UGGUUCUUGGGCAAUUUCUCGUUCUGAA	KLDASQGTKRSYEQMET
	GATGTCCTGAGCAAGAAACTGGA	GAAAGACGAGGACCGGUUUAGAGAAUGG	DGERQ
	AGAACGAGGATTCGATAAAGCCT	CUGAAAGACACUUGCGGCGCUAAUGCUA	
	ACGTGGTTCTTGGGCAATTTCTC	AACAGAGUAGAGACUGUUUUGGCUGUCU	
	GTTCTGAAGAAAGACGAGGACCG	CCGGGAAUGGUGUGAUGCUUUCCUCUGG	
	GTTTAGAGAATGGCTGAAAGACA	AGUCAUCCUCAGUUCGAAAAAUGGAGCC	
	CTTGCGGCGCTAATGCTAAACAG	AUCCUCAGUUCGAGAAAUGGUCUCAUCC	
	AGTAGAGACTGTTTTGGCTGTCT	ACAGUUUGAGAAACCUGCCGCCAAAAGA	

	CCGGGAATGGTGTGAUGCTTTCC	GUCAAGUUAGACGCUAGCCAGGGUACGA	
	TCTGGAGTCATCCTCAGTTCGAA	AGAGAAGCUAUGAACAGAUGGAAACCGA	
	AAATGGAGCCATCCTCAGTTCGA	UGGAGAAAGGCAGUAAAGACCAGACAGC	
	GAAATGGTCTCATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCTGCCGCCAAAAGAGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGTTAGACGCTAGCCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	GAAGAGAAGCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGAGAAAGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
29	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGEVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCAGUCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGACUUUGUCGCUGAGCCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCAGT	GAGAAAAGCCAGUUGGGUCUUUGGCUGG	ANAKQSRDCFGCLREWC
	CAGAAACATAGAGACTTTGTCGC	GAUCGGCGAAGUGCUCAGCAAGCGACUC	DAFLWSHPQFEKWSHPQ
	TGAGCCCATGGGAGAAAAGCCAG	GAGGAGAGAGGAUUUGACAAGGCCUAUG	FEKWSHPQFEKPAAKRV
	TTGGGTCTTTGGCTGGGATCGGC	UCGUCUUGGGACAAUUUUUGGUACUUAA	KLDASQGTKRSYEQMET
	GAAGTGCTCAGCAAGCGACTCGA	GAAGGACGAGGAUCGAUUUAGAGAGUGG	DGERQ
	GGAGAGAGGATTTGACAAGGCCT	CUCAAAGAUACCUGCGGAGCCAACGCCA	
	ATGTCGTCTTGGGACAATTTTTG	AGCAAUCUAGGGACUGCUUCGGAUGCUU	
	GTACTTAAGAAGGACGAGGATCG	GCGGGAGUGGUGCGACGCUUUCUUAUGG	
	ATTTAGAGAGTGGCTCAAAGATA	AGUCAUCCCCAAUUCGAGAAAUGGAGCC	
	CCTGCGGAGCCAACGCCAAGCAA	ACCCGCAGUUCGAGAAAUGGUCCCACCC	
	TCTAGGGACTGCTTCGGATGCTT	GCAAUUCGAGAAGCCCGCUGCUAAACGA	
	GCGGGAGTGGTGCGACGCTTTCT	GUUAAACUGGACGCUAGCCAAGGCACUA	
	TATGGAGTCATCCCCAATTCGAG	AGAGAUCCUAUGAACAAAUGGAGACAGA	
	AAATGGAGCCACCCGCAGTTCGA	CGGAGAGAGGCAAUAAAGACCAGACAGC	
	GAAATGGTCCCACCCGCAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCTGCTAAACGAGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGACGCTAGCCAAGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGAGATCCTATGAACAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACAGACGGAGAGAGGCAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
30	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLGKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAUCACAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGAUUUUGUGGCAGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACATCA	GCGAAAAGCCCGUUGGGUCAUUGGCUGG	ANAKQSRDCFGCLREWC
	CAAAAACACAGAGATTTTGTGGC	AAUUGGCGACGUGCUGGGAAAGAGAUUG	DAFLWSHPQFEKWSHPQ
	AGAACCTATGGGCGAAAAGCCCG	GAAGAAAGAGGGUUUGACAAGGCAUAUG	FEKWSHPQFEKPAAKRV
	TTGGGTCATTGGCTGGAATTGGC	UGGUCCUGGGACAGUUCUUGGUUUUGAA	KLDASQGTKRSYEQMET
	GACGTGCTGGGAAAGAGATTGGA	AAAAGAUGAAGACAGGUUUAGAGAGUGG	DGERQ
	AGAAAGAGGGTTTGACAAGGCAT	CUGAAAGACACCUGUGGGGCAAAUGCAA	
	ATGTGGTCCTGGGACAGTTCTTG	AACAGUCCCGCGACUGCUUCGGCUGUCU	
	GTTTTGAAAAAAGATGAAGACAG	CCGGGAGUGGUGUGACGCGUUUCUCUGG	

	GTTTAGAGAGTGGCTGAAAGACA	UCCCACCCUCAGUUUGAAAAAUGGUCCC	
	CCTGTGGGGCAAATGCAAAACAG	ACCCGCAGUUUGAAAAAUGGAGUCAUCC	
	TCCCGCGACTGCTTCGGCTGTCT	UCAGUUUGAAAAGCCUGCCGCAAAGCGA	
	CCGGGAGTGGTGTGACGCGTTTC	GUUAAACUGGAUGCUAGUCAGGGUACCA	
	TCTGGTCCCACCCTCAGTTTGAA	AGAGGAGUUACGAACAGAUGGAAACCGA	
	AAATGGTCCCACCCGCAGTTTGA	CGGAGAGCGACAAUAAAGACCAGACAGC	
	AAAATGGAGTCATCCTCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCTGCCGCAAAGCGAGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGATGCTAGTCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGAGGAGTTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGACGGAGAGCGACAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
31	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLKGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCAGCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGAUUUUGUGGCUGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCAGC	GAGAAAAACCUGUCGGAAGUCUUAAAGG	ANAKQSRDCFGCLREWC
	CAAAAGCATAGAGATTTTGTGGC	AAUCGGUGAUGUCUUAAGUAAACGCCUG	DAFLWSHPQFEKWSHPQ
	TGAGCCTATGGGAGAAAAACCTG	GAGGAACGGGGAUUCGAUAAGGCAUACG	FEKWSHPQFEKPAAKRV
	TCGGAAGTCTTAAAGGAATCGGT	UUGUCCUGGGCCAGUUCCUUGUUCUGAA	KLDASQGTKRSYEQMET
	GATGTCTTAAGTAAACGCCTGGA	AAAGGACGAAGACAGGUUCAGAGAGUGG	DGERQ
	GGAACGGGGATTCGATAAGGCAT	UUGAAGGACACCUGUGGCGCUAACGCAA	
	ACGTTGTCCTGGGCCAGTTCCTT	AACAGAGCAGGGACUGUUUCGGCUGCCU	
	GTTCTGAAAAAGGACGAAGACAG	GAGAGAGUGGUGCGAUGCGUUCCUCUGG	
	GTTCAGAGAGTGGTTGAAGGACA	AGCCACCCUCAGUUCGAAAAAUGGUCCC	
	CCTGTGGCGCTAACGCAAAACAG	AUCCACAAUUUGAAAAGUGGAGCCAUCC	
	AGCAGGGACTGTTTCGGCTGCCT	ACAGUUUGAAAAGCCAGCAGCAAAAAGG	
	GAGAGAGTGGTGCGATGCGTTCC	GUCAAGCUCGAUGCUUCUCAGGGAACUA	
	TCTGGAGCCACCCTCAGTTCGAA	AAAGGUCUUACGAGCAGAUGGAAACGGA	
	AAATGGTCCCATCCACAATTTGA	CGGCGAGAGGCAGUAAAGACCAGACAGC	
	AAAGTGGAGCCATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCAGCAGCAAAAAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTCGATGCTTCTCAGGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAAAGGTCTTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACGGACGGCGAGAGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
32	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDRLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAAGCCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGAUUUUGUGGCCGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACAAGC	GCGAGAAGCCAGUAGGCUCCCUUGCAGG	ANAKQSRDCFGCLREWC
	CAAAAACACCGAGATTTTGTGGC	AAUUGGCGAUAGACUGUCAAAACGUCUU	DAFLWSHPQFEKWSHPQ
	CGAACCAATGGGCGAGAAGCCAG	GAGGAGCGGGGAUUUGAUAAAGCCUACG	FEKWSHPQFEKPAAKRV
	TAGGCTCCCTTGCAGGAATTGGC	UGGUAUUGGGUCAGUUUCUGGUGCUUAA	KLDASQGTKRSYEQMET
	GATAGACTGTCAAAACGTCTTGA	GAAGGAUGAGGACAGAUUCCGGGAAUGG	DGERQ

	GGAGCGGGGATTTGATAAAGCCT	CUCAAGGACACAUGCGGCGCUAAUGCCA	
	ACGTGGTATTGGGTCAGTTTCTG	AACAGUCUAGGGACUGCUUCGGGUGUCU	
	GTGCTTAAGAAGGATGAGGACAG	GCGAGAGUGGUGCGAUGCUUUUCUCUGG	
	ATTCCGGGAATGGCTCAAGGACA	AGCCAUCCCCAGUUUGAAAAAUGGUCUC	
	CATGCGGCGCTAATGCCAAACAG	AUCCGCAGUUCGAGAAGUGGUCUCACCC	
	TCTAGGGACTGCTTCGGGTGTCT	CCAGUUCGAAAAACCAGCAGCCAAACGA	
	GCGAGAGTGGTGCGATGCTTTTC	GUAAAGUUGGAUGCCUCACAGGGAACCA	
	TCTGGAGCCATCCCCAGTTTGAA	AACGGAGUUACGAACAGAUGGAAACCGA	
	AAATGGTCTCATCCGCAGTTCGA	CGGAGAAAGACAGUAAAGACCAGACAGC	
	GAAGTGGTCTCACCCCCAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCAGCAGCCAAACGAGTA	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGTTGGATGCCTCACAGGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAACGGAGTTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGACGGAGAAAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
33	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACUAGCCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGAUUUUGUUGCAGAACCGAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACTAGC	GGGAGAAGCCAGUCGGCUCACUGGCAGG	ANKKQSRDCFGCLREWC
	CAGAAGCACAGAGATTTTGTTGC	AAUUGGAGACGUAUUAUCUAAGAGAUUA	DAFLWSHPQFEKWSHPQ
	AGAACCGATGGGGGAGAAGCCAG	GAAGAACGCGGGUUUGAUAAGGCUUACG	FEKWSHPQFEKPAAKRV
	TCGGCTCACTGGCAGGAATTGGA	UUGUCUUGGGCCAGUUUCUGGUACUGAA	KLDASQGTKRSYEQMET
	GACGTATTATCTAAGAGATTAGA	GAAAGAUGAGGACCGCUUUCGGGAGUGG	DGERQ
	AGAACGCGGGTTTGATAAGGCTT	CUGAAGGACACAUGUGGGGCUAAUAAGA	
	ACGTTGTCTTGGGCCAGTTTCTG	AGCAGAGUAGGGAUUGUUUCGGGUGCCU	
	GTACTGAAGAAAGATGAGGACCG	UAGAGAGUGGUGUGAUGCCUUUCUGUGG	
	CTTTCGGGAGTGGCTGAAGGACA	AGUCAUCCCCAGUUUGAAAAGUGGUCUC	
	CATGTGGGGCTAATAAGAAGCAG	ACCCACAAUUUGAGAAGUGGUCUCAUCC	
	AGTAGGGATTGTTTCGGGTGCCT	ACAAUUCGAGAAACCCGCCGCAAAAAGA	
	TAGAGAGTGGTGTGATGCCTTTC	GUGAAGUUGGAUGCUAGUCAGGGCACAA	
	TGTGGAGTCATCCCCAGTTTGAA	AGCGCAGCUACGAGCAGAUGGAGACUGA	
	AAGTGGTCTCACCCACAATTTGA	UGGAGAAAGACAAUAAAGACCAGACAGC	
	GAAGTGGTCTCATCCACAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCCGCCGCAAAAAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGTTGGATGCTAGTCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGCGCAGCTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACTGATGGAGAAAGACAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	

34	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLKGIGDRLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACUAGUCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGAGAUUUUGUAGCAGAGCCGAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACTAGT	GGGAGAAGCCAGUGGGCAGCUUAAAGGG	ANKKQSRDCFGCLREWC
	CAGAAGCATCGAGATTTTGTAGC	UAUUGGUGACAGGUUAAGUAAGAGGUUG	DAFLWSHPQFEKWSHPQ
	AGAGCCGATGGGGGAGAAGCCAG	GAAGAGAGAGGGUUCGAUAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TGGGCAGCTTAAAGGGTATTGGT	UGGUCCUCGGGCAGUUCCUUGUUCUCAA	KLDASQGTKRSYEQMET
	GACAGGTTAAGTAAGAGGTTGGA	AAAGGAUGAAGAUCGGUUCAGGGAAUGG	DGERQ
	AGAGAGAGGGTTCGATAAGGCCT	CUGAAAGAUACCUGCGGCGCUAAUAAGA	
	ACGTGGTCCTCGGGCAGTTCCTT	AACAGUCAAGAGAUUGUUUCGGAUGCUU	
	GTTCTCAAAAAGGATGAAGATCG	AAGGGAGUGGUGCGACGCCUUCCUGUGG	
	GTTCAGGGAATGGCTGAAAGATA	UCUCACCCACAGUUUGAAAAAUGGAGUC	
	CCTGCGGCGCTAATAAGAAACAG	ACCCACAGUUUGAGAAGUGGAGCCAUCC	
	TCAAGAGATTGTTTCGGATGCTT	GCAGUUCGAGAAGCCCGCCGCCAAAAGA	
	AAGGGAGTGGTGCGACGCCTTCC	GUAAAGCUGGAUGCAUCUCAGGGAACCA	
	TGTGGTCTCACCCACAGTTTGAA	AACGUUCUUACGAGCAGAUGGAAACAGA	
	AAATGGAGTCACCCACAGTTTGA	UGGCGAACGCCAGUAAAGACCAGACAGC	
	GAAGTGGAGCCATCCGCAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCCGCCAAAAGAGTA	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTGGATGCATCTCAGGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAACGTTCTTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGATGGCGAACGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
35	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCUCUCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGAUUUUGUCGCAGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCTCT	GAGAAAAGCCUGUCGGGUCCCUGGCCGG	ANAKQSRDCFRCLREWC
	CAAAAACATAGAGATTTTGTCGC	AAUAGGGGAUGUGUUGAGCAAAAGACUG	DAFLWSHPQFEKWSHPQ
	AGAACCAATGGGAGAAAAGCCTG	GAGGAGCGGGGCUUCGAUAAGGCAUACG	FEKWSHPQFEKPAAKRV
	TCGGGTCCCTGGCCGGAATAGGG	UCGUAUUGGGACAGUUUCUGGUGUUAAA	KLDASQGTKRSYEQMET
	GATGTGTTGAGCAAAAGACTGGA	AAAGGAUGAAGAUAGAUUUAGAGAAUGG	DGERQ
	GGAGCGGGGCTTCGATAAGGCAT	CUGAAGGAUACUUGUGGUGCCAAUGCAA	
	ACGTCGTATTGGGACAGTTTCTG	AACAAUCUCGCGACUGCUUCCGGUGUCU	
	GTGTTAAAAAAGGATGAAGATAG	CAGAGAGUGGUGCGAUGCUUUCCUGUGG	
	ATTTAGAGAATGGCTGAAGGATA	UCUCACCCGCAGUUUGAGAAAUGGUCAC	
	CTTGTGGTGCCAATGCAAAACAA	ACCCACAGUUCGAAAAAUGGAGUCAUCC	
	TCTCGCGACTGCTTCCGGTGTCT	UCAAUUUGAGAAACCCGCUGCAAAGCGA	
	CAGAGAGTGGTGCGATGCTTTCC	GUAAAGCUCGACGCUUCACAAGGUACCA	
	TGTGGTCTCACCCGCAGTTTGAG	AGAGGUCUUACGAACAGAUGGAAACGGA	
	AAATGGTCACACCCACAGTTCGA	CGGCGAACGGCAGUAAAGACCAGACAGC	
	AAAATGGAGTCATCCTCAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCCGCTGCAAAGCGAGTA	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTCGACGCTTCACAAGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGAGGTCTTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACGGACGGCGAACGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	
36	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACAACCTCA CAGAAGCACCGCGACTTCGTGGC TGAACCCATGGGAGAAAAACCTG TGGGATCACTCGCGGGCATAGGC GACGTACTCTCAAAGCGATTGGA GGAAAGAGGATTCGATAAGGCAT ATGTCGTGCTCGGTCAATTTTTG GTACTTAAAAAGGACGAGGATCG ATTTAGAGAGTGGCTGAAAGATA CTTGTGGGGCTAACGCAAAGCAA AGCCGTGACTGCTTCGGATGTTT AAGAGAATGGTGCGACGCTTTTC TTTGGAGTCACCCCCAGTTTGAG AAATGGTCCCACCCTCAGTTCGA AAAGTGGAGTCACCCACAATTTG AGAAGGCCACGGTTAAGTTTAAG TATAAAGGCGAAGAAAAAGAAGT GGACATCTCCAAAATCAAAAAGG TATGGAGAGTCGGCAAGATGATA TCCTTCACTTACGATGAAGGAGG AGGCAAGACTGGGAGGGGTGCAG TTTCAGAGAAGGACGCGCCAAAA GAACTGCTTCAAATGCTCGAGAA ACAGAAGAAATGGTCACATCCTC AATTCGAGAAATGGAGCCACCCC CAATTCGAGAAATGGTCTCATCC ACAATTTGAGAAGCCGGCTGCCA AGAGAGTGAAACTTGATGCAAGC CAGGGCACTAAACGCTCTTATGA ACAAATGGAAACCGACGGGGAAA GGCAGTAAAGACCAGACAGCTGG AGCCTCGGTGGCCATGCTTCTTG CCCCTTGGGCCTCCCCCCAGCCC CTCCTCCCCTTCCTGCACCCGTA CCCCCGTGGTCTTTGAATAAAGT CTGAGTGGGCGGCAAGCTTTTCT TCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACCUCACAGAAG CACCGCGACUUCGUGGCUGAACCCAUGG GAGAAAAACCUGUGGGAUCACUCGCGGG CAUAGGCGACGUACUCUCAAAGCGAUUG GAGGAAAGAGGAUUCGAUAAGGCAUAUG UCGUGCUCGGUCAAUUUUUGGUACUUAA AAAGGACGAGGAUCGAUUUAGAGAGUGG CUGAAAGAUACUUGUGGGGCUAACGCAA AGCAAAGCCGUGACUGCUUCGGAUGUUU AAGAGAAUGGUGCGACGCUUUUCUUUGG AGUCACCCCCAGUUUGAGAAAUGGUCCC ACCCUCAGUUCGAAAAGUGGAGUCACCC ACAAUUUGAGAAGGCCACGGUUAAGUUU AAGUAUAAAGGCGAAGAAAAAGAAGUGG ACAUCUCCAAAAUCAAAAAGGUAUGGAG AGUCGGCAAGAUGAUAUCCUUCACUUAC GAUGAAGGAGGAGGCAAGACUGGGAGGG GUGCAGUUUCAGAGAAGGACGCGCCAAA AGAACUGCUUCAAAUGCUCGAGAAACAG AAGAAAUGGUCACAUCCUCAAUUCGAGA AAUGGAGCCACCCCCAAUUCGAGAAAUG GUCUCAUCCACAAUUUGAGAAGCCGGCU GCCAAGAGAGUGAAACUUGAUGCAAGCC AGGGCACUAAACGCUCUUAUGAACAAAU GGAAACCGACGGGGAAAGGCAGUAAAGA CCAGACAGCUGGAGCCUCGGUGGCCAUG CUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGCAAGCUUUUCUUCUCUCCCUUGCA CCUAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKATVKFK YKGEEKEVDISKIKKVW RVGKMISFTYDEGGGKI GRGAVSEKDAPKELLQM LEKQKKWSHPQFEKWSH PQFEKWSHPQFEKPAAK RVKLDASQGTKRSYEQM ETDGERQ

37	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGAGCACTTCC CAGAAACATAGAGATTTTGTTGC AGAACCAATGGGGGAAAAAGCAG TGACAGCCCTGGCCGGAATTGGA GAGACTCTTGGAAGTAGACTCTG TTCCTTAGGATTCGATAAAGCTT ACGTGGTACTGGGGCAGTTTTTA GTGCTGAAGAAAGATAAAGAGAG ATTCGAGGAATGGCTGAAAGACG CCTGCAATGCTAATACGAAATAC TCTAAAGACTGCTATAATTGCTT AAAAGAATGGTGCGATGCATTCT TGTGGTCTCACCCGCAGTTTGAG AAATGGTCTCATCCCCAGTTTGA AAAATGGTCACACCCCCAGTTCG AAAAACCTGCCGCCAAGCGGGTT AAGCTGGATGCATCCCAAGGCAC GAAGCGAAGTTATGAACAGATGG AGACAGATGGAGAACGACAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAGCACUUCCCAGAAA CAUAGAGAUUUUGUUGCAGAACCAAUGG GGGAAAAAGCAGUGACAGCCCUGGCCGG AAUUGGAGAGACUCUUGGAAGUAGACUC UGUUCCUUAGGAUUCGAUAAAGCUUACG UGGUACUGGGGCAGUUUUUAGUGCUGAA GAAAGAUAAAGAGAGAUUCGAGGAAUGG CUGAAAGACGCCUGCAAUGCUAAUACGA AAUACUCUAAAGACUGCUAUAAUUGCUU AAAAGAAUGGUGCGAUGCAUUCUUGUGG UCUCACCCGCAGUUUGAGAAAUGGUCUC AUCCCCAGUUUGAAAAAUGGUCACACCC CCAGUUCGAAAAACCUGCCGCCAAGCGG GUUAAGCUGGAUGCAUCCCAAGGCACGA AGCGAAGUUAUGAACAGAUGGAGACAGA UGGAGAACGACAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MSTSQKHRDFVAEPMGE KAVTALAGIGETLGSRL CSLGFDKAYVVLGQFLV LKKDKERFEEWLKDACN ANTKYSKDCYNCLKEWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
38	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGCTAGCCAA AAGCACCGTGACTTTGTGGCAGA ACCGATGGGAGATAAGCCCGTGA CTGAACTCGCCGGCATTGGCAGA GTGTTAGGTGAACGTCTGAACGA GCATGGGTTTGACCAGGCCTACG TGGTACTCGGACAGTTTTTACTG CTCAGAAAAGATGAAGAGAGATT CACTAGATGGCTTAAAGACACAA CCCCAGCAAATAAGGAGCAGGCA CGCGATTGTTACACATGCTTGCG TGACTGGTGTCAAGCATTCCTCT GGAGCCACCCCCAGTTCGAGAAG TGGAGTCACCCACAGTTTGAAAA ATGGAGTCATCCCCAATTCGAGA AGCCTGCTGCCAAGAGAGTGAAG TTAGACGCTTCACAGGGTACTAA AAGGAGCTATGAACAGATGGAAA CAGACGGAGAAAGGCAGTAAAGA CCAGACAGCTGGAGCCTCGGTGG CCATGCTTCTTGCCCCTTGGGCC TCCCCCCAGCCCCTCCTCCCCTT CCTGCACCCGTACCCCCGTGGTC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCUAGCCAAAAGCAC CGUGACUUUGUGGCAGAACCGAUGGGAG AUAAGCCCGUGACUGAACUCGCCGGCAU UGGCAGAGUGUUAGGUGAACGUCUGAAC GAGCAUGGGUUUGACCAGGCCUACGUGG UACUCGGACAGUUUUUACUGCUCAGAAA AGAUGAAGAGAGAUUCACUAGAUGGCUU AAAGACACAACCCCAGCAAAUAAGGAGC AGGCACGCGAUUGUUACACAUGCUUGCG UGACUGGUGUCAAGCAUUCCUCUGGAGC CACCCCCAGUUCGAGAAGUGGAGUCACC CACAGUUUGAAAAAUGGAGUCAUCCCCA AUUCGAGAAGCCUGCUGCCAAGAGAGUG AAGUUAGACGCUUCACAGGGUACUAAAA GGAGCUAUGAACAGAUGGAAACAGACGG AGAAAGGCAGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCAAGCUUU UCUUCUCUCCCUUGCACCUAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MASQKHRDFVAEPMGDK PVTELAGIGRVLGERLN EHGFDQAYVVLGQFLLL RKDEERFTRWLKDTTPA NKEQARDCYTCLRDWCQ AFLWSHPQFEKWSHPQF EKWSHPQFEKPAAKRVK LDASQGTKRSYEQMETD GERQ

	TTTGAATAAAGTCTGAGTGGGCG GCAAGCTTTTCTTCTCTCCCTTG CACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	
39	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGAGCAGCACC TCTCAGAAGCACAGAAATTTCGT GGGTGAGCCCATGGGCCAAAAGC CCGTCACAGAGTTAGCAGGAATC GGCGAGGCCCTCGGTGGGAGATT AGCTGCAAAGGGGTTTGACAAGG CGTATGTTGTGCTGGGGCAATTC TTGCTGCTGAAAATGAATGAAGA AAGGTTCAAGGAGTGGCTCAACG CAGAATGTGGTGCTAATAGTAAA CAGCAGTCAGATTGTTATAATTG CCTGCGTGATTGGTGCGACGCTT TCCTGTGGTCTCACCCCCAATTT GAAAAATGGTCACATCCCCAGTT TGAAAAATGGAGTCACCCACAAT TCGAGAAACCCGCCGCGAAAAGA GTGAAGCTCGATGCCAGCCAGGG GACCAAGCGTAGCTACGAACAGA TGGAGACCGATGGTGAAAGACAG TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAGCAGCACCUCUCAG AAGCACAGAAAUUUCGUGGGUGAGCCCA UGGGCCAAAAGCCCGUCACAGAGUUAGC AGGAAUCGGCGAGGCCCUCGGUGGGAGA UUAGCUGCAAAGGGGUUUGACAAGGCGU AUGUUGUGCUGGGGCAAUUCUUGCUGCU GAAAAUGAAUGAAGAAAGGUUCAAGGAG UGGCUCAACGCAGAAUGUGGUGCUAAUA GUAAACAGCAGUCAGAUUGUUAUAAUUG CCUGCGUGAUUGGUGCGACGCUUUCCUG UGGUCUCACCCCCAAUUUGAAAAAUGGU CACAUCCCCAGUUUGAAAAAUGGAGUCA CCCACAAUUCGAGAAACCCGCCGCGAAA AGAGUGAAGCUCGAUGCCAGCCAGGGGA CCAAGCGUAGCUACGAACAGAUGGAGAC CGAUGGUGAAAGACAGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA A	MSSTSQKHRNFVGEPMG QKPVTELAGIGEALGGR LAAKGFDKAYVVLGQFL LLKMNEERFKEWLNAEC GANSKQQSDCYNCLRDW CDAFLWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ
40	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGTCATCAACG AGCCAGAAACACCGGAACTTTGT AGCAGAGCCTATGGGCGAGAAAT CCGTCACCGAGTTATCTGGGATT GGCGACACACTGGGTAAACGGCT CTCTGACGATGGATTTGATAAGG CTTACGTGGTGCTGGGACAGTTT TTGCTGCTGAAAAAAGATGAAGA CAGATTCAAAGCCTGGTTAAAGG ACAGCTGTGGAGCCAATGCAAAG CAACAGAAAGACTGCCATACATG TCTGAAGGAGTGGTGCGATTCTT TTTTGTGGTCTCACCCTCAGTTT GAGAAATGGTCACATCCACAGTT TGAAAAGTGGTCACACCCACAGT TTGAAAAACCTGCAGCCAAGCGG GTAAAATTAGATGCCTCACAGGG GACCAAAAGATCATATGAACAAA TGGAAACCGACGGCGAGAGGCAA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUCAUCAACGAGCCAG AAACACCGGAACUUUGUAGCAGAGCCUA UGGGCGAGAAAUCCGUCACCGAGUUAUC UGGGAUUGGCGACACACUGGGUAAACGG CUCUCUGACGAUGGAUUUGAUAAGGCUU ACGUGGUGCUGGGACAGUUUUUGCUGCU GAAAAAAGAUGAAGACAGAUUCAAAGCC UGGUUAAAGGACAGCUGUGGAGCCAAUG CAAAGCAACAGAAAGACUGCCAUACAUG UCUGAAGGAGUGGUGCGAUUCUUUUUUG UGGUCUCACCCUCAGUUUGAGAAAUGGU CACAUCCACAGUUUGAAAAAGUGGUCACA CCCACAGUUUGAAAAACCUGCAGCCAAG CGGGUAAAAUUAGAUGCCUCACAGGGGA CCAAAAGAUCAUAUGAACAAAUGGAAAC CGACGGCGAGAGGCAAUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA	MSSTSQKHRNFVAEPMG EKSVTELSGIGDTLGKR LSDDGFDKAYVVLGQFL LLKKDEDRFKAWLKDSC GANAKQQKDCHTCLKEW CDSFLWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ

	TAAAGACCAGACAGC□GGAGCCT CGGTGGCCATGCTTC□TGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTT□CTTCTCT CCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA A	
41	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGTCAAGTACC AGCCAAAAACACCGCAATTTTGT TGCTGAACCAATGGG□GAGAAAC CCGTGATGGCGTTAGCCGGGATT GGTGACGTCTTGGGTAAAAGACT GGAAGACAAGGGCTTCGACAAAG CTTACGTCGTTCTCGGGCAATTC CTTGTACTGAAAAAAGATGAAGA AAGATTTAATGACTGGATGAAGG AGACCTGCGGCGCGAACACAAAG CAGAGGGGAGATTGT□TTGTTTG TTTGAAGGAATGGAGCGATACAT TCGTGTGGAGCCACCCCCAGTTC GAGAAATGGAGCCACCCTCAGTT CGAAAAGTGGTCTCACCCTCAGT TCGAGAAGCCAGCAGCTAAGAGA GTTAAATTAGATGCC□CCCAAGG GACAAAACGGAGCTACGAGCAGA T□GAAACT□GACGGCGAGCGCCAG TAAAGACCAGACAGC□GGAGCCT CGGTGGCCATGCTTC□TGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTT□CTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUCAAGUACCAGCCAA AAACACCGCAAUUUUGUUGCUGAACCAA UGGGUGAGAAACCCGUGAUGGCGUUAGC CGGGAUUGGUGACGUCUUGGGUAAAAGA CUGGAAGACAAGGGCUUCGACAAAGCUU ACGUCGUUCUCGGGCAAUUCCUUGUACU GAAAAAAGAUGAAGAAAGAUUUAAUGAC UGGAUGAAGGAGACCUGCGGCGCGAACA CAAAGCAGAGGGGAGAUUGUUUUGUUUG UUUGAAGGAAUGGAGCGAUACAUUCGUG UGGAGCCACCCCCAGUUCGAGAAAUGGA GCCACCCUCAGUUCGAAAAGUGGUCUCA CCCUCAGUUCGAGAAGCCAGCAGCUAAG AGAGUUAAAUUAGAUGCCUCCCAAGGGA CAAAACGGAGCUACGAGCAGAUGGAAAC UGACGGCGAGCGCCAGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA A	MSSTSQKHRNFVAEPMG EKPVMALAGIGDVLGKR LEDKGFDKAYVVLGQFL VLKKDEERFNDWMKETC GANTKQRGDCFVCLKEW SDTFVWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ
42	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGC□CTCAGCA AATAAACTGGATCCAATGTCAAG TACGAGCCAGAAACA□AACGACT TTGTGGCGGAGCCAA□GGGCGAT AAGCCAGTTACTGCA□TACCCGG AATAGGGGAAATTCTCGGAGGCC GATTGAAGGAGAAAGGCTTCGGA AAAGCCTACAACGTACTGGGCAA CTTTCTGGCACTGAA□AGAGACG AAAAGCGCTTCCGGGATTGGCTG AAGGACACTTGCCAAGCAAGCTC CGAGCAGCAGAGCGACTGCTACG GGTGCCTGAAAGACTGGTGTGCT AACAATTTGTGGTCACATCCTCA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCUCUCAGCAAAUAAA CUGGAUCCAAUGUCAAGUACGAGCCAGA AACAUAACGACUUUGUGGCGGAGCCAAU GGGCGAUAAGCCAGUUACUGCAUUACCC GGAAUAGGGGAAAUUCUCGGAGGCCGAU UGAAGGAGAAAGGCUUCGGAAAAGCCUA CAACGUACUGGGCAACUUUCUGGCACUG AAUAGAGACGAAAAGCGCUUCCGGGAUU GGCUGAAGGACACUUGCCAAGCAAGCUC CGAGCAGCAGAGCGACUGCUACGGGUGC CUGAAAGACUGGUGUGCUAACAAUUUGU GGUCACAUCCUCAAUUCGAGAAAUGGUC UCAUCCCCAGUUCGAAAAAUGGUCCCAC CCACAGUUCGAAAAACCAGCCGCCAAAC GCGUGAAGUUAGAUGCUAGCCAGGGAAC GAAAAGAAGCUACGAGCAGAUGGAGACA	MLSANKLDPMSSTSQKH NDFVAEPMGDKPVTALP GIGEILGGRLKEKGFGK AYNVLGNFLA□NRDEKR FRDWLKDTCQASSEQQS DCYGCLKDWCANNLWSH PQFEKWSHPQFEKWSHP QFEKPAAKRVKLDASQG TKRSYEQMETDGERQ

	ATTCGAGAAATGGTCTCATCCCC AGTTCGAAAAATGGTCCCACCCA CAGTTCGAAAAACCAGCCGCCAA ACGCGTGAAGTTAGATGCTAGCC AGGGAACGAAAAGAAGCTACGAG CAGATGGAGACAGACGGCGAACG TCAGTAAAGACCAGACAGCTGGA GCCTCGGTGGCCATGCTTCTTGC CCCTTGGGCCTCCCCCCAGCCCC TCCTCCCCTTCCTGCACCCGTAC CCCCGTGGTCTTTGAATAAAGTC TGAGTGGGCGGCAAGCTTTTCTT CTCTCCCTTGCACCT	GACGGCGAACGUCAGUAAAGACCAGACA GCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUGGGCCUCCCCCCAGCCCCUCCU CCCCUUCCUGCACCCGUACCCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCAA GCUUUUCUUCUCUCCCUUGCACCUAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
43	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGATCATATG TCCCCACGGTTGCGGGCATTCCT GTCTGAACCTATTGGCGAAAAAG ACGTGGCTTGGGTGGATGGCATC TCCCGCGAGCTTGCTATCAATTT GGTGACAAAAGGATTCAACAAAG CGTATATACTGCTGGGGCAATTT CTGCTCATGCACAAGAACGAAGC AAGATTTCAAAAATGGCTGATAT GTTGCTGTGGGGCAACTGAGTGT GAAGCCCAGCAGAGCTCTAACTG TCTTAAGGAGTGGTGTGCATGTT TTTTGTGGTCTCATCCCCAGTTT GAAAAATGGTCACATCCTCAGTT TGAAAAATGGAGCCATCCCCAAT TTGAGAAACCTGCAGCAAAAAGA GTTAAATTGGACGCCTCACAAGG AACAAAAAGGAGTTACGAACAGA TGGAAACCGACGGAGAGAGACAG TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGAUCAUAUGUCCCCA CGGUUGCGGGCAUUCCUGUCUGAACCUA UUGGCGAAAAAGACGUGGCUUGGGUGGA UGGCAUCUCCCGCGAGCUUGCUAUCAAU UUGGUGACAAAAGGAUUCAACAAAGCGU AUAUACUGCUGGGGCAAUUUCUGCUCAU GCACAAGAACGAAGCAAGAUUUCAAAAA UGGCUGAUAUGUUGCUGUGGGGCAACUG AGUGUGAAGCCCAGCAGAGCUCUAACUG UCUUAAGGAGUGGUGUGCAUGUUUUUUG UGGUCUCAUCCCCAGUUUGAAAAAUGGU CACAUCCUCAGUUUGAAAAAUGGAGCCA UCCCCAAUUUGAGAAACCUGCAGCAAAA AGAGUUAAAUUGGACGCCUCACAAGGAA CAAAAAGGAGUUACGAACAGAUGGAAAC CGACGGAGAGAGACAGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA A	MDHMSPRLRAFLSEPIG EKDVAWVDGISRELAIN LVTKGFNKAYILLGQFL LMHKNEARFQKWLICCC GATECEAQQSSNCLKEW CACFLWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ
44	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGTCAACCAGT GTGAAGCATCGTGAATTTATTGG TGAGCCCATGGGAGAAAAGGAGG TGACCGCCATAGCAGGAATCGGG CCTACATATGGAAAAAAACTTAC AGAAAAAGGTTTCGACAAGGCTT ATGTGTTATTTGGACAGTTCTTG CTCCTGAAGAAAGATAAGGAGCG CTTCATTGATTGGTTAAAAGAAG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUCAACCAGUGUGAAG CAUCGUGAAUUUAUUGGUGAGCCCAUGG GAGAAAAGGAGGUGACCGCCAUAGCAGG AAUCGGGCCUACAUAUGGAAAAAAACUU ACAGAAAAAGGUUUCGACAAGGCUUAUG UGUUAUUUGGACAGUUCUUGCUCCUGAA GAAAGAUAAGGAGCGCUUCAUUGAUUGG UUAAAAGAAGAAGCCGGCGUAAGUGCAA AUCAUGCAAAGAGUGCAUACAAUUGUCU CAACGAAUGGGCCGAGCAGUUCAUAUGG AGCCAUCCACAGUUUGAAAAAUGGUCCC	MSTSVKHREFIGEPMGE KEVTAIAGIGPTYGKKL TEKGFDKAYV_FGQFLL LKKDKERFIDWLKEEAG VSANHAKSAYNCLNEWA EQFIWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	AAGCCGGCGTAAGTGCAAATCAT	ACCCACAGUUCGAAAAAUGGAGUCAUCC	
	GCAAAGAGTGCATACAATTGTCT	GCAGUUUGAAAAGCCAGCUGCGAAACGC	
	CAACGAATGGGCCGAGCAGTTCA	GUUAAACUCGAUGCUAGCCAGGGAACUA	
	TATGGAGCCATCCACAGTTTGAA	AGAGGUCUUACGAGCAGAUGGAAACAGA	
	AAATGGTCCCACCCACAGTTCGA	CGGGGAGAGGCAAUAAAGACCAGACAGC	
	AAAATGGAGTCATCCGCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCAGCTGCGAAACGCGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTCGATGCTAGCCAGGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGAGGTCTTACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGACGGGGAGAGGCAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
45	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MSGNKKGSPTTSQKHRN
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	FVSEPMGNRSVRDVPGI
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGUCAGGCAACAAAAAA	GPAHGRRLEEKGMPRAD
	GCGGTAGACGCACGGCATAGCAT	GGAUCUCCUACAACUAGUCAAAAGCACA	QMLGDYLKRGRDQGRFQ
	CCAACGCCACCATGTCAGGCAAC	GAAACUUUGUGAGUGAGCCUAUGGGUAA	DNLKNTTGANAKQQRDA
	AAAAAAGGATCTCCTACAACTAG	UCGCUCCGUGCGCGACGUGCCAGGUAUC	YNGMREWTDNNMWSHPQ
	TCAAAAGCACAGAAACTTTGTGA	GGCCCCGCACAUGGCCGUAGGUUGGAAG	FEKWSHPQFEKWSHPQF
	GTGAGCCTATGGGTAATCGCTCC	AAAAAGGAAUGCCCCGUGCCGAUCAGAU	EKPAAKRVKLDASQGTK
	GTGCGCGACGTGCCAGGTATCGG	GCUGGGAGACUAUUUGAAACGUGGACGG	RSYEQMETDGERQ
	CCCCGCACATGGCCGTAGGTTGG	GACCAGGGUCGUUUCCAAGACAAUUUAA	
	AAGAAAAAGGAATGCCCCGTGCC	AGAAUACAACCGGAGCCAAUGCCAAACA	
	GATCAGATGCTGGGAGACTATTT	GCAGAGAGACGCAUACAAUGGAAUGCGA	
	GAAACGTGGACGGGACCAGGGTC	GAAUGGACCGACAACAAUAUGUGGAGCC	
	GTTTCCAAGACAATTTAAAGAAT	ACCCCCAGUUUGAAAAGUGGUCUCAUCC	
	ACAACCGGAGCCAATGCCAAACA	UCAGUUCGAGAAGUGGUCUCAUCCACAG	
	GCAGAGAGACGCATACAATGGAA	UUCGAAAAGCCGGCAGCCAAACGGGUUA	
	TGCGAGAATGGACCGACAACAAT	AACUGGAUGCAUCUCAGGGAACCAAAAG	
	ATGTGGAGCCACCCCCAGTTTGA	AAGUUAUGAGCAAAUGGAGACGGACGGC	
	AAAGTGGTCTCATCCTCAGTTCG	GAGCGGCAGUAAAGACCAGACAGCUGGA	
	AGAAGTGGTCTCATCCACAGTTC	GCCUCGGUGGCCAUGCUUCUUGCCCCUU	
	GAAAAGCCGGCAGCCAAACGGGT	GGGCCUCCCCCCCAGCCCCUCCUCCCCUU	
	TAAACTGGATGCATCTCAGGGAA	CCUGCACCCGUACCCCCGUGGUCUUUGA	
	CCAAAAGAAGTTATGAGCAAATG	AUAAAGUCUGAGUGGGCGGCAAGCUUUU	
	GAGACGGACGGCGAGCGGCAGTA	CUUCUCUCCCUUGCACCUAAAAAAAAAA	
	AAGACCAGACAGCTGGAGCCTCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGCCATGCTTCTTGCCCCTTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGCCTCCCCCCAGCCCCTCCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCTTCCTGCACCCGTACCCCCGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGTCTTTGAATAAAGTCTGAGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGCGGCAAGCTTTTCTTCTCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGCACCT	AAAAAAAAAAAAAAAAAAAAAA	
46	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MSSSTTKKFREFMAEPL
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	GDKAIRDVPGIGDVLGT
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGAGCUCAAGCACUACA	KLSEAGYEKAYTLFGKY
	GCGGTAGACGCACGGCATAGCAT	AAGAAGUUUCGAGAAUUUAUGGCUGAAC	LLCNKDVERFQDWIQTQ
	CCAACGCCACCATGAGCTCAAGC	CACUUGGCGACAAGGCUAUCAGGGACGU	CGANSHQAKTAAQALSE
	ACTACAAAGAAGTTTCGAGAATT	GCCCGGGAUUGGGGAUGUACUCGGAACA	WAEAFIWSHPQFEKWSH
	TATGGCTGAACCACTTGGCGACA	AAAUUAAGCGAAGCAGGUUAUGAAAAGG	PQFEKWSHPQFEKPAAK
	AGGCTATCAGGGACGTGCCCGGG	CAUACACUCUGUUCGGUAAAUAUCUGCU	RVKLDASQGTKRSYEQM
	ATTGGGGATGTACTCGGAACAAA	GUGCAAUAAGGACGUAGAGAGAUUUCAG	ETDGERQ

	ATTAAGCGAAGCAGGTTATGAAA AGGCATACACTCTGTTCGGTAAA TATCTGCTGTGCAATAAGGACGT AGAGAGATTTCAGGACTGGATTC AGACCCAGTGTGGTGCAAATAGC CACCAGGCCAAAACGGCTGCCCA GGCCCTGTCCGAATGGGCCGAAG CTTTCATCTGGAGCCACCCGCAG TTTGAGAAATGGTCACACCCCCA GTTCGAGAAGTGGTCCCACCCCC AGTTCGAAAAACCCGCCGCAAAG AGAGTAAAACTTGATGCATCCCA AGGCACCAAACGATCCTACGAGC AGATGGAAACGGACGGTGAACGC CAATAAAGACCAGACAGCTGGAG CCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACC CCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCAAGCTTTTCTTC TCTCCCTTGCACCT	GACUGGAUUCAGACCCAGUGUGGUGCAA AUAGCCACCAGGCCAAAACGGCUGCCCA GGCCCUGUCCGAAUGGGCCGAAGCUUUC AUCUGGAGCCACCCGCAGUUUGAGAAAU GGUCACACCCCCAGUUCGAGAAGUGGUC CCACCCCCAGUUCGAAAAACCCGCCGCA AAGAGAGUAAAACUUGAUGCAUCCCAAG GCACCAAACGAUCCUACGAGCAGAUGGA AACGGACGGUGAACGCCAAUAAAGACCA GACAGCUGGAGCCUCGGUGGCCAUGCUU CUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCG UGGUCUUUGAAUAAAGUCUGAGUGGGCG GCAAGCUUUUCUUCUCUCCCUUGCACCU AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAA	
47	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTTCACGA AAACATAGGCAGTTTTGCTTAGA ACCTATGGGAGACAAATCCGTTT ATAACCTGCCCGGGATCGGACGA GTATTAGGCGGGAGGCTGCAGGA TAGCGGAATCATTCATGCTCGCG ATATCGAAGGACGATTCCTGGTC TTCAATCAGAATCAGGAAAAATT TGGAAATTGGTTAAGAGATACAT GCGGGGCTAATGCAAAACAACAG CGAGACTGCTATAACGGATTGCG CGAGTGGACTAGGAATAATCTGT GGTCACATCCACAATTCGAAAAA TGGTCACACCCACAGTTCGAGAA GTGGTCACATCCCCAGTTTGAGA AACCGGCCGCTAAGAGAGTCAAG CTCGACGCTTCTCAGGGAACAAA ACGTTCTTATGAGCAGATGGAAA CCGACGGCGAACGGCAGTAAAGA CCAGACAGCTGGAGCCTCGGTGG CCATGCTTCTTGCCCCTTGGGCC TCCCCCCAGCCCCTCCTCCCCTT CCTGCACCCGTACCCCCGTGGTC TTTGAATAAAGTCTGAGTGGGCG GCAAGCTTTTCTTCTCTCCCTTG CACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUUCACGAAAACAU AGGCAGUUUUGCUUAGAACCUAUGGGAG ACAAAUCCGUUUAUAACCUGCCCGGGAU CGGACGAGUAUUAGGCGGGAGGCUGCAG GAUAGCGGAAUCAUUCAUGCUCGCGAUA UCGAAGGACGAUUCCUGGUCUUCAAUCA GAAUCAGGAAAAAUUUGGAAAUUGGUUA AGAGAUACAUGCGGGGCUAAUGCAAAAC AACAGCGAGACUGCUAUAACGGAUUGCG CGAGUGGACUAGGAAUAAUCUGUGGUCA CAUCCACAAUUCGAAAAAUGGUCACACC CACAGUUCGAGAAGUGGUCACAUCCCCA GUUUGAGAAACCGGCCGCUAAGAGAGUC AAGCUCGACGCUUCUCAGGGAACAAAAC GUUCUUAUGAGCAGAUGGAAACCGACGG CGAACGGCAGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCAAGCUUU UCUUCUCUCCCUUGCACCUAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	MTSRKHRQFCLEPMGDK SVYNLPGIGRVLGGRLQ DSGIIHARDIEGRFLVF NQNQEKFGNWLRDTCGA NAKQQRDCYNGLREWTR NNLWSHPQFEKWSHPQF EKWSHPQFEKPAAKRVK LDASQGTKRSYEQMETD GERQ

48	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGTCCACATCT GTGAAACATGCTAACTTTGTTTC CGAAACCATGGGGGACAAGCCAG TCACAGAGGTAGCCGGCATAGGG GAAGTGTACGGCAAGAAACTTAA GGACGACGGATTTAAAAAAGCAT ATCATCTCTATGGTCAATTCCTG ATTCTGGATAAGGACAAAGAGCG GTTCAGTGAGTGGCTCAAGGACA CCTACAGCGTGAACGCTAAGCAT GCCAATGCTTGTGCCGAATGCCT GTCAGAATATGCTGAACAGCACT TATGGTCCCACCCACAGTTCGAA AAATGGAGTCACCCCCAGTTTGA GAAATGGTCTCATCCCCAATTCG AGAAACCTGCAGCTAAAAGAGTT AAATTGGACGCGTCCCAAGGTAC AAAGAGGAGCTATGAACAAATGG AGACCGACGGAGAAAGGCAGTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUCCACAUCUGUGAAA CAUGCUAACUUUGUUUCCGAAACCAUGG GGGAUAAGCCAGUCACAGAGGUAGCCGG CAUAGGGGAAGUGUACGGCAAGAAACUU AAGGACGACGGAUUUAAAAAAGCAUAUC AUCUCUAUGGUCAAUUCCUGAUUCUGGA UAAGGACAAAGAGCGGUUCAGUGAGUGG CUCAAGGACACCUACAGCGUGAACGCUA AGCAUGCCAAUGCUUGUGCCGAAUGCCU GUCAGAAUAUGCUGAACAGCACUUAUGG UCCCACCCACAGUUCGAAAAAUGGAGUC ACCCCCAGUUUGAGAAAUGGUCUCAUCC CCAAUUCGAGAAACCUGCAGCUAAAAGA GUUAAAUUGGACGCGUCCCAAGGUACAA AGAGGAGCUAUGAACAAAUGGAGACCGA CGGAGAAAGGCAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	MSTSVKHANFVSETMGD KPVTEVAGIGEVYGKKL KDDGFKKAYHLYGQFLI LDKDKERFSEWLKDTYS VNAKHANACAECLSEYA EQHLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
49	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGCTACCAGG TCCCAGAAGTTTAAAGACTTTGT CTCAGAGCCTATGGGGAACAAGA CAGTGACCGAGGTCGATGGTATC GATGAAGAGCTTGGGCCTAAGTT GGCAGCCGAGGGCTTTGATAAAG CTTATATACTGCTCGGACAGTTT TTGTTGTTAAAGAAAGATGCCTT GAGATTCCAACAGTGGTTGAAGG AAACCTTCGGGGCCTCTAGTCGC CAGGCTGTACAATGTGCAACATG CCTCACGGAGTGGTGCTACACCA TACTCTGGTCACACCCTCAATTT GAAAAATGGTCCCACCCTCAGTT CGAAAAATGGAGCCACCCACAGT TCGAAAAACCAGCAGCAAAAAGA GTGAAGCTCGACGCCAGCCAAGG TACGAAGCGGAGTTATGAACAGA TGGAAACAGACGGTGAAAGACAA TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCUACCAGGUCCCAG AAGUUUAAAGACUUUGUCUCAGAGCCUA UGGGGAACAAGACAGUGACCGAGGUCGA UGGUAUCGAUGAAGAGCUUGGGUCUAAG UUGGCAGCCGAGGGCUUUGAUAAAGCUU AUAUACUGCUCGGACAGUUUUUGUUGUU AAAGAAAGAUGCCUUGAGAUUCCAACAG UGGUUGAAGGAAACCUUCGGGGCCUCUA GUCGCCAGGCUGUACAAUGUGCAACAUG CCUCACGGAGUGGUGUUACACCAUACUC UGGUCACACCCUCAAUUUGAAAAAUGGU CCCACCCUCAGUUCGAAAAAUGGAGCCA CCCACAGUUCGAAAAACCAGCAGCAAAA AGAGUGAAGCUCGACGCCAGCCAAGGUA CGAAGCGGAGUUAUGAACAGAUGGAAAC AGACGGUGAAAGACAAUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	MATRSQKFKDFVSEPMG NKTVTEVDGIDEELGSK LAAEGFDKAYILLGQFL LLKKDALRFQQWLKETF GASSRQAVQCATCLTEW CYTILWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ

	GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA A	
50	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGATGGCCACG AGGACCCAGAAATTTAAGGCCTT TGTCAGCGAACCTATGGGCGATA AGGATATTACTGCTGTTGACGGC ATAAATGACGAATTGGGTCTTAA GTTGGCCGCAAAGGGATTTGACA AAGCCTACATCCTTCTGGGACAA TTTTTATTACTCAAGAAAGATTC CGCCCGGTTCCAAAGATGGCTTA AAGGGACATACGGCGCCAGTCCT TCCGAAGCCCAACAATGTGCCTC CTGCCTTCTGGAGTGGTGTTATG CTTTCCTTTGGTCCCACCCCCAA TTTGAAAAGTGGAGCCATCCCCA ATTCGAAAAATGGAGCCACCCCC AGTTTGAGAAGCCTGCCGCAAAG AGAGTGAAACTGGACGCTAGTCA GGGCACGAAGAGATCCTATGAAC AGATGGAGACAGACGGTGAACGA CAGTAAAGACCAGACAGCTGGAG CCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACC CCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCAAGCTTTTCTTC TCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAUGGCCACGAGGACC CAGAAAUUUAAGGCCUUUGUCAGCGAAC CUAUGGGCGAUAAGGAUAUUACUGCUGU UGACGGCAUAAAUGACGAAUUGGGUCUU AAGUUGGCCGCAAAGGGAUUUGACAAAG CCUACAUCCUUCUGGGACAAUUUUUAUU ACUCAAGAAAGAUUCCGCCCGGUUCCAA AGAUGGCUUAAAGGGACAUACGGCGCCA GUCCUUCCGAAGCCCAACAAUGUGCCUC CUGCCUUCUGGAGUGGUGUUAUGCUUUC CUUUGGUCCCACCCCCAAUUUGAAAAGU GGAGCCAUCCCCCAAUUCGAAAAAUGGAG CCACCCCCAGUUUGAGAAGCCUGCCGCA AAGAGAGUGAAACUGGACGCUAGUCAGG GCACGAAGAGAUCCUAUGAACAGAUGGA GACAGACGGUGAACGACAGUAAAGACCA GACAGCUGGAGCCUCGGUGGCCAUGCUU CUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCG UGGUCUUUGAAUAAAGUCUGAGUGGGCG GCAAGCUUUUCUUCUCUCCCUUGCACCU AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAA	
51	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACGTCA CAAGCACACAGAGACTTCGTGGC GGAGCCTATGGGCGAGAAACCTG TGGGTTCCTTGAAAGGCATTGGT GATGTACTGAGTAAGAGGTTGGA GGAACGTGGATTCGACAAAGCCT ATGTCGTGCTTGGGCAGTTTTTG GTACTTAAAAAGGATGAGGATAG ATTTAGAGAGTGGCTTAAGGATA CTTGTGGCGCTAACGCCAAGCAG TCACGAGATTGTTTCGGATGCCT TCGAGAGTGGTGCGATGCTTTTC TCTGGTCCCACCCACAGTTCGAA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACGUCACAAGCA CACAGAGACUUCGUGGCGGAGCCUAUGG GCGAGAAACCUGUGGGUUCCUUGAAAGG CAUUGGUGAUGUACUGAGUAAGAGGUUG GAGGAACGUGGAUUCGACAAAGCCUAUG UCGUGCUUGGGCAGUUUUUGGUACUUAA AAAGGAUGAGGAUAGAUUUAGAGAGUGG CUUAAGGAUACUUGUGGCGCUAACGCCA AGCAGUCACGAGAUUGUUUCGGAUGCCU UCGAGAGUGGUGCGAUGCUUUUCUCUGG UCCCACCCACAGUUCGAAAAAUGGAGCC AUCCCCAGUUUGAGAAAUGGUCCCAUCC ACAGUUCGAAAAACCAGCCGCUAAAAGA GUGAAGCUUGACGCUAGCCAGGGUACUA AGAGAUCCUAUGAACAGAUGGAAACCGA	MTTSQAHRDFVAEPMGE KPVGSLKGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	AAATGGAGCCATCCCCAGTTTGA	UGGAGAACGCCAGUAAAGACCAGACAGC	
	GAAATGGTCCCATCCACAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCAGCCGCTAAAAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTTGACGCTAGCCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGAGATCCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGAGAACGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAA	
52	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTASQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCGCCUCACAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGACUUUGUGGCUGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCGCCTCA	GAGAGAAGCCCGUCGGAAGCCUGGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACATAGAGACTTTGTGGC	UAUCGGGGACGUGCUGUCCAAGCGAUUA	DAFLWSHPQFEKWSHPQ
	TGAACCCATGGGAGAGAAGCCCG	GAAGAGAGAGGAUUCGACAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TCGGAAGCCTGGCCGGTATCGGG	UGGUCCUCGGUCAGUUUCUUGUACUGAA	KLDASQGTKRSYEQMET
	GACGTGCTGTCCAAGCGATTAGA	AAAAGACGAAGAUCGAUUCAGAGAGUGG	DGERQ
	AGAGAGAGGATTCGACAAGGCCT	CUCAAAGAUACCUGCGGGGCUAACGCUA	
	ACGTGGTCCTCGGTCAGTTTCTT	AGCAGAGCCGUGACUGCUUUGGCUGCUU	
	GTACTGAAAAAAGACGAAGATCG	AAGAGAAUGGUGCGAUGCUUUCUUGUGG	
	ATTCAGAGAGTGGCTCAAAGATA	AGCCACCCCCAGUUCGAGAAGUGGUCCC	
	CCTGCGGGGCTAACGCTAAGCAG	ACCCCCAGUUCGAAAAGUGGUCUCAUCC	
	AGCCGTGACTGCTTTGGCTGCTT	ACAGUUUGAGAAGCCCGCCGCAAAAAGG	
	AAGAGAATGGTGCGATGCTTTCT	GUGAAACUUGAUGCCUCACAAGGAACAA	
	TGTGGAGCCACCCCCAGTTCGAG	AAAGGAGCUAUGAGCAGAUGGAAACCGA	
	AAGTGGTCCCACCCCCAGTTCGA	UGGGGAACGGCAGUAAAGACCAGACAGC	
	AAAGTGGTCTCATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCCGCAAAAAGGGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTTGATGCCTCACAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAAAGGAGCTATGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGATGGGGAACGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAA	
53	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTAQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAGCACAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGACUUUGUGGCUGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACAGCA	GGGAGAAACCUGUCGGAAGUCUCGCGGG	ANAKQSRDCFGCLREWC
	CAAAAACATAGAGACTTTGTGGC	UAUUGGUGACGUGUUAUCCAAGAGGUUA	DAFLWSHPQFEKWSHPQ
	TGAGCCTATGGGGGAGAAACCTG	GAGGAGCGGGGUUUCGACAAAGCCUACG	FEKWSHPQFEKPAAKRV
	TCGGAAGTCTCGCGGGTATTGGT	UAGUGCUCGGGCAAUUUCUGGUGCUGAA	KLDASQGTKRSYEQMET
	GACGTGTTATCCAAGAGGTTAGA	GAAGGAUGAGGACCGGUUUAGGGAGUGG	DGERQ
	GGAGCGGGGTTTCGACAAAGCCT	CUGAAAGACACAUGUGGAGCCAACGCUA	
	ACGTAGTGCTCGGGCAATTTCTG	AACAAAGCAGGGAUUGUUUUGGGUGUUU	
	GTGCTGAAGAAGGATGAGGACCG	AAGAGAAUGGUGUGACGCCUUCCUUUGG	
	GTTTAGGGAGTGGCTGAAAGACA	UCUCACCCGCAGUUCGAGAAGUGGAGCC	
	CATGTGGAGCCAACGCTAAACAA	ACCCUCAAUUUGAGAAGUGGUCACAUCC	

	AGCAGGGATTGTTTTGGGTGTTT	ACAGUUUGAGAAACCAGCAGCAAAGAGG	
	AAGAGAATGGTGTGACGCCTTCC	GUUAAAUUAGACGCUUCACAGGGGACCA	
	TTTGGTCTCACCCGCAGTTCGAG	AGAGAUCCUAUGAACAGAUGGAGACCGA	
	AAGTGGAGCCACCCTCAATTTGA	CGGCGAACGACAGUAAAGACCAGACAGC	
	GAAGTGGTCACATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCAGCAGCAAAGAGGGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTAGACGCTTCACAGGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGAGATCCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACCGACGGCGAACGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
54	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSAKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCAGCGCCAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGACUUCGUGGCAGAGCCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCAGC	GAGAAAAGCCCGUCGGGUCUUUGGCUGG	ANAKQSRDCFGCLREWC
	GCCAAGCACAGGGACTTCGTGGC	CAUCGGUGACGUAUUAAGUAAAAGACUG	DAFLWSHPQFEKWSHPQ
	AGAGCCAATGGGAGAAAAGCCCG	GAAGAAAGAGGCUUUGAUAAGGCGUAUG	FEKWSHPQFEKPAAKRV
	TCGGGTCTTTGGCTGGCATCGGT	UGGUCUUGGGGCAGUUCCUUGUGCUUAA	KLDASQGTKRSYEQMET
	GACGTATTAAGTAAAAGACTGGA	GAAAGACGAAGAUCGGUUUAGAGAAUGG	DGERQ
	AGAAAGAGGCTTTGATAAGGCGT	CUGAAGGAUACUUGCGGUGCAAACGCUA	
	ATGTGGTCTTGGGGCAGTTCCTT	AACAGAGUCGCGAUUGUUUCGGAUGUUU	
	GTGCTTAAGAAAGACGAAGATCG	GCGAGAGUGGUGUGAUGCCUUCCUGUGG	
	GTTTAGAGAATGGCTGAAGGATA	AGUCACCCCCAGUUCGAAAAAUGGAGUC	
	CTTGCGGTGCAAACGCTAAACAG	ACCCUCAGUUCGAGAAAUGGUCUCAUCC	
	AGTCGCGATTGTTTCGGATGTTT	ACAGUUUGAGAAACCUGCAGCCAAACGG	
	GCGAGAGTGGTGTGATGCCTTCC	GUGAAACUUGACGCAUCACAAGGAACUA	
	TGTGGAGTCACCCCCAGTTCGAA	AACGGAGUUAUGAACAGAUGGAAACAGA	
	AAATGGAGTCACCCTCAGTTCGA	UGGAGAGAGACAGUAAAGACCAGACAGC	
	GAAATGGTCTCATCCACAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCTGCAGCCAAACGGGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTTGACGCATCACAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAACGGAGTTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGATGGAGAGAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
55	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKARDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACCUCACAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	GCAAGAGAUUUUGUGGCUGAGCCGAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACCTCA	GUGAGAAGCCAGUAGGAAGUCUGGCCGG	ANAKQSRDCFGCLREWC
	CAGAAAGCAAGAGATTTTGTGGC	CAUCGGUGAUGUGUUGUCCAAACGCCUG	DAFLWSHPQFEKWSHPQ
	TGAGCCGATGGGTGAGAAGCCAG	GAAGAACGAGGCUUCGACAAAGCCUAUG	FEKWSHPQFEKPAAKRV
	TAGGAAGTCTGGCCGGCATCGGT	UCGUCUUGGGACAGUUUUUGGUUUUAAA	KLDASQGTKRSYEQMET
	GATGTGTTGTCCAAACGCCTGGA	AAAAGAUGAGGACCGCUUCCGGGAGUGG	DGERQ
	AGAACGAGGCTTCGACAAAGCCT	CUGAAAGAUACUUGUGGUGCAAAUGCCA	
	ATGTCGTCTTGGGACAGTTTTTG	AACAGUCUAGAGAUUGCUUCGGUUGCCU	

	GTTTTAAAAAAAGATGAGGACCG	UAGGGAAUGGUGCGAUGCGUUUUUAUGG	
	CTTCCGGGAGTGGCTGAAAGATA	UCACACCCACAGUUCGAGAAAUGGAGCC	
	CTTGTGGTGCAAATGCCAAACAG	AUCCCCAGUUUGAGAAAUGGUCUCAUCC	
	TCTAGAGATTGCTTCGGTTGCCT	ACAAUUCGAGAAGCCCGCUGCCAAAAGA	
	TAGGGAATGGTGCGATGCGTTTT	GUAAAACUGGAUGCUUCUCAGGGCACAA	
	TATGGTCACACCCACAGTTCGAG	AACGAUCCUAUGAACAGAUGGAGACUGA	
	AAATGGAGCCATCCCCAGTTTGA	CGGCGAACGCCAGUAAAGACCAGACAGC	
	GAAATGGTCTCATCCACAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCTGCCAAAAGAGTA	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGATGCTTCTCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAACGATCCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACTGACGGCGAACGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAA	
56	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHADFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACGUCUCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACGCCGAUUUCGUUGCCGAGCCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACGTCT	GGGAAAAACCUGUCGGAUCACUGGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACACGCCGATTTCGTTGC	UAUCGGUGAUGUGUUAAGUAAAAGACUG	DAFLWSHPQFEKWSHPQ
	CGAGCCAATGGGGGAAAAACCTG	GAGGAAAGGGGAUUCGACAAAGCUUACG	FEKWSHPQFEKPAAKRV
	TCGGATCACTGGCCGGTATCGGT	UAGUUUUAGGCCAGUUUCUGGUCCUCAA	KLDASQGTKRSYEQMET
	GATGTGTTAAGTAAAAGACTGGA	AAAGGAUGAAGAUAGGUUCCGAGAGUGG	DGERQ
	GGAAAGGGGATTCGACAAAGCTT	UUGAAAGACACCUGCGGCGCCAACGCAA	
	ACGTAGTTTTAGGCCAGTTTCTG	AACAGUCACGAGAUUGCUUUGGUUGUCU	
	GTCCTCAAAAAGGATGAAGATAG	GAGAGAGUGGUGCGAUGCAUUUCUUUGG	
	GTTCCGAGAGTGGTTGAAAGACA	AGUCAUCCUCAGUUUGAAAAAUGGUCAC	
	CCTGCGGCGCCAACGCAAAACAG	ACCCUCAGUUUGAGAAAUGGAGCCACCC	
	TCACGAGATTGCTTTGGTTGTCT	GCAGUUUGAGAAACCGGCCGCUAAGAGA	
	GAGAGAGTGGTGCGATGCATTTC	GUUAAAUUAGACGCCAGUCAAGGAACUA	
	TTTGGAGTCATCCTCAGTTTGAA	AGCGAUCAUACGAACAGAUGGAAACAGA	
	AAATGGTCACACCCTCAGTTTGA	CGGGGAGCGGCAGUAAAGACCAGACAGC	
	GAAATGGAGCCACCCGCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCGGCCGCTAAGAGAGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTAGACGCCAGTCAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGCGATCATACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGACGGGGAGCGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAA	

57	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRAFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCUCCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGCCUUUGUUGCCGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCTCC	GAGAGAAACCAGUCGGCUCUCUCGCAGG	ANAKQSRDCFGCLREWC
	CAAAAGCACCGAGCCTTTGTTGC	UAUUGGCGAUGUGCUUAGCAAGCGGUUA	DAFLWSHPQFEKWSHPQ
	CGAACCCATGGGAGAGAAACCAG	GAAGAGAGGGGCUUUGAUAAAGCUUAUG	FEKWSHPQFEKPAAKRV
	TCGGCTCTCTCGCAGGTATTGGC	UGGUUCUCGGACAGUUUCUGGUCCUUAA	KLDASQGTKRSYEQMET
	GATGTGCTTAGCAAGCGGTTAGA	GAAGGAUGAAGACAGAUUCAGGGAAUGG	DGERQ
	AGAGAGGGGCTTTGATAAAGCTT	CUGAAGGACACCUGUGGAGCUAACGCCA	
	ATGTGGTTCTCGGACAGTTTCTG	AACAGAGUAGGGACUGCUUCGGGUGCCU	
	GTCCTTAAGAAGGATGAAGACAG	CCGGGAGUGGUGCGACGCGUUUCUUUGG	
	ATTCAGGGAATGGCTGAAGGACA	UCCCACCCACAGUUUGAAAAGUGGUCAC	
	CCTGTGGAGCTAACGCCAAACAG	ACCCUCAGUUCGAAAAAUGGUCACAUCC	
	AGTAGGGACTGCTTCGGGTGCCT	UCAGUUUGAGAAGCCUGCUGCUAAAAGG	
	CCGGGAGTGGTGCGACGCGTTTC	GUUAAACUCGACGCCUCUCAAGGAACAA	
	TTTGGTCCCACCCACAGTTTGAA	AGCGUAGUUAUGAACAAAUGGAGACUGA	
	AAGTGGTCACACCCTCAGTTCGA	CGGUGAGAGACAGUAAAGACCAGACAGC	
	AAAATGGTCACATCCTCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCTGCTGCTAAAAGGGTT	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTCGACGCCTCTCAAGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAGCGTAGTTATGAACAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACTGACGGTGAGAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
58	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDAVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAAGUCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGACGCAGUGGCAGAGCCGAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACAAGT	GAGAAAAGCCCGUGGGAAGCCUCGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACACAGAGACGCAGTGGC	CAUUGGAGACGUUCUGAGUAAGAGACUC	DAFLWSHPQFEKWSHPQ
	AGAGCCGATGGGAGAAAAGCCCG	GAAGAGAGGGGCUUCGAUAAAGCUUAUG	FEKWSHPQFEKPAAKRV
	TGGGAAGCCTCGCCGGCATTGGA	UAGUGCUGGGGCAAUUUCUGGUCCUGAA	KLDASQGTKRSYEQMET
	GACGTTCTGAGTAAGAGACTCGA	AAAAGACGAAGAUAGAUUUAGAGAAUGG	DGERQ
	AGAGAGGGGCTTCGATAAAGCTT	CUGAAAGAUACCUGCGGGGCAAAUGCUA	
	ATGTAGTGCTGGGGCAATTTCTG	AGCAAAGCCGAGAUUGUUUUGGGUGUCU	
	GTCCTGAAAAAAGACGAAGATAG	CAGGGAAUGGUGCGACGCCUUCCUGUGG	
	ATTTAGAGAATGGCTGAAAGATA	UCCCACCCACAAUUUGAAAAGUGGAGUC	
	CCTGCGGGGCAAATGCTAAGCAA	ACCCCCAGUUUGAAAAGUGGAGUCACCC	
	AGCCGAGATTGTTTTGGGTGTCT	ACAAUUUGAAAAACCCGCAGCUAAACGA	
	CAGGGAATGGTGCGACGCCTTCC	GUCAAACUUGAUGCUAGCCAAGGGACAA	
	TGTGGTCCCACCCACAATTTGAA	AAAGAAGCUAUGAACAGAUGGAAACAGA	
	AAGTGGAGTCACCCCCAGTTTGA	UGGAGAGCGCCAGUAAAGACCAGACAGC	
	AAAGTGGAGTCACCCACAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCCGCAGCTAAACGAGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTTGATGCTAGCCAAGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAAAGAAGCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGATGGAGAGCGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
59	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACAGCATCA CAGAAGCACAGAGACTTTGTGGC CGAACCAATGGGCGAGAAACCTG TCGGTTCCCTCAAAGGCATAGGA GACGTTTTGTCCAAAAGGTTAGA AGAAAGAGGCTTTGATAAAGCAT ATGTGGTTCTTGGCCAGTTTTTG GTTTTGAAGAAGGATGAAGATAG ATTTCGGGAGTGGCTGAAAGACA CATGCGGTGCCAACGCAAAGCAG TCTAGAGACTGTTTCGGATGTCT GAGGGAATGGTGTGACGCCTTTC TGTGGTCACACCCACAGTTTGAA AAATGGTCACACCCACAATTCGA GAAATGGTCTCATCCACAATTTG AGAAACCTGCAGCTAAAAGAGTT AAATTAGACGCTAGTCAAGGCAC CAAGAGGTCATACGAGCAAATGG AGACTGATGGAGAGCGCCAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAGCAUCACAGAAG CACAGAGACUUUGUGGCCGAACCAAUGG GCGAGAAACCUGUCGGUUCCCUCAAAGG CAUAGGAGACGUUUUGUCCAAAAGGUUA GAAGAAAGAGGCUUUGAUAAAGCAUAUG UGGUUCUUGGCCAGUUUUUGGUUUUGAA GAAGGAUGAAGAUAGAUUUCGGGAGUGG CUGAAAGACACAUGCGGUGCCAACGCAA AGCAGUCUAGAGACUGUUUCGGAUGUCU GAGGGAAUGGUGUGACGCCUUUCUGUGG UCACACCCACAGUUUGAAAAAUGGUCAC ACCCACAAUUCGAGAAAUGGUCUCAUCC ACAAUUUGAGAAACCUGCAGCUAAAAGA GUUAAAUUAGACGCUAGUCAAGGCACCA AGAGGUCAUACGAGCAAAUGGAGACUGA UGGAGAGCGCCAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	MTASQKHRDFVAEPMGE KPVGSLKGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
60	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACCAGC CAGAAGCACCGAGATTTTGTAAC TGAACCAATGGGAGAAAAGCCTG TGGGCTCTCTCGCGGGAATTGGC GATGTGCTTTCAAAGAGACTGGA AGAACGGGGATTCGATAAAGCAT ACGTTGTGCTGGGACAATTTCTC GTATTAAAGAAGGACGAAGACCG ATTTAGAGAGTGGTTAAAAGATA CATGTGGAGCTAACGCCAAGCAG AGCCGGGACTGTTTCGGATGTCT GAGGGAGTGGTGCGACGCCTTTC TTTGGAGTCATCCACAGTTCGAA AAATGGAGTCACCCCCAGTTCGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACCAGCCAGAAG CACCGAGAUUUUGUAACUGAACCAAUGG GAGAAAAGCCUGUGGGCUCUCUCGCGGG AAUUGGCGAUGUGCUUUCAAAGAGACUG GAAGAACGGGGAUUCGAUAAAGCAUACG UUGUGCUGGGACAAUUUCUCGUAUUAAA GAAGGACGAAGACCGAUUUAGAGAGUGG UUAAAAGAUACAUGUGGAGCUAACGCCA AGCAGAGCCGGGACUGUUUCGGAUGUCU GAGGGAGUGGUGCGACGCCUUUCUUUGG AGUCAUCCACAGUUCGAAAAAUGGAGUC ACCCCCAGUUCGAGAAGUGGUCCCAUCC ACAGUUCGAAAAACCAGCGGCCAAGCGC GUGAAAUUAGAUGCAUCUCAAGGGACCA AACGAUCCUAUGAGCAGAUGGAAACUGA UGGCGAGCGUCAGUAAAGACCAGACAGC	MTTSQKHRDFVTEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	GAAGTGGTCCCATCCACAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCAGCGGCCAAGCGCGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAATTAGATGCATCTCAAGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAACGATCCTATGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACTGATGGCGAGCGTCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
61	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAUCUCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGAUUUCGUGGCAGAACCGAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACATCT	GGGAGAAGCCGGUUGGCUCCCUCGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACACCGAGATTTCGTGGC	CAUUGGAGAUGUGCUGAGCAAGCGGCUC	DAFLWSHPQFEKWSHPQ
	AGAACCGATGGGGGAGAAGCCGG	GAGGAGAGAGGAUUCGACAAAGCAUAUG	FEKWSHPQFEKPAAKRA
	TTGGCTCCCTCGCCGGCATTGGA	UUGUGCUGGGACAGUUCCUGGUCCUCAA	KLDASQGTKRSYEQMET
	GATGTGCTGAGCAAGCGGCTCGA	GAAGGACGAAGAUAGAUUUAGAGAAUGG	DGERQ
	GGAGAGAGGATTCGACAAAGCAT	UUAAAGGAUACUUGUGGCGCGAAUGCCA	
	ATGTTGTGCTGGGACAGTTCCTG	AGCAGAGCAGAGAUUGCUUUGGGUGCCU	
	GTCCTCAAGAAGGACGAAGATAG	GCGGGAAUGGUGCGAUGCUUUCCUGUGG	
	ATTTAGAGAATGGTTAAAGGATA	AGCCACCCCCAGUUCGAAAAGUGGAGUC	
	CTTGTGGCGCGAATGCCAAGCAG	ACCCUCAGUUUGAAAAAUGGUCACACCC	
	AGCAGAGATTGCTTTGGGTGCCT	CCAAUUCGAAAAACCUGCAGCUAAAAGG	
	GCGGGAATGGTGCGATGCTTTCC	GCCAAACUGGACGCCAGCCAGGGUACCA	
	TGTGGAGCCACCCCCAGTTCGAA	AAAGGUCCUAUGAACAGAUGGAGACAGA	
	AAGTGGAGTCACCCTCAGTTTGA	UGGAGAAAGACAGUAAAGACCAGACAGC	
	AAAATGGTCACACCCCCAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCTGCAGCTAAAAGGGCC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGACGCCAGCCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAAAGGTCCTATGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACAGATGGAGAAAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
62	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAAGCCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGAGAUUUUGUGGCCGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACAAGC	GGGAGAAACCCGUCGGAAGUCUGGCUGG	ANAKQSRDCFGCLREWC
	CAGAAACATCGAGATTTGTGGC	UAUUGGUGAUGUGUUAAGUAAGAGAACUG	DAFLWSHPQFEKWSHPQ
	CGAACCCATGGGGGAGAAACCCG	GAAGAACGCGGCUUCGAUAAAGCUUAUG	FEKWSHPQFEKPAAKRV
	TCGGAAGTCTGGCTGGTATTGGT	UGGUAUUAGGCCAGUUCCUGGUGCUGAA	ALDASQGTKRSYEQMET
	GATGTGTTAAGTAAGAGACTGGA	AAAGGAUGAGGACCGUUUUCGCGAAUGG	DGERQ
	AGAACGCGGCTTCGATAAAGCTT	CUCAAGGAUACCUGUGGCGCUAACGCUA	
	ATGTGGTATTAGGCCAGTTCCTG	AGCAGAGCAGGGAUUGUUUUGGAUGCCU	
	GTGCTGAAAAAGGATGAGGACCG	GCGGGAGUGGUGCGACGCGUUUUUGUGG	
	TTTTCGCGAATGGCTCAAGGATA	UCACAUCCCCAGUUCGAAAAGUGGUCCC	
	CCTGTGGCGCTAACGCTAAGCAG	ACCCCCAAUUCGAAAAAUGGAGUCACCC	
	AGCAGGGATTGTTTTGGATGCCT	ACAAUUUGAGAAGCCAGCCGCAAAAAGA	

	GCGGGAGTGGTGCGACGCGTTTT	GUCGCACUCGACGCGUCCCAGGGUACUA	
	TGTGGTCACATCCCCAGTTCGAA	AGAGAUCAUAUGAGCAAAUGGAAACAGA	
	AAGTGGTCCCACCCCCAATTCGA	UGGCGAGAGACAGUAAAGACCAGACAGC	
	AAAATGGAGTCACCCACAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCAGCCGCAAAAAGAGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	GCACTCGACGCGTCCCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGAGATCATATGAGCAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGATGGCGAGAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
63	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACUUCCCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGAUUUCGUUGCUGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACTTCC	GAGAAAAGCCAGUGGGCUCACUCGCAGG	ANAKQSRDCFGCLREWC
	CAAAAACACCGAGATTTCGTTGC	GAUUGGAGAUGUUCUCUCCAAAAGACUG	DAFLWSHPQFEKWSHPQ
	TGAACCTATGGGAGAAAAGCCAG	GAGGAGAGAGGGUUCGACAAAGCUUAUG	FEKWSHPQFEKPAAKRV
	TGGGCTCACTCGCAGGGATTGGA	UAGUCUUGGGACAGUUCCUGGUGUUAAA	KADASQGTKRSYEQMET
	GATGTTCTCTCCAAAAGACTGGA	GAAGGAUGAGGAUAGAUUUAGAGAAUGG	DGERQ
	GGAGAGAGGGTTCGACAAAGCTT	CUCAAGGACACGUGCGGCGCCAAUGCUA	
	ATGTAGTCTTGGGACAGTTCCTG	AACAGAGUAGAGAUUGCUUUGGCUGUCU	
	GTGTTAAAGAAGGATGAGGATAG	UCGGGAAUGGUGUGACGCUUUUCUCUGG	
	ATTTAGAGAATGGCTCAAGGACA	AGCCACCCUCAGUUCGAAAAGUGGUCCC	
	CGTGCGGCGCCAATGCTAAACAG	ACCCUCAGUUCGAAAAAUGGUCACACCC	
	AGTAGAGATTGCTTTGGCTGTCT	GCAAUUUGAGAAGCCCGCCGCUAAGCGA	
	TCGGGAATGGTGTGACGCTTTTC	GUGAAAGCUGACGCAUCACAAGGCACAA	
	TCTGGAGCCACCCTCAGTTCGAA	AACGCUCCUACGAACAGAUGGAGACCGA	
	AAGTGGTCCCACCCTCAGTTCGA	CGGUGAACGCCAGUAAAGACCAGACAGC	
	AAAATGGTCACACCCGCAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCCGCTAAGCGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAAGCTGACGCATCACAAGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAACGCTCCTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACCGACGGTGAACGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
64	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACACAUCUCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGCGAUUUCGUGGCCGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACATCT	GCGAAAAGCCUGUCGGAAGCUUAGCAGG	ANAKQSRDCFGCLREWC
	CAGAAGCACCGCGATTTCGTGGC	CAUAGGCGAUGUUCUGAGUAAACGCCUG	DAFLWSHPQFEKWSHPQ
	CGAACCTATGGGCGAAAAGCCTG	GAAGAAAGAGGGUUUGAUAAGGCAUAUG	FEKWSHPQFEKPAAARV
	TCGGAAGCTTAGCAGGCATAGGC	UGGUGCUGGGGCAGUUUCUGGUCUUGAA	KLDASQGTKRSYEQMET
	GATGTTCTGAGTAAACGCCTGGA	AAAAGAUGAGGACCGAUUCCGAGAGUGG	DGERQ
	AGAAAGAGGGTTTGATAAGGCAT	UUGAAGGAUACCUGCGGAGCCAAUGCCA	
	ATGTGGTGCTGGGGCAGTTTCTG	AGCAAUCCCGAGAUUGCUUCGGAUGUUU	
	GTCTTGAAAAAAGATGAGGACCG	GCGAGAAUGGUGUGACGCAUUUCUCUGG	

	ATTCCGAGAGTGGTTGAAGGATA	AGUCACCCCCAGUUCGAGAAAUGGAGCC	
	CCTGCGGAGCCAATGCCAAGCAA	AUCCCCAAUUCGAAAAGUGGAGUCACCC	
	TCCCGAGATTGCTTCGGATGTTT	CCAGUUUGAGAAGCCCGCCGCUGCCAGA	
	GCGAGAATGGTGTGACGCATTTC	GUUAAGUUGGAUGCAAGCCAGGGCACAA	
	TCTGGAGTCACCCCCAGTTCGAG	AAAGAUCAUACGAGCAGAUGGAAACCGA	
	AAATGGAGCCATCCCCAATTCGA	CGGCGAACGGCAAUAAAGACCAGACAGC	
	AAAGTGGAGTCACCCCCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAGCCCGCCGCTGCCAGAGTT	CCUUGGGCCUCCCCCCCAGCCCCUCCUCC	
	AAGTTGGATGCAAGCCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAAAGATCATACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACCGACGGCGAACGGCAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
65	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAUCACAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGAUUUCGUAGCUGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACATCA	GCGAAAAACCCGUGGGCAGUCUGGCCGG	ANAKQSRDCFGCLREWC
	CAAAAACACCGAGATTTCGTAGC	CAUUGGGGAUGUGCUGUCAAAGCGACUG	DAFLWSHPQFEKWSHPQ
	TGAACCAATGGGCGAAAAACCCG	GAAGAACGAGGCUUCGACAAAGCGUAUG	FEKWSHPQFEKPAAKAV
	TGGGCAGTCTGGCCGGCATTGGG	UGGUUCUGGGGCAAUUUUUGGUACUUAA	KLDASQGTKRSYEQMET
	GATGTGCTGTCAAAGCGACTGGA	AAAAGAUGAGGAUAGGUUCCGGGAGUGG	DGERQ
	AGAACGAGGCTTCGACAAAGCGT	UUGAAGGACACUUGUGGUGCCAAUGCCA	
	ATGTGGTTCTGGGGCAATTTTTG	AGCAAUCUAGAGACUGUUUUGGAUGUCU	
	GTACTTAAAAAAGATGAGGATAG	CAGAGAAUGGUGUGAUGCAUUUUUAUGG	
	GTTCCGGGAGTGGTTGAAGGACA	AGCCACCCCCAGUUUGAAAAAUGGAGCC	
	CTTGTGGTGCCAATGCCAAGCAA	ACCCACAAUUUGAAAAAUGGUCCCAUCC	
	TCTAGAGACTGTTTTGGATGTCT	UCAAUUCGAAAAGCCCGCCGCAAAGGCC	
	CAGAGAATGGTGTGATGCATTTT	GUCAAGCUGGAUGCUUCACAGGGGACUA	
	TATGGAGCCACCCCCAGTTTGAA	AACGAAGCUACGAGCAAAUGGAGACAGA	
	AAATGGAGCCACCCACAATTTGA	CGGAGAGCGCCAGUAAAGACCAGACAGC	
	AAAATGGTCCCATCCTCAATTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAGCCCGCCGCAAAGGCCGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGCTGGATGCTTCACAGGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAACGAAGCTACGAGCAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACAGACGGAGAGCGCCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
66	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACUUCCCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGCGAUUUCGUUGCCGAGCCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACTTCC	GCGAGAAACCAGUCGGAUCCUUAGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACATCGCGATTTCGTTGC	AAUCGGUGACGUGUUGUCCAAGCGGCUG	DAFLWSHPQFEKWSHPQ
	CGAGCCCATGGGCGAGAAACCAG	GAAGAGAGAGGAUUUGAUAAAGCUUACG	FEKWSHPQFEKPSSDDE
	TCGGATCCTTAGCCGGAATCGGT	UGGUCCUGGGCCAGUUUCUUGUUCUCAA	ATADSQHAAPPKKKRKV
	GACGTGTTGTCCAAGCGGCTGGA	GAAGGAUGAGGAUCGAUUUCGAGAAUGG	EDPKDFPSELLS

	AGAGAGAGGATTTGATAAAGCTT	UUGAAAGAUACCUGUGGCGCAAACGCAA	
	ACGTGGTCCTGGGCCAGTTTCTT	AGCAGAGCCGAGAUUGUUUUGGGUGCCU	
	GTTCTCAAGAAGGATGAGGATCG	GAGGGAGUGGUGUGACGCCUUCUUAUGG	
	ATTTCGAGAATGGTTGAAAGATA	AGUCAUCCUCAAUUUGAAAAGUGGAGCC	
	CCTGTGGCGCAAACGCAAAGCAG	ACCCACAGUUUGAAAAGUGGAGUCAUCC	
	AGCCGAGATTGTTTTGGGTGCCT	CCAAUUCGAAAAGCCAAGUUCAGAUGAU	
	GAGGGAGTGGTGTGACGCCTTCT	GAGGCCACAGCAGAUUCUCAGCACGCAG	
	TATGGAGTCATCCTCAATTTGAA	CCCCACCCAAAAAGAAAAGGAAGGUAGA	
	AAGTGGAGCCACCCACAGTTTGA	AGAUCCCAAGGAUUUUCCUUCUGAACUC	
	AAAGTGGAGTCATCCCCAATTCG	CUGAGUUAAAGACCAGACAGCUGGAGCC	
	AAAAGCCAAGTTCAGATGATGAG	UCGGUGGCCAUGCUUCUUGCCCCUUGGG	
	GCCACAGCAGATTCTCAGCACGC	CCUCCCCCCAGCCCCUCCUCCCCUUCCU	
	AGCCCCACCCAAAAAGAAAAGGA	GCACCCGUACCCCCGUGGUCUUUGAAUA	
	AGGTAGAAGATCCCAAGGATTTT	AAGUCUGAGUGGGCGGCAAGCUUUUCUU	
	CCTTCTGAACTCCTGAGTTAAAG	CUCUCCCUUGCACCUAAAAAAAAAAAAAA	
	ACCAGACAGCTGGAGCCTCGGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCATGCTTCTTGCCCCTTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCCCAGCCCCTCCTCCCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCTGCACCCGTACCCCCGTGGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTTGAATAAAGTCTGAGTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGCAAGCTTTTCTTCTCTCCCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCACCT	AAAAAAAAAAAAAAAAAA	
67	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACUAGCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGACUUCGUCGCCGAGCCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACTAGC	GGGAGAAACCCGUUGGCAGCCUGGCCGG	ANAKQSRDCFGCLREWC
	CAAAAGCACAGAGACTTCGTCGC	UAUUGGAGACGUGCUGUCAAAGCGACUG	DAFLWSHPQFEKWSHPQ
	CGAGCCCATGGGGGAGAAACCCG	GAAGAGAGGGGGUUUGACAAGGCUUAUG	FEKWSHPQFEKPKKKRK
	TTGGCAGCCTGGCCGGTATTGGA	UAGUGCUCGGGCAAUUCCUCGUUCUUAA	VEDPASQGTKRSYEQME
	GACGTGCTGTCAAAGCGACTGGA	AAAAGAUGAGGAUAGAUUUAGAGAAUGG	TDGERQ
	AGAGAGGGGGTTTGACAAGGCTT	CUGAAAGACACUUGCGGCGCCAACGCUA	
	ATGTAGTGCTCGGGCAATTCCTC	AGCAAAGCAGGGACUGUUUUGGUUGCCU	
	GTTCTTAAAAAAGATGAGGATAG	UAGAGAAUGGUGCGAUGCAUUCUUGUGG	
	ATTTAGAGAATGGCTGAAAGACA	UCCCACCCGCAGUUUGAGAAGUGGUCCC	
	CTTGCGGCGCCAACGCTAAGCAA	ACCCACAAUUUGAAAAGUGGUCACAUCC	
	AGCAGGGACTGTTTTGGTTGCCT	CCAGUUUGAAAAACCCAAGAAGAAGAGG	
	TAGAGAATGGTGCGATGCATTCT	AAGGUGGAGGACCCAGCAAGUCAGGGCA	
	TGTGGTCCCACCCGCAGTTTGAG	CAAAAAGAUCAUACGAGCAGAUGGAAAC	
	AAGTGGTCCCACCCACAATTTGA	AGAUGGCGAAAGACAGUAAAGACCAGAC	
	AAAGTGGTCACATCCCCAGTTTG	AGCUGGAGCCUCGGUGGCCAUGCUUCUU	
	AAAAACCCAAGAAGAAGAGGAAG	GCCCCUUGGGCCUCCCCCCAGCCCCUCC	
	GTGGAGGACCCAGCAAGTCAGGG	UCCCCUUCCUGCACCCGUACCCCCGUGG	
	CACAAAAAGATCATACGAGCAGA	UCUUUGAAUAAAGUCUGAGUGGGCGGCA	
	TGGAAACAGATGGCGAAAGACAG	AGCUUUUCUUCUCUCCCUUGCACCUAAA	
	TAAAGACCAGACAGCTGGAGCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGGTGGCCATGCTTCTTGCCCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGGCCTCCCCCCAGCCCCTCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCTTCCTGCACCCGTACCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGTCTTTGAATAAAGTCTGAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGGCGGCAAGCTTTTCTTCTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
		A	

68	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACCUCACAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGACUUUGUCGCAGAACCAAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACCTCA	GAGAAAAGCCCGUCGGCAGUCUCGCAGG	ANAKQSRDCFGCLREWC
	CAGAAACACCGAGACTTTGTCGC	AAUUGGAGAUGUCUUAAGCAAGAGGCUU	DAFLWSHPQFEKWSHPQ
	AGAACCAATGGGAGAAAAGCCCG	GAAGAGAGAGGUUUUGAUAAAGCAUAUG	FEKWSHPQFEKPAAKRV
	TCGGCAGTCTCGCAGGAATTGGA	UUGUGCUGGGCCAGUUUCUUGUCCUCAA	KLDASQGTKRSYEQMET
	GATGTCTTAAGCAAGAGGCTTGA	AAAGGACGAGGAUCGAUUCAGAGAAUGG	DGERQPAAKRVKLD
	AGAGAGAGGTTTTGATAAAGCAT	CUGAAAGACACCUGCGGUGCAAACGCAA	
	ATGTTGTGCTGGGCCAGTTTCTT	AACAAAGCAGGGACUGUUUUGGAUGCCU	
	GTCCTCAAAAAGGACGAGGATCG	CCGAGAAUGGUGCGAUGCCUUUCUGUGG	
	ATTCAGAGAATGGCTGAAAGACA	AGCCACCCACAAUUCGAAAAGUGGUCAC	
	CCTGCGGTGCAAACGCAAAACAA	ACCCUCAAUUUGAAAAAUGGUCCCACCC	
	AGCAGGGACTGTTTTGGATGCCT	ACAAUUCGAGAAACCUGCCGCUAAGAGG	
	CCGAGAATGGTGCGATGCCTTTC	GUGAAGCUCGACGCCUCUCAGGGUACCA	
	TGTGGAGCCACCCACAATTCGAA	AGAGAUCCUACGAGCAGAUGGAAACCGA	
	AAGTGGTCACACCCTCAATTTGA	UGGAGAGAGGCAGCCCGCUGCGAAAAGA	
	AAAATGGTCCCACCCACAATTCG	GUGAAACUUGAUUAAAGACCAGACAGCU	
	AGAAACCTGCCGCTAAGAGGGTG	GGAGCCUCGGUGGCCAUGCUUCUUGCCC	
	AAGCTCGACGCCTCTCAGGGTAC	CUUGGGCCUCCCCCCCAGCCCCUCCUCCC	
	CAAGAGATCCTACGAGCAGATGG	CUUCCUGCACCCGUACCCCCGUGGUCUU	
	AAACCGATGGAGAGAGGCAGCCC	UGAAUAAAGUCUGAGUGGGCGGCAAGCU	
	GCTGCGAAAAGAGTGAAACTTGA	UUUCUUCUCUCCCUUGCACCUAAAAAAA	
	TTAAAGACCAGACAGCTGGAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCGGTGGCCATGCTTCTTGCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGGGCCTCCCCCCAGCCCCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCCCTTCCTGCACCCGTACCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGTGGTCTTTGAATAAAGTCTGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGGCGGCAAGCTTTTCTTCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
69	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACCACAUCACAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGAUUUCGUGGCAGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACCACATCA	GGGAAAAACCUGUGGGAAGCUUAGCUGG	ANAKQSRDCFGCLREWC
	CAAAAACACAGGGATTTCGTGGC	GAUCGGCGACGUGCUGUCAAAAAGACUG	DAFLWSHPQFEKWSHPQ
	AGAACCCATGGGGGAAAAACCTG	GAGGAGCGAGGAUUUGACAAGGCCUAUG	FEKWSHPQFEKPAAKRV
	TGGGAAGCTTAGCTGGGATCGGC	UAGUCCUCGGCCAAUUUUUGGUUCUUAA	KLDATQGTKRSYEQMET
	GACGTGCTGTCAAAAAGACTGGA	AAAGGACGAGGAUCGUUUCCGGGAAUGG	DGERQ
	GGAGCGAGGATTTGACAAGGCCT	UUAAAAGACACAUGUGGUGCCAACGCUA	
	ATGTAGTCCTCGGCCAATTTTTG	AGCAGAGCCGGGACUGCUUUGGUUGUCU	
	GTTCTTAAAAAGGACGAGGATCG	GCGCGAGUGGUGCGAUGCCUUCUUGUGG	
	TTTCCGGGAATGGTTAAAAGACA	UCACAUCCUCAGUUUGAGAAAUGGUCAC	
	CATGTGGTGCCAACGCTAAGCAG	ACCCGCAAUUUGAAAAGUGGUCUCACCC	
	AGCCGGGACTGCTTTGGTTGTCT	UCAGUUCGAAAAACCUGCCGCCAAAAGG	
	GCGCGAGTGGTGCGATGCCTTCT	GUCAAGUUAGAUGCAACUCAAGGGACUA	
	TGTGGTCACATCCTCAGTTTGAG	AGCGCUCCUACGAACAAAUGGAAACUGA	
	AAATGGTCACACCCGCAATTTGA	CGGUGAACGACAGUAAAGACCAGACAGC	
	AAAGTGGTCTCACCCTCAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCTGCCGCCAAAAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAGTTAGATGCAACTCAAGGGAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAGCGCTCCTACGAACAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACTGACGGTGAACGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
70	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACATCC CAAAAGCATAGAGACTTTGTAGC CGAGCCTATGGGAGAAAAGCCTG TGGGATCCCTGGCCCGCATTGGA GACGTGCTCAGCAAGCGGTTGGA GGAGAGAGGTTTTGACAAAGCTT ACGTAGTTTTGGGACAGTTTCTT GTACTGAAGAAGGATGAAGACCG GTTTAGAGAATGGCTGAAGGACA CATGTGGCGCAAATGCCAAACAA AGTAGGGATTGTTTCGGGTGCTT GCGTGAATGGTGCGATGCATTTC TTTGGAGTCACCCCCAGTTCGAG AAGTGGAGCCATCCTCAGTTTGA GAAATGGTCTCACCCACAATTCG AAAAACCTGCCGCTAAGAGGGTG AAGCTCGACGCAACAAAGGGCAC AAAAAGATCCTACGAGCAAATGG AAACTGACGGTGAAAGACAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACAUCCCAAAAG CAUAGAGACUUUGUAGCCGAGCCUAUGG GAGAAAAGCCUGUGGGAUCCCUGGCCGG CAUUGGAGACGUGCUCAGCAAGCGGUUG GAGGAGAGAGGUUUUGACAAAGCUUACG UAGUUUUGGGACAGUUUCUUGUACUGAA GAAGGAUGAAGACCGGUUUAGAGAAUGG CUGAAGGACACAUGUGGCGCAAAUGCCA AACAAAGUAGGGAUUGUUUCGGGUGCUU GCGUGAAUGGUGCGAUGCAUUUCUUUGG AGUCACCCCCAGUUCGAGAAGUGGAGCC AUCCUCAGUUUGAGAAAUGGUCUCACCC ACAAUUCGAAAAACCUGCCGCUAAGAGG GUGAAGCUCGACGCAACAAAGGGCACAA AAAGAUCCUACGAGCAAAUGGAAACUGA CGGUGAAAGACAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDATKGTKRSYEQMET DGERQ
71	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACGACTTCA CAGAAACATCGAGATTTTGTAGC CGAGCCAATGGGCGAAAAGCCAG TCGGGTCACTTGCTGGGATCGGT GATGTGCTTAGTAAGCGATTGGA GGAACGCGGATTTGATAAAGCCT ACGTAGTGCTGGGCCAGTTTCTG GTGTTAAAGAAAGACGAGGACCG GTTTAGAGAGTGGTTGAAGGATA CATGTGGAGCTAACGCTAAACAA AGCAGAGACTGTTTTGGATGTCT GAGAGAGTGGTGCGATGCTTTTC TCTGGAGCCATCCGCAATTCGAG AAATGGTCCCATCCCCAGTTTGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACGACUUCACAGAAA CAUCGAGAUUUUGUAGCCGAGCCAAUGG GCGAAAAGCCAGUCGGGUCACUUGCUGG GAUCGGUGAUGUGCUUAGUAAGCGAUUG GAGGAACGCGGAUUUGAUAAAGCCUACG UAGUGCUGGGCCAGUUUCUGGUGUUAAA GAAAGACGAGGACCGGUUUAGAGAGUGG UUGAAGGAUACAUGUGGAGCUAACGCUA AACAAAGCAGAGACUGUUUUGGAUGUCU GAGAGAGUGGUGCGAUGCUUUUCUCUGG AGCCAUCCGCAAUUCGAGAAAUGGUCCC AUCCCCAGUUUGAAAAAUGGAGUCACCC UCAGUUUGAGAAACCCGACACCAGAAAG CCACCUGCCGCUAAGCGUGUGAAACUGG ACGCAAGCCAGGGGACAAAAAGAAGUUA UGAACAGAUGGAAACCGAUGGUGAAAGG	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPDTRKP PAAKRVKLDASQGTKRS YEQMETDGERQ

	AAAATGGAGTCACCCTCAGTTTG AGAAACCCGACACCAGAAAGCCA CCTGCCGCTAAGCGTGTGAAACT GGACGCAAGCCAGGGGACAAAAA GAAGTTATGAACAGATGGAAACC GATGGTGAAAGGCAGTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC AAGCTTTTCTTCTCTCCCTTGCA CCT	CAGUAAAGACCAGACAGCUGGAGCCUCG GUGGCCAUGCUUCUUGCCCCUUGGGCCU CCCCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGCAAGCUUUUCUUCUC UCCCUUGCACCUAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
72	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACGACGAGC CAGAAGCATCGGGATTTTGTCGC CGAACCTATGGGAGAGAAGCCAG TCGGGAGTCTCGCTGGCATCGGA GATGTACTGTCTAAACGGCTGGA AGAACGTGGATTTGACAAGGCAT ATGTTGTCTTGGGGCAGTTTCTG GTACTTAAGAAGGATGAAGATAG GTTTAGAGAATGGTTAAAAGATA CATGTGGGGCCAATGCTAAACAG AGTCGGGATTGTTTTGGCTGTTT GAGAGAGTGGTGTGATGCATTTC TGTGGTCCCACCCTCAGTTCGAA AAATGGAGCCACCCGCAATTTGA AAAGTGGTCTCATCCTCAGTTTG AGAAGCCAGACACTCGGAAACCT CCCGCAGCCAAGCGCGTGAAGCT CGACGCTTCCCAAGGCACCAAGA GAAGCTATGAGCAGATGGAAACT GACGGGGAAAGGCAGGATGACGA TGATACCGAAGAGTAAAGACCAG ACAGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCCTTGGGCCTCCC CCCAGCCCCTCCTCCCCTTCCTG CACCCGTACCCCCGTGGTCTTTG AATAAAGTCTGAGTGGGCGGCAA GCTTTTCTTCTCTCCCTTGCACC T	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACGACGAGCCAGAAG CAUCGGGAUUUUGUCGCCGAACCUAUGG GAGAGAAGCCAGUCGGGAGUCUCGCUGG CAUCGGAGAUGUACUGUCUAAACGGCUG GAAGAACGUGGAUUUGACAAGGCAUAUG UUGUCUUGGGGCAGUUUCUGGUACUUAA GAAGGAUGAAGAUAGGUUUAGAGAAUGG UUAAAAGAUACAUGUGGGGCCAAUGCUA AACAGAGUCGGGAUUGUUUUGGCUGUUU GAGAGAGUGGUGUGAUGCAUUUCUGUGG UCCCACCCUCAGUUCGAAAAAUGGAGCC ACCCGCAAUUUGAAAAGUGGUCUCAUCC UCAGUUUGAGAAGCCAGACACUCGGAAA CCUCCCGCAGCCAAGCGCGUGAAGCUCG ACGCUUCCCAAGGCACCAAGAGAAGCUA UGAGCAGAUGGAAACUGACGGGGAAAGG CAGGAUGACGAUGAUACCGAAGAGUAAA GACCAGACAGCUGGAGCCUCGGUGGCCA UGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCAAGCUUUUCUUCUCUCCCUUG CACCUAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPDTRKP PAAKRVKLDASQGTKRS YEQMETDGERQDDDDTE E
73	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACAACTAGT CAAAAACACAGAGATTTGTTGC CGAACCCATGGGTGAGAAACCTG TTGGGAGTTTAGCAGGCATAGGC GACGTTCTCTCAAAACGATTAGA GGAGAGGGGGTTCGATAAAGCAT ATGTCGTACTGGGGCAGTTCCTG GTGTTGAAAAAAGATGAAGACCG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACUAGUCAAAAA CACAGAGAUUUUGUUGCCGAACCCAUGG GUGAGAAACCUGUUGGGAGUUUAGCAGG CAUAGGCGACGUUCUCUCAAAACGAUUA GAGGAGAGGGGGUUCGAUAAAGCAUAUG UCGUACUGGGGCAGUUCCUGGUGUUGAA AAAAGAUGAAGACCGGUUUAGGGAAUGG CUCAAAGACACGUGCGGAGCAAAUGCGA AACAGAGCCGAGAUUGUUUUGGCUGUCU GAGGGAGUGGUGUGACGCUUUCCUGUGG	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRAFEQMET DGERQ

	GTTTAGGGAATGGCTCAAAGACA	AGCCACCCCCAAUUUGAGAAGUGGUCUC	
	CGTGCGGAGCAAATGCGAAACAG	AUCCACAGUUUGAGAAAUGGUCCCAUCC	
	AGCCGAGATTGTTTTGGCTGTCT	UCAAUUUGAGAAACCCGCCGCAAAGAGA	
	GAGGGAGTGGTGTGACGCTTTCC	GUGAAACUGGAUGCCUCCCAGGGAACUA	
	TGTGGAGCCACCCCCAATTTGAG	AAAGGGCGUUUGAGCAGAUGGAAACAGA	
	AAGTGGTCTCATCCACAGTTTGA	CGGAGAGCGGCAGUAAAGACCAGACAGC	
	GAAATGGTCCCATCCTCAATTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCCGCCGCAAAGAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGATGCCTCCCAGGGAAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAAAGGGCGTTTGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGACGGAGAGCGGCAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
74	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACUUCACAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGAGACUUCGUGGCAGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACTTCA	GAGAAAAACCUGUCGGCAGUCUGGCUGG	ANAKQSRDCFGCLREWC
	CAGAAGCACCGAGACTTCGTGGC	CAUCGGUGAUGUGUUAAGUAAGAGGUUG	DAFLWSHPQFEKWSHPQ
	AGAACCTATGGGAGAAAAACCTG	GAAGAAAGGGGCUUCGAUAAGGCGUAUG	FEKWSHPQFEKPDDDDE
	TCGGCAGTCTGGCTGGCATCGGT	UCGUAUUGGGACAGUUCCUCGUACUGAA	ATADSQHAAPPKKKRKV
	GATGTGTTAAGTAAGAGGTTGGA	GAAGGAUGAGGAUAGGUUUAGAGAAUGG	EDPKDFPSELLS
	AGAAAGGGGCTTCGATAAGGCGT	CUCAAAGAUACAUGCGGUGCGAACGCUA	
	ATGTCGTATTGGGACAGTTCCTC	AGCAAAGCAGGGAUUGUUUUGGCUGUCU	
	GTACTGAAGAAGGATGAGGATAG	CCGGGAGUGGUGCGACGCCUUCUUGUGG	
	GTTTAGAGAATGGCTCAAAGATA	AGCCAUCCCCAGUUCGAAAAAUGGUCAC	
	CATGCGGTGCGAACGCTAAGCAA	AUCCCCAGUUUGAAAAAUGGAGUCACCC	
	AGCAGGGATTGTTTTGGCTGTCT	UCAGUUUGAGAAGCCUGACGACGAUGAC	
	CCGGGAGTGGTGCGACGCCTTCT	GAAGCAACUGCAGAUUCUCAGCACGCUG	
	TGTGGAGCCATCCCCAGTTCGAA	CUCCACCGAAAAAAAAAGCGCAAAGUUGA	
	AAATGGTCACATCCCCAGTTTGA	GGACCCGAAAGAUUUUCCUAGCGAACUC	
	AAAATGGAGTCACCCTCAGTTTG	CUGAGCUAAAGACCAGACAGCUGGAGCC	
	AGAAGCCTGACGACGATGACGAA	UCGGUGGCCAUGCUUCUUGCCCCUUGGG	
	GCAACTGCAGATTCTCAGCACGC	CCUCCCCCCCAGCCCCUCCUCCCCUUCCU	
	TGCTCCACCGAAAAAAAAGCGCA	GCACCCGUACCCCCGUGGUCUUUGAAUA	
	AAGTTGAGGACCCGAAAGATTTT	AAGUCUGAGUGGGCGGCAAGCUUUUCUU	
	CCTAGCGAACTCCTGAGCTAAAG	CUCUCCCUUGCACCUAAAAAAAAAAAAA	
	ACCAGACAGCTGGAGCCTCGGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCATGCTTCTTGCCCCTTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCCCAGCCCCTCCTCCCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCTGCACCCGTACCCCCGTGGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTTGAATAAAGTCTGAGTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGCAAGCTTTTCTTCTCTCCCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCACCT	AAAAAAAAAAAAAAAAAA	

75	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACCUCCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGAGACUUUGUGGCCGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACCTCC	GGGAGAAGCCUGUCGGCAGUUUAGCAGG	ANAKQSRDCFGCLREWC
	CAAAAGCATAGAGACTTTGTGGC	GAUUGGGGAUGUACUCAGCAAGCGUCUG	DAFLWSHPQFEKWSHPQ
	CGAGCCTATGGGGGAGAAGCCTG	GAAGAGAGGGGCUUUGAUAAGGCAUACG	FEKWSHPQFEKPAAKRV
	TCGGCAGTTTAGCAGGGATTGGG	UGGUUUUAGGCCAGUUUCUGGUCCUCAA	KLDSQGTKRSYEQMTKR
	GATGTACTCAGCAAGCGTCTGGA	AAAAGAUGAGGAUAGAUUUCGAGAGUGG	SYEQMSQGTKRSYEQMT
	AGAGAGGGGCTTTGATAAGGCAT	CUCAAGGACACAUGCGGGGCCAACGCCA	KRSYEQM
	ACGTGGTTTTAGGCCAGTTTCTG	AGCAGAGCAGAGAUUGUUUUGGCUGCUU	
	GTCCTCAAAAAAGATGAGGATAG	GAGAGAAUGGUGCGAUGCUUUUUUAUGG	
	ATTTCGAGAGTGGCTCAAGGACA	UCUCAUCCGCAGUUCGAGAAGUGGUCCC	
	CATGCGGGGCCAACGCCAAGCAG	AUCCUCAGUUCGAGAAGUGGUCUCAUCC	
	AGCAGAGATTGTTTTGGCTGCTT	ACAAUUCGAAAAGCCUGCCGCUAAACGA	
	GAGAGAATGGTGCGATGCTTTTT	GUUAAACUGGACAGCCAAGGAACAAAGC	
	TATGGTCTCATCCGCAGTTCGAG	GUUCUUAUGAGCAGAUGACCAAACGCAG	
	AAGTGGTCCCATCCTCAGTTCGA	UUACGAGCAAAUGAGCCAAGGUACAAAA	
	GAAGTGGTCTCATCCACAATTCG	AGAAGCUAUGAACAGAUGACAAAAAGAU	
	AAAAGCCTGCCGCTAAACGAGTT	CAUACGAACAAAUGUAAAGACCAGACAG	
	AAACTGGACAGCCAAGGAACAAA	CUGGAGCCUCGGUGGCCAUGCUUCUUGC	
	GCGTTCTTATGAGCAGATGACCA	CCCUUGGGCCUCCCCCCAGCCCCUCCUC	
	AACGCAGTTACGAGCAAATGAGC	CCCUUCCUGCACCCGUACCCCCGUGGUC	
	CAAGGTACAAAAAGAAGCTATGA	UUUGAAUAAAGUCUGAGUGGGCGGCAAG	
	ACAGATGACAAAAAGATCATACG	CUUUUCUUCUCUCCCUUGCACCUAAAAA	
	AACAAATGTAAAGACCAGACAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGAGCCTCGGTGGCCATGCTTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCCCCTTGGGCCTCCCCCCAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCTCCTCCCCTTCCTGCACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTACCCCCGTGGTCTTTGAATAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTCTGAGTGGGCGGCAAGCTTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTTCTCTCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAA	
76	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAAGCCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUCGGGAUUUUGUGGCCGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACAAGC	GAGAAAAACCAGUCGGAUCCCUCGCAGG	ANAKQSRDCFGCLREWC
	CAGAAACATCGGGATTTTGTGGC	GAUCGGCGAUGUGCUUAGCAAGCGACUG	DAFLWSHPQFEKWSHPQ
	CGAGCCTATGGGAGAAAAACCAG	GAAGAACGCGGAUUUGACAAAGCCUACG	FEKWSHPQFEKAARLHR
	TCGGATCCCTCGCAGGGATCGGC	UUGUACUCGGGCAGUUCUUGGUAUUAAA	FKNKGKDSTEMRRRRIE
	GATGTGCTTAGCAAGCGACTGGA	GAAAGACGAGGACCGAUUCCGGGAAUGG	VNVELRKAKKDDQMLKR
	AGAACGCGGATTTGACAAAGCCT	UUGAAAGAUACUUGCGGCGCCAAUGCUA	RNVS
	ACGTTGTACTCGGGCAGTTCTTG	AACAAAGCAGGGACUGUUUCGGCUGUCU	
	GTATTAAAGAAAGACGAGGACCG	UCGGGAAUGGUGUGAUGCAUUCCUCUGG	
	ATTCCGGGAATGGTTGAAAGATA	AGUCAUCCGCAGUUUGAGAAGUGGUCAC	
	CTTGCGGCGCCAATGCTAAACAA	ACCCACAAUUCGAGAAAUGGUCUCACCC	
	AGCAGGGACTGTTTCGGCTGTCT	ACAAUUCGAAAAAGCCGCCAGACUGCAU	
	TCGGGAATGGTGTGATGCATTCC	CGGUUCAAAAACAAGGGUAAAGACUCAA	
	TCTGGAGTCATCCGCAGTTTGAG	CAGAGAUGAGGAGGAGACGAAUUGAAGU	
	AAGTGGTCACACCCACAATTCGA	UAACGUGGAACUGAGAAAGGCUAAAAAG	
	GAAATGGTCTCACCCACAATTCG	GAUGAUCAGAUGUUAAAGCGAAGGAACG	
	AAAAAGCCGCCAGACTGCATCGG	UCUCAUAAAGACCAGACAGCUGGAGCCU	
	TTCAAAAACAAGGGTAAAGACTC	CGGUGGCCAUGCUUCUUGCCCCUUGGGC	
	AACAGAGATGAGGAGGAGACGAA	CUCCCCCCAGCCCCUCCUCCCCUUCCUG	
	TTGAAGTTAACGTGGAACTGAGA	CACCCGUACCCCCGUGGUCUUUGAAUAA	
	AAGGCTAAAAAGGATGATCAGAT	AGUCUGAGUGGGCGGCAAGCUUUUCUUC	
	GTTAAAGCGAAGGAACGTCTCAT	UCUCCCUUGCACCUAAAAAAAAAAAAAA	

	AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCAAGCTTTTCTTCTCTC CCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	
77	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACAAGT CAAAAACACAGGGACTTCGTAGC CGAGCCTATGGGAGAAAAGCCTG TGGGCTCCTTAGCTGGAATAGGT GACGTTCTGTCAAAACGACTGGA GGAACGTGGATTTGATAAGGCTT ATGTTGTACTCGGGCAATTTTTG GTCTTGAAGAAAGACGAAGATAG ATTCAGAGAGTGGCTTAAAGATA CTTGTGGTGCAAACGCCAAGCAG TCCCGCGATTGTTTCGGATGTTT ACGAGAGTGGTGTGATGCTTTTC TGTGGTCACATCCTCAGTTTGAG AAATGGAGCCATCCCCAGTTCGA GAAATGGTCTCACCCTCAGTTCG AAAAAGCTGCTGCAAAGAGGGTG AAACTGGACGCTAGTCAGGGAAC GAAACGTAGCTACGAACAAATGG AAACCGACGGAGAGCGCCAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACAAGUCAAAAA CACAGGGACUUCGUAGCCGAGCCUAUGG GAGAAAAGCCUGUGGGCUCCUUAGCUGG AAUAGGUGACGUUCUGUCAAAACGACUG GAGGAACGUGGAUUUGAUAAGGCUUAUG UUGUACUCGGGCAAUUUUUGGUCUUGAA GAAAGACGAAGAUAGAUUCAGAGAGUGG CUUAAAGAUACUUGUGGUGCAAACGCCA AGCAGUCCCGCGAUUGUUUCGGAUGUUU ACGAGAGUGGUGUGAUGCUUUUCUGUGG UCACAUCCUCAGUUUGAGAAAUGGAGCC AUCCCCAGUUCGAGAAAUGGUCUCACCC UCAGUUCGAAAAAGCUGCUGCAAAGAGG GUGAAACUGGACGCUAGUCAGGGAACGA AACGUAGCUACGAACAAAUGGAAACCGA CGGAGAGCGCCAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKAAAKRV KLDASQGTKRSYEQMET DGERQ
78	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACCAGC CAGAAGCATAGGGATTTCGTCGC AGAGCCCATGGGAGAGAAACCAG TGGGATCACTCGCAGGAATCGGG GATGTGCTGTCCAAGCGACTTGA AGAGCGAGGATTCGACAAAGCCT ATGTTGTGCTTGGCCAATTTCTC GTGTTAAAGAAGGATGAGGATCG GTTCAGGGAATGGTTAAAGGATA CCTGCGGCGCCAATGCCAAACAA TCTAGAGACTGTTTCGGATGTCT TAGGGAATGGTGTGATGCTTTTT TGTGGTCCCACCCACAATTTGAA AAATGGAGTCACCCACAGTTCGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACCAGCCAGAAG CAUAGGGAUUUCGUCGCAGAGCCCAUGG GAGAGAAACCAGUGGGAUCACUCGCAGG AAUCGGGGAUGUGCUGUCCAAGCGACUU GAAGAGCGAGGAUUCGACAAAGCCUAUG UUGUGCUUGGCCAAUUUCUCGUGUUAAA GAAGGAUGAGGAUCGGUUCAGGGAAUGG UUAAAGGAUACCUGCGGCGCCAAUGCCA AACAAUCUAGAGACUGUUUCGGAUGUCU UAGGGAAUGGUGUGAUGCUUUUUUGUGG UCCCACCCACAAUUUGAAAAAUGGAGUC ACCCACAGUUCGAAAAGUGGAGUCAUCC ACAGUUUGAAAAGAAAGAUGAACAAUAU GAGCAGAUGGUGAAGGAGCUGCAAGAGA UUAGAUUAGAAUAAAGACCAGACAGCUG GAGCCUCGGUGGCCAUGCUUCUUGCCCC	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKKDEEYE QMVKELQEIRLE

	AAAGTGGAGTCATCCACAGTTTG	UUGGGCCUCCCCCCAGCCCCUCCUCCCC	
	AAAAGAAAGATGAAGAATATGAG	UUCCUGCACCCGUACCCCCGUGGUCUUU	
	CAGATGGTGAAGGAGCTGCAAGA	GAAUAAAGUCUGAGUGGGCGGCAAGCUU	
	GATTAGATTAGAATAAAGACCAG	UUCUUCUCUCCCUUGCACCUAAAAAAAAA	
	ACAGCTGGAGCCTCGGTGGCCAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTCTTGCCCCTTGGGCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCAGCCCCTCCTCCCCTTCCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CACCCGTACCCCCGTGGTCTTTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATAAAGTCTGAGTGGGCGGCAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTTTCTTCTCTCCCTTGCACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	T	AAAAAAAAAAAAAAAAAAAAAAAA	
79	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACAAGCCAAAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGAUUUCGUCGCAGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACAAGC	GAGAGAAGCCAGUAGGGAGUCUCGCAGG	ANAKQSRDCFGCLREWC
	CAAAAGCACAGAGATTCGTCGC	UAUUGGAGACGUUUUAAGCAAACGGCUG	DAFLWSHPQFEKWSHPQ
	AGAACCTATGGGAGAGAAGCCAG	GAGGAGCGAGGUUUUGACAAAGCAUACG	FEKWSHPQFEKPWFPQE
	TAGGGAGTCTCGCAGGTATTGGA	UGGUUCUUGGCCAGUUUUUGGUAUUAAA	GTIDIKRWRRVG
	GACGTTTTAAGCAAACGGCTGGA	GAAAGAUGAAGACCGUUUUCGGGAGUGG	
	GGAGCGAGGTTTTGACAAAGCAT	CUGAAAGAUACUUGUGGAGCUAAUGCUA	
	ACGTGGTTCTTGGCCAGTTTTTG	AACAAAGCAGGGAUUGUUUCGGAUGUCU	
	GTATTAAAGAAAGATGAAGACCG	GAGGGAAUGGUGCGAUGCCUUCCUGUGG	
	TTTTCGGGAGTGGCTGAAAGATA	UCCCAUCCCCAAUUUGAAAAAUGGUCUC	
	CTTGTGGAGCTAATGCTAAACAA	ACCCUCAGUUUGAAAAAUGGAGCCAUCC	
	AGCAGGGATTGTTTCGGATGTCT	CCAAUUUGAAAAGCCCUGGUUUCCUCAG	
	GAGGGAATGGTGCGATGCCTTCC	GAGGGGACCAUCGAUAUUAAACGCUGGA	
	TGTGGTCCCATCCCCAATTTGAA	GGCGGGUGGGGUAAAGACCAGACAGCUG	
	AAATGGTCTCACCCTCAGTTTGA	GAGCCUCGGUGGCCAUGCUUCUUGCCCC	
	AAAATGGAGCCATCCCCAATTTG	UUGGGCCUCCCCCCAGCCCCUCCUCCCC	
	AAAAGCCCTGGTTTCCTCAGGAG	UUCCUGCACCCGUACCCCCGUGGUCUUU	
	GGGACCATCGATATTAAACGCTG	GAAUAAAGUCUGAGUGGGCGGCAAGCUU	
	GAGGCGGGTGGGGTAAAGACCAG	UUCUUCUCUCCCUUGCACCUAAAAAAAAA	
	ACAGCTGGAGCCTCGGTGGCCAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTCTTGCCCCTTGGGCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCAGCCCCTCCTCCCCTTCCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CACCCGTACCCCCGTGGTCTTTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATAAAGTCTGAGTGGGCGGCAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTTTCTTCTCTCCCTTGCACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	T	AAAAAAAAAAAAAAAAAAAAAAAA	
80	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAACAAGCCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGAGACUUUGUGGCCGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACAACAAGC	GAGAAAAACCAGUGGGGAGCCUUGCGGG	ANAKQSRDCFGCLREWC
	CAGAAACACAGAGACTTTGTGGC	AAUCGGUGACGUUCUGUCUAAACGGUUG	DAFLWSHPQFEKWSHPQ
	CGAGCCTATGGGAGAAAAACCAG	GAGGAGAGAGGGUUUGACAAAGCUUAUG	FEKWSHPQFEKPAAKRV
	TGGGGAGCCTTGCGGGAATCGGT	UUGUUCUGGGACAGUUUCUGGUAUUGAA	KLDFGNYNNQSSNFGPM
	GACGTTCTGTCTAAACGGTTGGA	GAAGGACGAGGAUCGCUUCAGAGAAUGG	KGGNFGGRSSGPY
	GGAGAGAGGGTTTGACAAAGCTT	CUGAAGGAUACAUGUGGGGCUAAUGCUA	
	ATGTTGTTCTGGGACAGTTTCTG	AACAGAGUCGUGACUGCUUUGGCUGCCU	
	GTATTGAAGAAGGACGAGGATCG	GAGGGAAUGGUGCGAUGCAUUCUUAUGG	
	CTTCAGAGAATGGCTGAAGGATA	UCUCAUCCUCAAUUCGAGAAAUGGAGCC	
	CATGTGGGGCTAATGCTAAACAG	AUCCGCAAUUUGAGAAGUGGAGCCAUCC	
	AGTCGTGACTGCTTTGGCTGCCT	GCAGUUCGAGAAACCCGCCGCCAAACGA	
	GAGGGAATGGTGCGATGCATTCT	GUGAAACUCGAUUUUGGCAAUUACAAUA	
	TATGGTCTCATCCTCAATTCGAG	AUCAAUCCAGUAAUUUUGGUCCCAUGAA	

	AAATGGAGCCATCCGCAATTTGA GAAGTGGAGCCATCCGCAGTTCG AGAAACCCGCCGCCAAACGAGTG AAACTCGATTTTGGCAATTACAA TAATCAATCCAGTAATTTTGGTC CCATGAAAGGAGGAAATTTTGGT GGACGGAGCTCCGGCCCCTACTA AAGACCAGACAGCTGGAGCCTCG GTGGCCATGCTTCTTGCCCCTTG GGCCTCCCCCCAGCCCCTCCTCC CCTTCCTGCACCCGTACCCCCGT GGTCTTTGAATAAAGTCTGAGTG GGCGGCAAGCTTTTCTTCTCTCC CTTGCACCT	AGGAGGAAAUUUUGGUGGACGGAGCUCC GGCCCCUACUAAAGACCAGACAGCUGGA GCCUCGGUGGCCAUGCUUCUUGCCCCUU GGGCCUCCCCCCAGCCCCUCCUCCCCUU CCUGCACCCGUACCCCCGUGGUCUUUGA AUAAAGUCUGAGUGGGCGGCAAGCUUUU CUUCUCUCCCUUGCACCUAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
81	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACAACCAGT CAAAAACACCGCGACTTTGTTGC CGAGCCCATGGGAGAGAAGCCCG TAGGATCCCTCGCTGGAATTGGA GACGTGTTGTCCAAGAGACTGGA GGAAAGGGGCTTCGACAAGGCAT ACGTGGTCCTTGGACAATTTTTG GTGCTGAAAAAAGACGAAGATCG GTTCAGAGAATGGTTGAAAGATA CCTGCGGTGCGAATGCCAAGCAG AGTAGGGATTGTTTTGGATGCCT TCGAGAGTGGTGCGATGCTTTCC TTTGGAGTCACCCACAATTCGAG AAATGGTCTCATCCACAGTTTGA GAAGTGGAGCCACCCACAGTTTG AAAAACCTGCCGCAAAGAGAGTT AAGCTGGATTAAAGACCAGACAG CTGGAGCCTCGGTGGCCATGCTT CTTGCCCCTTGGGCCTCCCCCCA GCCCCTCCTCCCCTTCCTGCACC CGTACCCCCGTGGTCTTTGAATA AAGTCTGAGTGGGCGGCAAGCTT TTCTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACCAGUCAAAAA CACCGCGACUUUGUUGCCGAGCCCAUGG GAGAGAAGCCCGUAGGAUCCCUCGCUGG AAUUGGAGACGUGUUGUCCAAGAGACUG GAGGAAAGGGGCUUCGACAAGGCAUACG UGGUCCUUGGACAAUUUUUGGUGCUGAA AAAAGACGAAGAUCGGUUCAGAGAAUGG UUGAAAGAUACCUGCGGUGCGAAUGCCA AGCAGAGUAGGGAUUGUUUUGGAUGCCU UCGAGAGUGGUGCGAUGCUUUCCUUUGG AGUCACCCACAAUUCGAGAAAUGGUCUC AUCCACAGUUUGAGAAGUGGAGCCACCC ACAGUUUGAAAAACCUGCCGCAAAGAGA GUUAAGCUGGAUUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCAAGCU UUUCUUCUCUCCCUUGCACCUAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLD
82	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACAACGTCC CAAAAGCACCGAGACTTTGTAGC TGAACCAATGGGCGAAAAACCTG TGGGCTCTCTTGCCGGCATAGGG GATGTGTTATCCAAGCGATTAGA GGAGAGAGGCTTTGATAAAGCAT ATGTTGTGCTTGGGCAGTTTCTT GTCCTGAAGAAAGATGAAGATAG GTTTCGTGAATGGCTGAAAGATA CATGCGGCGCGAACGCGAAGCAA AGTAGAGATTGCTTCGGCTGTCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACGUCCCAAAAG CACCGAGACUUUGUAGCUGAACCAAUGG GCGAAAAACCUGUGGGCUCUCUUGCCGG CAUAGGGGAUGUGUUAUCCAAGCGAUUA GAGGAGAGAGGCUUUGAUAAAGCAUAUG UUGUGCUUGGGCAGUUUCUUGUCCUGAA GAAAGAUGAAGAUAGGUUUCGUGAAUGG CUGAAAGAUACAUGCGGCGCGAACGCGA AGCAAAGUAGAGAUUGCUUCGGCUGUCU UCGGGAAUGGUGUGACGCAUUUCUGUGG AGCCACCCCCAGUUUGAGAAAUGGUCUC AUCCACAGUUCGAGAAAUGGUCCCACCC CCAAUUUGAAAAGCCAAAAAAGAAGAGG	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPKKKRK VEDP

	TCGGGAATGGTGTGACGCATTTC TGTGGAGCCACCCCCAGTTTGAG AAATGGTCTCATCCACAGTTCGA GAAATGGTCCCACCCCCAATTTG AAAAGCCAAAAAAGAAGAGGAAG GTTGAAGATCCCTAAAGACCAGA CAGCTGGAGCCTCGGTGGCCATG CTTCTTGCCCCTTGGGCCTCCCC CCAGCCCCTCCTCCCCTTCCTGC ACCCGTACCCCCGTGGTCTTTGA ATAAAGTCTGAGTGGGCGGCAAG CTTTTCTTCTCTCCCTTGCACCT	AAGGUUGAAGAUCCCUAAAGACCAGACA GCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUGGGCCUCCCCCCAGCCCCUCCU CCCCUUCCUGCACCCGUACCCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCAA GCUUUUCUUCUCUCCCUUGCACCUAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
83	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACCACCAGT CAGAAACATAGAGATTTCGTAGC TGAGCCGATGGGAGAAAAGCCAG TTGGATCACTGGCAGGGATTGGT GATGTGTTGTCAAAAAGACTGGA GGAGAGAGGATTCGACAAGGCCT ATGTGGTACTGGGCCAGTTCCTC GTCCTTAAGAAGGATGAAGATAG ATTCCGAGAGTGGCTCAAAGACA CCTGTGGCGCTAACGCTAAGCAA TCCCGCGACTGTTTTGGCTGTCT CAGAGAGTGGTGCGACGCTTTCT TGTGGAGTCACCCCCAATTTGAA AAATGGAGCCACCCCCAGTTCGA GAAGTGGAGTCACCCTCAGTTTG AAAAAGGCAAAATCAGCAAGCAT TGGACCGGTATCGCTAGCCAGGG AACAAAACGATCTTACGAGCAGA TGGAGACGGATGGAGAACGACAG TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGUCAGAAA CAUAGAGAUUUCGUAGCUGAGCCGAUGG GAGAAAAGCCAGUUGGAUCACUGGCAGG GAUUGGUGAUGUGUUGUCAAAAAGACUG GAGGAGAGAGGAUUCGACAAGGCCUAUG UGGUACUGGGCCAGUUCCUCGUCCUUAA GAAGGAUGAAGAUAGAUUCCGAGAGUGG CUCAAAGACACCUGUGGCGCUAACGCUA AGCAAUCCCGCGACUGUUUUGGCUGUCU CAGAGAGUGGUGCGACGCUUUCUUGUGG AGUCACCCCCAAUUUGAAAAAUGGAGCC ACCCCCAGUUCGAGAAGUGGAGUCACCC UCAGUUUGAAAAAGGCAAAAUCAGCAAG CAUUGGACCGGUAUCGCUAGCCAGGGAA CAAAACGAUCUUACGAGCAGAUGGAGAC GGAUGGAGAACGACAGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA A	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKGKISKH WTGIASQGTKRSYEQME TDGERQ
84	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACTTCC CAGAAACACCGAGACTTGTAGC CGAGCCCATGGGAGAGAAGCCAG TGGGGTCTTTGGCTGGGATTGGA GACGTACTCTCCAAAAGACTGGA AGAAAGAGGTTTTGATAAGGCTT ATGTTGTTCTTGGTCAATTTCTG GTGTTGAAGAAAGATGAAGACCG CTTCCGCGAATGGCTCAAAGATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACUUCCCAGAAA CACCGAGACUUUGUAGCCGAGCCCAUGG GAGAGAAGCCAGUGGGGUCUUUGGCUGG GAUUGGAGACGUACUCUCCAAAAGACUG GAAGAAAGAGGUUUUGAUAAGGCUUAUG UUGUUCUUGGUCAAUUUCUGGUGUUGAA GAAAGAUGAAGACCGCUUCCGCGAAUGG CUCAAAGAUACUUGUGGAGCUAAUGCCA AGCAGAGUCGGGAUUGCUUCGGAUGCCU GAGAGAAUGGUGUGACGCCUUCUUGUGG UCUCACCCACAAUUUGAGAAGUGGAGCC	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKKHHLGL EEPKKLRPPPAR

	CTTGTGGAGCTAATGCCAAGCAG	AUCCCCAAUUCGAAAAGUGGUCCCACCC	
	AGTCGGGATTGCTTCGGATGCCT	ACAAUUUGAGAAGAAGCACCAUCUUGGU	
	GAGAGAATGGTGTGACGCCTTCT	UUAGAAGAACCCAAGAAGCUUAGACCAC	
	TGTGGTCTCACCCACAATTTGAG	CCCCCGCAAGAUAAAGACCAGACAGCUG	
	AAGTGGAGCCATCCCCAATTCGA	GAGCCUCGGUGGCCAUGCUUCUUGCCCC	
	AAAGTGGTCCCACCCACAATTTG	UUGGGCCUCCCCCCAGCCCCUCCUCCCC	
	AGAAGAAGCACCATCCTGGTTTA	UUCCUGCACCCGUACCCCCGUGGUCUUU	
	GAAGAACCCAAGAAGCTTAGACC	GAAUAAAGUCUGAGUGGGCGGCAAGCUU	
	ACCCCCCGCAAGATAAAGACCAG	UUCUUCUCUCCCCUUGCACCUAAAAAAAA	
	ACAGCTGGAGCCTCGGTGGCCAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTCTTGCCCCTTGGGCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCAGCCCCTCCTCCCCTTCCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CACCCGTACCCCCGTGGTCTTTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATAAAGTCTGAGTGGGCGGCAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTTTTCTTCTCTCCCTTGCACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	T	AAAAAAAAAAAAAAAAAAAAAAA	
85	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACUAGUCAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACAGGGACUUCGUGGCCGAACCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACTAGT	GAGAAAAGCCUGUGGGGAGCCUUGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACACAGGGACCTCGTGGC	GAUAGGGGAUGUGCUCUCAAAGCGGUUA	DAFLWSHPQFEKWSHPQ
	CGAACCTATGGGAGAAAAGCCTG	GAGGAAAGAGGAUUUGAUAAGGCAUAUG	FEKWSHPQFEKRSGGNH
	TGGGGAGCCTTGCCGGGATAGGG	UAGUUUUGGGGCAGUUUUUAGUGCUGAA	RRNGRGGRGGYNRRNNG
	GATGTGCTCTCAAAGCGGTTAGA	GAAAGAUGAGGACAGGUUCCGUGAAUGG	YHPY
	GGAAAGAGGATTTGATAAGGCAT	CUUAAAGAUACAUGUGGCGCCAACGCAA	
	ATGTAGTTTTGGGGCAGTTTTTA	AGCAGUCAAGAGAUUGCUUUGGCUGUCU	
	GTGCTGAAGAAAGATGAGGACAG	UAGGGAGUGGUGUGAUGCAUUUUUGUGG	
	GTTCCGTGAATGGCTTAAAGATA	AGCCACCCUCAGUUUGAGAAGUGGAGUC	
	CATGTGGCGCCAACGCAAAGCAG	ACCCGCAAUUCGAAAAGUGGAGCCACCC	
	TCAAGAGATTGCTTTGGCTGTCT	UCAAUUCGAAAAGCGCUCAGGCGGAAAU	
	TAGGGAGTGGTGTGATGCATTTT	CAUAGGAGGAAUGGAAGAGGAGGCCGCG	
	TGTGGAGCCACCCTCAGTTTGAG	GCGGUUACAAUAGAAGGAAUAACGGUUA	
	AAGTGGAGTCACCCGCAATTCGA	CCAUCCUUAUUAAAGACCAGACAGCUGG	
	AAAGTGGAGCCACCCTCAATTCG	AGCCUCGGUGGCCAUGCUUCUUGCCCCU	
	AAAAGCGCTCAGGCGGAAATCAT	UGGGCCUCCCCCCAGCCCCUCCUCCCCU	
	AGGAGGAATGGAAGAGGAGGCCG	UCCUGCACCCGUACCCCCGUGGUCUUUG	
	CGGCGGTTACAATAGAAGGAATA	AAUAAAGUCUGAGUGGGCGGCAAGCUUU	
	ACGGTTACCATCCTTATTAAAGA	UCUUCUCUCCCCUUGCACCUAAAAAAAAA	
	CCAGACAGCTGGAGCCTCGGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCATGCTTCTTGCCCCTTGGGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCCCCCAGCCCCTCCTCCCCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCTGCACCCGTACCCCCGTGGTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTTGAATAAAGTCTGAGTGGGCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCAAGCTTTTCTTCTCTCCCTTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CACCT	AAAAAAAAAAAAAAAAAAAAAAA	
86	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTTSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUACAAGCCAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGGGACUUUGUCGCCGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTACAAGC	GAGAGAAGCCGGUCGGGUCAUUGGCAGG	ANAKQSRDCFGCLREWC
	CAAAAACATAGGGACCTTGTCGC	UAUAGGAGACGUCUUGUCAAAGAGACUU	DAFLWSHPQFEKWSHPQ
	CGAACCCATGGGAGAGAAGCCGG	GAGGAGAGGGGGUUCGACAAGGCUUACG	FEKWSHPQFEKPKTRRR
	TCGGGTCATTGGCAGGTATAGGA	UUGUGCUGGGGCAAUUUCUGGUAUUGAA	PRRSQRKRPPT
	GACGTCTTGTCAAAGAGACTTGA	GAAAGAUGAGGACAGGUUUAGAGAAUGG	
	GGAGAGGGGGTTCGACAAGGCTT	UUGAAAGAUACAUGUGGCGCAAAUGCCA	
	ACGTTGTGCTGGGGCAATTTCTG	AGCAGUCCCGGGACUGUUUCGGCUGUCU	

	GTATTGAAGAAAGATGAGGACAG GTTTAGAGAATGGTTGAAAGATA CATGTGGCGCAAATGCCAAGCAG TCCCGGGACTGTTTCGGCTGTCT GCGAGAATGGTGTGATGCTTTCC TGTGGAGCCACCCCCAATTCGAA AAATGGAGTCATCCACAGTTTGA GAAGTGGTCTCATCCTCAATTCG AAAAACCCAAGACCCGCCGTAGA CCACGCAGAAGTCAAAGAAAGAG GCCGCCCACCTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCTTCCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCAAGCT TTTCTTCTCTCCCTTGCACCT	GCGAGAAUGGUGUGAUGCUUUCCUGUGG AGCCACCCCCAAUUCGAAAAAUGGAGUC AUCCACAGUUUGAGAAGUGGUCUCAUCC UCAAUUCGAAAAACCCAAGACCCGCCGU AGACCACGCAGAAGUCAAAGAAAGAGGC CGCCCACCUAAAGACCAGACAGCUGGAG CCUCGGUGGCCAUGCUUCUUGCCCCUUG GGCCUCCCCCCCAGCCCCUCCUCCCCUUC CUGCACCCGUACCCCCGUGGUCUUUGAA UAAAGUCUGAGUGGGCGGCAAGCUUUUC UUCUCUCCCUUGCACCUAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAA	
87	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACCACGAGT CAGAAGCATCGAGATTTCGTCGC TGAACCGATGGGCGAAAAACCTG TAGGAAGTCTTGCTGGAATTGGC GACGTTTTATCCAAGCGACTTGA AGAACGTGGCTTTGACAAAGCTT ATGTCGTACTTGGCCAGTTTTTA GTTCTCAAGAAAGATGAAGATAG ATTCAGGGAGTGGCTTAAGGACA CCTGTGGTGCTAATGCTAAACAG TCTCGGGATTGTTTTGGGTGCCT TCGTGAATGGTGCGACGCATTCC TCTGGTCCCACCCACAGTTTGAA AAGTGGTCACACCCCCAGTTCGA GAAGTGGAGCCATCCCCAGTTTG AGAAACCAGCAGCTAAGAGGGTG AAATTGGATGCTTCTCAGGGTAC CAAGCGGTCTTATGAACAGATGG AAACCGACGGTGAGCGACAAGGA GGCAGTCCTAGCTCCGACGACGA AGCAACTGCAGATTCCCAACACG CAGCGCCCCCAAAAAAAAAGAGG AAAGTGGAAGATCCCAAAGATTT TCCCTCCGAGTTGCTGTCCTAAA GACCAGACAGCTGGAGCCTCGGT GGCCATGCTTCTTGCCCCTTGGG CCTCCCCCCAGCCCCTCCTCCCC TTCCTGCACCCGTACCCCCGTGG TCTTTGAATAAAGTCCTGAGTGGG CGGCAAGCTTTTCTTCTCTCCCT TGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACGAGUCAGAAG CAUCGAGAUUUCGUCGCUGAACCGAUGG GCGAAAAACCUGUAGGAAGUCUUGCUGG AAUUGGCGACGUUUUAUCCAAGCGACUU GAAGAACGUGGCUUUGACAAAGCUUAUG UCGUACUUGGCCAGUUUUUAGUUCUCAA GAAAGAUGAAGAUAGAUUCAGGGAGUGG CUUAAGGACACCUGUGGUGCUAAUGCUA AACAGUCUCGGGAUUGUUUUGGGUGCCU UCGUGAAUGGUGCGACGCAUUCCUCUGG UCCCACCCACAGUUUGAAAAGUGGUCAC ACCCCCAGUUCGAGAAGUGGAGCCAUCC CCAGUUUGAGAAACCAGCAGCUAAGAGG GUGAAAUUGGAUGCUUCUCAGGGUACCA AGCGGUCUUAUGAACAGAUGGAAACCGA CGGUGAGCGACAAGGAGGCAGUCCUAGC UCCGACGACGAAGCAACUGCAGAUUCCC AACACGCAGCGCCCCCAAAAAAAAAGAG GAAAGUGGAAGAUCCCAAAGAUUUUCCC UCCGAGUUGCUGUCCUAAAGACCAGACA GCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUGGGCCUCCCCCCAGCCCCUCCU CCCCUUCCUGCACCCGUACCCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCAA GCUUUUCUUCUCUCCCUUGCACCUAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQGGSPSSDDEATA DSQHAAPPKKKRKVEDP KDFPSELLS

88	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MPSSDDEATADSQHAAP
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PKKKRKVEDPKDFPSEL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGCCUUCCUCCGAUGAC	LSWSHPQFEKWSHPQFE
	GCGGTAGACGCACGGCATAGCAT	GAAGCCACAGCAGACUCUCAACACGCCG	KWSHPQFEKTTSQKHRD
	CCAACGCCACCATGCCTTCCTCC	CACCACCCAAAAAAAAGAGAAAGGUUGA	FVAEPMGEKPVGSLAGI
	GATGACGAAGCCACAGCAGACTC	AGAUCCUAAAGAUUUCCCAUCCGAGUUA	GDVLSKRLEERGFDKAY
	TCAACACGCCGCACCACCCAAAA	CUCUCCUGGUCACACCCACAGUUCGAGA	VVLGQFLVLKKDEDRFR
	AAAAGAGAAAGGTTGAAGATCCT	AGUGGAGCCACCCUCAAUUCGAGAAAUG	EWLKDTCGANAKQSRDC
	AAAGATTTCCCATCCGAGTTACT	GUCACAUCCCCAAUUCGAAAAGACCACU	FGCLREWCDAFL
	CTCCTGGTCACACCCACAGTTCG	AGUCAGAAACACAGAGAUUUCGUUGCUG	
	AGAAGTGGAGCCACCCTCAATTC	AGCCGAUGGGAGAGAAACCAGUGGGCAG	
	GAGAAATGGTCACATCCCCAATT	UCUCGCGGGCAUAGGUGAUGUACUGUCC	
	CGAAAAGACCACTAGTCAGAAAC	AAACGACUUGAAGAACGUGGAUUCGACA	
	ACAGAGATTTCGTTGCTGAGCCG	AAGCUUAUGUGGUCCUGGGGCAAUUCCU	
	ATGGGAGAGAAACCAGTGGGCAG	GGUGCUGAAGAAAGACGAAGACCGAUUC	
	TCTCGCGGGCATAGGTGATGTAC	CGAGAGUGGCUCAAGGAUACCUGCGGUG	
	TGTCCAAACGACTTGAAGAACGT	CAAAUGCGAAACAGAGUCGUGACUGCUU	
	GGATTCGACAAAGCTTATGTGGT	UGGGUGUUUACGGGAAUGGUGUGAUGCU	
	CCTGGGGCAATTCCTGGTGCTGA	UUUCUGUAAAGACCAGACAGCUGGAGCC	
	AGAAAGACGAAGACCGATTCCGA	UCGGUGGCCAUGCUUCUUGCCCCUUGGG	
	GAGTGGCTCAAGGATACCTGCGG	CCUCCCCCCAGCCCCUCCUCCCCUUCCU	
	TGCAAATGCGAAACAGAGTCGTG	GCACCCGUACCCCCGUGGUCUUUGAAUA	
	ACTGCTTTGGGTGTTTACGGGAA	AAGUCUGAGUGGGCGGCAAGCUUUUCUU	
	TGGTGTGATGCTTTTCTGTAAAG	CUCUCCCUUGCACCUAAAAAAAAAAAAA	
	ACCAGACAGCTGGAGCCTCGGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCATGCTTCTTGCCCCTTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCCCAGCCCCTCCTCCCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCTGCACCCGTACCCCCGTGGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTTGAATAAAGTCTGAGTGGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGCAAGCTTTTCTTCTCCCCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCACCT	AAAAAAAAAAAAAAAAAA	
89	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTSSQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACAUCCUCACAGAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CAUAGGGAUUUCGUGGCCGAACCCAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACATCCTCA	GAGAAAAACCAGUGGGUAGCCUGGCCGG	ANAKQSRDCFGCLREWC
	CAGAAACATAGGGATTCGTGGC	AAUAGGUGAUGUUCUCAGUAAGCGCCUC	DAFLWSHPQFEKWSHPQ
	CGAACCCATGGGAGAAAAACCAG	GAGGAACGGGGGUUCGAUAAAGCGUACG	FEKWSHPQFEKPAAKRV
	TGGGTAGCCTGGCCGGAATAGGT	UAGUGUUAGGACAAUUCCUCGUCUUAAA	KLDASQGTKRSYEQMET
	GATGTTCTCAGTAAGCGCCTCGA	GAAAGACGAAGAUAGGUUCCGUGAGUGG	DGERQ
	GGAACGGGGGTTCGATAAAGCGT	CUCAAAGAUACUUGUGGUGCUAAUGCUA	
	ACGTAGTGTTAGGACAATTCCTC	AACAAUCACGCGACUGUUUCGGAUGCCU	
	GTCTTAAAGAAAGACGAAGATAG	GCGUGAGUGGUGCGAUGCCUUCCUUUGG	
	GTTCCGTGAGTGGCTCAAAGATA	AGUCACCCCCAGUUCGAAAAAUGGUCAC	
	CTTGTGGTGCTAATGCTAAACAA	AUCCACAAUUUGAAAAAUGGUCACACCC	
	TCACGCGACTGTTTCGGATGCCT	CCAGUUCGAGAAACCUGCCGCAAAGAGG	
	GCGTGAGTGGTGCGATGCCTTCC	GUCAAACUGGACGCGAGUCAAGGCACAA	
	TTTGGAGTCACCCCCAGTTCGAA	AAAGAUCAUACGAGCAGAUGGAAACUGA	
	AAATGGTCACATCCACAATTTGA	UGGAGAACGCCAAUAAAGACCAGACAGC	
	AAAATGGTCACACCCCCAGTTCG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCTGCCGCAAAGAGGGTC	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTGGACGCGAGTCAAGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	AAAAAGATCATACGAGCAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACTGATGGAGAACGCCAATAA	UUUUCUUCUCUCCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
90	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGCCACCTCA CAGAAGCACCGTGATTTCGTTGC CGAACCTATGGGGGAAAAACCAG TTGGAAGCTTGGCTGGAATCGGT GATGTGTTGTCCAAACGGCTGGA AGAGCGCGGTTTCGACAAAGCCT ACGTGGTGCTTGGCCAGTTCCTT GTGCTTAAGAAGGACGAAGATAG GTTCAGAGAATGGCTTAAGGATA CCTGCGGAGCCAACGCAAAACAA AGTAGAGACTGTTTTGGATGCCT GCGTGAATGGTGCGACGCCTTCT TATGGTCTCATCCACAGTTTGAG AAGTGGTCACATCCCCAGTTCGA GAAATGGAGTCATCCCCAGTTTG AAAAGCCAGCCGCTAAACGGGTC AAGTTGGATGCCTCACAGGGAAC AAAGCGATCATATGAACAGATGG AGACAGACGGCGAGCGTCAGTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCACCUCACAGAAG CACCGUGAUUUCGUUGCCGAACCUAUGG GGGAAAAACCAGUUGGAAGCUUGGCUGG AAUCGGUGAUGUGUUGUCCAAACGGCUG GAAGAGCGCGGUUUCGACAAAGCCUACG UGGUGCUUGGCCAGUUCCUUGUCCUUAA GAAGGACGAAGAUAGGUUCAGAGAAUGG CUUAAGGAUACCUGCGGAGCCAACGCAA AACAAAGUAGAGACUGUUUUGGAUGCCU GCGUGAAUGGUGCGACGCCUUCUUAUGG UCUCAUCCACAGUUUGAGAAGUGGUCAC AUCCCCAGUUCGAGAAAUGGAGUCAUCC CCAGUUUGAAAAGCCAGCCGCUAAACGG GUCAAGUUGGAUGCCUCACAGGGAACAA AGCGAUCAUAUGAACAGAUGGAGACAGA CGGCGAGCGUCAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	MATSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
91	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGGCTGCTAGC CAAAAGCATAGAGATTTTGTTGC TGAACCAATGGGCGAGAAGCCTG TGGGCTCTCTCGCGGGAATTGGG GACGTCCTGAGTAAAAGGTTGGA GGAGCGAGGGTTTGATAAGGCAT ATGTTGTCCTGGGACAGTTCCTT GTGCTGAAAAAAGACGAAGACCG GTTTAGAGAGTGGCTCAAAGATA CATGTGGCGCTAATGCAAAGCAA TCTAGAGACTGTTTTGGTTGTTT AAGGGAGTGGTGCGATGCTTTCC TTTGGAGTCACCCACAGTTCGAG AAGTGGTCTCACCCACAGTTTGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCUGCUAGCCAAAAG CAUAGAGAUUUUGUUGCUGAACCAAUGG GCGAGAAGCCUGUGGGCUCUCUCGCGGG AAUUGGGGACGUCCUGAGUAAAAGGUUG GAGGAGCGAGGGUUUGAUAAGGCAUAUG UUGUCCUGGGACAGUUCCUUGUGCUGAA AAAAGACGAAGACCGGUUUAGAGAGUGG CUCAAAGAUACAUGUGGCGCUAAUGCAA AGCAAUCUAGAGACUGUUUUGGUUGUUU AAGGGAGUGGUGCGAUGCUUUCCUUUGG AGUCACCCACAGUUCGAGAAGUGGUCUC ACCCACAGUUUGAAAAGUGGUCCCACCC CCAGUUUGAAAAACCAGCAGCCAAACGA GUGAAACUUGAUGCAUCUCAGGGUACCA AGAGAUCCUACGAACAGAUGGAAACAGA CGGGGAGAGACAAUAAAGACCAGACAGC	MAASQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	AAAGTGGTCCCACCCCCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AAAAACCAGCAGCCAAACGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTTGATGCATCTCAGGGTAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	CAAGAGATCCTACGAACAGATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AAACAGACGGGGAGAGACAATAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
92	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MTAAQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGACUGCCGCACAAAAA	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGGGAUUUUGUGGCAGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGACTGCCGCA	GCGAAAAACCGGUGGGAUCUUUGGCCGG	ANAKQSRDCFGCLREWC
	CAAAAACACCGGGATTTTGTGGC	AAUUGGUGAUGUGUUAUCAAAGAGGCUG	DAFLWSHPQFEKWSHPQ
	AGAGCCTATGGGCGAAAAACCGG	GAGGAAAGAGGUUUUGAUAAGGCCUACG	FEKWSHPQFEKPAAKRV
	TGGGATCTTTGGCCGGAATTGGT	UUGUAUUGGGACAAUUUCUGGUGCUGAA	KLDASQGTKRSYEQMET
	GATGTGTTATCAAAGAGGCTGGA	GAAGGACGAAGACAGGUUCCGAGAAUGG	DGERQ
	GGAAAGAGGTTTTGATAAGGCCT	CUUAAGGAUACCUGCGGAGCAAACGCUA	
	ACGTTGTATTGGGACAATTTCTG	AGCAAUCACGUGACUGUUUUGGAUGUCU	
	GTGCTGAAGAAGGACGAAGACAG	GCGUGAAUGGUGUGACGCCUUUCUGUGG	
	GTTCCGAGAATGGCTTAAGGATA	UCACACCCACAAUUUGAAAAGUGGAGCC	
	CCTGCGGAGCAAACGCTAAGCAA	ACCCCCAAUUUGAGAAAUGGUCACAUCC	
	TCACGTGACTGTTTTGGATGTCT	UCAGUUUGAGAAACCCGCUGCUAAGAGA	
	GCGTGAATGGTGTGACGCCTTTC	GUGAAACUCGAUGCAAGUCAGGGCACUA	
	TGTGGTCACACCCACAATTTGAA	AACGAAGUUAUGAGCAAAUGGAGACUGA	
	AAGTGGAGCCACCCCCAATTTGA	UGGAGAAAGACAGUAAAGACCAGACAGC	
	GAAATGGTCACATCCTCAGTTTG	UGGAGCCUCGGUGGCCAUGCUUCUUGCC	
	AGAAACCCGCTGCTAAGAGAGTG	CCUUGGGCCUCCCCCCAGCCCCUCCUCC	
	AAACTCGATGCAAGTCAGGGCAC	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	TAAACGAAGTTATGAGCAAATGG	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGACTGATGGAGAAAGACAGTAA	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAA	
93	GACATTGGTCTGTACATTTAATA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MATAQKHRDFVAEPMGE
	CGACTCACTATAAGGGGATAATA	GUCAGCGGUAGACGCACGGCAUAGCAUC	KPVGSLAGIGDVLSKRL
	GGAAAAATCTATAGCAGAAGTCA	CAACGCCACCAUGGCAACAGCUCAGAAG	EERGFDKAYVVLGQFLV
	GCGGTAGACGCACGGCATAGCAT	CACCGGGAUUUUGUGGCAGAGCCUAUGG	LKKDEDRFREWLKDTCG
	CCAACGCCACCATGGCAACAGCT	GGGAGAAACCGGUCGGGAGUUUAGCCGG	ANAKQSRDCFGCLREWC
	CAGAAGCACCGGGATTTTGTGGC	CAUUGGAGACGUUUUAAGUAAACGACUG	DAFLWSHPQFEKWSHPQ
	AGAGCCTATGGGGGAGAAACCGG	GAGGAAAGAGGAUUUGACAAAGCCUAUG	FEKWSHPQFEKPAAKRV
	TCGGGAGTTTAGCCGGCATTGGA	UGGUGCUCGGUCAGUUUCUCGUGCUGAA	KLDASQGTKRSYEQMET
	GACGTTTTAAGTAAACGACTGGA	GAAGGACGAAGAUAGGUUUAGGGAAUGG	DGERQ
	GGAAAGAGGATTTGACAAAGCCT	UUAAAAGAUACAUGUGGCGCUAACGCGA	
	ATGTGGTGCTCGGTCAGTTTCTC	AACAGAGCAGAGAUUGCUUCGGCUGCUU	
	GTGCTGAAGAAGGACGAAGATAG	GAGAGAGUGGUGUGACGCCUUUCUCUGG	
	GTTTAGGGAATGGTTAAAAGATA	UCCCACCCUCAGUUUGAGAAAUGGUCCC	
	CATGTGGCGCTAACGCGAAACAG	AUCCCCAGUUUGAGAAGUGGUCACACCC	
	AGCAGAGATTGCTTCGGCTGCTT	GCAGUUUGAGAAGCCUGCAGCCAAGCGC	

	GAGAGAGTGGTGTGACGCCTTTC TCTGGTCCCACCCTCAGTTTGAG AAATGGTCCCATCCCCAGTTTGA GAAGTGGTCACACCCGCAGTTTG AGAAGCCTGCAGCCAAGCGCGTA AAACTGGACGCATCACAAGGGAC TAAACGCTCCTATGAGCAGATGG AGACCGACGGCGAACGGCAATAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	GUAAAACUGGACGCAUCACAAGGGACUA AACGCUCCUAUGAGCAGAUGGAGACCGA CGGCGAACGGCAAUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
94	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACTACAACT TCTCAGAAACACAGAGACTTTGT CGCTGAGCCCATGGGTGAAAAGC CAGTAGGAAGCTTAGCAGGAATT GGAGATGTACTTTCAAAGCGTCT GGAAGAGAGGGGATTTGATAAAG CTTACGTTGTTTTGGGGCAGTTT CTCGTGTTAAAGAAGGACGAAGA CCGATTTCGGGAGTGGCTTAAAG ATACATGTGGTGCCAACGCCAAA CAATCACGTGACTGCTTCGGTTG TCTGAGAGAGTGGTGCGATGCTT TTTTGTGGTCACATCCGCAGTTT GAGAAATGGAGCCATCCCCAGTT TGAGAAATGGAGCCACCCACAGT TCGAGAAACCAGCCGCCAAGCGT GTGAAACTGGATGCTTCTCAGGG TACCAAACGAAGCTATGAACAGA TGGAAACAGACGGAGAACGGCAA TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACUACAACUUCUCAG AAACACAGAGACUUUGUCGCUGAGCCCA UGGGUGAAAAGCCAGUAGGAAGCUUAGC AGGAAUUGGAGAUGUACUUUCAAAGCGU CUGGAAGAGAGGGGAUUUGAUAAAGCUU ACGUUGUUUUGGGGCAGUUUCUCGUGUU AAAGAAGGACGAAGACCGAUUUCGGGAG UGGCUUAAAGAUACAUGUGGUGCCAACG CCAAACAAUCACGUGACUGCUUCGGUUG UCUGAGAGAGUGGUGCGAUGCUUUUUUG UGGUCACAUCCGCAGUUUGAGAAAUGGA GCCAUCCCCAGUUUGAGAAAUGGAGCCA CCCACAGUUCGAGAAACCAGCCGCCAAG CGUGUGAAACUGGAUGCUUCUCAGGGUA CCAAACGAAGCUAUGAACAGAUGGAAAC AGACGGAGAACGGCAAUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	MTTTSQKHRDFVAEPMG EKPVGSLAGIGDVLSKR LEERGFDKAYVVLGQFL VLKKDEDRFREWLKDTC GANAKQSRDCFGCLREW CDAFLWSHPQFEKWSHP QFEKWSHPQFEKPAAKR VKLDASQGTKRSYEQME TDGERQ
95	GACATTGGTCTGTACATTTAATA CGACTCACTATAAGGGGATAATA GGAAAAATCTATAGCAGAAGTCA GCGGTAGACGCACGGCATAGCAT CCAACGCCACCATGACCACAAGT CAGAAACATCGGGACTTTGTTGC TGAACCAATGGGGGAAAAACCTG TGGGCTCCCTGGCCGGGATCGGT GATGTATTATCTAAGAGGCTGGA GGAACGCGGTTTCGATAAGGCTT ATGTGGTATTGGGACAATTCCTC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACAAGUCAGAAA CAUCGGGACUUUGUUGCUGAACCAAUGG GGGAAAAACCUGUGGGCUCCCUGGCCGG GAUCGGUGAUGUAUUAUCUAAGAGGCUG GAGGAACGCGGUUUCGAUAAGGCUUAUG UGGUAUUGGGACAAUUCCUCGUGCUCAA GAAGGACGAAGAUAGGUUUAGAGAGUGG UUAAAGGAUACCUGCGGCGCCAACGCAA AGCAAAGCAGGGACUGUUUCGGUUGUCU	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	GTGCTCAAGAAGGACGAAGATAG GTTTAGAGAGTGGTTAAAGGATA CCTGCGGCGCCAACGCAAAGCAA AGCAGGGACTGTTTCGGTTGTCT TCGAGAGTGGTGTGATGCCTTTT TATGGTCACATCCTCAGTTCGAG AAATGGAGCCACCCTCAGTTTGA AAAATGGTCCCACCCCCAATTTG AAAAGCCCGCCGCCAAGAGAGTA AAACTGGACGCGAGCCAAGGTAC AAAGAGAAGTTATGAGCAGATGG AGACCGACGGAGAACGGCAGTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCAAGCTTTTCTTCTCTCCC TTGCACCT	UCGAGAGUGGUGUGAUGCCUUUUUAUGG UCACAUCCUCAGUUCGAGAAAUGGAGCC ACCCUCAGUUUGAAAAAUGGUCCCACCC CCAAUUUGAAAAGCCCGCCGCCAAGAGA GUAAAACUGGACGCGAGCCAAGGUACAA AGAGAAGUUAUGAGCAGAUGGAGACCGA CGGAGAACGGCAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
96	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
97	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACCGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGATGACCACC AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACCG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GAUGACCACCAGCCAGAAGCACAGGGAC UUCGUGGCCGAGCCCAUGGGCGAGAAGC CCGUGGGCAGCCUGGCCGGCAUCGGCGA CGUGCUGAGCAAGAGGCUGGAGGAGAGG GGCUUCGACAAGGCCUACGUGGUGCUGG	MASQGTKRSYEQMETDG ERQPAAKRVKLDWSHPQ FEKWSHPQFEKWSHPQF EKMTTSQKHRDFVAEPM GEKPVGSLAGIGDVLSK RLEERGFDKAYVVLGQF LVLKKDEDRFREWLKDT CGANAKQSRDCFGCLRE WCDAFL

	CCGTGGGCAGCCTGGCCGGCATC GGCGACGTGCTGAGCAAGAGGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CAGGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAAGACCAGACAGCTGG AGCCTCGGTGGCCATGCTTCTTG CCCCTTGGGCCTCCCCCCAGCCC CTCCTCCCCTTCCTGCACCCGTA CCCCCGTGGTCTTTGAATAAAGT CTGAGTGGGCGGCAAGCTTTTCT TCTCTCCCTTGCACCT	GCCAGUUCCUGGUGCUGAAGAAGGACGA GGACAGGUUCAGGGAGUGGCUGAAGGAC ACCUGCGGCGCCAACGCCAAGCAGAGCA GGGACUGCUUCGGCUGCCUGAGGGAGUG GUGCGACGCCUUCCUGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
98	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCG CCGCCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGTGGAGCCACCCCCA GTTCGAGAAGCCCGCCGCCAAGA GGGTGAAGCTGGACGCCAGCCAG GGCACCAAGAGGAGCTACGAGCA GATGGAGACCGACGGCGAGAGGC AGTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCAAGCTTTTCTTCT CTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCGCCGCCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG AGCCACCCCCAGUUCGAGAAGUGGAGCC ACCCCCAGUUCGAGAAGUGGAGCCACCC CCAGUUCGAGAAGCCCGCCGCCAAGAGG GUGAAGCUGGACGCCAGCCAGGGCACCA AGAGGAGCUACGAGCAGAUGGAGACCGA CGGCGAGAGGCAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MAAAQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ
99	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACCGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACCG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GAUGGCCGCCGCCCAGAAGCACAGGGAC	MASQGTKRSYEQMETDG ERQPAAKRVKLDWSHPQ FEKWSHPQFEKWSHPQF EKMAAAQKHRDFVAEPM GEKPVGSLAGIGDVLSK RLEERGFDKAYVVLGQF LVLKKDEDRFREWLKDT CGANAKQSRDCFGCLRE WCDAFL

	CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGATGGCCGCC GCCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGACGTGCTGAGCAAGAGGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CAGGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAAGACCAGACAGCTGG AGCCTCGGTGGCCATGCTTCTTG CCCCTTGGGCCTCCCCCCAGCCC CTCCTCCCCTTCCTGCACCCGTA CCCCCGTGGTCTTTGAATAAAGT CTGAGTGGGCGGCAAGCTTTTCT TCTCTCCCTTGCACCT	UUCGUGGCCGAGCCCAUGGGCGAGAAGC CCGUGGGCAGCCUGGCCGGCAUCGGCGA CGUGCUGAGCAAGAGGCUGGAGGAGAGG GGCUUCGACAAGGCCUACGUGGUGCUGG GCCAGUUCCUGGUGCUGAAGAAGGACGA GGACAGGUUCAGGGAGUGGCUGAAGGAC ACCUGCGGCGCCAACGCCAAGCAGAGCA GGGACUGCUUCGGCUGCCUGAGGGAGUG GUGCGACGCCUUCCUGUAAAGACCAGAC AGCUGGAGCCUCGGUGGCCAUGCUUCUU GCCCCUUGGGCCUCCCCCCAGCCCCUCC UCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGCA AGCUUUUCUUCUCUCCCUUGCACCUAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
100	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGAGGTGCTGGGCAAGAAG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGAGGUGCUGGGCAAGAAGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
101	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

	TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
102	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGAGGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGAGGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIEDVLSKRL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
103	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGAGCAGT TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGAGCAGUUCCUGGUGCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLEQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

	TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GCACCUAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
104	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGTGGAGCCACCCCCA GTTCGAGAAGTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCTTCCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCAAGCT TTTCTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG AGCCACCCCCAGUUCGAGAAGUGGAGCC ACCCCCAGUUCGAGAAGUGGAGCCACCC CCAGUUCGAGAAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEK
105	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGACCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGACCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVTEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

	GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
106	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACG ACGACCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACGACGACCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MDDDQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
107	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCG CCGCCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCGCCGCCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MAAAQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
108	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCG CCGCCCAGAAGCACAGGGACTTC GTGACCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCGCCGCCCAGAAG CACAGGGACUUCGUGACCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MAAAQKHRDFVTEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
109	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGAACGAGCTGGCCCTGA AGCTGGCCGGCCTGGACATCAAC AAGACCTAAAGACCAGACAGCTG GAGCCTCGGTGGCCATGCTTCTT GCCCCTTGGGCCTCCCCCCAGCC CCTCCTCCCCTTCCTGCACCCGT ACCCCCGTGGTCTTTGAATAAAG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGAAC GAGCUGGCCCUGAAGCUGGCCGGCCUGG ACAUCAACAAGACCUAAAGACCAGACAG CUGGAGCCUCGGUGGCCAUGCUUCUUGC CCCUUGGGCCUCCCCCCAGCCCCUCCUC CCCUUCCUGCACCCGUACCCCCGUGGUC UUUGAAUAAAGUCUGAGUGGGCGGCAAG CUUUUCUUCUCUCCCUUGCACCUAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLNELALKLAGLDIN KT

	TCTGAGTGGGCGGCAAGCTTTTC TTCTCTCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	
110	TAATACGACTCACTA‾AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGC‾TCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAG‾GGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGG‾GCGACGC CTTCCTGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGTGGAGCCACCCCCA GTTCGAGAAGAACGAGCTGGCCC TGAAGCTGGCCGGCC‾GGACATC AACAAGACCTAAAGACCAGACAG CTGGAGCCTCGGTGGCCATGCTT CTTGCCCCTTGGGCC‾CCCCCCA GCCCCTCCTCCCCTTCCTGCACC CGTACCCCCGTGGTC‾TTGAATA AAGTCTGAGTGGGCGGCAAGCTT TTCTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG AGCCACCCCCAGUUCGAGAAGUGGAGCC ACCCCCAGUUCGAGAAGUGGAGCCACCC CCAGUUCGAGAAGAACGAGCUGGCCCUG AAGCUGGCCGGCCUGGACAUCAACAAGA CCUAAAGACCAGACAGCUGGAGCCUCGG UGGCCAUGCUUCUUGCCCCUUGGGCCUC CCCCCAGCCCCUCCUCCCCUUCCUGCAC CCGUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGCAAGCUUUUCUUCUCU CCCUUGCACCUAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	
111	TAATACGACTCACTA‾AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGC‾TCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAG‾GGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGG‾GCGACGC CTTCCTGCTGCCCCCCCTGGAGC GCCTGACCCTGTAAAGACCAGAC AGCTGGAGCCTCGGTGGCCATGC TTCTTGCCCCTTGGGCCTCCCCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGCUG CCCCCCCUGGAGCGCCUGACCCUGUAAA GACCAGACAGCUGGAGCCUCGGUGGCCA UGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCAAGCUUUUCUUCUCUCCCUUG CACCUAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKNELALK LAGLDINKT
	TAATACGACTCACTA‾AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGC‾TCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAG‾GGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGG‾GCGACGC CTTCCTGCTGCCCCCCCTGGAGC GCCTGACCCTGTAAAGACCAGAC AGCTGGAGCCTCGGTGGCCATGC TTCTTGCCCCTTGGGCCTCCCCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGCUG CCCCCCCUGGAGCGCCUGACCCUGUAAA GACCAGACAGCUGGAGCCUCGGUGGCCA UGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCAAGCUUUUCUUCUCUCCCUUG CACCUAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLLPPLER‾TL

	CAGCCCCTCCTCCCCTTCCTGCA CCCGTACCCCCGTGGTCTTTGAA TAAAGTCTGAGTGGGCGGCAAGC TTTTCTTCTCTCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAA	
112	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGTGGAGCCACCCCCA GTTCGAGAAGCTGCCCCCCCTGG AGCGCCTGACCCTGTAAAGACCA GACAGCTGGAGCCTCGGTGGCCA TGCTTCTTGCCCCTTGGGCCTCC CCCCAGCCCCTCCTCCCCTTCCT GCACCCGTACCCCCGTGGTCTTT GAATAAAGTCTGAGTGGGCGGCA AGCTTTTCTTCTCTCCCTTGCAC CT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG AGCCACCCCCAGUUCGAGAAGUGGAGCC ACCCCCAGUUCGAGAAGUGGAGCCACCC CCAGUUCGAGAAGCUGCCCCCCCUGGAG CGCCUGACCCUGUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCAAGCU UUUCUUCUCUCCCUUGCACCUAAAAAAA AAAAAAAAAAAAAAAAAAAA

	GTCTGAGTGGGCGGCCGGATCAG TACTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	
114	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACG ACGACCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCCGGATCAG TACTAGACTGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACGACGACCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCCGGAUCAGUACUAGACUGAU AGCUUAUUGGGAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	MDDDQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
115	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGTGCA TCATCTGCATGGACGAGCTGATC CACAGCCCCAACCAGCAGACCTG GAAGAACAAGAACAAGAAGCCCA AGAGGCTGCCCTGCGGCCACATC CTGCACCTGAGCTGCCTGAAGAA CTGGATGGAGAGGAGCCAGACCT GCCCCATCTGCAGGGGCGGCGGC AGCGGCGGCGGCAGCACCACCAG CCAGAAGCACAGGGACTTCGTGG CCGAGCCCATGGGCGAGAAGCCC GTGGGCAGCCTGGCCGGCATCGG CGACGTGCTGAGCAAGAGGCTGG AGGAGAGGGGCTTCGACAAGGCC TACGTGGTGCTGGGCCAGTTCCT GGTGCTGAAGAAGGACGAGGACA GGTTCAGGGAGTGGCTGAAGGAC ACCTGCGGCGCCAACGCCAAGCA GAGCAGGGACTGCTTCGGCTGCC TGAGGGAGTGGTGCGACGCCTTC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUGCAUCAUCUGCAUG GACGAGCUGAUCCACAGCCCCAACCAGC AGACCUGGAAGAACAAGAACAAGAAGCC CAAGAGGCUGCCCUGCGGCCACAUCCUG CACCUGAGCUGCCUGAAGAACUGGAUGG AGAGGAGCCAGACCUGCCCCAUCUGCAG GGGCGGCGGCAGCGGCGGCGGCAGCACC ACCAGCCAGAAGCACAGGGACUUCGUGG CCGAGCCCAUGGGCGAGAAGCCCGUGGG CAGCCUGGCCGGCAUCGGCGACGUGCUG AGCAAGAGGCUGGAGGAGAGGGGCUUCG ACAAGGCCUACGUGGUGCUGGGCCAGUU CCUGGUGCUGAAGAAGGACGAGGACAGG UUCAGGGAGUGGCUGAAGGACACCUGCG GCGCCAACGCCAAGCAGAGCAGGGACUG CUUCGGCUGCCUGAGGGAGUGGUGCGAC GCCUUCCUGUAAAGACCAGACAGCUGGA GCCUCCGCUGGCCAUGCUUCUUGCCCCUU GGGCCUCCCCCCAGCCCCUCCUCCCCUU CCUGCACCCGUACCCCCGUGGUCUUUGA AUAAAGUCUGAGUGGGCGGCCGGAUCAG	MCIICMDELIHSPNQQT WKNKNKKPKRLPCGHIL HLSCLKNWMERSQTCPI CRGGGSGGGSTTSQKHR DFVAEPMGEKPVGSLAG IGDVLSKRLEERGFDKA YVVLGQFLVLKKDEDRF REWLKDTCGANAKQSRD CFGCLREWCDAFL

	CTGTAAAGACCAGACAGCTGGAG CCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCCAGCCCCT CCTCCCCTTCCTGCACCCGTACC CCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCCGGATCAGTACT AGACTGATAGCTTATTGGG	UACUAGACUGAUAGCUUAUUGGGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
116	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGTGCA TCATCTGCAGGGAGGAGATGGTG ACCGGCGCCAAGAGGCTGCCCTG CAACCACATCTTCCACACCAGCT GCCTGAGGAGCTGGTTCCAGAGG CAGCAGACCTGCCCCACCTGCAG GGGCGGCGGCAGCGGCGGCGGCA GCACCACCAGCCAGAAGCACAGG GACTTCGTGGCCGAGCCCATGGG CGAGAAGCCCGTGGGCAGCCTGG CCGGCATCGGCGACGTGCTGAGC AAGAGGCTGGAGGAGAGGGGCTT CGACAAGGCCTACGTGGTGCTGG GCCAGTTCCTGGTGCTGAAGAAG GACGAGGACAGGTTCAGGGAGTG GCTGAAGGACACCTGCGGCGCCA ACGCCAAGCAGAGCAGGGACTGC TTCGGCTGCCTGAGGGAGTGGTG CGACGCCTTCCTGTAAAGACCAG ACAGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCCTTGGGCCTCCC CCCAGCCCCTCCTCCCCTTCCTG CACCCGTACCCCCGTGGTCTTTG AATAAAGTCTGAGTGGGCGGCCG GATCAGTACTAGACTGATAGCTT ATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUGCAUCAUCUGCAGG GAGGAGAUGGUGACCGGCGCCAAGAGGC UGCCCUGCAACCACAUCUUCCACACCAG CUGCCUGAGGAGCUGGUUCCAGAGGCAG CAGACCUGCCCCACCUGCAGGGGCGGCG GCAGCGGCCGGCGGCAGCACCACCAGCCA GAAGCACAGGGACUUCGUGGCCGAGCCC AUGGGCGAGAAGCCCGUGGGCAGCCUGG CCGGCAUCGGCGACGUGCUGAGCAAGAG GCUGGAGGAGAGGGGCUUCGACAAGGCC UACGUGGUGCUGGGCCAGUUCCUGGUGC UGAAGAAGGACGAGGACAGGUUCAGGGA GUGGCUGAAGGACACCUGCGGCGCCAAC GCCAAGCAGAGCAGGGACUGCUUCGGCU GCCUGAGGGAGUGGUGCGACGCCUUCCU GUAAAGACCAGACAGCUGGAGCCUCGGU GGCCAUGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGCCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MCIICREEMVTGAKRLP CNHIFHTSCLRSWFQRQ QTCPTCRGGGSGGGSTT SQKHRDFVAEPMGEKPV GSLAGIGDVLSKRLEER GFDKAYVVLGQFLVLKK DEDRFREWLKDTCGANA KQSRDCFGCLREWCDAF L
117	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA GCGGCATCCACAGCACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GACGTGCTGAGCAAGAGGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACAG GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAGCGGCAUCCAC AGCACCACCAGCCAGAAGCACAGGACU UCGUGGCCGAGCCCAUGGGCGAGAAGCC CGUGGGCAGCCUGGCCGGCAUCGGCGAC GUGCUGAGCAAGAGGCUGGAGGAGAGGG GCUUCGACAAGGCCUACGUGGUGCUGGG CCAGUUCCUGGUGCUGAAGAAGGACGAG GACAGGUUCAGGGAGUGGCUGAAGGACA CCUGCGGCGCCAACGCCAAGCAGAGCAG GGACUGCUUCGGCUGCCUGAGGGAGUGG UGCGACGCCUUCCUGUAAAGACCAGACA GCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUCGGCCUCCCCCCAGCCCCUCCU CCCCUUCCUGCACCCGUACCCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCCG GAUCAGUACUAGACUGAUAGCUUAUUGG	MDSGIHSTTSQKHRDFV AEPMGEKPVGSLAGIGD VLSKRLEERGFDKAYVV LGQFLVLKKDEDRFREW LKDTCGANAKQSRDCFG CLREWCDAFL

	CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCCGGATCAGTACTA GACTGATAGCTTATTGGG	GAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA AAAAA	
118	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCTGC CCACCCCCCCCACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAC GTGCTGAGCAAGAGGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCUGCCCACCCCCCCC ACCACCAGCCAGAAGCACAGGGACUUCG UGGCCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGACGUG CUGAGCAAGAGGCUGGAGGAGAGGGCU UCGACAAGGCCUACGUGGUGCUCGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCCGGAU CAGUACUAGACUGAUAGCUUAUUGGGAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AA	MLPTPPTTSQKHRDFVA EPMGEKPVGSLAGIGDV LSKRLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL
119	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCCCG AGAGCACCGAGCCCACCAGCCCC ACCAGCGAGACCAGCCAGAAGCA CAGGGACTTCGTGGCCGAGCCCA TGGGCGAGAAGCCCGTGGGCAGC CTGGCCGGCATCGGCGACGTGCT GAGCAAGAGGCTGGAGGAGAGGG GCTTCGACAAGGCCTACGTGGTG CTGGGCCAGTTCCTGGTGCTGAA GAAGGACGAGGACAGGTTCAGGG AGTGGCTGAAGGACACCTGCGGC GCCAACGCCAAGCAGAGCAGGGA CTGCTTCGGCTGCCTGAGGGAGT GGTGCGACGCCTTCCTGTAAAGA CCAGACAGCTGGAGCCTCGGTGG CCATGCTTCTTGCCCCTTGGGCC TCCCCCCAGCCCCTCCTCCCCTT CCTGCACCCGTACCCCCGTGGTC TTTGAATAAAGTCTGAGTGGGCG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCCCGAGAGCACCGAG CCCACCAGCCCCACCAGCGAGACCAGCC AGAAGCACAGGGACUUCGUGGCCGAGCC CAUGGGCGAGAAGCCCGUGGGCAGCCUG GCCGGCAUCGGCGACGUGCUGAGCAAGA GGCUGGAGGAGAGGGGCUUCGACAAGGC CUACGUGGUGCUGGGCCAGUUCCUGGUG CUGAAGAAGGACGAGGACAGGUUCAGGG AGUGGCUGAAGGACACCUGCGGCGCCAA CGCCAAGCAGAGCAGGGACUGCUUCGGC UGCCUGAGGGAGUGGUGCGACGCCUUCC UGUAAAGACCAGACAGCUGGAGCCUCGG UGGCCAUGCUUCUUGCCCCUUGGGCCUC CCCCCAGCCCCUCCUCCCCUUCCUGCAC CCGUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGCCGGAUCAGUACUAGA CUGAUACCUUAUUGGGAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	MPESTEPTSPTSETSQK HRDFVAEPMGEKPVGSL AGIGDVLSKRLEERGFD KAYVVLGQFLVLKKDED RFREWLKDTCGANAKQS RDCFGCLREWCDAFL

	GCCGGATCAGTACTAGACTGATA GCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
120	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGCCCGAGAGCACCGAGC CCACCAGCCCCACCAGCGAGTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCCGGATCAGTACTAGACTG ATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGCCC GAGAGCACCGAGCCCACCAGCCCCACCA GCGAGUAAAGACCAGACAGCUGGAGCCU CGGUGGCCAUGCUUCUUGCCCCUUGGGC CUCCCCCCAGCCCCUCCUCCCCUUCCUG CACCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGCCGGAUCAGUACU AGACUGAUAGCUUAUUGGGAAAAAAAAA AAAAAAAAAAAAAAAAAAA

	GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AA	
122	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGCAGTTCGAGAGGCAGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGCAG UUCGAGAGGCAGUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCCGGAU CAGUACUAGACUGAUAGCUUAUUGGGAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLQFERQ
123	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGGACTACAAGGACCACG ACGGCGACTACAAGGACCACGAC ATCGACTACAAGGACGACGACGA CAAGTAAAGACCAGACAGCTGGA GCCTCGGTGGCCATGCTTCTTGC CCCTTGGGCCTCCCCCCAGCCCC TCCTCCCCTTCCTGCACCCGTAC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGGAC UACAAGGACCACGACGGCGACUACAAGG ACCACGACAUCGACUACAAGGACGACGA CGACAAGUAAAGACCAGACAGCUGGAGC CUCGGUGGCCAUGCUUCUUGCCCCUUGG GCCUCCCCCCAGCCCCUCCUCCCCUUCC UGCACCCGUACCCCCGUGGUCUUUGAAU AAAGUCUGAGUGGGCGGCCGGAUCAGUA CUAGACUGAUAGCUUAUUGGGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLDYKDHDGDYKDHD IDYKDDDDK

	CCCCGTGGTCTTTGAATAAAGTC TGAGTGGGCGGCCGGATCAGTAC TAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
124	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGCACCACCACCACCACC ACTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCCGGATCAGTACTA GACTGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGCAC CACCACCACCACCACUAAAGACCAGACA GCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUGGGCCUCCCCCCAGCCCCUCCU CCCCUUCCUGCACCCGUACCCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCCG GAUCAGUACUAGACUGAUAGCUUAUUGG GAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLHHHHHH
125	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACATGAGCCCCAGGCTGAGGGCC TTCCTGAGCGAGCCCATCGGCGA GAAGGACGTGTGCTGGGTGGACG GCATCAGCCACGAGCTGGCCATC AACCTGGTGACCAAGGGCATCAA CAAGGCCTACATCCTGCTGGGCC AGTTCCTGCTGATGCACAAGAAC GAGGCCGAGTTCCAGAGGTGGCT GATCTGCTGCTTCGGCGCCACCG AGTGCGAGGCCCAGCAGACCAGC CACTGCCTGAAGGAGTGGTGCGC CTGCTTCCTGTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCTTCCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCCGGAT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACAUGAGCCCC AGGCUGAGGGCCUUCCUGAGCGAGCCCA UCGGCGAGAAGGACGUGUGCUGGGUGGA CGGCAUCAGCCACGAGCUGGCCAUCAAC CUGGUGACCAAGGGCAUCAACAAGGCCU ACAUCCUGCUGGGCCAGUUCCUGCUGAU GCACAAGAACGAGGCCGAGUUCCAGAGG UGGCUGAUCUGCUGCUUCGGCGCCACCG AGUGCGAGGCCCAGCAGACCAGCCACUG CCUGAAGGAGUGGUGCGCCUGCUUCCUG UAAAGACCAGACAGCUGGAGCCUCGGUG GCCAUGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCU GAGUGGGCGGCCGGAUCAGUACUAGACU GAUAGCUUAUUGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	MDNMSPRLRAFLSEPIG EKDVCWVDGISHELAIN LVTKGINKAYILLGQFL LMHKNEAEFQRWLICCF GATECEAQQTSHCLKEW CACFL

	CAGTACTAGACTGATAGCTTATT GGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	
126	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCTGA GGGCCTTCCTGAGCGAGCCCATC GGCGAGAAGGACGTGTGCTGGGT GGACGGCATCAGCCACGAGCTGG CCATCAACCTGGTGACCAAGGGC ATCAACAAGGCCTACATCCTGCT GGGCCAGTTCCTGCTGATGCACA AGAACGAGGCCGAGTTCCAGAGG TGGCTGATCTGCTGCTTCGGCGC CACCGAGTGCGAGGCCCAGCAGA CCAGCCACTGCCTGAAGGAGTGG TGCGCCTGCTTCCTGTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC CGGATCAGTACTAGACTGATAGC TTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCUGAGGGCCUUCCUG AGCGAGCCCAUCGGCGAGAAGGACGUGU GCUGGGUGGACGGCAUCAGCCACGAGCU GGCCAUCAACCUGGUGACCAAGGGCAUC AACAAGGCCUACAUCCUGCUGGGCCAGU UCCUGCUGAUGCACAAGAACGACGCCGA GUUCCAGAGGUGGCUGAUCUGCUGCUUC GGCGCCACCGAGUGCGAGGCCCAGCAGA CCAGCCACUGCCUGAAGGAGUGGUGCGC CUGCUUCCUGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MLRAFLSEPIGEKDVCW VDGISHELAINLVTKGI NKAYILLGQFLMHKNE AEFQRWLICCFGATECE AQQTSHCLKEWCACFL
127	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGGCA ACGAGCTGGCCCTGAAGCTGGCC GGCCTGGACATCACCACCAGCCA GAAGCACAGGGACTTCGTGGCCG AGCCCATGGGCGAGAAGCCCGTG GGCAGCCTGGCCGGCATCGGCGA CGTGCTGAGCAAGAGGCTGGAGG AGAGGGGCTTCGACAAGGCCTAC GTGGTGCTGGGCCAGTTCCTGGT GCTGAAGAAGGACGAGGACAGGT TCAGGGAGTGGCTGAAGGACACC TGCGGCGCCAACGCCAAGCAGAG CAGGGACTGCTTCGGCTGCCTGA GGGAGTGGTGCGACGCCTTCCTG TAAAGACCAGACAGCTGGAGCCT CGGTGGCCATGCTTCTTGCCCCT TGGGCCTCCCCCCAGCCCCTCCT CCCCTTCCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAGCTTTTCTTCTCT CCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGGCAACGAGCUGGCC CUGAAGCUGGCCGGCCUGGACAUCACCA CCAGCCAGAAGCACAGGGACUUCGUGGC CGAGCCCAUGGGCGAGAAGCCCGUGGGC AGCCUGGCCGGCAUCGGCGACGUGCUGA GCAAGAGGCUGGAGGAGAGGGGCUUCGA CAAGGCCUACGUGGUGCUGGGCCAGUUC CUGGUGCUGAAGAAGGACGAGGACAGGU UCAGGGAGUGGCUGAAGGACACCUGCGG CGCCAACGCCAAGCAGAGCAGGGACUGC UUCGGCUGCCUGAGGGAGUGGUGCGACG CCUUCCUGUAAAGACCAGACAGCUGGAG CCUCGGUGGCCAUGCUUCUUGCCCCUUG GGCCUCCCCCCAGCCCCUCCUCCCCUUC CUGCACCCGUACCCCCGUGGUCUUUGAA UAAAGUCUGAGUGGGCGGCAAGCUUUUC UUCUCUCCCUUGCACCUAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MGNELALKLAGLDITTS QKHRDFVAEPMGEKPVG SLAGIGDVLSKRLEERG FDKAYVVLGQFLVLKKD EDRFREWLKDTCGANAK QSRDCFGCLREWCDAFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
128	TAATACGACTCACTA‾AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACTACGCCGACCTGAGCGACACC GAGCTGACCACCCTGCTGAGGAG GTACAACATCCCCCACGGCCCCG TGGTGGGCAGCACCAGGAGGCTG TACGAGAAGAAGATC‾TCGAGTA CGAGACCCAGAGGAGGAGGTAAA GACCAGACAGCTGGAGCCTCGGT GGCCATGCTTCTTGCCCCTTGGG CCTCCCCCCAGCCCC‾CCTCCCC TTCCTGCACCCGTACCCCCGTGG TCTTTGAATAAAGTC‾GAGTGGG CGGCCGGATCAGTAC‾AGACTGA TAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACUACGCCGAC CUGAGCGACACCGAGCUGACCACCCUGC UGAGGAGGUACAACAUCCCCCACGGCCC CGUGGUGGGCAGCACCAGGAGGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGAGGAGGAGGUAAAGACCAGACAGCUG GAGCCUCGGUGGCCAUGCUUCUUGCCCC UUGGGCCUCCCCCCAGCCCCUCCUCCCC UUCCUGCACCCGUACCCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGCCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	MDNYADLSDTELTTLLR RYNIPHGPVVGSTRRLY EKKIFEYETQRRR
129	TAATACGACTCACTA‾AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACG ACTACGCCGTGCTGAGCGACACC GAGCTGGCCGCCGTGCTGAGGCA GTACAACATCCCCCACGGCCCCG TGGTGGGCAGCACCAGGAAGCTG TACGAGAAGAAGATC‾TCGAGTA CGAGACCCAGTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCT‾CCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCCGGAT CAGTACTAGACTGATAGCTTATT GGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACGACUACGCCGUG CUGAGCGACACCGAGCUGGCCGCCGUGC UGAGGCAGUACAACAUCCCCCACGGCCC CGUGGUGGGCAGCACCAGGAAGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGUAAAGACCAGACAGCUGGAGCCUCGG UGGCCAUGCUUCUUGCCCCUUGGGCCUC CCCCCAGCCCCUCCUCCCCUUCCUGCAC CCGUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGCCGGAUCAGUACUAGA CUGAUAGCUUAUUGGGAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	MDDYAVLSDTELAAVLR QYNIPHGPVVGSTRKLY EKKIFEYETQ

130	TAATACGACTCACTA⁻AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACC TGGACGTGACCGAGC⁻GACCAAC GAGGACCTGCTGGACCAGCTGGT GAAGTACGGCGTGAACCCCGGCC CCATCGTGGGCACCACCAGGAAG CTGTACGAGAAGAAGCTGCTGAA GCTGAGGGAGCAGTAAAGACCAG ACAGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCCTTGGGCCTCCC CCCAGCCCCTCCTCCCCTTCCTG CACCCGTACCCCCGTGGTCTTTG AATAAAGTCTGAGTGGGCGGCCG GATCAGTACTAGACTGATAGCTT ATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACCUGGACGUGACC GAGCUGACCAACGAGGACCUGCUGGACC AGCUGGUGAAGUACGGCGUGAACCCCGG CCCCAUCGUGGGCACCACCAGGAAGCUG UACGAGAAGAAGCUGCUGAAGCUGAGGG AGCAGUAAAGACCAGACAGCUGGAGCCU CGGUGGCCAUGCUUCUUGCCCCUUGGGC CUCCCCCCAGCCCCUCCUCCCCUUCCUG CACCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGCCGGAUCAGUACU AGACUGAUAGCUUAUUGGGAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	MDLDVTELTNEDLLDQL VKYGVNPGPIVGTTRKL YEKKLLKLREQ
131	TAATACGACTCACTA⁻AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACCGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGGACAACTAC GCCGACCTGAGCGACACCGAGCT GACCACCCTGCTGAGGAGGTACA ACATCCCCCACGGCCCCGTGGTG GGCAGCACCAGGAGGCTGTACGA GAAGAAGATCTTCGAGTACGAGA CCCAGAGGAGGAGGTAAAGACCA GACAGCTGGAGCCTCGGTGGCCA TGCTTCTTGCCCCTTGGGCCTCC CCCCAGCCCCTCCTCCCCTTCCT GCACCCGTACCCCCG⁻GGTCTTT GAATAAAGTCTGAGTGGGCGGCC GGATCAGTACTAGAC⁻GATAGCT TATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACCG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GGACAACUACGCCGACCUGAGCGACACC GAGCUGACCACCCUGCUGAGGAGGUACA ACAUCCCCCACGGCCCCGUGGUGGGCAG CACCAGGAGGCUGUACGAGAAGAAGAUC UUCGAGUACGAGACCCAGAGGAGGAGGU AAAGACCAGACAGCUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGCACCCG UACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGCCGGAUCAGUACUAGACUG AUAGCUUAUUGGGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MASQGTKRSYEQMETDG ERQPAAKRVK⌴DWSHPQ FEKWSHPQFEKWSHPQF EKDNYADLSDTELTTLL RRYNIPHGPVVGSTRRL YEKKIFEYETQRRR
132	TAATACGACTCACTA⁻AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACTACGCCGACCTGAGCGACACC GAGCTGACCACCCTGCTGAGGAG GTACAACATCCCCCACGGCCCCG TGGTGGGCAGCACCAGGAGGCTG TACGAGAAGAAGATC⁻TCGAGTA CGAGACCCAGAGGAGGAGGTGGA GCCACCCCCAGTTCGAGAAGTGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACUACGCCGAC CUGAGCGACACCGAGCUGACCACCCUGC UGAGGAGGUACAACAUCCCCCACGGCCC CGUGGUGGGCAGCACCAGGAGGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGAGGAGGAGGUGGAGCCACCCCCAGUU CGAGAAGUGGAGCCACCCCCAGUUCGAG AAGUGGAGCCACCCCCAGUUCGAGAAGC CCGCCGCCAAGAGGGUGAAGCUGGACGC	MDNYADLSDTELTTLLR RYNIPHGPVVGSTRRLY EKKIFEYETQRRRWSHP QFEKWSHPQFEKWSHPQ FEKPAAKRVK⌴DASQGT KRSYEQMETDGERQSA

	AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGC CCGCCGCCAAGAGGGTGAAGCTG GACGCCAGCCAGGGCACCAAGAG GAGCTACGAGCAGATGGAGACCG ACGGCGAGAGGCAGAGCGCCTAA AGACCAGACAGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCCTTGG GCCTCCCCCCAGCCCCTCCTCCC CTTCCTGCACCCGTACCCCCGTG GTCTTTGAATAAAGTCTGAGTGG GCGGCCGGATCAGTACTAGACTG ATAGCTTATTGGG	CAGCCAGGGCACCAAGAGGAGCUACGAG CAGAUGGAGACCGACGGCGAGAGGCAGA GCGCCUAAAGACCAGACAGCUGGAGCCU CGGUGGCCAUGCUUCUUGCCCCUUGGGC CUCCCCCCAGCCCCUCCUCCCCUUCCUG CACCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGCCGGAUCAGUACU AGACUGAUAGCUUAUUGGGAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	
133	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGGCA ACGAGCTGGCCCTGAAGCTGGCC GGCCTGGACATCTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGTGGAGCCAC CCCCAGTTCGAGAAGGACAACTA CGCCGACCTGAGCGACACCGAGC TGACCACCCTGCTGAGGAGGTAC AACATCCCCCACGGCCCCGTGGT GGGCAGCACCAGGAGGCTGTACG AGAAGAAGATCTTCGAGTACGAG ACCCAGAGGAGGAGGTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC CGGATCAGTACTAGACTGATAGC TTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGGCAACGAGCUGGCC CUGAAGCUGGCCGGCCUGGACAUCUGGA GCCACCCCCAGUUCGAGAAGUGGAGCCA CCCCCAGUUCGAGAAGUGGAGCCACCCC CAGUUCGAGAAGGACAACUACGCCGACC UGAGCGACACCGAGCUGACCACCCUGCU GAGGAGGUACAACAUCCCCCACGGCCCC GUGGUGGGCAGCACCAGGAGGCUGUACG AGAAGAAGAUCUUCGAGUACGAGACCCA GAGGAGGAGGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MGNELALKLAGLDIWSH PQFEKWSHPQFEKWSHP QFEKDNYADLSDTELTT LLRRYNIPHGPVVGSTR RLYEKKIFEYETQRRR
134	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACTACGCCGACCTGAGCGACACC GAGCTGACCACCCTGCTGAGGAG GTACAACATCCCCCACGGCCCCG TGGTGGGCAGCACCAGGAGGCTG TACGAGAAGAAGATCTTCGAGTA CGAGACCCAGAGGAGGAGGTGGA GCCACCCCCAGTTCGAGAAGTGG AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGG GCAACGAGCTGGCCCTGAAGCTG GCCGGCCTGGACATCTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACUACGCCGAC CUGAGCGACACCGAGCUGACCACCCUGC UGAGGAGGUACAACAUCCCCCACGGCCC CGUGGUGGGCAGCACCAGGAGGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGAGGAGGAGGUGGAGCCACCCCCAGUU CGAGAAGUGGAGCCACCCCCAGUUCGAG AAGUGGAGCCACCCCCAGUUCGAGAAGG GCAACGAGCUGGCCCUGAAGCUGGCCGG CCUGGACAUCUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MDNYADLSDTELTTLLR RYNIPHGPVVGSTRRLY EKKIFEYETQRRRWSHP QFEKWSHPQFEKWSHPQ FEKGNELALKLAGLDI

	TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC CGGATCAGTACTAGACTGATAGC TTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	
135	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGTGGA GCCACCCCCAGTTCGAGAAGTGG AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGG ACAACTACGCCGACCTGAGCGAC ACCGAGCTGACCACCCTGCTGAG GAGGTACAACATCCCCCACGGCC CCGTGGTGGGCAGCACCAGGAGG CTGTACGAGAAGAAGATCTTCGA GTACGAGACCCAGAGGAGGAGGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GGACAACUACGCCGACCUGAGCGACACC GAGCUGACCACCCUGCUGAGGAGGUACA ACAUCCCCCACGGCCCCGUGGUCGGCAG CACCAGGAGGCUGUACGAGAAGAAGAUC UUCGAGUACGAGACCCAGAGGAGGAGGU AAAGACCAGACAGCUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGCACCCG UACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGCCGGAUCAGUACUAGACUG AUAGCUUAUUGGGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MWSHPQFEKWSHPQFEK WSHPQFEKDNYADLSDT ELTTLLRRYNIPHGPVV GSTRRLYEKKIFEYETQ RRR
136	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACTACGCCGACCTGAGCGACACC GAGCTGACCACCCTGCTGAGGAG GTACAACATCCCCCACGGCCCCG TGGTGGGCAGCACCAGGAGGCTG TACGAGAAGAAGATCTTCGAGTA CGAGACCCAGAGGAGGAGGTGGA GCCACCCCCAGTTCGAGAAGTGG AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACUACGCCGAC CUGAGCGACACCGAGCUGACCACCCUGC UGAGGAGGUACAACAUCCCCCACGGCCC CGUGGUGGGCAGCACCAGGAGGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGAGGAGGAGGUGGAGCCACCCCCAGUU CGAGAAGUGGAGCCACCCCCAGUUCGAG AAGUGGAGCCACCCCCAGUUCGAGAAGU AAAGACCAGACAGCUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGCACCCG UACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGCCGGAUCAGUACUAGACUG AUAGCUUAUUGGGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MDNYADLSDTELTTLLR RYNIPHGPVVGSTRRLY EKKIFEYETQRRRWSHP QFEKWSHPQFEKWSHPQ FEK

137	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGC ACGCCAGGACCAGCGGCAGGGTG GCCGTGGAGGAGGTGGACGAGGA GGGCAAGTTCGTGAGGCTGAGGA ACAAGAGCAACGAGGACCAGAGC ATGGGCAACTGGCAGATCAAGAG GCAGAACGGCGACGACCCCCTGC TGACCTACAGGTTCCCCCCCAAG TTCACCCTGAAGGCCGGCCAGGT GGTGACCATCTGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCC ACCGACCTGGTGTGGAAGGCCCA GAACACCTGGGGCTGCGGCAACA GCCTGAGGACCGCCCTGATCAAC AGCACCGGCGAGGAGGTGGCCAT GAGGAAGCTGGTGAGGAGCTAAA GACCAGACAGCTGGAGCCTCGGT GGCCATGCTTCTTGCCCCTTGGG CCTCCCCCCAGCCCCTCCTCCCC TTCCTGCACCCGTACCCCCGTGG TCTTTGAATAAAGTCTGAGTGGG CGGCCGGATCAGTACTAGACTGA TAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGCACGCCAGGACC AGCGGCAGGGUGGCCGUGGAGGAGGUGG ACGAGGAGGGCAAGUUCGUGAGGCUGAG GAACAAGAGCAACGAGGACCAGAGCAUG GGCAACUGGCAGAUCAAGAGGCAGAACG GCGACGACCCCCUGCUGACCUACAGGUU CCCCCCCAAGUUCACCCUGAAGGCCGGC CAGGUGGUGACCAUCUGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCCACCGA CCUGGUGUGGAAGGCCCAGAACACCUGG GGCUGCGGCAACAGCCUGAGGACCGCCC UGAUCAACAGCACCGGCGAGGAGGUGGC CAUGAGGAAGCUGGUGAGGAGCUAAAGA CCAGACAGCUGGAGCCUCGGUGGCCAUG CUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGCCGGAUCAGUACUAGACUGAUAGC UUAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAA	MQHARTSGRVAVEEVDE EGKFVRLRNKSNEDQSM GNWQIKRQNGDDPLLTY RFPPKFTLKAGQVVTIW AAGAGATHSPPTDLVWK AQNTWGCGNS_RTALIN STGEEVAMRK_VRS
138	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGC ACGCCAGGACCAGCGGCAGGGTG GCCGTGGAGGAGGTGGACGAGGA GGGCAAGTTCGTGAGGCTGAGGA ACAAGAGCAACGAGGACCAGAGC ATGGGCAACTGGCAGATCAAGAG GCAGAACGGCGACGACCCCCTGC TGACCTACAGGTTCCCCCCCAAG TTCACCCTGAAGGCCGGCCAGGT GGTGACCATCTGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCC ACCGACCTGGTGTGGAAGGCCCA GAACACCTGGGGCTGCGGCAACA GCCTGAGGACCGCCCTGATCAAC AGCACCGGCGAGGAGGTGGCCAT GAGGAAGCTGGTGAGGAGCTGGA GCCACCCCCAGTTCGAGAAGTGG AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAGTACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGCACGCCAGGACC AGCGGCAGGGUGGCCGUGGAGGAGGUGG ACGAGGAGGGCAAGUUCGUGAGGCUGAG GAACAAGAGCAACGAGGACCAGAGCAUG GGCAACUGGCAGAUCAAGAGGCAGAACG GCGACGACCCCCUGCUGACCUACAGGUU CCCCCCCAAGUUCACCCUGAAGGCCGGC CAGGUGGUGACCAUCUGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCCACCGA CCUGGUGUGGAAGGCCCAGAACACCUGG GGCUGCGGCAACAGCCUGAGGACCGCCC UGAUCAACAGCACCGGCGAGGAGGUGGC CAUGAGGAAGCUGGUGAGGAGCUGGAGC CACCCCCAGUUCGAGAAGUGGAGCCACC CCCAGUUCGAGAAGUGGAGCCACCCCCA GUUCGAGAAGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MQHARTSGRVAVEEVDE EGKFVRLRNKSNEDQSM GNWQIKRQNGDDPLLTY RFPPKFTLKAGQVVTIW AAGAGATHSPPTDLVWK AQNTWGCGNS_RTALIN STGEEVAMRK_VRSWSH PQFEKWSHPQFEKWSHP QFEK

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
139	TAATACGACTCACTA⁻AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGCCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGC⁻TCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAG⁻GGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGG⁻GCGACGC CTTCCTGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGTGGAGCCACCCCCA GTTCGAGAAGTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCT⁻CCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCCGGAT CAGTACTAGACTGATAGCTTATT GGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUCCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG AGCCACCCCCAGUUCGAGAAGUGGAGCC ACCCCCAGUUCGAGAAGUGGAGCCACCC CCAGUUCGAGAAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCCGGA UCAGUACUAGACUGAUAGCUUAUUGGGA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEK	
140	TAATACGACTCACTA⁻AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACG ACTACGCCGTGCTGAGCGACACC GAGCTGGCCGCCGTGCTGAGGCA GTACAACATCCCCCACGGCCCCA TCGTGGGCAGCACCAGGAAGCTG TACGAGAAGAAGATC⁻TCGAGTA CGAGACCCAGAGGAGGAGGTAAA GACCAGACAGCTGGAGCCTCGGT GGCCATGCTTCTTGCCCCTTGGG CCTCCCCCCAGCCCCC⁻CCTCCCC TTCCTGCACCCGTACCCCCGTGG TCTTTGAATAAAGTC⁻GAGTGGG CGGCCGGATCAGTAC⁻AGACTGA TAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACGACUACGCCGUG CUGAGCGACACCGAGCUGGCCGCCGUGC UGAGGCAGUACAACAUCCCCCACGGCCC CAUCGUGGGCAGCACCAGGAAGCUGUAC GAGAAGAAGAUCUUCGAGUACGAGACCC AGAGGAGGAGGUAAAGACCAGACAGCUG GAGCCUCGGUGGCCAUGCUUCUUGCCCC UUGGGCCUCCCCCCCAGCCCCUCCUCCCC UUCCUGCACCCGUACCCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGCCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MDDYAVLSDTELAAVLR QYNIPHGPIVGSTRKLY EKKIFEYETQRRR	

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
141	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGGCA ACGAGCTGGCCCTGAAGCTGGCC GGCCTGGACATCTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGTGGAGCCAC CCCCAGTTCGAGAAGCAGCACGC CAGGACCAGCGGCAGGGTGGCCG TGGAGGAGGTGGACGAGGAGGGC AAGTTCGTGAGGCTGAGGAACAA GAGCAACGAGGACCAGAGCATGG GCAACTGGCAGATCAAGAGGCAG AACGGCGACGACCCCCTGCTGAC CTACAGGTTCCCCCCCAAGTTCA CCCTGAAGGCCGGCCAGGTGGTG ACCATCTGGGCCGCCGGCGCCGG CGCCACCCACAGCCCCCCCACCG ACCTGGTGTGGAAGGCCCAGAAC ACCTGGGGCTGCGGCAACAGCCT GAGGACCGCCCTGATCAACAGCA CCGGCGAGGAGGTGGCCATGAGG AAGCTGGTGAGGAGCTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC CGGATCAGTACTAGACTGATAGC TTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGGCAACGAGCUGGCC CUGAAGCUGGCCGGCCUGGACAUCUGGA GCCACCCCCAGUUCGAGAAGUGGAGCCA CCCCCAGUUCGAGAAGUGGAGCCACCCC CAGUUCGAGAAGCAGCACGCCAGGACCA GCGGCAGGGUGGCCGUGGAGGACGUGGA CGAGGAGGGCAAGUUCGUGAGGCUGAGG AACAAGAGCAACGAGGACCAGAGCAUGG GCAACUGGCAGAUCAAGAGGCAGAACGG CGACGACCCCCUGCUGACCUACAGGUUC CCCCCCAAGUUCACCCUGAAGGCCGGCC AGGUGGUGACCAUCUGGGCCGCCGGCGC CGGCGCCACCCACAGCCCCCCCACCGAC CUGGUGUGGAAGGCCCAGAACACCUGGG GCUGCGGCAACAGCCUGAGGACCGCCCU GAUCAACAGCACCGGCGAGGAGGUGGCC AUGAGGAAGCUGGUGAGGAGCUAAAGAC CAGACAGCUGGAGCCUCGGUGGCCAUGC UUCUUGCCCCUUGGGCCUCCCCCCAGCC CCUCCUCCCCUUCCUGCACCCGUACCCC CGUGGUCUUUGAAUAAAGUCUGAGUGGG CGGCCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
142	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGC ACGCCAGGACCAGCGGCAGGGTG GCCGTGGAGGAGGTGGACGAGGA GGGCAAGTTCGTGAGGCTGAGGA ACAAGAGCAACGAGGACCAGAGC ATGGGCAACTGGCAGATCAAGAG GCAGAACGGCGACGACCCCCTGC TGACCTACAGGTTCCCCCCCAAG TTCACCCTGAAGGCCGGCCAGGT GGTGACCATCTGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCC ACCGACCTGGTGTGGAAGGCCCA GAACACCTGGGGCTGCGGCAACA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGCACGCCAGGACC AGCGGCAGGGUGGCCGUGGAGGAGGUGG ACGAGGAGGGCAAGUUCGUGAGGCUGAG GAACAAGAGCAACGAGGACCAGAGCAUG GGCAACUGGCAGAUCAAGAGGCAGAACG GCGACGACCCCCUGCUGACCUACAGGUU CCCCCCCAAGUUCACCCUGAAGGCCGGC CAGGUGGUGACCAUCUGGGCCGCCGGCG CCGGCGCCACCCACAGCCCCCCCACCGA CCUGGUGUGGAAGGCCCAGAACACCUGG GGCUGCCGCAACAGCCUGAGGACCGCCC UGAUCAACAGCACCGGCGAGGAGGUGGC CAUGAGGAAGCUGGUGAGGAGCUGGAGC CACCCCCAGUUCGAGAAGUGGAGCCACC	MQHARTSGRVAVEEVDE EGKFVRLRNKSNEDQSM GNWQIKRQNGDDPLLTY RFPPKFTLKAGQVVTIW AAGAGATHSPPTDLVWK AQNTWGCGNSLRTALIN STGEEVAMRKLVRSWSH PQFEKWSHPQFEKWSHP QFEKGNELALKLAGLDI

	GCCTGAGGACCGCCCTGATCAAC AGCACCGGCGAGGAGGTGGCCAT GAGGAAGCTGGTGAGGAGCTGGA GCCACCCCCAGTTCGAGAAGTGG AGCCACCCCCAGTTCGAGAAGTG GAGCCACCCCCAGTTCGAGAAGG GCAACGAGCTGGCCCTGAAGCTG GCCGGCCTGGACATCTAATAAAG ACCAGACAGCTGGAGCCTCGGTG GCCATGCTTCTTGCCCCTTGGGC CTCCCCCCAGCCCCTCCTCCCCT TCCTGCACCCGTACCCCCGTGGT CTTTGAATAAAGTCTGAGTGGGC GGCCGGATCAGTACTAGACTGAT AGCTTATTGGG	CCCAGUUCGAGAAGUGGAGCCACCCCCA GUUCGAGAAGGGCAACGAGCUGGCCCUG AAGCUGGCCGGCCUGGACAUCUAAUAAA GACCAGACAGCUGGAGCCUCGGUGGCCA UGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCCGGAUCAGUACUAGACUGAUA GCUUAUUGGGAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	
143	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGCCCTGGGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCCGGATCAG TACTAGACTGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGCCCUGGGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCCGGAUCAGUACUAGACUGAU AGCUUAUUGGGAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDALGKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
144	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC CTGAGCGAGCCCATCGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGAACGAG GCCGAGTTCCAGAGGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCCUGAGCGAGCCCAUCG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGAACGAGGCCGAGUUCCAGAGGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG	MTTSQKHRDFLSEPIGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKNEAEFQRWLKDTCG ANAKQSRDCFGCLREWC DAFL

	GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCCGGATCAG TACTAGACTGATAGCTTATTGGG	UGGGCGGCCGGAUCAGUACUAGACUGAU AGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	
145	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACAA CCTCCCAAAAGCACAGGGACTTC CTGAGCGAGCCCATCGGCGAGAA GGACGTGTGCTGGCTGGCCGGGA TTGGTGACGTCCTGAGCAAGAGG CTGGAGGAAAGGGGCATCAACAA GGCCTACATCCTGCTGGGCCAGT TCCTGCTGATGCACAAGAACGAG GCCGAGTTCCAGAGGTGGCTGAT CTGCTGCTTCGGCGCCAATGCCA AGCAGTCCCAGCAGACCAGCCAC TGCCTGAAGGAGTGGTGCGCCTG CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCCGGATCAG TACTAGACTGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACCUCCCAAAAG CACAGGGACUUCCUGAGCGAGCCCAUCG GCGAGAAGGACGUGUGCUGGCUGGCCGG GAUUGGUGACGUCCUGAGCAAGAGGCUG GAGGAAAGGGGCAUCAACAAGGCCUACA UCCUGCUGGGCCAGUUCCUGCUGAUGCA CAAGAACGAGGCCGAGUUCCAGAGGUGG CUGAUCUGCUGCUUCGGCGCCAAUGCCA AGCAGUCCCAGCAGACCAGCCACUGCCU GAAGGAGUGGUGCGCCUGCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCCGGAUCAGUACUAGACUGAU AGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	MTTSQKHRDFLSEPIGE KDVCWLAGIGDVLSKRL EERGINKAYILLGQFLL MHKNEAEFQRWLICCFG ANAKQSQQTSHCLKEWC ACFL
146	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGAGGG ACTTCCTGAGCGAGCCCATCGGC GAGAAGGACGTGTGCTGGCTGGC CGGGATTGGTGACGTCCTGAGCA AGAGGCTGGAGGAAAGGGGCATC AACAAGGCCTACATCCTGCTGGG CCAGTTCCTGCTGATGCACAAGA ACGAGGCCGAGTTCCAGAGGTGG CTGATCTGCTGCTTCGGCGCCAA TGCCAAGCAGTCCCAGCAGACCA GCCACTGCCTGAAGGAGTGGTGC GCCTGCTTCCTGTAAAGACCAGA CAGCTGGAGCCTCGGTGGCCATG CTTCTTGCCCCTTGGGCCTCCCC CCAGCCCCTCCTCCCCTTCCTGC ACCCGTACCCCCGTGGTCTTTGA ATAAAGTCTGAGTGGGCGGCCGG ATCAGTACTAGACTGATAGCTTA TTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAGGGACUUCCUGAGC GAGCCCAUCGGCGAGAAGGACGUGUGCU GGCUGGCCGGGAUUGGUGACGUCCUGAG CAAGAGGCUGGAGGAAAGGGGCAUCAAC AAGGCCUACAUCCUGCUGGGCCAGUUCC UGCUGAUGCACAAGAACGAGGCCGAGUU CCAGAGGUGGCUGAUCUGCUGCUUCGGC GCCAAUGCCAAGCAGUCCCAGCAGACCA GCCACUGCCUGAAGGAGUGGUGCGCCUG CUUCCUGUAAAGACCAGACAGCUGGAGC CUCGGUGGCCAUGCUUCUUGCCCCUUGG GCCUCCCCCCAGCCCCUCCUCCCCUUCC UGCACCCGUACCCCCGUGGUCUUUGAAU AAAGUCUGAGUGGGCGGCCGGAUCAGUA CUAGACUGAUAGCUUAUUGGGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	MRDFLSEPIGEKDVCWL AGIGDVLSKRLEERGIN KAYILLGQFLLMHKNEA EFQRWLICCFGANAKQS QQTSHCLKEWCACFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA	
147	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACA ACACAACCTCCCAAAAGCACAGG GACTTCCTGAGCGAGCCCATCGG CGAGAAGGACGTGTGCTGGGTGG ACGGCATCAGCCACCGAGCTGGCC ATCAACCTGGTGACCAAGGGCAT CAACAAGGCCTACATCCTGCTGG GCCAGTTCCTGCTGATGCACAAG AACGAGGCCGAGTTCCAGAGGTG GCTGATCTGCTGCTTCGGCGCCA ATGCCAAGCAGTCCCAGCAGACC AGCCACTGCCTGAAGGAGTGGTG CGCCTGCTTCCTGTAAAGACCAG ACAGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCCTTGGGCCTCCC CCCAGCCCCTCCTCCCCTTCCTG CACCCGTACCCCCGTGGTCTTTG AATAAAGTCTGAGTGGGCGGCCG GATCAGTACTAGACTGATAGCTT ATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACAACACAACCUCC CAAAAGCACAGGGACUUCCUGAGCGAGC CCAUCGGCGAGAAGGACGUGUGCUGGGU GGACGGCAUCAGCCACGAGCUGGCCAUC AACCUGGUGACCAAGGGCAUCAACAAGG CCUACAUCCUGCUGGGCCAGUUCCUGCU GAUGCACAAGAACGAGGCCGAGUUCCAG AGGUGGCUGAUCUGCUGCUUCGGCGCCA AUGCCAAGCAGUCCCAGCAGACCAGCCA CUGCCUGAAGGAGUGGUGCGCCUGCUUC CUGUAAAGACCAGACAGCUGGAGCCUCG GUGGCCAUGCUUCUUGCCCCUUGGGCCU CCCCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGCCGGAUCAGUACUAG ACUGAUAGCUUAUUGGGAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	MDNTTSQKHRDFLSEPI GEKDVCWVDGISHELAI NLVTKGINKAYILLGQF LLMHKNEAEFQRWLICC FGANAKQSQQTSHCLKE WCACFL
148	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGAGGG ACTTCCTGAGCGAGCCCATCGGC GAGAAGGACGTGTGCTGGGTGGA CGGCATCAGCCACGAGCTGGCCA TCAACCTGGTGACCAAGGGCATC AACAAGGCCTACATCCTGCTGGG CCAGTTCCTGCTGATGCACAAGA ACGAGGCCGAGTTCCAGAGGTGG CTGATCTGCTGCTTCGGCGCCAA TGCCAAGCAGTCCCAGCAGACCA GCCACTGCCTGAAGGAGTGGTGC GCCTGCTTCCTGTAAAGACCAGA CAGCTGGAGCCTCGGTGGCCATG CTTCTTGCCCCTTGGGCCTCCCC CCAGCCCCTCCTCCCCTTCCTGC ACCCGTACCCCCGTGGTCTTTGA ATAAAGTCTGAGTGGGCGGCCGG ATCAGTACTAGACTGATAGCTTA TTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAGGGACUUCCUGAGC GAGCCCAUCGGCGAGAAGGACGUGUGCU GGGUGGACGGCAUCAGCCACGAGCUGGC CAUCAACCUGGUGACCAAGGGCAUCAAC AAGGCCUACAUCCUGCUGGGCCAGUUCC UGCUGAUGCACAAGAACGAGGCCGAGUU CCAGAGGUGGCUGAUCUGCUGCUUCGGC GCCAAUGCCAAGCAGUCCCAGCAGACCA GCCACUGCCUGAAGGAGUGGUGCGCCUG CUUCCUGUAAAGACCAGACAGCUGGAGC CUCGGUGGCCAUGCUUCUUGCCCCUUGG GCCUCCCCCCAGCCCCUCCUCCCCUUCC UGCACCCGUACCCCCGUGGUCUUUGAAU AAAGUCUGAGUGGGCGGCCGGAUCAGUA CUAGACUGAUAGCUUAUUGGGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MRDFLSEPIGEKDVCWV DGISHELAINLVTKGIN KAYILLGQFLMHKNEA EFQRWLICCFGANAKQS QQTSHCLKEWCACFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA	
149	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGACG ACATGAGCCCCAGGCTGAGGGCC TTCCTGAGCGAGCCCATCGGCGA GAAGGACGTGGCCTGGGTGGACG GCATCAGCAGGGAGCTGGCCATC AACCTGGTGACCAAGGGCTTCAA CAAGGCCTACGTGCTGCTGGGCC AGTTCCTGCTGATGCACAAGAAC GAGGCCGAGTTCCAGAGGTGGAT CATCTGCTGCTGCGGCGCCACCG AGTGCGAGGCCAGGCAGAGCAGC ACCTGCCTGAAGGAGTGGTGCAG CTGCTTCCTGTAAAGACCAGACA GCTGGAGCCTCGGTGGCCATGCT TCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCTTCCTGCAC CCGTACCCCCGTGGTCTTTGAAT AAAGTCTGAGTGGGCGGCCGGAT CAGTACTAGACTGATAGCTTATT GGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGACGACAUGAGCCCC AGGCUGAGGGCCUUCCUGAGCGAGCCCA UCGGCGAGAAGGACGUGGCCUGGGUGGA CGGCAUCAGCAGGGAGCUGGCCAUCAAC CUGGUGACCAAGGGCUUCAACAAGGCCU ACGUGCUGCUGGGCCAGUUCCUGCUGAU GCACAAGAACGAGGCCGAGUUCCAGAGG UGGAUCAUCUGCUGCUGCGGCGCCACCG AGUGCGAGGCCAGGCAGAGCAGCACCUG CCUGAAGGAGUGGUGCAGCUGCUUCCUG UAAAGACCAGACAGCUGGAGCCUCGGUG GCCAUGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCU GAGUGGGCGGCCGGAUCAGUACUAGACU GAUAGCUUAUUGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MDDMSPRLRAFLSEPIG EKDVAWVDGISRELAIN LVTKGFNKAYVLLGQFL LMHKNEAEFQRWIICCC GATECEARQSSTCLKEW CSCFL
150	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCTGA GGGCCTTCCTGAGCGAGCCCATC GGCGAGAAGGACGTGGCCTGGGT GGACGGCATCAGCAGGGAGCTGG CCATCAACCTGGTGACCAAGGGC TTCAACAAGGCCTACGTGCTGCT GGGCCAGTTCCTGCTGATGCACA AGAACGAGGCCGAGTTCCAGAGG TGGATCATCTGCTGCTGCGGCGC CACCGAGTGCGAGGCCAGGCAGA GCAGCACCTGCCTGAAGGAGTGG TGCAGCTGCTTCCTGTAAAGACC AGACAGCTGGAGCCTCGGTGGCC ATGCTTCTTGCCCCTTGGGCCTC CCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGC CGGATCAGTACTAGACTGATAGC TTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCUGAGGGCCUUCCUG AGCGAGCCCAUCGGCGAGAAGGACGUGG CCUGGGUGGACGGCAUCAGCAGGGAGCU GGCCAUCAACCUGGUGACCAAGGGCUUC AACAAGGCCUACGUGCUGCUGGGCCAGU UCCUGCUGAUGCACAAGAACGAGGCCGA GUUCCAGAGGUGGAUCAUCUGCUGCUGC GGCGCCACCGAGUGCGAGGCCAGGCAGA GCAGCACCUGCCUGAAGGAGUGGUGCAG CUGCUUCCUGUAAAGACCAGACAGCUGG AGCCUCGGUGGCCAUGCUUCUUGCCCCU UGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUG AAUAAAGUCUGAGUGGGCGGCCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MLRAFLSEPIGEKDVAW VDGISRELAINLVTKGF NKAYVLLGQFLMHKNE AEFQRWIICCCGATECE ARQSSTCLKEWCSCFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
151	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGACCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGAGGTGCTGGGCAAGAAG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGACCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGAGGUGCUGGGCAAGAAGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MTTSQKHRDFVTEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
152	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGACCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGAGGTGCTGGGCAAGAAG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACCTGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCAACGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTAAAGACCAGACAGCT GGAGCCTCGGTGGCCATGCTTCT TGCCCCTTGGGCCTCCCCCCAGC CCCTCCTCCCCTTCCTGCACCCG TACCCCCGTGGTCTTTGAATAAA GTCTGAGTGGGCGGCAAGCTTTT CTTCTCTCCCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGACCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGAGGUGCUGGGCAAGAAGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACCUGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCAACGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MTTSQKHRDFVTEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

		AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
153	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGCGACCT GGAGACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGCCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYY

	CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA A	
154	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGTGGAGCCACCCCCAGTT CGAGAAGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGAACTTCCAGGGCCT GGTGCTGACCGACAACTGCAAGA ACCAGTGGGTGGTGGGCCCCCTG ATCGGCAAGGGCGGCTTCGGCAG CATCTACACCACCAACGACAACA ACTACGTGGTGAAGATCGAGCCC AAGGCCAACGGCAGCCTGTTCAC CGAGCAGGCCTTCTACACCAGGG TGCTGAAGCCCAGCGTGATCGAG GAGTGGAAGAAGAGCCACAACAT CAAGCACGTGGGCCTGATCACCT GCAAGGCCTTCGGCCTGTACAAG AGCATCAACGTGGAGTACAGGTT CCTGGTGATCAACAGGCTGGGCG CCGACCTGGACGCCGTGATCAGG GCCAACAACAACAGGCTGCCCAA GAGGAGCGTGATGCTGATCGGCA TCGAGATCCTGAACACCATCCAG TTCATGCACGAGCAGGGCTACAG CCACGGCGACATCAAGGCCAGCA ACATCGTGCTGGACCAGATCGAC AAGAACAAGCTGTACCTGGTGGA CTACGGCCTGGTGAGCAAGTTCA TGAGCAACGGCGAGCACGTGCCC TTCATCAGGAACCCCAACAAGAT GGACAACGGCACCCTGGAGTTCA CCCCCATCGACAGCCACAAGGGC TACGTGGTGAGCAGGAGGGGCGA CCTGGAGACCCTGGGCTACTGCA TGATCAGGTGGCTGGGCGGCATC CTGCCCTGGACCAAGATCAGCGA GACCAAGAACTGCGCCCTGGTGA GCGCCACCAAGCAGAAGTACGTG AACAACACCGCCACCCTGCTGAT GACCAGCCTGCAGTACGCCCCCA GGGAGCTGCTGCAGTACATCACC ATGGTGAACAGCCTGACCTACTT CGAGGAGCCCAACTACGACAAGT	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGUGGAGCCACCCCCAGUUCGA GAAGUGGAGCCACCCCCAGUUCGAGAAG UGGAGCCACCCCCAGUUCGAGAAGAACU UCCAGGGCCUGGUGCUGACCGACAACUG CAAGAACCAGUGGGUGGUGGGCCCCCUG AUCGGCAAGGGCGGCUUCGGCAGCAUCU ACACCACCAACGACAACAACUACGUGGU GAAGAUCGAGCCCAAGGCCAACGGCAGC CUGUUCACCGAGCAGGCCUUCUACACCA GGGUGCUGAAGCCCAGCGUGAUCGAGGA GUGGAAGAAGAGCCACAACAUCAAGCAC GUGGGCCUGAUCACCUGCAAGGCCUUCG GCCUGUACAAGAGCAUCAACGUGGAGUA CAGGUUCCUGGUGAUCAACAGGCUGGGC GCCGACCUGGACGCCGUGAUCAGGGCCA ACAACAACAGGCUGCCCAAGAGGAGCGU GAUGCUGAUCGGCAUCGAGAUCCUGAAC ACCAUCCAGUUCAUGCACGAGCAGGGCU ACAGCCACGGCGACAUCAAGGCCAGCAA CAUCGUGCUGGACCAGAUCGACAAGAAC AAGCUGUACCUGGUGGACUACGGCCUGG UGAGCAAGUUCAUGAGCAACGGCGAGCA CGUGCCCUUCAUCAGGAACCCCAACAAG AUGGACAACGGCACCCUGGAGUUCACCC CCAUCGACAGCCACAAGGGCUACGUGGU GAGCAGGAGGGGCGACCUGGAGACCCUG GGCUACUGCAUGAUCAGGUGGCUGGGCG GCAUCCUGCCCUGGACCAAGAUCAGCGA GACCAAGAACUGCGCCCUGGUGAGCGCC ACCAAGCAGAAGUACGUGAACAACACCG CCACCCUGCUGAUGACCAGCCUGCAGUA CGCCCCCAGGGAGCUGCUGCAGUACAUC ACCAUGGUGAACAGCCUGACCUACUUCG AGGAGCCCAACUACGACAAGUUCAGGCA CAUCCUGAUGCAGGGCGUGUACUACUAA GAUCUGGUUACCACUAAACCAGCCUCAA GAACACCCGAAUGGAGUCUCUAAGCUAC AUAAUACCAACUUACACUUACAAAAUGU UGUCCCCCAAAAUGUAGCCAUUCGUAUC UGCUCCUAAUAAAAAGAAAGUUUCUUCA CAUUUCUUCUCUCCCUUGCACCUCGGAU CAGUACUAGACUGAUAGCUUAUUGGGAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	MWSHPQFEKWSHPQFEK WSHPQFEKNFQGLVLTD NCKNQWVVGPLIGKGGF GSIYTTNDNNYVVKIEP KANGSLFTEQAFYTRVL KPSVIEEWKKSHNIKHV GLITCKAFGLYKSINVE YRFLVINRLGADLDAVI RANNNRLPKRSVMLIGI EILNTIQFMHEQGYSHG DIKASNIVLDQIDKNKL YLVDYGLVSKFMSNGEH VPFIRNPNKMDNGTLEF TPIDSHKGYVVSRRGDL ETLGYCMIRWLGGILPW TKISETKNCALVSATKQ KYVNNTATLLMTSLQYA PRELLQYITMVNSLTYF EEPNYDKFRHILMQGVY Y

	TCAGGCACATCCTGATGCAGGGC GTGTACTACTAAGATCTGGTTAC CACTAAACCAGCCTCAAGAACAC CCGAATGGAGTCTCTAAGCTACA TAATACCAACTTACACTTACAAA ATGTTGTCCCCCAAAATGTAGCC ATTCGTATCTGCTCCTAATAAAA AGAAAGTTTCTTCACATTTCTTC TCTCCCTTGCACCTCGGATCAGT ACTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AA	
155	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGAGTC TATTATTGGTCACATCCACAATT	GGGAUAAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGA GUCUAUUAUUGGUCACAUCCACAAUUUG AAAAGUGGUCUCAUCCACAAUUUGAAAA GUGGUCUCAUCCUCAAUUUGAAAAGUGA GAUCUGGUUACCACUAAACCAGCCUCAA GAACACCCGAAUGGAGUCUCUAAGCUAC AUAAUACCAACUUACACUUACAAAAUGU UGUCCCCCAAAAUGUAGCCAUUCGUAUC UGCUCCUAAUAAAAAGAAAGUUUCUUCA CAUUUCUUCUCUCCCUUGCACCUCGGAU CAGUACUAGACUGAUAGCUUAUUGGGAA	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYWSHPQF EKWSHPQFEKWSHPQFE K

	TGAAAAGTGGTCTCAUCCACAAT	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGAAAAGTGGTCTCATCCTCAA	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTTGAAAAGTGAGATCTGGTTAC	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CACTAAACCAGCCTCAAGAACAC	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGAATGGAGTCTCTAAGCTACA	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAATACCAACTTACACTTACAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTTGTCCCCCAAAATGTAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATTCGTATCTGCTCCUAATAAAA	AA	
	AGAAAGTTTCTTCACATTTCTTC		
	TCTCCCTTGCACCTCGGATCAGT		
	ACTAGACTGATAGCTUATTGGG		
156	TAATACGACTCACTAUAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MDYKDHDGDYKDHDIDY
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	KDDDDKNFQGLVLTDNC
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	KNQWVVGPLIGKGGFGS
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	IYTTNDNNYVVKIEPKA
	AATTTAAATCATTTCUTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	NGSLFTEQAFYTRVLKP
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGACUACAAGGACCACGACGG	SVIEEWKKSHNIKHVGL
	TCACCATTTACGAACGATAGGCC	CGACUACAAGGACCACGACAUCGACUAC	ITCKAFGLYKSINVEYR
	ACCATGGACTACAAGGACCACGA	AAGGACGACGACGACAAGAACUUCCAGG	FLVINRLGADLDAVIRA
	CGGCGACTACAAGGACCACGACA	GCCUGGUGCUGACCGACAACUGCAAGAA	NNNRLPKRSVMLIGIEI
	TCGACTACAAGGACGACGACGAC	CCAGUGGGUGGUGGGCCCCCUGAUCGGC	LNTIQFMHEQGYSHGDI
	AAGAACTTCCAGGGCCTGGTGCT	AAGGGCGGCUUCGGCAGCAUCUACACCA	KASNIVLDQIDKNKLYL
	GACCGACAACTGCAAGAACCAGT	CCAACGACAACAACUACGUGGUGAAGAU	VDYGLVSKFMSNGEHVP
	GGGTGGTGGGCCCCCUGATCGGC	CGAGCCCAAGGCCAACGGCAGCCUGUUC	FIRNPNKMDNGTLEFTP
	AAGGGCGGCTTCGGCAGCATCTA	ACCGAGCAGGCCUUCUACACCAGGGUGC	IDSHKGYVVSRRGDLET
	CACCACCAACGACAACAACTACG	UGAAGCCCAGCGUGAUCGAGGAGUGGAA	LGYCMIRWLGGILPWTK
	TGGTGAAGATCGAGCCCAAGGCC	GAAGAGCCACAACAUCAAGCACGUGGGC	ISETKNCALVSATKQKY
	AACGGCAGCCTGTTCACCGAGCA	CUGAUCACCUGCAAGGCCUUCGGCCUGU	VNNTATLLMTSLQYAPR
	GGCCTTCTACACCAGGGTGCTGA	ACAAGAGCAUCAACGUGGAGUACAGGUU	ELLQYITMVNSLTYFEE
	AGCCCAGCGTGATCGAGGAGTGG	CCUGGUGAUCAACAGGCUGGGCGCCGAC	PNYDKFRHILMQGVYY
	AAGAAGAGCCACAACATCAAGCA	CUGGACGCCGUGAUCAGGGCCAACAACA	
	CGTGGGCCTGATCACCTGCAAGG	ACAGGCUGCCCAAGAGGAGCGUGAUGCU	
	CCTTCGGCCTGTACAAGAGCATC	GAUCGGCAUCGAGAUCCUGAACACCAUC	
	AACGTGGAGTACAGGTTCCTGGT	CAGUUCAUGCACGAGCAGGGCUACAGCC	
	GATCAACAGGCTGGGCGCCGACC	ACGGCGACAUCAAGGCCAGCAACAUCGU	
	TGGACGCCGTGATCAGGGCCAAC	GCUGGACCAGAUCGACAAGAACAAGCUG	
	AACAACAGGCTGCCCAAGAGGAG	UACCUGGUGGACUACGGCCUGGUGAGCA	
	CGTGATGCTGATCGGCATCGAGA	AGUUCAUGAGCAACGGCGAGCACGUGCC	
	TCCTGAACACCATCCAGTTCATG	CUUCAUCAGGAACCCCAACAAGAUGGAC	
	CACGAGCAGGGCTACAGCCACGG	AACGGCACCCUGGAGUUCACCCCCAUCG	
	CGACATCAAGGCCAGCAACATCG	ACAGCCACAAGGGCUACGUGGUGAGCAG	
	TGCTGGACCAGATCGACAAGAAC	GAGGGGCGACCUGGAGACCCUGGGCUAC	
	AAGCTGTACCTGGTGGACTACGG	UGCAUGAUCAGGUGGCUGGGCGGCAUCC	
	CCTGGTGAGCAAGTTCATGAGCA	UGCCCUGGACCAAGAUCAGCGAGACCAA	
	ACGGCGAGCACGTGCCCTTCATC	GAACUGCGCCCUGGUGAGCGCCACCAAG	
	AGGAACCCCAACAAGATGGACAA	CAGAAGUACGUGAACAACACCGCCACCC	
	CGGCACCCTGGAGTTCACCCCCA	UGCUGAUGACCAGCCUGCAGUACGCCCC	
	TCGACAGCCACAAGGGCTACGTG	CAGGGAGCUGCUGCAGUACAUCACCAUG	
	GTGAGCAGGAGGGGCGACCTGGA	GUGAACAGCCUGACCUACUUCGAGGAGC	
	GACCCTGGGCTACTGCATGATCA	CCAACUACGACAAGUUCAGGCACAUCCU	
	GGTGGCTGGGCGGCAUCCTGCCC	GAUGCAGGGCGUGUACUACUAAGAUCUG	
	TGGACCAAGATCAGCGAGACCAA	GUUACCACUAAACCAGCCUCAAGAACAC	
	GAACTGCGCCCTGGTGAGCGCCA	CCGAAUGGAGUCUCUAAGCUACAUAAUA	
	CCAAGCAGAAGTACGUGAACAAC	CCAACUUACACUUACAAAAUGUUGUCCC	
	ACCGCCACCCTGCTGATGACCAG	CCAAAAUGUAGCCAUUCGUAUCUGCUCC	
	CCTGCAGTACGCCCCCAGGGAGC	UAAUAAAAAGAAAGUUUCUUCACAUUUC	
	TGCTGCAGTACATCACCATGGTG	UUCUCUCCCUUGCACCUCGGAUCAGUAC	

	AACAGCCTGACCTACTTCGAGGA GCCCAACTACGACAAGTTCAGGC ACATCCTGATGCAGGGCGTGTAC TACTAAGATCTGGTTACCACTAA ACCAGCCTCAAGAACACCCGAAT GGAGTCTCTAAGCTACATAATAC CAACTTACACTTACAAAATGTTG TCCCCCAAAATGTAGCCATTCGT ATCTGCTCCTAATAAAAAGAAAG TTTCTTCACATTTCTTCTCTCCC TTGCACCTCGGATCAGTACTAGA CTGATAGCTTATTGGG	UAGACUGAUAGCUUAUUGGGAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	
157	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA	GGGAUAAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACGACUACAAGGACCACGACG GCGACUACAAGGACCACGACAUCGACUA CAAGGACGACGACGACAAGUAACUAGUG ACUGACUAGGAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAAUACCAACUUACACUU ACAAAAUGUUGUCCCCCAAAAUGUAGCC AUUCGUAUCUGCUCCUAAUAAAAAGAAA	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYDYKDHD GDYKDHDIDYKDDDDK

	GGCACATCCTGATGCAGGGCGTG TACTACGACTACAAGGACCACGA CGGCGACTACAAGGACCACGACA TCGACTACAAGGACGACGACGAC AAGTAACTAGTGACTGACTAGGA TCTGGTTACCACTAAACCAGCCT CAAGAACACCCGAATGGAGTCTC TAAGCTACATAATACCAACTTAC ACTTACAAAATGTTGTCCCCCAA AATGTAGCCATTCGTATCTGCTC CTAATAAAAAGAAAGTTTCTTCA CATTTCTTCTCTCCCTTGCACCT CGGATCAGTACTAGACTGATAGC TTATTGGG	GUUUCUUCACAUUUCUUCUCUCCCUUGC ACCUCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
158	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CGAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDE FRHILMQGVYY

	GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACGAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
159	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CGCCUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDE FRHLLMQGVYY

	GGAGCCCAACTACGACGCCTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
160	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGGGTGAAGGCCGC CCAGGCCGGCAGGCAGAGCAGCG CCAAGAGGCACCTGGCCGAGCAG TTCGCCGTGGGCGAGATCATCAC CGACATGGCCAAGAAGGAGTGGA AGGTGGGCCTGCCCATCGGCCAG GGCGGCTTCGGCTGCATCTACCT GGCCGACATGAACAGCAGCGAGA GCGTGGGCAGCGACGCCCCCTGC GTGGTGAAGGTGGAGCCCAGCGA CAACGGCCCCCTGTTCACCGAGC TGAAGTTCTACCAGAGGGCCGCC AAGCCCGAGCAGATCCAGAAGTG GATCAGGACCAGGAAGCTGAAGT ACCTGGGCGTGCCCAAGTACTGG GGCAGCGGCCTGCACGACAAGAA CGGCAAGAGCTACAGGTTCATGA TCATGGACAGGTTCGGCAGCGAC CTGCAGAAGATCTACGAGGCCAA CGCCAAGAGGTTCAGCAGGAAGA CCGTGCTGCAGCTGAGCCTGAGG ATCCTGGACATCCTGGAGTACAT CCACGAGCACGAGTACGTGCACG GCGACATCAAGGCCAGCAACCTG CTGCTGAACTACAAGAACCCCGA CCAGGTGTACCTGGTGGACTACG GCCTGGCCTACAGGTACTGCCCC GAGGGCGTGCACAAGGAGTACAA GGAGGACCCCAAGAGGTGCCACG ACGGCACCATCGAGTTCACCAGC ATCGACGCCCACAACGGCGTGGC CCCCAGCAGGAGGGGCGACCTGG AGATCCTGGGCTACTGCATGATC CAGTGGCTGACCGGCCACCTGCC CTGGGAGGACAACCTGAAGGACC CCAAGTACGTGAGGGACAGCAAG ATCAGGTACAGGGAGAACATCGC CAGCCTGATGGACAAGTGCTTCC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGGGUGAAGGCCGCCCA GGCCGGCAGGCAGAGCAGCGCCAAGAGG CACCUGGCCGAGCAGUUCGCCGUGGGCG AGAUCAUCACCGACAUGGCCAAGAAGGA GUGGAAGGUGGGCCUGCCCAUCGGCCAG GGCGGCUUCGGCUGCAUCUACCUGGCCG ACAUGAACAGCAGCGAGAGCGUGGGCAG CGACGCCCCCUGCGUGGUGAAGGUGGAG CCCAGCGACAACGGCCCCCUGUUCACCG AGCUGAAGUUCUACCAGAGGGCCGCCAA GCCCGAGCAGAUCCAGAAGUGGAUCAGG ACCAGGAAGCUGAAGUACCUGGGCGUGC CCAAGUACUGGGGCAGCGGCCUGCACGA CAAGAACGGCAAGAGCUACAGGUUCAUG AUCAUGGACAGGUUCGGCAGCGACCUGC AGAAGAUCUACGAGGCCAACGCCAAGAG GUUCAGCAGGAAGACCGUGCUGCAGCUG AGCCUGAGGAUCCUGGACAUCCUGGAGU ACAUCCACGAGCACGAGUACGUGCACGG CGACAUCAAGGCCAGCAACCUGCUGCUG AACUACAAGAACCCCGACCAGGUGUACC UGGUGGACUACGGCCUGGCCUACAGGUA CUGCCCCGAGGGCGUGCACAAGGAGUAC AAGGAGGACCCCAAGAGGUGCCACGACG GCACCAUCGAGUUCACCAGCAUCGACGC CCACAACGGCGUGGCCCCCAGCAGGAGG GGCGACCUGGAGAUCCUGGGCUACUGCA UGAUCCAGUGGCUGACCGGCCACCUGCC CUGGGAGGACAACCUGAAGGACCCCAAG UACGUGAGGGACAGCAAGAUCAGGUACA GGGAGAACAUCGCCAGCCUGAUGGACAA GUGCUUCCCCGAGAAGAACAAGCCCGGC GAGAUCGCCAAGUACAUGGAGACCGUGA AGCUGCUGGACUACACCGAGAAGCCCCU GUACGAGAACCUGAGGGACAUCCUGCUG CAGGGCCUGAAGGCCAUCGGCAGCAAGG ACGACGGCAAGCUGGACCUGAGCGUGGU GGAGAACGGCGGCCUGAAGGCCAAGACC AUCACCAAGAAGAGGAAGAAGGAGAUCG AGGAGAGCAAGGAGCCCGGCGUGGAGGA	MPRVKAAQAGRQSSAKR HLAEQFAVGEIITDMAK KEWKVGLPIGQGGFGCI YLADMNSSESVGSDAPC VVKVEPSDNGPLFTELK FYQRAAKPEQIQKWIRT RKLKYLGVPKYWGSGLH DKNGKSYRFMIMDRFGS DLQKIYEANAKRFSRKT VLQLSLRILDILEYIHE HEYVHGDIKASNLLLNY KNPDQVYLVDYGLAYRY CPEGVHKEYKEDPKRCH DGTIEFTSIDAHNGVAP SRRGDLEILGYCMIQWL TGHLPWEDNLKDPKYVR DSKIRYRENIASLMDKC FPEKNKPGEIAKYMETV KLLDYTEKPLYENLRDI LLQGLKAIGSKDDGKLD LSVVENGGLKAKTITKK RKKEIEESKEPGVEDTE WSNTQTEEAIQTRSRTR KRVQK

	CCGAGAAGAACAAGCCCGGCGAG ATCGCCAAGTACATGGAGACCGT GAAGCTGCTGGACTACACCGAGA AGCCCCTGTACGAGAACCTGAGG GACATCCTGCTGCAGGGCCTGAA GGCCATCGGCAGCAAGGACGACG GCAAGCTGGACCTGAGCGTGGTG GAGAACGGCGGCCTGAAGGCCAA GACCATCACCAAGAAGAGGAAGA AGGAGATCGAGGAGAGCAAGGAG CCCGGCGTGGAGGACACCGAGTG GAGCAACACCCAGACCGAGGAGG CCATCCAGACCAGGAGCAGGACC AGGAAGAGGGTGCAGAAGTAAAT CTGGTTACCACTAAACCAGCCTC AAGAACACCCGAATGGAGTCTCT AAGCTACATAATACCAACTTACA CTTACAAAATGTTGTCCCCCAAA ATGTAGCCATTCGTATCTGCTCC TAATAAAAAGAAAGTTTCTTCAC ATTTCTTCTCTCCCTTGCACCTC GGATCAGTACTAGACTGATAGCT TATTGGG	CACCGAGUGGAGCAACACCCAGACCGAG GAGGCCAUCCAGACCAGGAGCAGGACCA GGAAGAGGGUGCAGAAGUAAAUCUGGUU ACCACUAAACCAGCCUCAAGAACACCCG AAUGGAGUCUCUAAGCUACAUAAUACCA ACUUACACUUACAAAAAUGUUGUCCCCCA AAAUGUAGCCAUUCGUAUCUGCUCCUAA UAAAAAGAAAGUUUCUUCACAUUUCUUC UCUCCCUUGCACCUCGGAUCAGUACUAG ACUGAUAGCUUAUUGGGAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
161	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGGGTGAAGGCCGC CCAGGCCGGCAGGCAGAGCAGCG CCAAGAGGCACCTGGCCGAGCAG TTCGCCGTGGGCGAGATCATCAC CGACATGGCCAAGAAGGAGTGGA AGGTGGGCCTGCCCATCGGCCAG GGCGGCTTCGGCTGCATCTACCT GGCCGACATGAACAGCAGCGAGA GCGTGGGCAGCGACGCCCCCTGC GTGGTGAAGGTGGAGCCCAGCGA CAACGGCCCCCTGTTCACCGAGC TGAAGTTCTACCAGAGGGCCGCC AAGCCCGAGCAGATCCAGAAGTG GATCAGGACCAGGAAGCTGAAGT ACCTGGGCGTGCCCAAGTACTGG GGCAGCGGCCTGCACGACAAGAA CGGCAAGAGCTACAGGTTCATGA TCATGGACAGGTTCGGCAGCGAC CTGCAGAAGATCTACGAGGCCAA CGCCAAGAGGTTCAGCAGGAAGA CCGTGCTGCAGCTGAGCCTGAGG ATCCTGGACATCCTGGAGTACAT CCACGAGCACGAGTACGTGCACG GCGACATCAAGGCCAGCAACCTG CTGCTGAACTACAAGAACCCCGA CCAGGTGTACCTGGTGGACTACG GCCTGGCCTACAGGTACTGCCCC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGGGUGAAGGCCGCCCA GGCCGGCAGGCAGAGCAGCGCCAAGAGG CACCUGGCCGAGCAGUUCGCCGUGGGCG AGAUCAUCACCGACAUGGCCAAGAAGGA GUGGAAGGUGGGCCUGCCCAUCGGCCAG GGCGGCUUCGGCUGCAUCUACCUGGCCG ACAUGAACAGCAGCGAGAGCGUGGGCAG CGACGCCCCCUGCGUGGUGAAGGUGGAG CCCAGCGACAACGGCCCCCUGUUCACCG AGCUGAAGUUCUACCAGAGGGCCGCCAA GCCCGAGCAGAUCCAGAAGUGGAUCAGG ACCAGGAAGCUGAAGUACCUGGGCGUGC CCAAGUACUGGGGCAGCGGCCUGCACGA CAAGAACGGCAAGAGCUACAGGUUCAUG AUCAUGGACAGGUUCGGCAGCGACCUGC AGAAGAUCUACGAGGCCAACGCCAAGAG GUUCAGCAGGAAGACCGUGCUGCAGCUG AGCCUGAGGAUCCUGGACAUCCUGGAGU ACAUCCACGAGCACGAGUACGUGCACGG CGACAUCAAGGCCAGCAACCUGCUGCUG AACUACAAGAACCCCGACCAGGUGUACC UGGUGGACUACGGCCUGGCCUACAGGUA CUGCCCCGAGGGCGUGCACAAGGAGUAC AAGGAGGACCCCAAGAGGUGCCACGACG GCACCAUCGAGUUCACCAGCAUCGACGC CCACAACGGCGUGGCCCCCAGCAGGAGG GGCGACCUGGAGAUCCUGGGCUACUGCA UGAUCCAGUGGCUGACCGGCCACCUGCC CUGGGAGGACAACCUGAAGGACCCCAAG	MPRVKAAQAGRQSSAKR HLAEQFAVGEIITDMAK KEWKVGLPIGQGGFGCI YLADMNSSESVGSDAPC VVKVEPSDNGPLFTELK FYQRAAKPEQIQKWIRT RKLKYLGVPKYWGSGLH DKNGKSYRFMIMDRFGS DLQKIYEANAKRFSRKT VLQLSLRILDILEYIHE HEYVHGDIKASNLLLNY KNPDQVYLVDYGLAYRY CPEGVHKEYKEDPKRCH DGTIEFTSIDAHNGVAP SRRGDLEILGYCMIQWL TGHLPWEDNLKDPKYVR DSKIRYRENIASLMDKC FPEKNKPGEIAKYMETV KLLDYTEKPLYENLRDI LLQG

	GAGGGCGTGCACAAGGAGTACAA GGAGGACCCCAAGAGGTGCCACG ACGGCACCATCGAGTTCACCAGC ATCGACGCCCACAACGGCGTGGC CCCCAGCAGGAGGGGCGACCTGG AGATCCTGGGCTACTGCATGATC CAGTGGCTGACCGGCCACCTGCC CTGGGAGGACAACCTGAAGGACC CCAAGTACGTGAGGGACAGCAAG ATCAGGTACAGGGAGAACATCGC CAGCCTGATGGACAAGTGCTTCC CCGAGAAGAACAAGCCCGGCGAG ATCGCCAAGTACATGGAGACCGT GAAGCTGCTGGACTACACCGAGA AGCCCCTGTACGAGAACCTGAGG GACATCCTGCTGCAGGGCTAAAT CTGGTTACCACTAAACCAGCCTC AAGAACACCCGAATGGAGTCTCT AAGCTACATAATACCAACTTACA CTTACAAAATGTTGTCCCCCAAA ATGTAGCCATTCGTATCTGCTCC TAATAAAAAGAAAGTTTCTTCAC ATTTCTTCTCTCCCTTGCACCTC GGATCAGTACTAGACTGATAGCT TATTGGG	UACGUGAGGGACAGCAAGAUCAGGUACA GGGAGAACAUCGCCAGCCUGAUGGACAA GUGCUUCCCCGAGAAGAACAAGCCCGGC GAGAUCGCCAAGUACAUGGAGACCGUGA AGCUGCUGGACUACACCGAGAAGCCCCU GUACGAGAACCUGAGGGACAUCCUGCUG CAGGGCUAAAUCUGGUUACCACUAAACC AGCCUCAAGAACACCCGAAUGGAGUCUC UAAGCUACAUAAUACCAACUUACACUUA CAAAAUGUUGUCCCCCAAAAUGUAGCCA UUCGUAUCUGCUCCUAAUAAAAAGAAAG UUUCUUCACAUUUCUUCUCUCCCUUGCA CCUCGGAUCAGUACUAGACUGAUAGCUU AUUGGGAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
162	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGCCGTGGGCGAGATCAT CACCGACATGGCCAAGAAGGAGT GGAAGGTGGGCCTGCCCATCGGC CAGGGCGGCTTCGGCTGCATCTA CCTGGCCGACATGAACAGCAGCG AGAGCGTGGGCAGCGACGCCCCC TGCGTGGTGAAGGTGGAGCCCAG CGACAACGGCCCCCTGTTCACCG AGCTGAAGTTCTACCAGAGGGCC GCCAAGCCCGAGCAGATCCAGAA GTGGATCAGGACCAGGAAGCTGA AGTACCTGGGCGTGCCCAAGTAC TGGGGCAGCGGCCTGCACGACAA GAACGGCAAGAGCTACAGGTTCA TGATCATGGACAGGTTCGGCAGC GACCTGCAGAAGATCTACGAGGC CAACGCCAAGAGGTTCAGCAGGA AGACCGTGCTGCAGCTGAGCCTG AGGATCCTGGACATCCTGGAGTA CATCCACGAGCACGAGTACGTGC ACGGCGACATCAAGGCCAGCAAC CTGCTGCTGAACTACAAGAACCC CGACCAGGTGTACCTGGTGGACT ACGGCCTGGCCTACAGGTACTGC CCCGAGGGCGTGCACAAGGAGTA	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGCCGUGGGCGAGAUCAUCAC CGACAUGGCCAAGAAGGAGUGGAAGGUG GGCCUGCCCAUCGGCCAGGGCGGCUUCG GCUGCAUCUACCUGGCCGACAUGAACAG CAGCGAGAGCGUGGGCAGCGACGCCCCC UGCGUGGUGAAGGUGGAGCCCAGCGACA ACGGCCCCCUGUUCACCGAGCUGAAGUU CUACCAGAGGGCCGCCAAGCCCGAGCAG AUCCAGAAGUGGAUCAGGACCAGGAAGC UGAAGUACCUGGGCGUGCCCAAGUACUG GGGCAGCGGCCUGCACGACAAGAACGGC AAGAGCUACAGGUUCAUGAUCAUGGACA GGUUCGGCAGCGACCUGCAGAAGAUCUA CGAGGCCAACGCCAAGAGGUUCAGCAGG AAGACCGUGCUGCAGCUGAGCCUGAGGA UCCUGGACAUCCUGGAGUACAUCCACGA GCACGAGUACGUGCACGGCGACAUCAAG GCCAGCAACCUGCUGCUGAACUACAAGA ACCCCGACCAGGUGUACCUGGUGGACUA CGGCCUGGCCUACAGGUACUGCCCCGAG GGCGUGCACAAGGAGUACAAGGAGGACC CCAAGAGGUGCCACGACGGCACCAUCGA GUUCACCAGCAUCGACGCCCACAACGGC GUGGCCCCCAGCAGGAGGGGCGACCUGG AGAUCCUGGGCUACUGCAUGAUCCAGUG GCUGACCGGCCACCUGCCCUGGGAGGAC AACCUGAAGGACCCCAAGUACGUGAGGG	MAVGEIITDMAKKEWKV GLPIGQGGFGCIYLADM NSSESVGSDAPCVVKVE PSDNGPLFTELKFYQRA AKPEQIQKWIRTRKLKY LGVPKYWGSGLHDKNGK SYRFMIMDRFGSDLQKI YEANAKRFSRKTVLQLS LRILDILEYIHEHEYVH GDIKASNLLLNYKNPDQ VYLVDYGLAYRYCPEGV HKEYKEDPKRCHDGTIE FTSIDAHNGVAPSRRGD LEILGYCMIQWLTGHLP WEDNLKDPKYVRDSKIR YRENIASLMDKCFPEKN KPGEIAKYMETVKLLDY TEKPLYENLRDILLQGL KAIGSKDDGKLDLSVVE NGGLKAKTITKKRKKEI EESKEPGVEDTEWSNTQ TEEAIQTRSRTRKRVQK

	CAAGGAGGACCCCAAGAGGTGCC	ACAGCAAGAUCAGGUACAGGGAGAACAU	
	ACGACGGCACCATCGAGTTCACC	CGCCAGCCUGAUGGACAAGUGCUUCCCC	
	AGCATCGACGCCCACAACGGCGT	GAGAAGAACAAGCCCGGCGAGAUCGCCA	
	GGCCCCCAGCAGGAGGGGCGACC	AGUACAUGGAGACCGUGAAGCUGCUGGA	
	TGGAGATCCTGGGCTACTGCATG	CUACACCGAGAAGCCCCUGUACGAGAAC	
	ATCCAGTGGCTGACCGGCCACCT	CUGAGGGACAUCCUGCUGCAGGGCCUGA	
	GCCCTGGGAGGACAACCTGAAGG	AGGCCAUCGGCAGCAAGGACGACGGCAA	
	ACCCCAAGTACGTGAGGGACAGC	GCUGGACCUGAGCGUGGUGGAGAACGGC	
	AAGATCAGGTACAGGGAGAACAT	GGCCUGAAGGCCAAGACCAUCACCAAGA	
	CGCCAGCCTGATGGACAAGTGCT	AGAGGAAGAAGGAGAUCGAGGAGAGCAA	
	TCCCCGAGAAGAACAAGCCCGGC	GGAGCCCGGCGUGGAGGACACCGAGUGG	
	GAGATCGCCAAGTACATGGAGAC	AGCAACACCCAGACCGAGGAGGCCAUCC	
	CGTGAAGCTGCTGGACTACACCG	AGACCAGGAGCAGGACCAGGAAGAGGGU	
	AGAAGCCCCTGTACGAGAACCTG	GCAGAAGUAAAUCUGGUUACCACUAAAC	
	AGGGACATCCTGCTGCAGGGCCT	CAGCCUCAAGAACACCCGAAUGGAGUCU	
	GAAGGCCATCGGCAGCAAGGACG	CUAAGCUACAUAAUACCAACUUACACUU	
	ACGGCAAGCTGGACCTGAGCGTG	ACAAAAUGUUGUCCCCCAAAAUGUAGCC	
	GTGGAGAACGGCGGCCTGAAGGC	AUUCGUAUCUGCUCCUAAUAAAAAGAAA	
	CAAGACCATCACCAAGAAGAGGA	GUUUCUUCACAUUUCUUCUCUCCCUUGC	
	AGAAGGAGATCGAGGAGAGCAAG	ACCUCGGAUCAGUACUAGACUGAUAGCU	
	GAGCCCGGCGTGGAGGACACCGA	UAUUGGGAAAAAAAAAAAAAAAAAAAAA	
	GTGGAGCAACACCCAGACCGAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGGCCATCCAGACCAGGAGCAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCAGGAAGAGGGTGCAGAAGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATCTGGTTACCACTAAACCAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCAAGAACACCCGAATGGAGTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTAAGCTACATAATACCAACTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACACTTACAAAATGTTGTCCCCC	AAAAAAAAAAA	
	AAAATGTAGCCATTCGTATCTGC		
	TCCTAATAAAAAGAAAGTTTCTT		
	CACATTTCTTCTCTCCCTTGCAC		
	CTCGGATCAGTACTAGACTGATA		
	GCTTATTGGG		
163	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MAVGEIITDMAKKEWKV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	GLPIGQGGFGCIYLADM
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	NSSESVGSDAPCVVKVE
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	PSDNGPLFTELKFYQRA
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	AKPEQIQKWIRTRKLKY
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGCCGUGGGCGAGAUCAUCAC	LGVPKYWGSGLHDKNGK
	TCACCATTTACGAACGATAGGCC	CGACAUGGCCAAGAAGGAGUGGAAGGUG	SYRFMIMDRFGSDLQKI
	ACCATGGCCGTGGGCGAGATCAT	GGCCUGCCCAUCGGCCAGGGCGGCUUCG	YEANAKRFSRKTVLQLS
	CACCGACATGGCCAAGAAGGAGT	GCUGCAUCUACCUGGCCGACAUGAACAG	LRILDILEYIHEHEYVH
	GGAAGGTGGGCCTGCCCATCGGC	CAGCGAGAGCGUGGGCAGCGACGCCCCC	GDIKASNLLLNYKNPDQ
	CAGGGCGGCTTCGGCTGCATCTA	UGCGUGGUGAAGGUGGAGCCCAGCGACA	VYLVDYGLAYRYCPEGV
	CCTGGCCGACATGAACAGCAGCG	ACGGCCCCCUGUUCACCGAGCUGAAGUU	HKEYKEDPKRCHDGTIE
	AGAGCGTGGGCAGCGACGCCCCC	CUACCAGAGGGCCGCCAAGCCCGAGCAG	FTSIDAHNGVAPSRRGD
	TGCGTGGTGAAGGTGGAGCCCAG	AUCCAGAAGUGGAUCAGGACCAGGAAGC	LEILGYCMIQWLTGHLP
	CGACAACGGCCCCCTGTTCACCG	UGAAGUACCUGGGCGUGCCCAAGUACUG	WEDNLKDPKYVRDSKIR
	AGCTGAAGTTCTACCAGAGGGCC	GGGCAGCGGCCUGCACGACAAGAACGGC	YRENIASLMDKCFPEKN
	GCCAAGCCCGAGCAGATCCAGAA	AAGAGCUACAGGUUCAUGAUCAUGGACA	KPGEIAKYMETVKLLDY
	GTGGATCAGGACCAGGAAGCTGA	GGUUCGGCAGCGACCUGCAGAAGAUCUA	TEKPLYENLRDILLQG
	AGTACCTGGGCGTGCCCAAGTAC	CGAGGCCAACGCCAAGAGGUUCAGCAGG	
	TGGGGCAGCGGCCTGCACGACAA	AAGACCGUGCUGCAGCUGAGCCUGAGGA	
	GAACGGCAAGAGCTACAGGTTCA	UCCUGGACAUCCUGGAGUACAUCCACGA	
	TGATCATGGACAGGTTCGGCAGC	GCACGAGUACGUGCACGGCGACAUCAAG	
	GACCTGCAGAAGATCTACGAGGC	GCCAGCAACCUGCUGCUGAACUACAAGA	
	CAACGCCAAGAGGTTCAGCAGGA	ACCCCGACCAGGUGUACCUGGUGGACUA	

	AGACCGTGCTGCAGCTGAGCCTG AGGATCCTGGACATCCTGGAGTA CATCCACGAGCACGAGTACGTGC ACGGCGACATCAAGGCCAGCAAC CTGCTGCTGAACTACAAGAACCC CGACCAGGTGTACCTGGTGGACT ACGGCCTGGCCTACAGGTACTGC CCCGAGGGCGTGCACAAGGAGTA CAAGGAGGACCCCAAGAGGTGCC ACGACGGCACCATCGAGTTCACC AGCATCGACGCCCACAACGGCGT GGCCCCCAGCAGGAGGGGCGACC TGGAGATCCTGGGCTACTGCATG ATCCAGTGGCTGACCGGCCACCT GCCCTGGGAGGACAACCTGAAGG ACCCCAAGTACGTGAGGGACAGC AAGATCAGGTACAGGGAGAACAT CGCCAGCCTGATGGACAAGTGCT TCCCCGAGAAGAACAAGCCCGGC GAGATCGCCAAGTACATGGAGAC CGTGAAGCTGCTGGACTACACCG AGAAGCCCCTGTACGAGAACCTG AGGGACATCCTGCTGCAGGGCTA AATCTGGTTACCACTAAACCAGC CTCAAGAACACCCGAATGGAGTC TCTAAGCTACATAATACCAACTT ACACTTACAAAATGTTGTCCCCC AAAATGTAGCCATTCGTATCTGC TCCTAATAAAAAGAAAGTTTCTT CACATTTCTTCTCTCCCTTGCAC CTCGGATCAGTACTAGACTGATA GCTTATTGGG	CGGCCUGGCCUACAGGUACUGCCCCGAG GGCGUGCACAAGGAGUACAAGGAGGACC CCAAGAGGUGCCACGACGGCACCAUCGA GUUCACCAGCAUCGACGCCCACAACGGC GUGGCCCCCAGCAGGAGGGGCGACCUGG AGAUCCUGGGCUACUGCAUGAUCCAGUG GCUGACCGGCCACCUGCCCUGGGAGGAC AACCUGAAGGACCCCAAGUACGUGAGGG ACAGCAAGAUCAGGUACAGGGAGAACAU CGCCAGCCUGAUGGACAAGUGCUUCCCC GAGAAGAACAAGCCCGGCGAGAUCGCCA AGUACAUGGAGACCGUGAAGCUGCUGGA CUACACCGAGAAGCCCCUGUACGAGAAC CUGAGGGACAUCCUGCUGCAGGGCUAAA UCUGGUUACCACUAAACCAGCCUCAAGA ACACCCGAAUGGAGUCUCUAAGCUACAU AAUACCAACUUACACUUACAAAAUGUUG UCCCCCAAAAUGUAGCCAUUCGUAUCUG CUCCUAAUAAAAAGAAAGUUUCUUCACA UUUCUUCUCUCCCUUGCACCUCGGAUCA GUACUAGACUGAUAGCUUAUUGGGAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
164	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGGGTGAAGGCCGC CCAGGCCGGCAGGCAGAGCAGCG CCAAGAGGCACCTGGCCGAGCAG TTCGCCGTGGGCGAGATCATCAC CGACATGGCCAAGAAGGAGTGGA AGGTGGGCCTGCCCATCGGCCAG GGCGGCTTCGGCTGCATCTACCT GGCCGACATGAACAGCAGCGAGA GCGTGGGCAGCGACGCCCCCTGC GTGGTGAAGGTGGAGCCCAGCGA CAACGGCCCCCTGTTCACCGAGC TGAAGTTCTACCAGAGGGCCGCC AAGCCCGAGCAGATCCAGAAGTG GATCAGGACCAGGAAGCTGAAGT ACCTGGGCGTGCCCAAGTACTGG GGCAGCGGCCTGCACGACAAGAA CGGCAAGAGCTACAGGTTCATGA TCATGGACAGGTTCGGCAGCGAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGGGUGAAGGCCGCCCA GGCCGGCAGGCAGAGCAGCGCCAAGAGG CACCUGGCCGAGCAGUUCGCCGUGGGCG AGAUCAUCACCGACAUGGCCAAGAAGGA GUGGAAGGUGGGCCUGCCCAUCGGCCAG GGCGGCUUCGGCUGCAUCUACCUGGCCG ACAUGAACAGCAGCGAGAGCGUGGGCAG CGACGCCCCCUGCGUGGUGAAGGUGGAG CCCAGCGACAACGGCCCCCUGUUCACCG AGCUGAAGUUCUACCAGAGGGCCGCCAA GCCCGAGCAGAUCCAGAAGUGGAUCAGG ACCAGGAAGCUGAAGUACCUGGGCGUGC CCAAGUACUGGGGCAGCGGCCUGCACGA CAAGAACGGCAAGAGCUACAGGUUCAUG AUCAUGGACAGGUUCGGCAGCGACCUGC AGAAGAUCUACGAGGCCAACGCCAAGAG GUUCAGCAGGAAGACCGUGCUGCAGCUG AGCCUGAGGAUCCUGGACAUCCUGGAGU ACAUCCACGAGCACGAGUACGUGCACGG CGACAUCAAGGCCAGCAACCUGCUGCUG	MPRVKAAQAGRQSSAKR HLAEQFAVGEIITDMAK KEWKVGLPIGQGGFGCI YLADMNSSESVGSDAPC VVKVEPSDNGPLFTELK FYQRAAKPEQIQKWIRT RKLKYLGVPKYWGSGLH DKNGKSYRFMIMDRFGS DLQKIYEANAKRFSRKI VLQLSLRILDILEYIHE HEYVHGDIKASNLLLNY KNPDQVYLVDYGLAYRY CPEGVHKEYKEDPKRCH DGTIEFTSIDAHNGVAP SRRGDLEILGYCMIQWL TGHLPWEDNLKDPKYVR DSKIRYRENIASLMDKC FPEKNKPGEIAKYMETV KLLDYTEKPLYKNLRDI LLQGLKAIGSKDDGKLD LSVVENGGLKAKTITKK RKKEIEESKEPGVEDTE WSNTQTEEAIQTRSRTR KRVQK

	CTGCAGAAGATCTACGAGGCCAA CGCCAAGAGGTTCAGCAGGAAGA CCGTGCTGCAGCTGAGCCTGAGG ATCCTGGACATCCTGGAGTACAT CCACGAGCACGAGTACGTGCACG GCGACATCAAGGCCAGCAACCTG CTGCTGAACTACAAGAACCCCGA CCAGGTGTACCTGGTGGACTACG GCCTGGCCTACAGGTACTGCCCC GAGGGCGTGCACAAGGAGTACAA GGAGGACCCCAAGAGGTGCCACG ACGGCACCATCGAGTTCACCAGC ATCGACGCCCACAACGGCGTGGC CCCCAGCAGGAGGGGCGACCTGG AGATCCTGGGCTACTGCATGATC CAGTGGCTGACCGGCCACCTGCC CTGGGAGGACAACCTGAAGGACC CCAAGTACGTGAGGGACAGCAAG ATCAGGTACAGGGAGAACATCGC CAGCCTGATGGACAAGTGCTTCC CCGAGAAGAACAAGCCCGGCGAG ATCGCCAAGTACATGGAGACCGT GAAGCTGCTGGACTACACCGAGA AGCCCCTGTACAAGAACCTGAGG GACATCCTGCTGCAGGGCCTGAA GGCCATCGGCAGCAAGGACGACG GCAAGCTGGACCTGAGCGTGGTG GAGAACGGCGGCCTGAAGGCCAA GACCATCACCAAGAAGAGGAAGA AGGAGATCGAGGAGAGCAAGGAG CCCGGCGTGGAGGACACCGAGTG GAGCAACACCCAGACCGAGGAGG CCATCCAGACCAGGAGCAGGACC AGGAAGAGGGTGCAGAAGTAAAT CTGGTTACCACTAAACCAGCCTC AAGAACACCCGAATGGAGTCTCT AAGCTACATAATACCAACTTACA CTTACAAAATGTTGTCCCCCAAA ATGTAGCCATTCGTATCTGCTCC TAATAAAAAGAAAGTTTCTTCAC ATTTCTTCTCTCCCTTGCACCTC GGATCAGTACTAGACTGATAGCT TATTGGG	AACUACAAGAACCCCGACCAGGUGUACC UGGUGGACUACGGCCUGGCCUACAGGUA CUGCCCCGAGGGCGUGCACAAGGAGUAC AAGGAGGACCCCAAGAGGUGCCACGACG GCACCAUCGAGUUCACCAGCAUCGACGC CCACAACGGCGUGGCCCCCAGCAGGAGG GGCGACCUGGAGAUCCUGGGCUACUGCA UGAUCCAGUGGCUGACCGGCCACCUGCC CUGGGAGGACAACCUGAAGGACCCCAAG UACGUGAGGGACAGCAAGAUCAGGUACA GGGAGAACAUCGCCAGCCUGAUGGACAA GUGCUUCCCCGAGAAGAACAAGCCCGGC GAGAUCGCCAAGUACAUGGAGACCGUGA AGCUGCUGGACUACACCGAGAAGCCCCU GUACAAGAACCUGAGGGACAUCCUGCUG CAGGGCCUGAAGGCCAUCGGCAGCAAGG ACGACGGCAAGCUGGACCUGAGCGUGGU GGAGAACGGCGGCCUGAAGGCCAAGACC AUCACCAAGAAGAGGAAGAAGGAGAUCG AGGAGAGCAAGGAGCCCGGCGUGGAGGA CACCGAGUGGAGCAACACCCAGACCGAG GAGGCCAUCCAGACCAGGAGCAGGACCA GGAAGAGGGUGCAGAAGUAAAUCUGGUU ACCACUAAACCAGCCUCAAGAACACCCG AAUGGAGUCUCUAAGCUACAUAAUACCA ACUUACACUUACAAAAUGUUGUCCCCCA AAAUGUAGCCAUUCGUAUCUGCUCCUAA UAAAAAGAAAGUUUCUUCACAUUUCUUC UCUCCCUUGCACCUCGGAUCAGUACUAG ACUGAUAGCUUAUUGGGAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
165	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGGGTGAAGGCCGC CCAGGCCGGCAGGCAGAGCAGCG CCAAGAGGCACCTGGCCGAGCAG TTCGCCGTGGGCGAGATCATCAC CGACATGGCCAAGAAGGAGTGGA AGGTGGGCCTGCCCATCGGCCAG GGCGGCTTCGGCTGCATCTACCT	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGGGUGAAGGCCGCCCA GGCCGGCAGGCAGAGCAGCGCCAAGAGG CACCUGGCCGAGCAGUUCGCCGUGGGCG AGAUCAUCACCGACAUGGCCAAGAAGGA GUGGAAGGUGGGCCUGCCCAUCGGCCAG GGCGGCUUCGGCUGCAUCUACCUGGCCG ACAUGAACAGCAGCGAGAGCGUGGGCAG CGACGCCCCCUGCGUGGUGAAGGUGGAG CCCAGCGACAACGGCCCCCUGUUCACCG	MPRVKAAQAGRQSSAKR HLAEQFAVGEIITDMAK KEWKVGLPIGQGGFGCI YLADMNSSESVGSDAPC VVKVEPSDNGPLFTELK FYQRAAKPEQIQKWIRT RKLKYLGVPKYWGSGLH DKNGKSYRFMIMDRFGS DLQKIYEANAKRFSRKT VLQLSLRILDILEYIHE HEYVHGDIKASNLLLNY KNPDQVYLVDYGLAYRY CPEGVHKEYKEDPKRCH DGTIEFTSIDAHNGVAP

	GGCCGACATGAACAGCAGCGAGA	AGCUGAAGUUCUACCAGAGGGCCGCCAA	SRRGDLEILGYCMIQWL
	GCGTGGGCAGCGACGCCCCCTGC	GCCCGAGCAGAUCCAGAAGUGGAUCAGG	TGHLPWEDNLKDPKYVR
	GTGGTGAAGGTGGAGCCCAGCGA	ACCAGGAAGCUGAAGUACCUGGGCGUGC	DSKIRYRENIASLMDKC
	CAACGGCCCCCTGTTCACCGAGC	CCAAGUACUGGGGCAGCGGCCUGCACGA	FPEKNKPGEIAKYMETV
	TGAAGTTCTACCAGAGGGCCGCC	CAAGAACGGCAAGAGCUACAGGUUCAUG	KLLDYTEKPLYKNLRDI
	AAGCCCGAGCAGATCCAGAAGTG	AUCAUGGACAGGUUCGGCAGCGACCUGC	LLQG
	GATCAGGACCAGGAAGCTGAAGT	AGAAGAUCUACGAGGCCAACGCCAAGAG	
	ACCTGGGCGTGCCCAAGTACTGG	GUUCAGCAGGAAGACCGUGCUGCAGCUG	
	GGCAGCGGCCTGCACGACAAGAA	AGCCUGAGGAUCCUGGACAUCCUGGAGU	
	CGGCAAGAGCTACAGGTTCATGA	ACAUCCACGAGCACGAGUACGUGCACGG	
	TCATGGACAGGTTCGGCAGCGAC	CGACAUCAAGGCCAGCAACCUGCUGCUG	
	CTGCAGAAGATCTACGAGGCCAA	AACUACAAGAACCCCGACCAGGUGUACC	
	CGCCAAGAGGTTCAGCAGGAAGA	UGGUGGACUACGGCCUGGCCUACAGGUA	
	CCGTGCTGCAGCTGAGCCTGAGG	CUGCCCCGAGGGCGUGCACAAGGAGUAC	
	ATCCTGGACATCCTGGAGTACAT	AAGGAGGACCCCAAGAGGUGCCACGACG	
	CCACGAGCACGAGTACGTGCACG	GCACCAUCGAGUUCACCAGCAUCGACGC	
	GCGACATCAAGGCCAGCAACCTG	CCACAACGGCGUGGCCCCCAGCAGGAGG	
	CTGCTGAACTACAAGAACCCCGA	GGCGACCUGGAGAUCCUGGGCUACUGCA	
	CCAGGTGTACCTGGTGGACTACG	UGAUCCAGUGGCUGACCGGCCACCUGCC	
	GCCTGGCCTACAGGTACTGCCCC	CUGGGAGGACAACCUGAAGGACCCCAAG	
	GAGGGCGTGCACAAGGAGTACAA	UACGUGAGGGACAGCAAGAUCAGGUACA	
	GGAGGACCCCAAGAGGTGCCACG	GGGAGAACAUCGCCAGCCUGAUGGACAA	
	ACGGCACCATCGAGTTCACCAGC	GUGCUUCCCCGAGAAGAACAAGCCCGGC	
	ATCGACGCCCACAACGGCGTGGC	GAGAUCGCCAAGUACAUGGAGACCGUGA	
	CCCCAGCAGGAGGGGCGACCTGG	AGCUGCUGGACUACACCGAGAAGCCCCU	
	AGATCCTGGGCTACTGCATGATC	GUACAAGAACCUGAGGGACAUCCUGCUG	
	CAGTGGCTGACCGGCCACCTGCC	CAGGGCUAAAUCUGGUUACCACUAAACC	
	CTGGGAGGACAACCTGAAGGACC	AGCCUCAAGAACACCCGAAUGGAGUCUC	
	CCAAGTACGTGAGGGACAGCAAG	UAAGCUACAUAAUACCAACUUACACUUA	
	ATCAGGTACAGGGAGAACATCGC	CAAAAUGUUGUCCCCCAAAAUGUAGCCA	
	CAGCCTGATGGACAAGTGCTTCC	UUCGUAUCUGCUCCUAAUAAAAAGAAAG	
	CCGAGAAGAACAAGCCCGGCGAG	UUUCUUCACAUUUCUUCUCUCCCUUGCA	
	ATCGCCAAGTACATGGAGACCGT	CCUCGGAUCAGUACUAGACUGAUAGCUU	
	GAAGCTGCTGGACTACACCGAGA	AUUGGGAAAAAAAAAAAAAAAAAAAAAA	
	AGCCCCTGTACAAGAACCTGAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GACATCCTGCTGCAGGGCTAAAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTGGTTACCACTAAACCAGCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAGAACACCCGAATGGAGTCTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAGCTACATAATACCAACTTACA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTACAAAATGTTGTCCCCCAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAGCCATTCGTATCTGCTCC	AAAAAAAAAA	
	TAATAAAAAGAAAGTTTCTTCAC		
	ATTTCTTCTCTCCCTTGCACCTC		
	GGATCAGTACTAGACTGATAGCT		
	TATTGGG		
166	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MAVGEIITDMAKKEWKV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	GLPIGQGGFGCIYLADM
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	NSSESVGSDAPCVVKVE
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	PSDNGPLFTELKFYQRA
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	AKPEQIQKWIRTRKLKY
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGCCGUGGGCGAGAUCAUCAC	LGVPKYWGSGLHDKNGK
	TCACCATTTACGAACGATAGGCC	CGACAUGGCCAAGAAGGAGUGGAAGGUG	SYRFMIMDRFGSDLQKI
	ACCATGGCCGTGGGCGAGATCAT	GGCCUGCCCAUCGGCCAGGGCGGCUUCG	YEANAKRFSRKTVLQLS
	CACCGACATGGCCAAGAAGGAGT	GCUGCAUCUACCUGGCCGACAUGAACAG	LRILDILEYIHEHEYVH
	GGAAGGTGGGCCTGCCCATCGGC	CAGCGAGAGCGUGGGCAGCGACGCCCCC	GDIKASNLLLNYKNPDQ
	CAGGGCGGCTTCGGCTGCATCTA	UGCGUGGUGAAGGUGGAGCCCAGCGACA	VYLVDYGLAYRYCPEGV
	CCTGGCCGACATGAACAGCAGCG	ACGGCCCCCUGUUCACCGAGCUGAAGUU	HKEYKEDPKRCHDGTIE

	AGAGCGTGGGCAGCGACGCCCCC	CUACCAGAGGGCCGCCAAGCCCGAGCAG	FTSIDAHNGVAPSRRGD
	TGCGTGGTGAAGGTGGAGCCCAG	AUCCAGAAGUGGAUCAGGACCAGGAAGC	LEILGYCMIQWLTGHLP
	CGACAACGGCCCCCTGTTCACCG	UGAAGUACCUGGGCGUGCCCAAGUACUG	WEDNLKDPKYVRDSKIR
	AGCTGAAGTTCTACCAGAGGGCC	GGGCAGCGGCCUGCACGACAAGAACGGC	YRENIASLMDKCFPEKN
	GCCAAGCCCGAGCAGATCCAGAA	AAGAGCUACAGGUUCAUGAUCAUGGACA	KPGEIAKYMETVKLLDY
	GTGGATCAGGACCAGGAAGCTGA	GGUUCGGCAGCGACCUGCAGAAGAUCUA	TEKPLYKNLRDILLQG
	AGTACCTGGGCGTGCCCAAGTAC	CGAGGCCAACGCCAAGAGGUUCAGCAGG	
	TGGGGCAGCGGCCTGCACGACAA	AAGACCGUGCUGCAGCUGAGCCUGAGGA	
	GAACGGCAAGAGCTACAGGTTCA	UCCUGGACAUCCUGGAGUACAUCCACGA	
	TGATCATGGACAGGTTCGGCAGC	GCACGAGUACGUGCACGGCGACAUCAAG	
	GACCTGCAGAAGATCTACGAGGC	GCCAGCAACCUGCUGCUGAACUACAAGA	
	CAACGCCAAGAGGTTCAGCAGGA	ACCCCGACCAGGUGUACCUGGUGGACUA	
	AGACCGTGCTGCAGCTGAGCCTG	CGGCCUGGCCUACAGGUACUGCCCCGAG	
	AGGATCCTGGACATCCTGGAGTA	GGCGUGCACAAGGAGUACAAGGAGGACC	
	CATCCACGAGCACGAGTACGTGC	CCAAGAGGUGCCACGACGGCACCAUCGA	
	ACGGCGACATCAAGGCCAGCAAC	GUUCACCAGCAUCGACGCCCACAACGGC	
	CTGCTGCTGAACTACAAGAACCC	GUGGCCCCCAGCAGGAGGGGCGACCUGG	
	CGACCAGGTGTACCTGGTGGACT	AGAUCCUGGGCUACUGCAUGAUCCAGUG	
	ACGGCCTGGCCTACAGGTACTGC	GCUGACCGGCCACCUGCCCUGGGAGGAC	
	CCCGAGGGCGTGCACAAGGAGTA	AACCUGAAGGACCCCAAGUACGUGAGGG	
	CAAGGAGGACCCCAAGAGGTGCC	ACAGCAAGAUCAGGUACAGGGAGAACAU	
	ACGACGGCACCATCGAGTTCACC	CGCCAGCCUGAUGGACAAGUGCUUCCCC	
	AGCATCGACGCCCACAACGGCGT	GAGAAGAACAAGCCCGGCGAGAUCGCCA	
	GGCCCCCAGCAGGAGGGGCGACC	AGUACAUGGAGACCGUGAAGCUGCUGGA	
	TGGAGATCCTGGGCTACTGCATG	CUACACCGAGAAGCCCCUGUACAAGAAC	
	ATCCAGTGGCTGACCGGCCACCT	CUGAGGGACAUCCUGCUGCAGGGCUAAA	
	GCCCTGGGAGGACAACCTGAAGG	UCUGGUUACCACUAAACCAGCCUCAAGA	
	ACCCCAAGTACGTGAGGGACAGC	ACACCCGAAUGGAGUCUCUAAGCUACAU	
	AAGATCAGGTACAGGGAGAACAT	AAUACCAACUUACACUUACAAAAUGUUG	
	CGCCAGCCTGATGGACAAGTGCT	UCCCCCAAAAUGUAGCCAUUCGUAUCUG	
	TCCCCGAGAAGAACAAGCCCGGC	CUCCUAAUAAAAAGAAAGUUUCUUCACA	
	GAGATCGCCAAGTACATGGAGAC	UUUCUUCUCUCCCUUGCACCUCGGAUCA	
	CGTGAAGCTGCTGGACTACACCG	GUACUAGACUGAUAGCUUAUUGGGAAAA	
	AGAAGCCCCTGTACAAGAACCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGGGACATCCTGCTGCAGGGCTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATCTGGTTACCACTAAACCAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCAAGAACACCCGAATGGAGTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTAAGCTACATAATACCAACTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACACTTACAAAATGTTGTCCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAAATGTAGCCATTCGTATCTGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCTAATAAAAAGAAAGTTTCTT		
	CACATTTCTTCTCTCCCTTGCAC		
	CTCGGATCAGTACTAGACTGATA		
	GCTTATTGGG		
167	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MPPKRNEKYKLPIPFPE
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	GKVLDDMEGNQWVLGKK
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	IGSGGFGLIYLAFPTNK
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	PEKDARHVVKVEYQENG
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	PLFSELKFYQRVAKKDC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGCCCCCCAAGAGGAACGAGAA	IKKWIERKQLDYLGIPL
	TCACCATTTACGAACGATAGGCC	GUACAAGCUGCCCAUCCCCUUCCCCGAG	FYGSGLTEFKGRSYRFM
	ACCATGCCCCCCAAGAGGAACGA	GGCAAGGUGCUGGACGACAUGGAGGGCA	VMERLGIDLQKISGQNG
	GAAGTACAAGCTGCCCATCCCCT	ACCAGUGGGUGCUGGGCAAGAAGAUCGG	TFKKSTVLQLGIRMLDV
	TCCCCGAGGGCAAGGTGCTGGAC	CAGCGGCGGCUUCGGCCUGAUCUACCUG	LEYIHENEYVHGDIKAA
	GACATGGAGGGCAACCAGTGGGT	GCCUUCCCCACCAACAAGCCCGAGAAGG	NLLLGYKNPDQVYLADY
	GCTGGGCAAGAAGATCGGCAGCG	ACGCCAGGCACGUGGUGAAGGUGGAGUA	GLSYRYCPNGNHKQYQE
	GCGGCTTCGGCCTGATCTACCTG	CCAGGAGAACGGCCCCCUGUUCAGCGAG	NPRKGHNGTIEFTSLDA

	GCCTTCCCCACCAACAAGCCCGA	CUGAAGUUCUACCAGAGGGUGGCCAAGA	HKGVALSRRSDVEILGY
	GAAGGACGCCAGGCACGTGGTGA	AGGACUGCAUCAAGAAGUGGAUCGAGAG	CMLRWLCGKLPWEQNLK
	AGGTGGAGTACCAGGAGAACGGC	GAAGCAGCUGGACUACCUGGGCAUCCCC	DPVAVQTAKTNLLDELP
	CCCCTGTTCAGCGAGCTGAAGTT	CUGUUCUACGGCAGCGGCCUGACCGAGU	QSVLKWAPSGSSCCEIA
	CTACCAGAGGGTGGCCAAGAAGG	UCAAGGGCAGGAGCUACAGGUUCAUGGU	QFLVCAHSLAYDEKPNY
	ACTGCATCAAGAAGTGGATCGAG	GAUGGAGAGGCUGGGCAUCGACCUGCAG	QALKKILNPHGIPLGPL
	AGGAAGCAGCTGGACCACCTGGG	AAGAUCAGCGGCCAGAACGGCACCUUCA	DFSTKGQSINVHTPNSQ
	CATCCCCCTGTTCTACGGCAGCG	AGAAGAGCACCGUGCUGCAGCUGGGCAU	KVDSQKAATKQVNKAHN
	GCCTGACCGAGTTCAAGGGCAGG	CAGGAUGCUGGACGUGCUGGAGUACAUC	RLIEKKVHSERSAESCA
	AGCTACAGGTTCATGGTGATGGA	CACGAGAACGAGUACGUGCACGGCGACA	TWKVQKEEKLIGLMNNE
	GAGGCTGGGCATCGACCTGCAGA	UCAAGGCCGCCAACCUGCUGCUGGGCUA	AAQVEA
	AGATCAGCGGCCAGAACGGCACC	CAAGAACCCCGACCAGGUGUACCUGGCC	
	TTCAAGAAGAGCACCGTGCTGCA	GACUACGGCCUGAGCUACAGGUACUGCC	
	GCTGGGCATCAGGATGCTGGACG	CCAACGGCAACCACAAGCAGUACCAGGA	
	TGCTGGAGTACATCCACGAGAAC	GAACCCCAGGAAGGGCCACAACGGCACC	
	GAGTACGTGCACGGCGACATCAA	AUCGAGUUCACCAGCCUGGACGCCCACA	
	GGCCGCCAACCTGCTGCTGGGCT	AGGGCGUGGCCCUGAGCAGGAGGAGCGA	
	ACAAGAACCCCGACCAGGTGTAC	CGUGGAGAUCCUGGGCUACUGCAUGCUG	
	CTGGCCGACTACGGCCTGAGCTA	AGGUGGCUGUGCGGCAAGCUGCCCUGGG	
	CAGGTACTGCCCCAACGGCAACC	AGCAGAACCUGAAGGACCCCGUGGCCGU	
	ACAAGCAGTACCAGGAGAACCCC	GCAGACCGCCAAGACCAACCUGCUGGAC	
	AGGAAGGGCCACAACGGCACCAT	GAGCUGCCCCAGAGCGUGCUGAAGUGGG	
	CGAGTTCACCAGCCTGGACGCCC	CCCCCAGCGGCAGCAGCUGCUGCGAGAU	
	ACAAGGGCGTGGCCCTGAGCAGG	CGCCCAGUUCCUGGUGUGCGCCCACAGC	
	AGGAGCGACGTGGAGATCCTGGG	CUGGCCUACGACGAGAAGCCCAACUACC	
	CTACTGCATGCTGAGGTGGCTGT	AGGCCCUGAAGAAGAUCCUGAACCCCCA	
	GCGGCAAGCTGCCCTGGGAGCAG	CGGCAUCCCCCUGGGCCCCCUGGACUUC	
	AACCTGAAGGACCCCGTGGCCGT	AGCACCAAGGGCCAGAGCAUCAACGUGC	
	GCAGACCGCCAAGACCAACCTGC	ACACCCCCAACAGCCAGAAGGUGGACAG	
	TGGACGAGCTGCCCCAGAGCGTG	CCAGAAGGCCGCCACCAAGCAGGUGAAC	
	CTGAAGTGGGCCCCCAGCGGCAG	AAGGCCCACAACAGGCUGAUCGAGAAGA	
	CAGCTGCTGCGAGATCGCCCAGT	AGGUGCACAGCGAGAGGAGCGCCGAGAG	
	TCCTGGTGTGCGCCCACAGCCTG	CUGCGCCACCUGGAAGGUGCAGAAGGAG	
	GCCTACGACGAGAAGCCCAACTA	GAGAAGCUGAUCGGCCUGAUGAACAACG	
	CCAGGCCCTGAAGAAGATCCTGA	AGGCCGCCCAGGUGGAGGCCUAAAUCUG	
	ACCCCCACGGCATCCCCCTGGGC	GUUACCACUAAACCAGCCUCAAGAACAC	
	CCCCTGGACTTCAGCACCAAGGG	CCGAAUGGAGUCUCUAAGCUACAUAAUA	
	CCAGAGCATCAACGTGCACACCC	CCAACUUACACUUACAAAAUGUUGUCCC	
	CCAACAGCCAGAAGGTGGACAGC	CCAAAAUGUAGCCAUUCGUAUCUGCUCC	
	CAGAAGGCCGCCACCAAGCAGGT	UAAUAAAAAGAAAGUUUCUUCACAUUUC	
	GAACAAGGCCCACAACAGGCTGA	UUCUCUCCCUUGCACCUCGGAUCAGUAC	
	TCGAGAAGAAGGTGCACAGCGAG	UAGACUGAUAGCUUAUUGGGAAAAAAAA	
	AGGAGCGCCGAGAGCTGCGCCAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTGGAAGGTGCAGAAGGAGGAGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGCTGATCGGCCTGATGAACAAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGGCCGCCCAGGTGGAGGCCTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATCTGGTTACCACTAAACCAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCAAGAACACCCGAATGGAGTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTAAGCTACATAATACCAACTT	AAAAAAAAAAAAAAAAAAAAAA	
	ACACTTACAAAATGTTGTCCCCC		
	AAAATGTAGCCATTCGTATCTGC		
	TCCTAATAAAAAGAAAGTTTCTT		
	CACATTTCTTCTCTCCCTTGCAC		
	CTCGGATCAGTACTAGACTGATA		
	GCTTATTGGG		

168	TAATACGACTCACTA-AAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MPPKRNEKYK-PIPFPE
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	GKVLDDMEGNQWVLGKK
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	IGSGGFGLIY-AFPTNK
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	PEKDARHVVKVEYQENG
	AATTTAAATCATTTC-TTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	PLFSELKFYQRVAKKDC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGCCCCCCAAGAGGAACGAGAA	IKKWIERKQLDYLGIPL
	TCACCATTTACGAACGATAGGCC	GUACAAGCUGCCCAUCCCCUUCCCCGAG	FYGSGLTEFKGRSYRFM
	ACCATGCCCCCCCAAGAGGAACGA	GGCAAGGUGCUGGACGACAUGGAGGGCA	VMERLGIDLQKISGQNG
	GAAGTACAAGCTGCCCATCCCCT	ACCAGUGGGUGCUGGGCAAGAAGAUCGG	TFKKSTVLQLGIRMLDV
	TCCCCGAGGGCAAGG-GCTGGAC	CAGCGGCGGCUUCGGCCUGAUCUACCUG	LEYIHENEYVHGDIKAA
	GACATGGAGGGCAACCAGTGGGT	GCCUUCCCCACCAACAAGCCCGAGAAGG	NLLLGYKNPDQVYLADY
	GCTGGGCAAGAAGATCGGCAGCG	ACGCCAGGCACGUGGUGAAGGUGGAGUA	GLSYRYCPNGNHKQYQE
	GCGGCTTCGGCCTGA-CTACCTG	CCAGGAGAACGGCCCCCUGUUCAGCGAG	NPRKGHNGTIEFTSLDA
	GCCTTCCCCACCAACAAGCCCGA	CUGAAGUUCUACCAGAGGGUGGCCAAGA	HKGVALSRRSDVEILGY
	GAAGGACGCCAGGCACGTGGTGA	AGGACUGCAUCAAGAAGUGGAUCGAGAG	CMLRWLCGKLPWEQNLK
	AGGTGGAGTACCAGGAGAACGGC	GAAGCAGCUGGACUACCUGGGCAUCCCC	DPVAVQTAKTNLLDELP
	CCCCTGTTCAGCGAGCTGAAGTT	CUGUUCUACGGCAGCGGCCUGACCGAGU	QSVLKWAPSGSSCCEIA
	CTACCAGAGGGTGGCCAAGAAGG	UCAAGGGCAGGAGCUACAGGUUCAUGGU	QFLVCAHSLAYDEKPNY
	ACTGCATCAAGAAGTGGATCGAG	GAUGGAGAGGCUGGGCAUCGACCUGCAG	QALKKILNPHGIPLGPL
	AGGAAGCAGCTGGAC-ACCTGGG	AAGAUCAGCGGCCAGAACGGCACCUUCA	DFSTKGQSINVHTPNSQ
	CATCCCCCTGTTCTACGGCAGCG	AGAAGAGCACCGUGCUGCAGCUGGGCAU	KVDSQKAATKQVNKAHN
	GCCTGACCGAGTTCAAGGGCAGG	CAGGAUGCUGGACGUGCUGGAGUACAUC	RLIEKKVHSERSAESCA
	AGCTACAGGTTCATGGTGATGGA	CACGAGAACGAGUACGUGCACGGCGACA	TWKVQKEEKLIGLMNNE
	GAGGCTGGGCATCGACCTGCAGA	UCAAGGCCGCCAACCUGCUGCUGGGCUA	AAQVEA
	AGATCAGCGGCCAGAACGGCACC	CAAGAACCCCGACCAGGUGUACCUGGCC	
	TTCAAGAAGAGCACCGTGCTGCA	GACUACGGCCUGAGCUACAGGUACUGCC	
	GCTGGGCATCAGGATGCTGGACG	CCAACGGCAACCACAAGCAGUACCAGGA	
	TGCTGGAGTACATCCACGAGAAC	GAACCCCAGGAAGGGCCACAACGGCACC	
	GAGTACGTGCACGGCGACATCAA	AUCGAGUUCACCAGCCUGGACGCCCACA	
	GGCCGCCAACCTGCTGCTGGGCT	AGGGCGUGGCCCUGAGCAGGAGGAGCGA	
	ACAAGAACCCCGACCAGGTGTAC	CGUGGAGAUCCUGGGCUACUGCAUGCUG	
	CTGGCCGACTACGGCCTGAGCTA	AGGUGGCUGUGCGGCAAGCUGCCCUGGG	
	CAGGTACTGCCCCAACGGCAACC	AGCAGAACCUGAAGGACCCCGUGGCCGU	
	ACAAGCAGTACCAGGAGAACCCC	GCAGACCGCCAAGACCAACCUGCUGGAC	
	AGGAAGGGCCACAACGGCACCAT	GAGCUGCCCCAGAGCGUGCUGAAGUGGG	
	CGAGTTCACCAGCCTGGACGCCC	CCCCCAGCGGCAGCAGCUGCUGCGAGAU	
	ACAAGGGCGTGGCCC-GAGCAGG	CGCCCAGUUCCUGGUGUGCGCCCACAGC	
	AGGAGCGACGTGGAGATCCTGGG	CUGGCCUACGACGAGAAGCCCAACUACA	
	CTACTGCATGCTGAGGTGGCTGT	AGGCCCUGAAGAAGAUCCUGAACCCCCA	
	GCGGCAAGCTGCCCTGGGAGCAG	CGGCAUCCCCCUGGGCCCCCUGGACUUC	
	AACCTGAAGGACCCCGTGGCCGT	AGCACCAAGGGCCAGAGCAUCAACGUGC	
	GCAGACCGCCAAGACCAACCTGC	ACACCCCCAACAGCCAGAAGGUGGACAG	
	TGGACGAGCTGCCCCAGAGCGTG	CCAGAAGGCCGCCACCAAGCAGGUGAAC	
	CTGAAGTGGGCCCCCAGCGGCAG	AAGGCCCACAACAGGCUGAUCGAGAAGA	
	CAGCTGCTGCGAGATCGCCCAGT	AGGUGCACAGCGAGAGGAGCGCCGAGAG	
	TCCTGGTGTGCGCCCACAGCCTG	CUGCGCCACCUGGAAGGUGCAGAAGGAG	
	GCCTACGACGAGAAGCCCAACTA	GAGAAGCUGAUCGGCCUGAUGAACAACG	
	CAAGGCCCTGAAGAAGATCCTGA	AGGCCGCCCAGGUGGAGGCCUAAAUCUG	
	ACCCCCACGGCATCCCCCTGGGC	GUUACCACUAAACCAGCCUCAAGAACAC	
	CCCCTGGACTTCAGCACCAAGGG	CCGAAUGGAGUCUCUAAGCUACAUAAUA	
	CCAGAGCATCAACGTGCACACCC	CCAACUUACACUUACAAAAUGUUGUCCC	
	CCAACAGCCAGAAGG-GGACAGC	CCAAAAUGUAGCCAUUCGUAUCUGCUCC	
	CAGAAGGCCGCCACCAAGCAGGT	UAAUAAAAAGAAAGUUUCUUCACAUUUC	
	GAACAAGGCCCACAACAGGCTGA	UUCUCUCCCUUGCACCUCGGAUCAGUAC	
	TCGAGAAGAAGGTGCACAGCGAG	UAGACUGAUAGCUUAUUGGGAAAAAAAA	
	AGGAGCGCCGAGAGC-GCGCCAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTGCAAGGTGCAGAAGGAGCAGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	AGCTGATCGGCCTGATGAACAAC GAGGCCGCCCAGGTGGAGGCCTA AATCTGGTTACCACTAAACCAGC CTCAAGAACACCCGAATGGAGTC TCTAAGCTACATAATACCAACTT ACACTTACAAAATGTTGTCCCCC AAAATGTAGCCATTCGTATCTGC TCCTAATAAAAAGAAAGTTTCTT CACATTTCTTCTCTCCCTTGCAC CTCGGATCAGTACTAGACTGATA GCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
169	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGAGCCCGCCGCCGGCTT CCTGAGCCCCAGGCCCTTCCAGA GGACCGCCGCCGCCACCGCCCCC CCCGCCGGCCCCGGCCCCCCCCC CAGCGCCCTGAGGGGCCCCGAGC TGGAGATGCTGGCCGGCCTGCCC ACCAGCGACCCCGGCAGGCTGAT CACCGACCCCAGGAGCGGCAGGA CCTACCTGAAGGGCAGGCTGCTG GGCAAGGGCGGCTTCGCCAGGTG CTACGAGGCCACCGACACCGAGA CCGGCAGCGCCTACGCCGTGAAG GTGATCCCCCAGAGCAGGGTGGT GAAGCCCCACCAGAGGGAGAAGA TCCTGAACGAGATCGAGCTGCAC AGGGACCTGCAGCACAGGCACAT CGTGAGGTTCAGCCACCACTTCG AGGACGCCGACAACATCTACATC TTCCTGGAGCTGTGCAGCAGGAA GAGCCTGGCCCACATCTGGAAGG CCAGGCACACCCTGCTGGAGCCC GAGGTGAGGTACTACCTGAGGCA GATCCTGAGCGGCCTGAAGTACC TGCACCAGAGGGGCATCCTGCAC AGGGACCTGAAGCTGGGCAACTT CTTCATCACCGAGAACATGGAGC TGAAGGTGGGCGACTTCGGCCTG GCCGCCAGGCTGGAGCCCCCCGA GCAGAGGAAGAAGACCATCTGCG GCACCCCCAACTACGTGGCCCCC GAGGTGCTGCTGAGGCAGGGCCA CGGCCCCGAGGCCGACGTGTGGA GCCTGGGCTGCGTGATGTACACC CTGCTGTGCGGCAGCCCCCCCTT CGAGACCGCCGACCTGAAGGAGA CCTACAGGTGCATCAAGCAGGTG CACTACACCCTGCCCGCCAGCCT GAGCCTGCCCGCCAGGCAGCTGC TGGCCGCCATCCTGAGGGCCAGC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGAGCCCGCCGCCGGCUUCCU GAGCCCCAGGCCCUUCCAGAGGACCGCC GCCGCCACCGCCCCCCCCGCCGGCCCCG GCCCCCCCCCCAGCGCCCUGAGGGGCCC CGAGCUGGAGAUGCUGGCCGGCCUGCCC ACCAGCGACCCCGGCAGGCUGAUCACCG ACCCCAGGAGCGGCAGGACCUACCUGAA GGGCAGGCUGCUGGGCAAGGGCGGCUUC GCCAGGUGCUACGAGGCCACCGACACCG AGACCGGCAGCGCCUACGCCGUGAAGGU GAUCCCCCAGAGCAGGGUGGUGAAGCCC CACCAGAGGGAGAAGAUCCUGAACGAGA UCGAGCUGCACAGGGACCUGCAGCACAG GCACAUCGUGAGGUUCAGCCACCACUUC GAGGACGCCGACAACAUCUACAUCUUCC UGGAGCUGUGCAGCAGGAAGAGCCUGGC CCACAUCUGGAAGGCCAGGCACACCCUG CUGGAGCCCGAGGUGAGGUACUACCUGA GGCAGAUCCUGAGCGGCCUGAAGUACCU GCACCAGAGGGGCAUCCUGCACAGGGAC CUGAAGCUGGGCAACUUCUUCAUCACCG AGAACAUGGAGCUGAAGGUGGGCGACUU CGGCCUGGCCGCCAGGCUGGAGCCCCCC GAGCAGAGGAAGAAGACCAUCUGCGGCA CCCCCAACUACGUGGCCCCCGAGGUGCU GCUGAGGCAGGGCCACGGCCCCGAGGCC GACGUGUGGAGCCUGGGCUGCGUGAUGU ACACCCUGCUGUGCGGCAGCCCCCCCUU CGAGACCGCCGACCUGAAGGAGACCUAC AGGUGCAUCAAGCAGGUGCACUACACCC UGCCCGCCAGCCUGAGCCUGCCCGCCAG GCAGCUGCUGGCCGCCAUCCUGAGGGCC AGCCCCAGGGACAGGCCCAGCAUCGACC AGAUCCUGAGGCACGACUUCUUCACCAA GGGCUACACCCCCGACAGGCUGCCCAUC AGCAGCUGCGUGACCGUGCCCGACCUGA CCCCCCCCAACCCCGCCAGGAGCCUGUU CGCCAAGGUGACCAAGAGCCUGUUCGUG AGGAAGAAGAAGAGCAAGAACCACGCCC AGGAGAGGGACGAGGUGAGCGGCCUGGU GAGCGGCCUGAUGAGGACCAGCGUGGGC	MEPAAGFLSPRPFQRTA AATAPPAGPGPPPSALR GPELEMLAGLPTSDPGR LITDPRSGRTYLKGRLL GKGGFARCYEATDTETG SAYAVKVIPQSRVVKPH QREKILNEIELHRDLQH RHIVRFSHHFEDADNIY IFLELCSRKSLAHIWKA RHTLLEPEVRYYLRQIL SGLKYLHQRGILHRDLK LGNFFITENMELKVGDF GLAARLEPPEQRKKTIC GTPNYVAPEVLLRQGHG PEADVWSLGCVMYTLLC GSPPFETADLKETYRCI KQVHYTLPASLSLPARQ LLAAILRASPRDRPSID QILRHDFFTKGYTPDRL PISSCVTVPDLTPPNPA RSLFAKVTKSLFVRKKK SKNHAQERDEVSGLVSG LMRTSVGHQDARPEAPA ASGPAPVSLVETAPEDS SPRGTLASSGHGFEEGL TVATVVESALCALRNCI AFMPPAEQNPAPLAQPE PLVWFSEWVGFSNKFGF GYQLSSRRVAVLFNDGT HMALSANRKTVHYNPTS TKHFSFSVGAVRRALQP QLGILRYFASYMEQHLM KGGDLPSVEEVEVPAPP LLLQWVKTDQALLMLFS DGTVQVNFYGDHTKLIL SGWEPLLVTFVARNRSA CTYLASHLRQLGCSPDL RQRLRYALRLLRDRSPA

	CCCAGGGACAGGCCCAGCATCGA	CACCAGGACGCCAGGCCCGAGGCCCCCG
	CCAGATCCTGAGGCACGACTTCT	CCGCCAGCGGCCCCGCCCCCGUGAGCCU
	TCACCAAGGGCTACACCCCCGAC	GGUGGAGACCGCCCCCGAGGACAGCAGC
	AGGCTGCCCATCAGCAGCTGCGT	CCCAGGGGCACCCUGGCCAGCAGCGGCC
	GACCGTGCCCGACCTGACCCCCC	ACGGCUUCGAGGAGGGCCUGACCGUGGC
	CCAACCCCGCCAGGAGCCTGTTC	CACCGUGGUGGAGAGCGCCCUGUGCGCC
	GCCAAGGTGACCAAGAGCCTGTT	CUGAGGAACUGCAUCGCCUUCAUGCCCC
	CGTGAGGAAGAAGAAGAGCAAGA	CCGCCGAGCAGAACCCCGCCCCCCUGGC
	ACCACGCCCAGGAGAGGGACGAG	CCAGCCCGAGCCCCUGGUGUGGUUCAGC
	GTGAGCGGCCTGGTGAGCGGCCT	GAGUGGGUGGGCUUCAGCAACAAGUUCG
	GATGAGGACCAGCGTGGGCCACC	GCUUCGGCUACCAGCUGAGCAGCAGGAG
	AGGACGCCAGGCCCGAGGCCCCC	GGUGGCCGUGCUGUUCAACGACGGCACC
	GCCGCCAGCGGCCCCGCCCCCGT	CACAUGGCCCUGAGCGCCAACAGGAAGA
	GAGCCTGGTGGAGACCGCCCCCG	CCGUGCACUACAACCCCACCAGCACCAA
	AGGACAGCAGCCCCAGGGGCACC	GCACUUCAGCUUCAGCGUGGGCGCCGUG
	CTGGCCAGCAGCGGCCACGGCTT	AGGAGGGCCCUGCAGCCCCAGCUGGGCA
	CGAGGAGGGCCTGACCGTGGCCA	UCCUGAGGUACUUCGCCAGCUACAUGGA
	CCGTGGTGGAGAGCGCCCTGTGC	GCAGCACCUGAUGAAGGGCGGCGACCUG
	GCCCTGAGGAACTGCATCGCCTT	CCCAGCGUGGAGGAGGUGGAGGUGCCCG
	CATGCCCCCCGCCGAGCAGAACC	CCCCCCCCCUGCUGCUGCAGUGGGUGAA
	CCGCCCCCCTGGCCCAGCCCGAG	GACCGACCAGGCCCUGCUGAUGCUGUUC
	CCCCTGGTGTGGTTCAGCGAGTG	AGCGACGGCACCGUGCAGGUGAACUUCU
	GGTGGGCTTCAGCAACAAGTTCG	ACGGCGACCACACCAAGCUGAUCCUGAG
	GCTTCGGCTACCAGCTGAGCAGC	CGGCUGGGAGCCCCUGCUGGUGACCUUC
	AGGAGGGTGGCCGTGCTGTTCAA	GUGGCCAGGAACAGGAGCGCCUGCACCU
	CGACGGCACCCACATGGCCCTGA	ACCUGGCCAGCCACCUGAGGCAGCUGGG
	GCGCCAACAGGAAGACCGTGCAC	CUGCAGCCCCGACCUGAGGCAGAGGCUG
	TACAACCCCACCAGCACCAAGCA	AGGUACGCCCUGAGGCUGCUGAGGGACA
	CTTCAGCTTCAGCGTGGGCGCCG	GGAGCCCCGCCUAAAUCUGGUUACCACU
	TGAGGAGGGCCCTGCAGCCCCAG	AAACCAGCCUCAAGAACACCCGAAUGGA
	CTGGGCATCCTGAGGTACTTCGC	GUCUCUAAGCUACAUAAUACCAACUUAC
	CAGCTACATGGAGCAGCACCTGA	ACUUACAAAAUGUUGUCCCCCAAAAUGU
	TGAAGGGCGGCGACCTGCCCAGC	AGCCAUUCGUAUCUGCUCCUAAUAAAAA
	GTGGAGGAGGTGGAGGTGCCCGC	GAAAGUUUCUUCACAUUUCUUCUCUCCC
	CCCCCCCCTGCTGCTGCAGTGGG	UUGCACCUCGGAUCAGUACUAGACUGAU
	TGAAGACCGACCAGGCCCTGCTG	AGCUUAUUGGGAAAAAAAAAAAAAAAAA
	ATGCTGTTCAGCGACGGCACCGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	GCAGGTGAACTTCTACGGCGACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	ACACCAAGCTGATCCTGAGCGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	TGGGAGCCCCTGCTGGTGACCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	CGTGGCCAGGAACAGGAGCGCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	GCACCTACCTGGCCAGCCACCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
	AGGCAGCTGGGCTGCAGCCCCGA	AAAAAAAAAAAAAAA
	CCTGAGGCAGAGGCTGAGGTACG	
	CCCTGAGGCTGCTGAGGGACAGG	
	AGCCCCGCCTAAATCTGGTTACC	
	ACTAAACCAGCCTCAAGAACACC	
	CGAATGGAGTCTCTAAGCTACAT	
	AATACCAACTTACACTTACAAAA	
	TGTTGTCCCCCAAAATGTAGCCA	
	TTCGTATCTGCTCCTAATAAAAA	
	GAAAGTTTCTTCACATTTCTTCT	
	CTCCCTTGCACCTCGGATCAGTA	
	CTAGACTGATAGCTTATTGGG	

170	TAATACGACTCACTA<u>T</u>AAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCTGC CCACCCCCCCCACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCA<u>T</u>CGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGT<u>T</u>CCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGC<u>T</u>GCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCT<u>T</u>GCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCG<u>T</u>ACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCCGGATCAG<u>T</u>ACTAGAC TGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCUGCCCACCCCCCCC ACCACCAGCCAGAAGCACAGGGACUUCG UGGCCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGAGGUG CUGGGCAAGAAGCUGGAGGAGAGGGGCU UCGACAAGGCCUACGUGGUGCUGGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAGACCAGACAGCU GGAGCCUCGGUGGCCAUGCUUCUUGCCC CUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCCGGAU CAGUACUAGACUGAUAGCUUAUUGGGAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AA	MLPTPPTTSQKHRDFVA EPMGEKPVGS<u>L</u>AGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL
171	TAATACGACTCACTA<u>T</u>AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAAT<u>C</u>ATTTC<u>T</u>TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTG<u>T</u>TCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGA<u>T</u>CGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGA<u>T</u>CAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCA<u>T</u>CCAGTTC ATGCACGAGCAGGGC<u>T</u>ACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGA<u>T</u>CGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACG<u>T</u>GCCCTTC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACGCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACGAAGAACUGCGCC	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHN<u>I</u>KHVGL<u>I</u>TCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVML<u>I</u>GIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYEPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYY

	ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACGCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGAGAC GAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGAGCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAAATCTGGTTACCACT AAACCAGCCTCAAGAACACCCGA ATGGAGTCTCTAAGCTACATAAT ACCAACTTACACTTACAAAATGT TGTCCCCCAAAATGTAGCCATTC GTATCTGCTCCTAATAAAAAGAA AGTTTCTTCACATTTCTTCTCTC CCTTGCACCTCGGATCAGTACTA GACTGATAGCTTATTGGG	CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGAGCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAAAUCUGGUUACCACUAA ACCAGCCUCAAGAACACCCGAAUGGAGU CUCUAAGCUACAUAAUACCAACUUACAC UUACAAAAUGUUGUCCCCCAAAAUGUAG CCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUUCUUCUCCCCUU GCACCUCGGAUCAGUACUAGACUGAUAG CUUAUUGGGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	
172	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACGCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGAGACGAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKISETK NCALVSATKQKYVNNTA TLLMTSLQYEPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYY

	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CGAGUUCAGGCACAUCCUGAUGCAGGGC	
	CCCTGGACCAAGATCAGCGAGAC	GUGUACUACUAAAUCUGGUUACCACUAA	
	GAAGAACTGCGCCCTGGTGAGCG	ACCAGCCUCAAGAACACCCGAAUGGAGU	
	CCACCAAGCAGAAGTACGTGAAC	CUCUAAGCUACAUAAUACCAACUUACAC	
	AACACCGCCACCCTGCTGATGAC	UUACAAAAUGUUGUCCCCCAAAAUGUAG	
	CAGCCTGCAGTACGCCCCCAGGG	CCAUUCGUAUCUGCUCCUAAUAAAAAGA	
	AGCTGCTGCAGTACATCACCATG	AAGUUUCUUCACAUUUCUUCUCUCCCUU	
	GTGAACAGCCTGACCTACTTCGA	GCACCUCGGAUCAGUACUAGACUGAUAG	
	GGAGCCCAACTACGACGAGTTCA	CUUAUUGGGAAAAAAAAAAAAAAAAAAA	
	GGCACATCCTGATGCAGGGCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TACTACTAAATCTGGTTACCACT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAACCAGCCTCAAGAACACCCGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGGAGTCTCTAAGCTACATAAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCAACTTACACTTACAAAATGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGTCCCCCAAAATGTAGCCATTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTATCTGCTCCTAATAAAAAGAA	AAAAAAAAAAAAA	
	AGTTTCTTCACATTTCTTCTCTC		
	CCTTGCACCTCGGATCAGTACTA		
	GACTGATAGCTTATTGGG		
173	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCCTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKTRHISHVGVITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINVEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PKRSVMLIGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQMDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGACCAGG	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	CACAUCAGCCACGUGGGCGUGAUCACCU	IRWLGGILPWTKISETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSAIKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACGUGGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGACCAGGCACATCAG	AACAGGCUGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYNE
	CCACGTGGGCGTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRRVLMNGVM
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGAUCGGCAUC	
	ATCAACGTGGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUGGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGATCGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATGGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCAGCGAGACGAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCAUCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCGCCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	

	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACAA	
	TCAGGTGGCTGGGCGGCATCCTG	CGAGUUCAGGAGGGUUCUGAUGAACGGC	
	CCCTGGACCAAGATCAGCGAGAC	GUGAUGUAAAUCUGGUUACCACUAAACC	
	GAAGAACTGCGCCCTGGTGAGCG	AGCCUCAAGAACACCCGAAUGGAGUCUC	
	CCATCAAGCAGAAGTACGTGAAC	UAAGCUACAUAAUACCAACUUACACUUA	
	AACACCGCCACCCTGCTGATGAC	CAAAAUGUUGUCCCCCAAAAUGUAGCCA	
	CAGCCTGCAGTACGCCCCCAGGG	UUCGUAUCUGCUCCUAAUAAAAAGAAAG	
	AGCTGCTGCAGTACATCACCATG	UUUCUUCACAUUUCUUCUCUCCCUUGCA	
	GTGAACAGCCTGACCTACTTCGA	CCUCGGAUCAGUACUAGACUGAUAGCUU	
	GGAGCCCAACTACAACGAGTTCA	AUUGGGAAAAAAAAAAAAAAAAAAAAAA	
	GGAGGGTTCTGATGAACGGCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAAATCTGGTTACCACTAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAGCCTCAAGAACACCCGAATG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGTCTCTAAGCTACATAATACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AACTTACACTTACAAAATGTTGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCCAAAATGTAGCCATTCGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTGCTCCTAATAAAAAGAAAGT	AAAAAAAAAA	
	TTCTTCACATTTCTTCTCTCCCT		
	TGCACCTCGGATCAGTACTAGAC		
	TGATAGCTTATTGGG		
174	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MKFQGLVLIDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHNIKHVGLITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAAGUUCCAGGGCCUGGUGCU	FGLYKSINVEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GAUCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAAGTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PERSVMLIGIEILNTIQ
	GCTGATCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQIDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	AACAUCAAGCACGUGGGCCUGAUCACCU	IRWLGGILPWTKISETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NSALVSAAKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACGUGGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACAACATCAA	AACAGGCUGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYDE
	GCACGTGGGCCTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRRVLMNGVM
	AGGCCTTCGGCCTGTACAAGAGC	CGAGAGGAGCGUGAUGCUGAUCGGCAUC	
	ATCAACGTGGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCGAGAG	AUCGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGATCGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATCGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCAGCGAGACGAAGAACAGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCGCCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCGCCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CGAGUUCAGGAGGGUUCUGAUGAACGGC	

	CCCTGGACCAAGATCAGCGAGAC	GUGAUGUAAAUCUGGUUACCACUAAACC	
	GAAGAACAGCGCCCTGGTGAGCG	AGCCUCAAGAACACCCGAAUGGAGUCUC	
	CCGCCAAGCAGAAGTACGTGAAC	UAAGCUACAUAAUACCAACUUACACUUA	
	AACACCGCCACCCTGCTGATGAC	CAAAAUGUUGUCCCCCAAAAUGUAGCCA	
	CAGCCTGCAGTACGCCCCCAGGG	UUCGUAUCUGCUCCUAAUAAAAAGAAAG	
	AGCTGCTGCAGTACATCACCATG	UUUCUUCACAUUUCUUCUCUCCCUUGCA	
	GTGAACAGCCTGACCTACTTCGA	CCUCGGAUCAGUACUAGACUGAUAGCUU	
	GGAGCCCAACTACGACGAGTTCA	AUUGGGAAAAAAAAAAAAAAAAAAAAAA	
	GGAGGGTTCTGATGAACGGCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAAATCTGGTTACCACTAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAGCCTCAAGAACACCCGAATG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGTCTCTAAGCTACATAATACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AACTTACACTTACAAAATGTTGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCCAAAATGTAGCCATTCGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTGCTCCTAATAAAAAGAAAGT	AAAAAAAAAA	
	TTCTTCACATTTCTTCTCTCCCT		
	TGCACCTCGGATCAGTACTAGAC		
	TGATAGCTTATTGGG		
175	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MKFQGLVLIDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHNIKHVGLITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAAGUUCCAGGGCCUGGUGCU	FGLYKSINVEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GAUCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAAGTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PERSVMLIGIEILNTIQ
	GCTGATCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQIDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	AACAUCAAGCACGUGGGCCUGAUCACCU	IRWLGGILPWTKISETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NSALVSAAKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACGUGGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACAACATCAA	AACAGGCUGGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYDE
	GCACGTGGGCCTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRRVLMNGVMKNFC
	AGGCCTTCGGCCTGTACAAGAGC	CGAGAGGAGCGUGAUGCUGAUCGGCAUC	
	ATCAACGTGGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCGAGAG	AUCGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGATCGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATCGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCAGCGAGACGAAGAACAGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCGCCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCGCCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CGAGUUCAGGAGGGUUCUGAUGAACGGC	
	CCCTGGACCAAGATCAGCGAGAC	GUGAUGAAGAACUUCUGCUAAAUCUGGU	
	GAAGAACAGCGCCCTGGTGAGCG	UACCACUAAACCAGCCUCAAGAACACCC	

	CCGCCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACGAGTTCA GGAGGGTTCTGATGAACGGCGTG ATGAAGAACTTCTGCTAAATCTG GTTACCACTAAACCAGCCTCAAG AACACCCGAATGGAGTCTCTAAG CTACATAATACCAACTTACACTT ACAAAATGTTGTCCCCCAAAATG TAGCCATTCGTATCTGCTCCTAA TAAAAAGAAAGTTTCTTCACATT TCTTCTCTCCCTTGCACCTCGGA TCAGTACTAGACTGATAGCTTAT TGGG	GAAUGGAGUCUCUAAGCUACAUAAUACC AACUUACACUUACAAAAUGUUGUCCCCC AAAAUGUAGCCAUUCGUAUCUGCUCCUA AUAAAAAGAAAGUUUCUUCACAUUUCUU CUCUCCCUUGCACCUCGGAUCAGUACUA GACUGAUAGCUUAUUGGGAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
176	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA CCGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATGGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGAGACGCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATGAGCGAGAC GAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCACCGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUGGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGAGACGCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUGAGCGAGACGAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAAGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CGAGUUCAGGAGGGUUCUGAUGAACGGC GUGAUGUAAAUCUGGUUACCACUAAACC AGCCUCAAGAACACCCGAAUGGAGUCUC UAAGCUACAUAAUACCAACUUACACUUA	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCTA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQMDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLETLGYCM IRWLGGILPWTKMSETK NCALVSATKQKYVNNTA TLLMTSLQYAPKELLQY ITMVNSLTYFEEPNYDE FRRVLMNGVM

	AACACCGCCACCCTGCTGATGAC	CAAAAUGUUGUCCCCCAAAAUGUAGCCA	
	CAGCCTGCAGTACGCCCCCAAGG	UUCGUAUCUGCUCCUAAUAAAAAGAAAG	
	AGCTGCTGCAGTACATCACCATG	UUUCUUCACAUUUCUUCUCUCCCUUGCA	
	GTGAACAGCCTGACCTACTTCGA	CCUCGGAUCAGUACUAGACUGAUAGCUU	
	GGAGCCCAACTACGACGAGTTCA	AUUGGGAAAAAAAAAAAAAAAAAAAAAA	
	GGAGGGTTCTGATGAACGGCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAAATCTGGTTACCACTAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAGCCTCAAGAACACCCGAATG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGTCTCTAAGCTACATAATACC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AACTTACACTTACAAAATGTTGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCCAAAATGTAGCCATTCGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTGCTCCTAATAAAAAGAAAGT	AAAAAAAAAA	
	TTCTTCACATTTCTTCTCTCCCT		
	TGCACCTCGGATCAGTACTAGAC		
	TGATAGCTTATTGGG		
177	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHNIKHVGLITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINVEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PKRSVMLIGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQIDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	AACAUCAAGCACGUGGGCCUGAUCACCU	IRWLGGILPWTKISETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSATKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACGUGGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACAACATCAA	AACAGGCUGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYDK
	GCACGTGGGCCTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRHILMQGVYY
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGAUCGGCAUC	
	ATCAACGTGGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUCGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGATCGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATCGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCAGCGAGACGAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCACCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCGCCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CAAGUUCAGGCACAUCCUGAUGCAGGGC	
	CCCTGGACCAAGATCAGCGAGAC	GUGUACUACUAACUAGUGACUGACUAGG	
	GAAGAACTGCGCCCTGGTGAGCG	AUCUGGUUACCACUAAACCAGCCUCAAG	
	CCACCAAGCAGAAGTACGTGAAC	AACACCCGAAUGGAGUCUCUAAGCUACA	
	AACACCGCCACCCTGCTGATGAC	UAAUACCAACUUACACUUACAAAAUGUU	
	CAGCCTGCAGTACGCCCCCAGGG	GUCCCCCAAAAUGUAGCCAUUCGUAUCU	

	AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
178	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGACAGCGGCATCCACAG CACCACCAGCCAGAAGCACAGGG ACTTCGTGGCCGAGCCCATGGGC GAGAAGCCCGTGGGCAGCCTGGC CGGCATCGGCGAGGTGCTGGGCA AGAAGCTGGAGGAGAGGGGCTTC GACAAGGCCTACGTGGTGCTGGG CCAGTTCCTGGTGCTGAAGAAGG ACGAGGACAGGTTCAGGGAGTGG CTGAAGGACACCTGCGGCGCCAA CGCCAAGCAGAGCAGGGACTGCT TCGGCTGCCTGAGGGAGTGGTGC GACGCCTTCCTGTAAATCTGGTT ACCACTAAACCAGCCTCAAGAAC ACCCGAATGGAGTCTCTAAGCTA CATAATACCAACTTACACTTACA AAATGTTGTCCCCCAAAATGTAG CCATTCGTATCTGCTCCTAATAA AAAGAAAGTTTCTTCACATTTCT TCTCTCCCTTGCACCTCGGATCA GTACTAGACTGATAGCTTATTGG G	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGACAGCGGCAUCCACAGCAC CACCAGCCAGAAGCACAGGGACUUCGUG GCCGAGCCCAUGGGCGAGAAGCCCGUGG GCAGCCUGGCCGGCAUCGGCGAGGUGCU GGGCAAGAAGCUGGAGGAGAGGGGCUUC GACAAGGCCUACGUGGUGCUGGGCCAGU UCCUGGUGCUGAAGAAGGACGAGGACAG GUUCAGGGAGUGGCUGAAGGACACCUGC GGCGCCAACGCCAAGCAGAGCAGGGACU GCUUCGGCUGCCUGAGGGAGUGGUGCGA CGCCUUCCUGUAAAUCUGGUUACCACUA AACCAGCCUCAAGAACACCCGAAUGGAG UCUCUAAGCUACAUAAUACCAACUUACA CUUACAAAAUGUUGUCCCCCAAAAUGUA GCCAUUCGUAUCUGCUCCUAAUAAAAAG AAAGUUUCUUCACAUUUCUUCUCUCCCU UGCACCUCGGAUCAGUACUAGACUGAUA GCUUAUUGGGAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	MDSGIHSTTSQKHRDFV AEPMGEKPVGSLAGIGE VLGKKLEERGFDKAYVV LGQFLVLKKDEDRFREW LKDTCGANAKQSRDCFG CLREWCDAFL
179	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGT TCGAGAGGCAGACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGUUCGAGAGGCAG ACCACCAGCCAGAAGCACAGGGACUUCG UGGCCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGAGGUG CUGGGCAAGAAGCUGGAGGAGAGGGGCU UCGACAAGGCCUACGUGGUGCUGGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAUCUGGUUACCAC	MQFERQTTSQKHRDFVA EPMGEKPVGS_AGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL

	GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	UAAACCAGCCUCAAGAACACCCGAAUGG AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
180	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGGGTGAAGGCCGC CCAGGCCGGCAGGCAGAGCAGCG CCAAGAGGCACCTGGCCGAGCAG TTCGCCGTGGGCGAGATCATCAC CGACATGGCCAAGAAGGAGTGGA AGGTGGGCCTGCCCATCGGCCAG GGCGGCTTCGGCTGCATCTACCT GGCCGACATGAACAGCAGCGAGA GCGTGGGCAGCGACGCCCCCTGC GTGGTGAAGGTGGAGCCCAGCGA CAACGGCCCCCTGTTCACCGAGC TGAAGTTCTACCAGAGGGCCGCC AAGCCCGAGCAGATCCAGAAGTG GATCAGGACCAGGAAGCTGAAGT ACCTGGGCGTGCCCAAGTACTGG GGCAGCGGCCTGCACGACAAGAA CGGCAAGAGCTACAGGTTCATGA TCATGGACAGGTTCGGCAGCGAC CTGCAGAAGATCTACGAGGCCAA CGCCAAGAGGTTCAGCAGGAAGA CCGTGCTGCAGCTGAGCCTGAGG ATCCTGGACATCCTGGAGTACAT CCACGAGCACGAGTACGTGCACG GCGACATCAAGGCCAGCAACCTG CTGCTGAACTACAAGAACCCCGA CCAGGTGTACCTGGTGGACTACG GCCTGGCCTACAGGTACTGCCCC GAGGGCGTGCACAAGGAGTACAA GGAGGACCCCAAGAGGTGCCACG ACGGCACCATCGAGTTCACCAGC ATCGACGCCCACAACGGCGTGGC CCCCAGCAGGAGGGGCGACCTGG AGATCCTGGGCTACTGCATGATC CAGTGGCTGACCGGCCACCTGCC CTGGGAGGACAACCTGAAGGACC CCAAGTACGTGAGGGACAGCAAG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGGGUGAAGGCCGCCCA GGCCGGCAGGCAGAGCAGCGCCAAGAGG CACCUGGCCGAGCAGUUCGCCGUGGGCG AGAUCAUCACCGACAUGGCCAAGAAGGA GUGGAAGGUGGGCCUGCCCAUCGGCCAG GGCGGCUUCGGCUGCAUCUACCUGGCCG ACAUGAACAGCAGCGAGAGCGUGGGCAG CGACGCCCCCUGCGUGGUGAAGGUGGAG CCCAGCGACAACGGCCCCCUGUUCACCG AGCUGAAGUUCUACCAGAGGGCCGCCAA GCCCGAGCAGAUCCAGAAGUGGAUCAGG ACCAGGAAGCUGAAGUACCUGGGCGUGC CCAAGUACUGGGGCAGCGGCCUGCACGA CAAGAACGGCAAGAGCUACAGGUUCAUG AUCAUGGACAGGUUCGGCAGCGACCUGC AGAAGAUCUACGAGGCCAACGCCAAGAG GUUCAGCAGGAAGACCGUGCUGCAGCUG AGCCUGAGGAUCCUGGACAUCCUGGAGU ACAUCCACGAGCACGAGUACGUGCACGG CGACAUCAAGGCCAGCAACCUGCUGCUG AACUACAAGAACCCCGACCAGGUGUACC UGGUGGACUACGGCCUGGCCUACAGGUA CUGCCCCGAGGGCGUGCACAAGGAGUAC AAGGAGGACCCCAAGAGGUGCCACGACG GCACCAUCGAGUUCACCAGCAUCGACGC CCACAACGGCGUGGCCCCCAGCAGGAGG GGCGACCUGGAGAUCCUGGGCUACUGCA UGAUCCAGUGGCUGACCGGCCACCUGCC CUGGGAGGACAACCUGAAGGACCCCAAG UACGUGAGGGACAGCAAGAUCAGGUACA GGGAGAACAUCGCCAGCCUGAUGGACAA GUGCUUCCCCGAGAAGAACAAGCCCGGC GAGAUCGCCAAGUACAUGGAGACCGUGA AGCUGCUGGACUACACCGAGAAGCCCCU GUACGAGAACCUGAGGGACAUCCUGCUG CAGGGCCUGAAGGCCAUCGGCAGCAAGG ACGACGGCAAGCUGGACCUGAGCGUGGU GGAGAACGGCGGCCUGAAGGCCAAGACC	MPRVKAAQAGRQSSAKR HLAEQFAVGEIITDMAK KEWKVGLPIGQGGFGCI YLADMNSSESVGSDAPC VVKVEPSDNGPLFTELK FYQRAAKPEQIQKWIRT RKLKYLGVPKYWGSGLH DKNGKSYRFMIMDRFGS DLQKIYEANAKRFSRKT VLQLSLRILDILEYIHE HEYVHGDIKASNLLLNY KNPDQVYLVDYGLAYRY CPEGVHKEYKEDPKRCH DGTIEFTSIDAHNGVAP SRRGDLEILGYCMIQWL TGHLPWEDNLKDPKYVR DSKIRYRENIASLMDKC FPEKNKPGEIAKYMETV KLLDYTEKPLYENLRDI LLQGLKAIGSKDDGKLD LSVVENGGLKAKTITKK RKKEIEESKEPGVEDTE WSNTQTEEAIQT

	ATCAGGTACAGGGAGAACATCGC	AUCACCAAGAAGAGGAAGAAGGAGAUCG	
	CAGCCTGATGGACAAGTGCTTCC	AGGAGAGCAAGGAGCCCGGCGUGGAGGA	
	CCGAGAAGAACAAGCCCGGCGAG	CACCGAGUGGAGCAACACCCAGACCGAG	
	ATCGCCAAGTACATGGAGACCGT	GAGGCCAUCCAGACCUAAAUCUGGUUAC	
	GAAGCTGCTGGACTACACCGAGA	CACUAAACCAGCCUCAAGAACACCCGAA	
	AGCCCCTGTACGAGAACCTGAGG	UGGAGUCUCUAAGCUACAUAAUACCAAC	
	GACATCCTGCTGCAGGGCCTGAA	UUACACUUACAAAAUGUUGUCCCCCAAA	
	GGCCATCGGCAGCAAGGACGACG	AUGUAGCCAUUCGUAUCUGCUCCUAAUA	
	GCAAGCTGGACCTGAGCGTGGTG	AAAAGAAAGUUUCUUCACAUUUCUUCUC	
	GAGAACGGCGGCCTGAAGGCCAA	UCCCUUGCACCUCGGAUCAGUACUAGAC	
	GACCATCACCAAGAAGAGGAAGA	UGAUAGCUUAUUGGGAAAAAAAAAAAAA	
	AGGAGATCGAGGAGAGCAAGGAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCGGCGTGGAGGACACCGAGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGCAACACCCAGACCGAGGAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCATCCAGACCTAAATCTGGTTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCACTAAACCAGCCTCAAGAACA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCGAATGGAGTCTCTAAGCTAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATAATACCAACTTACACTTACAA	AAAAAAAAAAAAAAAAAA	
	AATGTTGTCCCCCAAAATGTAGC		
	CATTCGTATCTGCTCCTAATAAA		
	AAGAAAGTTTCTTCACATTTCTT		
	CTCTCCCTTGCACCTCGGATCAG		
	TACTAGACTGATAGCTTATTGGG		
181	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHHISHVGVITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINTEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGVDLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PKRSVMLVGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGNIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQMDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	CACAUCAGCCACGUGGGCGUGAUCACCU	IRWLGGILPWTKIAETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSATKQKYVNNTT
	TGAAGCCCAGCGTGATCGAGGAG	CAACACCGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACCACATCAG	AACAGGCUGGGCGUGGACCUGGACGCCG	ITMVNSLTYFEEPNYDK
	CCACGTGGGCGTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRHILMQGAYY
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGGUGGGCAUC	
	ATCAACACCGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGTGG	ACGAGCAGGGCUACAGCCACGGCAACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUGGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGGTGGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCAACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATGGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACCCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCGCCGAGACCAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCACCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCACCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	

	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACCCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CAAGUUCAGGCACAUCCUGAUGCAGGGC	
	CCCTGGACCAAGATCGCCGAGAC	GCCUACUACUAAAUCUGGUUACCACUAA	
	CAAGAACTGCGCCCTGGTGAGCG	ACCAGCCUCAAGAACACCCGAAUGGAGU	
	CCACCAAGCAGAAGTACGTGAAC	CUCUAAGCUACAUAAUACCAACUUACAC	
	AACACCACCACCCTGCTGATGAC	UUACAAAAUGUUGUCCCCCAAAAUGUAG	
	CAGCCTGCAGTACGCCCCCAGGG	CCAUUCGUAUCUGCUCCUAAUAAAAAGA	
	AGCTGCTGCAGTACATCACCATG	AAGUUUCUUCACAUUUCUUCUCUCCCUU	
	GTGAACAGCCTGACCTACTTCGA	GCACCUCGGAUCAGUACUAGACUGAUAG	
	GGAGCCCAACTACGACAAGTTCA	CUUAUUGGGAAAAAAAAAAAAAAAAAAA	
	GGCACATCCTGATGCAGGGCGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TACTACTAAATCTGGTTACCACT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAACCAGCCTCAAGAACACCCGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGGAGTCTCTAAGCTACATAAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCAACTTACACTTACAAAATGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGTCCCCCAAAATGTAGCCATTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTATCTGCTCCTAATAAAAAGAA	AAAAAAAAAAAAA	
	AGTTTCTTCACATTTCTTCTCTC		
	CCTTGCACCTCGGATCAGTACTA		
	GACTGATAGCTTATTGGG		
182	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHHISHVGVITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINTEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGVDLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PKRSVMLVGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQMDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMYNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	CACAUCAGCCACGUGGGCGUGAUCACCU	IRWLGGILPWTKIAETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSATKQKYVNNTT
	TGAAGCCCAGCGTGATCGAGGAG	CAACACCGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACCACATCAG	AACAGGCUGGGCGUGGACCUGGACGCCG	ITMVNSLTYFEEPNYDK
	CCACGTGGGCGTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRHILMQGAYY
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGGUGGGCAUC	
	ATCAACACCGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGTGG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUGGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGGTGGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGUA	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATGGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGT	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	ACAACGGCGAGCACGTGCCCTTC	CAAGAUCGCCGAGACGAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCACCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCACCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	

	TCAGGTGGCTGGGCGGCATCCTG	CAAGUUCAGGCACAUCCUGAUGCAGGGC	
	CCCTGGACCAAGATCGCCGAGAC	GCCUACUACUAAAUCUGGUUACCACUAA	
	GAAGAACTGCGCCCTGGTGAGCG	ACCAGCCUCAAGAACACCCGAAUGGAGU	
	CCACCAAGCAGAAGTACGTGAAC	CUCUAAGCUACAUAAUACCAACUUACAC	
	AACACCACCACCCTGCTGATGAC	UUACAAAAUGUUGUCCCCCAAAAUGUAG	
	CAGCCTGCAGTACGCCCCCAGGG	CCAUUCGUAUCUGCUCCUAAUAAAAAGA	
	AGCTGCTGCAGTACATCACCATG	AAGUUUCUUCACAUUUCUUCUCUCCCUU	
	GTGAACAGCCTGACCTACTTCGA	GCACCUCGGAUCAGUACUAGACUGAUAG	
	GGAGCCCAACTACGACAAGTTCA	CUUAUUGGGAAAAAAAAAAAAAAAAAAA	
	GGCACATCCTGATGCAGGGCGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TACTACTAAATCTGGTTACCACT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAACCAGCCTCAAGAACACCCGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGGAGTCTCTAAGCTACATAAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCAACTTACACTTACAAAATGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGTCCCCCAAAATGTAGCCATTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTATCTGCTCCTAATAAAAAGAA	AAAAAAAAAAAAA	
	AGTTTCTTCACATTTCTTCTCTC		
	CCTTGCACCTCGGATCAGTACTA		
	GACTGATAGCTTATTGGG		
183	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLIGKGGFGSIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	NNNYVVKIEPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHHISHVGVITCKA
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINTEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGAUCGGCAAGGGCGGCU	PKRSVMLVGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGGCAGCAUCUACACCACCAACAACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGATC	CAACUACGUGGUGAAGAUCGAGCCCAAG	VLDQMDKNKLYLVDYGL
	GGCAAGGGCGGCTTCGGCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACAACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGTGAAGATCGAGCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLETLGYCM
	GCCAACGGCAGCCTGTTCACCGA	CACAUCAGCCACGUGGGCGUGAUCACCU	IRWLGGILPWTKIAETK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSATKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACACCGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACCACATCAG	AACAGGCUGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYDK
	CCACGTGGGCGTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRHILMQGAYY
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGGUGGGCAUC	
	ATCAACACCGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUGGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGGTGGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATGGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGAGACGCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCGCCGAGACGAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCACCAAGCAGAAGUACG	
	CAACGGCACCCTGGAGTTCACCC	UGAACAACACCGCCACCCUGCUGAUGAC	
	CCATCGACAGCCACAAGGGCTAC	CAGCCUGCAGUACGCCCCCAGGGAGCUG	
	GTGGTGAGCAGGAGGGGCGACCT	CUGCAGUACAUCACCAUGGUGAACAGCC	
	GGAGACGCTGGGCTACTGCATGA	UGACCUACUUCGAGGAGCCCAACUACGA	
	TCAGGTGGCTGGGCGGCATCCTG	CAAGUUCAGGCACAUCCUGAUGCAGGGC	
	CCCTGGACCAAGATCGCCGAGAC	GCCUACUACUAAAUCUGGUUACCACUAA	

	GAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGCC TACTACTAAATCTGGTTACCACT AAACCAGCCTCAAGAACACCCGA ATGGAGTCTCTAAGCTACATAAT ACCAACTTACACTTACAAAATGT TGTCCCCCAAAATGTAGCCATTC GTATCTGCTCCTAATAAAAAGAA AGTTTCTTCACATTTCTTCTCTC CCTTGCACCTCGGATCAGTACTA GACTGATAGCTTATTGGG	ACCAGCCUCAAGAACACCCGAAUGGAGU CUCUAAGCUACAUAAUACCAACUUACAC UUACAAAAUGUUGUCCCCCAAAAUGUAG CCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUUCUUCUCUCCCUU GCACCUCGGAUCAGUACUAGACUGAUAG CUUAUUGGGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAA	
184	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAAGTTCGAGAGGCAGAA CTTCCAGGGCCTGGTGCTGACCG ACAACTGCAAGAACCAGTGGGTG GTGGGCCCCCTGATCGGCAAGGG CGGCTTCGGCAGCATCTACACCA CCAACGACAACAACTACGTGGTG AAGATCGAGCCCAAGGCCAACGG CAGCCTGTTCACCGAGCAGGCCT TCTACACCAGGGTGCTGAAGCCC AGCGTGATCGAGGAGTGGAAGAA GAGCCACAACATCAAGCACGTGG GCCTGATCACCTGCAAGGCCTTC GGCCTGTACAAGAGCATCAACGT GGAGTACAGGTTCCTGGTGATCA ACAGGCTGGGCGCCGACCTGGAC GCCGTGATCAGGGCCAACAACAA CAGGCTGCCCAAGAGGAGCGTGA TGCTGATCGGCATCGAGATCCTG AACACCATCCAGTTCATGCACGA GCAGGGCTACAGCCACGGCGACA TCAAGGCCAGCAACATCGTGCTG GACCAGATCGACAAGAACAAGCT GTACCTGGTGGACTACGGCCTGG TGAGCAAGTTCATGAGCAACGGC GAGCACGTGCCCTTCATCAGGAA CCCCAACAAGATGGACAACGGCA CCCTGGAGTTCACCCCCATCGAC AGCCACAAGGGCTACGTGGTGAG CAGGAGGGGCGACCTGGACACCC TGGGCTACTGCATGATCAGGTGG CTGGGCGGCATCCTGCCCTGGAC CAAGATCAGCGACACCAAGAACT GCGCCCTGGTGAGCGCCACCAAG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAAGUUCGAGAGGCAGAACUU CCAGGGCCUGGUGCUGACCGACAACUGC AAGAACCAGUGGGUGGUGGGCCCCCUGA UCGGCAAGGGCGGCUUCGGCAGCAUCUA CACCACCAACGACAACAACUACGUGGUG AAGAUCGAGCCCAAGGCCAACGGCAGCC UGUUCACCGAGCAGGCCUUCUACACCAG GGUGCUGAAGCCCAGCGUGAUCGAGGAG UGGAAGAAGAGCCACAACAUCAAGCACG UGGGCCUGAUCACCUGCAAGGCCUUCGG CCUGUACAAGAGCAUCAACGUGGAGUAC AGGUUCCUGGUGAUCAACAGGCUGGGCG CCGACCUGGACGCCGUGAUCAGGGCCAA CAACAACAGGCUGCCCAAGAGGAGCGUG AUGCUGAUCGGCAUCGAGAUCCUGAACA CCAUCCAGUUCAUGCACGAGCAGGGCUA CAGCCACGGCGACAUCAAGGCCAGCAAC AUCGUGCUGGACCAGAUCGACAAGAACA AGCUGUACCUGGUGGACUACGGCCUGGU GAGCAAGUUCAUGAGCAACGGCGAGCAC GUGCCCUUCAUCAGGAACCCCAACAAGA UGGACAACGGCACCCUGGAGUUCACCCC CAUCGACAGCCACAAGGGCUACGUGGUG AGCAGGAGGGGCGACCUGGACACCCUGG GCUACUGCAUGAUCAGGUGGCUGGGCGG CAUCCUGCCCUGGACCAAGAUCAGCGAC ACCAAGAACUGCGCCCUGGUGAGCGCCA CCAAGCAGAAGUACGUGAACAACACCGC CACCCUGCUGAUGACCAGCCUGCAGUAC GCCCCCAGGGAGCUGCUGCAGUACAUCA CCAUGGUGAACAGCCUGACCUACUUCGA GGAGCCCAACUACGACAAGUUCAGGCAC AUCCUGAUGCAGGGCGUGUACUACUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MKFERQNFQGLVLTDNC KNQWVVGPLIGKGGFGS IYTTNDNNYVVKIEPKA NGSLFTEQAFYTRVLKP SVIEEWKKSHNIKHVGL ITCKAFGLYKSINVEYR FLVINRLGADLDAVIRA NNNRLPKRSVMLIGIEI LNTIQFMHEQGYSHGDI KASNIVLDQIDKNKLYL VDYGLVSKFMSNGEHVP FIRNPNKMDNGTLEFTP IDSHKGYVVSRRGDLDT LGYCMIRWLGGILPWTK ISDTKNCALVSATKQKY VNNTATLLMTSLQYAPR ELLQYITMVNSLTYFEE PNYDKFRHILMQGVYY

	CAGAAGTACGTGAACAACACCGC CACCCTGCTGATGACCAGCCTGC AGTACGCCCCCAGGGAGCTGCTG CAGTACATCACCATGGTGAACAG CCTGACCTACTTCGAGGAGCCCA ACTACGACAAGTTCAGGCACATC CTGATGCAGGGCGTGTACTACTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
185	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGACAGCGGCATCCACAG CAACTTCCAGGGCCTGGTGCTGA CCGACAACTGCAAGAACCAGTGG GTGGTGGGCCCCCTGATCGGCAA GGGCGGCTTCGGCAGCATCTACA CCACCAACGACAACAACTACGTG GTGAAGATCGAGCCCAAGGCCAA CGGCAGCCTGTTCACCGAGCAGG CCTTCTACACCAGGGTGCTGAAG CCCAGCGTGATCGAGGAGTGGAA GAAGAGCCACAACATCAAGCACG TGGGCCTGATCACCTGCAAGGCC TTCGGCCTGTACAAGAGCATCAA CGTGGAGTACAGGTTCCTGGTGA TCAACAGGCTGGGCGCCGACCTG GACGCCGTGATCAGGGCCAACAA CAACAGGCTGCCCAAGAGGAGCG TGATGCTGATCGGCATCGAGATC CTGAACACCATCCAGTTCATGCA CGAGCAGGGCTACAGCCACGGCG ACATCAAGGCCAGCAACATCGTG CTGGACCAGATCGACAAGAACAA GCTGTACCTGGTGGACTACGGCC TGGTGAGCAAGTTCATGAGCAAC GGCGAGCACGTGCCCTTCATCAG GAACCCCAACAAGATGGACAACG GCACCCTGGAGTTCACCCCCATC GACAGCCACAAGGGCTACGTGGT GAGCAGGAGGGGCGACCTGGACA CCCTGGGCTACTGCATGATCAGG TGGCTGGGCGGCATCCTGCCCTG GACCAAGATCAGCGACACCAAGA ACTGCGCCCTGGTGAGCGCCACC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGACAGCGGCAUCCACAGCAA CUUCCAGGGCCUGGUGCUGACCGACAAC UGCAAGAACCAGUGGGUGGUGGGCCCCC UGAUCGGCAAGGGCGGCUUCGGCAGCAU CUACACCACCAACGACAACAACUACGUG GUGAAGAUCGAGCCCAAGGCCAACGGCA GCCUGUUCACCGAGCAGGCCUUCUACAC CAGGGUGCUGAAGCCCAGCGUGAUCGAG GAGUGGAAGAAGAGCCACAACAUCAAGC ACGUGGGCCUGAUCACCUGCAAGGCCUU CGGCCUGUACAAGAGCAUCAACGUGGAG UACAGGUUCCUGGUGAUCAACAGGCUGG GCGCCGACCUGGACGCCGUGAUCAGGGC CAACAACAACAGGCUGCCCAAGAGGAGC GUGAUGCUGAUCGGCAUCGAGAUCCUGA ACACCAUCCAGUUCAUGCACGAGCAGGG CUACAGCCACGGCGACAUCAAGGCCAGC AACAUCGUGCUGGACCAGAUCGACAAGA ACAAGCUGUACCUGGUGGACUACGGCCU GGUGAGCAAGUUCAUGAGCAACGGCGAG CACGUGCCCUUCAUCAGGAACCCCAACA AGAUGGACAACGGCACCCUGGAGUUCAC CCCCAUCGACAGCCACAAGGGCUACGUG GUGAGCAGGAGGGGCGACCUGGACACCC UGGGCUACUGCAUGAUCAGGUGGCUGGG CGGCAUCCUGCCCUGGACCAAGAUCAGC GACACCAAGAACUGCGCCCUGGUGAGCG CCACCAAGCAGAAGUACGUGAACAACAC CGCCACCCUGCUGAUGACCAGCCUGCAG UACGCCCCCAGGGAGCUGCUGCAGUACA UCACCAUGGUGAACAGCCUGACCUACUU CGAGGAGCCCAACUACGACAAGUUCAGG CACAUCCUGAUGCAGGGCGUGUACUACU AACUAGUGACUGACUAGGAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA	MDSGIHSNFQGLVLTDN CKNQWVVGPLIGKGGFG SIYTTNDNNYVVKIEPK ANGSLFTEQAFYTRVLK PSVIEEWKKSHNIKHVG LITCKAFGLYKSINVEY RFLVINRLGADLDAVIR ANNNRLPKRSVMLIGIE ILNTIQFMHEQGYSHGD IKASNIVLDQIDKNKLY LVDYGLVSKFMSNGEHV PFIRNPNKMDNGTLEFT PIDSHKGYVVSRRGDLD TLGYCMIRWLGGILPWT KISDTKNCALVSATKQK YVNNTATLLMTSLQYAP RELLQYITMVNSLTYFE EPNYDKFRHILMQGVYY

	AAGCAGAAGTACGTGAACAACAC CGCCACCCTGCTGATGACCAGCC TGCAGTACGCCCCCAGGGAGCTG CTGCAGTACATCACCATGGTGAA CAGCCTGACCTACTTCGAGGAGC CCAACTACGACAAGTTCAGGCAC ATCCTGATGCAGGGCGTGTACTA CTAACTAGTGACTGACTAGGATC TGGTTACCACTAAACCAGCCTCA AGAACACCCGAATGGAGTCTCTA AGCTACATAATACCAACTTACAC TTACAAAATGTTGTCCCCCAAAA TGTAGCCATTCGTATCTGCTCCT AATAAAAAGAAAGTTTCTTCACA TTTCTTCTCTCCCTTGCACCTCG GATCAGTACTAGACTGATAGCTT ATTGGG	UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	
186	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCTGCCCACCCCCCCCAA CTTCCAGGGCCTGGTGCTGACCG ACAACTGCAAGAACCAGTGGGTG GTGGGCCCCCTGATCGGCAAGGG CGGCTTCGGCAGCATCTACACCA CCAACGACAACAACTACGTGGTG AAGATCGAGCCCAAGGCCAACGG CAGCCTGTTCACCGAGCAGGCCT TCTACACCAGGGTGCTGAAGCCC AGCGTGATCGAGGAGTGGAAGAA GAGCCACAACATCAAGCACGTGG GCCTGATCACCTGCAAGGCCTTC GGCCTGTACAAGAGCATCAACGT GGAGTACAGGTTCCTGGTGATCA ACAGGCTGGGCGCCGACCTGGAC GCCGTGATCAGGGCCAACAACAA CAGGCTGCCCAAGAGGAGCGTGA TGCTGATCGGCATCGAGATCCTG AACACCATCCAGTTCATGCACGA GCAGGGCTACAGCCACGGCGACA TCAAGGCCAGCAACATCGTGCTG GACCAGATCGACAAGAACAAGCT GTACCTGGTGGACTACGGCCTGG TGAGCAAGTTCATGAGCAACGGC GAGCACGTGCCCTTCATCAGGAA CCCCAACAAGATGGACAACGGCA CCCTGGAGTTCACCCCCATCGAC AGCCACAAGGGCTACGTGGTGAG CAGGAGGGGCGACCTGGACACCC TGGGCTACTGCATGATCAGGTGG CTGGGCGGCATCCTGCCCTGGAC CAAGATCAGCGACACCAAGAACT GCGCCCTGGTGAGCGCCACCAAG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCUGCCCACCCCCCCCAACUU CCAGGGCCUGGUGCUGACCGACAACUGC AAGAACCAGUGGGUGGUGGGCCCCCUGA UCGGCAAGGGCGGCUUCGGCAGCAUCUA CACCACCAACGACAACAACUACGUGGUG AAGAUCGAGCCCAAGGCCAACGGCAGCC UGUUCACCGAGCAGGCCUUCUACACCAG GGUGCUGAAGCCCAGCGUGAUCGAGGAG UGGAAGAAGAGCCACAACAUCAAGCACG UGGGCCUGAUCACCUGCAAGGCCUUCGG CCUGUACAAGAGCAUCAACGUGGAGUAC AGGUUCCUGGUGAUCAACAGGCUGGGCG CCGACCUGGACGCCGUGAUCAGGGCCAA CAACAACAGGCUGCCCAAGAGGAGCGUG AUGCUGAUCGGCAUCGAGAUCCUGAACA CCAUCCAGUUCAUGCACGAGCAGGGCUA CAGCCACGGCGACAUCAAGGCCAGCAAC AUCGUGCUGGACCAGAUCGACAAGAACA AGCUGUACCUGGUGGACUACGGCCUGGU GAGCAAGUUCAUGAGCAACGGCGAGCAC GUGCCCUUCAUCAGGAACCCCAACAAGA UGGACAACGGCACCCUGGAGUUCACCCC CAUCGACAGCCACAAGGGCUACGUGGUG AGCAGGAGGGGCGACCUGGACACCCUGG GCUACUGCAUGAUCAGGUGGCUGGGCGG CAUCCUGCCCUGGACCAAGAUCAGCGAC ACCAAGAACUGCGCCCUGGUGAGCGCCA CCAAGCAGAAGUACGUGAACAACACCGC CACCCUGCUGAUGACCAGCCUGCAGUAC GCCCCCAGGGAGCUGCUGCAGUACAUCA CCAUGGUGAACAGCCUGACCUACUUCGA GGAGCCCAACUACGACAAGUUCAGGCAC AUCCUGAUGCAGGGCGUGUACUACUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MLPTPPNFQGLVLTDNC KNQWVVGPLIGKGGFGS IYTTNDNNYVVKIEPKA NGSLFTEQAFYTRVLKP SVIEEWKKSHNIKHVGL ITCKAFGLYKSINVEYR FLVINRLGADLDAVIRA NNNRLPKRSVMLIGIEI LNTIQFMHEQGYSHGDI KASNIVLDQIDKNKLYL VDYGLVSKFMSNGEHVP FIRNPNKMDNGTLEFTP IDSHKGYVVSRRGDLDT LGYCMLRWLGGLLPWTK ISDTKNCALVSATKQKY VNNTATLLMTSLQYAPR ELLQYITMVNSLTYFEE PNYDKFRHILMQGVYY

	CAGAAGTACGTGAACAACACCGC CACCCTGCTGATGACCAGCCTGC AGTACGCCCCCAGGGAGCTGCTG CAGTACATCACCATGGTGAACAG CCTGACCTACTTCGAGGAGCCCA ACTACGACAAGTTCAGGCACATC CTGATGCAGGGCGTGTACTACTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
187	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCAGTTCGAGAGGCAGAA CTTCCAGGGCCTGGTGCTGACCG ACAACTGCAAGAACCAGTGGGTG GTGGGCCCCCTGATCGGCAAGGG CGGCTTCGGCAGCATCTACACCA CCAACGACAACAACTACGTGGTG AAGATCGAGCCCAAGGCCAACGG CAGCCTGTTCACCGAGCAGGCCT TCTACACCAGGGTGCTGAAGCCC AGCGTGATCGAGGAGTGGAAGAA GAGCCACAACATCAAGCACGTGG GCCTGATCACCTGCAAGGCCTTC GGCCTGTACAAGAGCATCAACGT GGAGTACAGGTTCCTGGTGATCA ACAGGCTGGGCGCCGACCTGGAC GCCGTGATCAGGGCCAACAACAA CAGGCTGCCCAAGAGGAGCGTGA TGCTGATCGGCATCGAGATCCTG AACACCATCCAGTTCATGCACGA GCAGGGCTACAGCCACGGCGACA TCAAGGCCAGCAACATCGTGCTG GACCAGATCGACAAGAACAAGCT GTACCTGGTGGACTACGGCCTGG TGAGCAAGTTCATGAGCAACGGC GAGCACGTGCCCTTCATCAGGAA CCCCAACAAGATGGACAACGGCA CCCTGGAGTTCACCCCCATCGAC AGCCACAAGGGCTACGTGGTGAG CAGGAGGGGCGACCTGGACACCC TGGGCTACTGCATGATCAGGTGG CTGGGCGGCATCCTGCCCTGGAC CAAGATCAGCGACACCAAGAACT GCGCCCTGGTGAGCGCCACCAAG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCAGUUCGAGAGGCAGAACUU CCAGGGCCUGGUGCUGACCGACAACUGC AAGAACCAGUGGGUGGUGGGCCCCCUGA UCGGCAAGGGCGGCUUCGGCAGCAUCUA CACCACCAACGACAACAACUACGUGGUG AAGAUCGAGCCCAAGGCCAACGGCAGCC UGUUCACCGAGCAGGCCUUCUACACCAG GGUGCUGAAGCCCAGCGUGAUCGAGGAG UGGAAGAAGAGCCACAACAUCAAGCACG UGGGCCUGAUCACCUGCAAGGCCUUCGG CCUGUACAAGAGCAUCAACGUGGAGUAC AGGUUCCUGGUGAUCAACAGGCUGGGCG CCGACCUGGACGCCGUGAUCAGGGCCAA CAACAACAGGCUGCCCAAGAGGAGCGUG AUGCUGAUCGGCAUCGAGAUCCUGAACA CCAUCCAGUUCAUGCACGAGCAGGGCUA CAGCCACGGCGACAUCAAGGCCAGCAAC AUCGUGCUGGACCAGAUCGACAAGAACA AGCUGUACCUGGUGGACUACGGCCUGGU GAGCAAGUUCAUGAGCAACGGCGAGCAC GUGCCCUUCAUCAGGAACCCCAACAAGA UGGACAACGGCACCCUGGAGUUCACCCC CAUCGACAGCCACAAGGGCUACGUGGUG AGCAGGAGGGGCGACCUGGACACCCUGG GCUACUGCAUGAUCAGGUGGCUGGGCGG CAUCCUGCCCUGGACCAAGAUCAGCGAC ACCAAGAACUGCGCCCUGGUGAGCGCCA CCAAGCAGAAGUACGUGAACAACACCGC CACCCUGCUGAUGACCAGCCUGCAGUAC GCCCCCAGGGAGCUGCUGCAGUACAUCA CCAUGGUGAACAGCCUGACCUACUUCGA GGAGCCCAACUACGACAAGUUCAGGCAC AUCCUGAUGCAGGGCGUGUACUACUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MQFERQNFQGLVLTDNC KNQWVVGPLIGKGGFGS IYTTNDNNYVVKIEPKA NGSLFTEQAFYTRVLKP SVIEEWKKSHNIKHVGL ITCKAFGLYKSINVEYR FLVINRLGADLDAVIRA NNNRLPKRSVMLIGIEI LNTIQFMHEQGYSHGDI KASNIVLDQIDKNKLYL VDYGLVSKFMSNGEHVP FIRNPNKMDNGTLEFTP IDSHKGYVVSRRGDLDT LGYCMIRWLGGILPWTK ISDTKNCALVSATKQKY VNNTATLLMTSLQYAPR ELLQYITMVNSLTYFEE PNYDKFRHILMQGVYY

	CAGAAGTACGTGAACAACACCGC CACCCTGCTGATGACCAGCCTGC AGTACGCCCCCAGGGAGCTGCTG CAGTACATCACCATGGTGAACAG CCTGACCTACTTCGAGGAGCCCA ACTACGACAAGTTCAGGCACATC CTGATGCAGGGCGTGTACTACTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
188	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACAAGUUCGAGAGGCAGUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYKFERQ

	AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACAAGTTCGAGAGGCAGTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
189	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACGACAGCGGCAUCCACAGCU AACUAGUGACUGACUAGGAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYDSGIHS

	AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACGACAGCGGCATCCACAG CTAACTAGTGACTGACTAGGATC TGGTTACCACTAAACCAGCCTCA AGAACACCCGAATGGAGTCTCTA AGCTACATAATACCAACTTACAC TTACAAAATGTTGTCCCCCAAAA TGTAGCCATTCGTATCTGCTCCT AATAAAAAGAAAGTTTCTTCACA TTTCTTCTCTCCCTTGCACCTCG GATCAGTACTAGACTGATAGCTT ATTGGG	UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	
190	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACCUGCCCACCCCCCCCUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYLPTPP

	AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACCTGCCCACCCCCCCCTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
191	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACAGCCA CGGCGACATCAAGGCCAGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGTGAGCAAGTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACAGCCACGGCGACAU CAAGGCCAGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGUGAGCAAGUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACCAGUUCGAGAGGCAGUAAC UAGUGACUGACUAGGAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYSHGDIKASNI VLDQIDKNKLYLVDYGL VSKFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYYLPTPP

	AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACCAGTTCGAGAGGCAGTA ACTAGTGACTGACTAGGATCTGG TTACCACTAAACCAGCCTCAAGA ACACCCGAATGGAGTCTCTAAGC TACATAATACCAACTTACACTTA CAAAATGTTGTCCCCCAAAATGT AGCCATTCGTATCTGCTCCTAAT AAAAAGAAAGTTTCTTCACATTT CTTCTCTCCCTTGCACCTCGGAT CAGTACTAGACTGATAGCTTATT GGG	AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
192	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGAAGT TCGAGAGGCAGACCACCAGCCAG AAGCACAGGGACTTCGTGACCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAAGUUCGAGAGGCAG ACCACCAGCCAGAAGCACAGGGACUUCG UGACCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGAGGUG CUGGGCAAGAAGCUGGAGGAGAGGGGCU UCGACAAGGCCUACGUGGUGCUGGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	MKFERQTTSQKHRDFVT EPMGEKPVGSLAGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL
193	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGACAGCGGCATCCACAG CACCACCAGCCAGAAGCACAGGG ACTTCGTGACCGAGCCCATGGGC GAGAAGCCCGTGGGCAGCCTGGC CGGCATCGGCGAGGTGCTGGGCA AGAAGCTGGAGGAGAGGGGCTTC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGACAGCGGCAUCCACAGCAC CACCAGCCAGAAGCACAGGGACUUCGUG ACCGAGCCCAUGGGCGAGAAGCCCGUGG GCAGCCUGGCCGGCAUCGGCGAGGUGCU GGGCAAGAAGCUGGAGGAGAGGGGCUUC GACAAGGCCUACGUGGUGCUGGGCCAGU UCCUGGUGCUGAAGAAGGACGAGGACAG GUUCAGGGAGUGGCUGAAGGACACCUGC	MDSGIHSTTSQKHRDFV TEPMGEKPVGSLAGIGE VLGKKLEERGFDKAYVV LGQFLVLKKDEDRFREW LKDTCGANAKQSRDCFG CLREWCDAFL

	GACAAGGCCTACGTGGTGCTGGG CCAGTTCCTGGTGCTGAAGAAGG ACGAGGACAGGTTCAGGGAGTGG CTGAAGGACACCTGCGGCGCCAA CGCCAAGCAGAGCAGGGACTGCT TCGGCTGCCTGAGGGAGTGGTGC GACGCCTTCCTGTAAATCTGGTT ACCACTAAACCAGCCTCAAGAAC ACCCGAATGGAGTCTCTAAGCTA CATAATACCAACTTACACTTACA AAATGTTGTCCCCCAAAATGTAG CCATTCGTATCTGCTCCTAATAA AAAGAAAGTTTCTTCACATTTCT TCTCTCCCTTGCACCTCGGATCA GTACTAGACTGATAGCTTATTGG G	GGCGCCAACGCCAAGCAGAGCAGGGACU GCUUCGGCUGCCUGAGGGAGUGGUGCGA CGCCUUCCUGUAAAUCUGGUUACCACUA AACCAGCCUCAAGAACACCCGAAUGGAG UCUCUAAGCUACAUAAUACCAACUUACA CUUACAAAAUGUUGUCCCCCAAAAUGUA GCCAUUCGUAUCUGCUCCUAAUAAAAAG AAAGUUUCUUCACAUUUCUUCUCUCCCU UGCACCUCGGAUCAGUACUAGACUGAUA GCUUAUUGGGAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	
194	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCTGGACAGCGGCATCCA CAGCGGCGCCACCACCACCGCCC CCAGCCTGAGCACCACCAGCCAG AAGCACAGGGACTTCGTGACCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCUGGACAGCGGCAUCCACAG CGGCGCCACCACCACCGCCCCCAGCCUG AGCACCACCAGCCAGAAGCACAGGGACU UCGUGACCGAGCCCAUGGGCGAGAAGCC CGUGGGCAGCCUGGCCGGCAUCGGCGAG GUGCUGGGCAAGAAGCUGGAGGAGAGGG GCUUCGACAAGGCCUACGUGGUGCUGGG CCAGUUCCUGGUGCUGAAGAAGGACGAG GACAGGUUCAGGGAGUGGCUGAAGGACA CCUGCGGCGCCAACGCCAAGCAGAGCAG GGACUGCUUCGGCUGCCUGAGGGAGUGG UGCGACGCCUUCCUGUAAAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MLDSGIHSGATTTAPSL STTSQKHRDFVTEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
195	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCTGCCCACCCCCCCCAC CACCAGCCAGAAGCACAGGGACT	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCUGCCCACCCCCCCCACCAC CAGCCAGAAGCACAGGGACUUCGUGACC GAGCCCAUGGGCGAGAAGCCCGUGGGCA GCCUGGCCGGCAUCGGCGAGGUGCUGGG	MLPTPPTTSQKHRDFVT EPMGEKPVGSLAGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL

	TCGTGACCGAGCCCATGGGCGAG AAGCCCGTGGGCAGCCTGGCCGG CATCGGCGAGGTGCTGGGCAAGA AGCTGGAGGAGAGGGGCTTCGAC AAGGCCTACGTGGTGCTGGGCCA GTTCCTGGTGCTGAAGAAGGACG AGGACAGGTTCAGGGAGTGGCTG AAGGACACCTGCGGCGCCAACGC CAAGCAGAGCAGGGACTGCTTCG GCTGCCTGAGGGAGTGGTGCGAC GCCTTCCTGTAAATCTGGTTACC ACTAAACCAGCCTCAAGAACACC CGAATGGAGTCTCTAAGCTACAT AATACCAACTTACACTTACAAAA TGTTGTCCCCCAAAATGTAGCCA TTCGTATCTGCTCCTAATAAAAA GAAAGTTTCTTCACATTTCTTCT CTCCCTTGCACCTCGGATCAGTA CTAGACTGATAGCTTATTGGG	CAAGAAGCUGGAGGAGAGGGGCUUCGAC AAGGCCUACGUGGUGCUGGGCCAGUUCC UGGUGCUGAAGAAGGACGAGGACAGGUU CAGGGAGUGGCUGAAGGACACCUGCGGC GCCAACGCCAAGCAGAGCAGGGACUGCU UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUAAAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAAUACCAACUUACACUU ACAAAAUGUUGUCCCCCAAAAUGUAGCC AUUCGUAUCUGCUCCUAAUAAAAAGAAA GUUUCUUCACAUUUCUUCUCCCCUUGC ACCUCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAA	
196	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGT TCGAGAGGCAGACCACCAGCCAG AAGCACAGGGACTTCGTGACCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGUUCGAGAGGCAG ACCACCAGCCAGAAGCACAGGGACUUCG UGACCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGAGGUG CUGGGCAAGAAGCUGGAGGAGAGGGGCU UCGACAAGGCCUACGUGGUGCUGGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	MQFERQTTSQKHRDFVT EPMGEKPVGS_AGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL
197	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGAAGT TCGAGAGGCAGACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGAAGUUCGAGAGGCAG ACCACCAGCCAGAAGCACAGGGACUUCG UGGCCGAGCCCAUGGGCGAGAAGCCCGU GGGCAGCCUGGCCGGCAUCGGCGAGGUG CUGGGCAAGAAGCUGGAGGAGAGGGGCU UCGACAAGGCCUACGUGGUGCUGGGCCA GUUCCUGGUGCUGAAGAAGGACGAGGAC	MKFERQTTSQKHRDFVA EPMGEKPVGS_AGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDRFREWL KDTCGANAKQSRDCFGC LREWCDAFL

	GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	AGGUUCAGGGAGUGGCUGAAGGACACCU GCGGCGCCAACGCCAAGCAGAGCAGGGA CUGCUUCGGCUGCCUGAGGGAGUGGUGC GACGCCUUCCUGUAAAUCUGGUUACCAC UAAACCAGCCUCAAGAACACCCGAAUGG AGUCUCUAAGCUACAUAAUACCAACUUA CACUUACAAAAUGUUGUCCCCCAAAAUG UAGCCAUUCGUAUCUGCUCCUAAUAAAA AGAAAGUUUCUUCACAUUUCUUCUCUCC CUUGCACCUCGGAUCAGUACUAGACUGA UAGCUUAUUGGGAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
198	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGACAGCGGCATCCACAG CGACAGCGGCATCCACAGCGACA GCGGCATCCACAGCACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GAGGTGCTGGGCAAGAAGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACAG GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAATCTGGTTACCACTAAAC CAGCCTCAAGAACACCCGAATGG AGTCTCTAAGCTACATAATACCA ACTTACACTTACAAAATGTTGTC CCCCAAAATGTAGCCATTCGTAT CTGCTCCTAATAAAAAGAAAGTT TCTTCACATTTCTTCTCTCCCTT GCACCTCGGATCAGTACTAGACT GATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGACAGCGGCAUCCACAGCGA CAGCGGCAUCCACAGCGACAGCGGCAUC CACAGCACCACCAGCCAGAAGCACAGGG ACUUCGUGGCCGAGCCCAUGGGCGAGAA GCCCGUGGGCAGCCUGGCCGGCAUCGGC GAGGUGCUGGGCAAGAAGCUGGAGGAGA GGGGCUUCGACAAGGCCUACGUGGUGCU GGGCCAGUUCCUGGUGCUGAAGAAGGAC GAGGACAGGUUCAGGGAGUGGCUGAAGG ACACCUGCGGCGCCAACGCCAAGCAGAG CAGGGACUGCUUCGGCUGCCUGAGGGAG UGGUGCGACGCCUUCCUGUAAAUCUGGU UACCACUAAACCAGCCUCAAGAACACCC GAAUGGAGUCUCUAAGCUACAUAAUACC AACUUACACUUACAAAAUGUUGUCCCCC AAAAUGUAGCCAUUCGUAUCUGCUCCUA AUAAAAAGAAAGUUUCUUCACAUUUCUU CUCUCCCUUGCACCUCGGAUCAGUACUA GACUGAUAGCUUAUUGGGAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	MDSGIHSDSGIHSDSGI HSTTSQKHRDFVAEPMG EKPVGSLAGIGEVLGKK LEERGFDKAYVVLGQFL VLKKDEDRFREWLKDTC GANAKQSRDCFGCLREW CDAFL

199	TAATACGACTCACTA_AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC_TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCTGGACAGCGGCATCCA CAGCGGCGCCACCACCACCGCCC CCAGCCTGAGCACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCA_CGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGT_CCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGC_GCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCAC_AAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAA_ACCAACT TACACTTACAAAATG_TGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCUGGACAGCGGCAUCCACAG CGGCGCCACCACCACCGCCCCCAGCCUG AGCACCACCAGCCAGAAGCACAGGGACU UCGUGGCCGAGCCCAUGGGCGAGAAGCC CGUGGGCAGCCUGGCCGGCAUCGGCGAG GUGCUGGGCAAGAAGCUGGAGGAGAGGG GCUUCGACAAGGCCUACGUGGUGCUGGG CCAGUUCCUGGUGCUGAAGAAGGACGAG GACAGGUUCAGGGAGUGGCUGAAGGACA CCUGCGGCGCCAACGCCAAGCAGAGCAG GGACUGCUUCGGCUGCCUGAGGGAGUGG UGCGACGCCUUCCUGUAAAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MLDSGIHSGATTTAPSL STTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL
200	TAATACGACTCACTA_AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC_TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACAGCGGCGACAGCGG CTACCCCAGCGAGACGACCAGCC AGAAGCACAGGGACTT_CGTGGCC GAGCCCATGGGCGAGAAGCCCGT GGGCAGCCTGGCCGGCATCGGCG AGGTGCTGGGCAAGAAGCTGGAG GAGAGGGGCTTCGACAAGGCCTA CGTGGTGCTGGGCCAGTTCCTGG TGCTGAAGAAGGACGAGGACAGG TTCAGGGAGTGGCTGAAGGACAC CTGCGGCGCCAACGCCAAGCAGA GCAGGGACTGCTTCGGCTGCCTG AGGGAGTGGTGCGACGCCTTCCT GTAAATCTGGTTACCACTAAACC AGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAA CTTACACTTACAAAA_GTTGTCC CCCAAAATGTAGCCA_TCGTATC TGCTCCTAATAAAAAGAAAGTTT	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACAGCGGCGACAGCGGCUA CCCCAGCGAGACGACCAGCCAGAAGCAC AGGGACUUCGUGGCCGAGCCCAUGGGCG AGAAGCCCGUGGGCAGCCUGGCCGGCAU CGGCGAGGUGCUGGGCAAGAAGCUGGAG GAGAGGGGCUUCGACAAGGCCUACGUGG UGCUGGGCCAGUUCCUGGUGCUGAAGAA GGACGAGGACAGGUUCAGGGAGUGGCUG AAGGACACCUGCGGCGCCAACGCCAAGC AGAGCAGGGACUGCUUCGGCUGCCUGAG GGAGUGGUGCGACGCCUUCCUGUAAAUC UGGUUACCACUAAACCAGCCUCAAGAAC ACCCGAAUGGAGUCUCUAAGCUACAUAA UACCAACUUACACUUACAAAAUGUUGUC CCCCAAAAUGUAGCCAUUCGUAUCUGCU CCUAAUAAAAAGAAAGUUUCUUCACAUU UCUUCUCUCCCUUGCACCUCGGAUCAGU ACUAGACUGAUAGCUUAUUGGGAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MNSGDSGYPSETTSQKH RDFVAEPMGEKPVGSLA GIGEVLGKKLEERGFDK AYVVLGQFLV_KKDEDR FREWLKDTCGANAKQSR DCFGCLREWCDAFL

	CTTCACATTTCTTCTCTCCCTTG CACCTCGGATCAGTACTAGACTG ATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	
201	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAGCGACAACGACAGCGG CATCTGCATGAGCACCACCAGCC AGAAGCACAGGGACTTCGTGGCC GAGCCCATGGGCGAGAAGCCCGT GGGCAGCCTGGCCGGCATCGGCG AGGTGCTGGGCAAGAAGCTGGAG GAGAGGGGCTTCGACAAGGCCTA CGTGGTGCTGGGCCAGTTCCTGG TGCTGAAGAAGGACGAGGACAGG TTCAGGGAGTGGCTGAAGGACAC CTGCGGCGCCAACGCCAAGCAGA GCAGGGACTGCTTCGGCTGCCTG AGGGAGTGGTGCGACGCCTTCCT GTAAATCTGGTTACCACTAAACC AGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAA CTTACACTTACAAAATGTTGTCC CCCAAAATGTAGCCATTCGTATC TGCTCCTAATAAAAAGAAAGTTT CTTCACATTTCTTCTCTCCCTTG CACCTCGGATCAGTACTAGACTG ATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAGCGACAACGACAGCGGCAU CUGCAUGAGCACCACCAGCCAGAAGCAC AGGGCACUUCGUGGCCGAGCCCAUGGGCG AGAAGCCCGUGGGCAGCCUGGCCGGCAU CGGCGAGGUGCUGGGCAAGAAGCUGGAG GAGAGGGGCUUCGACAAGGCCUACGUGG UGCUGGGCCAGUUCCUGGUGCUGAAGAA GGACGAGGACAGGUUCAGGGAGUGGCUG AAGGACACCUGCGGCGCCAACGCCAAGC AGAGCAGGGACUGCUUCGGCUGCCUGAG GGAGUGGUGCGACGCCUUCCUGUAAAUC UGGUUACCACUAAACCAGCCUCAAGAAC ACCCGAAUGGAGUCUCUAAGCUACAUAA UACCAACUUACACUUACAAAAUGUUGUC CCCCAAAAUGUAGCCAUUCGUAUCUGCU CCUAAUAAAAAGAAAGUUUCUUCACAUU UCUUCUCUCCCUUGCACCUCGGAUCAGU ACUAGACUGAUAGCUUAUUGGGAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	MSDNDSGICMSTTSQKH RDFVAEPMGEKPVGSLA GIGEVLGKKLEERGFDK AYVVLGQFLVLKKDEDR FREWLKDTCGANAKQSR DCFGCLREWCDAFL
202	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCAGCGAGGACAGCGG CTACGGCATGGAGACGACCAGCC AGAAGCACAGGGACTTCGTGGCC GAGCCCATGGGCGAGAAGCCCGT GGGCAGCCTGGCCGGCATCGGCG AGGTGCTGGGCAAGAAGCTGGAG GAGAGGGGCTTCGACAAGGCCTA CGTGGTGCTGGGCCAGTTCCTGG TGCTGAAGAAGGACGAGGACAGG TTCAGGGAGTGGCTGAAGGACAC CTGCGGCGCCAACGCCAAGCAGA	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCAGCGAGGACAGCGGCUA CGGCAUGGAGACGACCAGCCAGAAGCAC AGGGACUUCGUGGCCGAGCCCAUGGGCG AGAAGCCCGUGGGCAGCCUGGCCGGCAU CGGCGAGGUGCUGGGCAAGAAGCUGGAG GAGAGGGGCUUCGACAAGGCCUACGUGG UGCUGGGCCAGUUCCUGGUGCUGAAGAA GGACGAGGACAGGUUCAGGGAGUGGCUG AAGGACACCUGCGGCGCCAACGCCAAGC AGACCACGGACUGCUUCGGCUGCCUGAG GGAGUGGUGCGACGCCUUCCUGUAAAUC UGGUUACCACUAAACCAGCCUCAAGAAC ACCCGAAUGGAGUCUCUAAGCUACAUAA	MPSEDSGYGMETTSQKH RDFVAEPMGEKPVGSLA GIGEVLGKKLEERGFDK AYVVLGQFLVLKKDEDR FREWLKDTCGANAKQSR DCFGCLREWCDAFL

	GCAGGGACTGCTTCGGCTGCCTG AGGGAGTGGTGCGACGCCTTCCT GTAAATCTGGTTACCACTAAACC AGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAA CTTACACTTACAAAA-GTTGTCC CCCAAAATGTAGCCA-TCGTATC TGCTCCTAATAAAAAGAAAGTTT CTTCACATTTCTTCTCTCCCTTG CACCTCGGATCAGTACTAGACTG ATAGCTTATTGGG	UACCAACUUACACUUACAAAAUGUUGUC CCCCAAAAUGUAGCCAUUCGUAUCUGCU CCUAAUAAAAAGAAAGUUUCUUCACAUU UCUUCUCUCCCUUGCACCUCGGAUCAGU ACUAGACUGAUAGCUUAUUGGGAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	
203	TAATACGACTCACTA-AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC-TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAGCTACCTGGACAGCGG CATCCACAGCGGCACCACCAGCC AGAAGCACAGGGACTTCGTGGCC GAGCCCATGGGCGAGAAGCCCGT GGGCAGCCTGGCCGGCATCGGCG AGGTGCTGGGCAAGAAGCTGGAG GAGAGGGGCTTCGACAAGGCCTA CGTGGTGCTGGGCCAGTTCCTGG TGCTGAAGAAGGACGAGGACAGG TTCAGGGAGTGGCTGAAGGACAC CTGCGGCGCCAACGCCAAGCAGA GCAGGGACTGCTTCGGCTGCCTG AGGGAGTGGTGCGACGCCTTCCT GTAAATCTGGTTACCACTAAACC AGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAA CTTACACTTACAAAA-GTTGTCC CCCAAAATGTAGCCA-TCGTATC TGCTCCTAATAAAAAGAAAGTTT CTTCACATTTCTTCTCTCCCTTG CACCTCGGATCAGTACTAGACTG ATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAGCUACCUGGACAGCGGCAU CCACAGCGGCACCACCAGCCAGAAGCAC AGGGACUUCGUGGCCGAGCCCAUGGGCG AGAAGCCCGUGGGCAGCCUGGCCGGCAU CGGCGAGGUGCUGGGCAAGAAGCUGGAG GAGAGGGGCUUCGACAAGGCCUACGUGG UGCUGGGCCAGUUCCUGGUGCUGAAGAA GGACGAGGACAGGUUCAGGGAGUGGCUG AAGGACACCUGCGGCGCCAACGCCAAGC AGAGCAGGGACUGCUUCGGCUGCCUGAG GGAGUGGUGCGACGCCUUCCUGUAAAUC UGGUUACCACUAAACCAGCCUCAAGAAC ACCCGAAUGGAGUCUCUAAGCUACAUAA UACCAACUUACACUUACAAAAUGUUGUC CCCCAAAAUGUAGCCAUUCGUAUCUGCU CCUAAUAAAAAGAAAGUUUCUUCACAUU UCUUCUCUCCCUUGCACCUCGGAUCAGU ACUAGACUGAUAGCUUAUUGGGAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	MSYLDSGIHSGTTSQKH RDFVAEPMGEKPVGSLA GIGEVLGKKLEERGFDK AYVVLGQFLV_KKDEDR FREWLKDTCGANAKQSR DCFGCLREWCDAFL
204	TAATACGACTCACTA-AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC-TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAGCTACCTGGACAGCGG CATCCACAGCGGCGCCACCACCA CCGCCCCCAGCCTGAGCACCACC AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGAGGTGCTGGGCAAGAAGCT GGAGGAGAGGGGCTTCGACAAGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAGCUACCUGGACAGCGGCAU CCACAGCGGCGCCACCACCACCGCCCCC AGCCUGAGCACCACCAGCCAGAAGCACA GGGACUUCGUGGCCGAGCCCAUGGGCGA GAAGCCCGUGGGCAGCCUGGCCGGCAUC GGCGAGGUGCUGGGCAAGAAGCUGGAGG AGAGGGGCUUCGACAAGGCCUACGUGGU GCUGGGCCAGUUCCUGGUGCUGAAGAAG GACGAGGACAGGUUCAGGGAGUGGCUGA AGGACACCUGCGGCGCCAACGCCAAGCA	MSYLDSGIHSGATTTAP SLSTTSQKHRDFVAEPM GEKPVGSLAGIGEVLGK KLEERGFDKAYVVLGQF LVLKKDEDRFREWLKDT CGANAKQSRDCFGCLRE WCDAFL

	CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CAGGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAATCTGGTTACCACTA AACCAGCCTCAAGAACACCCGAA TGGAGTCTCTAAGCTACATAATA CCAACTTACACTTACAAAATGTT GTCCCCCAAAATGTAGCCATTCG TATCTGCTCCTAATAAAAAGAAA GTTTCTTCACATTTCTTCTCTCC CTTGCACCTCGGATCAGTACTAG ACTGATAGCTTATTGGG	GAGCAGGGACUGCUUCGGCUGCCUGAGG GAGUGGUGCGACGCCUUCCUGUAAAUCU GGUUACCACUAAACCAGCCUCAAGAACA CCCGAAUGGAGUCUCUAAGCUACAUAAU ACCAACUUACACUUACAAAAUGUUGUCC CCCAAAAUGUAGCCAUUCGUAUCUGCUC CUAAUAAAAAGAAAGUUUCUUCACAUUU CUUCUCUCCCUUGCACCUCGGAUCAGUA CUAGACUGAUAGCUUAUUGGGAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA	
205	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAAGGCCGCCGTGAGCCA CTGGCAGCAGCAGAGCTACCTGG ACAGCGGCATCCACAGCGGCGCC ACCACCACCGCCCCCAGCCTGAG CACCACCAGCCAGAAGCACAGGG ACTTCGTGGCCGAGCCCATGGGC GAGAAGCCCGTGGGCAGCCTGGC CGGCATCGGCGAGGTGCTGGGCA AGAAGCTGGAGGAGAGGGGCTTC GACAAGGCCTACGTGGTGCTGGG CCAGTTCCTGGTGCTGAAGAAGG ACGAGGACAGGTTCAGGGAGTGG CTGAAGGACACCTGCGGCGCCAA CGCCAAGCAGAGCAGGGACTGCT TCGGCTGCCTGAGGGAGTGGTGC GACGCCTTCCTGTAAATCTGGTT ACCACTAAACCAGCCTCAAGAAC ACCCGAATGGAGTCTCTAAGCTA CATAATACCAACTTACACTTACA AAATGTTGTCCCCCAAAATGTAG CCATTCGTATCTGCTCCTAATAA AAAGAAAGTTTCTTCACATTTCT TCTCTCCCTTGCACCTCGGATCA GTACTAGACTGATAGCTTATTGG G	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAAGGCCGCCGUGAGCCACUG GCAGCAGCAGAGCUACCUGGACAGCGGC AUCCACAGCGGCGCCACCACCACCGCCC CCAGCCUGAGCACCACCAGCCAGAAGCA CAGGGACUUCGUGGCCGAGCCCAUGGGC GAGAAGCCCGUGGGCAGCCUGGCCGGCA UCGGCGAGGUGCUGGGCAAGAAGCUGGA GGAGAGGGGCUUCGACAAGGCCUACGUG GUGCUGGGCCAGUUCCUGGUGCUGAAGA AGGACGAGGACAGGUUCAGGGAGUGGCU GAAGGACACCUGCGGCGCCAACGCCAAG CAGAGCAGGGACUGCUUCGGCUGCCUGA GGGAGUGGUGCGACGCCUUCCUGUAAAU CUGGUUACCACUAAACCAGCCUCAAGAA CACCCGAAUGGAGUCUCUAAGCUACAUA AUACCAACUUACACUUACAAAAUGUUGU CCCCCAAAAUGUAGCCAUUCGUAUCUGC UCCUAAUAAAAAGAAAGUUUCUUCACAU UUCUUCUCUCCCUUGCACCUCGGAUCAG UACUAGACUGAUAGCUUAUUGGGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA	MKAAVSHWQQQSYLDSG IHSGATTTAPSLSTTSQ KHRDFVAEPMGEKPVGS LAGIGEVLGKKLEERGF DKAYVVLGQFLVLKKDE DRFREWLKDTCGANAKQ SRDCFGCLREWCDAFL
206	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCTGCCCACCCCCCCCCT GCCCACCCCCCCCCTGCCCACCC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCUGCCCACCCCCCCCCUGCC CACCCCCCCCCUGCCCACCCCCCCCACC ACCAGCCAGAAGCACAGGGACUUCGUGG CCGAGCCCAUGGGCGAGAAGCCCGUGGG	MLPTPPLPTPPLPTPPT TSQKHRDFVAEPMGEKP VGSLAGIGEVLGKKLEE RGFDKAYVVLGQFLVLK KDEDRFREWLKDTCGAN AKQSRDCFGCLREWCDA FL

	CCCCCACCACCAGCCAGAAGCAC AGGGACTTCGTGGCCGAGCCCAT GGGCGAGAAGCCCGTGGGCAGCC TGGCCGGCATCGGCGAGGTGCTG GGCAAGAAGCTGGAGGAGAGGGG CTTCGACAAGGCCTACGTGGTGC TGGGCCAGTTCCTGGTGCTGAAG AAGGACGAGGACAGGTTCAGGGA GTGGCTGAAGGACACCTGCGGCG CCAACGCCAAGCAGAGCAGGGAC TGCTTCGGCTGCCTGAGGGAGTG GTGCGACGCCTTCCTGTAAATCT GGTTACCACTAAACCAGCCTCAA GAACACCCGAATGGAGTCTCTAA GCTACATAATACCAACTTACACT TACAAAATGTTGTCCCCCAAAAT GTAGCCATTCGTATCTGCTCCTA ATAAAAAGAAAGTTTCTTCACAT TTCTTCTCTCCCTTGCACCTCGG ATCAGTACTAGACTGATAGCTTA TTGGG	CAGCCUGGCCGGCAUCGGCGAGGUGCUG GGCAAGAAGCUGGAGGAGAGGGGCUUCG ACAAGGCCUACGUGGUGCUGGGCCAGUU CCUGGUGCUGAAGAAGGACGAGGACAGG UUCAGGGAGUGGCUGAAGGACACCUGCG GCGCCAACGCCAAGCAGAGCAGGGACUG CUUCGGCUGCCUGAGGGAGUGGUGCGAC GCCUUCCUGUAAAUCUGGUUACCACUAA ACCAGCCUCAAGAACACCCGAAUGGAGU CUCUAAGCUACAUAAUACCAACUUACAC UUACAAAAUGUUGUCCCCCAAAAUGUAG CCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUUCUUCUCUCCCUU GCACCUCGGAUCAGUACUAGACUGAUAG CUUAUUGGGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAA	
207	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCCTGCCCAGCGGCCT GCTGACCCCCCCCCAGAGCCCCA AGAAGCAGAGCACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCCUGCCCAGCGGCCUGCU GACCCCCCCCCCAGAGCCCCAAGAAGCAG AGCACCACCAGCCAGAAGCACAGGGACU UCGUGGCCGAGCCCAUGGGCGAGAAGCC CGUGGGCAGCCUGGCCGGCAUCGGCGAG GUGCUGGGCAAGAAGCUGGAGGAGAGGG GCUUCGACAAGGCCUACGUGGUGCUGGG CCAGUUCCUGGUGCUGAAGAAGGACGAG GACAGGUUCAGGGAGUGGCUGAAGGACA CCUGCGGCGCCAACGCCAAGCAGAGCAG GGACUGCUUCGGCUGCCUGAGGGAGUGG UGCGACGCCUUCCUGUAAAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MPLPSGLLTPPQSPKKQ STTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

208	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGCCGCCAGGGGCCTGCT GACCCCCCCCGCCAGCCCCCTGG AGCTGCTGGAGGCCACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GAGGTGCTGGGCAAGAAGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACAG GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAATCTGGTTACCACTAAAC CAGCCTCAAGAACACCCGAATGG AGTCTCTAAGCTACATAATACCA ACTTACACTTACAAAATGTTGTC CCCCAAAATGTAGCCATTCGTAT CTGCTCCTAATAAAAAGAAAGTT TCTTCACATTTCTTCTCTCCCTT GCACCTCGGATCAGTACTAGACT GATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGCCGCCAGGGGCCUGCUGAC CCCCCCCGCCAGCCCCCUGGAGCUGCUG GAGGCCACCACCAGCCAGAAGCACAGGG ACUUCGUGGCCGAGCCCAUGGGCGAGAA GCCCGUGGGCAGCCUGGCCGGCAUCGGC GAGGUGCUGGGCAAGAAGCUGGAGGAGA GGGGCUUCGACAAGGCCUACGUGGUGCU GGGCCAGUUCCUGGUGCUGAAGAAGGAC GAGGACAGGUUCAGGGAGUGGCUGAAGG ACACCUGCGGCGCCAACGCCAAGCAGAG CAGGGACUGCUUCGGCUGCCUGAGGGAG UGGUGCGACGCCUUCCUGUAAAUCUGGU UACCACUAAACCAGCCUCAAGAACACCC GAAUGGAGUCUCUAAGCUACAUAAUACC AACUUACACUUACAAAAUGUUGUCCCCC AAAAUGUAGCCAUUCGUAUCUGCUCCUA AUAAAAAGAAAGUUUCUUCACAUUUCUU CUCUCCCUUGCACCUCGGAUCAGUACUA GACUGAUAGCUUAUUGGGAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
209	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGGCCAGAGCCAGGGCCC CCCCACCCCCCCCACCACCCCCA AGACCGACGTGACCACCAGCCAG AAGCACAGGGACTTCGTGGCCGA GCCCATGGGCGAGAAGCCCGTGG GCAGCCTGGCCGGCATCGGCGAG GTGCTGGGCAAGAAGCTGGAGGA GAGGGGCTTCGACAAGGCCTACG TGGTGCTGGGCCAGTTCCTGGTG CTGAAGAAGGACGAGGACAGGTT CAGGGAGTGGCTGAAGGACACCT GCGGCGCCAACGCCAAGCAGAGC AGGGACTGCTTCGGCTGCCTGAG GGAGTGGTGCGACGCCTTCCTGT AAATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGGCCAGAGCCAGGGCCCCCC CACCCCCCCCACCACCCCCAAGACCGAC GUGACCACCAGCCAGAAGCACAGGGACU UCGUGGCCGAGCCCAUGGGCGAGAAGCC CGUGGGCAGCCUGGCCGGCAUCGGCGAG GUGCUGGGCAAGAAGCUGGAGGAGAGGG GCUUCGACAAGGCCUACGUGGUGCUGGG CCAGUUCCUGGUGCUGAAGAAGGACGAG GACAGGUUCAGGGAGUGGCUGAAGGACA CCUGCGGCGCCAACGCCAAGCAGAGCAG GGACUGCUUCGGCUGCCUGAGGGAGUGG UGCGACGCCUUCCUGUAAAUCUGGUUAC CACUAAACCAGCCUCAAGAACACCCGAA UGGAGUCUCUAAGCUACAUAAUACCAAC UUACACUUACAAAAUGUUGUCCCCCAAA AUGUAGCCAUUCGUAUCUGCUCCUAAUA AAAAGAAAGUUUCUUCACAUUUCUUCUC UCCCUUGCACCUCGGAUCAGUACUAGAC UGAUAGCUUAUUGGGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MGQSQGPPTPPTTPKTD VTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

	CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	
210	TAATACGACTCACTA⁻AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC⁻TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGACCCACGGCATGATGCT GACCCCCCCCAGCAGCCCCCTGG AGCTGACCACCAGCCAGAAGCAC AGGGACTTCGTGGCCGAGCCCAT GGGCGAGAAGCCCGTGGGCAGCC TGGCCGGCATCGGCGAGGTGCTG GGCAAGAAGCTGGAGGAGAGGGG CTTCGACAAGGCCTACGTGGTGC TGGGCCAGTTCCTGG⁻GCTGAAG AAGGACGAGGACAGG⁻TCAGGGA GTGGCTGAAGGACACCTGCGGCG CCAACGCCAAGCAGAGCAGGGAC TGCTTCGGCTGCCTGAGGGAGTG GTGCGACGCCTTCCTGTAAATCT GGTTACCACTAAACCAGCCTCAA GAACACCCGAATGGAGTCTCTAA GCTACATAATACCAACTTACACT TACAAAATGTTGTCCCCCAAAAT GTAGCCATTCGTATC⁻GCTCCTA ATAAAAAGAAAGTTTCTTCACAT TTCTTCTCTCCCTTGCACCTCGG ATCAGTACTAGACTGATAGCTTA TTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGACCCACGGCAUGAUGCUGAC CCCCCCCAGCAGCCCCCUGGAGCUGACC ACCAGCCAGAAGCACAGGGACUUCGUGG CCGAGCCCAUGGGCGAGAAGCCCGUGGG CAGCCUGGCCGGCAUCGGCGAGGUGCUG GGCAAGAAGCUGGAGGAGAGGGGCUUCG ACAAGGCCUACGUGGUGCUGGGCCAGUU CCUGGUGCUGAAGAAGGACGAGGACAGG UUCAGGGAGUGGCUGAAGGACACCUGCG GCGCCAACGCCAAGCAGAGCAGGGACUG CUUCGGCUGCCUGAGGGAGUGGUGCGAC GCCUUCCUGUAAAUCUGGUUACCACUAA ACCAGCCUCAAGAACACCCGAAUGGAGU CUCUAAGCUACAUAAUACCAACUUACAC UUACAAAAUGUUGUCCCCCAAAAUGUAG CCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUUCUUCUCUCCCUU GCACCUCGGAUCAGUACUAGACUGAUAG CUUAUUGGGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	MTHGMMLTPPSSPLELT TSQKHRDFVAEPMGEKP VGSLAGIGEV⌴GKKLEE RGFDKAYVVLGQFLVLK KDEDRFREWLKDTCGAN AKQSRDCFGC⌴REWCDA FL
211	TAATACGACTCACTA⁻AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC⁻TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAGCGGCTTCAAGCCCCT GACCCCCCCCACCACCCCCCTGA GCCCCACCACCAGCCAGAAGCAC AGGGACTTCGTGGCCGAGCCCAT GGGCGAGAAGCCCGTGGGCAGCC TGGCCGGCATCGGCGAGGTGCTG GGCAAGAAGCTGGAGGAGAGGGG CTTCGACAAGGCCTACGTGGTGC TGGGCCAGTTCCTGG⁻GCTGAAG AAGGACGAGGACAGG⁻TCAGGGA	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAGCGGCUUCAAGCCCCUGAC CCCCCCCACCACCCCCCUGAGCCCCACC ACCAGCCAGAAGCACAGGGACUUCGUGG CCGAGCCCAUGGGCGAGAAGCCCGUGGG CAGCCUGGCCGGCAUCGGCGAGGUGCUG GGCAAGAAGCUGGAGGAGAGGGGCUUCG ACAAGGCCUACGUGGUGCUGGGCCAGUU CCUGGUGCUGAAGAAGGACGAGGACAGG UUCAGGCAGUGGCUGAAGGACACCUGCG GCGCCAACGCCAAGCAGAGCAGGGACUG CUUCGGCUGCCUGAGGGAGUGGUGCGAC GCCUUCCUGUAAAUCUGGUUACCACUAA	MSGFKPLTPPTTPLSPT TSQKHRDFVAEPMGEKP VGSLAGIGEV⌴GKKLEE RGFDKAYVVLGQFLVLK KDEDRFREWLKDTCGAN AKQSRDCFGC⌴REWCDA FL

	GTGGCTGAAGGACACCTGCGGCG CCAACGCCAAGCAGAGCAGGGAC TGCTTCGGCTGCCTGAGGGAGTG GTGCGACGCCTTCCTGTAAATCT GGTTACCACTAAACCAGCCTCAA GAACACCCGAATGGAGTCTCTAA GCTACATAATACCAACTTACACT TACAAAATGTTGTCCCCCAAAAT GTAGCCATTCGTATCTGCTCCTA ATAAAAAGAAAGTTTCTTCACAT TTCTTCTCTCCCTTGCACCTCGG ATCAGTACTAGACTGATAGCTTA TTGGG	ACCAGCCUCAAGAACACCCGAAUGGAGU CUCUAAGCUACAUAAUACCAACUUACAC UUACAAAAUGUUGUCCCCCAAAAUGUAG CCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUUCUUCUCUCCCUU GCACCUCGGAUCAGUACUAGACUGAUAG CUUAUUGGGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	
212	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGCAGT TCGAGAGGCAGCAGTTCGAGAGG CAGCAGTTCGAGAGGCAGACCAC CAGCCAGAAGCACAGGGACTTCG TGGCCGAGCCCATGGGCGAGAAG CCCGTGGGCAGCCTGGCCGGCAT CGGCGAGGTGCTGGGCAAGAAGC TGGAGGAGAGGGGCTTCGACAAG GCCTACGTGGTGCTGGGCCAGTT CCTGGTGCTGAAGAAGGACGAGG ACAGGTTCAGGGAGTGGCTGAAG GACACCTGCGGCGCCAACGCCAA GCAGAGCAGGGACTGCTTCGGCT GCCTGAGGGAGTGGTGCGACGCC TTCCTGTAAATCTGGTTACCACT AAACCAGCCTCAAGAACACCCGA ATGGAGTCTCTAAGCTACATAAT ACCAACTTACACTTACAAAATGT TGTCCCCCAAAATGTAGCCATTC GTATCTGCTCCTAATAAAAAGAA AGTTTCTTCACATTTCTTCTCTC CCTTGCACCTCGGATCAGTACTA GACTGATAGCTTATTGGG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGCAGUUCGAGAGGCAG CAGUUCGAGAGGCAGCAGUUCGAGAGGC AGACCACCAGCCAGAAGCACAGGGACUU CGUGGCCGAGCCCAUGGGCGAGAAGCCC GUGGGCAGCCUGGCCGGCAUCGGCGAGG UGCUGGGCAAGAAGCUGGAGGAGAGGGG CUUCGACAAGGCCUACGUGGUGCUGGGC CAGUUCCUGGUGCUGAAGAAGGACGAGG ACAGGUUCAGGGAGUGGCUGAAGGACAC CUGCGGCGCCAACGCCAAGCAGAGCAGG GACUGCUUCGGCUGCCUGAGGGAGUGGU GCGACGCCUUCCUGUAAAUCUGGUUACC ACUAAACCAGCCUCAAGAACACCCGAAU GGAGUCUCUAAGCUACAUAAUACCAACU UACACUUACAAAAUGUUGUCCCCCAAAA UGUAGCCAUUCGUAUCUGCUCCUAAUAA AAAGAAAGUUUCUUCACAUUUCUUCUCU CCCUUGCACCUCGGAUCAGUACUAGACU GAUAGCUUAUUGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MQFERQQFERQQFERQT TSQKHRDFVAEPMGEKP VGSLAGIGEVLGKKLEE RGFDKAYVVLGQFLVLK KDEDRFREWLKDTCGAN AKQSRDCFGCLREWCDA FL
213	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGACCACCAGCCAGAAGCA CAGGGACTTCGTGGCCGAGCCCA TGGGCGAGAAGCCCGTGGGCAGC CTGGCCGGCATCGGCGAGGTGCT GGGCAAGAAGCTGGAGGAGAGGG GCTTCGACAAGGCCTACGTGGTG CTGGGCCAGTTCCTGGTGCTGAA GAAGGACGAGGACCTGTTCAGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGACCACCAGCCAGAAGCACAG GGACUUCGUGGCCGAGCCCAUGGGCGAG AAGCCCGUGGGCAGCCUGGCCGGCAUCG GCGAGGUGCUGGGCAAGAAGCUGGAGGA GAGGGGCUUCGACAAGGCCUACGUGGUG CUGGGCCAGUUCCUGGUGCUGAAGAAGG ACGAGGACCUGUUCAGGGAGUGGCUGAA GGACACCUGCGGCGCCAACGCCAAGCAG AGCAGGGACUGCUUCGGCUGCCUGAGGG AGUGGUGCGACGCCUUCCUGUAAAUCUG	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFL

	AGTGGCTGAAGGACACCTGCGGC GCCAACGCCAAGCAGAGCAGGGA CTGCTTCGGCTGCCTGAGGGAGT GGTGCGACGCCTTCCTGTAAATC TGGTTACCACTAAACCAGCCTCA AGAACACCCGAATGGAGTCTCTA AGCTACATAATACCAACTTACAC TTACAAAATGTTGTCCCCCAAAA TGTAGCCATTCGTATCTGCTCCT AATAAAAAGAAAGTTCTTCACA TTTCTTCTCTCCCTTGCACCTCG GATCAGTACTAGACTGATAGCTT ATTGGG	GUUACCACUAAACCAGCCUCAAGAACAC CCGAAUGGAGUCUCUAAGCUACAUAAUA CCAACUUACACUUACAAAAUGUUGUCCC CCAAAAUGUAGCCAUUCGUAUCUGCUCC UAAUAAAAAGAAAGUUUCUUCACAUUUC UUCUCUCCCUUGCACCUCGGAUCAGUAC UAGACUGAUAGCUUAUUGGGAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	
214	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAAGGACGAGGAGTACGA GCAGATGGTGAAGGAGCTGCAGG AGATCAGGCTGGAGTGGAGCCAC CCCCAGTTCGAGAAGTGGAGCCA CCCCCAGTTCGAGAAGTGGAGCC ACCCCCAGTTCGAGAAGACCACC AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGAGGTGCTGGGCAAGAAGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CCTGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAATCTGGTTACCACTA AACCAGCCTCAAGAACACCCGAA TGGAGTCTCTAAGCTACATAATA CCAACTTACACTTACAAAATGTT GTCCCCCAAAATGTAGCCATTCG TATCTGCTCCTAATAAAAAGAAA GTTTCTTCACATTTCTTCTCTCC CTTGCACCTCGGATCAGTACTAG ACTGATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAAGGACGAGGAGUACGAGCA GAUGGUGAAGGAGCUGCAGGAGAUCAGG CUGGAGUGGAGCCACCCCCAGUUCGAGA AGUGGAGCCACCCCCAGUUCGAGAAGUG GAGCCACCCCCAGUUCGAGAAGACCACC AGCCAGAAGCACAGGGACUUCGUGGCCG AGCCCAUGGGCGAGAAGCCCGUGGGCAG CCUGGCCGGCAUCGGCGAGGUGCUGGGC AAGAAGCUGGAGGAGAGGGGCUUCGACA AGGCCUACGUGGUGCUGGGCCAGUUCCU GGUGCUGAAGAAGGACGAGGACCUGUUC AGGGAGUGGCUGAAGGACACCUGCGGCG CCAACGCCAAGCAGAGCAGGGACUGCUU CGGCUGCCUGAGGGAGUGGUGCGACGCC UUCCUGUAAAUCUGGUUACCACUAAACC AGCCUCAAGAACACCCGAAUGGAGUCUC UAAGCUACAUAAUACCAACUUACACUUA CAAAAUGUUGUCCCCCAAAAUGUAGCCA UUCGUAUCUGCUCCUAAUAAAAAGAAAG UUUCUUCACAUUUCUUCUCUCCCUUGCA CCUCGGAUCAGUACUAGACUGAUAGCUU AUUGGGAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MKDEEYEQMVKELQEIR LEWSHPQFEKWSHPQFE KWSHPQFEKTTSQKHRD FVAEPMGEKPVGSLAGI GEVLGKKLEERGFDKAY VVLGQFLVLKKDEDLFR EWLKDTCGANAKQSRDC FGCLREWCDAFL
215	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCTGGTTCCCCCAGGA GGGCACCATCGACATCAAGAGGT	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCUGGUUCCCCCAGGAGGG CACCAUCGACAUCAAGAGGUGGAGGAGG GUGGGCUGGAGCCACCCCCAGUUCGAGA AGUGGAGCCACCCCCAGUUCGAGAAGUG	MPWFPQEGTIDIKRWRR VGWSHPQFEKWSHPQFE KWSHPQFEKTTSQKHRD FVAEPMGEKPVGSLAGI GEVLGKKLEERGFDKAY VVLGQFLVLKKDEDLFR EWLKDTCGANAKQSRDC FGCLREWCDAFL

	GGAGGAGGGTGGGCTGGAGCCAC CCCCAGTTCGAGAAGTGGAGCCA CCCCCAGTTCGAGAAGTGGAGCC ACCCCCAGTTCGAGAAGACCACC AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGAGGTGCTGGGCAAGAAGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CCTGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAATCTGGTTACCACTA AACCAGCCTCAAGAACACCCGAA TGGAGTCTCTAAGCTACATAATA CCAACTTACACTTACAAAATGTT GTCCCCCAAAATGTAGCCATTCG TATCTGCTCCTAATAAAAAGAAA GTTTCTTCACATTTCTTCTCTCC CTTGCACCTCGGATCAGTACTAG ACTGATAGCTTATTGGG	GAGCCACCCCCAGUUCGAGAAGACCACC AGCCAGAAGCACAGGGACUUCGUGGCCG AGCCCAUGGGCGAGAAGCCCGUGGGCAG CCUGGCCGGCAUCGGCGAGGUGCUGGGC AAGAAGCUGGAGGAGAGGGGCUUCGACA AGGCCUACGUGGUGCUGGGCCAGUUCCU GGUGCUGAAGAAGGACGAGGACCUGUUC AGGGAGUGGCUGAAGGACACCUGCGGCG CCAACGCCAAGCAGAGCAGGGACUGCUU CGGCUGCCUGAGGGAGUGGUGCGACGCC UUCCUGUAAAUCUGGUUACCACUAAACC AGCCUCAAGAACACCCGAAUGGAGUCUC UAAGCUACAUAAUACCAACUUACACUUA CAAAAUGUUGUCCCCCAAAAUGUAGCCA UUCGUAUCUGCUCCUAAUAAAAAGAAAG UUUCUUCACAUUUCUUCUCUCCCUUGCA CCUCGGAUCAGUACUAGACUGAUAGCUU AUUGGGAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
216	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAAGTTCGAGAGGCAGAC CACCAGCCAGAAGCACAGGGACT TCGTGGCCGAGCCCATGGGCGAG AAGCCCGTGGGCAGCCTGGCCGG CATCGGCGAGGTGCTGGGCAAGA AGCTGGAGGAGAGGGGCTTCGAC AAGGCCTACGTGGTGCTGGGCCA GTTCCTGGTGCTGAAGAAGGACG AGGACCTGTTCAGGGAGTGGCTG AAGGACACCTGCGGCGCCAACGC CAAGCAGAGCAGGGACTGCTTCG GCTGCCTGAGGGAGTGGTGCGAC GCCTTCCTGTAAATCTGGTTACC ACTAAACCAGCCTCAAGAACACC CGAATGGAGTCTCTAAGCTACAT AATACCAACTTACACTTACAAAA TGTTGTCCCCCAAAATGTAGCCA TTCGTATCTGCTCCTAATAAAAA GAAAGTTTCTTCACATTTCTTCT CTCCCTTGCACCTCGGATCAGTA CTAGACTGATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAAGUUCGAGAGGCAGACCAC CAGCCAGAAGCACAGGGACUUCGUGGCC GAGCCCAUGGGCGAGAAGCCCGUGGGCA GCCUGGCCGGCAUCGGCGAGGUGCUGGG CAAGAAGCUGGAGGAGAGGGGCUUCGAC AAGGCCUACGUGGUGCUGGGCCAGUUCC UGGUGCUGAAGAAGGACGAGGACCUGUU CAGGGAGUGGCUGAAGGACACCUGCGGC GCCAACGCCAAGCAGAGCAGGGACUGCU UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUAAAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAAUACCAACUUACACUU ACAAAAUGUUGUCCCCCAAAAUGUAGCC AUUCGUAUCUGCUCCUAAUAAAAAGAAA GUUUCUUCACAUUUCUUCUCUCCCUUGC ACCUCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MKFERQTTSQKHRDFVA EPMGEKPVGSLAGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDLFREWL KDTCGANAKQSRDCFGC LREWCDAFL

217	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MGNELALKLAGLDIWSH
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	PQFEKWSHPQFEKWSHP
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	QFEKTTSQKHRDFVAEP
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	MGEKPVGSLAGIGEVLG
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	KKLEERGFDKAYVVLGQ
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGGCAACGAGCUGGCCCUGAA	FLVLKKDEDLFREWLKD
	TCACCATTTACGAACGATAGGCC	GCUGGCCGGCCUGGACAUCUGGAGCCAC	TCGANAKQSRDCFGCLR
	ACCATGGGCAACGAGCTGGCCCT	CCCCAGUUCGAGAAGUGGAGCCACCCCC	EWCDAFL
	GAAGCTGGCCGGCCTGGACATCT	AGUUCGAGAAGUGGAGCCACCCCCAGUU	
	GGAGCCACCCCCAGTTCGAGAAG	CGAGAAGACCACCAGCCAGAAGCACAGG	
	TGGAGCCACCCCCAGTTCGAGAA	GACUUCGUGGCCGAGCCCAUGGGCGAGA	
	GTGGAGCCACCCCCAGTTCGAGA	AGCCCGUGGGCAGCCUGGCCGGCAUCGG	
	AGACCACCAGCCAGAAGCACAGG	CGAGGUGCUGGGCAAGAAGCUGGAGGAG	
	GACTTCGTGGCCGAGCCCATGGG	AGGGGCUUCGACAAGGCCUACGUGGUGC	
	CGAGAAGCCCGTGGGCAGCCTGG	UGGGCCAGUUCCUGGUGCUGAAGAAGGA	
	CCGGCATCGGCGAGGTGCTGGGC	CGAGGACCUGUUCAGGGAGUGGCUGAAG	
	AAGAAGCTGGAGGAGAGGGGCTT	GACACCUGCGGCGCCAACGCCAAGCAGA	
	CGACAAGGCCTACGTGGTGCTGG	GCAGGGACUGCUUCGGCUGCCUGAGGGA	
	GCCAGTTCCTGGTGCTGAAGAAG	GUGGUGCGACGCCUUCCUGUAAAUCUGG	
	GACGAGGACCTGTTCAGGGAGTG	UUACCACUAAACCAGCCUCAAGAACACC	
	GCTGAAGGACACCTGCGGCGCCA	CGAAUGGAGUCUCUAAGCUACAUAAUAC	
	ACGCCAAGCAGAGCAGGGACTGC	CAACUUACACUUACAAAAUGUUGUCCCC	
	TTCGGCTGCCTGAGGGAGTGGTG	CAAAAUGUAGCCAUUCGUAUCUGCUCCU	
	CGACGCCTTCCTGTAAATCTGGT	AAUAAAAAGAAAGUUUCUUCACAUUUCU	
	TACCACTAAACCAGCCTCAAGAA	UCUCUCCCUUGCACCUCGGAUCAGUACU	
	CACCCGAATGGAGTCCCTAAGCT	AGACUGAUAGCUUAUUGGGAAAAAAAAA	
	ACATAATACCAACTTACACTTAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAAATGTTGTCCCCCAAAATGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCATTCGTATCTGCTCCTAATA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAAAGAAAGTTTCTTCACATTTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTCTCTCCCTTGCACCTCGGATC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTACTAGACTGATAGCTTATTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GG	AAAAAAAAAAAAAAAAAAAAAA	
218	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MGNELALKLAGLDIGSG
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	SGSTTSQKHRDFVAEPM
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	GEKPVGSLAGIGEVLGK
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	KLEERGFDKAYVVLGQF
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	LVLKKDEDLFREWLKDT
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGGCAACGAGCUGGCCCUGAA	CGANAKQSRDCFGCLRE
	TCACCATTTACGAACGATAGGCC	GCUGGCCGGCCUGGACAUCGGCAGCGGC	WCDAFLWSHPQFEKWSH
	ACCATGGGCAACGAGCTGGCCCT	AGCGGCAGCACCACCAGCCAGAAGCACA	PQFEKWSHPQFEKPAAK
	GAAGCTGGCCGGCCTGGACATCG	GGGACUUCGUGGCCGAGCCCAUGGGCGA	RVKLDASQGTKRSYEQM
	GCAGCGGCAGCGGCAGCACCACC	GAAGCCCGUGGGCAGCCUGGCCGGCAUC	ETDGERQSA
	AGCCAGAAGCACAGGGACTTCGT	GGCGAGGUGCUGGGCAAGAAGCUGGAGG	
	GGCCGAGCCCATGGGCGAGAAGC	AGAGGGGCUUCGACAAGGCCUACGUGGU	
	CCGTGGGCAGCCTGGCCGGCATC	GCUGGGCCAGUUCCUGGUGCUGAAGAAG	
	GGCGAGGTGCTGGGCAAGAAGCT	GACGAGGACCUGUUCAGGGAGUGGCUGA	
	GGAGGAGAGGGGCTTCGACAAGG	AGGACACCUGCGGCGCCAACGCCAAGCA	
	CCTACGTGGTGCTGGGCCAGTTC	GAGCAGGGACUGCUUCGGCUGCCUGAGG	
	CTGGTGCTGAAGAAGGACGAGGA	GAGUGGUGCGACGCCUUCCUGUGGUCAC	
	CCTGTTCAGGGAGTGGCTGAAGG	AUCCUCAAUUUGAGAAAUGGUCCCAUCC	
	ACACCTGCGGCGCCAACGCCAAG	ACAAUUUGAAAAGUGGUCCCAUCCUCAA	
	CAGAGCAGGGACTGCCTCGGCTG	UUUGAAAAGCCCGCCGCCAAGAGGGUGA	
	CCTGAGGGAGTGGTGCGACGCCT	AGCUGGACGCCAGCCAGGGCACCAAGAG	
	TCCTGTGGTCACATCCTCAATTT	GAGCUACGAGCAGAUGGAGACGGACGGC	
	GAGAAATGGTCCCATCCACAATT	GAGAGGCAGAGCGCCUAAUAAAUCUGGU	
	TGAAAAGTGGTCCCATCCTCAAT	UACCACUAAACCAGCCUCAAGAACACCC	

	TTGAAAAGCCCGCCGCCAAGAGG	GAAUGGAGUCUCUAAGCUACAUAAUACC	
	GTGAAGCTGGACGCCAGCCAGGG	AACUUACACUUACAAAAUGUUGUCCCCC	
	CACCAAGAGGAGCTACGAGCAGA	AAAAUGUAGCCAUUCGUAUCUGCUCCUA	
	TGGAGACGGACGGCGAGAGGCAG	AUAAAAAGAAAGUUUCUUCACAUUUCUU	
	AGCGCCTAATAAATCTGGTTACC	CUCUCCCUUGCACCUCGGAUCAGUACUA	
	ACTAAACCAGCCTCAAGAACACC	GACUGAUAGCUUAUUGGGAAAAAAAAAA	
	CGAATGGAGTCTCTAAGCTACAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATACCAACTTACACTTACAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGTTGTCCCCCAAAATGTAGCCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTCGTATCTGCTCCTAATAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAAAGTTTCTTCACATTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCTCGGATCAGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAA	
219	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MDSGIHSTTSQKHRDFV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	AEPMGEKPVGSLAGIGE
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	VLGKKLEERGFDKAYVV
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	LGQFLVLKKDEDLFREW
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	LKDTCGANAKQSRDCFG
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGACAGCGGCAUCCACAGCAC	CLREWCDAFL
	TCACCATTTACGAACGATAGGCC	CACCAGCCAGAAGCACAGGGACUUCGUG	
	ACCATGGACAGCGGCATCCACAG	GCCGAGCCCAUGGGCGAGAAGCCCGUGG	
	CACCACCAGCCAGAAGCACAGGG	GCAGCCUGGCCGGCAUCGGCGAGGUGCU	
	ACTTCGTGGCCGAGCCCATGGGC	GGGCAAGAAGCUGGAGGAGAGGGGCUUC	
	GAGAAGCCCGTGGGCAGCCTGGC	GACAAGGCCUACGUGGUGCUGGGCCAGU	
	CGGCATCGGCGAGGTGCTGGGCA	UCCUGGUGCUGAAGAAGGACGAGGACCU	
	AGAAGCTGGAGGAGAGGGGCTTC	GUUCAGGGAGUGGCUGAAGGACACCUGC	
	GACAAGGCCTACGTGGTGCTGGG	GGCGCCAACGCCAAGCAGAGCAGGGACU	
	CCAGTTCCTGGTGCTGAAGAAGG	GCUUCGGCUGCCUGAGGGAGUGGUGCGA	
	ACGAGGACCTGTTCAGGGAGTGG	CGCCUUCCUGUAAAUCUGGUUACCACUA	
	CTGAAGGACACCTGCGGCGCCAA	AACCAGCCUCAAGAACACCCGAAUGGAG	
	CGCCAAGCAGAGCAGGGACTGCT	UCUCUAAGCUACAUAAUACCAACUUACA	
	TCGGCTGCCTGAGGGAGTGGTGC	CUUACAAAAUGUUGUCCCCCAAAAUGUA	
	GACGCCTTCCTGTAAATCTGGTT	GCCAUUCGUAUCUGCUCCUAAUAAAAAG	
	ACCACTAAACCAGCCTCAAGAAC	AAAGUUUCUUCACAUUUCUUCUCUCCCU	
	ACCCGAATGGAGTCTCTAAGCTA	UGCACCUCGGAUCAGUACUAGACUGAUA	
	CATAATACCAACTTACACTTACA	GCUUAUUGGGAAAAAAAAAAAAAAAAAA	
	AAATGTTGTCCCCCAAAATGTAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCATTCGTATCTGCTCCTAATAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AAAGAAAGTTTCTTCACATTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTCTCCCTTGCACCTCGGATCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTACTAGACTGATAGCTTATTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	G	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
		AAAAAAAAAAAAAA	
220	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MQFERQTTSQKHRDFVA
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	EPMGEKPVGSLAGIGEV
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	LGKKLEERGFDKAYVVL
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	GQFLVLKKDEDLFREWL
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	KDTCGANAKQSRDCFGC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGCAGUUCGAGAGGCAGACCAC	LREWCDAFL
	TCACCATTTACGAACGATAGGCC	CAGCCAGAAGCACAGGGACUUCGUGGCC	
	ACCATGCAGTTCGAGAGGCAGAC	GAGCCCAUGGGCGAGAAGCCCGUGGGCA	
	CACCAGCCAGAAGCACAGGGACT	GCCUGGCCGGCAUCGGCGAGGUGCUGGG	
	TCGTGGCCGAGCCCATGGGCGAG	CAAGAAGCUGGAGGAGAGGGGCUUCGAC	
	AAGCCCGTGGGCAGCCTGGCCGG	AAGGCCUACGUGGUGCUGGGCCAGUUCC	
	CATCGGCGAGGTGCTGGGCAAGA	UGGUGCUGAAGAAGGACGAGGACCUGUU	
	AGCTGGAGGAGAGGGGCTTCGAC	CAGGGAGUGGCUGAAGGACACCUGCGGC	
	AAGGCCTACGTGGTGCTGGGCCA	GCCAACGCCAAGCAGAGCAGGGACUGCU	

	GTTCCTGGTGCTGAAGAAGGACG AGGACCTGTTCAGGGAGTGGCTG AAGGACACCTGCGGCGCCAACGC CAAGCAGAGCAGGGACTGCTTCG GCTGCCTGAGGGAGTGGTGCGAC GCCTTCCTGTAAATCTGGTTACC ACTAAACCAGCCTCAAGAACACC CGAATGGAGTCTCTAAGCTACAT AATACCAACTTACACTTACAAAA TGTTGTCCCCCAAAATGTAGCCA TTCGTATCTGCTCCTAATAAAAA GAAAGTTTCTTCACATTTCTTCT CTCCCTTGCACCTCGGATCAGTA CTAGACTGATAGCTTATTGGG	UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUAAAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAAUACCAACUUACACUU ACAAAAAUGUUGUCCCCCAAAAUGUAGCC AUUCGUAUCUGCUCCUAAUAAAAAGAAA GUUUCUUCACAUUUCUUCUCUCCCUUGC ACCUCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
221	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGACCACCAGCCAGAAGCA CAGGGACTTCGTGGCCGAGCCCA TGGGCGAGAAGCCCGTGGGCAGC CTGGCCGGCATCGGCGAGGTGCT GGGCAAGAAGCTGGAGGAGAGGG GCTTCGACAAGGCCTACGTGGTG CTGGGCCAGTTCCTGGTGCTGAA GAAGGACGAGGACCTGTTCAGGG AGTGGCTGAAGGACACCTGCGGC GCCAACGCCAAGCAGAGCAGGGA CTGCTTCGGCTGCCTGAGGGAGT GGTGCGACGCCTTCCTGTGGAGT CATCCTCAATTTGAGAAATGGTC ACATCCTCAATTTGAGAAATGGT CCCATCCACAATTTGAGAAACCC GCCGCCAAGAGGGTGAAGCTGGA CGCCAGCCAGGGCACCAAGAGGA GCTACGAGCAGATGGAGACCGAC GGCGAGAGGCAGAGCGCCTAAAT CTGGTTACCACTAAACCAGCCTC AAGAACACCCGAATGGAGTCTCT AAGCTACATAATACCAACTTACA CTTACAAAATGTTGTCCCCCAAA ATGTAGCCATTCGTATCTGCTCC TAATAAAAAGAAAGTTTCTTCAC ATTTCTTCTCTCCCTTGCACCTC GGATCAGTACTAGACTGATAGCT TATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGACCACCAGCCAGAAGCACAG GGACUUCGUGGCCGAGCCCAUGGGCGAG AAGCCCGUGGGCAGCCUGGCCGGCAUCG GCGAGGUGCUGGGCAAGAAGCUGGAGGA GAGGGGCUUCGACAAGGCCUACGUGGUG CUGGGCCAGUUCCUGGUGCUGAAGAAGG ACGAGGACCUGUUCAGGGAGUGGCUGAA GGACACCUGCGGCGCCAACGCCAAGCAG AGCAGGGACUGCUUCGGCUGCCUGAGGG AGUGGUGCGACGCCUUCCUGUGGAGUCA UCCUCAAUUUGAGAAAUGGUCACAUCCU CAAUUUGAGAAAUGGUCCCAUCCACAAU UUGAGAAACCCGCCGCCAAGAGGGUGAA GCUGGACGCCAGCCAGGGCACCAAGAGG AGCUACGAGCAGAUGGAGACCGACGGCG AGAGGCAGAGCGCCUAAAUCUGGUUACC ACUAAACCAGCCUCAAGAACACCCGAAU GGAGUCUCUAAGCUACAUAAUACCAACU UACACUUACAAAAUGUUGUCCCCCAAAA UGUAGCCAUUCGUAUCUGCUCCUAAUAA AAAGAAAGUUUCUUCACAUUUCUUCUCU CCCUUGCACCUCGGAUCAGUACUAGACU GAUAGCUUAUUGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQSA

222	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MKFERQTTSQKHRDFVA
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	EPMGEKPVGSLAGIGEV
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	LGKKLEERGFDKAYVVL
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	GQFLVLKKDEDLFREWL
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	KDTCGANAKQSRDCFGC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAAGUUCGAGAGGCAGACCAC	LREWCDAFLWSHPQFEK
	TCACCATTTACGAACGATAGGCC	CAGCCAGAAGCACAGGGACUUCGUGGCC	WSHPQFEKWSHPQFEKP
	ACCATGAAGTTCGAGAGGCAGAC	GAGCCCAUGGGCGAGAAGCCCGUGGGCA	AAKRVKLDASQGTKRSY
	CACCAGCCAGAAGCACAGGGACT	GCCUGGCCGGCAUCGGCGAGGUGCUGGG	EQMETDGERQSA
	TCGTGGCCGAGCCCATGGGCGAG	CAAGAAGCUGGAGGAGAGGGGCUUCGAC	
	AAGCCCGTGGGCAGCCTGGCCGG	AAGGCCUACGUGGUGCUGGGCCAGUUCC	
	CATCGGCGAGGTGCTGGGCAAGA	UGGUGCUGAAGAAGGACGAGGACCUGUU	
	AGCTGGAGGAGAGGGGCTTCGAC	CAGGGAGUGGCUGAAGGACACCUGCGGC	
	AAGGCCTACGTGGTGCTGGGCCA	GCCAACGCCAAGCAGAGCAGGGACUGCU	
	GTTCCTGGTGCTGAAGAAGGACG	UCGGCUGCCUGAGGGAGUGGUGCGACGC	
	AGGACCTGTTCAGGGAGTGGCTG	CUUCCUGUGGAGCCACCCCCAGUUCGAG	
	AAGGACACCTGCGGCGCCAACGC	AAGUGGAGCCACCCCCAGUUCGAGAAGU	
	CAAGCAGAGCAGGGACTGCTTCG	GGAGCCACCCCCAGUUCGAGAAGCCCGC	
	GCTGCCTGAGGGAGTGGTGCGAC	CGCCAAGAGGGUGAAGCUGGACGCCAGC	
	GCCTTCCTGTGGAGCCACCCCCA	CAGGGCACCAAGAGGAGCUACGAGCAGA	
	GTTCGAGAAGTGGAGCCACCCCC	UGGAGACGGACGGCGAGAGGCAGAGCGC	
	AGTTCGAGAAGTGGAGCCACCCC	CUAAAUCUGGUUACCACUAAACCAGCCU	
	CAGTTCGAGAAGCCCGCCGCCAA	CAAGAACACCCGAAUGGAGUCUCUAAGC	
	GAGGGTGAAGCTGGACGCCAGCC	UACAUAAUACCAACUUACACUUACAAAA	
	AGGGCACCAAGAGGAGCTACGAG	UGUUGUCCCCCAAAAUGUAGCCAUUCGU	
	CAGATGGAGACGGACGGCGAGAG	AUCUGCUCCUAAUAAAAAGAAAGUUUCU	
	GCAGAGCGCCTAAATCTGGTTAC	UCACAUUUCUUCUCUCCCUUGCACCUCG	
	CACTAAACCAGCCTCAAGAACAC	GAUCAGUACUAGACUGAUAGCUUAUUGG	
	CCGAATGGAGTCTCTAAGCTACA	GAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAATACCAACTTACACTTACAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTTGTCCCCCAAAATGTAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATTCGTATCTGCTCCTAATAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGAAAGTTTCTTCACATTTCTTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTCCCTTGCACCTCGGATCAGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
		AAAAA	
223	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MDSGIHSTTSQKHRDFV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	AEPMGEKPVGSLAGIGE
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	VLGKKLEERGFDKAYVV
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	LGQFLVLKKDEDLFREW
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	LKDTCGANAKQSRDCFG
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGACAGCGGCAUCCACAGCAC	CLREWCDAFLWSHPQFE
	TCACCATTTACGAACGATAGGCC	CACCAGCCAGAAGCACAGGGACUUCGUG	KWSHPQFEKWSHPQFEK
	ACCATGGACAGCGGCATCCACAG	GCCGAGCCCAUGGGCGAGAAGCCCGUGG	PAAKRVKLDASQGTKRS
	CACCACCAGCCAGAAGCACAGGG	GCAGCCUGGCCGGCAUCGGCGAGGUGCU	YEQMETDGERQSA
	ACTTCGTGGCCGAGCCCATGGGC	GGGCAAGAAGCUGGAGGAGAGGGGCUUC	
	GAGAAGCCCGTGGGCAGCCTGGC	GACAAGGCCUACGUGGUGCUGGGCCAGU	
	CGGCATCGGCGAGGTGCTGGGCA	UCCUGGUGCUGAAGAAGGACGAGGACCU	
	AGAAGCTGGAGGAGAGGGGCTTC	GUUCAGGGAGUGGCUGAAGGACACCUGC	
	GACAAGGCCTACGTGGTGCTGGG	GGCGCCAACGCCAAGCAGAGCAGGGACU	
	CCAGTTCCTGGTGCTGAAGAAGG	GCUUCGGCUGCCUGAGGGAGUGGUGCGA	
	ACGAGGACCTGTTCAGGGAGTGG	CGCCUUCCUGUGGUCACAUCCACAAUUU	
	CTGAAGGACACCTGCGGCGCCAA	GAAAAGUGGAGUCAUCCGCAAUUUGAGA	
	CGCCAAGCAGAGCAGGGACTGCT	AAUGGAGUCAUCCGCAAUUUGAGAAACC	
	TCGGCTGCCTGAGGGAGTGGTGC	CGCCGCCAAGAGGGUGAAGCUGGACGCC	
	GACGCCTTCCTGTGGTCACATCC	AGCCAGGGCACCAAGAGGAGCUACGAGC	
	ACAATTTGAAAAGTGGAGTCATC	AGAUGGAGACGGACGGCGAGAGGCAGAG	

	CGCAATTTGAGAAATGGAGTCAT CCGCAATTTGAGAAACCCGCCGC CAAGAGGGTGAAGCTGGACGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACGGACGGCGA GAGGCAGAGCGCCTAAATCTGGT TACCACTAAACCAGCCTCAAGAA CACCCGAATGGAGTCTCTAAGCT ACATAATACCAACTTACACTTAC AAAATGTTGTCCCCCAAAATGTA GCCATTCGTATCTGCTCCTAATA AAAAGAAAGTTTCTTCACATTTC TTCTCTCCCTTGCACCTCGGATC AGTACTAGACTGATAGCTTATTG GG	CGCCUAAAUCUGGUUACCACUAAACCAG CCUCAAGAACACCCGAAUGGAGUCUCUA AGCUACAUAAUACCAACUUACACUUACA AAAUGUUGUCCCCCAAAAUGUAGCCAUU CGUAUCUGCUCCUAAUAAAAAGAAAGUU UCUUCACAUUUCUUCUCUCCCUUGCACC UCGGAUCAGUACUAGACUGAUAGCUUAU UGGGAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAA	
224	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCAGTTCGAGAGGCAGAC CACCAGCCAGAAGCACAGGGACT TCGTGGCCGAGCCCATGGGCGAG AAGCCCGTGGGCAGCCTGGCCGG CATCGGCGAGGTGCTGGGCAAGA AGCTGGAGGAGAGGGGCTTCGAC AAGGCCTACGTGGTGCTGGGCCA GTTCCTGGTGCTGAAGAAGGACG AGGACCTGTTCAGGGAGTGGCTG AAGGACACCTGCGGCGCCAACGC CAAGCAGAGCAGGGACTGCTTCG GCTGCCTGAGGGAGTGGTGCGAC GCCTTCCTGTGGTCACATCCTCA ATTTGAGAAATGGTCACATCCGC AATTTGAAAAGTGGAGTCATCCT CAATTTGAAAAGCCCGCCGCCAA GAGGGTGAAGCTGGACGCCAGCC AGGGCACCAAGAGGAGCTACGAG CAGATGGAGACGGACGGCGAGAG GCAGAGCGCCTAAATCTGGTTAC CACTAAACCAGCCTCAAGAACAC CCGAATGGAGTCTCTAAGCTACA TAATACCAACTTACACTTACAAA ATGTTGTCCCCCAAAATGTAGCC ATTCGTATCTGCTCCTAATAAAA AGAAAGTTTCTTCACATTTCTTC TCTCCCTTGCACCTCGGATCAGT ACTAGACTGATAGCTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCAGUUCGAGAGGCAGACCAC CAGCCAGAAGCACAGGGACUUCGUGGCC GAGCCCAUGGGCGAGAAGCCCGUGGGCA GCCUGGCCGGCAUCGGCGAGGUGCUGGG CAAGAAGCUGGAGGAGAGGGGCUUCGAC AAGGCCUACGUGGUGCUGGGCCAGUUCC UGGUGCUGAAGAAGGACGAGGACCUGUU CAGGGAGUGGCUGAAGGACACCUGCGGC GCCAACGCCAAGCAGAGCAGGGACUGCU UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUGGUCACAUCCUCAAUUUGAG AAAUGGUCACAUCCGCAAUUUGAAAAGU GGAGUCAUCCUCAAUUUGAAAAGCCCGC CGCCAAGAGGGUGAAGCUGGACGCCAGC CAGGGCACCAAGAGGAGCUACGAGCAGA UGGAGACGGACGGCGAGAGGCAGAGCGC CUAAAUCUGGUUACCACUAAACCAGCCU CAAGAACACCCGAAUGGAGUCUCUAAGC UACAUAAUACCAACUUACACUUACAAAA UGUUGUCCCCCAAAAUGUAGCCAUUCGU AUCUGCUCCUAAUAAAAAGAAAGUUUCU UCACAUUUCUUCUCUCCCUUGCACCUCG GAUCAGUACUAGACUGAUAGCUUAUUGG GAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAA	MQFERQTTSQKHRDFVA EPMGEKPVGSLAGIGEV LGKKLEERGFDKAYVVL GQFLVLKKDEDLFREWL KDTCGANAKQSRDCFGC LREWCDAFLWSHPQFEK WSHPQFEKWSHPQFEKP AAKRVKLDASQGTKRSY EQMETDGERQSA

225	TAATACGACTCACTA□AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC□TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGTGGAGCCACCCCCAGTT CGAGAAGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGACCACCAGCCAGAA GCACAGGGACTTCGTGGCCGAGC CCATGGGCGAGAAGCCCGTGGGC AGCCTGGCCGGCATCGGCGAGGT GCTGGGCAAGAAGCTGGAGGAGA GGGGCTTCGACAAGGCCTACGTG GTGCTGGGCCAGTTCCTGGTGCT GAAGAAGGACGAGGACCTGTTCA GGGAGTGGCTGAAGGACACCTGC GGCGCCAACGCCAAGCAGAGCAG GGACTGCTTCGGCTGCCTGAGGG AGTGGTGCGACGCCTTCCTGTAA ATCTGGTTACCACTAAACCAGCC TCAAGAACACCCGAA□GGAGTCT CTAAGCTACATAATACCAACTTA CACTTACAAAATGTTGTCCCCCA AAATGTAGCCATTCG□ATCTGCT CCTAATAAAAAGAAAGTTTCTTC ACATTTCTTCTCTCCCTTGCACC TCGGATCAGTACTAGACTGATAG CTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGUGGAGCCACCCCCAGUUCGA GAAGUGGAGCCACCCCCAGUUCGAGAAG UGGAGCCACCCCCAGUUCGAGAAGACCA CCAGCCAGAAGCACAGGGACUUCGUGGC CGAGCCCAUGGGCGAGAAGCCCGUGGGC AGCCUGGCCGGCAUCGGCGAGGUGCUGG GCAAGAAGCUGGAGGAGAGGGGCUUCGA CAAGGCCUACGUGGUGCUGGGCCAGUUC CUGGUGCUGAAGAAGGACGAGGACCUGU UCAGGGAGUGGCUGAAGGACACCUGCGG CGCCAACGCCAAGCAGAGCAGGGACUGC UUCGGCUGCCUGAGGGAGUGGUGCGACG CCUUCCUGUAAAUCUGGUUACCACUAAA CCAGCCUCAAGAACACCCGAAUGGAGUC UCUAAGCUACAUAAUACCAACUUACACU UACAAAAUGUUGUCCCCCAAAAUGUAGC CAUUCGUAUCUGCUCCUAAUAAAAAGAA AGUUUCUUCACAUUUCUUCUCUCCCUUG CACCUCGGAUCAGUACUAGACUGAUAGC UUAUUGGGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAA	
226	TAATACGACTCACTA□AAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTC□TTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGACCACCAGCCAGAAGCA CAGGGACTTCGTGGCCGAGCCCA TGGGCGAGAAGCCCG□GGGCAGC CTGGCCGGCATCGGCGAGGTGCT GGGCAAGAAGCTGGAGGAGAGGG GCTTCGACAAGGCCTACGTGGTG CTGGGCCAGTTCCTGGTGCTGAA GAAGGACGAGGACCTGTTCAGGG AGTGGCTGAAGGACACCTGCGGC GCCAACGCCAAGCAGAGCAGGGA CTGCTTCGGCTGCCTGAGGGAGT GGTGCGACGCCTTCC□GTGGAGC CACCCCCAGTTCGAGAAGTGGAG CCACCCCCAGTTCGAGAAGTGGA GCCACCCCCAGTTCGAGAAGCCC GCCGCCAAGAGGGTGAAGCTGGA CGCCACCAAGGGCACCAAGAGGA GCTACGAGCAGATGGAGACCGAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGACCACCAGCCAGAAGCACAG GGACUUCGUGGCCGAGCCCAUGGGCGAG AAGCCCGUGGGCAGCCUGGCCGGCAUCG GCGAGGUGCUGGGCAAGAAGCUGGAGGA GAGGGGCUUCGACAAGGCCUACGUGGUG CUGGGCCAGUUCCUGGUGCUGAAGAAGG ACGAGGACCUGUUCAGGGAGUGGCUGAA GGACACCUGCGGCGCCAACGCCAAGCAG AGCAGGGACUGCUUCGGCUGCCUGAGGG AGUGGUGCGACGCCUUCCUGUGGAGCCA CCCCCAGUUCGAGAAGUGGAGCCACCCC CAGUUCGAGAAGUGGAGCCACCCCCAGU UCGAGAAGCCCGCCGCCAAGAGGGUGAA GCUGGACGCCACCAAGGGCACCAAGAGG AGCUACGAGCAGAUGGAGACCGACGGCG AGAGGCAGAGCGCCUAAAUCUGGUUACC ACUAAACCAGCCUCAAGAACACCCGAAU GGAGUCUCUAAGCUACAUAAUACCAACU UACACUUACAAAAUGUUGUCCCCCAAAA UGUAGCCAUUCGUAUCUGCUCCUAAUAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDATKGTKRSYEQMET DGERQSA

	GGCGAGAGGCAGAGCGCCTAAAT CTGGTTACCACTAAACCAGCCTC AAGAACACCCGAATGGAGTCTCT AAGCTACATAATACCAACTTACA CTTACAAAATGTTGTCCCCCAAA ATGTAGCCATTCGTATCTGCTCC TAATAAAAAGAAAGTTTCTTCAC ATTTCTTCTCTCCCTTGCACCTC GGATCAGTACTAGACTGATAGCT TATTGGG	AAAGAAAGUUUCUUCACAUUUCUUCUCU CCCUUGCACCUCGGAUCAGUACUAGACU GAUAGCUUAUUGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA	
227	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAAAGACGAGGAGTACGA GCAGATGGTCAAGGAGCTGCAGG AGATCCGCCTCGAGTGGAGCCAC CCCCAGTTCGAGAAGTGGAGCCA CCCCCAGTTCGAGAAGTGGAGCC ACCCCCAGTTCGAGAAGACCACC AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGAGGTGCTGGGCAAGAAGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CAGGGTTCAGGGAGTGGCTGAAG GACACCTGCGGCGCCAACGCCAA GCAGAGCAGGGACTGCTTCGGCT GCCTGAGGGAGTGGTGCGACGCC TTCCTGTAAATCTGGTTACCACT AAACCAGCCTCAAGAACACCCGA ATGGAGTCTCTAAGCTACATAAT ACCAACTTACACTTACAAAATGT TGTCCCCCAAAATGTAGCCATTC GTATCTGCTCCTAATAAAAAGAA AGTTTCTTCACATTTCTTCTCTC CCTTGCACCTCGGATCAGTACTA GACTGATAGCTTATTGGG	GGGAUAAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAAAGACGAGGAGUACGAGCA GAUGGUCAAGGAGCUGCAGGAGAUCCGC CUCGAGUGGAGCCACCCCCAGUUCGAGA AGUGGAGCCACCCCCAGUUCGAGAAGUG GAGCCACCCCCAGUUCGAGAAGACCACC AGCCAGAAGCACAGGGACUUCGUGGCCG AGCCCAUGGGCGAGAAGCCCGUGGGCAG CCUGGCCGGCAUCGGCGAGGUGCUGGGC AAGAAGCUGGAGGAGAGGGGCUUCGACA AGGCCUACGUGGUGCUGGGCCAGUUCCU GGUGCUGAAGAAGGACGAGGACAGGGUU CAGGGAGUGGCUGAAGGACACCUGCGGC GCCAACGCCAAGCAGAGCAGGGACUGCU UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUAAAUCUGGUUACCACUAAAC CAGCCUCAAGAACACCCGAAUGGAGUCU CUAAGCUACAUAAUACCAACUUACACUU ACAAAAUGUUGUCCCCCAAAAUGUAGCC AUUCGUAUCUGCUCCUAAUAAAAAGAAA GUUUCUUCACAUUUCUUCUCCCCUUGC ACCUCGGAUCAGUACUAGACUGAUAGCU UAUUGGGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
228	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCCCTGGTTCCCCCAGGA GGGCACCATCGACATCAAGAGGT GGAGGAGGGTCGGCTGGAGCCAC CCCCAGTTCGAGAAGTGGAGCCA CCCCCAGTTCGAGAAGTGGAGCC ACCCCCAGTTCGAGAAGACCACC	GGGAUAAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCCCUGGUUCCCCCAGGAGGG CACCAUCGACAUCAAGAGGUGGAGGAGG GUCGGCUGGAGCCACCCCCAGUUCGAGA AGUGGAGCCACCCCCAGUUCGAGAAGUG GAGCCACCCCCAGUUCGAGAAGACCACC AGCCAGAAGCACAGGGACUUCGUGGCCG AGCCCAUGGGCGAGAAGCCCGUGGGCAG CCUGGCCGGCAUCGGCGAGGUGCUGGGC	MPWFPQEGTIDIKRWRR VGWSHPQFEKWSHPQFE KWSHPQFEKTTSQKHRD FVAEPMGEKPVGSLAGI GEVLGKKLEERGFDKAY VVLGQFLVLKKDEDRFR EWLKDTCGANAKQSRDC FGCLREWCDAFL

	AGCCAGAAGCACAGGGACTTCGT GGCCGAGCCCATGGGCGAGAAGC CCGTGGGCAGCCTGGCCGGCATC GGCGAGGTGCTGGGCAAGAAGCT GGAGGAGAGGGGCTTCGACAAGG CCTACGTGGTGCTGGGCCAGTTC CTGGTGCTGAAGAAGGACGAGGA CAGGTTCAGGGAGTGGCTGAAGG ACACCTGCGGCGCCAACGCCAAG CAGAGCAGGGACTGCTTCGGCTG CCTGAGGGAGTGGTGCGACGCCT TCCTGTAAATCTGGTTACCACTA AACCAGCCTCAAGAACACCCGAA TGGAGTCTCTAAGCTACATAATA CCAACTTACACTTACAAAATGTT GTCCCCCAAAATGTAGCCATTCG TATCTGCTCCTAATAAAAAGAAA GTTTCTTCACATTTCTTCTCTCC CTTGCACCTCGGATCAGTACTAG ACTGATAGCTTATTGGG	AAGAAGCUGGAGGAGAGGGGCUUCGACA AGGCCUACGUGGUGCUGGGCCAGUUCCU GGUGCUGAAGAAGGACGAGGACAGGUUC AGGGAGUGGCUGAAGGACACCUGCGGCG CCAACGCCAAGCAGAGCAGGGACUGCUU CGGCUGCCUGAGGGAGUGGUGCGACGCC UUCCUGUAAAUCUGGUUACCACUAAACC AGCCUCAAGAACACCCGAAUGGAGUCUC UAAGCUACAUAAUACCAACUUACACUUA CAAAAUGUUGUCCCCCAAAAUGUAGCCA UUCGUAUCUGCUCCUAAUAAAAAGAAAG UUUCUUCACAUUUCUUCUCUCCCUUGCA CCUCGGAUCAGUACUAGACUGAUAGCUU AUUGGGAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	
229	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGGCAACGAGCTGGCCCT GAAGCTGGCCGGCCTGGACATCT GGAGCCACCCCCAGTTCGAGAAG TGGAGCCACCCCCAGTTCGAGAA GTGGAGCCACCCCCAGTTCGAGA AGACCACCAGCCAGAAGCACAGG GACTTCGTGGCCGAGCCCATGGG CGAGAAGCCCGTGGGCAGCCTGG CCGGCATCGGCGAGGTGCTGGGC AAGAAGCTGGAGGAGAGGGGCTT CGACAAGGCCTACGTGGTGCTGG GCCAGTTCCTGGTGCTGAAGAAG GACGAGGACAGGTTCAGGGAGTG GCTGAAGGACACCTGCGGCGCCA ACGCCAAGCAGAGCAGGGACTGC TTCGGCTGCCTGAGGGAGTGGTG CGACGCCTTCCTGTAAATCTGGT TACCACTAAACCAGCCTCAAGAA CACCCGAATGGAGTCCTAAGCT ACATAATACCAACTTACACTTAC AAAATGTTGTCCCCCAAAATGTA GCCATTCGTATCTGCTCCTAATA AAAAGAAAGTTTCTTCACATTTC TTCTCTCCCTTGCACCTCGGATC AGTACTAGACTGATAGCTTATTG GG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGGCAACGAGCUGGCCCUGAA GCUGGCCGGCCUGGACAUCUGGAGCCAC CCCCAGUUCGAGAAGUGGAGCCACCCCC AGUUCGAGAAGUGGAGCCACCCCCAGUU CGAGAAGACCACCAGCCAGAAGCACAGG GACUUCGUGGCCGAGCCCAUGGGCGAGA AGCCCGUGGGCAGCCUGGCCGGCAUCGG CGAGGUGCUGGGCAAGAAGCUGGAGGAG AGGGGCUUCGACAAGGCCUACGUGGUGC UGGGCCAGUUCCUGGUGCUGAAGAAGGA CGAGGACAGGUUCAGGGAGUGGCUGAAG GACACCUGCGGCGCCAACGCCAAGCAGA GCAGGGACUGCUUCGGCUGCCUGAGGGA GUGGUGCGACGCCUUCCUGUAAAUCUGG UUACCACUAAACCAGCCUCAAGAACACC CGAAUGGAGUCUCUAAGCUACAUAAUAC CAACUUACACUUACAAAAUGUUGUCCCC CAAAAUGUAGCCAUUCGUAUCUGCUCCU AAUAAAAAGAAAGUUUCUUCACAUUUCU UCUCUCCCUUGCACCUCGGAUCAGUACU AGACUGAUAGCUUAUUGGGAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA	MGNELALKLAGLDIWSH PQFEKWSHPQFEKWSHP QFEKTTSQKHRDFVAEP MGEKPVGSLAGIGEVLG KKLEERGFDKAYVVLGQ FLVLKKDEDRFREWLKD TCGANAKQSRDCFGCLR EWCDAFL

230	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MGNELALKLAGLDIGSG
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	SGSTTSQKHRDFVAEPM
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	GEKPVGSLAGIGEVLGK
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	KLEERGFDKAYVVLGQF
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	LVLKKDEDRFREWLKDT
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGGGCAACGAGCUGGCCCUGAA	CGANAKQSRDCFGCLRE
	TCACCATTTACGAACGATAGGCC	GCUGGCCGGCCUGGACAUCGGCAGCGGC	WCDAFLWSHPQFEKWSH
	ACCATGGGCAACGAGCTGGCCCT	AGCGGCAGCACCACCAGCCAGAAGCACA	PQFEKWSHPQFEKPAAK
	GAAGCTGGCCGGCCTGGACATCG	GGGACUUCGUGGCCGAGCCCAUGGGCGA	RVKLDASQGTKRSYEQM
	GCAGCGGCAGCGGCAGCACCACC	GAAGCCCGUGGGCAGCCUGGCCGGCAUC	ETDGERQSA
	AGCCAGAAGCACAGGGACTTCGT	GGCGAGGUGCUGGGCAAGAAGCUGGAGG	
	GGCCGAGCCCATGGGCGAGAAGC	AGAGGGGCUUCGACAAGGCCUACGUGGU	
	CCGTGGGCAGCCTGGCCGGCATC	GCUGGGCCAGUUCCUGGUGCUGAAGAAG	
	GGCGAGGTGCTGGGCAAGAAGCT	GACGAGGACAGGUUCAGGGAGUGGCUGA	
	GGAGGAGAGGGGCTTCGACAAGG	AGGACACCUGCGGCGCCAACGCCAAGCA	
	CCTACGTGGTGCTGGGCCAGTTC	GAGCAGGGACUGCUUCGGCUGCCUGAGG	
	CTGGTGCTGAAGAAGGACGAGGA	GAGUGGUGCGACGCCUUCCUGUGGAGCC	
	CAGGTTCAGGGAGTGGCTGAAGG	ACCCCCAGUUCGAGAAGUGGAGCCACCC	
	ACACCTGCGGCGCCAACGCCAAG	CCAGUUCGAGAAGUGGAGCCACCCCCAG	
	CAGAGCAGGGACTGCTTCGGCTG	UUCGAGAAGCCCGCCGCCAAGAGGGUGA	
	CCTGAGGGAGTGGTGCGACGCCT	AGCUGGACGCCAGCCAGGGCACCAAGAG	
	TCCTGTGGAGCCACCCCCAGTTC	GAGCUACGAGCAGAUGGAGACGGACGGC	
	GAGAAGTGGAGCCACCCCCAGTT	GAGAGGCAGAGCGCCUAAUAAAUCUGGU	
	CGAGAAGTGGAGCCACCCCCAGT	UACCACUAAACCAGCCUCAAGAACACCC	
	TCGAGAAGCCCGCCGCCAAGAGG	GAAUGGAGUCUCUAAGCUACAUAAUACC	
	GTGAAGCTGGACGCCAGCCAGGG	AACUUACACUUACAAAAUGUUGUCCCCC	
	CACCAAGAGGAGCTACGAGCAGA	AAAAUGUAGCCAUUCGUAUCUGCUCCUA	
	TGGAGACGGACGGCGAGAGGCAG	AUAAAAAGAAAGUUUCUUCACAUUUCUU	
	AGCGCCTAATAAATCTGGTTACC	CUCUCCCUUGCACCUCGGAUCAGUACUA	
	ACTAAACCAGCCTCAAGAACACC	GACUGAUAGCUUAUUGGGAAAAAAAAAA	
	CGAATGGAGTCTCTAAGCTACAT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AATACCAACTTACACTTACAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGTTGTCCCCCAAAATGTAGCCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTCGTATCTGCTCCTAATAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAAAGTTTCTTCACATTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCTCGGATCAGTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAA	
231	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MTTSQKHRDFVAEPMGE
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	KPVGSLAGIGEVLGKKL
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	EERGFDKAYVVLGQFLV
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	LKKDEDRFREWLKDTCG
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	ANAKQSRDCFGCLREWC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGACCACCAGCCAGAAGCACAG	DAFLWSHPQFEKWSHPQ
	TCACCATTTACGAACGATAGGCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	FEKWSHPQFEKPAAKRV
	ACCATGACCACCAGCCAGAAGCA	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	KLDASQGTKRSYEQMET
	CAGGGACTTCGTGGCCGAGCCCA	GCGAGGUGCUGGGCAAGAAGCUGGAGGA	DGERQSA
	TGGGCGAGAAGCCCGTGGGCAGC	GAGGGGCUUCGACAAGGCCUACGUGGUG	
	CTGGCCGGCATCGGCGAGGTGCT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	
	GGGCAAGAAGCTGGAGGAGAGGG	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	GCTTCGACAAGGCCTACGTGGTG	GGACACCUGCGGCGCCAACGCCAAGCAG	
	CTGGGCCAGTTCCTGGTGCTGAA	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GAAGGACGAGGACAGGTTCAGGG	AGUGGUGCGACGCCUUCCUGUGGAGUCA	
	AGTGGCTGAAGGACACCTGCGGC	UCCUCAAUUUGAAAAGUGGUCUCAUCCG	
	GCCAACGCCAAGCAGAGCAGGGA	CAAUUUGAAAAGUGGUCACAUCCUCAAU	
	CTGCTTCGGCTGCCTGAGGGAGT	UUGAGAAACCCGCCGCCAAGAGGGUGAA	
	GGTGCGACGCCTTCCTGTGGAGT	GCUGGACGCCAGCCAGGGCACCAAGAGG	
	CATCCTCAATTTGAAAAGTGGTC	AGCUACGAGCAGAUGGAGACGGACGGCG	

	TCATCCGCAATTTGAAAAGTGGT	AGAGGCAGAGCGCCUAAAUCUGGUUACC	
	CACATCCTCAATTTGAGAAACCC	ACUAAACCAGCCUCAAGAACACCCGAAU	
	GCCGCCAAGAGGGTGAAGCTGGA	GGAGUCUCUAAGCUACAUAAUACCAACU	
	CGCCAGCCAGGGCACCAAGAGGA	UACACUUACAAAAUGUUGUCCCCCAAAA	
	GCTACGAGCAGATGGAGACGGAC	UGUAGCCAUUCGUAUCUGCUCCUAAUAA	
	GGCGAGAGGCAGAGCGCCTAAAT	AAAGAAAGUUUCUUCACAUUUCUUCUCU	
	CTGGTTACCACTAAACCAGCCTC	CCCUUGCACCUCGGAUCAGUACUAGACU	
	AAGAACACCCGAATGGAGTCTCT	GAUAGCUUAUUGGGAAAAAAAAAAAAAA	
	AAGCTACATAATACCAACTTACA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTACAAAATGTTGTCCCCCAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAGCCATTCGTATCTGCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAATAAAAAGAAAGTTTCTTCAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATTTCTTCTCTCCCTTGCACCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGATCAGTACTAGACTGATAGCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TATTGGG	AAAAAAAAAAAAAAAAA	
232	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MKFERQTTSQKHRDFVA
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	EPMGEKPVGSLAGIGEV
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	LGKKLEERGFDKAYVVL
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	GQFLVLKKDEDRFREWL
	AATTTAAATCATTTCTTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	KDTCGANAKQSRDCFGC
	CAAAAGCAATTTTCTGAAAATTT	CCACCAUGAAGUUCGAGAGGCAGACCAC	LREWCDAFLWSHPQFEK
	TCACCATTTACGAACGATAGGCC	CAGCCAGAAGCACAGGGACUUCGUGGCC	WSHPQFEKWSHPQFEKP
	ACCATGAAGTTCGAGAGGCAGAC	GAGCCCAUGGGCGAGAAGCCCGUGGGCA	AAKRVKLDASQGTKRSY
	CACCAGCCAGAAGCACAGGGACT	GCCUGGCCGGCAUCGGCGAGGUGCUGGG	EQMETDGERQSA
	TCGTGGCCGAGCCCATGGGCGAG	CAAGAAGCUGGAGGAGAGGGGCUUCGAC	
	AAGCCCGTGGGCAGCCTGGCCGG	AAGGCCUACGUGGUGCUGGGCCAGUUCC	
	CATCGGCGAGGTGCTGGGCAAGA	UGGUGCUGAAGAAGGACGAGGACAGGUU	
	AGCTGGAGGAGAGGGGCTTCGAC	CAGGGAGUGGCUGAAGGACACCUGCGGC	
	AAGGCCTACGTGGTGCTGGGCCA	GCCAACGCCAAGCAGAGCAGGGACUGCU	
	GTTCCTGGTGCTGAAGAAGGACG	UCGGCUGCCUGAGGGAGUGGUGCGACGC	
	AGGACAGGTTCAGGGAGTGGCTG	CUUCCUGUGGAGCCACCCCCAGUUCGAG	
	AAGGACACCTGCGGCGCCAACGC	AAGUGGAGCCACCCCCAGUUCGAGAAGU	
	CAAGCAGAGCAGGGACTGCTTCG	GGAGCCACCCCCAGUUCGAGAAGCCCGC	
	GCTGCCTGAGGGAGTGGTGCGAC	CGCCAAGAGGGUGAAGCUGGACGCCAGC	
	GCCTTCCTGTGGAGCCACCCCCA	CAGGGCACCAAGAGGAGCUACGAGCAGA	
	GTTCGAGAAGTGGAGCCACCCCC	UGGAGACGGACGGCGAGAGGCAGAGCGC	
	AGTTCGAGAAGTGGAGCCACCCC	CUAAAUCUGGUUACCACUAAACCAGCCU	
	CAGTTCGAGAAGCCCGCCGCCAA	CAAGAACACCCGAAUGGAGUCUCUAAGC	
	GAGGGTGAAGCTGGACGCCAGCC	UACAUAAUACCAACUUACACUUACAAAA	
	AGGGCACCAAGAGGAGCTACGAG	UGUUGUCCCCCAAAAUGUAGCCAUUCGU	
	CAGATGGAGACGGACGGCGAGAG	AUCUGCUCCUAAUAAAAAGAAAGUUUCU	
	GCAGAGCGCCTAAATCTGGTTAC	UCACAUUUCUUCUCUCCCUUGCACCUCG	
	CACTAAACCAGCCTCAAGAACAC	GAUCAGUACUAGACUGAUAGCUUAUUGG	
	CCGAATGGAGTCTCTAAGCTACA	GAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAATACCAACTTACACTTACAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTTGTCCCCCAAAATGTAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATTCGTATCTGCTCCTAATAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGAAAGTTTCTTCACATTTCTTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCTCCCTTGCACCTCGGATCAGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACTAGACTGATAGCTTATTGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
		AAAAA	

233	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGGACAGCGGCATCCACAG CACCACCAGCCAGAAGCACAGGG ACTTCGTGGCCGAGCCCATGGGC GAGAAGCCCGTGGGCAGCCTGGC CGGCATCGGCGAGGTGCTGGGCA AGAAGCTGGAGGAGAGGGGCTTC GACAAGGCCTACGTGGTGCTGGG CCAGTTCCTGGTGCTGAAGAAGG ACGAGGACAGGTTCAGGGAGTGG CTGAAGGACACCTGCGGCGCCAA CGCCAAGCAGAGCAGGGACTGCT TCGGCTGCCTGAGGGAGTGGTGC GACGCCTTCCTGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGTGGAGCCAC CCCCAGTTCGAGAAGCCCGCCGC CAAGAGGGTGAAGCTGGACGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACGGACGGCGA GAGGCAGAGCGCCTAAATCTGGT TACCACTAAACCAGCCTCAAGAA CACCCGAATGGAGTCTCTAAGCT ACATAATACCAACTTACACTTAC AAAATGTTGTCCCCCAAAATGTA GCCATTCGTATCTGCTCCTAATA AAAAGAAAGTTTCTTCACATTTC TTCTCTCCCTTGCACCTCGGATC AGTACTAGACTGATAGCTTATTG GG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGGACAGCGGCAUCCACAGCAC CACCAGCCAGAAGCACAGGGACUUCGUG GCCGAGCCCAUGGGCGAGAAGCCCGUGG GCAGCCUGGCCGGCAUCGGCGAGGUGCU GGGCAAGAAGCUGGAGGAGAGGGGCUUC GACAAGGCCUACGUGGUGCUGGGCCAGU UCCUGGUGCUGAAGAAGGACGAGGACAG GUUCAGGGAGUGGCUGAAGGACACCUGC GGCGCCAACGCCAAGCAGAGCAGGGACU GCUUCGGCUGCCUGAGGGAGUGGUGCGA CGCCUUCCUGUGGAGCCACCCCCAGUUC GAGAAGUGGAGCCACCCCCAGUUCGAGA AGUGGAGCCACCCCCAGUUCGAGAAGCC CGCCGCCAAGAGGGUGAAGCUGGACGCC AGCCAGGGCACCAAGAGGAGCUACGAGC AGAUGGAGACGGACGGCGAGAGGCAGAG CGCCUAAAUCUGGUUACCACUAAACCAG CCUCAAGAACACCCGAAUGGAGUCUCUA AGCUACAUAAUACCAACUUACACUUACA AAAUGUUGUCCCCCAAAAUGUAGCCAUU CGUAUCUGCUCCUAAUAAAAAGAAAGUU UCUUCACAUUUCUUCUCUCCCUUGCACC UCGGAUCAGUACUAGACUGAUAGCUUAU UGGGAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAA	
234	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGCAGTTCGAGAGGCAGAC CACCAGCCAGAAGCACAGGGACT TCGTGGCCGAGCCCATGGGCGAG AAGCCCGTGGGCAGCCTGGCCGG CATCGGCGAGGTGCTGGGCAAGA AGCTGGAGGAGAGGGGCTTCGAC AAGGCCTACGTGGTGCTGGGCCA GTTCCTGGTGCTGAAGAAGGACG AGGACAGGTTCAGGGAGTGGCTG AAGGACACCTGCGGCGCCAACGC CAAGCAGAGCAGGGACTGCTTCG GCTGCCTGAGGGAGTGGTGCGAC GCCTTCCTGTGGTCACATCCACA ATTTGAAAAGTGGTCACATCCTC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGCAGUUCGAGAGGCAGACCAC CAGCCAGAAGCACAGGGACUUCGUGGCC GAGCCCAUGGGCGAGAAGCCCGUGGGCA GCCUGGCCGGCAUCGGCGAGGUGCUGGG CAAGAAGCUGGAGGAGAGGGGCUUCGAC AAGGCCUACGUGGUGCUGGGCCAGUUCC UGGUGCUGAAGAAGGACGAGGACAGGUU CAGGGAGUGGCUGAAGGACACCUGCGGC GCCAACGCCAAGCAGAGCAGGGACUGCU UCGGCUGCCUGAGGGAGUGGUGCGACGC CUUCCUGUGGUCACAUCCACAAUUUGAA AAGUGGUCACAUCCUCAAUUUGAGAAAU GGUCCCAUCCUCAAUUUGAAAAGCCCGC CGCCAAGAGGGUGAAGCUGGACGCCAGC CAGGGCACCAAGAGGAGCUACGAGCAGA UGGAGACGGACGGCGAGAGGCAGAGCGC	

	AATTTGAGAAATGGTCCCATCCT CAATTTGAAAAGCCCGCCGCCAA GAGGGTGAAGCTGGACGCCAGCC AGGGCACCAAGAGGAGCTACGAG CAGATGGAGACGGACGGCGAGAG GCAGAGCGCCTAAATCTGGTTAC CACTAAACCAGCCTCAAGAACAC CCGAATGGAGTCTCTAAGCTACA TAATACCAACTTACACTTACAAA ATGTTGTCCCCCAAAATGTAGCC ATTCGTATCTGCTCCTAATAAAA AGAAAGTTTCTTCACATTTCTTC TCTCCCTTGCACCTCGGATCAGT ACTAGACTGATAGCTTATTGGG	CUAAAUCUGGUUACCACUAAACCAGCCU CAAGAACACCCGAAUGGAGUCUCUAAGC UACAUAAUACCAACUUACACUUACAAAA UGUUGUCCCCCAAAAUGUAGCCAUUCGU AUCUGCUCCUAAUAAAAAGAAAGUUUCU UCACAUUUCUUCUCUCCCUUGCACCUCG GAUCAGUACUAGACUGAUAGCUUAUUGG GAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAA	
235	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGTGGAGCCACCCCCAGTT CGAGAAGTGGAGCCACCCCCAGT TCGAGAAGTGGAGCCACCCCCAG TTCGAGAAGACCACCAGCCAGAA GCACAGGGACTTCGTGGCCGAGC CCATGGGCGAGAAGCCCGTGGGC AGCCTGGCCGGCATCGGCGAGGT GCTGGGCAAGAAGCTGGAGGAGA GGGGCTTCGACAAGGCCTACGTG GTGCTGGGCCAGTTCCTGGTGCT GAAGAAGGACGAGGACAGGTTCA GGGAGTGGCTGAAGGACACCTGC GGCGCCAACGCCAAGCAGAGCAG GGACTGCTTCGGCTGCCTGAGGG AGTGGTGCGACGCCTTCCTGTAA ATCTGGTTACCACTAAACCAGCC TCAAGAACACCCGAATGGAGTCT CTAAGCTACATAATACCAACTTA CACTTACAAAATGTTGTCCCCCA AAATGTAGCCATTCGTATCTGCT CCTAATAAAAAGAAAGTTTCTTC ACATTTCTTCTCTCCCTTGCACC TCGGATCAGTACTAGACTGATAG CTTATTGGG	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGUGGAGCCACCCCCAGUUCGA GAAGUGGAGCCACCCCCAGUUCGAGAAG UGGAGCCACCCCCAGUUCGAGAAGACCA CCAGCCAGAAGCACAGGGACUUCGUGGC CGAGCCCAUGGGCGAGAAGCCCGUGGGC AGCCUGGCCGGCAUCGGCGAGGUGCUGG GCAAGAAGCUGGAGGAGAGGGGCUUCGA CAAGGCCUACGUGGUGCUGGGCCAGUUC CUGGUGCUGAAGAAGGACGAGGACAGGU UCAGGGAGUGGCUGAAGGACACCUGCGG CGCCAACGCCAAGCAGAGCAGGGACUGC UUCGGCUGCCUGAGGGAGUGGUGCGACG CCUUCCUGUAAAUCUGGUUACCACUAAA CCAGCCUCAAGAACACCCGAAUGGAGUC UCUAAGCUACAUAAUACCAACUUACACU UACAAAAUGUUGUCCCCCAAAAUGUAGC CAUUCGUAUCUGCUCCUAAUAAAAAGAA AGUUUCUUCACAUUUCUUCUCUCCCUUG CACCUCGGAUCAGUACUAGACUGAUAGC UUAUUGGGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAA	
236	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGACCACCAGCCAGAAGCA CAGGGACTTCGTGGCCGAGCCCA TGGGCGAGAAGCCCGTGGGCAGC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGACCACCAGCCAGAAGCACAG GGACUUCGUGGCCGAGCCCAUGGGCGAG AAGCCCGUGGGCAGCCUGGCCGGCAUCG GCGAGGUGCUGGGCAAGAAGCUGGAGGA GAGGGGCUUCGACAAGGCCUACGUGGUG	MTTSQKHRDFVAEPMGE KPVGSLAGIGEVLGKKL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ANAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDATKGTKRSYEQMET DGERQSA

	CTGGCCGGCATCGGCGAGGTGCT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	
	GGGCAAGAAGCTGGAGGAGAGGG	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	GCTTCGACAAGGCCTACGTGGTG	GGACACCUGCGGCGCCAACGCCAAGCAG	
	CTGGGCCAGTTCCTGGTGCTGAA	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GAAGGACGAGGACAGGTTCAGGG	AGUGGUGCGACGCCUUCCUGUGGAGCCA	
	AGTGGCTGAAGGACACCTGCGGC	CCCCCAGUUCGAGAAGUGGAGCCACCCC	
	GCCAACGCCAAGCAGAGCAGGGA	CAGUUCGAGAAGUGGAGCCACCCCCAGU	
	CTGCTTCGGCTGCCTGAGGGAGT	UCGAGAAGCCCGCCGCCAAGAGGGUGAA	
	GGTGCGACGCCTTCCTGTGGAGC	GCUGGACGCCACCAAGGGCACCAAGAGG	
	CACCCCCAGTTCGAGAAGTGGAG	AGCUACGAGCAGAUGGAGACGGACGGCG	
	CCACCCCCAGTTCGAGAAGTGGA	AGAGGCAGAGCGCCUAAAUCUGGUUACC	
	GCCACCCCCAGTTCGAGAAGCCC	ACUAAACCAGCCUCAAGAACACCCGAAU	
	GCCGCCAAGAGGGTGAAGCTGGA	GGAGUCUCUAAGCUACAUAAUACCAACU	
	CGCCACCAAGGGCACCAAGAGGA	UACACUUACAAAAUGUUGUCCCCCAAAA	
	GCTACGAGCAGATGGAGACGGAC	UGUAGCCAUUCGUAUCUGCUCCUAAUAA	
	GGCGAGAGGCAGAGCGCCTAAAT	AAAGAAAGUUUCUUCACAUUUCUUCUCU	
	CTGGTTACCACTAAACCAGCCTC	CCCUUGCACCUCGGAUCAGUACUAGACU	
	AAGAACACCCGAATGGAGTCTCT	GAUAGCUUAUUGGGAAAAAAAAAAAAAA	
	AAGCTACATAATACCAACTTACA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTACAAAATGTTGTCCCCCAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGTAGCCATTCGTATCTGCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAATAAAAAGAAAGTTTCTTCAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATTTCTTCTCTCCCTTGCACCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGATCAGTACTAGACTGATAGCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TATTGGG	AAAAAAAAAAAAAAAAA	
237	TAATACGACTCACTATAAGGGGA	GGGAUAAUAGUCAACACAACAUAUACAA	MNFQGLVLTDNCKNQWV
	TAATAGTCAACACAACATATACA	AACAAACGAAUCUCAAGCAAUCAAGCAU	VGPLAGKAGFASIYTTN
	AAACAAACGAATCTCAAGCAATC	UCUACUUCUAUUGCAGCAAUUUAAAUCA	DNNYVAGAGPKANGSLF
	AAGCATTCTACTTCTATTGCAGC	UUUCUUUUAAAGCAAAAGCAAUUUUCUG	TEQAFYTRVLKPSVIEE
	AATTTAAATCATTTCTTTAAAG	AAAAUUUUCACCAUUUACGAACGAUAGG	WKKSHNIKHVGLITCKA
	CAAAAGCAATTTTTCTGAAAATTT	CCACCAUGAACUUCCAGGGCCUGGUGCU	FGLYKSINVEYRFLVIN
	TCACCATTTACGAACGATAGGCC	GACCGACAACUGCAAGAACCAGUGGGUG	RLGADLDAVIRANNNRL
	ACCATGAACTTCCAGGGCCTGGT	GUGGGCCCCCUGGCCGGCAAGGCCGGCU	PKRSVMLIGIEILNTIQ
	GCTGACCGACAACTGCAAGAACC	UCGCCAGCAUCUACACCACCAACGACAA	FMHEQGYSHGDIKASNI
	AGTGGGTGGTGGGCCCCCTGGCC	CAACUACGUGGCCGGCGCCGGCCCCAAG	VLDQIDKNKLYLVDYGL
	GGCAAGGCCGGCTTCGCCAGCAT	GCCAACGGCAGCCUGUUCACCGAGCAGG	VSKFMSNGEHVPFIRNP
	CTACACCACCAACGACAACAACT	CCUUCUACACCAGGGUGCUGAAGCCCAG	NKMDNGTLEFTPIDSHK
	ACGTGGCCGGCGCCGGCCCCAAG	CGUGAUCGAGGAGUGGAAGAAGAGCCAC	GYVVSRRGDLDTLGYCM
	GCCAACGGCAGCCTGTTCACCGA	AACAUCAAGCACGUGGGCCUGAUCACCU	IRWLGGILPWTKISDTK
	GCAGGCCTTCTACACCAGGGTGC	GCAAGGCCUUCGGCCUGUACAAGAGCAU	NCALVSATKQKYVNNTA
	TGAAGCCCAGCGTGATCGAGGAG	CAACGUGGAGUACAGGUUCCUGGUGAUC	TLLMTSLQYAPRELLQY
	TGGAAGAAGAGCCACAACATCAA	AACAGGCUGGGCGCCGACCUGGACGCCG	ITMVNSLTYFEEPNYDK
	GCACGTGGGCCTGATCACCTGCA	UGAUCAGGGCCAACAACAACAGGCUGCC	FRHILMQGVYY
	AGGCCTTCGGCCTGTACAAGAGC	CAAGAGGAGCGUGAUGCUGAUCGGCAUC	
	ATCAACGTGGAGTACAGGTTCCT	GAGAUCCUGAACACCAUCCAGUUCAUGC	
	GGTGATCAACAGGCTGGGCGCCG	ACGAGCAGGGCUACAGCCACGGCGACAU	
	ACCTGGACGCCGTGATCAGGGCC	CAAGGCCAGCAACAUCGUGCUGGACCAG	
	AACAACAACAGGCTGCCCAAGAG	AUCGACAAGAACAAGCUGUACCUGGUGG	
	GAGCGTGATGCTGATCGGCATCG	ACUACGGCCUGGUGAGCAAGUUCAUGAG	
	AGATCCTGAACACCATCCAGTTC	CAACGGCGAGCACGUGCCCUUCAUCAGG	
	ATGCACGAGCAGGGCTACAGCCA	AACCCCAACAAGAUGGACAACGGCACCC	
	CGGCGACATCAAGGCCAGCAACA	UGGAGUUCACCCCCAUCGACAGCCACAA	
	TCGTGCTGGACCAGATCGACAAG	GGGCUACGUGGUGAGCAGGAGGGGCGAC	
	AACAAGCTGTACCTGGTGGACTA	CUGGACACCCUGGGCUACUGCAUGAUCA	
	CGGCCTGGTGAGCAAGTTCATGA	GGUGGCUGGGCGGCAUCCUGCCCUGGAC	
	GCAACGGCGAGCACGTGCCCTTC	CAAGAUCAGCGACACCAAGAACUGCGCC	
	ATCAGGAACCCCAACAAGATGGA	CUGGUGAGCGCCACCAAGCAGAAGUACG	

	CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
238	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGATC GGCAAGGGCGGCTTCGGCAGCAT CTACACCACCAACGACAACAACT ACGTGGTGAAGATCGAGCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACGCCCA CGGCGACATCGGCGCCGGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGGCGCCGGCTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGAUCGGCAAGGGCGGCU UCGGCAGCAUCUACACCACCAACGACAA CAACUACGUGGUGAAGAUCGAGCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACGCCCACGGCGACAU CGGCGCCGGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGGCGCCGGCUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC	MNFQGLVLTDNCKNQWV VGPLIGKGGFGSIYTTN DNNYVVKIEPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLIGIEILNTIQ FMHEQGYAHGDIGAGNI VLDQIDKNKLYLVDYGL GAGFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYY

	CCATCGACAGCCACAAGGGCTAC GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	CAGCCUGCAGUACGCCCCCAGGGAGCUG CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
239	TAATACGACTCACTATAAGGGGA TAATAGTCAACACAACATATACA AAACAAACGAATCTCAAGCAATC AAGCATTCTACTTCTATTGCAGC AATTTAAATCATTTCTTTTAAAG CAAAAGCAATTTTCTGAAAATTT TCACCATTTACGAACGATAGGCC ACCATGAACTTCCAGGGCCTGGT GCTGACCGACAACTGCAAGAACC AGTGGGTGGTGGGCCCCCTGGCC GGCAAGGCCGGCTTCGCCAGCAT CTACACCACCAACGACAACAACT ACGTGGCCGGCGCCGGCCCCAAG GCCAACGGCAGCCTGTTCACCGA GCAGGCCTTCTACACCAGGGTGC TGAAGCCCAGCGTGATCGAGGAG TGGAAGAAGAGCCACAACATCAA GCACGTGGGCCTGATCACCTGCA AGGCCTTCGGCCTGTACAAGAGC ATCAACGTGGAGTACAGGTTCCT GGTGATCAACAGGCTGGGCGCCG ACCTGGACGCCGTGATCAGGGCC AACAACAACAGGCTGCCCAAGAG GAGCGTGATGCTGATCGGCATCG AGATCCTGAACACCATCCAGTTC ATGCACGAGCAGGGCTACGCCCA CGGCGACATCGGCGCCGGCAACA TCGTGCTGGACCAGATCGACAAG AACAAGCTGTACCTGGTGGACTA CGGCCTGGGCGCCGGCTTCATGA GCAACGGCGAGCACGTGCCCTTC ATCAGGAACCCCAACAAGATGGA CAACGGCACCCTGGAGTTCACCC CCATCGACAGCCACAAGGGCTAC	GGGAUAAUAGUCAACACAACAUAUACAA AACAAACGAAUCUCAAGCAAUCAAGCAU UCUACUUCUAUUGCAGCAAUUUAAAUCA UUUCUUUUAAAGCAAAAGCAAUUUUCUG AAAAUUUUCACCAUUUACGAACGAUAGG CCACCAUGAACUUCCAGGGCCUGGUGCU GACCGACAACUGCAAGAACCAGUGGGUG GUGGGCCCCCUGGCCGGCAAGGCCGGCU UCGCCAGCAUCUACACCACCAACGACAA CAACUACGUGGCCGGCGCCGGCCCCAAG GCCAACGGCAGCCUGUUCACCGAGCAGG CCUUCUACACCAGGGUGCUGAAGCCCAG CGUGAUCGAGGAGUGGAAGAAGAGCCAC AACAUCAAGCACGUGGGCCUGAUCACCU GCAAGGCCUUCGGCCUGUACAAGAGCAU CAACGUGGAGUACAGGUUCCUGGUGAUC AACAGGCUGGGGCGCCGACCUGGACGCCG UGAUCAGGGCCAACAACAACAGGCUGCC CAAGAGGAGCGUGAUGCUGAUCGGCAUC GAGAUCCUGAACACCAUCCAGUUCAUGC ACGAGCAGGGCUACGCCCACGGCGACAU CGGCGCCGGCAACAUCGUGCUGGACCAG AUCGACAAGAACAAGCUGUACCUGGUGG ACUACGGCCUGGGCGCCGGCUUCAUGAG CAACGGCGAGCACGUGCCCUUCAUCAGG AACCCCAACAAGAUGGACAACGGCACCC UGGAGUUCACCCCCAUCGACAGCCACAA GGGCUACGUGGUGAGCAGGAGGGGCGAC CUGGACACCCUGGGCUACUGCAUGAUCA GGUGGCUGGGCGGCAUCCUGCCCUGGAC CAAGAUCAGCGACACCAAGAACUGCGCC CUGGUGAGCGCCACCAAGCAGAAGUACG UGAACAACACCGCCACCCUGCUGAUGAC CAGCCUGCAGUACGCCCCCAGGGAGCUG	MNFQGLVLTDNCKNQWV VGPLAGKAGFASIYTTN DNNYVAGAGPKANGSLF TEQAFYTRVLKPSVIEE WKKSHNIKHVGLITCKA FGLYKSINVEYRFLVIN RLGADLDAVIRANNNRL PKRSVMLTGIEILNTIQ FMHEQGYAHGDIGAGNI VLDQIDKNKLYLVDYGL GAGFMSNGEHVPFIRNP NKMDNGTLEFTPIDSHK GYVVSRRGDLDTLGYCM IRWLGGILPWTKISDTK NCALVSATKQKYVNNTA TLLMTSLQYAPRELLQY ITMVNSLTYFEEPNYDK FRHILMQGVYY

	GTGGTGAGCAGGAGGGGCGACCT GGACACCCTGGGCTACTGCATGA TCAGGTGGCTGGGCGGCATCCTG CCCTGGACCAAGATCAGCGACAC CAAGAACTGCGCCCTGGTGAGCG CCACCAAGCAGAAGTACGTGAAC AACACCGCCACCCTGCTGATGAC CAGCCTGCAGTACGCCCCCAGGG AGCTGCTGCAGTACATCACCATG GTGAACAGCCTGACCTACTTCGA GGAGCCCAACTACGACAAGTTCA GGCACATCCTGATGCAGGGCGTG TACTACTAACTAGTGACTGACTA GGATCTGGTTACCACTAAACCAG CCTCAAGAACACCCGAATGGAGT CTCTAAGCTACATAATACCAACT TACACTTACAAAATGTTGTCCCC CAAAATGTAGCCATTCGTATCTG CTCCTAATAAAAAGAAAGTTTCT TCACATTTCTTCTCTCCCTTGCA CCTCGGATCAGTACTAGACTGAT AGCTTATTGGG	CUGCAGUACAUCACCAUGGUGAACAGCC UGACCUACUUCGAGGAGCCCAACUACGA CAAGUUCAGGCACAUCCUGAUGCAGGGC GUGUACUACUAACUAGUGACUGACUAGG AUCUGGUUACCACUAAACCAGCCUCAAG AACACCCGAAUGGAGUCUCUAAGCUACA UAAUACCAACUUACACUUACAAAAUGUU GUCCCCCAAAAUGUAGCCAUUCGUAUCU GCUCCUAAUAAAAAGAAAGUUUCUUCAC AUUUCUUCUCUCCCUUGCACCUCGGAUC AGUACUAGACUGAUAGCUUAUUGGGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	
240	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACGGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GACGTGCTGAGCAAGAGGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACAG GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCAAGCTTTTCTTCT CTCCCTTGCACCTTAATGAGAT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACGG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GACCACCAGCCAGAAGCACAGGGACUUC GUGGCCGAGCCCAUGGGCGAGAAGCCCG UGGGCAGCCUGGCCGGCAUCGGCGACGU GCUGAGCAAGAGGCUGGAGGAGAGGGGC UUCGACAAGGCCUACGUGGUGCUGGGCC AGUUCCUGGUGCUGAAGAAGGACGAGGA CAGGUUCAGGGAGUGGCUGAAGGACACC UGCGGCGCCAACGCCAAGCAGAGCAGGG ACUGCUUCGGCUGCCUGAGGGAGUGGUG CGACGCCUUCCUGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUUAAUGA GAUAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAA	MASQGTKRSYEQMETDG ERQPAAKRVKLDWSHPQ FEKWSHPQFEKWSHPQF EKTTSQKHRDFVAEPMG EKPVGSLAGIGDVLSKR LEERGFDKAYVVLGQFL VLKKDEDRFREWLKDTC GANAKQSRDCFGCLREW CDAFL

241	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACGGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GAGGTGCTGGGCAAGAAGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACCT GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCAAGCTTTTCTTCT CTCCCTTGCACCTTAATGAGAT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACGG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GACCACCAGCCAGAAGCACAGGGACUUC GUGGCCGAGCCCAUGGGCGAGAAGCCCG UGGGCAGCCUGGCCGGCAUCGGCGAGGU GCUGGGCAAGAAGCUGGAGGAGAGGGGC UUCGACAAGGCCUACGUGGUGCUGGGCC AGUUCCUGGUGCUGAAGAAGGACGAGGA CCUGUUCAGGGAGUGGCUGAAGGACACC UGCGGCGCCAACGCCAAGCAGAGCAGGG ACUGCUUCGGCUGCCUGAGGGAGUGGUG CGACGCCUUCCUGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUUAAUGA GAUAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAA	MASQGTKRSYEQMETDG ERQPAAKRVKLDWSHPQ FEKWSHPQFEKWSHPQF EKTTSQKHRDFVAEPMG EKPVGSLAGIGEVLGKK LEERGFDKAYVVLGQFL VLKKDEDLFREWLKDTC GANAKQSRDCFGCLREW CDAFL
242	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGGCCA GCCAGGGCACCAAGAGGAGCTAC GAGCAGATGGAGACGGACGGCGA GAGGCAGCCCGCCGCCAAGAGGG TGAAGCTGGACTGGAGCCACCCC CAGTTCGAGAAGTGGAGCCACCC CCAGTTCGAGAAGTGGAGCCACC CCCAGTTCGAGAAGACCACCAGC CAGAAGCACAGGGACTTCGTGGC CGAGCCCATGGGCGAGAAGCCCG TGGGCAGCCTGGCCGGCATCGGC GAGGTGCTGGGCAAGAAGCTGGA GGAGAGGGGCTTCGACAAGGCCT ACGTGGTGCTGGGCCAGTTCCTG GTGCTGAAGAAGGACGAGGACAG GTTCAGGGAGTGGCTGAAGGACA CCTGCGGCGCCAACGCCAAGCAG AGCAGGGACTGCTTCGGCTGCCT GAGGGAGTGGTGCGACGCCTTCC TGTAAAGACCAGACAGCTGGAGC CTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCCAGCCCCTC CTCCCCTTCCTGCACCCGTACCC	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGGCCAGCCAGGGCACC AAGAGGAGCUACGAGCAGAUGGAGACGG ACGGCGAGAGGCAGCCCGCCGCCAAGAG GGUGAAGCUGGACUGGAGCCACCCCCAG UUCGAGAAGUGGAGCCACCCCCAGUUCG AGAAGUGGAGCCACCCCCAGUUCGAGAA GACCACCAGCCAGAAGCACAGGGACUUC GUGGCCGAGCCCAUGGGCGAGAAGCCCG UGGGCAGCCUGGCCGGCAUCGGCGAGGU GCUGGGCAAGAAGCUGGAGGAGAGGGGC UUCGACAAGGCCUACGUGGUGCUGGGCC AGUUCCUGGUGCUGAAGAAGGACGAGGA CAGGUUCAGGGAGUGGCUGAAGGACACC UGCGGCGCCAACGCCAAGCAGAGCAGGG ACUGCUUCGGCUGCCUGAGGGAGUGGUG CGACGCCUUCCUGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC CCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCAAGC UUUUCUUCUCUCCCUUGCACCUUAAUGA GAUAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	MASQGTKRSYEQMETDG ERQPAAKRVKLDWSHPQ FEKWSHPQFEKWSHPQF EKTTSQKHRDFVAEPMG EKPVGSLAGIGEVLGKK LEERGFDKAYVVLGQFL VLKKDEDRFREWLKDTC GANAKQSRDCFGCLREW CDAFL

	CCGTGGTCTTTGAATAAAGTCTG AGTGGGCGGCAAGCTTTTCTTCT CTCCCTTGCACCTTAATGAGAT	AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAA	
243	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACAA CCTCCCAAAAGCACCGAGACTTC GTGGCAGAGCCCATGGGGGAAAA GCCAGTGGGGAGCCTGGCCGGGA TTGGTGACGTCCTGAGCAAGAGG CTGGAGGAAAGGGGCTTTGACAA GGCTTATGTGGTCCTTGGCCAGT TTCTGGTGCTAAAGAAAGATGAA GACCTCTTCCGAGAATGGCTGAA GGATACATGTGGTGCCAATGCCA AGCAGTCCCGGGACTGCTTTGGG TGCCTTCGAGAATGGTGTGATGC CTTCTTGCCCGCCGCCAAGAGGG TGAAGCTGGACGCCAGCCAGGGC ACCAAGAGGAGCTACGAGCAGAT GGAGACCGACGGCGAGAGGCAGT AAAGACCAGACAGCTGGAGCCTC GGTGGCCATGCTTCTTGCCCCTT GGGCCTCCCCCCAGCCCCTCCTC CCCTTCCTGCACCCGTACCCCCG TGGTCTTTGAATAAAGTCTGAGT GGGCGGCAAGCTTTTCTTCTCTC CCTTGCACCT	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACAACCUCCCAAAAG CACCGAGACUUCGUGGCAGAGCCCAUGG GGGAAAAGCCAGUGGGGAGCCUGGCCGG GAUUGGUGACGUCCUGAGCAAGAGGCUG GAGGAAAGGGGCUUUGACAAGGCUUAUG UGGUCCUUGGCCAGUUUCUGGUCCUAAA GAAAGAUGAAGACCUCUUCCGAGAAUGG CUGAAGGAUACAUGUGGUGCCAAUGCCA AGCAGUCCCGGGACUGCUUUGGGUGCCU UCGAGAAUGGUGUGAUGCCUUCUUGCCC GCCGCCAAGAGGGUGAAGCUGGACGCCA GCCAGGGCACCAAGAGGAGCUACGAGCA GAUGGAGACCGACGGCGAGAGGCAGUAA AGACCAGACAGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCCCCC AGCCCCUCCUCCCCUUCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCAAGCUUUUCUUCUCUCCCUU GCACCUAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDLFREWLKDTCG ANAKQSRDCFGCLREWC DAFLPAAKRVKLDASQG TKRSYEQMETDGERQ
244	TAATACGACTCACTATAAGGGGA TAATAGGAAAAATCTATAGCAGA AGTCAGCGGTAGACGCACGGCAT AGCATCCAACGCCACCATGACCA CCAGCCAGAAGCACAGGGACTTC GTGGCCGAGCCCATGGGCGAGAA GCCCGTGGGCAGCCTGGCCGGCA TCGGCGACGTGCTGAGCAAGAGG CTGGAGGAGAGGGGCTTCGACAA GGCCTACGTGGTGCTGGGCCAGT TCCTGGTGCTGAAGAAGGACGAG GACAGGTTCAGGGAGTGGCTGAA GGACACCTGCGGCGCCACCGCCA AGCAGAGCAGGGACTGCTTCGGC TGCCTGAGGGAGTGGTGCGACGC CTTCCTGTGGTCTCATCCACAAT TTGAGAAATGGTCACATCCACAA TTTGAGAAATGGAGCCACCCCCA GTTCGAGAAGCCCGCCGCCAAGA GGGTGAAGCTGGACGCCAGCCAG	GGGAUAAUAGGAAAAAUCUAUAGCAGAA GUCAGCGGUAGACGCACGGCAUAGCAUC CAACGCCACCAUGACCACCAGCCAGAAG CACAGGGACUUCGUGGCCGAGCCCAUGG GCGAGAAGCCCGUGGGCAGCCUGGCCGG CAUCGGCGACGUGCUGAGCAAGAGGCUG GAGGAGAGGGGCUUCGACAAGGCCUACG UGGUGCUGGGCCAGUUCCUGGUGCUGAA GAAGGACGAGGACAGGUUCAGGGAGUGG CUGAAGGACACCUGCGGCGCCACCGCCA AGCAGAGCAGGGACUGCUUCGGCUGCCU GAGGGAGUGGUGCGACGCCUUCCUGUGG UCUCAUCCACAAUUUGAGAAAUGGUCAC AUCCACAAUUUGAGAAAUGGAGCCACCC CCAGUUCGAGAAGCCCGCCGCCAAGAGG GUGAAGCUGGACGCCAGCCAGGGCACCA AGACGACCUACGAGCAGAUGGACACCGA CGGCGAGAGGCAGUAAAGACCAGACAGC UGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCC	MTTSQKHRDFVAEPMGE KPVGSLAGIGDVLSKRL EERGFDKAYVVLGQFLV LKKDEDRFREWLKDTCG ATAKQSRDCFGCLREWC DAFLWSHPQFEKWSHPQ FEKWSHPQFEKPAAKRV KLDASQGTKRSYEQMET DGERQ

	GGCACCAAGAGGAGCTACGAGCA	CCUUCCUGCACCCGUACCCCCGUGGUCU	
	GATGGAGACCGACGGCGAGAGGC	UUGAAUAAAGUCUGAGUGGGCGGCAAGC	
	AGTAAAGACCAGACAGCTGGAGC	UUUUCUUCUCUCCCUUGCACCUAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCTTGCACCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
		AAAAAAAAAAAAAAAAAAAAAAAAAA	
245	GCGTCAATTGACATTGGTCTGTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MATVKFKYKGEEKEVDI
	CATTTAATACGACTCACTATAAG	GUCAGCGGUAGACGCACGGCAUAGCAUC	SKIKKVWRVGKMISFTY
	GGGATAATAGGAAAAATCTATAG	CAACGCCACCAUGGCCACCGUGAAGUUC	DEGGGKTGRGAVSEKDA
	CAGAAGTCAGCGGTAGACGCACG	AAGUACAAGGGCGAGGAGAAGGAGGUGG	PKELLQMLEKQKKGGGS
	GCATAGCATCCAACGCCACCATG	ACAUCAGCAAGAUCAAGAAGGUGUGGAG	GGGSGGGSTTSQKHRDF
	GCCACCGTGAAGTTCAAGTACAA	GGUGGGCAAGAUGAUCAGCUUCACCUAC	VAEPMGEKPVGSLAGIG
	GGGCGAGGAGAAGGAGGTGGACA	GACGAGGGCGGCGGCAAGACCGGCAGGG	DVLSKRLEERGFDKAYV
	TCAGCAAGATCAAGAAGGTGTGG	GCGCCGUGAGCGAGAAGGACGCCCCCAA	VLGQFLVLKKDEDRFRE
	AGGGTGGGCAAGATGATCAGCTT	GGAGCUGCUGCAGAUGCUGGAGAAGCAG	WLKDTCGANAKQSRDCF
	CACCTACGACGAGGGCGGCGGCA	AAGAAGGGCGGCGGCAGCGGCGGCGGCA	GCLREWCDAF_WSHPQF
	AGACCGGCAGGGGCGCCGTGAGC	GCGGCGGCGGCAGCACCACCAGCCAGAA	EKWSHPQFEKWSHPQFE
	GAGAAGGACGCCCCCAAGGAGCT	GCACAGGGACUUCGUGGCCGAGCCCAUG	KPAAKRVKLDASQGTKR
	GCTGCAGATGCTGGAGAAGCAGA	GGCGAGAAGCCCGUGGGCAGCCUGGCCG	SYEQMETDGERQ
	AGAAGGGCGGCGGCAGCGGCGGC	GCAUCGGCGACGUGCUGAGCAAGAGGCU	
	GGCAGCGGCGGCGGCAGCACCAC	GGAGGAGAGGGGCUUCGACAAGGCCUAC	
	CAGCCAGAAGCACAGGGACTTCG	GUGGUGCUGGGCCAGUUCCUGGUGCUGA	
	TGGCCGAGCCCATGGGCGAGAAG	AGAAGGACGAGGACAGGUUCAGGGAGUG	
	CCCGTGGGCAGCCTGGCCGGCAT	GCUGAAGGACACCUGCGGCGCCAACGCC	
	CGGCGACGTGCTGAGCAAGAGGC	AAGCAGAGCAGGGACUGCUUCGGCUGCC	
	TGGAGGAGAGGGGCTTCGACAAG	UGAGGGAGUGGUGCGACGCCUUCCUGUG	
	GCCTACGTGGTGCTGGGCCAGTT	GAGCCACCCCCAGUUCGAGAAGUGGAGC	
	CCTGGTGCTGAAGAAGGACGAGG	CACCCCCAGUUCGAGAAGUGGAGCCACC	
	ACAGGTTCAGGGAGTGGCTGAAG	CCCAGUUCGAGAAGCCCGCCGCCAAGAG	
	GACACCTGCGGCGCCAACGCCAA	GGUGAAGCUGGACGCCAGCCAGGGCACC	
	GCAGAGCAGGGACTGCTTCGGCT	AAGAGGAGCUACGAGCAGAUGGAGACCG	
	GCCTGAGGGAGTGGTGCGACGCC	ACGGCGAGAGGCAGUAAAGACCAGACAG	
	TTCCTGTGGAGCCACCCCCAGTT	CUGGAGCCUCGGUGGCCAUGCUUCUUGC	
	CGAGAAGTGGAGCCACCCCCAGT	CCCUUGGGCCUCCCCCCAGCCCCUCCUC	
	TCGAGAAGTGGAGCCACCCCCAG	CCCUUCCUGCACCCGUACCCCCGUGGUC	
	TTCGAGAAGCCCGCCGCCAAGAG	UUUGAAUAAAGUCUGAGUGGGCGGCAAG	
	GGTGAAGCTGGACGCCAGCCAGG	CUUUUCUUCUCUCCCUUGCACCUAAAAA	
	GCACCAAGAGGAGCTACGAGCAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATGGAGACCGACGGCGAGAGGCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTAAAGACCAGACAGCTGGAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCGGTGGCCATGCTTCTTGCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGGGCCTCCCCCCAGCCCCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCCCTTCCTGCACCCGTACCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGTGGTCTTTGAATAAAGTCTGA	AAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGGCGGCAAGCTTTTCTTCTC		
	TCCCTTGCACCT		

246	GCGTCAATTGACATTGGTCTGTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MASQGTKRSYEQMETDG
	CATTTAATACGACTCACTATAAG	GUCAGCGGUAGACGCACGGCAUAGCAUC	ERQPAAKRVKLDGGGSG
	GGGATAATAGGAAAAATCTATAG	CAACGCCACCAUGGCCAGCCAGGGCACC	GGSGGGSTTSQKHRDFV
	CAGAAGTCAGCGGTAGACGCACG	AAGAGGAGCUACGAGCAGAUGGAGACGG	AEPMGEKPVGSLAGIGD
	GCATAGCATCCAACGCCACCATG	ACGGCGAGAGGCAGCCCGCCGCCAAGAG	VLSKRLEERGFDKAYVV
	GCCAGCCAGGGCACCAAGAGGAG	GGUGAAGCUGGACGGCGGCGGCAGCGGC	LGQFLVLKKDEDRFREW
	CTACGAGCAGATGGAGACGGACG	GGCGGCAGCGGCGGCGGCAGCACCACCA	LKDTCGANAKQSRDCFG
	GCGAGAGGCAGCCCGCCGCCAAG	GCCAGAAGCACAGGGACUUCGUGGCCGA	CLREWCDAFLWSHPQFE
	AGGGTGAAGCTGGACGGCGGCGG	GCCCAUGGGCGAGAAGCCCGUGGGCAGC	KWSHPQFEKWSHPQFEK
	CAGCGGCGGCGGCAGCGGCGGCG	CUGGCCGGCAUCGGCGACGUGCUGAGCA	ATVKFKYKGEEKEVDIS
	GCAGCACCACCAGCCAGAAGCAC	AGAGGCUGGAGGAGAGGGGCUUCGACAA	KIKKVWRVGKMISFTYD
	AGGGACTTCGTGGCCGAGCCCAT	GGCCUACGUGGUGCUGGGCCAGUUCCUG	EGGGKTGRGAVSEKDAP
	GGGCGAGAAGCCCGTGGGCAGCC	GUGCUGAAGAAGGACGAGGACAGGUUCA	KELLQMLEKQKK
	TGGCCGGCATCGGCGACGTGCTG	GGGAGUGGCUGAAGGACACCUGCGGCGC	
	AGCAAGAGGCTGGAGGAGAGGGG	CAACGCCAAGCAGAGCAGGGACUGCUUC	
	CTTCGACAAGGCCTACGTGGTGC	GGCUGCCUGAGGGAGUGGUGCGACGCCU	
	TGGGCCAGTTCCTGGTGCTGAAG	UCCUGUGGAGCCACCCCCAGUUCGAGAA	
	AAGGACGAGGACAGGTTCAGGGA	GUGGAGCCACCCCCAGUUCGAGAAGUGG	
	GTGGCTGAAGGACACCTGCGGCG	AGCCACCCCCAGUUCGAGAAGGCCACCG	
	CCAACGCCAAGCAGAGCAGGGAC	UGAAGUUCAAGUACAAGGGCGAGGAGAA	
	TGCTTCGGCTGCCTGAGGGAGTG	GGAGGUGGACAUCAGCAAGAUCAAGAAG	
	GTGCGACGCCTTCCTGTGGAGCC	GUGUGGAGGGUGGGCAAGAUGAUCAGCU	
	ACCCCCAGTTCGAGAAGTGGAGC	UCACCUACGACGAGGGCGGCGGCAAGAC	
	CACCCCCAGTTCGAGAAGTGGAG	CGGCAGGGGCGCCGUGAGCGAGAAGGAC	
	CCACCCCCAGTTCGAGAAGGCCA	GCCCCCAAGGAGCUGCUGCAGAUGCUGG	
	CCGTGAAGTTCAAGTACAAGGGC	AGAAGCAGAAGAAGUAAAGACCAGACAG	
	GAGGAGAAGGAGGTGGACATCAG	CUGGAGCCUCGGUGGCCAUGCUUCUUGC	
	CAAGATCAAGAAGGTGTGGAGGG	CCCUUGGGCCUCCCCCCAGCCCCUCCUC	
	TGGGCAAGATGATCAGCTTCACC	CCCUUCCUGCACCCGUACCCCCGUGGUC	
	TACGACGAGGGCGGCGGCAAGAC	UUUGAAUAAAGUCUGAGUGGGCGGCAAG	
	CGGCAGGGGCGCCGTGAGCGAGA	CUUUUCUUCUCUCCCUUGCACCUAAAAA	
	AGGACGCCCCCAAGGAGCTGCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CAGATGCTGGAGAAGCAGAAGAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTAAAGACCAGACAGCTGGAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCGGTGGCCATGCTTCTTGCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGGGCCTCCCCCCAGCCCCTCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TCCCCTTCCTGCACCCGTACCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGTGGTCTTTGAATAAAGTCTGA	AAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGGCGGCAAGCTTTTCTTCTC		
	TCCCTTGCACCT		
247	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQWSHPQFEKWSH
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PQFEKWSHPQFEKTTSQ
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KHRDFVAEPMGEKPVGS
	ACTCACTATAAGGGGATAATAGG	UGGAGCCACCCCCAGUUCGAGAAGUGGA	LAGIGEVLGKRLEERGF
	AAAAATCTATAGCAGAAGTCAGC	GCCACCCCCAGUUCGAGAAGUGGAGCCA	DKAYVVLGQFLVLKKDE
	GGTAGACGCACGGCATAGCATCC	CCCCCAGUUCGAGAAGACCACCAGCCAG	DRFREWLKDTCGANAKQ
	AACGCCACCATGAAGTTCGAGAG	AAGCACAGGGACUUCGUGGCCGAGCCCA	SRDCFGCLREWCDAFLW
	GCAGTGGAGCCACCCCCAGTTCG	UGGGCGAGAAGCCCGUGGGCAGCCUGGC	SHPQFEKWSHPQFEKWS
	AGAAGTGGAGCCACCCCCAGTTC	CGGCAUCGGCGAGGUGCUGGGCAAGAGG	HPQFEKPAAKRVKLD
	GAGAAGTGGAGCCACCCCCAGTT	CUGGAGGAGAGGGGCUUCGACAAGGCCU	
	CGAGAAGACCACCAGCCAGAAGC	ACGUGGUGCUGGGCCAGUUCCUGGUGCU	
	ACAGGGACTTCGTGGCCGAGCCC	GAAGAAGGACGAGGACAGGUUCAGGGAG	
	ATGGGCGAGAAGCCCGTGGGCAG	UGGCUGAAGGACACCUGCGGCGCCAACG	
	CCTGGCCGGCATCGGCGAGGTGC	CCAAGCAGAGCAGGGACUGCUUCGGCUG	
	TGGGCAAGAGGCTGGAGGAGAGG	CCUGAGGGAGUGGUGCGACGCCUUCCUG	
	GGCTTCGACAAGGCCTACGTGGT	UGGAGCCACCCCCAGUUCGAGAAGUGGA	
	GCTGGGCCAGTTCCTGGTGCTGA	GCCACCCCCAGUUCGAGAAGUGGCAGCCA	

	AGAAGGACGAGGACAGGTTCAGG	CCCCCAGUUCGAGAAGCCCGCCGCCAAG	
	GAGTGGCTGAAGGACACCTGCGG	AGGGUGAAGCUGGACUAAAGACCAGACA	
	CGCCAACGCCAAGCAGAGCAGGG	GCUGGAGCCUCGGUGGCCAUGCUUCUUG	
	ACTGCTTCGGCTGCCTGAGGGAG	CCCCUUGGGCCUCCCCCCAGCCCCUCCU	
	TGGTGCGACGCCTTCCTGTGGAG	CCCCUUCCUGCACCCGUACCCCCGUGGU	
	CCACCCCCAGTTCGAGAAGTGGA	CUUUGAAUAAAGUCUGAGUGGGCGGCAA	
	GCCACCCCCAGTTCGAGAAGTGG	GCUUUUCUUCUCUCCCUUGCACCUAAAA	
	AGCCACCCCCAGTTCGAGAAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGCCGCCAAGAGGGTGAAGCTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACTAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT		
	CTCCCTTGCACCT		
248	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQWSHPQFEKWSH
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PQFEKWSHPQFEKTTSQ
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KHRDFVAEPMGEKPVGS
	ACTCACTATAAGGGGATAATAGG	UGGAGCCACCCCCAGUUCGAGAAGUGGA	LAGIGEVLGKRLEERGF
	AAAAATCTATAGCAGAAGTCAGC	GCCACCCCCAGUUCGAGAAGUGGAGCCA	DKAYVVLGQFLVLKKDE
	GGTAGACGCACGGCATAGCATCC	CCCCCAGUUCGAGAAGACCACCAGCCAG	DRFREWLKDTCGANAKQ
	AACGCCACCATGAAGTTCGAGAG	AAGCACAGGGACUUCGUGGCCGAGCCCA	SRDCFGCLREWCDAFLW
	GCAGTGGAGCCACCCCCAGTTCG	UGGGCGAGAAGCCCGUGGGCAGCCUGGC	SHPQFEKWSHPQFEKWS
	AGAAGTGGAGCCACCCCCAGTTC	CGGCAUCGGCGAGGUGCUGGGCAAGAGG	HPQFEKPDTRKPPAAKR
	GAGAAGTGGAGCCACCCCCAGTT	CUGGAGGAGAGGGGCUUCGACAAGGCCU	VKLD
	CGAGAAGACCACCAGCCAGAAGC	ACGUGGUGCUGGGCCAGUUCCUGGUGCU	
	ACAGGGACTTCGTGGCCGAGCCC	GAAGAAGGACGAGGACAGGUUCAGGGAG	
	ATGGGCGAGAAGCCCGTGGGCAG	UGGCUGAAGGACACCUGCGGCGCCAACG	
	CCTGGCCGGCATCGGCGAGGTGC	CCAAGCAGAGCAGGGACUGCUUCGGCUG	
	TGGGCAAGAGGCTGGAGGAGAGG	CCUGAGGGAGUGGUGCGACGCCUUCCUG	
	GGCTTCGACAAGGCCTACGTGGT	UGGAGCCACCCCCAGUUCGAGAAGUGGA	
	GCTGGGCCAGTTCCTGGTGCTGA	GCCACCCCCAGUUCGAGAAGUGGAGCCA	
	AGAAGGACGAGGACAGGTTCAGG	CCCCCAGUUCGAGAAGCCCGACACCAGG	
	GAGTGGCTGAAGGACACCTGCGG	AAGCCCCCCGCCGCCAAGAGGGUGAAGC	
	CGCCAACGCCAAGCAGAGCAGGG	UGGACUAAAGACCAGACAGCUGGAGCCU	
	ACTGCTTCGGCTGCCTGAGGGAG	CGGUGGCCAUGCUUCUUGCCCCUUGGGC	
	TGGTGCGACGCCTTCCTGTGGAG	CUCCCCCCAGCCCCUCCUCCCCUUCCUG	
	CCACCCCCAGTTCGAGAAGTGGA	CACCCGUACCCCCGUGGUCUUUGAAUAA	
	GCCACCCCCAGTTCGAGAAGTGG	AGUCUGAGUGGGCGGCAAGCUUUUCUUC	
	AGCCACCCCCAGTTCGAGAAGCC	UCUCCCUUGCACCUAAAAAAAAAAAAAAA	
	CGACACCAGGAAGCCCCCCGCCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAAGAGGGTGAAGCTGGACTAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGACCAGACAGCTGGAGCCTCGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGCCATGCTTCTTGCCCCTTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCTCCCCCCAGCCCCTCCTCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCCTGCACCCGTACCCCCGTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTTTGAATAAAGTCTGAGTGG	AAAAAAAAAAAAAAAAAA	
	GCGGCAAGCTTTTCTTCTCTCCC		
	TTGCACCT		

249	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQWSHPQFEKWSH
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PQFEKWSHPQFEKTTSQ
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KHRDFVAEPMGEKPVGS
	ACTCACTATAAGGGGATAATAGG	UGGAGCCACCCCCAGUUCGAGAAGUGGA	LAGIGEVLGKRLEERGF
	AAAAATCTATAGCAGAAGTCAGC	GCCACCCCCAGUUCGAGAAGUGGAGCCA	DKAYVVLGQFLVLKKDE
	GGTAGACGCACGGCATAGCATCC	CCCCCAGUUCGAGAAGACCACCAGCCAG	DRFREWLKDTCGANKKQ
	AACGCCACCATGAAGTTCGAGAG	AAGCACAGGGACUUCGUGGCCGAGCCCA	SRDCFRCLREWCDAFLW
	GCAGTGGAGCCACCCCCAGTTCG	UGGGCGAGAAGCCCGUGGGCAGCCUGGC	SHPQFEKWSHPQFEKWS
	AGAAGTGGAGCCACCCCCAGTTC	CGGCAUCGGCGAGGUGCUGGGCAAGAGG	HPQFEKPAAKRVKLD
	GAGAAGTGGAGCCACCCCCAGTT	CUGGAGGAGAGGGGCUUCGACAAGGCCU	
	CGAGAAGACCACCAGCCAGAAGC	ACGUGGUGCUGGGCCAGUUCCUGGUGCU	
	ACAGGGACTTCGTGGCCGAGCCC	GAAGAAGGACGAGGACAGGUUCAGGGAG	
	ATGGGCGAGAAGCCCGTGGGCAG	UGGCUGAAGGACACCUGCGGCGCCAACA	
	CCTGGCCGGCATCGGCGAGGTGC	AGAAGCAGAGCAGGGACUGCUUCAGGUG	
	TGGGCAAGAGGCTGGAGGAGAGG	CCUGAGGGAGUGGUGCGACGCCUUCCUG	
	GGCTTCGACAAGGCCTACGTGGT	UGGAGCCACCCCCAGUUCGAGAAGUGGA	
	GCTGGGCCAGTTCCTGGTGCTGA	GCCACCCCCAGUUCGAGAAGUGGAGCCA	
	AGAAGGACGAGGACAGGTTCAGG	CCCCCAGUUCGAGAAGCCCGCCGCCAAG	
	GAGTGGCTGAAGGACACCTGCGG	AGGGUGAAGCUGGACUAAAGACCAGACA	
	CGCCAACAAGAAGCAGAGCAGGG	GCUGGAGCCUCGGUGGCCAUGCUUCUUG	
	ACTGCTTCAGGTGCCTGAGGGAG	CCCCUUGGGCCUCCCCCCAGCCCCUCCU	
	TGGTGCGACGCCTTCCTGTGGAG	CCCCUUCCUGCACCCGUACCCCCGUGGU	
	CCACCCCCAGTTCGAGAAGTGGA	CUUUGAAUAAAGUCUGAGUGGGCGGCAA	
	GCCACCCCCAGTTCGAGAAGTGG	GCUUUUCUUCUCUCCCUUGCACCUAAAA	
	AGCCACCCCCAGTTCGAGAAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGCCGCCAAGAGGGTGAAGCTGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACTAAAGACCAGACAGCTGGAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCGGTGGCCATGCTTCTTGCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTGGGCCTCCCCCCAGCCCCTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTCCCCTTCCTGCACCCGTACCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCGTGGTCTTTGAATAAAGTCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGTGGGCGGCAAGCTTTTCTTCT		
	CTCCCTTGCACCT		
250	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQWSHPQFEKWSH
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	PQFEKWSHPQFEKTTSQ
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KHRDFVAEPMGEKPVGS
	ACTCACTATAAGGGGATAATAGG	UGGAGCCACCCCCAGUUCGAGAAGUGGA	LAGIGEVLGKRLEERGF
	AAAAATCTATAGCAGAAGTCAGC	GCCACCCCCAGUUCGAGAAGUGGAGCCA	DKAYVVLGQFLVLKKDE
	GGTAGACGCACGGCATAGCATCC	CCCCCAGUUCGAGAAGACCACCAGCCAG	DRFREWLKDTCGANAKQ
	AACGCCACCATGAAGTTCGAGAG	AAGCACAGGGACUUCGUGGCCGAGCCCA	SRDCFGCLREWCDAFLW
	GCAGTGGAGCCACCCCCAGTTCG	UGGGCGAGAAGCCCGUGGGCAGCCUGGC	SHPQFEKWSHPQFEKWS
	AGAAGTGGAGCCACCCCCAGTTC	CGGCAUCGGCGAGGUGCUGGGCAAGAGG	HPQFEKPDTRKPPAAKR
	GAGAAGTGGAGCCACCCCCAGTT	CUGGAGGAGAGGGGCUUCGACAAGGCCU	VKLDASQGTKRSYEQME
	CGAGAAGACCACCAGCCAGAAGC	ACGUGGUGCUGGGCCAGUUCCUGGUGCU	TDGERQDDDDTEE
	ACAGGGACTTCGTGGCCGAGCCC	GAAGAAGGACGAGGACAGGUUCAGGGAG	
	ATGGGCGAGAAGCCCGTGGGCAG	UGGCUGAAGGACACCUGCGGCGCCAACG	
	CCTGGCCGGCATCGGCGAGGTGC	CCAAGCAGAGCAGGGACUGCUUCGGCUG	
	TGGGCAAGAGGCTGGAGGAGAGG	CCUGAGGGAGUGGUGCGACGCCUUCCUG	
	GGCTTCGACAAGGCCTACGTGGT	UGGAGCCACCCCCAGUUCGAGAAGUGGA	
	GCTGGGCCAGTTCCTGGTGCTGA	GCCACCCCCAGUUCGAGAAGUGGAGCCA	
	AGAAGGACGAGGACAGGTTCAGG	CCCCCAGUUCGAGAAGCCCGACACCAGG	
	GAGTGGCTGAAGGACACCTGCGG	AAGCCCCCCGCCGCCAAGAGGGUGAAGC	
	CGCCAACGCCAAGCAGAGCAGGG	UGGACGCCAGCCAGGGCACCAAGAGGAG	
	ACTGCTTCGGCTGCCTGAGGGAG	CUACGAGCAGAUGGAGACUGACGGCGAG	
	TGGTGCGACGCCTTCCTGTGGAG	AGGCAGGACGACGACGACACUGAGGAGU	
	CCACCCCCAGTTCGAGAAGTGGA	AAAGACCAGACAGCUGGAGCCUCGGUGG	
	GCCACCCCCAGTTCGAGAAGTGG	CCAUGCUUCUUGCCCCUUGGGCCUCCCC	

	AGCCACCCCCAGTTCGAGAAGCC	CCAGCCCCUCCUCCCCUUCCUGCACCCG	
	CGACACCAGGAAGCCCCCCGCCG	UACCCCCGUGGUCUUUGAAUAAAGUCUG	
	CCAAGAGGGTGAAGCTGGACGCC	AGUGGGCGGCAAGCUUUUCUUCUCUCCC	
	AGCCAGGGCACCAAGAGGAGCTA	UUGCACCUAAAAAAAAAAAAAAAAAAAAA	
	CGAGCAGATGGAGACtGACGGCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGAGGCAGGACGACGACGACACt	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGGAGTAAAGACCAGACAGCTG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GAGCCTCGGTGGCCATGCTTCTT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCCCCTTGGGCCTCCCCCCAGCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCTCCTCCCCTTCCTGCACCCGT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCCCCGTGGTCTTTGAATAAAG	AAAAAAAAAAAA	
	TCTGAGTGGGCGGCAAGCTTTTC		
	TTCTCTCCCTTGCACCT		
251	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQGGGSGGGSGGG
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	STTSQKHRDFVAEPMGE
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KPVGSLAGIGDVLSKRL
	ACTCACTATAAGGGGATAATAGG	GGCGGCGGCAGCGGCGGCGGCAGCGGCG	EERGFDKAYVVLGQFLV
	AAAAATCTATAGCAGAAGTCAGC	GCGGCAGCACCACCAGCCAGAAGCACAG	LKKDEDRFREWLKDTCG
	GGTAGACGCACGGCATAGCATCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	ANAKQSRDCFGCLREWC
	AACGCCACCATGAAGTTCGAGAG	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	DAFLWSHPQFEKWSHPQ
	GCAGGGCGGCGGCAGCGGCGGCG	GCGACGUGCUGAGCAAGAGGCUGGAGGA	FEKWSHPQFEKPAAKRV
	GCAGCGGCGGCGGCAGCACCACC	GAGGGGCUUCGACAAGGCCUACGUGGUG	KLDASQGTKRSYEQMET
	AGCCAGAAGCACAGGGACTTCGT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	DGERQ
	GGCCGAGCCCATGGGCGAGAAGC	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	CCGTGGGCAGCCTGGCCGGCATC	GGACACCUGCGGCGCCAACGCCAAGCAG	
	GGCGACGTGCTGAGCAAGAGGCT	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GGAGGAGAGGGGCTTCGACAAGG	AGUGGUGCGACGCCUUCCUGUGGUCUCA	
	CCTACGTGGTGCTGGGCCAGTTC	UCCACAAUUUGAGAAAUGGUCACAUCCA	
	CTGGTGCTGAAGAAGGACGAGGA	CAAUUUGAGAAAUGGAGCCACCCCCAGU	
	CAGGTTCAGGGAGTGGCTGAAGG	UCGAGAAGCCCGCCGCCAAGAGGGUGAA	
	ACACCTGCGGCGCCAACGCCAAG	GCUGGACGCCAGCCAGGGCACCAAGAGG	
	CAGAGCAGGGACTGCTTCGGCTG	AGCUACGAGCAGAUGGAGACCGACGGCG	
	CCTGAGGGAGTGGTGCGACGCCT	AGAGGCAGUAAAGACCAGACAGCUGGAG	
	TCCTGTGGTCTCATCCACAATTT	CCUCGGUGGCCAUGCUUCUUGCCCCUUG	
	GAGAAATGGTCACATCCACAATT	GGCCUCCCCCCAGCCCCUCCUCCCCUUC	
	TGAGAAATGGAGCCACCCCCAGT	CUGCACCCGUACCCCCGUGGUCUUUGAA	
	TCGAGAAGCCCGCCGCCAAGAGG	UAAAGUCUGAGUGGGCGGCAAGCUUUUC	
	GTGAAGCTGGACGCCAGCCAGGG	UUCUCUCCCUUGCACCUAAAAAAAAAAAA	
	CACCAAGAGGAGCTACGAGCAGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGAGACCGACGGCGAGAGGCAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TAAAGACCAGACAGCTGGAGCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGGTGGCCATGCTTCTTGCCCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGGGCCTCCCCCCAGCCCCTCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCTTCCTGCACCCGTACCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGGTCTTTGAATAAAGTCTGAG	AAAAAAAAAAAAAAAAAAAA	
	TGGGCGGCAAGCTTTTCTTCTCT		
	CCCTTGCACCT		
252	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MAYISKIPHCTKEEDAP
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	SDNDSGICGGGSGGGSG
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGGCCUACAUCAGCAAG	GGSTTSQKHRDFVAEPM
	ACTCACTATAAGGGGATAATAGG	AUCCCCCACUGCACCAAGGAGGAGGACG	GEKPVGSLAGIGDVLSK
	AAAAATCTATAGCAGAAGTCAGC	CCCCCAGCGACAACGACAGCGGCAUCUG	RLEERGFDKAYVVLGQF
	GGTAGACGCACGGCATAGCATCC	CGGCGGCGGCAGCGGCGGCGGCAGCGGC	LVLKKDEDRFREWLKDT
	AACGCCACCATGGCCTACATCAG	GGCGGCAGCACCACCAGCCAGAAGCACA	CGANAKQSRDCFGCLRE
	CAAGATCCCCCACTGCACCAAGG	GGGACUUCGUGGCCGAGCCCAUGGGCGA	WCDAFLWSHPQFEKWSH
	AGGAGGACGCCCCCAGCGACAAC	GAAGCCCGUGGGCAGCCUGGCCGGCAUC	PQFEKWSHPQFEKPAAK
	GACAGCGGCATCTGCGGCGCCGG	GGCCACCUGCUGAGCAAGAGGCUGGAGG	

	CAGCGGCGGCGGCAGCGGCGGCG	AGAGGGGCUUCGACAAGGCCUACGUGGU	RVKLDASQGTKRSYEQM
	GCAGCACCACCAGCCAGAAGCAC	GCUGGGCCAGUUCCUGGUGCUGAAGAAG	ETDGERQ
	AGGGACTTCGTGGCCGAGCCCAT	GACGAGGACAGGUUCAGGGAGUGGCUGA	
	GGGCGAGAAGCCCGTGGGCAGCC	AGGACACCUGCGGCGCCAACGCCAAGCA	
	TGGCCGGCATCGGCGACGTGCTG	GAGCAGGGACUGCUUCGGCUGCCUGAGG	
	AGCAAGAGGCTGGAGGAGAGGGG	GAGUGGUGCGACGCCUUCCUGUGGUCUC	
	CTTCGACAAGGCCTACGTGGTGC	AUCCACAAUUUGAGAAAUGGUCACAUCC	
	TGGGCCAGTTCCTGGTGCTGAAG	ACAAUUUGAGAAAUGGAGCCACCCCCAG	
	AAGGACGAGGACAGGTTCAGGGA	UUCGAGAAGCCCGCCGCCAAGAGGGUGA	
	GTGGCTGAAGGACACCTGCGGCG	AGCUGGACGCCAGCCAGGGCACCAAGAG	
	CCAACGCCAAGCAGAGCAGGGAC	GAGCUACGAGCAGAUGGAGACCGACGGC	
	TGCTTCGGCTGCCTGAGGGAGTG	GAGAGGCAGUAAAGACCAGACAGCUGGA	
	GTGCGACGCCTTCCTGTGGTCTC	GCCUCGGUGGCCAUGCUUCUUGCCCCUU	
	ATCCACAATTTGAGAAATGGTCA	GGGCCUCCCCCCAGCCCCUCCUCCCCUU	
	CATCCACAATTTGAGAAATGGAG	CCUGCACCCGUACCCCCGUGGUCUUUGA	
	CCACCCCCAGTTCGAGAAGCCCG	AUAAAGUCUGAGUGGGCGGCAAGCUUUU	
	CCGCCAAGAGGGTGAAGCTGGAC	CUUCUCUCCCUUGCACCUAAAAAAAAAAA	
	GCCAGCCAGGGCACCAAGAGGAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTACGAGCAGATGGAGACCGACG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCGAGAGGCAGTAAAGACCAGAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	AGCTGGAGCCTCGGTGGCCATGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTCTTGCCCCTTGGGCCTCCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CAGCCCCTCCTCCCCTTCCTGCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCGTACCCCCGTGGTCTTTGAA	AAAAAAAAAAAAAAAAAAAA	
	TAAAGTCTGAGTGGGCGGCAAGC		
	TTTTCTTCTCTCCCTTGCACCT		
253	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQGGGSGGGSGGG
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	STTSQKHRDFVAEPMGE
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KPVGSLAGIGEVLGKRL
	ACTCACTATAAGGGGATAATAGG	GGCGGCGGCAGCGGCGGCGGCAGCGGCG	EERGFDKAYVVLGQFLV
	AAAAATCTATAGCAGAAGTCAGC	GCGGCAGCACCACCAGCCAGAAGCACAG	LKKDEDRFREWLKDTCG
	GGTAGACGCACGGCATAGCATCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	ANAKQSRDCFGCLREWC
	AACGCCACCATGAAGTTCGAGAG	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	DAFLWSHPQFEKWSHPQ
	GCAGGGCGGCGGCAGCGGCGGCG	GCGAGGUGCUGGGCAAGAGGCUGGAGGA	FEKWSHPQFEKPAAKRV
	GCAGCGGCGGCGGCAGCACCACC	GAGGGGCUUCGACAAGGCCUACGUGGUG	KLD
	AGCCAGAAGCACAGGGACTTCGT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	
	GGCCGAGCCCATGGGCGAGAAGC	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	CCGTGGGCAGCCTGGCCGGCATC	GGACACCUGCGGCGCCAACGCCAAGCAG	
	GGCGAGGTGCTGGGCAAGAGGCT	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GGAGGAGAGGGGCTTCGACAAGG	AGUGGUGCGACGCCUUCCUGUGGAGCCA	
	CCTACGTGGTGCTGGGCCAGTTC	CCCCCAGUUCGAGAAGUGGAGCCACCCC	
	CTGGTGCTGAAGAAGGACGAGGA	CAGUUCGAGAAGUGGAGCCACCCCCAGU	
	CAGGTTCAGGGAGTGGCTGAAGG	UCGAGAAGCCCGCCGCCAAGAGGGUGAA	
	ACACCTGCGGCGCCAACGCCAAG	GCUGGACUAAAGACCAGACAGCUGGAGC	
	CAGAGCAGGGACTGCTTCGGCTG	CUCGGUGGCCAUGCUUCUUGCCCCUUGG	
	CCTGAGGGAGTGGTGCGACGCCT	GCCUCCCCCCAGCCCCUCCUCCCCUUCC	
	TCCTGTGGAGCCACCCCCAGTTC	UGCACCCGUACCCCCGUGGUCUUUGAAU	
	GAGAAGTGGAGCCACCCCCAGTT	AAAGUCUGAGUGGGCGGCAAGCUUUUCU	
	CGAGAAGTGGAGCCACCCCCAGT	UCUCUCCCUUGCACCUAAAAAAAAAAAA	
	TCGAGAAGCCCGCCGCCAAGAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGAAGCTGGACTAAAGACCAGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CAGCTGGAGCCTCGGTGGCCATG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCTTGCCCCTTGGGCCTCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAGCCCCTCCTCCCCTTCCTGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCCGTACCCCCGTGGTCTTTGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATAAAGTCTGAGTGGGCGGCAAG	AAAAAAAAAAAAAAAAAAAA	
	CTTTTCTTCTCTCCCTTGCACCT		

254	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQGGGSGGGSGGG
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	STTSQKHRDFVAEPMGE
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KPVGSLAGIGEVLGKRL
	ACTCACTATAAGGGGATAATAGG	GGCGGCGGCAGCGGCGGCGGCAGCGGCG	EERGFDKAYVVLGQFLV
	AAAAATCTATAGCAGAAGTCAGC	GCGGCAGCACCACCAGCCAGAAGCACAG	LKKDEDRFREWLKDTCG
	GGTAGACGCACGGCATAGCATCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	ANAKQSRDCFGCLREWC
	AACGCCACCATGAAGTTCGAGAG	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	DAFLWSHPQFEKWSHPQ
	GCAGGGCGGCGGCAGCGGCGGCG	GCGAGGUGCUGGGCAAGAGGCUGGAGGA	FEKWSHPQFEKPDTRKP
	GCAGCGGCGGCGGCAGCACCACC	GAGGGGCUUCGACAAGGCCUACGUGGUG	PAAKRVKLD
	AGCCAGAAGCACAGGGACTTCGT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	
	GGCCGAGCCCATGGGCGAGAAGC	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	CCGTGGGCAGCCTGGCCGGCATC	GGACACCUGCGGCGCCAACGCCAAGCAG	
	GGCGAGGTGCTGGGCAAGAGGCT	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GGAGGAGAGGGGCTTCGACAAGG	AGUGGUGCGACGCCUUCCUGUGGAGCCA	
	CCTACGTGGTGCTGGGCCAGTTC	CCCCCAGUUCGAGAAGUGGAGCCACCCC	
	CTGGTGCTGAAGAAGGACGAGGA	CAGUUCGAGAAGUGGAGCCACCCCCAGU	
	CAGGTTCAGGGAGTGGCTGAAGG	UCGAGAAGCCCGACACCAGGAAGCCCCC	
	ACACCTGCGGCGCCAACGCCAAG	CGCCGCCAAGAGGGUGAAGCUGGACUAA	
	CAGAGCAGGGACTGCTTCGGCTG	AGACCAGACAGCUGGAGCCUCGGUGGCC	
	CCTGAGGGAGTGGTGCGACGCCT	AUGCUUCUUGCCCCUUGGGCCUCCCCCC	
	TCCTGTGGAGCCACCCCCAGTTC	AGCCCCUCCUCCCCUUCCUGCACCCGUA	
	GAGAAGTGGAGCCACCCCCAGTT	CCCCCGUGGUCUUUGAAUAAAGUCUGAG	
	CGAGAAGTGGAGCCACCCCCAGT	UGGGCGGCAAGCUUUUCUUCUCUCCCUU	
	TCGAGAAGCCCGACACCAGGAAG	GCACCUAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCCCGCCGCCAAGAGGGTGAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GCTGGACTAAAGACCAGACAGCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GGAGCCTCGGTGGCCATGCTTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TGCCCCTTGGGCCTCCCCCCAGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCTCCTCCCCTTCCTGCACCCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TACCCCCGTGGTCTTTGAATAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTCTGAGTGGGCGGCAAGCTTTT	AAAAAAAAAA	
	CTTCTCTCCCTTGCACCT		
255	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQGGGSGGGSGGG
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	STTSQKHRDFVAEPMGE
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KPVGSLAGIGEVLGKRL
	ACTCACTATAAGGGGATAATAGG	GGCGGCGGCAGCGGCGGCGGCAGCGGCG	EERGFDKAYVVLGQFLV
	AAAAATCTATAGCAGAAGTCAGC	GCGGCAGCACCACCAGCCAGAAGCACAG	LKKDEDRFREWLKDTCG
	GGTAGACGCACGGCATAGCATCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	ANKKQSRDCFRCLREWC
	AACGCCACCATGAAGTTCGAGAG	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	DAFLWSHPQFEKWSHPQ
	GCAGGGCGGCGGCAGCGGCGGCG	GCGAGGUGCUGGGCAAGAGGCUGGAGGA	FEKWSHPQFEKPAAKRV
	GCAGCGGCGGCGGCAGCACCACC	GAGGGGCUUCGACAAGGCCUACGUGGUG	KLD
	AGCCAGAAGCACAGGGACTTCGT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	
	GGCCGAGCCCATGGGCGAGAAGC	ACGAGGACAGGUUCAGGGAGUGGCUGAA	
	CCGTGGGCAGCCTGGCCGGCATC	GGACACCUGCGGCGCCAACAAGAAGCAG	
	GGCGAGGTGCTGGGCAAGAGGCT	AGCAGGGACUGCUUCAGGUGCCUGAGGG	
	GGAGGAGAGGGGCTTCGACAAGG	AGUGGUGCGACGCCUUCCUGUGGAGCCA	
	CCTACGTGGTGCTGGGCCAGTTC	CCCCCAGUUCGAGAAGUGGAGCCACCCC	
	CTGGTGCTGAAGAAGGACGAGGA	CAGUUCGAGAAGUGGAGCCACCCCCAGU	
	CAGGTTCAGGGAGTGGCTGAAGG	UCGAGAAGCCCGCCGCCAAGAGGGUGAA	
	ACACCTGCGGCGCCAACAAGAAG	GCUGGACUAAAGACCAGACAGCUGGAGC	
	CAGAGCAGGGACTGCTTCAGGTG	CUCGGUGGCCAUGCUUCUUGCCCCUUGG	
	CCTGAGGGAGTGGTGCGACGCCT	GCCUCCCCCCAGCCCCUCCUCCCCUUCC	
	TCCTGTGGAGCCACCCCCAGTTC	UGCACCCGUACCCCCGUGGUCUUUGAAU	
	GAGAAGTGGAGCCACCCCCAGTT	AAAGUCUGAGUGGGCGGCAAGCUUUUCU	
	CGAGAAGTGGAGCCACCCCCAGT	UCUCUCCCUUGCACCUAAAAAAAAAAAA	
	TCGAGAAGCCCGCCGCCAAGAGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	GTGAAGCTGGACTAAAGACCAGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

	CAGCTGGAGCCTCGGTGGCCATG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTTCTTGCCCCTTGGGCCTCCCC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCAGCCCCTCCTCCCCTTCCTGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACCCGTACCCCCGTGGTCTTTGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ATAAAGTCTGAGTGGGCGGCAAG	AAAAAAAAAAAAAAAAAAAA	
	CTTTTCTTCTCTCCCTTGCACCT		
256	AAGACTAGAGGTAGAGCAGCCTA	GGGAUAAUAGGAAAAAUCUAUAGCAGAA	MKFERQGGGSGGGSGGG
	ATACGACTCTACGCGTCAATTGA	GUCAGCGGUAGACGCACGGCAUAGCAUC	STTSQKHRDFVAEPMGE
	CATTGGTCTGTACATTTAATACG	CAACGCCACCAUGAAGUUCGAGAGGCAG	KPVGSLAGIGEVLGKRL
	ACTCACTATAAGGGGATAATAGG	GGCGGCGGCAGCGGCGGCGGCAGCGGCG	EERGFDKAYVVLGQFLV
	AAAAATCTATAGCAGAAGTCAGC	GCGGCAGCACCACCAGCCAGAAGCACAG	LKKDEDRFREWLKDTCG
	GGTAGACGCACGGCATAGCATCC	GGACUUCGUGGCCGAGCCCAUGGGCGAG	ANAKQSRDCFGCLREWC
	AACGCCACCATGAAGTTCGAGAG	AAGCCCGUGGGCAGCCUGGCCGGCAUCG	DAFLWSHPQFEKWSHPQ
	GCAGGGCGGCGGCAGCGGCGGCG	GCGAGGUGCUGGGCAAGAGGCUGGAGGA	FEKWSHPQFEKPDTRKP
	GCAGCGGCGGCGGCAGCACCACC	GAGGGGCUUCGACAAGGCCUACGUGGUG	PAAKRVKLDASQGTKRS
	AGCCAGAAGCACAGGGACTTCGT	CUGGGCCAGUUCCUGGUGCUGAAGAAGG	YEQMETDGERQDDDDTE
	GGCCGAGCCCATGGGCGAGAAGC	ACGAGGACAGGUUCAGGGAGUGGCUGAA	E
	CCGTGGGCAGCCTGGCCGGCATC	GGACACCUGCGGCGCCAACGCCAAGCAG	
	GGCGAGGTGCTGGGCAAGAGGCT	AGCAGGGACUGCUUCGGCUGCCUGAGGG	
	GGAGGAGAGGGGCTTCGACAAGG	AGUGGUGCGACGCCUUCCUGUGGAGCCA	
	CCTACGTGGTGCTGGGCCAGTTC	CCCCCAGUUCGAGAAGUGGAGCCACCCC	
	CTGGTGCTGAAGAAGGACGAGGA	CAGUUCGAGAAGUGGAGCCACCCCCAGU	
	CAGGTTCAGGGAGTGGCTGAAGG	UCGAGAAGCCCGACACCAGGAAGCCCCC	
	ACACCTGCGGCGCCAACGCCAAG	CGCCGCCAAGAGGGUGAAGCUGGACGCC	
	CAGAGCAGGGACTGCTTCGGCTG	AGCCAGGGCACCAAGAGGAGCUACGAGC	
	CCTGAGGGAGTGGTGCGACGCCT	AGAUGGAGACUGACGGCGAGAGGCAGGA	
	TCCTGTGGAGCCACCCCCAGTTC	CGACGACGACACUGAGGAGUAAAGACCA	
	GAGAAGTGGAGCCACCCCCAGTT	GACAGCUGGAGCCUCGGUGGCCAUGCUU	
	CGAGAAGTGGAGCCACCCCCAGT	CUUGCCCCUUGGGCCUCCCCCCAGCCCC	
	TCGAGAAGCCCGACACCAGGAAG	UCCUCCCCUUCCUGCACCCGUACCCCCG	
	CCCCCCGCCGCCAAGAGGGTGAA	UGGUCUUUGAAUAAAGUCUGAGUGGGCG	
	GCTGGACGCCAGCCAGGGCACCA	GCAAGCUUUUCUUCUCUCCCUUGCACCU	
	AGAGGAGCTACGAGCAGATGGAG	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	ACTGACGGCGAGAGGCAGGACGA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CGACGACACTGAGGAGTAAAGAC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CAGACAGCTGGAGCCTCGGTGGC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CATGCTTCTTGCCCCTTGGGCCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CCCCCCAGCCCCTCCTCCCCTTC	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	CTGCACCCGTACCCCCGTGGTCT	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	
	TTGAATAAAGTCTGAGTGGGCGG	AAAA	
	CAAGCTTTTCTTCTCTCCCTTGC		
	ACCT		

In at least some of the previously described embodiments, one or more elements used in an embodiment can interchangeably be used in another embodiment unless such a replacement is not technically feasible. It will be appreciated by those skilled in the art that various other omissions, additions and modifications may be made to the methods and structures described above without departing from the scope of the claimed subject matter. All such modifications and changes are intended to fall within the scope of the subject matter, as defined by the appended claims.

It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (*e.g.*, bodies of the appended claims) are generally intended as "open" terms (*e.g.*, the term "including" should be interpreted as "including but not limited to," the term "having" should be interpreted as "having at least," the term "includes" should be interpreted as "includes but is not limited to," etc.). It will be further understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory phrases "at least one" and "one or more" to introduce claim recitations. However, the use of such phrases should not be construed to imply that the introduction of a claim recitation by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim recitation to embodiments containing only one such recitation, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an" (*e.g.*, "a" and/or "an" should be interpreted to mean "at least one" or "one or more"); the same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation is explicitly recited, those skilled in the art will recognize that such recitation should be interpreted to mean at least the recited number (*e.g.*, the bare recitation of "two recitations," without other modifiers, means at least two recitations, or two or more recitations). Furthermore, in those instances where a convention analogous to "at least one of A, B, and C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, " a system having at least one of A, B, and C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). In those instances where a convention analogous to "at least one of A, B, or C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, " a system having at least one of A, B, or C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). It will be further understood by those within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the

description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase "A or B" will be understood to include the possibilities of "A" or "B" or "A and B."

In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

As will be understood by one skilled in the art, for any and all purposes, such as in terms of providing a written description, all ranges disclosed herein also encompass any and all possible sub-ranges and combinations of sub-ranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, etc. As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, etc. As will also be understood by one skilled in the art all language such as "up to," "at least," "greater than," "less than," and the like include the number recited and refer to ranges which can be subsequently broken down into sub-ranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 articles refers to groups having 1, 2, or 3 articles. Similarly, a group having 1-5 articles refers to groups having 1, 2, 3, 4, or 5 articles, and so forth.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Accordingly, the preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. Moreover,

nothing disclosed herein is intended to be dedicated to the public regardless of whether such disclosure is explicitly recited in the claims.

The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and described herein. Rather, the scope and spirit of present invention is embodied by the appended claims. In the claims, 35 U.S.C. §112(f) or 35 U.S.C. §112(6) is expressly defined as being invoked for a limitation in the claim only when the exact phrase "means for" or the exact phrase "step for" is recited at the beginning of such limitation in the claim; if such exact phrase is not used in a limitation in the claim, then 35 U.S.C. § 112 (f) or 35 U.S.C. §112(6) is not invoked.

WHAT IS CLAIMED IS:

1. A composition for the delivery of DNA to the nucleus of a cell, the composition comprising: a cytosolic delivery vehicle, a DNA, and an agent that modulates the protein Barrier to Autointegration Factor (BAF).

2. The composition according to claim 1, wherein the agent is a protein comprising an amino acid sequence having a sequence identity of 90% or more to wild type BAF, or an mRNA encoding the protein comprising an amino acid sequence having a sequence identity of 90% or more to wild type BAF.

3. The composition according to claim 2, wherein the protein comprises a mutation relative to wild type human BAF selected from the group consisting of T2D, T3D, S4D, A12T, E28D, G31S, K33R, E36D, R37K, L58R, A71K, and G79R.

4. The composition according to claim 2, wherein the protein comprises a heterologous UTR.

5. The composition according to claim 1, wherein the agent is a BAF kinase or an mRNA encoding a BAF kinase.

6. The composition according to claim 5, wherein the BAF kinase comprises an amino acid sequence having a sequence identity of 90% or more to vaccinia virus Ankara strain B1 protein.

7. The composition according to claim 6, wherein the BAF kinase comprises an amino acid sequence having a sequence identity of 90% or more to human VRK1, VRK1-X1, human VRK2A, human VRK2B, or human VRK3.

8. The composition according to claim 2, wherein the agent is a siRNA that is specific for BAF.

9. The composition according to claim 2, wherein the agent is a small molecule.

10. The composition according to claim 9, wherein the small molecule is selected from the group consisting of obtusilactone B, Mahubanolide, kotomolide B, epilitsenolide D2, and rabeprazole.

11. The composition according to claim 1, wherein the agent comprises a LEM domain or an mRNA encoding a LEM domain.

12. The composition according to claim 11, wherein the LEM domain comprises an animo acid sequence having a sequence identity of 90% or more to a sequence selected from the group consisting of:

a. an emerin LEM domain
MDNYADLSDTELTTLLRRYNIPHGPVVGSTRRLYEKKIFEYETQ;
b. a LAP2-beta LEM domain:
DLDVTELTNEDLLDQLVKYGVNPGPIVGTTRKLYEKKLLKLREQ; and
c. a MAN1 LEM domain:
ASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKKLKKLREE.

13. The composition according to claim 1, wherein the agent does not comprise a LEM domain or a nucleic acid encoding a LEM domain.

14. The composition according to any one of claims 1-13, wherein the cytosolic delivery vehicle comprises a non-viral delivery vehicle.

15. The composition according to claim 14, wherein the non-viral delivery vehicle is a nanoparticle.

16. The composition according to claim 15, wherein the nanoparticle is a lipid nanoparticle (LNP).

17. The composition according to claim 14, wherein the non-viral delivery vehicle is a vesicle.

18. The composition according to claim 17, wherein the vesicle is selected from the group consisting of an exosome, an extracellular vesicle, and a micelle.

19. The composition according to any one of claims 1-18, wherein the composition comprises a single cytosolic delivery vehicle comprising both the DNA and the agent.

20. The composition according to any one of claims 1-18, wherein the composition comprises a first cytosolic delivery vehicle comprising the DNA and a second cytosolic delivery vehicle comprising the agent.

21. The composition according to any one of claims 1-20, wherein the composition further comprises an mRNA encoding a nuclear targeting factor (NTF), and the DNA comprises a DNA nuclear targeting sequence (DTS).

22. The composition according to claim 21, wherein:

a. the NTF is a Gal4 NTF and the DTS is a Gal4 binding sequence; or

b. the NTF is a TetR NTF and the DTS is a TetR binding sequence.

23. The composition according to any one of claims 1-22, wherein the composition comprises a nuclease, an integrase or a transposase.

24. The composition according to claim 23, wherein the composition encodes a nuclease and the composition comprises a guide RNA (gRNA).

25. The composition according to any one of claims 1-24, wherein the DNA is 100 – 15,000 bp.

26. The composition according to any one of claims 1-25, wherein the DNA comprises a coding sequence.

27. The composition according to claim 26, wherein the DNA comprises an expression cassette comprising a promoter operably linked to the coding sequence.

28. The composition according to claims 26 or claim 27, wherein the DNA comprises sequences flanking the coding sequence of claim 26 or the expression cassette of claim 27 that are homologous to genomic sequences of the cell for which the DNA is configured to be used.

29. A method comprising contacting a cell with a DNA and an agent that modulates the activity of the protein Barrier to Autointegration Factor (BAF).

30. The method according to claim 29, wherein the agent is a protein comprising an amino acid sequence having a sequence identity of 90% or more to wild type BAF, or an mRNA encoding the protein comprising an amino acid sequence having a sequence identity of 90% or more to wild type BAF.

31. The method according to claim 30, wherein the protein comprises a mutation relative to wild type human BAF selected from the group consisting of T2D, T3D, S4D, A12T, E28D, G31S, K33R, E36D, R37K,L58R, A71K, and G79R.

32. The method according to claim 30, wherein the protein comprises a heterologous UTR.

33. The method according to claim 29, wherein the agent is a BAF kinase or an mRNA encoding a BAF kinase.

34. The method according to claim 33, wherein the BAF kinase comprises an amino acid sequence having a sequence identity of 90% or more to the vaccinia virus Ankara strain B1 protein.

35. The method according to claim 33, wherein the BAF kinase comprises a sequence having a sequence identity of 90% or more to human VRK1, VRK1-X1, human VRK2A, human VRK2B, or human VRK3.

36. The method according to claim 29, wherein the agent is a siRNA that is specific for BAF.

37. The method according to claim 29, wherein the agent is a small molecule.

38. The method according to claim 37, wherein the small molecule is selected from the group consisting of obtusilactone B, Mahubanolide, kotomolide B, epilitsenolide D2, and rabeprazole.

39. The method according to claim 29, wherein the agent comprises a LEM domain or an mRNA encoding a LEM domain.

40. The method according to claim 39, wherein the LEM domain comprises a sequence having a sequence identity of 90% or more to a sequence selected from the group consisting of:

a. an emerin LEM domain
MDNYADLSDTELTTLLRRYNIPHGPVVGSTRRLYEKKIFEYETQ;
b. a LAP2-beta LEM domain:
DLDVTELTNEDLLDQLVKYGVNPGPIVGTTRKLYEKKLLKLREQ; and
c. a MAN1 LEM domain:
ASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKKLKKLREE.

41. The method according to claim 29, wherein the agent does not comprise a LEM domain or a nucleic acid encoding a LEM domain.

42. The method according to any one of claims 29-41, wherein the cell is simultaneously or sequentially contacted with the DNA and the agent.

43. The method according to any one of claims 29-42, wherein the DNA and the agent are present in a cytosolic delivery vehicle.

44. The method according to claim 43, wherein the DNA and the agent are present in the same cytosolic delivery vehicle.

45. The method according to claim 43, wherein the DNA and the agent are present in different cytosolic delivery vehicles.

46. The method according to any one of claims 43-45, wherein the cytosolic delivery vehicle comprises a nanoparticle.

47. The method according to claim 46, wherein the nanoparticle is a lipid nanoparticle.

48. The method according to any one of claims 44-45, wherein the cytosolic delivery vehicle comprises a vesicle.

49. The method according to claim 48, wherein the vesicle is selected from the group consisting of an exosome, an extracellular vesicle, and a micelle.

50. The method according to any one of claims 29-42, wherein the method comprises electroporating the cell.

51. The method according to any one of claims 29-50, wherein the cell is *in vivo*.

52. The method according to claim 51, wherein the method comprises orally, parenterally, subretinally, intravitreally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal administration, intrapleurally, intravenously, intra-arterially, intraperitoneally, intracranially, subcutaneously, intramuscularly, intranasally, intrathecally and/or intraarticularly delivering the DNA and the agent.

53. The method according to any one of claims 29-52, wherein the cell is *ex vivo*.

54. The method according to any one of claims 29-53, wherein the cell is selected from the group consisting of hepatocytes, stellate cells, T lymphocytes, B lymphocytes, NK cells, hematopoietic stem cells, skeletal muscle cells, cardiomyocytes, neurons, astrocytes, oligodendrocytes, dendritic cells and skin cells.