



(51) International Patent Classification:

CI2N 15/86 (2006.01) C07K 14/475 (2006.01)

CI2N 15/113 (2010.01) A61K 48/00 (2006.01)

CI2N 15/65 (2006.01) A61P 27/16 (2006.01)

(21) International Application Number:

PCT/US2019/060324

(22) International Filing Date:

07 November 2019 (07.11.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/756,910 07 November 2018 (07.11.2018) US

62/888,105 16 August 2019 (16.08.2019) US

(71) Applicant: AKOUOS, INC. [US/US]; 7 Stetson Street, Brookline, Massachusetts 02446 (US).

(72) Inventors: SIMONS, Emmanuel J.; 7 Stetson Street, Brookline, Massachusetts 02446 (US). NG, Robert; 47 Lincoln Rd, Unit 1, Newton, Massachusetts 02130 (US). LENZ, Danielle; 47 Lincoln Rd, Unit 1, Brookline, Massachusetts 02446 (US). VALERO, Michelle; 650 Ocean Ave., Unit #612, Revere, Massachusetts 02151 (US).

(74) Agent: LYON, Charles E. et al.; Choate, Hall & Stewart LLP, Two International Place, Boston, Massachusetts 02110 (US).

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

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

(54) Title: COMPOSITIONS AND METHODS OF INDUCING DIFFERENTIATION OF A HAIR CELL

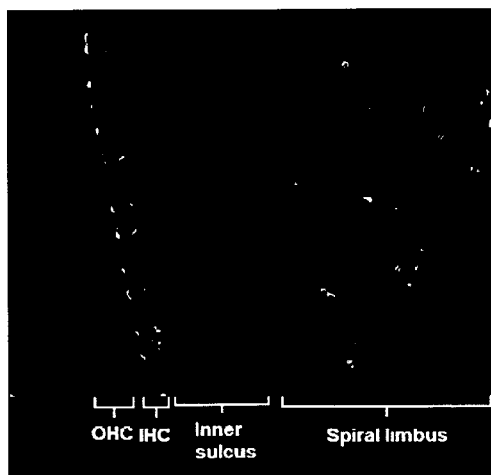


Figure 2B

(57) Abstract: Provided herein are compositions that include one or more adeno-associated virus (AAVs) vectors and methods of inducing differentiation of a hair cell using these vector(s).



Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

**COMPOSITIONS AND METHODS OF INDUCING DIFFERENTIATION
OF A HAIR CELL**

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims priority to U.S. Provisional Patent Application Serial No. 62/756,910, filed November 7, 2018 and U.S. Provisional Patent Application Serial No. 62/888,105, filed August 16, 2019; the entire contents of which are herein incorporated by reference.

TECHNICAL FIELD

10 The present disclosure relates to the fields of molecular biology, and more specifically, to the use of nucleic acids for treating hearing loss in a primate.

BACKGROUND OF THE INVENTION

15 Hearing loss can be conductive (arising from the ear canal or middle ear), sensorineural (arising from the inner ear or auditory nerve), or mixed. Most forms of non-syndromic deafness are associated with permanent hearing loss caused by damage to structures in the inner ear (sensorineural deafness), although some forms may involve changes in the middle ear (conductive hearing loss). The great majority of human sensorineural hearing loss is caused by abnormalities in the hair cells of the organ of Corti in the cochlea (poor hair cell function). The
20 hair cells may be abnormal at birth, or may be damaged during the lifetime of an individual (e.g., as a result of noise trauma or infection).

SUMMARY

25 The present invention is based on the discovery that administration of an AAV vector that includes a nucleic acid encoding a gene, to the inner ear of a primate, can result in the successful expression of a protein encoded by the gene in a supporting cell or hair cell in the inner ear of the primate. In view of this discovery, provided here are AAV vector(s) and methods of using these vectors to induce expression and/or activity of a hair cell differentiation protein in a supporting cell or hair cell in the inner ear of a primate or decreasing the expression and/or activity of a hair
30 cell differentiation suppressing gene in a supporting cell or hair cell in the inner ear of a primate.

Provided herein are compositions that include at least two different nucleic acid vectors, where: each of the at least two different adeno-associated virus (AAV) vectors includes a coding sequence that encodes a different portion of a hair cell differentiation protein, each of the encoded portions being at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes the full-length hair cell differentiation protein; at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons; and when introduced into a primate cell the at least two different vectors undergo concatamerization or homologous recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length hair cell differentiation protein that is expressed in the primate cell.

In some embodiments of any of the compositions described herein, the amino acid sequence of none of the encoded portions overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments of any of the compositions described herein, the amino acid sequence of each of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments of any of the compositions described herein, the overlapping amino acid sequence is between 30 amino acid residues to about 390 amino acid residues in length.

In some embodiments of any of the compositions described herein, the vectors include two different vectors, each of which includes a different segment of an intron, where the intron includes the nucleotide sequence of an intron that is present in a hair cell differentiation genomic DNA, and where the two different segments overlap in sequence by at least 100 nucleotides. In some embodiments of any of the compositions described herein, the two different intron segments overlap in sequence by about 100 nucleotides to about 800 nucleotides.

In some embodiments of any of the compositions described herein, the entire nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 10,000 nucleotides in length. In some embodiments of any of the compositions described herein, the entire nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 5,000 nucleotides in length.

In some embodiments of any of the compositions described herein, the number of different vectors in the composition is two. In some embodiments of any of the compositions described herein, a first of the two different vectors includes a coding sequence that encodes an N-terminal portion of the hair cell differentiation protein. In some embodiments of any of the compositions described herein, the N-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 750 amino acids in length. In some embodiments of any of the compositions described herein, the N-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 320 amino acids in length.

In some embodiments of any of the compositions described herein, the first vector further includes one or both of a promoter and a Kozak sequence. In some embodiments of any of the compositions described herein, the first vector includes a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.

In some embodiments of any of the compositions described herein, the second of the two different vectors includes a coding sequence that encodes a C-terminal portion of the hair cell differentiation protein. In some embodiments of any of the compositions described herein, the C-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 750 amino acids in length. In some embodiments of any of the compositions described herein, the C-terminal portion of the hair cell differentiation portion is between about 30 amino acids to about 320 amino acids in length.

In some embodiments of any of the compositions described herein, the second vector further includes a poly(dA) sequence. In some embodiments of any of the compositions described herein, the second vector further includes a destabilizing sequence. In some embodiments of any of the compositions described herein, the second vector further includes a FKB12 destabilizing sequence.

Also provided herein are compositions that include two different nucleic acid vectors, where: a first nucleic acid vector of the two different nucleic acid vectors includes a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, and a splicing donor signal sequence positioned at the 3' end of the first coding sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing

acceptor signal sequence, and a polyadenylation sequence at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the amino acid sequences of the encoded portions do not overlap, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell, to produce RNA transcripts, splicing occurs between the splicing donor signal sequence on one transcript and the splicing acceptor signal sequence on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons.

Also provided herein are compositions that include: a first nucleic acid vector including a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence; and a second nucleic acid vector, different from the first nucleic acid vector, including a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the respective amino acid sequences of the encoded portions do not overlap with each other, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell to produce RNA transcripts, splicing occurs between the splicing donor signal on one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the neighboring exons. In

some embodiments of any of the compositions described herein, the first or second detectable marker gene is alkaline phosphatase. In some embodiments of any of the compositions described herein, the first and second detectable marker genes are the same.

Also provided herein are compositions that include: a first nucleic acid vector including a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' to the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splicing donor signal sequence; and a second nucleic acid vector, different from the first nucleic acid vector, including a second F1 phage recombinogenic region, a splicing acceptor signal sequence positioned 3' of the second F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the respective amino acid sequences of the encoded portions do not overlap with each other, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell to produce RNA transcripts, splicing occurs between the splicing donor signal on one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons.

Also provided herein are compositions that include a single adeno-associated virus (AAV) vector, where the single AAV vector includes a nucleic acid sequence that encodes a hair cell differentiation protein; and when introduced into a mammalian cell (e.g., primate cell (e.g., a hair cell or a supporting cell of the inner ear), a nucleic acid encoding the hair cell differentiation protein is generated at the locus of the hair cell differentiation gene and the primate cell expresses the hair cell differentiation protein.

In some embodiments of any of the compositions described herein, the hair cell differentiation gene is selected from the group of: atonal bHLH transcription factor 1 (ATOH1),

POU Class 4 Homeobox 3 (POU4F3), catenin beta 1 (CTNNB1), Noggin (NOG), growth factor independent 1 transcriptional repressor (GFI-1), neurotrophin 3 (NTF3), and brain-derived neurotrophic factor (BDNF).

Also provided herein are compositions including two different nucleic acid vectors, wherein a first nucleic acid vector includes a first nucleic acid sequence that encodes a first hair cell differentiation protein (e.g., any of the hair cell differentiation proteins described herein); and a second nucleic acid vector includes a second nucleic acid sequence that encodes a second hair cell differentiation protein (e.g., any of the hair cell differentiation proteins described herein), and when introduced into a primate cell, the first nucleic acid and the second nucleic acid encoding the first hair cell differentiation protein and the second hair cell differentiation protein are generated at the locus of the hair cell differentiation gene and the primate cell expresses the first hair cell differentiation protein and the second hair cell differentiation protein.

In some embodiments, the first and the second hair cell differentiation proteins are selected from the group consisting of: atonal bHLH transcription factor 1 (ATOH1), POU Class 4 Homeobox 3 (POU4F3), catenin beta 1 (CTNNB1), Noggin (NOG), growth factor independent 1 transcriptional repressor (GFI-1), neurotrophin 3 (NTF3) and brain-derived neurotrophic factor (BDNF).

In some embodiments of any of the compositions described herein, the second nucleic acid vector further includes a destabilizing sequence.

In some embodiments, the second nucleic acid vector further includes a FKB12 destabilizing sequence.

Provided herein are compositions that include at least one adeno-associated virus (AAV) vector that encodes an inhibitory nucleic acid that decreases the expression of a hair cell differentiation-suppressing protein in a primate cell.

In some embodiments of any of the compositions described herein, the inhibitory nucleic acid is a short interfering RNA (siRNA), a short hairpin RNA (shRNA), an antisense oligonucleotide, or a ribozyme.

In some embodiments of any of the compositions described herein, the hair cell differentiation-suppressing gene is HES1, HES5, sex determining region Y-box 2 (SOX2), and p27kip (CDKN1B). In some embodiments of any of the compositions described herein, the composition further includes a pharmaceutically acceptable excipient. Also provided herein are

kits including any of the compositions described herein. In some embodiments of any of the kits described herein, the kit further includes a pre-loaded syringe containing the composition.

Also provided herein are methods of promoting differentiation of a supporting cell of an inner ear of a primate into a hair cell that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering promotes differentiation of the supporting cell of the inner ear of the primate into a hair cell.

Also provided herein are methods of increasing the expression level of a hair cell differentiation protein in a supporting cell or hair cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in an increase in the expression level of the hair cell differentiation protein in the supporting cell or hair cell of the inner ear of the primate. In some embodiments of any of the methods described herein, the hair cell differentiation protein is selected from the group of: Atoh1, Pou4f3, β -Catenin, Noggin, GFI-1, NTF3, and BDNF. In some embodiments of the methods described herein, the primate has previously been determined to have a defective hair cell differentiation gene.

Also provided herein are methods of decreasing the expression level of a hair cell differentiation-suppressing protein in a supporting cell or hair cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in a decrease in the expression level of the hair cell differentiation-suppressing protein in the supporting cell or hair cell of the inner ear of the primate.

Also provided herein are methods of increasing the number of functional hair cells in a primate in need thereof that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of improving hearing in a primate in need thereof that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

In some embodiments of any of the methods described herein, the method further includes prior to the administering step, determining that the primate has a defective hair cell differentiation gene.

Also provided herein are methods of repairing a hair cell toxicity-inducing mutation in an endogenous hair cell differentiation gene locus in a supporting cell or hair cell of an inner ear of a primate, that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in repair of the hair cell toxicity-inducing mutation in the endogenous hair cell differentiation gene locus in the supporting cell or hair cell of the inner ear of the primate.

Also provided herein are methods of decreasing the risk of hearing loss due to hair cell loss or dysfunction in a primate in need thereof that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

In some embodiments of any of the methods described herein, the primate has been previously identified as having a defective hair cell differentiation gene.

The term “a” and “an” refers to one or to more than one (i.e., at least one) of the grammatical object of the article.

The term “conservative mutation” refers to a mutation that does not change the amino acid encoded at the site of the mutation (due to codon degeneracy).

Modifications can be introduced into a nucleotide sequence by standard techniques known in the art, such as site-directed mutagenesis and PCR-mediated mutagenesis.

Conservative amino acid substitutions are ones in which the amino acid residue in a protein is replaced with an amino acid residue having a chemically-similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, and histidine), acidic side chains (e.g., aspartic acid and glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, and tryptophan), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, and methionine), beta-branched side chains (e.g., threonine, valine, and isoleucine), and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, and histidine).

Unless otherwise specified, a “nucleotide sequence encoding an amino acid sequence” includes all nucleotide sequences that are degenerate versions of each other and thus encode the same amino acid sequence.

The term “endogenous” refers to any material originating from within an organism, cell, or tissue.

The term “exogenous” refers to any material introduced from or originating from outside an organism, cell, or tissue that is not produced or does not originate from the same organism, cell, or tissue in which it is being introduced.

The term “isolated” means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term “transfected,” “transformed,” or “transduced” refers to a process by which exogenous nucleic acid is transferred or introduced into a cell. A “transfected,” “transformed,” or “transduced” primate cell is one that has been transfected, transformed, or transduced with exogenous nucleic acid.

The term “expression” refers to the transcription and/or translation of a particular nucleotide sequence encoding a protein.

The term “transient expression” refers to the expression of a non-integrated coding sequence for a short period of time (e.g., hours or days). The coding sequence that is transiently expressed in a cell (e.g., a primate cell) is lost upon multiple rounds of cell division.

The term “primate” is intended to include any primate (e.g., a human, a non-human primate (e.g., simian (e.g., a monkey (e.g., a marmoset, a baboon, a macaque), or an ape (e.g., a gorilla, a gibbon, an orangutan, or a chimpanzee). In some embodiments, the primate has or is at risk of having hearing loss. In some embodiments, the primate has been previously identified as having a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene. In some embodiments, the primate has been previously identified as having a mutation in a hair cell differentiation gene. In some embodiments, the primate has been previously identified as having a mutation in a hair cell differentiation-suppressing gene. In some embodiments, the primate has been identified as having a mutation in hair cell differentiation gene and/or a hair cell differentiation-suppressing gene and has been diagnosed with hearing loss. In some embodiments, the primate has been identified as having hearing loss.

A treatment is “therapeutically effective” when it results in a reduction in one or more of the number, severity, and frequency of one or more symptoms of a disease state (e.g., non-syndromic sensorineural hearing loss or syndromic sensorineural hearing loss) in a primate. In

some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active hair cell differentiation protein (e.g., a wildtype, full-length hair cell differentiation protein, or an active variant of a hair cell differentiation protein,) (e.g., as compared to the expression level prior to treatment with the composition). In some

5 embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active hair cell differentiation protein (e.g., a wildtype, full-length hair cell differentiation protein or active variant) in a target cell (e.g., a supporting cell of the inner ear or a hair cell (e.g., an outer hair cell or an inner hair cell) of the inner ear). In some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level
10 of an active hair cell differentiation protein (e.g., a wildtype, full-length hair cell differentiation protein or active variant), and/or an increase in one or more activities of a hair cell differentiation protein in a target cell (e.g., as compared to a reference level, such as the level(s) in a primate cell prior to treatment, the level(s) in a primate cell having a mutation in a hair cell differentiation gene, or the level(s) in a primate cell or a population of primate cells from a
15 subject having non-syndromic sensorineural hearing loss, or the level(s) in a primate cell or a population of primate cells from a subject having syndromic sensorineural hearing loss).

The term “nucleic acid” or “polynucleotide” refers to deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), or a combination thereof, in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of
20 natural nucleotides that have similar binding properties as the reference nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses complementary sequences as well as the sequence explicitly indicated. In some embodiments of any of the nucleic acids described herein, the nucleic acid is DNA. In some embodiments of any of the nucleic acids described herein, the nucleic acid is RNA.

25 The term “hair cell toxicity-inducing mutation” refers to a mutation in a hair cell differentiation gene that encodes a protein that when expressed (e.g., by a supporting cell or a hair cell) induces toxicity in a hair cell (e.g., in a primate).

The term “active hair cell differentiation protein” means a protein encoded by DNA that, if substituted for both wildtype alleles encoding full-length hair cell differentiation protein in
30 supporting cells of the inner ear of what is otherwise a wildtype primate, and if expressed in the supporting cells of that primate, results in that primate’s having a level of hearing approximating

the normal level of hearing of a similar primate that is entirely wildtype. Non-limiting examples of active hair cell differentiation proteins are full-length hair cell differentiation proteins (e.g., any of the full-length hair cell differentiation proteins described herein).

The term “inhibitory nucleic acid” refers to a nucleic acid sequence that hybridizes specifically to a target gene or a target mRNA (e.g., a hair cell differentiation-suppressing gene or a hair cell differentiation-suppressing mRNA) and thereby inhibits the expression and/or activity of the target gene or the target mRNA (e.g., a hair cell differentiation-suppressing gene or a hair cell differentiation-suppressing mRNA). In some embodiments, the inhibitory nucleic acid is a short interfering RNA (siRNA), a short hairpin RNA (shRNA), an antisense oligonucleotide, or a ribozyme. In some embodiments, the inhibitory nucleic acid is between about 10 nucleotides to about 30 nucleotides in length (e.g., about 10 nucleotides to about 28 nucleotides, about 10 nucleotides to about 26 nucleotides, about 10 nucleotides to about 24 nucleotides, about 10 nucleotides to about 22 nucleotides, about 10 nucleotides to about 20 nucleotides, about 10 nucleotides to about 18 nucleotides, about 10 nucleotides to about 16 nucleotides, about 10 nucleotides to about 14 nucleotides, about 10 nucleotides to about 12 nucleotides, about 12 nucleotides to about 30 nucleotides, about 12 nucleotides to about 28 nucleotides, about 12 nucleotides to about 26 nucleotides, about 12 nucleotides to about 24 nucleotides, about 12 nucleotides to about 22 nucleotides, about 12 nucleotides to about 20 nucleotides, about 12 nucleotides to about 18 nucleotides, about 12 nucleotides to about 16 nucleotides, about 12 nucleotides to about 14 nucleotides, about 16 nucleotides to about 30 nucleotides, about 16 nucleotides to about 28 nucleotides, about 16 nucleotides to about 26 nucleotides, about 16 nucleotides to about 24 nucleotides, about 16 nucleotides to about 22 nucleotides, about 16 nucleotides to about 20 nucleotides, about 16 nucleotides to about 18 nucleotides, about 18 nucleotides to about 30 nucleotides, about 18 nucleotides to about 28 nucleotides, about 18 nucleotides to about 26 nucleotides, about 18 nucleotides to about 24 nucleotides, about 18 nucleotides to about 22 nucleotides, about 18 nucleotides to about 20 nucleotides, about 20 nucleotides to about 30 nucleotides, about 20 nucleotides to about 28 nucleotides, about 20 nucleotides to about 26 nucleotides, about 20 nucleotides to about 24 nucleotides, about 20 nucleotides to about 22 nucleotides, about 22 nucleotides to about 30 nucleotides, about 22 nucleotides to about 28 nucleotides, about 22 nucleotides to about 26 nucleotides, about 22 nucleotides to about 24 nucleotides, about 24 nucleotides to about 30

nucleotides, about 24 nucleotides to about 28 nucleotides, about 24 nucleotides to about 26 nucleotides, about 26 nucleotides to about 30 nucleotides, about 26 nucleotides to about 28 nucleotides, about 28 nucleotides to about 30 nucleotides, or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 nucleotides).

Unless otherwise defined, 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. Methods and materials are described herein for use in the present invention; other suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1A is a representative image of Myo7a/Iba-1 immunofluorescent staining of cochlear tissue of a cynomolgus macaque (non-human primate) following administration of a single Anc80-GFP AAV vector directly into the inner ear through the round window.

Figure 1B is a representative image of Anc80-GFP immunofluorescent staining of the same cochlear tissue of the cynomolgus macaque as in Figure 1A.

Figure 1C is a representative image of a merged immunofluorescent staining of Myo7a/Iba-1 and Anc80-GFP of the same cochlear tissue of the cynomolgus macaque as in Figure 1A.

Figure 2A is a representative image of Anc80-GFP immunofluorescent staining of a NHP cochlear tissue showing the stria vascularis, the spiral ligament and the lateral wall.

Figure 2B is a representative image of Anc80-GFP immunofluorescent staining of the same NHP cochlear tissue as in Figure 2A showing the spiral limbus, the inner sulcus, inner hair cells (IHC) and outer hair cells (OHC).

Figure 3 is a simplified schematic diagram showing the proteins that play a role during the development of supporting cells and hair cells in the cochlea.

Figure 4A is an exemplary nucleic acid vector (SEQ ID NO: 66), that includes an ITR sequence (SEQ ID NO: 51), a CMV enhancer sequence (SEQ ID NO: 52), a CMV promoter

sequence (SEQ ID NO: 53), a human ATOH1 gene sequence (SEQ ID NO: 67), a 3xFlag sequence (SEQ ID NO: 62), a T2A sequence (SEQ ID NO: 63), a SV40-NLS sequence (SEQ ID NO: 54), a mScarlet gene sequence (SEQ ID NO: 55), a destabilizing domain (DD) sequence (SEQ ID NO: 59), a bGHpA sequence (SEQ ID NO: 56), and an ITR sequence (SEQ ID NO: 57).

Figure 4B is an exemplary nucleic acid vector (SEQ ID NO: 64), that includes an ITR sequence (SEQ ID NO: 51), a CMV enhancer sequence (SEQ ID NO: 52), a CMV promoter sequence (SEQ ID NO: 53), a human GFII1 gene sequence (SEQ ID NO: 65), a 3xFlag sequence (SEQ ID NO: 62), a T2A sequence (SEQ ID NO: 63), a SV40-NLS sequence (SEQ ID NO: 54), a mScarlet sequence (SEQ ID NO: 55), a destabilizing domain (DD) sequence (SEQ ID NO: 59), a bGHpA sequence (SEQ ID NO: 56), and an ITR sequence (SEQ ID NO: 57).

Figure 4C is an exemplary nucleic acid vector (SEQ ID NO: 60), that includes an ITR sequence (SEQ ID NO: 51), a CMV enhancer sequence (SEQ ID NO: 52), a CMV promoter sequence (SEQ ID NO: 53), a human POU4F3 gene sequence (SEQ ID NO: 61), a 3xFlag sequence (SEQ ID NO: 62), a T2A sequence (SEQ ID NO: 63), a SV40-NLS sequence (SEQ ID NO: 54), a mScarlet sequence (SEQ ID NO: 55), a destabilizing domain (DD) sequence (SEQ ID NO: 59), a bGHpA sequence (SEQ ID NO: 56), and an ITR sequence (SEQ ID NO: 57).

Figure 4D is an exemplary nucleic acid vector (SEQ ID NO: 68), that includes an ITR sequence (SEQ ID NO: 51), a CMV enhancer sequence (SEQ ID NO: 52), a CMV promoter sequence (SEQ ID NO: 53), a luciferase (Fluc) gene sequence (SEQ ID NO: 69), a T2A sequence (SEQ ID NO: 63), an mScarlet gene sequence (SEQ ID NO: 55), a SV40pA sequence (SEQ ID NO: 70), a U6 sequence (SEQ ID NO: 71), a short hairpin RNA (shRNA) sequence (SEQ ID NO: 72), and an ITR sequence (SEQ ID NO: 57).

Figure 5A is bar graph showing the relative quantification of Hes-1 RNA in HEK293FT cells transfected with combinations of dual and triple shRNA constructs (S3 (GAAAGTCATCAAAGCCTAT; SEQ ID NO: 73), S5 (ACTGCATGACCCAGATCAA; SEQ ID NO: 74), Kop (ACTGCATGACCCAGATCAA; SEQ ID NO: 75), S3 plus S5, S3 plus Kop, and S5 plus Kop) as determined by real time quantitative polymerase chain reaction (RTqPCR).

Figure 5B is bar graph showing the relative quantification of Hes-1 protein in HEK293FT cells transfected with combinations of dual and triple shRNA constructs (S3, S5, Kop, S3 plus S5, S3 plus Kop, and S5 plus Kop) as determined by Western blotting.

Figure 6A is a bar graph showing the relative quantification of ATOH1, POU4F3, and GFI1 (APG) RN in HEK293FT cells transfected with the individual plasmids of Figures 4A-C.

Figure 6B is an image of a Western blot showing the relative quantification of ATOH1, POU4F3 and GFI1 protein expression in HEK293FT cells transfected with the individual
5 plasmids of Figures 4A-C.

Figure 7A is an exemplary nucleic acid vector (SEQ ID NO: 76), that includes an ITR sequence (SEQ ID NO: 51), a CMV promoter sequence (SEQ ID NO: 53), a mScarlet sequence (SEQ ID NO: 55), a bGHpA sequence (SEQ ID NO: 56) and an ITR sequence (SEQ ID NO: 57).

Figure 7B is an exemplary nucleic acid vector (SEQ ID NO: 77), that includes an ITR
10 sequence (SEQ ID NO: 51), a CMV promoter sequence (SEQ ID NO: 53), a mScarlet sequence (SEQ ID NO: 55), a destabilizing domain (DD) sequence (SEQ ID NO: 59), a bGHpA sequence (SEQ ID NO: 56) and an ITR sequence (SEQ ID NO: 57).

Figure 8A is a dose response curve showing the functionality and reversibility of the destabilizing domain (DD) using fluorescence microscopy. Serial dilutions of TMP (0.1 μ M, 1
15 μ M, 10 μ M, 20 μ M and 100 μ M) were tested in the mScarlet and mScarlet-DD transfected HEK293FT cells.

Figure 8B is a graph showing the functionality and reversibility of the destabilizing domain (DD) by flow cytometry (Attune flow cytometer).

Figure 9A is an image showing mScarlet positive cells in a P1-P3 mouse cochlea explant
20 transfected with AAVanc80 vector at various MOIs. 10 μ M TMP was added at a later time point.

Figure 9B is an image showing mScarlet positive HEK293FT cells transfected with AAVanc80 vector at various MOIs. 10 μ M TMP was added at a later time point.

Figure 10 is an image showing mScarlet positive hair cells and supporting cells in a
25 cochlear explants infected with AAVanc80 with and without 10uM TMP that was added at a later time point.

Figure 11A is an exemplary nucleic acid vector (SEQ ID NO: 83), that includes an ITR sequence (SEQ ID NO: 51), a U6 sequence (SEQ ID NO: 84), a short hairpin HES1 RNA (shHES1) sequence (SEQ ID NO: 85), a CMV enhancer sequence (SEQ ID NO: 52), a CMV
30 promoter sequence (SEQ ID NO: 53), a 3xFlag sequence (SEQ ID NO: 86), a human ATOH1 gene sequence (SEQ ID NO: 87), a destabilizing domain (DD) sequence (SEQ ID NO: 88), a

T2A sequence (SEQ ID NO: 89), a human POU4F3 gene sequence (SEQ ID NO: 61), a bGHpA sequence (SEQ ID NO: 90), a U6 sequence (SEQ ID NO: 91), a short hairpin HES1 RNA (shHES1-2) sequence (SEQ ID NO: 92) and an ITR sequence (SEQ ID NO: 57).

Figure 11B is an exemplary nucleic acid vector (SEQ ID NO: 93), that includes an ITR sequence (SEQ ID NO: 51), a U6 sequence (SEQ ID NO: 84), a short hairpin HES1 RNA (shHES-1) sequence (SEQ ID NO: 85), an ATOH1 enhancer-promoter sequence (SEQ ID NO: 94), a 3xFlag sequence (SEQ ID NO: 86), a human ATOH1 gene sequence (SEQ ID NO: 67), a T2A sequence (SEQ ID NO: 63), a human POU4F3 gene sequence (SEQ ID NO: 95), a bGHpA sequence (SEQ ID NO: 90), a U6 sequence (SEQ ID NO: 84), a short hairpin HES1 RNA (shHES1-2) sequence (SEQ ID NO: 92) and an ITR sequence (SEQ ID NO: 57).

Figure 12A is a bar graph showing the relative quantification of ATOH1, POU4F3, and HES1 in HEK293FT cells transfected with the combined plasmids of Figures 11A-B.

Figure 12B is an image of a Western blot showing the relative quantification of 3xFlag-ATOH1 and HES1 protein expression in HEK293FT cells transfected with the combined plasmids of Figures 11A-B.

DETAILED DESCRIPTION

Provided herein are compositions including at least two different nucleic acid vectors, where: each of the at least two different adeno-associated virus (AAV) vectors comprises a coding sequence that encodes a different portion of a hair cell differentiation protein, each of the encoded portions being at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes the full-length hair cell differentiation protein; at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons; and when introduced into a primate cell (e.g., a hair cell or a supporting cell of the inner ear) the at least two different vectors undergo concatamerization or homologous recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length hair cell differentiation protein that is expressed in the primate cell.

Also provided herein are compositions including two different nucleic acid vectors, where: a first nucleic acid vector of the two different nucleic acid vectors includes a promoter, a

first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, and a splicing donor signal sequence positioned at the 3' end of the first coding sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the amino acid sequences of the encoded portions do not overlap, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell (e.g., a hair cell or a supporting cell of the inner ear), to produce RNA transcripts, splicing occurs between the splicing donor signal sequence on one transcript and the splicing acceptor signal sequence on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

Also provided herein are compositions including: a first nucleic acid vector including a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence; and a second nucleic acid vector, different from the first nucleic acid vector, including a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the respective amino acid sequences of the encoded portions do not overlap with each other, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell (e.g., a hair cell or a supporting cell of the inner ear) to produce RNA transcripts, splicing occurs between the splicing donor signal on one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

Also provided herein are compositions including: a first nucleic acid vector including a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' to the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splicing donor signal sequence; and a second nucleic acid vector, different from the first nucleic acid vector, including a second F1 phage recombinogenic region, a splicing acceptor signal sequence positioned 3' of the second F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the respective amino acid sequences of the encoded portions do not overlap with each other, where no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell (e.g., a hair cell or a supporting cell of the inner ear) to produce RNA transcripts, splicing occurs between the splicing donor signal on one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

Also provided herein are compositions including a single adeno-associated virus (AAV) vector, where the single AAV vector comprises a nucleic acid sequence that encodes a hair cell differentiation protein; and when introduced into a primate cell (e.g., a hair cell or a supporting cell of the inner ear), a nucleic acid encoding the hair cell differentiation protein is generated at the locus of the hair cell differentiation gene and the primate cell expresses the hair cell differentiation protein. Also provided herein are compositions including a single adeno-associated virus (AAV) vector that encodes an inhibitory nucleic acid that decreases the expression of a hair cell differentiation-suppressing protein in a primate cell (e.g., a hair cell or a supporting cell of the inner ear).

Also provided herein are methods of promoting differentiation of a supporting cell of an inner ear of a primate into a hair cell that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering promotes differentiation of the supporting cell of the inner ear of the primate into a hair cell. Also provided herein are methods of increasing the expression level of a hair cell

differentiation protein in a supporting cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in an increase in the expression level of the hair cell differentiation protein in the supporting cell of the inner ear of the primate.

5 Also provided herein are methods of decreasing the expression level of a hair cell differentiation-suppressing protein in a supporting cell or a hair cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in a decrease in the expression level of the hair cell differentiation-suppressing protein in the supporting cell or the
10 hair cell of the inner ear of the primate.

Also provided herein are methods of increasing the number of functional hair cells in a primate in need thereof that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein. Also provided herein are methods of improving hearing in a primate in need thereof, the method comprising
15 administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of repairing a hair cell toxicity-inducing mutation in an endogenous hair cell differentiation gene locus in a supporting cell or a hair cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective
20 amount of any of the compositions described herein, where the administering results in repair of the hair cell toxicity-inducing mutation in the endogenous hair cell differentiation gene locus in the supporting cell or the hair cell of the inner ear of the primate.

Also provided herein are methods of decreasing the risk of hearing loss due to hair cell loss or dysfunction in a primate in need thereof that include: administering to the inner ear of the
25 primate a therapeutically effective amount of any of the compositions described herein. Also provided herein are methods that include introducing into a cochlea of a mammal a therapeutically effective amount of any of the compositions described herein.

Also provided are kits that include any of the compositions described herein.

Additional non-limiting aspects of the compositions, kits, and methods are described
30 herein and can be used in any combination without limitation.

Hair Cell Differentiation Genes

The term “hair cell differentiation gene” refers to a gene encoding a protein (e.g., a transcription factor) that positively contributes, either directly or indirectly, to hair cell differentiation and viability in a primate (e.g., a human). Non-limiting examples of hair cell differentiation genes include: ATOH1, POU4F3, CTNNB1, NOG, GFI-1, NTF3, and BDNF.

The term “mutation in a hair cell differentiation gene” refers to a modification in a wildtype hair cell differentiation gene that results in the production of a hair cell differentiation protein having one or more of: a deletion in one or more amino acids, one or more amino acid substitutions, and one or more amino acid insertions as compared to the wildtype hair cell differentiation protein, and/or results in a decrease in the expressed level of the encoded hair cell differentiation protein in a primate cell as compared to the expressed level of the encoded hair cell differentiation protein in a primate cell not having a mutation. In some embodiments, a mutation can result in the production of a hair cell differentiation protein having a deletion in 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 embodiments, the mutation can result in a frameshift in the hair cell differentiation gene. The term “frameshift” is known in the art to encompass any mutation in a coding sequence that results in a shift in the reading frame of the coding sequence. In some embodiments, a frameshift can result in a nonfunctional protein. In some embodiments, a point mutation can be a nonsense mutation (i.e., results in a premature stop codon in an exon of the gene). A nonsense mutation can result in the production of a truncated protein (as compared to a corresponding wildtype protein) that may or may not be functional. In some embodiments, the mutation can result in the loss (or a decrease in the level) of expression of hair cell differentiation mRNA or hair cell differentiation protein, or both the mRNA and protein. In some embodiments, the mutation can result in the production of an altered hair cell differentiation protein having a loss or decrease in one or more biological activities (functions) as compared to a wildtype hair cell differentiation protein.

In some embodiments, the mutation is an insertion of one or more nucleotides into a hair cell differentiation gene. In some embodiments, the mutation is in a regulatory sequence of the hair cell differentiation gene, i.e., a portion of the gene that is not coding sequence. In some embodiments, a mutation in a regulatory sequence may be in a promoter or enhancer region and prevent or reduce the proper transcription of the hair cell differentiation gene.

For example, an active hair cell differentiation protein can include a sequence of a wildtype, full-length hair cell differentiation protein (e.g., a wildtype, human, full-length hair cell differentiation protein) including 1 amino acid substitution to about 160 amino acid substitutions, 1 amino acid substitution to about 155 amino acid substitutions, 1 amino acid substitution to about 150 amino acid substitutions, 1 amino acid substitution to about 145 amino acid substitutions, 1 amino acid substitution to about 140 amino acid substitutions, 1 amino acid substitution to about 135 amino acid substitutions, 1 amino acid substitution to about 130 amino acid substitutions, 1 amino acid substitution to about 125 amino acid substitutions, 1 amino acid substitution to about 120 amino acid substitutions, 1 amino acid substitution to about 115 amino acid substitutions, 1 amino acid substitution to about 110 amino acid substitutions, 1 amino acid substitution to about 105 amino acid substitutions, 1 amino acid substitution to about 100 amino acid substitutions, 1 amino acid substitution to about 95 amino acid substitutions, 1 amino acid substitution to about 90 amino acid substitutions, 1 amino acid substitution to about 85 amino acid substitutions, 1 amino acid substitution to about 80 amino acid substitutions, 1 amino acid substitution to about 75 amino acid substitutions, 1 amino acid substitution to about 70 amino acid substitutions, 1 amino acid substitution to about 65 amino acid substitutions, 1 amino acid substitution to about 60 amino acid substitutions, 1 amino acid substitution to about 55 amino acid substitutions, 1 amino acid substitution to about 50 amino acid substitutions, 1 amino acid substitution to about 45 amino acid substitutions, 1 amino acid substitution to about 40 amino acid substitutions, 1 amino acid substitution to about 35 amino acid substitutions, 1 amino acid substitution to about 30 amino acid substitutions, 1 amino acid substitution to about 25 amino acid substitutions, 1 amino acid substitution to about 20 amino acid substitutions, 1 amino acid substitution to about 15 amino acid substitutions, 1 amino acid substitution to about 10 amino acid substitutions, 1 amino acid substitution to about 9 amino acid substitutions, 1 amino acid substitution to about 8 amino acid substitutions, 1 amino acid substitution to about 7 amino acid substitutions, 1 amino acid substitution to about 6 amino acid substitutions, 1 amino acid substitution to about 5 amino acid substitutions, 1 amino acid substitution to about 4 amino acid substitutions, 1 amino acid substitution to about 3 amino acid substitutions, between about 2 amino acid substitutions to about 160 amino acid substitutions, about 2 amino acid substitutions to about 155 amino acid substitutions, about 2 amino acid substitutions to about 150 amino acid substitutions, about 2 amino acid substitutions to about 145 amino acid substitutions, about 2

5

10

15

20

25

30

5
10
15
20
25
30

23

24

5
10
15
20
25
30

[illegible]

5
10
15
20
25
30

5

10

15

20

25

30

30

5
10
15
20
25
30

5

10

15

20

25

30

[illegible]

5
10
15
20
25
30

34

substitutions to about 135 amino acid substitutions, between about 140 amino acid substitutions to about 160 amino acid substitutions, about 140 amino acid substitutions to about 155 amino acid substitutions, about 140 amino acid substitutions to about 150 amino acid substitutions, about 140 amino acid substitutions to about 145 amino acid substitutions, between about 150 amino acid substitutions to about 160 amino acid substitutions, or about 150 amino acid substitutions to about 155 amino acid substitutions. One skilled in the art would appreciate that amino acids that are not conserved between wildtype hair cell differentiation proteins from different species can be mutated without losing activity, while those amino acids that are conserved between wildtype hair cell differentiation proteins from different species should not be mutated as they are more likely (than amino acids that are not conserved between different species) to be involved in activity.

An active hair cell differentiation protein can include, e.g., a sequence of a wildtype, full-length hair cell differentiation protein (e.g., a wildtype, human, full-length hair cell differentiation protein) that has 1 amino acid to about 50 amino acids, 1 amino acid to about 45 amino acids, 1 amino acid to about 40 amino acids, 1 amino acid to about 35 amino acids, 1 amino acid to about 30 amino acids, 1 amino acid to about 25 amino acids, 1 amino acid to about 20 amino acids, 1 amino acid to about 15 amino acids, 1 amino acid to about 10 amino acids, 1 amino acid to about 9 amino acids, 1 amino acid to about 8 amino acids, 1 amino acid to about 7 amino acids, 1 amino acid to about 6 amino acids, 1 amino acid to about 5 amino acids, 1 amino acid to about 4 amino acids, 1 amino acid to about 3 amino acids, about 2 amino acids to about 50 amino acids, about 2 amino acids to about 45 amino acids, about 2 amino acids to about 40 amino acids, about 2 amino acids to about 35 amino acids, about 2 amino acids to about 30 amino acids, about 2 amino acids to about 25 amino acids, about 2 amino acids to about 20 amino acids, about 2 amino acids to about 15 amino acids, about 2 amino acids to about 10 amino acids, about 2 amino acids to about 9 amino acids, about 2 amino acids to about 8 amino acids, about 2 amino acids to about 7 amino acids, about 2 amino acids to about 6 amino acids, about 2 amino acids to about 5 amino acids, about 2 amino acids to about 4 amino acids, about 3 amino acids to about 50 amino acids, about 3 amino acids to about 45 amino acids, about 3 amino acids to about 40 amino acids, about 3 amino acids to about 35 amino acids, about 3 amino acids to about 30 amino acids, about 3 amino acids to about 25 amino acids, about 3 amino acids to about 20 amino acids, about 3 amino acids to about 15 amino acids, about 3

5
10
15
20
25
30

amino acids to about 25 amino acids, about 10 amino acids to about 20 amino acids, about 10 amino acids to about 15 amino acids, about 15 amino acids to about 50 amino acids, about 15 amino acids to about 45 amino acids, about 15 amino acids to about 40 amino acids, about 15 amino acids to about 35 amino acids, about 15 amino acids to about 30 amino acids, about 15 amino acids to about 25 amino acids, about 15 amino acids to about 20 amino acids, about 20 amino acids to about 50 amino acids, about 20 amino acids to about 45 amino acids, about 20 amino acids to about 40 amino acids, about 20 amino acids to about 35 amino acids, about 20 amino acids to about 30 amino acids, about 20 amino acids to about 25 amino acids, about 25 amino acids to about 50 amino acids, about 25 amino acids to about 45 amino acids, about 25 amino acids to about 40 amino acids, about 25 amino acids to about 35 amino acids, about 25 amino acids to about 30 amino acids, about 30 amino acids to about 50 amino acids, about 30 amino acids to about 45 amino acids, about 30 amino acids to about 40 amino acids, about 30 amino acids to about 35 amino acids, about 35 amino acids to about 50 amino acids, about 35 amino acids to about 45 amino acids, about 35 amino acids to about 40 amino acids, about 40 amino acids to about 50 amino acids, about 40 amino acids to about 45 amino acids, about 45 amino acids to about 50 amino acids, deleted. In some embodiments where two or more amino acids are deleted from the sequence of a wildtype, full-length hair cell differentiation protein, the two or more deleted amino acids can be contiguous in the sequence of the wildtype, full-length protein. In other examples where two or more amino acids are deleted from the sequence of a wildtype, full-length hair cell differentiation protein, the two or more deleted amino acids are not contiguous in the sequence of the wildtype, full-length protein. One skilled in the art would appreciate that amino acids that are not conserved between wildtype, full-length hair cell differentiation proteins from different species can be deleted without losing activity, while those amino acids that are conserved between wildtype, full-length hair cell differentiation proteins from different species should not be deleted as they are more likely (than amino acids that are not conserved between different species) to be involved in activity.

In some examples, an active hair cell differentiation protein can, e.g., include a sequence of a wildtype, full-length hair cell differentiation protein that has between 1 amino acid to about 100 amino acids, 1 amino acid to about 95 amino acids, 1 amino acid to about 90 amino acids, 1 amino acid to about 85 amino acids, 1 amino acid to about 80 amino acids, 1 amino acid to about 75 amino acids, 1 amino acid to about 70 amino acids, 1 amino acid to about 65 amino acids, 1

5
10
15
20
25
30

[illegible]

[illegible]

41

amino acids to about 75 amino acids, about 80 amino acids to about 100 amino acids, about 80 amino acid to about 95 amino acids, about 80 amino acids to about 90 amino acids, about 80 amino acids to about 85 amino acids, about 90 amino acids to about 100 amino acids, about 90 amino acids to about 95 amino acids, or about 95 amino acids to about 100 amino acids,

5 removed from its N-terminus and/or 1 amino acid to 100 amino acids (or any of the subranges of this range described herein) removed from its C-terminus.

In some embodiments, an active hair cell differentiation protein can, e.g., include the sequence of a wildtype, full-length hair cell differentiation protein where 1 amino acid to 50 amino acids, 1 amino acid to 45 amino acids, 1 amino acid to 40 amino acids, 1 amino acid to 35 amino acids, 1 amino acid to 30 amino acids, 1 amino acid to 25 amino acids, 1 amino acid to 20 amino acids, 1 amino acid to 15 amino acids, 1 amino acid to 10 amino acids, 1 amino acid to 9 amino acids, 1 amino acid to 8 amino acids, 1 amino acid to 7 amino acids, 1 amino acid to 6 amino acids, 1 amino acid to 5 amino acids, 1 amino acid to 4 amino acids, 1 amino acid to 3 amino acids, about 2 amino acids to 50 amino acids, about 2 amino acids to 45 amino acids, about 2 amino acids to 40 amino acids, about 2 amino acids to 35 amino acids, about 2 amino acids to 30 amino acids, about 2 amino acids to 25 amino acids, about 2 amino acids to 20 amino acids, about 2 amino acids to 15 amino acids, about 2 amino acids to 10 amino acids, about 2 amino acids to 9 amino acids, about 2 amino acids to 8 amino acids, about 2 amino acids to 7 amino acids, about 2 amino acids to 6 amino acids, about 2 amino acids to 5 amino acids, about 2 amino acids to 4 amino acids, about 3 amino acids to 50 amino acids, about 3 amino acids to 45 amino acids, about 3 amino acids to 40 amino acids, about 3 amino acids to 35 amino acids, about 3 amino acids to 30 amino acids, about 3 amino acids to 25 amino acids, about 3 amino acids to 20 amino acids, about 3 amino acids to 15 amino acids, about 3 amino acids to 10 amino acids, about 3 amino acids to 9 amino acids, about 3 amino acids to 8 amino acids, about 3 amino acids to 7 amino acids, about 3 amino acids to 6 amino acids, about 3 amino acids to 5 amino acids, about 4 amino acids to 50 amino acids, about 4 amino acids to 45 amino acids, about 4 amino acids to 40 amino acids, about 4 amino acids to 35 amino acids, about 4 amino acids to 30 amino acids, about 4 amino acids to 25 amino acids, about 4 amino acids to 20 amino acids, about 4 amino acids to 15 amino acids, about 4 amino acids to 10 amino acids, about 4 amino acids to 9 amino acids, about 4 amino acids to 8 amino acids, about 4 amino acids to 7 amino acids, about 4 amino acids to 6 amino acids, about 5 amino acids to 50 amino acids, about 5

amino acids to 45 amino acids, about 5 amino acids to 40 amino acids, about 5 amino acids to
35 amino acids, about 5 amino acids to 30 amino acids, about 5 amino acids to 25 amino acids,
about 5 amino acids to 20 amino acids, about 5 amino acids to 15 amino acids, about 5 amino
acids to 10 amino acids, about 5 amino acids to 9 amino acids, about 5 amino acids to 8 amino
5 acids, about 5 amino acids to 7 amino acids, about 6 amino acids to 50 amino acids, about 6
amino acids to 45 amino acids, about 6 amino acids to 40 amino acids, about 6 amino acids to
35 amino acids, about 6 amino acids to 30 amino acids, about 6 amino acids to 25 amino acids,
about 6 amino acids to 20 amino acids, about 6 amino acids to 15 amino acids, about 6 amino
acids to 10 amino acids, about 6 amino acids to 9 amino acids, about 6 amino acids to 8 amino
10 acids; about 7 amino acids to 50 amino acids, about 7 amino acids to 45 amino acids, about 7
amino acids to 40 amino acids, about 7 amino acids to 35 amino acids, about 7 amino acids to 30
amino acids, about 7 amino acids to 25 amino acids, about 7 amino acids to 20 amino acids,
about 7 amino acids to 15 amino acids, about 7 amino acids to 10 amino acids, about 7 amino
acids to 9 amino acids, about 8 amino acids to 50 amino acids, about 8 amino acids to 45 amino
15 acids, about 8 amino acids to 40 amino acids, about 8 amino acids to 35 amino acids, about 8
amino acids to 30 amino acids, about 8 amino acids to 25 amino acids, about 8 amino acids to 20
amino acids, about 8 amino acids to 15 amino acids, about 8 amino acids to 10 amino acids,
about 10 amino acids to 50 amino acids, about 10 amino acids to 45 amino acids, about 10
amino acids to 40 amino acids, about 10 amino acids to 35 amino acids, about 10 amino acids to
20 amino acids, about 10 amino acids to 25 amino acids, about 10 amino acids to 20 amino
20 acids, about 10 amino acids to 15 amino acids, about 15 amino acids to 50 amino acids, about 15
amino acids to 45 amino acids, about 15 amino acids to 40 amino acids, about 15 amino acids to
35 amino acids, about 15 amino acids to 30 amino acids, about 15 amino acids to 25 amino
acids, about 15 amino acids to 20 amino acids, about 20 amino acids to 50 amino acids, about 20
25 amino acids to 45 amino acids, about 20 amino acids to 40 amino acids, about 20 amino acids to
35 amino acids, about 20 amino acids to 30 amino acids, about 20 amino acids to 25 amino
acids, about 25 amino acids to 50 amino acids, about 25 amino acids to 45 amino acids, about 25
amino acids to 40 amino acids, about 25 amino acids to 35 amino acids, about 25 amino acids to
30 amino acids, about 30 amino acids to 50 amino acids, about 30 amino acids to 45 amino
30 acids, about 30 amino acids to 40 amino acids, about 30 amino acids to 35 amino acids, about 35
amino acids to 50 amino acids, about 35 amino acids to 45 amino acids, about 35 amino acids to

40 amino acids, about 40 amino acids to 50 amino acids, about 40 amino acids to 45 amino acids, or about 45 amino acids to about 50 amino acids, are inserted. In some examples, the 1 amino acid to 50 amino acids (or any subrange thereof) can be inserted as a contiguous sequence into the sequence of a wildtype, full-length protein. In some examples, the 1 amino acid to 50 amino acids (or any subrange thereof) are not inserted as a contiguous sequence into the sequence of a wildtype, full-length protein. As can be appreciated in the art, the 1 amino acid to 50 amino acids can be inserted into a portion of the sequence of a wildtype, full-length protein that is not well-conserved between species.

Atonal basic Helix-Loop-Helix transcription factor 1 (Atoh1)

The ATOH1 gene encodes atonal basic helix-loop-helix (bHLH) transcription factor 1. ATOH1 is a primary regulator of hair cell differentiation (Kawamoto et al., *J. Neurosci.* (2003) 23(11):4395-4400; Izumikawa et al. (2005) *Nat. Med.* 11(3): 271-276; Minoda et al. (2007) *Hear Res.* 232(1-2): 44-51; Atkinson et al. (2014) *PLoS One* 9(7): e102077; Kuo et al. (2015) *J. Neurosci* 35(30): 10786-10798; Walters et al. (2017) *Cell Rep* 19(2): 307-320).

The human ATOH1 gene is located on chromosome 4q22. It contains 1 exon encompassing ~2 kilobases (kb) (NCBI Accession No. NM_005172.1). The full-length wildtype ATOH1 protein expressed from the human ATOH1 gene is 354 amino acids in length.

Non-limiting examples of detecting techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype ATOH1 protein is or includes the sequence of SEQ ID NO: 1. Non-limiting examples of a nucleic acid encoding a wildtype ATOH1 protein is or includes SEQ ID NO: 4. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 4 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype ATOH1 Protein (SEQ ID NO: 1)

MSRLLHAEWEAEVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSLLDSTDPAWLAPTLOGICTARA
AQYLLHSPELGASEAAAPRDEVDRGELVRRSSGGASSKSPGPVKVREQLCKLKGGVVVDELGCSRQRAPSSKQVN
GVQKQRRLLAANARERRRMHGLNHAFDQLRNVI PSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASC
KSDHHHLRTAASYEGGAGNATAAGAAQASGGGSRPTPPGSCRTRFSAPASAGGYSVQLDALHFSTFEDSALTAMMAQ
KNLSPSLPGSILQPVQEENSKTSPRSHRSDGEFSPHSHYSDSDEAS

Mouse Full-Length Wildtype ATOH1 Protein (SEQ ID NO: 2)

MSRLLHAEWEAEVKELGDHHRHPQPHVPLTPQPPATLQARDLPVYPAELSLLDSTDPRAWLTPTLQGLCTARAAQ
 YLLHSPELGASEAAAPRDEADSQGELVRRSGCGGLSKSPGPVKVREQLCKLKGGVVVDELGCSRQRAPSSKQVNGVQ
 KQRRLAANARERRRMHGLNHAFDQLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPNVGEQPPPPTASCKND
 5 HHHLRRTASSYEGGAGASAVAGAPAPGGGPRPTPPGPCRTRFSGPASSGGYSVQLDALHFPAFEDRAL TAMMAQKDL
 SPSLPGGILQPVEDNSKTSRSHRSDGEFSPHSHYSDSDEAS

Rat Full-Length Wildtype ATOH1 Protein (SEQ ID NO: 3)

MSRLLHAEWEAEVKELGDHHRHPQPHHIPQLTPQPPATLQARDHPVYPAELSLLDSTDPRAWLTPTLQGLCTARAAQ
 10 YLLHSPELGASEAAAPGDEADGQGELVRRSGCGGLSKSPGPVKVREQLCKLKGGVVVDELGCSRQRAPSSKQVNGVQ
 KQRRLAANARERRRMHGLNHAFDQLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSVGEQPPPPAASCKND
 HHHLRAAA SYEGGAGASAVAGAPAPGGGPRPTPPGACRTRFSTPASSGGYSVQLDALHFPAFEDRAL TAMMAQKDL
 SPSLPGGILQVPEDSSKTSRSHRSDGEFSPHSHYSDSDEAS

Human Wildtype ATOH1 cDNA (SEQ ID NO: 4)

atgtcccgctgctgcatgcagaagagtgggctgaagtgaaggagtgggagaccaccatcgccagccccagccgca
 tcatctcccgaaccgcccgcgcgcagccaccttgcagggcagagagcatcccgctctaccgcctg
 agctgtccctcctggacagcaccgaccacgcgcctggctggctcccactttgcagggcatctgcacggcacgcgcc
 gccagtatcttctacattccccggagctgggtgcctcagaggccgctgcgccccgggacgaggtggacggccgggg
 20 ggagctggtaaggaggagcagcggcggtgccagcagcagcaagagccccgggcccgggtgaaagtgcgggaacagctgt
 gcaagctgaaaggcgggggtggtggtagacgagctgggctgcagccgccaacggggccccttcagcaaacaggtgaat
 ggggtgcagaagcagagacggctagcagccaacgccaggaggcggcgagcatgcatgggctgaaccacgccttcga
 ccagctgcgcaatgttatcccgctcggtcaacaacgacaagaagctgtccaaatatgagaccctgcagatggccaaaa
 tctacatcaacgccttgtccgagctgctacaaacgcccagcggaggggaacagccaccgcccgcctccagcctcctgc
 25 aaaagcgaccaccaccaccttcgcaccgcggcctcctatgaagggggcgcgggcaacgcgaccgcagctggggctca
 gcaggcttccggaggaggcagcggccgaccccgcccgggagttgccggactcgcttctcagccccagcttctgcgg
 gaggggtactcgggtgcagctggacgctctgcacttctcgactttcgaggacagcgccctgacagcgatgatggcgcaa
 aagaatttgtctccttctctccccgggagcatcttgcagccagtgaggagggaaaacagcaaaacttcgcctcggtc
 ccacagaagcgacggggaattttccccccattcccattacagtgcggtgaggcaagtttag

A non-limiting example of a human wildtype ATOH1 genomic DNA sequence is SEQ ID NO: 5. The exon in SEQ ID NO: 5 is: nucleotide positions 1-1065 (exon 1).

Human Wildtype ATOH1 Gene (SEQ ID NO: 5)

35 1 atgtcccgcc tgctgcatgc agaagagtgg gctgaagtga aggagttggg agaccacat
 61 cgccagcccc agccgcatca tctcccgcaa ccgcccgcgc cgccgcagcc acctgcaact

121 ttgcaggcga gagagcatcc cgtctaccgc cctgagctgt ccctcctgga cagcaccgac
 181 ccacgcgcct ggctggctcc cactttgcag ggcatctgca cggcacgcgc cgcccagtat
 241 ttgctacatt ccccgagct ggggtgcctca gaggccgctg cgccccggga cgagggtggac
 301 ggccgggggg agctggtaag gaggagcagc ggcggtgcca gcagcagcaa gagccccggg
 5 361 ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg ggggtggtgt agacgagctg
 421 ggctgcagcc gccaacgggc cccttcacgc aaacagggtga atgggggtgca gaagcagaga
 481 cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac
 541 cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagacc
 601 ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gccacgcgga
 10 661 ggggaacagc caccgccgcc tccagcctcc tgcaaaagcg accaccacca ccttcgcacc
 721 gcggcctcct atgaaggggg cgcgggcaac gcgaccgag ctgggggtca gcaggcttcc
 781 ggaggggagc agcggccgac cccgcccggg agttgccgga ctgcgttctc agccccagct
 841 tctgcgggag ggtactcggg gcagctggac gctctgcact tctcgacttt cgaggacagc
 901 gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg
 15 961 cagccagtgc aggaggaaaa cagcaaaact tcgcctcggg ccacagaaag cgacggggaa
 1021 ttttcccccc attcccatta cagtgactcg gatgaggcaa gtttag

POU Class 4 Homeobox 3 (Pou4f3)

The POU4F3 gene encodes POU class 4 homeobox 3, and acts as a transcriptional
 20 activator. POU4F3 activates ATOH1 transcription in early development and is later further
 activated by ATOH1 and required for hair cell survival after birth. POU4F3 activates NT3 and
 BDNF. Mutations in POU4F3 have been associated with hearing loss (Lee et al. (2010) *Biochem
 Biophys Res Commun* 396(3):626-630; Clough et al. (2004) *Biochem Biophys Res Commun*
 324(1):372-381; Costa et al. (2015) *Development* 142(11):1948-1959; and Walters et al. (2017)
 25 *Cell Rep* 19(2):307-320).

The human POU4F3 gene is located on chromosome 5q32. It contains 2 exons
 encompassing ~15 kilobases (kb) (NCBI Accession No. NG_011885.1). The full-length
 wildtype POU4F3 protein expressed from the human POU4F3 gene is 338 amino acids in length.

Various mutations in the POU4F3 gene have been associated with hearing loss due to
 30 hair cell degeneration. For example, a nonsense mutation c.337C>T in POU4F3 was identified
 to cause autosomal dominant hearing loss (Zhang et al. (2016) *Neural Plast*
 doi:10.1155/2016/1512831).

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype POU4F3 protein is or includes the sequence of SEQ ID NO: 6. Non-limiting examples of nucleic acid encoding a wildtype POU4F3 protein are or include SEQ ID NO: 9. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 9 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype POU4F3 Protein (SEQ ID NO: 6)

MMAMNSKQPFMHPVLQEPKFSSLSHSGSEAMRRVCLPAPQLQGNI FGSFDESLARAEALAAVDIVSHGKNHPFKPD
 ATYHTMSSVPCTSTSTVPISHPAALTSHPHHAVHQGLEGDLEHISPTLSVSGLGAPESHVMPAQIHPHHLGAMGH
 LHQAMGMSHPHTVAPHSAMPACLSDVESDPRELEAFAERFKQRRIKLGVQTADVGAALANLKIPGVGSLSQSTICRF
 ESLTSLSHNNMIALKPVLQAWLEEAEEAAYREKNSKPELFNGSERKRKRTSIAAPEKRSLEAYFAIQPRPSSEKIAAIA
 EKLDLKKNVVRVWFCNQRQKQKRMKYS AVH

Mouse Full-Length Wildtype POU4F3 Protein (SEQ ID NO: 7)

MMAMNAKQPFMHPVLQEPKFSSLSHSGSEAMRRVCLPAPQLQGNI FGSFDESLARAEALAAVDIVSHGKNHPFKPD
 ATYHTMSSVPCTSTSTPTVPISHPAALTSHPHHAVHQGLEGDLEHISPTLSVSGLGAPESHVMPAQIHPHHLGAMGH
 LHQAMGMSHPHAVAPHSAMPACLSDVESDPRELEAFAERFKQRRIKLGVQTADVGAALANLKIPGVGSLSQSTICRF
 ESLTSLSHNNMIALKPVLQAWLEEAEEAAYREKNSKPELFNGSERKRKRTSIAAPEKRSLEAYFAIQPRPSSEKIAAIA
 EKLDLKKNVVRVWFCNQRQKQKRMKYS AVH

Rat Full-Length Wildtype POU4F3 Protein (SEQ ID NO: 8)

MMAMNAKQPFMHPVLQEPKFSSLSHSGSEAMRRVCLPAPQLQGNI FGSFDESLARAEALAAVDIVSHGKNHPFKPD
 ATYHTMSSVPCTSTSTPTVPISHPAALTSHPHHPVHQGLEGDLEHISPTLSVSGLGAPESHVMPAQIHPHHLGAMGH
 LHQAMGMSHPHAVAPHSAMPACLSDVESDPRELEAFAERFKQRRIKLGVQTADVGAALANLKIPGVGSLSQSTICRF
 ESLTSLSHNNMIALKPVLQAWLEEAEEAAYREKNSKPELFNGSERKRKRTSIAAPEKRSLEAYFAIQPRPSSEKIAAIA
 EKLDLKKNVVRVWFCNQRQKQKRMKYS AVH

Human Wildtype POU4F3 cDNA (SEQ ID NO: 9)

atgatggccatgaactccaagcagcctttcggcatgcacccggtgctgcaagaacccaaattctccagtctgcactc
 tggctccgaggccatgcgcccagctctgtctcccagccccgcagctgcagggtaatatatttggagccttgatgaga
 gcctgctggcacgcgccgaagctctggcggcggtggatatcgtctcccacggcaagaaccatccgttcaagcccgcac
 gccacctaccataccatgagcagcgtgcctgcacgtccacttcgtccaccgtgcccattctcccaccagctgcgct
 cacctcacaccctcaccacgcctgtgcaccaggcctcgaaggcgacctgctggagcacatctcgccacgctgagtg

tgagcggcctgggcgctccggaacactcggtgatgcccgcacagatccatccacaccacctgggcgccatggggccac
 ctgcaccaggccatgggcatgagtcacccgcacaccgtggccccctcatagcgccatgcctgcatgcctcagcgacgt
 ggagtcagacccgcgcgagctggaagccttcgccgagcgcttcaagcagcgggcgcatcaagctgggggtgacccagg
 cggacgtgggcgcggtcttggttaatctcaagatccccggcggtgggtcgtgagccaaagcaccatctgcaggttc
 5 gagtctctcactctctcgcacaacaacatgatcgtcttcaagccggtgctccaggcctggttgaggaggccgagggc
 cgcctaccgagagaag aacagcaagccagagctcttcaacggcagcgaacggaagcgcaaacgcacgtccatcgcg
 gcgccggagaagcggttactcagaggcctatcttgcgtatccagccacgtccttcatctgagaaga|cgcgccatcg
 tgagaaactggaccttaaaaagaacgtggtgagagctctggttctgcaaccagagacagaaacagaaacgaatgaagt
 attcggtgtccactga

A non-limiting example of a human wildtype POU4F3 genomic DNA sequence is SEQ ID NO: 10. The exons in SEQ ID NO: 10 are: nucleotide positions 1-209 (exon 1) and nucleotide positions 525-1497 (exon 2). The intron in SEQ ID NO: 10 is: nucleotide positions 210-524 (intron 1).

Human Wildtype POU4F3 Gene (SEQ ID NO: 10)

1 cgctgagcag cgctcacttg gagagcggca agcaagctag acaagcctga ttccatgtca
 61 cccgctgccca ccctgccagg agcgcgaaga tgatggccat gaactccaag cagcctttcg
 121 gcatgcaccc ggtgctgcaa gaacccaaat tctccagtct gcaactctggc tccgaggcca
 20 181 tgcgccgagt ctgtctccca gccccgcagg tacgtagtgg agcataatta ccgctctaag
 241 gcacattttt tgacaggcac tagcttcatg ttttttcat gtcgccaga acaatcgccg
 301 ctgtctgaac ccctctcctt gtctcccccg cgttctctcc cggcgcgctc tctctctcat
 361 tcatgtctct gatccacacg tctgttccag cagagccgct gcctccgtat taatttttat
 421 gacctgggct ttgaggagag gcatctcggg tgcttgaaaa tgtgttttaa tctgtgttg
 25 481 acagtattcc ctactgaccg tgctgtgctc cttctcgctt gcagctgcag ggtaatatat
 541 ttggaagctt tgatgagagc ctgctggcac gcgccgaagc tctggcggcg gtggatatcg
 601 tctcccacgg caagaaccat ccgttcaagc ccgacgccac ctaccatacc atgagcagcg
 661 tgccctgcac gtccacttcg tccaccgtgc ccatctccca cccagctgcg ctcacctcac
 721 accctcacca cgccgtgcac cagggcctcg aaggcgacct gctggagcac atctcgccca
 30 781 cgctgagtgt gagcggcctg ggcgctccgg aacactcggg gatgcccgca cagatccatc
 841 cacaccacct gggcgccatg ggccacctgc accaggccat gggcatgagt caccgcaca
 901 ccgtggcccc tcatagcgcc atgcctgcat gcctcagcga cgtggagtca gacccgcgcg
 961 agctggaagc cttcgccgag cgcttcaagc agcggcgcat caagctgggg gtgacccagg
 1021 cggacgtggg cgcggctctg gctaattctca agatccccgg cgtgggctcg ctgagccaaa
 35 1081 gcaccatctg caggttcgag tctctcactc tctcgacaaa caacatgatc gctctcaagc
 1141 cgggtgtcca ggcttggtg gaggaggccg aggccgcta ccgagagaag aacagcaagc
 1201 cagagctctt caacggcagc gaacggaagc gcaaacgcac gtccatcgcg gcgccggaga

1261 agcggttcaact cgaggcctat ttcgctatcc agccacgtcc ttcacatctgag aagatcgcg
 1321 ccacgctga gaaactggac cttaaaaaga acgtggtgag agtctggttc tgcaaccaga
 1381 gacagaaaca gaaacgaatg aagtattcgg ctgtccactg attgcggcag ggcgcagcgt
 1441 cgggagccgg gagagcctag tgctcatccc tcccgggttc gggggatggt tatcggg

5

Catenin beta 1 (CTNNB1)

The CTNNB1 gene encodes catenin beta 1 (β -Catenin), a protein involved both in transcriptional activation and in adherens junctions. CTNNB1 is required for hair cell development and differentiation. β -Catenin activates ATOH1 through binding to its enhancer.

10 Overexpression or stabilization of CTNNB1 results in supporting cell proliferation and differentiation into hair cells (Shi et al. (2013) *Proc Natl Acad Sci U S A.* 110(34):13851-13856; Kuo et al. (2015) *J. Neurosci.* 35(30):10786-10798). Knock-out of CTNNB1 in early development prevents hair cell differentiation (Shi et al. (2013) *J. Neurosci.* 34(19):6470-6479). Overexpression of CTNNB1 induces ectopic hair cells.

15 The human CTNNB1 gene is located on chromosome 3p22. It contains 15 exons encompassing ~41 kilobases (kb) (NCBI Accession No. NG_013302.2). The full-length wildtype CTNNB1 protein expressed from the human CTNNB1 gene is 781 amino acids in length.

20 Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype CTNNB1 protein is or includes the sequence of SEQ ID NO: 11. Non-limiting examples of a nucleic acid encoding a wildtype POU4F3 protein is or includes SEQ ID NO: 14. As can be appreciated in the art, at least some or all of the codons in
 25 SEQ ID NO: 14 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype CTNNB1 Protein (SEQ ID NO: 11)

30 MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTTAPSLSGKGNPEEEDVDTSQVLYEWEQGFSQSFTQE
 QVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDDAAHPTNVQRLAEPQMLKHAVVNLINYQDDAELATRAIP
 ELTKLLNDEQVVVNKAAMVHQLSKKEASRAIMRSPQMVSIAIVRTMQNTNDVETARCTAGTLHNLSHHREGLLAI
 FKSGGIPALVKMLGSPVDSVLFYAITTLHNLLHQQEGAKMAVRLAGGLQKMVALLNKTNVKFLAITTDCLQILAYGN
 QESKLIILASGGPQALVNIMRTYTYEKLWTTSRVLKVLVSCSSNKPAIVEAGGMQALGLHLTDPSQRLVQNCLWTL

RNLSDAATKQEGMEGLLGLTLVQLLGSDDINVVTCAAGILSNLTCNNYKNKMMVCQVGGIEALVRTVLRAGDREDITE
 PAICALRHILTSRHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLALCPANHAPLREQGAI PRLVQL
 LVRAHQDTQRRITSMGGTQQQFVEGVRMEEIVEGCTGALHILARDVHNRI VIRGLNTIPLFVQLLYSPIENIQRVAAG
 VLCELAQDKEAAEAIEAEGATAPLTELLHSRNEGVATYAAAVLFRMSEDKPQDYKKRLSVELTSSLFRTEPMAWN
 5 ADLGLDIGAQGEPLGYRQDDPSYRSFHSGGYQDALGMDPMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGD
 SNQLAWFDTDL

Mouse Full-length Wildtype CTNNB1 Protein (SEQ ID NO: 12)

MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTAPSLSGKGNPEEEDVDTSQVLYEWEQGFSQSFTQE
 10 QVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDDAHPTNVQRLAEP SQMLKHAVVNLINYQDDAELATRAIP
 ELTKLLNDEDDQVVVNKAAMVHQLSKKEASRHAIMRSPQMVS AIVRTMQNTNDVETARCTAGTLHNLSSHREGLLAI
 FKSGGIPALVKMLGSPVDSVLFYAITTLHNL LLLHQEGAKMAVRLAGGLQKMVALLNKTNVKFLAITTDCLQILAYGN
 QESKLIIILASGGPQALVNIMRTYTYEKL LWTTSRVLKVL SVCSSNKPAIVEAGGMQALGLHLTDPSQRLVQNCLWTL
 RNLSDAATKQEGMEGLLGLTLVQLLGSDDINVVTCAAGILSNLTCNNYKNKMMVCQVGGIEALVRTVLRAGDREDITE
 15 PAICALRHILTSRHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLALCPANHAPLREQGAI PRLVQL
 LVRAHQDTQRRITSMGGTQQQFVEGVRMEEIVEGCTGALHILARDVHNRI VIRGLNTIPLFVQLLYSPIENIQRVAAG
 VLCELAQDKEAAEAIEAEGATAPLTELLHSRNEGVATYAAAVLFRMSEDKPQDYKKRLSVELTSSLFRTEPMAWN
 ADLGLDIGAQGEALGYRQDDPSYRSFHSGGYQDALGMDPMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGD
 20 SNQLAWFDTDL

Rat Full-length Wildtype CTNNB1 Protein (SEQ ID NO: 13)

MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTAPSLSGKGNPEEEDVDTSQVLYEWEQGFSQSFTQE
 QVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDDAHPTNVQRLAEP SQMLKHAVVNLINYQDDAELATRAIP
 ELTKLLNDEDDQVVVNKAAMVHQLSKKEASRHAIMRSPQMVS AIVRTMQNTNDVETARCTAGTLHNLSSHREGLLAI
 25 FKSGGIPALVKMLGSPVDSVLFYAITTLHNL LLLHQEGAKMAVRLAGGLQKMVALLNKTNVKFLAITTDCLQILAYGN
 QESKLIIILASGGPQALVNIMRTYTYEKL LWTTSRVLKVL SVCSSNKPAIVEAGGMQALGPHLTDPSQRLVQNCLWTL
 RNLSDAATKQEGMEGLLGLTLVQLLGSDDINVVTCAAGILSNLTCNNYKNKMMVCQVGGIEALVRTVLRAGDREDITE
 PAICALRHILTSRHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLALCPANHAPLREQGAI PRLVQL
 LVRAHQDTQRRITSMGGTQQQFVEGVRMEEIVEGCTGALHILARDVHNRI VIRGLNTIPLFVQLLYSPIENIQRVAAG
 30 VLCELAQDKEAAEAIEAEGATAPLTELLHSRNEGVATYAAAVLFRMSEDKPQDYKKRLSVELTSSLFRTEPMAWN
 ADLGLDIGAQGEALGYRQDDPSYRSFHSGGYQDALGMDPMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGD
 SNQLAWFDTDL

Human Wildtype CTNNB1 cDNA (SEQ ID NO: 14)

atggctactcaagctgatttgatggagttggacatggccatggaaccagacagaaaagcggctgttagtcactggca
 gcaacagtcttacctggactctggaatccattctggtgccactaccacagctccttctctgagtggttaaaggcaatc
 ctgaggaagaggatgtggatacctcccaagtcctgtatgagtggaacagggattttctcagtccttactcaagaa

caagtagctgatattgatggacagtatgcaatgactcgagctcagagggtagcagctgctatgttccctgagacatt
 agatgagggcatgcagatcccactctacacagtttgatgctgctcatcccactaatgtccagcgtttggtgaaccat
 cacagatgctgaaacatgcagttgtaaacttgattaactatcaagatgatgcagaacttgccacacgtgcaatccct
 gaactgacaaaactgctaaatgacgaggaccaggtggtggttaataaggctgcagttatggtccatcagctttctaa
 5 aaaggaagcttccagacacgctatcatgcgttctcctcagatgggtgtctgctattgtacgtaccatgcagaatacaa
 atgatgtagaaacagctcggtgtaccgctgggaccttgacataacctttcccatcatcgtgagggcttactggccatc
 ttttaagtctggaggcattcctgcccctggtgaaaatgcttggttcaccagtggttctgtgtgtgttttatgccattac
 aactctccacaaccttttattacatcaagaaggagctaaaatggcagtgcgtttagctgggtgggctgcagaaaatgg
 ttgcttgctcaacaaaacaaatgttaaattcttggtattacgacagactgccttcaaatttttagcttatggcaac
 10 caagaaagcaagctcatcactactggctagtgggtggacccaagcttttagtaaatataatgaggacctatacttacga
 aaaactactgtggaccacaagcagagtgctgaaggtgctatctgtctgctctagtaataagccggctattgtagaag
 ctggtggaatgcaagctttaggacttcacctgacagatccaagtcaacgtcttgttcagaactgtctttggactctc
 aggaatctttcagatgctgcaactaaacaggaagggtggaaggtctccttgggactcttgttcagcttctgggttc
 agatgatataaatgtggtcacctgtgcagctggaattctttctaacctcacttgcaataattataagaacaagatga
 15 tggctgccaagtgggtggtatagaggctcttgtgcgtactgtccttcgggctggtgacaggggaagacatcactgag
 cctgccatctgtgctcttcgtcatctgaccagccgacaccaagaagcagagatggccagaatgcagttcgcttca
 ctatggactaccagttgtggttaagctcttacaccaccatcccactggcctctgataaaggctactgttggattga
 ttcgaaatcttgccctttgtcccgcaatcatgcacctttgctgagcaggggtgccattccacgactagttcagttg
 cttgttcgtgcacatcaggatacccagcgccgtacgtccatgggtgggacacagcagcaatttgtggagggggccg
 20 catggaagaaatagttgaaggtgtaccggagcccttcacatcctagctcgggatgttcacaaccgaattgttatca
 gaggactaaataccattccattgtttgtgcagctgctttattctccattgaaaacatccaaagagtagctgcaggg
 gtcctctgtgaacttgctcaggacaaggaagctgcagaagctattgaagctgagggagccacagctcctctgacaga
 gttacttcactctaggaatgaaggtgtggcgacatatgcagctgctgttttgttccgaatgtctgaggacaagccac
 aagattacaagaaacggctttcagttgagctgaccagctctctcttcagaacagagccaatggcttggaaatgagact
 25 gctgatcttgacttgatattggtgcccagggagaacccttgatatacgccaggatgatcctagctatcgttcttt
 tcactctggtggatattggccaggatgccttgggtatggaccccatgatggaacatgagatgggtggccaccaccctg
 gtgctgactatccagttgatgggtgccagatctggggcatgccaggacctcatggatgggtgcctccaggtgac
 agcaatcagctggcctggtttgatactgacctgtaa

30 A non-limiting example of a human wildtype CTNNB1 genomic DNA sequence is SEQ
 ID NO: 15. The exons in SEQ ID NO: 15 are: nucleotide positions 1-220 (exon 1), nucleotide
 positions 24571-24631 (exon 2), nucleotide positions 25076-25303 (exon 3), nucleotide
 positions 25504-25757 (exon 4), nucleotide positions 25884-26122 (exon 5), nucleotide
 positions 26210-26411 (exon 6), nucleotide positions 27758-27902 (exon 7), nucleotide
 35 positions 33891-33994 (exon 8), nucleotide positions 34079-34417 (exon 9), nucleotide
 positions 34689-34847 (exon 10), nucleotide positions 36274-36393 (exon 11), nucleotide

positions 36899-37049 (exon 12), nucleotide positions 37138-37259 (exon 13), nucleotide positions 38566-38626 (exon 14), and nucleotide positions 39684-40998 (exon 15). The introns in SEQ ID NO: 15 are: nucleotide positions 221-24570 (intron 1), nucleotide positions 24632-25075 (intron 2), nucleotide positions 25304-25503 (intron 3), nucleotide positions 25758-24883 (intron 4), nucleotide positions 26123-26209 (intron 5), nucleotide positions 26412-27757 (intron 6), nucleotide positions 27903-33890 (intron 7), nucleotide positions 33995-34078 (intron 8), nucleotide positions 34418-34688 (intron 9), nucleotide positions 34848-36273 (intron 10), nucleotide positions 36394-36898 (intron 11), nucleotide positions 37050-37137 (intron 12), nucleotide position 37260-38565 (intron 13), and nucleotide position 38627-39683 (intron 14).

Human Wildtype CTNNB1 Gene (SEQ ID NO: 15)

	1	aggatacagc	ggcttctg	cgacttataa	gagctccttg	tgcggcgcca	ttttaagcct
	61	ctcgggtctgt	ggcagcagcg	ttggcccggc	cccgggagcg	gagagcgagg	ggaggcggag
15	121	acggaggaag	gtctgaggag	cagcttcagt	ccccgccgag	ccgccaccgc	aggtcgagga
	181	cggtcggact	cccgcggcgg	gaggagcctg	ttcccctgag	gtgcttgggc	gctcctttcc
	241	ttatccttcc	ggggctgctc	ccgcttcctc	tcggagccaa	acttcgtagc	aggcgcgcgg
	301	tccgggcggc	gggctggg	cagccgggag	gcctgggggt	gggagcggg	agctcagggtg
	361	ggggacgggtg	aggggtgggc	gcgcccggg	cgcggaaggc	ggcggccggg	cccggttcc
20	421	ggtcgcgctg	cctctctggg	gccctggggg	catcgcttgc	ggggagggg	cgccgcggg
	481	gcgcgtacag	gagcccggat	ggcaggcggg	gtgggggtg	gggtgggggt	ctgtggtttc
	541	cgtccggggc	tctggccttg	gccgagttt	ggggagggac	ccggtgcctc	gggatgcgcc
	601	gggccctggg	tggggggcgg	ggtggggacg	gggggctccg	ccttctcagc	tcttgcggcg
	661	agttgggggt	cgggcgctga	ggcagagacg	ccaccctaag	tcccatcagt	cctggggatc
25	721	ggaccagtgg	actttctctt	aagatttcct	ctttcattct	taagaataga	agtgttatta
	781	tttttttttaa	tgccctggct	atgtgagttt	gaatcgaagc	aactttaaac	cttagagcaa
	841	ctaaactcta	agtgcagcgg	gtgcgatgcg	tcagtagggg	gagcacataa	aaaatccatg
	901	tcttgacact	gtatttttagc	gtactatgca	ggtgagtga	agcagtggt	aatgtactgg
	961	gagcttatg	gatttatggt	agtgggtatg	agaccctggt	gaaataagg	ggtggaggaa
30	1021	ggcgaagggtg	atggcttact	gtttcttacc	aagtgaactg	caggattcag	cctctgactc
	1081	agaccgcttc	gagaattttt	ttcgtagaaa	taatttaaat	ttattcaaat	agtttgatgg
	1141	cagctaaaat	tgaattatag	agcacgtttt	cttttcagcg	gagtgaattt	ttccttcgct
	1201	ccaaagctgg	ccaaatggaa	ttcaagcatt	gcaacttctt	tcagtgtttt	gtctggagag
	1261	aggactttga	accgagactt	ttcgaagtta	agttcctata	gcctgcttct	gaatctgcca
35	1321	agcttgaaag	ctttggcagt	tgggtgtatg	tagttgttgc	cttcgttctc	ttcccttttg
	1381	gaggagcgt	tgtctcctac	tttgtatctt	ccagacatct	gtggtcttcc	ccccaccctt
	1441	cgagtttgtg	agtgggtgaat	gaagaaagac	taggctgctg	gtatgcagag	gtcggcaaaa
	1501	ggaatcgag	gagtgggtttt	agtgaatga	gagctttgta	tcataataaa	tgggtggcta
	1561	ggctagacat	caacttgaag	agacggcagc	atttcctttc	ataaagtcta	ggctaattgt
40	1621	tttcagatcg	ctaagtgtga	gtttgtctgg	aatttaggaa	gccatttcag	tatttgtcac
	1681	ttgggtgaacg	aacattcaat	accttcagat	gtcttcgtgt	tgacttgtat	tcatactaag
	1741	aaatagtaaa	tatagtctca	agtgttattt	atgttatact	gctggtttat	tctctgctta
	1801	aattattgac	ataaatttct	actttggagg	cttttcgttt	gaactaaggc	tgtgcggaat
	1861	ttatttttact	tttatattta	aatctttgaa	aaatctctga	ttaaaaaaaa	agtaccctta
45	1921	aaggtttgag	gatgtccttt	cacaccagac	aaaatttggt	taatttgccg	ccaatattca
	1981	ttactttgac	ctaacccttg	ttctgaaggc	cgtgtacaag	gacaaggccc	tgagattatt
	2041	gcaacagtaa	cttgaaaaac	tttcagaagt	ctattctgta	ggattaaagg	aatgctgaga

	2101	ctattcaagt	ttgaagtcct	gggggtgggg	aaaaataaaa	aacctgtgct	agaaagctta
	2161	gtatagcatg	taactttaga	gtcctgtgga	gtcctgagtc	tcccacagac	cagaacagtc
	2221	atttaaaagt	tttcaggaaa	aaccaactta	aaaaaaaaata	aggtggctaa	ttaaaaaaaaa
5	2281	atgaagcatt	taacagtgtt	caggtttcag	agtatggaag	aggggttttt	taaactgtta
	2341	tctgattatt	tctttttacca	acatgatata	gaaaagtgtg	tttccagtat	taaaatttat
	2401	cagactgagc	ttactgttcc	tgtaaatgac	tggaataaaa	attggcataa	atgagggctc
	2461	gtatgcttgt	tttaataaca	ccaccaccaa	gatagaaaac	gaggaggcaa	gtttctccaa
	2521	gggtattttg	aaatgtgtta	gcaaaactat	tcagataact	cgtttttgtt	ataggggtgag
10	2581	gtggggagag	gcgcagtgcta	agtattgttg	aaactaggga	tgtagagaat	taaaagtttg
	2641	aatataatta	ttttgtagtt	ataagtagca	gtgaaattaa	atctcctgca	atagactata
	2701	gaagtatat	tagccaaatg	aaacttcagt	gttattgaaa	tgaaataata	catctgtcct
	2761	gttacaagat	tattttttatt	tctctgtgtg	tttcctagct	tctgataatc	aataattgta
	2821	gatgagtagg	tggttaagttt	taagtttgta	ctttgagctt	agtcggaagc	atgcttgact
	2881	gccaaaccgg	ggcacaaggg	atgaaggctt	ttagaactgg	acaaacttct	aacaaaaggt
15	2941	atgtgcaact	cttttgtagt	gtgtcatgtt	gatttgtgac	attgtttttg	aaaatatgtg
	3001	ttaacttagt	tttctttagt	ccctcttttt	attggaactg	tggtatctat	tggtgaaact
	3061	gcttgactga	gaacattttt	ataccataaa	agtaaatagt	aaacatagcc	caggagcggc
	3121	ttctggtttg	tccatcgtat	gtagccattg	cctccttgta	ctctcattga	gaagatactg
	3181	atgtgcagat	tcagttgtcc	ttctctaaca	gactatttat	gtaatatgtc	agttgtgatt
20	3241	gtgataggta	agtgagccag	tcgggttaaaa	taaataactca	ggtttcacaa	aaggaaaata
	3301	atatgatttg	tggtgatcta	aatgagtata	ggagttaact	cctatagttt	ttcatcactt
	3361	aaactcaggg	gaaagtcttt	tatttcctct	gtttacttaa	gaatgctgct	tttgtgtttc
	3421	atgcaagact	gagcttgact	cagtttgaaa	cctaggctca	tctgttgagg	cctgaaccct
	3481	gctgtccttg	aagtatgcat	ataatttgct	tccttcctaa	ggaaaaataa	gctcttgaaa
25	3541	gataaagtca	atcacattag	gaacccattt	ttagggttta	gccacttttt	tttttttttt
	3601	tttttaactc	atgggcatct	cttctgttaa	gagacattcc	ccactctcca	agtttccttc
	3661	aagcctgaag	cagcagagtg	agtagtggtg	gagcatgttt	tcattgcatg	cttgggtcat
	3721	gttgagtgcc	ctccagtggg	tatagtataa	tgcttgatga	tttttttttt	ttaattccaa
30	3781	acaagtttat	gtgggatata	tttaggaata	gttctgatga	gggagaatca	actaagaaac
	3841	ctttgatttc	taaaataatt	aatatcatta	ctgctaatta	aaatacaggc	ttgagaaaat
	3901	gtcttctcag	ccaatatattg	cagtagaaaa	gtcggggagg	tttttaaggt	cacttttgagt
	3961	aggcagttct	gcttaaatat	atcataatga	taaaccagaa	tctcagtata	gtactttagg
	4021	aggtaaaaga	tcataaatatt	cagttatatt	gatgaattac	agcaactgaa	attctcagaa
	4081	aaaaattaat	gaaaatgtga	attgtcaatt	tgtctaaaat	cattcacaga	gtaaaacata
35	4141	agtgtctaac	ttgattatat	taggaaaatag	atagaaataa	aggtaattga	gccagtgtat
	4201	gtgacctaaa	atataatgcc	cttagtgacc	ataggggttg	tctcatttgt	acatagtggg
	4261	gggccatgat	gaactgtgtt	ttgccctttg	aatttttcct	taaaaagctt	tctctaggct
	4321	cctatgttca	tggtttttct	gttagtaata	ttattttctg	aaaatccatg	tttcaaatca
	4381	gaatctaatt	agcaacagga	atgaagctta	ttctaaatta	gtttttggaa	gttaaacggg
40	4441	cagcatatgg	aaatttttca	gggttttagat	ttttaaaaa	ttgtttttca	gaatatgttg
	4501	ctggaatgaa	aacgttagcg	tagggacgga	aatgacact	taccagtgat	tgctttactt
	4561	tgctgtgga	attcagtgtg	atttttgtgga	aacattggta	tatgattttt	tactacttaa
	4621	gaaatgtatt	gctatagtta	gggttttttt	ttttttaaa	gcaagaatgc	ctcaagtgtc
	4681	ttatgtgaat	gattatttca	ggatggatta	aatattcctc	catcaaggac	catacttgta
45	4741	aatcagtgat	ttccaagttg	gtgcttagta	tttacagcat	ttactgtcta	taagcttctg
	4801	ttctgatttt	tcaagagttt	tctgagaaat	gagagttagc	ttaaaagttc	tttgaaaaat
	4861	tatgtacata	caacttactg	aaaaaaattg	ctaccgggga	cttaatttgt	ctcttgaaat
	4921	gggctacttg	ccttcattaa	tgtagcatat	tacaatttga	tgttcaagat	atgttactaa
	4981	gaataagatc	gctttcagaa	gccttatata	ggatttgtct	tactacattg	tagtgggaat
50	5041	ggctactcaa	atgtctccag	ggccagttag	gtattgggta	aatgggacca	tgacagactat
	5101	taaaaattga	agtgcacatg	aagcagccag	tcataagcag	ctccagccac	tgtgtgggaa
	5161	tatagtttat	gttgccagat	catctgattt	ctttccccta	agtgggaaat	ccagatcaat
	5221	gtacatctct	tgatttgcaa	gtgttggtga	acaaaattca	tattttaaga	tgctgtattc
	5281	agcacaaatt	aaatacactt	atttgctgaa	tactgccagt	ttgtccctct	gcagtagtac
55	5341	catttgaaagt	acagtgtttt	cataatgatt	ctgtgaaatg	actggttctg	tgaatgtaca
	5401	taatttagca	gataacattg	ttaaattatt	aggtttgtat	ttatttaggc	acttgggaaa
	5461	tgcttgtgtg	caattgatta	tagattagga	gcttaaaagc	aagatttata	ttatcaactt

	5521	atttgtgaag	actgggaaac	ccacattttt	aaagttagga	attaagatgg	ccaggttcaa
	5581	ggaaaagggg	gagaagtaac	tttcttatta	ctcaaccatc	ttaaataagag	ttctttaagt
	5641	gtatttttaa	gaggtctcaa	aacttaatct	gaagggacgt	caaatgctgg	acaaattctg
5	5701	tgtatacaac	tcaagtcagc	ccccaatttt	actggtcttt	aaatcatgtc	ctttttacca
	5761	gaagtttgca	tttctaagct	aaactattac	tgtagacta	gatccaaaac	ttaaaaacag
	5821	tttaggtaat	taaaaattaa	ttgaatataa	acgttttact	taaattaatg	gcaaattggct
	5881	ttttggccaa	tttaagttta	tgtaggcagt	taaatcgatt	ttggttaaata	cttttgctgc
	5941	taacaaggta	tttccagatt	ttgaaaagtg	gggtggcctg	gtgcctgtag	taccagcact
10	6001	ttgggaggct	ggggagggtg	gatcacctga	ggtcaggagt	tcgaactag	cctggccgac
	6061	gtgggtgaata	caaaaattag	ccaggcatgg	tggcagggtg	ctgtgatccc	agctgcttgg
	6121	aagtctgaag	catgagaatt	gcttgaacct	gggaagcgga	ggttgacagt	agctgagatc
	6181	acgccactgc	actccagctg	gggcaacaga	gcgagactcc	atctcaagaa	agaaaagtgg
	6241	ggtgttttagt	cttcaaactc	cgtgtttaag	tgactggagt	gaaaatgtaa	atcataggcc
15	6301	ggtgttggtt	taaaaagcat	catctgaaaa	taatgctgta	gtctgcaatt	atttttatta
	6361	cgatacgatg	gtgtaaaaata	caagcagatc	agtgaaccat	tcatgaaaca	ttaatcctaa
	6421	aggcgtctca	ccccaaagtct	atcccacaat	ctccatgaga	cttcgtggaa	ccactgtaaa
	6481	gtttcttgtg	taatatccca	gaagtttctc	acctctggta	tcttttgaac	ttgttgaaaa
	6541	ggcttttcca	ccccctcttt	atgatggttt	gaagagtgtg	aacatctgaa	tgatgctggg
20	6601	gtgaaactgc	ttcataacac	ttccattttc	tcccctattt	atttccatat	ttttattttt
	6661	tcactaatat	ccccacggtt	ttacttctgt	tttagtaatt	cacatgttgc	tggactaatt
	6721	cttttttaact	gacttgtaac	agatatgtta	aaccgtttaa	aacttggggg	gtatttttaa
	6781	cctacttttaa	gttagttcaa	gttaatcagt	ctacatggca	tataaacctt	atgattaata
	6841	aatctttaat	gctggtagct	gagttggaag	ccaaagacgt	acaaaaaagc	tgaagtgtta
25	6901	ggttttagtgt	gataagcttc	tcttactaac	aggggtttgt	aatagcagaa	atagatatat
	6961	gcataatat	gtgcataatat	atagcatacc	ttattggatg	tccatataaa	aatgtgtgaag
	7021	aagttaaatt	tactgcaaaa	tttcttggga	gtgcaatttg	aagatgatct	taagtgggtga
	7081	tagtagtttg	ctacactggg	ggatagttgt	tgcaaaactgc	tcctaatttt	cctttactgt
	7141	gaagtaaact	gaacagctgt	aatagggtt	aggaactgta	ctccctctct	ctctttttta
30	7201	agtataatta	agtggttttg	gggtaagggt	gtaggagtg	agtgtctttg	aagttttgca
	7261	tactactagat	gaatgccaca	tgtataaggg	aggaacaagg	gattcttgga	aatatttttc
	7321	aatccaagta	actttggagg	cttccaagtg	gagttcattc	ccctgtgtag	gaaagtgtctg
	7381	gggtagaccc	ttaaattcct	ttctgagcca	ttgaaagaat	gtcctcaaac	ttcgcttata
	7441	ctttatagtt	catttagata	caaaagttac	aaactgaatg	ctatttagga	aacgtaatac
35	7501	actgacatac	cgctctttaa	atagattata	aatttagtat	atcaattttc	tggcattttg
	7561	ctgaattttta	ttgttttagtt	ttcaagccca	actatcttgt	tactttgtat	atcgtagttg
	7621	tcccccggtg	atcactgttt	cctgtttaat	tgtgctgtcg	tttttcttgg	gtcctgattc
	7681	agagtgtcag	cattctgttc	cccatagaat	aagaagaggc	tagaaagttt	acagatgaga
	7741	tatctaggaa	tgccagaaga	tcagggggtca	ccgttgaggc	agagtaatta	attatggtta
40	7801	aaatggtgtt	gctgataagt	gggtgctggg	aaataattaa	aatttgattt	tttagaagaa
	7861	tacttctcat	gcttgaagag	cgccctcatt	atatgctaaa	gggcctcagg	tttttcttta
	7921	ttgccattat	gctgcagatt	ctattacatt	tgtctgaaaa	gatctaagac	agaagggctg
	7981	tttaataacct	tcccttttct	cctgaacttc	ccctctcctc	tccccatca	ggagctaagt
	8041	aggaacccct	tcaccttggt	accatcagat	ttcatcaatg	gtctgtcttt	acaatgaagg
45	8101	aagtagtact	gcattctggg	cagaggccag	tcctgaggca	tgccttttca	aggacattgt
	8161	tacttttagtt	acactggctc	ttctgtttta	actcttatcc	cccagactct	aatcctgttg
	8221	cttttttttg	tccccatctc	ccacctttca	tcacttgaaa	tccattcatt	gtaacttctg
	8281	gaactcagtc	gttagaaaaat	cctttatatt	ctcaatcttg	tgaatgttcc	tttctttctt
	8341	attccagctg	taacctagcc	ttctcccaa	gaatgctact	tcccttgtag	ctctctcaag
50	8401	tgggtgaattt	ttcccttctt	gcacacctta	taacactgaa	ctaggagggtg	tgtggactaa
	8461	atgtctgctt	ttgttcctta	ttgtcacttc	ttgaccttta	ttttccaaaa	cttcaagctt
	8521	tgactttcat	gtgatcaaat	tataccaccc	actgcctgtc	tttattttcaa	gcacctgcaa
	8581	accttctctg	gtcattcaca	tccttctttg	ttcacttcat	tagctcttgg	ctcattgtca
	8641	ctgtctctta	tttctgtcat	aattcttggg	gacatcagta	tctatgtaga	gcaatactag
	8701	tgaagatgtg	gtctggtaac	tgttacctgt	atgaattaa	ataaggagt	atgccagaat
55	8761	ataagtcacc	tgtgtcacta	agtttactgt	ttagcttact	ttttttgtag	caagattttg
	8821	atgaaggacg	caatatgttg	atttacagtc	tggtaaaaat	tttgatgtag	aagatgcttc
	8881	caatatcctg	gtctcttagt	tccttgattt	cttctccagt	gatcttattt	tctaccctaa

8941 ctcaactaca tattcccatt gtcatatcct agaataatattt gtctttttatc tgtaactctg
 9001 ctctctttccc ccaatctcat ttcaagcatc ccacttttcta attcctctag taaatacgtc
 9061 agttccaaca gcccatcaat cccattggga cctacagttt atctatccaa gctttttccct
 9121 gttcctcacc ctcaacttcta tacagctgaa gtttcatact gaattataat cacttttctcg
 5 9181 tatacacggt taacaatctt gtccctccct ggcttcatgc ccagtgatct cttgtatcta
 9241 tgaccatgtc ctttatcttc tctctgtca ctggatgaac tgtagccttc caagataagg
 9301 ccaactcagtt cttttgtaca gcagattcca tcccctcttg ctctcaagaa tattactgtg
 9361 gtatctctct tttcttgtct ctactggctc tttccatgag caaacatggt attatcccat
 10 9421 tacaaaaaaa attttttctc cgtctctcct tccactcacc acctcagctc ctgcttctct
 9481 ttcccgc aaaataccttga aaaattgctt tatgtactcc cgttttcttt tgaacccctg
 9541 ccagtgacca ccacgttata aattttagt tgtcatctca cttaatctgt tagtagtatt
 9601 tggcaccatt gctacagttg cttgaaatgc cttttcattg gtttccaggc caccatgtct
 9661 gtttagcagct tttcctctta cttcactagc atttcttct tttgtttttc tgttatcttt
 15 9721 ctgacctctg ttggagtggc tgaaggttta gtccttgaat ctttttttgt tgtgcatatt
 9781 tactccagta tcatagcttt atacagatgg tatttacatc tgtttgctaa cgatttccaa
 9841 attggtatcc ttaaactggg atccagctat tttttggtca gcattttgga tgtctaagaa
 9901 gcttctcaaa ctaaaactgac ctcccggttt tccccaaagc tgcactctag tcttttccga
 9961 aatgcaattc tgtctttcca gttacctagc taaaaagctt gcagttcttg actcatcttt
 20 10021 ctctcatacc acgtatctga attctctctg caaaaaattg tctgttctcc cttcagaata
 10081 aagtcacgtg tcaatttatg atggggatag attcagaaat gcgtcattag gagataatca
 10141 tggttgtgtg aacatcagag tatacataga caaacctaga tggatatagc tactacacat
 10201 ctaggctata tgggtgtggc aattactatg atgaatactg taggtaattg taacataaag
 10261 gtaggatatt ttatctaaac gtattgaaac atagaaaaag tacagtaaaa aatatggtat
 25 10321 caaaaaataaa aaatgttaca actgtataag gcagttgtga tgaatggagc ttgcaggata
 10381 tgttgctctg ggtgagtcag cgagtgcga ttgaagggaac gtgaaagatg tgggacatca
 10441 ctgtacacta ctgtagactt tataaaact gtacacttg gctacactac atttttgtaa
 10501 ggttttaaaa gacttttttc tataataaac cttaaattac tgcactttt ttactttatg
 10561 aattcttaat tttttaaacg ttttctact tgaataaca cgtagcttaa aacatacatt
 30 10621 gtacagctgt acaaaaattt tctttatct tttataagct ttttatatt tttaaaatta
 10681 ctttttacct tttagctttt ttgttgaaaa actaagacat gggccaggcg cgggtggctca
 10741 cgcctgtaat cccagcactt tgggaggctg aggcaggcg atcacgaggt caggagataa
 10801 gagaccatcc tggctaacat ggtgaaaccc cgtctctact aaaaaataca aaaattagcc
 10861 gggcgtggtg gcgggcacct gtagtccgag ctacttgga ggctgaggca ggagaatggc
 35 10921 gtgaaccag gaggcggagt ttgcagttag ccgagatag gccactgcac tccagctctg
 10981 gcgacagagc ggaactccg tctcaaaaaa aaacaaacaa aaaactaaga catgaacaca
 11041 ttagcctagg cctacagagg gtcaggatca tcagtatcac tgtatttcca tctccacatc
 11101 ttgtccttct ggaatgtctt cagaggcagt aaacataaat ggagctgcca cctcctgtga
 11161 taacagtgc ttctggaata cctctgaag gacctacctg ttggtgtttt atagttaact
 40 11221 tttttttttt aagaagtaac agaaggatg cactctaag taaaaaagta tagtaagtac
 11281 ataaacctgt aacaatcatt atcattatca agtgtcatgt actggacata actgtatatg
 11341 ctatactttt tttttttgag atggcatctc actctgtcac ccaggctgga gtgcagtggg
 11401 gcgaggatag ctcaactgtaa cctcagactc ctgggtcaa gtgatcctcc tacctcagcc
 11461 tcccaagtag ctgggactac accaggcacc ccacatgcc tggctaatta aaaaaaattt
 45 11521 tttgtagaga cagggtctca ctctgttgcc agggctggcc ttgaattcct ggcatacagt
 11581 aatcctccca ctttggcctc acaaagtgcg aggattacag gtaagagcca ccatgtctgg
 11641 cccactgtac ttttatacaa ctgaagcaca gtaaacctac tgtggtttcg ttacaccag
 11701 catcaccaca aacaccatga gtagaacatt gtgctgcgac gtaaacgat gctacaacat
 11761 cactaggtga taggaatttt tcagctccat tataatctta tgagaccact gttgtatgtg
 11821 cagttcatca tccactgaaa tgtccttatg tgatgcatgt cttcatatcc aaaaatatta
 50 11881 atcattttct actgaagcca tgccatgcca tgccatcttt tgctgtatt attatttttc
 11941 agctttttatt ttagattcag ggtgtacatg tgcaggtttg ttagaaagag tatatcgat
 12001 gatgctgaag tttgggatac agttgaacca gtcaccagc tagtgagcat agtactcaat
 12061 agataacgtt ctaacattac tctccttcc ctccctgttc ttgtctctgt ctattgtatc
 55 12121 tttatgtcca tgtgtaccaa atgttttagc cattcttgtg agaacatgtg gcatttgatt
 12181 ttgtttctgt gttaatttgc ttacaataaa tagtctccag ctgcatccac attgctacaa
 12241 aggacatgat tttgttcttt tttataggct gcacatatt ccattggtga taggtaccac
 12301 attttcttga tccagtctac cgttcatggg catttgggtt gattgtatct ttgctattat

5 12361 ggatggccttt tgcctatatatt attggaaagg ccttctaact ggtgtccctg cttacaccgt
 12421 tttccccctt aaatgtgttt tcaacatggt agccagagta acccttttta taacaataaa
 12481 tcgtgtaact tttttgttca gaaacttaca gggcctacca ttccattcag taaaagctca
 12541 agctcctgta tagtcagacc atatccttca tcacctgtta cttttctcct ctgactcttc
 12601 agcctttttg tttttcctca aactgatgaa gccttcatgg ctgatgtcag atgttttgcc
 12661 cattgagatc ttccttggtg actcagttgc acttggtcat atgattttca tttatttggg
 12721 gtatctaatac ataatactgaa agttggctac ttattttttac ccctttgagg gtccttgccc
 12781 tgtttttgta tccctgatag cgggacagcc agatatctgg aacttacagg tgttcaataa
 12841 agtttttggtt aatgaatatt ctggaatcac ccaacctttt ttttccccct cacttatttt
 10 12901 tcttctccct ttcacggcct gaaagatgtc ctatgtatat ggttccactt atcactctca
 12961 tcccagtttg tgatatacta ttccattata ttactattat taatacaatt ccattgaact
 13021 tgctcttgct gacttcacca ctggacctac atgttggcca aatggatact ttataatttt
 13081 agtcttgacc cctgcctttg gcacatttct tacctctagc acagcactgt ccagtaatcc
 13141 acactttctg agacagtggg aatgttcagt atctgtgctg ttcagttggt agcaaccagc
 15 13201 taccatgcc tattaacat ttgaaatgtg gctgtgtgac tagtggaat tatgttgagg
 13261 agtacagttt tagaaaactcc tgtttttctt acatggcact acatttagta tcacaatcta
 13321 attgtgcaag ccagataggt aggagtcac tttattcctg ttatttaatt ttctcatct
 13381 actatatcca gttcatcaca tcaacagcgc ctgttgtttc tacctcctaa atatttcttt
 20 13441 agtctaacta ctacttgtcc ctagtccac caccatctat cagctggaat attgctatag
 13501 ctgccttaca ggttccctt ctttctgtt ctcttctagt tttttgaatt ttagtcagca
 13561 cgagatttta aaaaactcaaa taagattgtg ttattcacct gcttaaaacc tttcatgact
 13621 ttcagtgta cgtagaacag aaaacacttt tcttaccaaa ggctagagag ctctacgtga
 13681 tctggctatt tttaacgttt cattgcactc acccttttcc tctataatca aactactctg
 25 13741 atctcaaggg ttagttcttg aaagatgatc atgttcttta atgactttag gtttttgtgt
 13801 gttattttct atttctggga tgtttattct ctgttcctta catgctggcc cttttgcac
 13861 cttcttcagg tctcagctta catgttacct tcaagaagcc tttgaccact ctaagtgggc
 13921 ccttccctcc acttctgtctg tgtaatccca ctcccttctc ccacttgta attagttaca
 13981 tacttttttg taattgttta tttggttgct gtctccctct caagaatgca gggaccatgt
 30 14041 ctgcattctg cagtaatcac tactgcacac ccagaatcta ttacagatcc tggcatgtag
 14101 ctgatgcata aatatttgtt gaatgaaagt ctgtacattg tatttatgct attgggtattg
 14161 ctatgacctg aaactaaaag gagttgtgga aaagatttct tatggaacag aaatatccct
 14221 tttgattaat atcacaatct cgtaaatgga gaaaacaaaa aaatatatac tactggagca
 14281 ttcagtgata gttggagatt atgactcatt tattgggtgtg tttttggact cagaacaaaag
 35 14341 atgagggaaat attccttaaa gctctgtatt gaaataacga aaagcagtca cattttaata
 14401 atagaagctt cctagcttac tcttctgtta atcttctttt cctaaatgta agagagcctc
 14461 ataattatga ggcttattac tagagtaagg ctgtcaaagg cagcaaaatg tcttctgtt
 14521 tggaagaata acataaactt gacatgtatg gtgggggaca gaaggtttca aaagtttaag
 14581 aatctgtgtt gtcttaacaa atagatgctt ctcaaggagc ttacgctagt ggttactctg
 40 14641 tccagtcagg gttttttctt ctttaacttg ggttcatttc ctgatggcac acatgaagtt
 14701 tggatcatat gggttgactt tagctatggt ccttagctat ggggagcagc atcagcgacc
 14761 tgtgacatgt aaattaaaaa tacaatgcca gggcccttcc ccagccctc tgatagagaa
 14821 cctcttggcc atctgtattt ttagatgttc caggttagtc tgattaacac ccttggttaa
 14881 gaaccattgg gaggatctga ttgccagttt aaggggacct tcaagcctgt aggtctttat
 45 14941 agttaaaaaa aaaaaaagat tttaaaaatc atgcatatgt tgtggctgaa ttctggttta
 15001 gcacatactg cttttaatgg cctgaaatgt ttttcccaa taaattgtct tgttatagct
 15061 ttcagtggtg atttggtcca gcttctgtt ttgaagatac ttacggggg gaacactttg
 15121 tgatttctct tagtaacata ttaaccact taaaaaccct ttctattaca ggtcttcaca
 15181 tttaggctta atgtgcttaa ttcaaagtta aaaatacacc tgcctttgtt ctcagtgaag
 50 15241 gtatgtaata aataaatgag gggttggcaa actactgcc accatctgtt ttttatggc
 15301 ctatgaacta agaatcggtt tggatagcta aaaaaaaaa tcaaaaggat aattattttg
 15361 tgacgtgaaa attatatgaa attcaaattt cagtttctgt gaatgaagtt ttaatggaac
 15421 acagccatcc atgcttatgt aagtgtgcat attctctggc tgttttctact gcaatagcag
 15481 agttgagtag ttgtgacaaa gagtttatgg ccacacaaac ctaaaatatt tactttctga
 55 15541 tgctttacag aaaaagtttc ctgaacctta ttctagctat atgttgttca taaatgaatc
 15601 tttcgtggtt ctgaaggcat ttaagaatct cttaggttat aaattggctg ggcgcagtgg
 15661 ctacgcctg taatcccagc actttgggag gccgaggctg gtggatcacg agggcaggag
 15721 ttcaagatca gcctagccaa gatggtgaaa ccctgtctcc attaaaaaaa aaaaaaaaaa

5 15781 aaaaaaaata gctgggggttg gtgggtgggca gtaatcccag ctactcggga ggctgaggca
 15841 gagaattgct taaacccagg aggcggaggga tgcagtgagc caagatcgcg ccactgcact
 15901 ccagcctggg caacagagtg aaacaccatc tcaaaaaaaaa aaaaaaaaaa aaaacactct
 15961 taggttataa ataattgttg ttagctctcc aagcctccat attacatttt gtgtgttctc
 16021 ctgttcacat tttgagcatt ttatTTTTTT ttagcacatt cagttcatca ggtattttaag
 16081 agcttaatat atgccaaaagc atatattaag cgagaagctg tttctaaatg tactgtctca
 16141 gccctcacag agttcacttc attaggtctt ttaaaatttc tttcttttaa aggtcagcgt
 16201 gctggtatag tggggaaggg aaactcttac aacacgtcga gtagaggag gttatcatta
 16261 tgggatataa tttggaagtc attgagtacc tgccattaat tctgcctgta gtctgaatgt
 10 16321 agagttatac atgtagaaac ttttttgaaa taaaatcttc aatttctttg gcatacttag
 16381 tactgtctag ctaggcatat agtcaaagta tgggtgtatat ttcaagtatt aaaagtTTTT
 16441 ttgggctgta gtcactgttg aaaggatata gttctttact attacatgtg atacctttat
 16501 ataaaattgg ctaacccctg tctttcattt atctgcaaca ctgactgtta ccagttgtct
 16561 ctaactttgg tatggggggg ggaatatga ttagattgaa aggttacatg actgagccac
 15 16621 aagcagacct ggatttgaat tttaactgaa cggtttatta gctattctta cattaatact
 16681 gctaatacagt tttcttgtga tatgaggaat gatgtcttct ttatgaggtt gctaggaaga
 16741 ttcaatgaga taacatacta ggctcagaac tgaagttgct aggaatttaa ttatgtctacc
 16801 ttgttaaagt atgtcaaagg cagaattcag tgtttagctg ataccacaag gcagtatcct
 16861 aaaattatgc tgtaaaagat ataaagatgc tgtaagtga tcagaaacct agtgactttg
 20 16921 taatgcagtt gattcttaga atactgtcac tttaacagaa taggagctag gaatgaagaa
 16981 atagtattta aattactaaa atagaaaatt tattgacaca tgtaaatgta catttgctta
 17041 aatattgaaa aatttgtagt actatttcct tgcttttagaa aacattgggt accactTTTT
 17101 ttatttatag cagtttgttt ttgccttgag gcaagatggg tgactgagta gttgccacat
 17161 ttcttttgta caaagtccat ttcataggcc atctagcttt tatgtctaga aacatttcct
 25 17221 taacgttata tttcagtatt tggctaacct atataggggt aaattatata ggctaacttc
 17281 tcggacagat atttctaata atttatgtat ttggttctgc aaatgtatgc aaaaatatat
 17341 gtacaaagggt atgcagatgc cttgcatact tgatatatgt taaatttttt ttaatgtaga
 17401 cctttttcgt tctctttaat gactatatgg tattccacca tccccgctc acctggacaa
 30 17461 ctacagtaac ctctaaatg gtgtttctac ttgtctattg ccccttattg tcttttttcc
 17521 cctttatagc tgctggagtg aatttttagaa agcctaagtc atacatcaca ttgcttcatg
 17581 ggcattcccag tacactttgg attttatttt acatccttac tgatctgatt ctcatctctg
 17641 tctcttcatg gttctctgcc ttctagttag actggtgacc tttcaaaacc tttaccacat
 17701 tgagttcatt ccttactttt cactctttct ctgcctggag tgttctgccc catctttacg
 35 17761 tggccagctg ctctcctct gatgaaatgt ctcttcctca caggccttcc ctgaccaccc
 17821 actagagtag cacatcttct acctcataaa cttgtttatt agtatttctt actctaaatt
 17881 ttctttttaa ttgcttaatt ccctaacagt agaataaag cttcactgta tgatgatct
 17941 tgttgactct cttactcatt gttattgtaa taccagtaac aaagggtgtt taaaatttgt
 18001 tcagtggttg aatataatgt ccatttaagt gataaattat tttttattca gtctcctgtt
 40 18061 gatggacatt tgaataattt ccattttttt ctctatgaat gcctcacttg gactgttct
 18121 gacagtattg ccacagaata ctttctgttt ataaaaattg aatttttaag tcaaagggtg
 18181 gttacacttt aatggatagt ggcagcttac tatcaaaagt ttctgctagt ttcaccatat
 18241 ccttatttagc agtagatatt atcaatcttt tcaatctttg ccaatctgat aagcaaaaag
 18301 taaatgggtt taaacatcct ttgtatatat tcattgctca ctttatgttt ttcttttgaa
 45 18361 atgttatttc ttgttctttc cctgcagtat gattctttct ttttttgact tgttcccagt
 18421 tttttgtgta ctatggatat tagcctttaa ttatgttacg gatgttctag tatgttattt
 18481 tttgaattac ttcaaatgtg atttgttgct cagattttta aaactacata cacaaattat
 18541 ctcatgtttc ctttttttgt ttcaatttcg actcatgctt aatcagttca tcgattgggc
 18601 atggttttat tcttaatata taccggtatt ttatctcatt ttattttttt acgtgtaaat
 18661 atttggtgaa tataggttta attttaatgt aaaaataagga tgaaaaatga tagttggaat
 50 18721 tacaagcca tttctcctaa tacttttaat caagtaatcc actaattgaa atattacctt
 18781 cttcattttat gaaattgcc aattatatct ggggtgtttt ctgcctacta cagtctctta
 18841 cccatttctt tcctaataat acaatacttg aattgctgtg gttgttgatt tataatgtta
 18901 tcttaatgat aacattataa atgtgatgga actggttcct ccttatagtt cttcttaaat
 18961 caagaacaag acatatcttc ccatttactc tcgtatgtat ctcattttac tgttatgaat
 55 19021 gaaatctgtc ctatttgtgt ataggaaaat agtttttgta tgtaattgtg atatggccag
 19081 ttttattaaa aatttgggtt aactaagagt tgttttctgt tcagccttat catactataa
 19141 aatccacata aaatgggtat aaaagtgtcg caggacactg ggctcagatg attctcccac

5 19201 ctcagcttcc caagtagctg ggactacagc ggcatatgcc accacaccca gccaatTTTT
 19261 aaataagttt taaaaatagt attttttagta gagacagggg ttcaccatgt tgcccaggct
 19321 ggtcttgaac tcctggactc agacaatcca cctgccttgg cttcccaaag tgttgggatt
 19381 acaggtgtga gccaccacac cttgccgaat tgcagccata tttataactt tttccatcc
 19441 tattcccttt gctgccccca ggccctcctgt attgatagcc cgctattaag aagctagtgt
 19501 atattctttg catactttta cttcataaac tatatgaagc attgttctgt tttttaactt
 19561 aattggtata aaattatatt ttggaaattc agtatattct gtgaaaatta tttagaaaat
 19621 gtgcctctga gataaagcct attcaggatg tatcttaaag gagatagctg tgcTTTaaCa
 19681 ttatcagtct ttttggctgc ttatgttaat ataagttgga gaaaaacagt ctgctTTTTg
 10 19741 tgataatatg ttcttggaga tggagtgaag gattgtttta aaacattgtc ttttttttcc
 19801 cctgaagtac cagtatttat tttaggatta tgttactgat caaagatgct gtgtggagtt
 19861 actcattggt gagactaaca ataaatcaca catgcaaagg atgttaccat aatctaatta
 19921 ttttaaacag taaaattata ttctaagaca tccagttggc ctatatgtgc tatatcaatg
 19981 actatcaagg ggctTTTTat gtatactgta tacatgtact tcacaaaaat ataaaaggat
 15 20041 gacatcaaaa atctggcaag ccaaaagcct acattacatg tagcaataa ataagcatat
 20101 gaacttattg gaatttaaaa ccctgtagga tgggcgggtg atggtatgta tgttagatgt
 20161 gtggacatat ctattaaaag ttgtgtcaga taacagctgg tgctgacaag cccttggtaa
 20221 gatggcagca tgttcaatat gttctgtgaa aattatctca gtttatgatt ttcagtgatt
 20281 gtggagctat gcatgaaagg acttaaaatt cttaccctta aactcagtaa cagtgtttct
 20 20341 agaacttctg gtgatatggg aaattaagag aattatttat atgcaaagggt gtttattgca
 20401 gcattgttgg aataatagac aaaatgggga agaacaagct cagaatggag gaggtagctt
 20461 atagtataga catacgatac aatccagatg ataatatTTT ataatagtct tcacaaggaa
 20521 ttttatattt ttatttttaa aaatacatag cagtgaagttt aatataccaa acataccaaa
 20581 atgtcatcat ttactgtgtg gtggactcat atgatggaga tgataaataa aaatattaat
 25 20641 ttatttgagg catatatTTT tggctgagga aggaagacag ttatgaagaa cagctcattc
 20701 tggaaacata ctaatttttc ccagccataa agagatttcc tatttctttt ttttttccat
 20761 ttaccttctg tttcctacct gagaagattt catacttcta ataaccattt gtgtacctat
 20821 ttaaagacag taccaaaggc atacatttta gtgtttggag gaccaagggt catttgatgt
 20881 ttgatgctta ttgactattc gaggatgaca agacaccttg agaacacaca caccacacc
 30 20941 cacaccaca cctcaccaca cccacccac cccctcccc gaagaaagct gtgaaggaag
 21001 aaagcagaaa agaacctgga gtgagttgta acttaaaatg ttagtggtgc atgaagtgtg
 21061 ttaaaacagg aagatttgag gaaattgcat acattttcta gatggcaaag tattactggt
 21121 gacagttaat gaaaatgcat atgcatgtgt ttttagattt acaaatttta ctaagaactt
 21181 tttaaaaatc cctgaagggt tatcaaaagt ttatcatgct tatgaaatag agtagcattt
 35 21241 tctaacttta aaacggggaa taattctttg gatcttgatt attggaaaag tgaattatga
 21301 attgctagta taaaactgtg gttttaaaat atgtctgctt tatattttta ttagcagat
 21361 ttactcctag ttaataatac tcaaacttac tgaaaactaa ggtaattaag ataattctgt
 21421 cctgatggga agaggaaaaa taacttcagt gtgaaatcta ttatatatta ttatggcaa
 21481 gatttctccc attgactttg actggagaca tttatagggt taaaatcgga aatagcacgg
 40 21541 tgaattttga agtatccttg tagttggaaa gagtattatg ttcatattgc caaaaaaag
 21601 atgcatggat gcattagact ggatggaaaa tacatgagaa gttggctagc cccctcttg
 21661 tcaaaacatc acttggtggt gataaagctg ttggaaaaca cagcattcta atgtagtctg
 21721 tagtttaatg ataactgtg tcttgaaaca tttagcgtag tacttataca aacctagatg
 21781 gcatagtgtg ctgcatgcct agcctatata gtatagcctg ttgcttctag ggtgtaaagc
 45 21841 tgtatagcgt gttactatag gcagttgaaa cagtggattt tatgtatcct tttttttt
 21901 tttaaattct tttagagac agggctcttg tctgttgccc aggctggatg cattggtgtg
 21961 atcatagctc actataacct tgaactccta agtgatcctc tttgcctcag cctccccagt
 22021 ggctaggact acaggcacat actaccacac ctggctaatt tttaacattt tttgtagag
 22081 attgaatttc gctgtgttgt ccaggctggt cttggaactc ttgtgctgca gcaatccacc
 50 22141 cgctcccaa agtgttagaa ttacaagcca cttgccttg cttgtttacc taaacataga
 22201 aaagatccag taaaaatata gaattaaaat cttgtggggc cactgtagca tatgtagtcc
 22261 atcttgactg aaatgtcctt atgcagtgc tgattgtact tcataatttt taagcactcc
 22321 tccctcttga ttggtactta gtggatttta tcatttttgt ttcttcataa ttctttctga
 22381 aatgtctact ggttgacact ttgatctcct gaattgatcg tgatttcttc tgttgtattt
 55 22441 tttgtctttg tcattttttt gtactctagg cagttttctc aattttagtt tctattcaac
 22501 tttttgtttt tattttattct ctccagtatt tatggagata ctaaaattgaa gtgttctggt
 22561 tctctctcca ccctatccct agtttcaagt tttatctcag tttctatgga gtcagttttt

	22621	tcgttgccttt	aaaaaaaaaat	tttcctgaag	tgattggtaa	gttttggtta	attgggagca
	22681	ctagaattgg	gcccttaatg	gttggcaggg	tgtggtggag	gagagacagc	ccttagtcca
	22741	aaggctcagg	ccagaaaaag	aaagaggaag	gctttccttt	tcctttccgg	agcagggttc
5	22801	tgccctaggt	cttgcttggc	agtctatttg	atctcttttag	cagttaatgc	tcagtttttt
	22861	ggcatatgtg	gatctgcctc	cagagcaggt	acaagggtgag	tgagtctatg	ctgttaccta
	22921	attagatccc	catttctacc	ctttgttttt	acttctctat	ctactgatag	gtttttaccc
	22981	tccttcacct	catagggttg	cagtgaagag	caagatgaat	ttttatttat	gttgcataaa
	23041	ttttaaaagc	taaaaaatat	atatgtaatg	ttgggaagtc	ccagtgtaca	aatggctatt
10	23101	gtaaatttgg	aacatgaact	tgcttttttc	cattgtaaaa	atgaaatcat	tataaattgc
	23161	ggccaagtta	ctagggtcagc	ccacacagag	tttaccaggt	aatatgcgta	aatgttttgc
	23221	ctttgcatca	acaacaagga	aaaacagtac	tataaaaaaa	tgttcctgga	agccggatgt
	23281	atcaaagcac	ttctgaaata	gctatatagc	ctatagacat	gaccagttgg	tttctgagtc
	23341	tgttgacatt	ggccaaagga	gaagctcagt	gtagaacatg	tttgaggtct	ccttttgcag
	23401	aaatacattg	gaggctggag	tggggaacca	atttttcaga	aagggtggtga	agtagttaca
15	23461	tagccactct	tttaaagaca	gtcaaaaagat	agaaaactaag	gccagggtgtt	ggctcacatc
	23521	tgagatagga	aaatcacttg	aacctgggag	gcgagggttg	cagtgaagccc	agtatgcacc
	23581	tctgcactcc	agcctggttt	ggcaagagac	caaaaactctg	tctcaaaaaa	aaacaaaaa
	23641	tagttcacac	ttaaatattt	tattccatat	ctttacatac	ccaatatggt	aatttatagt
	23701	tcaagatgaa	cttgtttggg	acagattttg	taataaaagga	aatcgtgtta	ttagaaatat
20	23761	ctagaggcca	tgagccctta	aactgttcta	atgtgcaagt	agttccctgt	gtgatgcagt
	23821	ttttttcaat	attgcacaat	aaaggcaaaa	tacggacaaa	ttagatgata	agattttatat
	23881	aaatttttta	aatattgatc	aaaatatgta	tccatatttg	taatatattg	atttataata
	23941	aatcattgct	gtaaaattga	acttagaaaa	attttactaa	taaagggtgct	tttggtgtgc
	24001	aaactttcat	ttgaaaagta	atttttcttt	gtaccaaaaa	atctaaaaatt	cgctattcta
25	24061	gtcaccaaaa	tttgctttat	gaaaaataat	ttttgatggc	actatatcag	aaaacaactt
	24121	gttaaagaaa	atgtggagtt	tttaaaatcc	cactgtacct	ctgttatcca	aaggggatct
	24181	gtgaattttt	ctgtgaaaag	ttaaaaaagg	agagaccttt	aggaattcag	agagcagctg
	24241	atttttgaat	agtgttttcc	cctccctggc	ttttattatt	acaactctgt	gctttttcat
	24301	caccatcctg	aatatctata	attaatatatt	atactattaa	taaaaagaca	tttttggtta
30	24361	ggaggagtgt	tcactgaagt	tcagcagtg	tgaggctgtg	gttgaggtgt	ctggaggaga
	24421	ccatgaggtc	tgctgtttcac	taacctggta	aaagaggata	tggtgttttt	ttgtgggtgt
	24481	aatagtgcac	tttaacaggt	atcccagtg	cttaggagta	ttaatcaagc	taaatttaaa
	24541	tcctaattgac	ttttgattaa	cttttttttag	ggtatttgaa	gtataccata	caactgtttt
	24601	gaaaatccag	cgtggacaat	ggctactcaa	ggtttgtgtc	attaaatctt	tagttactga
35	24661	attggggctc	tgcttcgttg	ccattaagcc	agtctggctg	agatccccct	gctttcctct
	24721	ctccctgctt	acttgtcagg	ctaccttttg	ctccattttc	tgctcactcc	tcctaattggc
	24781	ttggtgaaat	agcaaaacaag	ccaccagcag	gaatctagtc	tggtgactgc	cttctggagc
	24841	ctggatgcag	taccattctt	ccactgattc	agtgagtaac	tgtaggtgg	ttccctaagg
	24901	gattaggtat	ttcatcactg	agctaacctt	ggctatcatt	ctgcttttct	tggtgtctct
40	24961	tcagatttga	ctttattttct	aaaaatattt	caatgggtca	tatcacagat	tctttttttt
	25021	taaattaaag	taacattttcc	aatctactaa	tgctaatact	gtttcgtatt	tatagctgat
	25081	ttgatggagt	tggaacatggc	catggaacca	gacagaaaag	cggctgttag	tcactggcag
	25141	caacagtctt	acctggactc	tggaatccat	tctggtgcca	ctaccacagc	tccttctctg
45	25201	agtggtaaag	gcaatcctga	ggaagaggat	gtggataacct	cccaagtcct	gtatgagtgg
	25261	gaacagggat	tttctcagtc	cttactcaaa	gaacaagtag	ctggtaagag	tattattttt
	25321	cattgcctta	ctgaaagtca	gaatgcagtt	ttgagaacta	aaaagttagt	gtataatagt
	25381	ttaaataaaa	tgttgtgggtg	aagaaaagag	agtaatagca	atgtcacttt	taccatttag
	25441	gatagcaaat	acttaggtaa	atgctgaact	gtggatagtg	agtgttgaat	taaccttttc
	25501	cagatattga	tggaacagtat	gcaatgactc	gagctcagag	ggtacgagct	gctatgttcc
50	25561	ctgagacatt	agatgagggc	atgcagatcc	catctacaca	gtttgatgct	gctcatccca
	25621	ctaattgtcca	gcgtttggct	gaaccatcac	agatgctgaa	acatgcagtt	gtaaaactga
	25681	ttaaactatca	agatgatgca	gaacttgcca	cacgtgcaat	ccctgaactg	acaaaactgc
	25741	taaatgacga	ggaccaggtg	agcaatgaca	tagctagctt	tttagtctgc	tttgaagtaa
	25801	atgctcaagg	ggagtgtttt	cagaatgtct	acccaataacc	agtacttgaa	aactaacgat
55	25861	gtttctgaat	tcctgtatta	cagggtgggtg	ttaataaggc	tgcatgttatg	gtccatcagc
	25921	tttctaaaaa	ggaagcttcc	agacacgcta	tcatgcgttc	tcctcagatg	gtgtctgcta
	25981	ttgtacgtac	catgcagaat	acaaatgatg	tagaaaacagc	tcgttggtacc	gctgggacct

	26041	tgcataacct	ttcccatcat	cgtgagggct	tactggccat	ctttaagtct	ggaggcattc
	26101	ctgccctggt	gaaaatgctt	gggtaagaaa	acatgtcaga	atgcttgaag	ctaaaaagta
	26161	gaagagtata	ctcacaatat	ttctgatgag	gcttttttct	tcttcccagt	tcaccagtgg
5	26221	attctgtgtt	gttttatgcc	attacaactc	tccacaacct	tttattacat	caagaaggag
	26281	ctaaaatggc	agtgcgttta	gctggtgggc	tgcagaaaaat	ggttgccctg	ctcaacaaaa
	26341	caaatgttaa	attcttggct	attacgacag	actgccttca	aatttttagct	tatggcaacc
	26401	aagaaagcaa	ggtaagagaa	ttattcttta	tgtggttttc	atggagcatt	ggacacctcc
	26461	agtgtcatgt	cattccatgc	agtgttcccta	accttttttg	caccagggac	cagtttcgtg
10	26521	gaaaacagtt	tttccatgaa	tgggttggtg	gaatggtttc	tggatgacac	cattccacct
	26581	cagataatca	ggcattagat	tctcataggg	agcgtgcagc	ctagatccct	cgcagtgtga
	26641	gtccacacta	gggtttctac	tcctatgaga	ctctcatggt	gcagttgatc	tgacaggagg
	26701	tagagctcaa	gccaggtaat	gctcgctcac	ctgccactta	cctcctgctg	tgacagcccag
	26761	ttcattttctg	ttctttttaa	tttttgagtt	tccatatgta	aagcactatg	cgaagtagta
15	26821	gggatattgt	aggcaagcct	ctcttcacac	ttttgttctt	agggtgggatg	tagatgttgg
	26881	gaataataac	ctaataattta	atthgtgtag	tgggaagaag	tggggctatg	agggcacata
	26941	acacaagttg	aaactgactc	tttttgaggg	ttcaaggaga	cctcttgagg	gaagtgatag
	27001	ttgagttcag	tgttcaagga	tgagaaggga	ttcactaggt	gaaggttagg	tgagaaaaca
	27061	acattctttga	aacgaaggaa	ggagatggaa	agttttggga	atttaagaaa	tactaatagt
20	27121	aaggaggaag	aaaggtttga	ggtgaggcta	ttgagataga	cttagcagat	ctcatagggc
	27181	tttgtagagc	atgttttaaaa	gcacaatggg	aaatttcagc	agaagcctga	aatgatgaaa
	27241	tttgttttta	gaaaattggg	gcagtgttga	aagggaagat	atacagggaa	tgaaaggaca
	27301	agcatgaatg	atcattttat	ggtatctgtt	tttaagggtg	atataattag	gaaaattaaa
	27361	gggccaaatg	atgaggagtt	aagtgccagt	tctggttcaa	attttcagtg	aatcagtttt
25	27421	gatataactt	tcatcttagg	gcattactct	tgccatacaa	catagtttct	aaattttttt
	27481	cttttggtgt	gatcactgtg	ggaagaagga	aattgggccc	aaactgatac	attgttttga
	27541	ggactgggat	gtctgaattt	gagtggaaatg	ctttaaaagg	acaagttgga	tagggcccca
	27601	gtatgggggt	ctgagtgatg	gggtccagga	atacatttag	gtccaatggc	aagctggctg
	27661	aaattcttgt	ataataaaa	aggttggtaa	tatggctctt	ctcagacatg	tgatcaagat
30	27721	tccttgacta	acaagatata	tatatataatc	tttctagctc	atcatactgg	ctagtgttgg
	27781	tccccaagct	ttagttaaata	taatgaggac	ctataacttac	gaaaaactac	tgtggaccac
	27841	aagcagagtg	ctgaagggtgc	tatctgtctg	ctctagtaat	aagccggcta	ttgtagaagc
	27901	tggtaagtat	atgtatctat	tctgagtctt	gtgtatagca	tctgcagttc	taattagatt
	27961	acttttctta	ggaaaagggtg	gtagaacttt	aactactgaa	aataaatggt	cctattcagt
35	28021	ttgcagccaa	gatttacatt	cagagtacct	gtcatctgga	ttgtagctaa	atatttaagg
	28081	ctagtttagg	tagagttcct	attatccatc	aaaaatgatg	gcataatggt	tgcttaataa
	28141	aatttgtttg	taatttcagt	tttgagtaaa	cctaagattt	gctaacagag	ctgtgaattt
	28201	ataggagaaa	agacaaattc	taatatagta	cagttttatg	taaagtattg	gctttattag
	28261	tagatgtcta	tgagcagttt	ttgttttgtt	tttaactttta	gggtccgggt	aatgtgcagg
40	28321	cttggtatata	aggtaaatg	catgtcacag	gggtttcgtg	tgagatttat	tttgtcacc
	28381	aggcagtaag	tattgtaccc	aataggtagt	ttttcagttc	tttacctccc	accgtaagt
	28441	aggccccagt	gtctgttgtt	cccttctttg	tgcccgtgtg	tactcagttg	ttacctccca
	28501	cttataaagt	agaacatgtg	gtatttggtt	ttctattcct	atgttagttt	gcttaggata
	28561	atggcctcca	gtcccatcca	tggtgctgag	gaagacatct	tggtattttt	ttatggctgc
45	28621	ttagtattcc	atagtatata	tgtaccacat	tttctttatc	tagtctacca	ttgatgggca
	28681	tttaggttaa	ttccatatct	ttgctattgt	gaataatgct	gcagtgaaca	tatgcatgca
	28741	tgtgtcttta	tggtaaaaag	atttcttttt	ctttgggcat	atacctaata	ataggattgc
	28801	tggattgaat	ggtaattctg	tcaggttttt	tgagaaatca	ccaaattgct	ttccacaatg
	28861	gctgaactaa	tttactttcc	caccagcagt	gtataagcat	tctcttttct	cagcaacctc
50	28921	accagcatct	gtcatttttt	gactttttat	tagtagccat	tctaactggt	gtgagacggt
	28981	atctcattgt	ggttttgatt	tgcattttct	taatgatcag	tgatgtcgag	cttttcttca
	29041	tatgtttctt	ggccacttgt	atgtcttctt	ttgaaaagtg	tctgttcag	tcctttgccc
	29101	actttttaat	ggggttggtc	ttttttgctt	gttaatttaa	gtttattgta	aactctggat
	29161	attagacctt	tgtcagatgc	atagtttgcc	agtactttct	cccatgccag	tactttctcc
55	29221	cattctgtag	gttgctgtgt	tactctgttg	atthcttttg	ctgcgcagaa	gctctttata
	29281	ctgtcccat	tgtcagtttt	tgtttttgtt	gcaacttctc	ttggcatctt	cgtcatgaaa
	29341	tctttgccag	gtcttatgtc	cagaatggta	tttccatagg	tatcttgtag	agtttttaca
	29401	gttttaagtt	ttatatttaa	gtctttaatc	cattctgagt	tgatttttgt	acatcatgta

	29461	aggatggggt	gcagtttcaa	tcttgatgt	ggctagccag	ttatcccagc	accattttatt
	29521	gaataggagg	tcctttcccc	attgcttggt	tttgtttact	tgtaggtgt	gcggcctaac
	29581	ttctgggctt	tcttttctgt	tccattgggtc	tctgtgtctg	tttgatatacc	agtaccatgc
5	29641	tgtgattgta	accttggtatt	aacagtatag	cttgaagttg	ggtaaagtga	ttcctccagt
	29701	tttgttcttt	ttgcttagga	ttgccttggc	tattcaggct	cttttttggg	ttcatatgaa
	29761	tttttaaata	gttttttttt	aattatgtga	agaatgccat	tggtagtttg	gtaggaatag
	29821	cattgaatct	gtgaattgct	ttgggctgta	tggccatttt	aacaatattg	attcttctctg
	29881	ccatgaaata	gaatgttttt	tcatttggtg	gtgtcatctc	tgatttcttt	gagcagtgtt
10	29941	ttttgtaatt	ctcattgtag	agatctttca	cctccctggg	tagttgtatt	cctagggtatt
	30001	ttattctttt	tgtggctttg	gtaaatggga	ttgcattctt	gatttggctt	gcagcttgga
	30061	tgttggttgg	gtctagaaat	gcttctgact	tttgtacatt	gatttttata	tcctgaaact
	30121	ttgctgaagt	ttattggatc	aaggagcttt	tgggcagaga	ttatgggggt	ttctaggtat
	30181	agaatcatat	tgtttgcaaa	cagacttcct	atttggatgc	attttctttc	tcttgctga
15	30241	ttatgagcag	tgttttgccc	tgatattctg	tattctcagt	gaatagatgt	cgtctaagta
	30301	tgagaaacaa	tttttttcta	ttctgagtat	ttttaagaag	gcaacttata	tgtggtactt
	30361	tgtatatatt	gtatggtggc	aattggggaa	aagaatagat	ggtttgtact	agggcctctt
	30421	gggttctgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtcat	gaaaacagtt	acttttttagc
	30481	taccaagcat	tttttctcct	ttcagtaacc	cacctaacaa	catttactca	gaatttcaaa
20	30541	gcaagcttca	aatcagtatt	gaaagaagga	aaaataataa	ggcatttaat	ggaagaaaat
	30601	gttgggaata	aagtataggg	ctggcaacac	ttacttttct	cacttattga	gagtaatttt
	30661	acttgggaat	ttatgagaga	gaaagacatt	atgattgtct	caggtaacta	ctggcagagg
	30721	aaccatagtc	ttggggatag	acaaatgtgg	ctgagttcat	atagaatgag	gggatgggat
	30781	gtaaattctg	tcagctgttc	cagcagtaac	ctgtaatgta	ggctaaaaat	acagattttg
25	30841	agatttattt	aatcagaatc	cctggagtgt	taatttttat	atcaagatct	catagtgttt
	30901	tatttgaaagt	gacagggagg	tctgtagata	gctggacatg	tatgggactg	gaagcttagg
	30961	aatctttaag	ttcttccagg	ttattcttat	gttcatttgt	ttattctgaa	aatagcatct
	31021	aatgtatttt	aagaaatgga	ataggcacat	agtatacat	gggtaacaca	acagataggg
	31081	tccccgtgct	taattcttag	tcttgtgaag	gtgacaaaaa	tacttaaaaa	tatgtgatcc
30	31141	taaattagaa	tgagtgttat	gggagaaatg	acagcaaata	gtgatgagaa	ttaatgggga
	31201	ggggaattgt	ctagatgaga	gggaaaaggt	ctccttgaaa	aggggatgtt	aagtgggact
	31261	gcaggatgag	agggaaaccgt	ctcttgtcta	tatgagaagt	gaggggttaa	cgttttccag
	31321	gtagagaaaa	ggaacacccat	gtgctatgtc	ttagaaccag	ggatatccag	tcttttggct
	31381	tccctgggcc	acattggaag	aagaataatt	gtcttgggct	acacacccaa	tacactaatg
35	31441	atagctgatg	agctaaaaca	aaaaaaaaatt	gcaaaaagaat	ctcataatgt	ttaagaaagt
	31501	ttacgaattt	gtgttgggct	acattcagag	ctgtcctagg	ccatgtggcc	catgggctgc
	31561	aggttggaca	agcttgcctt	agaaggaaaag	agattgggtca	ggcacgggtg	ctcacgctg
	31621	taattccagc	actttgggag	gctgaggtgg	gcggatcatg	aggtcaggag	atcgagacca
	31681	gcctggctaa	cacagtgaag	ccccatctct	actaaaaata	caaaaagtta	gccgggctgt
40	31741	gtggcaggcg	cctgtagtcc	cagctacttg	ggaggctgag	gcaggagaat	ggtgtgaacc
	31801	cgggagcg	agcttgagct	gagctgagat	agcgccactg	cacttcagcc	tgggcgacag
	31861	agtgagactc	tatctcaaaa	aaaaaaaaaa	gggaaagaga	ttgtggagat	ccaggtgctg
	31921	aagagaaggt	ctgcataaac	agaacttagt	aatgaggtgg	atggcctggg	atgaggttga
	31981	ggttaggtaa	gcagagccat	aacatgcagg	actttctagg	ttcctataag	atagttacta
45	32041	ctcatggagt	ttattcatgc	tttattccag	ctttggagcc	atagatacag	aatactttgg
	32101	tcagtttgga	aggctaggtg	ggatccaaat	tctaaacggt	tcctcagggt	tatactaaag
	32161	tattttctatt	atcttaaaag	gatgctgaga	cactttcgtat	ggttgtttat	caatagcaaa
	32221	gcatcacagt	ggtgtgttta	aaatattaat	aatagcattg	tatagattaa	cagtttgaat
	32281	gaccaaagc	tagaagacca	gactactgag	atgttacagg	cttttaggaa	tgaatatagt
	32341	tgctttttaga	actcaatagc	aaagggcaga	tgtctgagat	gcctgaaaga	atcatagaat
50	32401	gtaataatat	aggagctaag	ggagcaacca	aaaacggttt	gtggagggga	caacattggt
	32461	accatgaaga	taaatggaac	cctcagaagg	catccttaat	ttttgaacat	aataatttaa
	32521	gaagctgact	taaagtgact	taaaagggtca	gtaggtagct	ggaaatgtat	gatactagaa
	32581	tgcaagagag	gcaggctaga	gatttgggaag	tttccctctt	agtatatagg	ggtaagggca
	32641	gcagggaagg	ggaggtagag	gtgccacaga	gtcatctgta	tgggactttt	ttttttacc
55	32701	tagaactgct	gaatcagaat	gtgtgtgttt	taaagtctct	gtaggccatt	ctgatggaca
	32761	tctgggggtta	aaatccattc	tcttagagtt	aatagttatg	taaagggagg	gaatgaagtc
	32821	ttaaagaggg	gaaagaaggt	agtcatttca	caaatactga	gcatcctgat	catcagtcct

5 32881 acgcagatca ttctattagt agctggagct actatgaaaa aggaacccaa cagagggtgat
 32941 ctttgtcttg tagggaaagt ggagtaactt acactatgaa ggagaagtgc agggtagcat
 33001 aagaattaca gcagatagac ctcatctgag gaaataaaac agacccgaaa gatgaaggag
 33061 acaaggaaaa gtatctctta ctgcattcag aagtgtattta agttgaagat ggatgagcga
 33121 agttaatcta ctatgtgggc attgggcttc catttatact cctttgccag agtaaatgtc
 33181 ccccatTTaa gggtcctaaa ggatggaaga ttgtaaacct tggaacacat gttttgtagt
 33241 cagtgaattg tataaagtcc ctgacagtaa gtgttttcat gccgtctttc tggattgttc
 33301 ttaccccagg aatttaccta gcttcttttag gtcttttagtc agatgtcacc ttcacagtga
 10 33361 ggtgacctaa ttatctatTT aaaatcgag cccactcca ttatttttct ccatagccct
 33421 ttaatatcat ctgacatact gtatggTTTT agtttattgt atatttttct gcctcttcca
 33481 actagatcat aaattctgag ggtaggaact tctgaatatt tttgttcact ggtctatctg
 33541 cagctcagaa caggacctgg tactgaataa atatttttga aatgattgaa tggatgaaaa
 33601 gaaatgagta ataagaatat tacctaaggg ggacagtga gataacaaag gctttttcgg
 33661 cttaggaaag gaacagtagc tatttgagag tttgtcacta gtgagggtgaa ctggc aaaagt
 15 33721 gaaggaaaact gagcaacatt ctagaaaatg agaggaaaatc aaatacttag gtgaaaggaa
 33781 gtaaaactctg gaaatacaga aggacacctc ctaaggctag aacagatatt taggattgat
 33841 aggcacttct agctaattgac tagggcctta tatccttttt aattttctag gtggaatgca
 33901 agcttttagga cttcacctga cagatccaag tcaacgtctt gttcagaact gtcctttggac
 33961 tctcaggaat ctttcagatg ctgcaactaa acaggtaaat tctgagtaaa ctggtgccat
 20 34021 gggaatagag tcaagatgag tatgtgcttg tactgacct ctgtttttat ctccatagga
 34081 agggatggaa ggtctccttg ggactcttgt tcagcttctg ggttcagatg atataaatgt
 34141 ggtcacctgt gcagctggaa ttcttttctaa cctcacttgc aataattata agaacaagat
 34201 gatggtctgc caagtgggtg gtatagaggc tcttggtcgt actgtccttc gggctggtga
 34261 cagggaagac atcactgagc ctgccatctg tgctcttcgt catctgacca gccgacacca
 25 34321 agaagcagag atggcccaga atgcagttcg ccttcactat ggactaccag ttgtggttaa
 34381 gctcttacac ccaccatccc actggcctct gataaaggta aattgtcaaa gtagaattta
 34441 cctttgttgC agaattgaaa atgaagcatc tctagctgtt ggatggctgt ctaagcatag
 34501 tgatcaataa gtaggaattg tattccttag taagtaggaa gtaggctgc gatagggtga
 34561 agatttctgaa atgttttgt agtcagaact acttttagtt gataccaata gatttagtgt
 30 34621 ggtgggaatt ttagggttaag aaaatgattt tgttgagttg tatgccagtt cttccttctg
 34681 tttttcaggc tactgttgga ttgattcgaa atcttgccct ttgtcccga aatcatgcac
 34741 ctttgcgtga gcagggtgcc attccacgac tagttcagtt gcttgttcgt gcacatcagg
 34801 ataccagcg ccgtacgtcc atgggtggga cacagcagca atttgtggtg ggtaaatctt
 34861 tacagtgata cctggctatc taaaagggaat gcataaatcc aaaggatcct gaacttcttt
 35 34921 ctttggtcac ttgttcccc catccgtctt cctgaagagc taatgacaaa gtaataaat
 34981 aaataattac acatttctat ggctgcagag aaaataaggc atagtgtggc cccagtgata
 35041 tttccttgga cacgtccttc acatggtcag tcttaciaag gttgggttag gtgtttcata
 35101 aagtgttctc atttaattta cacaaaggcc cacttcctta ggaagaggtg gtagtcataat
 35161 ttgagatcaa atctgtgtaa ttctcagacc tcttaccctt gcctcatcat gcattttgac
 40 35221 tataaatatt tagcagtcgg ttttattatc ttttctgtga gttaaacttt tttcatggac
 35281 ctaagaatat tcagaaataa gtagtagcat ttctgtactc ttaaccacaa aaatctcaac
 35341 ctgaagcttt gatcaaaagt ttgtgtctta aaagtagctt cattaaaagt atagtctaata
 35401 gacatttctg atttctcaga ctttaagacc ttattaggtt agtttagaaa acaaagatgg
 35461 agcctaccag aacagatgtt aggaatctca ttttgctggt tgctttgtgt atgtactcat
 45 35521 attggggctt tggctttctt catttattac tgttgggtatt ggcccatctc catgagggtga
 35581 cttaatagaa cgttgagggc accttttatt ttaaactctt tttctaggaa gaagagagt
 35641 tttgtgtcct tgtaagaatc aagttattta taaaagctgc taaatgtagc agaataataa
 35701 ccccttttaa aactcaaatc cagaaacagg agaaacagat ggtacttaca tattgcaaaa
 50 35761 gctatcttcc ttctatacat gaggctgtca gctgaatagt cttggaagag tgaggagtga
 35821 atttttctgc tggcaactcg gttagtttta gcagttgggt ctaaaacttg gcaaagtttt
 35881 caccaaatac atggaagata tacaaaaata gagggggcat gtaaaagaaa aacgttgaca
 35941 tagtctgagc attactttct catcttctct ttttatatac cttttacca gaatgattgg
 36001 tgcccttact gtaggaaagt tgtctttggg attcagcgct gtatggaagc tctgttgac
 36061 tgtgtatggg ggaggggtgc tgctttgaat tagtgctgcc aggaggcctc ttttcagtga
 55 36121 cattcaagtt aatggaatcc ttcttccttc ctgaactaat tgcaagttac ggggaacttc
 36181 gggatatataa tgtaataaat tacagtctaa taattgttcc tcaaacttta cagaggagaa
 36241 tgccctgttt gtttaacctg tttcttttgg caggaggggg tccgcatgga agaaatagtt

5 36301 gaagggttgta ccggagccct tcacatccta gctcgggatg ttcacaaccg aattggtatc
 36361 agaggactaa ataccattcc attgtttgtg caggtatggt ttaagtgaag tgttctaggt
 36421 tttatgtcca taaaatttcc agattgtaat gactaataac atttcagaaa attagggacc
 36481 ataatagggg taccaacatt taattttatg aaaattccct acattttttg gtcagtaaga
 36541 gaaacattga gacttgagaa gagggaggag atttcacatt tcaactttat ggggtgcctag
 36601 aggggagagc tgacctgggc tgccagaggc agggcataga cccccaacca attctggggt
 36661 ttccaaatct tagatcagtt agagctgcct ctgaagaaaag gggttatagc taaaaaatat
 36721 tatggaaatc cagtgtcca gagcattaaa caccccaaga cataaaattc agagaatatt
 10 36781 atttactaca gtgtgaatgc ctcttgact ctgaattggg aatgtttgca ccacagtggg
 36841 gggcttgcca tgttttagct tttagatttaa ttaggttttg tttgtgtttt ctcttagct
 36901 gctttattct cccattgaaa acatccaaag agtagctgca ggggtcctct gtgaacttgc
 36961 tcaggacaag gaagctgcag aagctattga agctgaggga gccacagctc ctctgacaga
 37021 gttacttcac tctaggaatg aaggtgtggg taagtaaaaa ggaaccaaaag cctttagcag
 37081 atgtgtacat tgaagtctca gtttttctc aagggccttt ttctccttgt ctcttagcga
 15 37141 catatgcagc tgctgttttg ttccgaatgt ctgaggacaa gccacaagat tacaagaaac
 37201 ggctttcagt tgagctgacc agctctctct tcagaacaga gccaatggct tggaaatgagg
 37261 tagggaaatg tgagcagtta tttatctggt agtttcctag agcaggtatg gcagcttgtt
 37321 ctttctctc aaaacactta gtacacattc atttgcattg atgtttccct ggcttgagta
 37381 tttcttctt atgtgtctca gcaactgctc tgaggaagaa ctataataca agctttaaag
 20 37441 agtctgttca gaatcattac aaataagttg tggtatttaa aattataatt cataagggag
 37501 aaagatgaaa aatgttacca gattaaagaa gatTTTTTcaa aaggatgtaa ggaaagaggc
 37561 agtgTTaaac actgttaaga ggacagtTTa tcagtatttt ttactaaact ttaataaaac
 37621 ttttctattt gaatttctgc tatgaatttt tcttcagcat ttgtcctcag tacagggtgg
 37681 tccttgaAAC attgtttcta ataaaactag aacatcctga tattttatcc attctataga
 25 37741 gatcattgat ggtacacaga catacagtgg attatgtttg ttgagtgaat ggaaagagag
 37801 attgttaggt ttacaacgat gcagctcttg agaccggagt ttaagatcag cctgggcaac
 37861 atagtgaAAC cccatcttta gctgggcatg gagatggatg cctatagtcc tagctactgg
 37921 ggagacgggg gcaggaggat tgcttgaacc caggagttaa cagactgcac tcagtgcag
 37981 agccagactc caacacaaaa aaaaaaaaaa aaaaaaagca aattaccagt gagtgtgtg
 30 38041 ttacttgggt ttttaatagg catcttatta acatgttcca acttgagccc ttaactttct
 38101 ccacctaccc ccttccacaa acctgttttc actgtcttct ctgtcttagt taatgtcagc
 38161 tttgtctgtc cagctgtctca ggctaaaact tttctttcat ataacacatc ctatcagcag
 38221 ctctgttttg tgggtaggca ttttgccctt tttttttttt ttttttttaa actgctatat
 38281 ctctagcatg tagaacagt cctggcagca cataatagg gcttaatata atatttgttg
 35 38341 aaagaacaag tcagtgcagta tttttaatgt gaggtgcaaa gagaAAAAaa aatgtatctt
 38401 tgagggtgtg agttttgaag aacttccatt ttctaagcat ttgtgtaatg ttggagttac
 38461 ttgttccctt tgtaatctga aagtatgctt taaaaaaaat tagtgactt ttgagaattt
 38521 tcattttgct ttctattctt ccttgctttg tgcatgttta tctagactgc tgatcttggg
 38581 cttgatattg gtgcccagg agaaccctt ggatgcgcat aggatggat gtgtctcata
 40 38641 tttctcgatt aactccagat caagcttaag ttctaaaact tttatcagaa gagccgggtt
 38701 gctcatctgg gaaaccagt ttggcagaaa agtagtggct tcaattaaaa gcagttctta
 38761 aattccagtc agcaacagta tctttaatgg agcacaggga attcagagcc acacaatgag
 38821 tagcagtagg attacaccac caacaaatac atgctactgc taggcctctg cagtgcagga
 38881 tgttacaatt tacctggctt tttattctct ttttgccag aggactcata atacctttgt
 45 38941 ctacaagcta cccaaggaag ataggaaaac tcctgtttct aggtcagat ctcggtggg
 39001 tttttacata gttgcattat catcagggtt ttcttgaaaa gctaatttaa atctgggtaa
 39061 tgaacatgga ggatggcata gaccactaac aattataact gtcttacatt tataaccgca
 39121 tctgcttcta cctaattatg aaaccactaa agcgcagatt ctactgtga gaaataacat
 39181 gtcaacccta agataaaata tgttgaggtt tcatggaaat agtgcccttc cttagtactt
 50 39241 ttgtgggtgt cacttggcct ttttgtcaag atagattaca cctgccagac ctctattattg
 39301 tcttaatcct ccttccatg acttctcact gcctaggtgg tcacacagta gattcctgct
 39361 tcttctctc gggaacccca agtctcttga caggggtaaa tgcagagtgt tcagggttag
 39421 actaatgatg tgactaggcc ctgctggtgt gcctgtctga tggaaataga tgttatttgt
 39481 gtagtctcat ggggtggcctg gcactgagta attacttggc taaagaaagc tggaggttga
 55 39541 agaggctaga aagcgttgtt ttctgacaag tttgctgctg aactttggat gccctaacct
 39601 cagtgttaac gtctatgtct gcttctctcc tctctctttt gccttcttc ttgcctattt
 39661 tgttgacacc ctgactcttc tagatcctag ctatcgttct tttactctg gtggatatgg

39721 ccaggatgcc ttgggtatgg accccatgat ggaacatgag atgggtggcc accaccctgg
 39781 tgctgactat ccagttgatg ggctgccaga tctggggcat gccaggacc tcatggatgg
 39841 gctgcctcca ggtgacagca atcagctggc ctggtttgat actgacctgt aaatcatcct
 39901 ttaggtaaga agttttaaaa agccagtttg ggtaaaatac ttttactctg cctacagaac
 5 39961 ttcagaaaga cttggttggg aggggtgggag tggtttaggc tatttgtaaa tctgccacaa
 40021 aaacaggat atactttgaa aggagatgtc ttggaacatt ggaatgttct cagatttctg
 40081 gttgttatgt gatcatgtgt ggaagttatt aactttaatg ttttttgcca cagcttttgc
 40141 aacttaatac tcaaatgagt aacatttgct gttttaaaca ttaatagcag cctttctctc
 10 40201 tttatacagc tgtattgtct gaacttgcac tgtgattggc ctgtagagtt gctgagaggg
 40261 ctcgaggggt gggctgggtat ctcagaaagt gcctgacaca ctaaccaagc tgagtttctc
 40321 atgggaacaa ttgaagtaaa ctttttgttc tggctccttt tggctcgagga gtaacaatac
 40381 aaatggattt tgggagtgc tcaagaagtg aagaatgcac aagaatggat cacaagatgg
 40441 aatttatcaa accctagcct tgcttggtta atttttttt tttttttttt aagaatatct
 15 40501 gtaatggtac tgactttgct tgctttgaag tagctccttt tttttttttt tttttttttt
 40561 tgcagtaact gttttttaag tctctcgtag tgttaagtta tagtgaatac tgctacagca
 40621 atttctaatt ttaagaatt gagtaatggg ttagaacact aattcataat cactctaatt
 40681 aattgtaatc tgaataaagt gtaacaattg ttagaccttt ttgtataaaa tagacaaata
 40741 gaaaatggtc caattagttt cttttttaat atgcttaaaa taagcagggt gatctatttc
 20 40801 atgtttttga tcaaaaacta tttgggatat gtatgggtag ggtaaatcag taagaggtgt
 40861 tatttgaac cttgttttgg acagtttacc agttgccttt tatcccaaaag ttgttgtaac
 40921 ctgctgtgat acgatgcttc aagagaaaat gcggttataa aaaatggttc agaattaaac
 40981 ttttaattca ttcgattg

Noggin (Nog)

25 The NOG gene encodes the noggin protein, and is a bone morphogenetic protein 4 (BMP4) inhibitor. Activation of NOG in supporting cells inhibits BMP4 and induces hair cell regeneration (Lewis et al. (2018) *Hear Res.* 364:1-11).

The human NOG gene is located on chromosome 17q22. It contains 1 exon encompassing ~2 kilobases (kb) (NCBI Accession No. NG_011958.1). The full-length wildtype
 30 NOG protein expressed from the human NOG gene is 232 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype NOG protein is or includes the sequence of SEQ ID NO:
 35 16. Non-limiting examples of a nucleic acid encoding a wildtype NOG protein is or includes SEQ ID NO: 19. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 19 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype NOG Protein (SEQ ID NO: 16)

40 MERCPSLGVTLYALVVVLGLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDPI FDPKEKDLNETLLRSLLGGHYDPG
 FMATSPPEDRPGGGGGAAGGAEDLAELDQLLRQPSGAMPSEIKGLEFSEGLAQGKKQRLSKLRRKLQMWLWSQTF

CPVLYAWNDLGSRFWPRYVKVGSFCFSKRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGGQRCGWIPYPII SECKCS
C

Mouse Full-length Wildtype NOG Protein (SEQ ID NO: 17)

5 MERCPSLGVTLIALVVVLGLRAAPAGGQHYLHIRPAPSDNLPVLDLIEHPDPI FDPKEKDLNETLLRSLLGGHYDPG
FMATSPPEDRPGGGGGPAGGAEDLAELDQLLRQPSGAMPSEIKGLEFSEGLAQGKKQRLSKKLRRKLQMWLWSQTF
CPVLYAWNDLGSRFWPRYVKVGSFCFSKRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGGQRCGWIPYPII SECKCS
C

10 Rat Full-length Wildtype NOG Protein (SEQ ID NO: 18)

MERCPSLGVTLIALVVVLGLRAAPAGGQHYLHIRPAPSDNLPVLDLIEHPDPI FDPKEKDLNETLLRSLLGGHYDPG
FMATSPPEDRPGGGGGPAGGAEDLAELDQLLRQPSGAMPSEIKGLEFSEGLAQGKKQRLSKKLRRKLQMWLWSQTF
CPVLYAWNDLGSRFWPRYVKVGSFCFSKRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGGQRCGWIPYPII SECKCS
C

15

Human Wildtype NOG cDNA (SEQ ID NO: 19)

atggagcgcctgccccagcctaggggtcaccctctacgccctggtggtggtcctggggctgcggcgacaccggccgg
cggccagcactatctccacatccgcccggcaccagcgacaacctgcccctggtggacctcatcgaacaccagacc
ctatctttgaccccaaggaaagatctgaacgagacgctgctgcgctcgctgctcgggggccactacgaccaggc
20 ttcattggccacctcgcccccgaggaccggcccggcggggggtgcagctggggcgcgaggacctggcgga
gctggaccagctgctgcggcagcgccgtcgggggccatgccgagcgagatcaaagggttagagttctccgagggct
tgcccgagggaagaagcagcgccctaagcaagaagctgcggaggaagttacagatgtggctgtggctgcgacattc
tgccccgtgctgtacgcgtggaacgacctgggcagccgcttttgccgcgctacgtgaagggtggcagctgcttcag
taagcgctcgtgctccgtgcccagggcatggtgtgcaagccgtccaagtcctgtcacctcacgggtgctgcgggtggc
25 gctgtcagcgcgcgggggccagcgctgcggctggattcccatccagtaccccatcatttccgagtgcaagtgctcg
tgctag

A non-limiting example of a human wildtype NOG genomic DNA sequence is SEQ ID
NO: 20. The exons in SEQ ID NO: 20 are: nucleotide positions 1-1892 (exon 1).

30

Human Wildtype NOG Gene (SEQ ID NO: 20)

1 aaaccggtgc caacgtgcgc ggacgccgcc gccgccgccg ccgctggagt ccgccgggca
61 gagccgggccg cggagcccgg agcaggcgga ggaagtgcc ctagaacca gctcagccag
121 cggcgcttgc acagagcggc cggacgaaga gcagcgagag gaggagggga gagcggctcg
35 181 tccacgcgcc ctgcgccgcc gccggcccgg gaaggcagcg aggagccggc gcctcccgcg
241 ccccgcggtc gccctggagt aatttcggat gccagccgc ggccgccttc ccagtagac

301 ccgggagagg agttgcggcc aacttggtgtg cctttcttcc gccccggtgg gagccggcgc
 361 tgcgcgaagg gctctcccgg cggctcatgc tgccggccct gcgcctgcc agcctcgggt
 421 gagccgcctc cggagagacg ggggagcgcg gcggcgccgc gggctcggcg tgctctctc
 481 cggggacgcg ggacgaagca gcagccccg gcgcgcgcca gaggcattgga gcgtgcccc
 5 541 agcctagggg tcaccctcta cgccttggtg gtggtcctgg ggctgcgggc gacaccggcc
 601 ggcggccagc actatctcca catccggccg gcaccagcg acaacctgcc cctggtggac
 661 ctcatcgaac acccagaccc tatctttgac cccaaggaaa aggatctgaa cgagacgctg
 721 ctgcgctcgc tgctcggggg ccactacgac ccaggcttca tggccacctc gcccccgag
 781 gaccggccccg gcggggggcg ggggtgcagc gggggcgcg aggacctggc ggagctggac
 10 841 cagctgctgc ggcagcgcc gtcgggggcc atgccgagcg agatcaaagg gctagagttc
 901 tccgagggct tggcccagg caagaagcag cgcctaagca agaagctgcg gaggaagtta
 961 cagatgtggc tgtggtcgca gacattctgc cccgtgctgt acgctggaa cgacctgggc
 1021 agccgctttt ggccgcgcta cgtgaagggtg ggcagctgct tcagtaagcg ctctgctcc
 1081 gtgcccagg gcatggtgtg caagccgtcc aagtccgtgc acctcacgtg gctgcgggtg
 15 1141 cgctgtcagc ggcgcggggg ccagcgctgc ggctggattc ccatccagta cccatcatt
 1201 tccgagtga agtgctcgtg ctagaactcg ggggccccct gccgcaccc ggacacttga
 1261 tcgatcccca ccgacgcccc ctgcaccgcc tccaaccagt tccaccaccc tctagcgagg
 1321 gttttcaatg aactttttt tttttttt ttttttttc tgggctacag agacctagct
 1381 ttctggttcc tgtaatgcac tgtttaactg ttaggaatg tatatgtgtg tgtatatacg
 20 1441 gtcccagttt taatttactt attaaaaggt cagtattata cgttaaaagt taccggcttc
 1501 tactgtattt ttaaaaaaaaa gtaagcaaaa gaaaaaaaaa agaacagaga aaagagagac
 1561 ttattctggt tgttgcta atgtttaacc tgctatttat attccagtgc cttcgcagtg
 1621 gcgaagcagg ggggaaaagt ttttttttc ttgaagtaca aagagacggg ggaacttttg
 1681 tagaggactt tttaaaagct attttccatt cttcgaaaag tgttttggtt ttccttggac
 25 1741 ctcgaagaag ctatagagtt caatgttatt ttacagttat tgtaaataata gagaacaaat
 1801 ggaatgacta atcattgtaa attaagagta tctgctattt attctttata atatcccggtg
 1861 tagtaaatga gaaagaagtg cagagcagga tt

Growth factor independent 1 transcriptional repressor (GFI-1)

30 The GFI-1 gene encodes a nuclear zinc finger protein, and acts as a transcriptional repressor. GFI-1 is activated by Atoh1 and Pou4f3 in early development and is required for hair cell survival after birth (Hertzano et al. (2004) *Hum. Mol. Genet.* 13(18):2143-2153; Costa et al. (2015) *Genom Data* 6:77-80).

35 The human GFI-1 gene is located on chromosome 1p22. It contains 7 exons encompassing ~12 kilobases (kb) (NCBI Accession No. NG_007874.1). The full-length wildtype GFI-1 protein expressed from the human GFI-1 gene is 422 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype GFI-1 protein is or includes the sequence of SEQ ID NO:

- 5 21. Non-limiting examples of a nucleic acid encoding a wildtype GFI-1 protein is or includes SEQ ID NO: 24. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 24 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype GFI-1 Protein (SEQ ID NO: 21)

10 MPRSFLVKS KKAHSYHQPRSPGPDYSLRLENV PAPS RADSTSNAGGAKAEPRDRLSPESQLTEAPDRASASPDSC
SVCERSSEFEDFWRPPSPSPASEKSMCPSLDEAQPFP LFPKPYSWGLAGSDLRHLVQSYRPCGALERGAGLGLF
CEPAPEPGHPAALYGPKRAAGGAGAGAPGSCSAGAGATAGPGLGLYGDFGSAAAGLYERP TAAAGLLYPERGHGLHA
DKGAGVKVESELLCTRLLLGGGSYKCIKCSKVFSTPHGLEVHVRRSHSGTRPFACEMCGKTFGHAVSLEQHKAVHSQ
ERSFDCKICGKSFKRSSTLSTHLLIHSDTRPYPCQYCGKR FHQKSDMKKHTFIHTGEKPHKCQVCGKAFSQSSNLIT
15 HSRKHTGFKPFGCDLCGKGFQRKVDLRRHRETQHGLK

Mouse Full-length Wildtype GFI-1 Protein (SEQ ID NO: 22)

MPRSFLVKS KKAHSYHQPRSPGPDYSLRLETVPAPGRAEGGAVSAGESKMEPRERLSPDSQLTEAPDRASASPNSCE
GSVCDPCSEFEDFWRPPSPSPASEKSLCRSLDEAQPYTLPFKPYAWSGLAGSDLRHLVQSYRQCSALERSAGLSL
20 FCERGSEPGRPAARYGPEQAAGGAGAGQPGRCGVAGGATSAAGLGLYGDFAPAAAGLYERPSTAAGRLYQDHGHELH
ADKSVGKVESELLCTRLLLGGGSYKCIKCSKVFSTPHGLEVHVRRSHSGTRPFACEMCGKTFGHAVSLEQHKAVHS
QERSFDCKICGKSFKRSSTLSTHLLIHSDTRPYPCQYCGKR FHQKSDMKKHTFIHTGEKPHKCQVCGKAFSQSSNLI
THSRKHTGFKPFGCDLCGKGFQRKVDLRRHRETQHGLK

25 Rat Full-length Wildtype GFI-1 Protein (SEQ ID NO: 23)

MPRSFLVKS KKAHSYHQPRSPGPDYSLRLETVPVPGRADGGAVSAGESKMEPRERLSPESQLTEAPDRASASPNSCE
GSVCDPSSEFEDYWRPPSPSPASEKSLCRSLDEAQPYTLPFKPYAWSGLAGSDLRHLVQSYRQCSALERSAGLSL
FCERGAESGRPAARYGSEQAAGGAGAGQPGSCGAASGATSAGGLGLYGDFAPAAAGL FERPSTAAGRLYQDRGHELH
ADKSVGKVESELLCTRLLLGGGSYKCIKCSKVFSTPHGLEVHVRRSHSGTRPFACEMCGKTFGHAVSLEQHKAVHS
30 QERSFDCKICGKSFKRSSTLSTHLLIHSDTRPYPCQYCGKR FHQKSDMKKHTFIHTGEKPHKCQVCGKAFSQSSNLI
THSRKHTGFKPFGCDLCGKGFQRKVDLRRHRETQHGLK

Human Wildtype GFI-1 cDNA (SEQ ID NO: 24)

atgccgcgctcatttctcgtcaaaagcaagaaggctcacagctaccaccagccgcgctccccaggaccagactattc
35 cctccgtttagagaatgtaccggcgcctagccgagcagacagcattcaaatgcaggcggggcgaaggcggagcccc

gggaccgtttgtccccgaatcgcagctgaccgaagccccagacagagcctccgcatccccagacagctgcgaaggc
 agcgtctgcgaacggagctcggagtttgaggacttctggaggcccccgtcaccctccgctctccagcctcggagaa
 gtcaatgtgccatcgctggacgaagcccagcccttccccctgcctttcaaaccgtactcatggagcggcctggcgg
 gttctgacctgcggcacctggtgcagagctaccgaccgtgtggggccctggagcgtggcgtggcctgggcctcttc
 5 tgcgaacccgccccggagcctggccacccggccgcgtgtacggccccgaagcgggctgccggcggcggggggccgg
 ggcgccaggagctgcagcgcaggggcccgtgccaccgctggccctggcctagggctctacggcgacttcgggtctg
 cggcagccgggctgtatgagaggccacggcagcggcgggcttgctgtaccccgagcgtggccacgggctgcacgca
 gacaaggcgctggcgtcaaggtggagtcggagctgctgtgcaccgcctgctgctggcgcggtcctacaagtg
 10 catcaagtgcagcaaggtgttctccacgccgcagggctcgaggtgcacgtgcgcaggtcccacagcgggtaccagac
 cctttgcctgcgagatgtgcggcaagaccttcgggcacgcggtgagcctggagcagcacaagccgtgcactcgcag
 gaacggagcctttgactgtaagatctgtgggaagagcttcaagaggtcatccacactgtccacacacctgcttatcca
 ctacagacactcggccctacccctgtcagtactgtggcaagaggttcaccagaagtgcagacatgaagaaacacactt
 tcatccacactggtgagaagcctcacaagtgcaggtgtgcggcaaggcattcagccagagctccaacctcatcacc
 cacagccgcaaacacacaggcttcaagcccttcggctgcgacctctgtgggaagggtttccagaggaagggtggacct
 15 ccgaaggcaccgggagacgcagcatgggctcaaatga

A non-limiting example of a human wildtype GFI-1 genomic DNA sequence is SEQ ID
 NO: 25. The exons in SEQ ID NO: 25 are: nucleotide positions 1-151 (exon 1), nucleotide
 positions 3291-3504 (exon 2), nucleotide positions 3831-4013 (exon 3), nucleotide positions
 5789-6276 (exon 4), nucleotide positions 6392-6529 (exon 5), nucleotide positions 8124-8289
 20 (exon 6), and nucleotide positions 10670-12116 (exon 7). The introns in SEQ ID NO: 25 are:
 nucleotide positions 152-3290 (intron 1), nucleotide positions 3505-3830 (intron 2), nucleotide
 positions 4014-5788 (intron 3), nucleotide positions 6277-6391 (intron 4), nucleotide positions
 6530-8123 (intron 5), and nucleotide position 8290-10669 (intron 6).

25

Human Wildtype GFI-1 Gene (SEQ ID NO: 25)

1 gaggggtgcgc ccaccggtcc cgcggggcgc ccgcgggacg cgcgccagg gccctctccg
 61 ccggggggctc ggcgctcgcc cacctcttcc aaatttaacc attacctaaa tccgaaggga
 121 aatgagcaaa cctctcgat tgggtgtcaa ggtattttca gcctcgttgg gcgtatttat
 181 ccccaagtgt ttccacaaca agctatttcg gggcctgcgg ggcaggtttc gctctgcgga
 241 cgcctgtggc actcgccggg ctccaggccg gcggcaccgc gggccggtga ttcacggtcc
 301 cgacccgggg gtggtgcagc cctaggaggc ggccgggctc ggggtggggg gggcgggtga
 361 ccgaggcctg aggggcgggg agggctctcg gagcggggcg cccccaccc ctctctcgcc
 421 agtcaatctg tgcctcaat ctgtggcttc tctcgctgcg gaagtctccc tggagccaag
 481 aatagttcat tttctttcaa gtcatttcta gtgcctaagt gtccggacct ccaatttccc
 541 ccatcccctg ccgaccaca gggagagaac tgggaggact actaaggggc gcgcgggacg

35

5 601 ggctggaaaag gccaggcccc ccaccgcctg gccacttgcg caaaggagcg cgcccggccg
 661 cccgacgggg gttgggagca ggtctgggag ggctatgcga gcgactcagt aacgctcagg
 721 aagtgaagct tgtggttttg ggggctgagc tcggaaggag aatttttttt ttttttaagt
 781 cagagagata gagcgggtctc tcccgaaaagc aagattccgt ttgaaaactc tcctagcgcg
 841 gtgcccgcgc cgtgactccg caggtaggtc cgccgagcct gttctgcgcc tcctgccctg
 901 gtggggggcg ccgcggggac tcgcagagca ctggcactgc gggggcgatc agagggcgcg
 961 ggcggtttcc cacctgctgc ggaccgccgt gcgggagccc agagagctcc ggcagctggg
 10 1021 ttagggcgcg acccgcgagc tgtgactgga aatctggagc tggggcgcg cagcaggcg
 1081 tctggtggtt cggcagggga gccaaatcca ccagggaagg aaacatctgg tggggaggcg
 1141 gcggcagctg cgctcgggag gacgccccct tagcgccctc ggctccctcc ttcctgggcc
 1201 cggacggtga ggagaggcct gagcgcgcg aggggcccgc ccacctccc cgccagctgc
 1261 agcgcggggt tgccttccca cgcgctcggc ccaggccccg gggccctat ctcctccaaa
 1321 ctctgtcgct cccacagga accagcaaag cgggcccggg tgcgagagag gcgtgggctt
 1381 acagcccggg tggggaggcg gcctccgcgt ccgcctggtc tctggtggcg ccagcccaga
 15 1441 cccagctccg gcgctgacag ttaccccgcc cccatctgtc ccgctcccag ccaactggg
 1501 tccaagctgc agcgggaccc tcgggacaac gccactccgt ttttctttt ttcgagttt
 1561 cgtggctggt taaagaattg ggtttgggt ttgtggcgtc taattgtacg gacgagaagt
 1621 gcaggaagcg acaaagctct agccctagga gccaccccgg agggaggcg atggcactct
 1681 caccctagg ggcatctctg cgcttgggta gcgggaaact tcgcgggagc cccgcgacac
 20 1741 gtcccaggcg tcttttctcc caggtctatt cccattcctc cgagagaagg gacacaatgg
 1801 ggctggggat ctggagcagg gggcctgcac cctacaggga ccaaggcctg taggactcgt
 1861 ttgagctgag agcgccaacg gacagacgta gactgtgtgg cctgcatctt gcctaggaag
 1921 ccgaggggct ctagtccgg cagtggaaac agcgcgaaag cgaggactg caggtcctgc
 1981 cccggcccag agttcccagc accctcgttt ctgaaccagc cgaggccacg gagaactgct
 25 2041 gtactgcagc tcacgtgtga acccggtcac catcgcttc accccgggag gaaggcagat
 2101 tcgtttactc cagaccacct cgactgtggg gtaccgcccc cgagaccggc tggagcctag
 2161 cggcaggcgc agccacgccc tcccgtgcg ctcagatttc gacctggtat taggtgaact
 2221 gattgggggt taatgagagc gacgccccg gcagctagtt ccctcccggg cccgggcccg
 2281 acccccgctc tctgctaatt cagcctgcgc gctctggcgt cctgtctttt ttgtctgcta
 30 2341 aattgtcggg gcaactaccga ctcgggacac ctagcatttc ccagtcaacg ttcgtggatc
 2401 gggctccacc tccctaggac aagatttttt ggtgagcaga acggaaagtg cttttttccc
 2461 gggacctgat tcccagaggt aggtctccat ggtctgggat ggctcgccgc agcctcgacc
 2521 ggtgcccgcc gcagccggga gtccaagggc aaagtcttc ctacgtgggg cactagttag
 2581 gggcgagtgg gatcaccag atgcgagttt ctctggggc gggggttggt cgtctgttgt
 35 2641 tcccctcact ttctgtccc tttgctctcc cacctccttt ctctggcctc tgctgtcccc
 2701 aatccctctg ctgctgtcct ccgcgcgcc cacagtttta tcttgtgttc tgttcccccc
 2761 tccccccggt ctttctactc cagttggtag ctggctcttg gaggtcttg ctcttgccc
 2821 cttccgggtc ctcgaccact gggcatcccc ggcccctaaa ccgatccgcg tgtccccgcc
 2881 ctccctcgcc agccgtaaag cacagccagg caggcgatga gtagctgagt tggggtaacc
 40 2941 caccgatgg gaactacagc tctccaggga gtttgattgc cgagcgagc ttcgctagga
 3001 aaggggagga gctggggggc gtgggcaggg aggaggaaa gggcctgaga cagggccccg

3061 ggacaggttt taccgctgag ctgtgtcagt ggcggcggcg gcaacgacgg cgggttcgcy
 3121 ccacctgtcc aagtgccacc tggtaagcgc ggcgagcag ggtcaagccc ctctcccgt
 3181 gggccctctg cgcgcctccc tggccgcgc tctccctccg cctgggtgcc cagtccgccg
 3241 caccggagag tactgaccca cgtctccacc tggttttctt tccctctcag gtctcctccg
 5 3301 ggctggggct gagcaagccc tcggagtgc cgtgggtgac agcggctcca gggactcttg
 3361 gggcgagtg gggaaaagtgc cggaccacca tgccgcgctc atttctcgtc aaaagcaaga
 3421 aggtctcacag ctaccaccag ccgcgctccc caggaccaga ctattccctc cgtttagaga
 3481 atgtaccggc gcctagccga gcaggtgcga ggcgcgcgcg ggccaggcgg ggctgctccc
 3541 ccgatgcct actgcacctc ggacacacct tagtccggag ctgggagggg ctgccccaac
 10 3601 gtcccttttg ctgctgtttt tgtttctac tgtcctggtt cctccgggtt tgtctcctag
 3661 gtgccatggc ctctctgcgc ctgccctcgg atccgagagg gttcccgcc ggggtctggg
 3721 tggagagggg aagacgctcg gctgccctgg tcgggggatt gggggagcct tcagcacctc
 3781 cagactcaac cgggtcccagc ctgagcccct cacctgcctc ctctttgcag acagcacttc
 3841 aaatgcaggc ggggcgaagg cggagccccg ggaccgtttg tccccgaat cgcagctgac
 15 3901 cgaagcccca gacagagcct ccgcatcccc agacagctgc gaaggcagcg tctgcgaacg
 3961 gagctcggag tttgaggact tctggaggcc cccgtcacc tccgcgtctc caggtaggaa
 4021 cccactggga acctcttggg cgggagctgc agggacccgg cagtgcctgg ggggaattgg
 4081 cgcgaccttg ggcgtagaaa tgctaacggg gaggtagaga gtctttccgg gagaaggag
 4141 ctgattcgta ggggaaggag gcatccggct tctctgggac ttggacagct tgcccgtgg
 20 4201 ggctgctgcc tccatcccag gcggcaggac ctagctgct tgtogcttag attcgtttg
 4261 gcggagctgg ccagtgcagg aaaacaaacc agtcgtttcg actggcggca acgctgacct
 4321 ttcattttct gaccagacct gactgtttta taaagtgcga ggatcctgca atctagacct
 4381 ccaaacctca aacagagaac agggcagaac gggccaggca gaggagctag gcgctgggcy
 4441 gcagggaggg ggcaggacga aaatctcagc ccgcggcttg gtcttcacag gcgcagattg
 25 4501 ggggcctgtt tcatttttctg ttttgccggg ttaacctagc ctgaggggcy tgctctctgg
 4561 gtttcatttc cagcgagcaa tccagcttca ggcaaaactaa gtgaccacac gttgggtggg
 4621 ggcgtctcga gtcccggccg ggggaaggaa tgagcagacc agccggattc tgtcaagggc
 4681 cggttatatc cagaatatgt ttgctagttt tagaagatac caccaccgct cccacaatca
 4741 gtgagttgac ttggcgaaaa ccatagctcc agcaagtgtg tctgggagcc ggcggcggga
 30 4801 ggattcttcc tgccagggcy tcaagtggcc agacaaggat tgggcgcgcc ccgaaccct
 4861 ccgaacgaaa ctccgggtac agcctctcac tgaagtggcc agcctgaact ggagtgttgt
 4921 gcgcacacac acacacacac atttgtaaatt gccgtatgca ctacatgcy ttgggggtcac
 4981 tagtttttagc aaaattcacg tgggtggggg cgtagcaggc cgagaattca gagctgtctc
 5041 cttgcaggtg gtggctaaac cttatgagtt atagttattc tctgagaaat tcagggttccc
 35 5101 cgctccatc aaactgtaac aggaatggg agtatttggc tgtcaattta agccaaaag
 5161 cccctttcct gctgctcctt tgctacgtac actgggcact taacttcgtg aaatcttaat
 5221 accttcgggt ttattcagac agcagccttt cgggtagttc ggggccgcat ttatggacct
 5281 tctccctcct tcctcttggg ttctgggaag aaaaagaatt gaatgggaac atgtaggggc
 5341 tgggagagtg cctgcgctgg tggctggacc ctccgccct tgagtgtgt gagggccga
 40 5401 acggccgcca cttctcctt cttaacagct caactacggg catttataga tgcgccctc
 5461 cctgtaggat ctccaggtgc gcgggtccag ccagaaaaga tctcgggaac gccgagcgcc

5521 tccgctgcac tcgcacagaa ttacgacct cctctcccga ggtcttttca atgatctgtt
 5581 tactgttctg cctcctatag tggcctgcga ggccccaggg cccgggccac gttttaccct
 5641 ggggcgagcc tggcacctgg cgcacgcagt gttctacaag cgctgggtgc cccgcagtcc
 5701 gcgaacacgc cacgctcgca gccgcagccc ggcggcctcc gctctgccgt ctgaagcctg
 5 5761 accggacgct ccccttgtgc ctccacagcc tcggagaagt caatgtgccc atcgctggac
 5821 gaagcccagc ccttccccct gcctttcaaa ccgtactcat ggagcggcct ggcgggttct
 5881 gacctgcggc acctggtgca gagctaccga ccgtgtgggg ccctggagcg tggcgctggc
 5941 ctgggcctct tctgcgaacc cgccccggag cctggccacc cggccgcgct gtacggcccg
 6001 aagcgggctg ccggcggcgc gggggccggg gcgccaggga gctgcagcgc aggggcccgt
 10 6061 gccaccgctg gccctggcct agggctctac ggcgacttcg ggtctgcggc agccgggctg
 6121 tatgagaggc ccacggcagc ggcgggcttg ctgtaccccg agcgtggcca cgggctgcac
 6181 gcagacaagg gcgctggcgt caagggtggag tcggagctgc tgtgcacccg cctgctgctg
 6241 ggcggcggct cctacaagtg catcaagtgc agcaagggtga ggctcccag ctcaccacct
 6301 cgcttgcgt gcgcccgtt cccctacccg cgctcgcct gcgccccgcg gccctctca
 15 6361 gcggccttct ctctggcccc acccgctta ggtgttctcc acgcgcacg ggctcgaggt
 6421 gcacgtgcgc aggtcccaca gcggtaccag accctttgcc tgcgagatgt gcggaagac
 6481 cttcgggcac gcggtgagcc tggagcagca caaagccgtg cactcgcagg taagcgcggg
 6541 gcgcaccgcc gcgcgcggcc ctgctcgggg atcttctgca tctcctcgt gcagcaccag
 6601 ccaactctct cctggaagtt ttctcctcga ctccccccag ttctctcccc caagccctcc
 20 6661 gctgcgtccc cttgccctgg tgcaggtgtg tagggaaagg aggattgtgg ccggctcagg
 6721 ccttgaggca gccctggatt ttggtgtcac accactgtga gcctcgagag tgtgatcctc
 6781 attgttactt tgggcttgag gtaggtttgt atgcactgat tcgtgctgct gatatatcag
 6841 acttactagc tctgtttctt tgtgcctatt cttttacca aatggttgtc acttaatttg
 6901 cattgacccc tctcgactga aaaggcagga atctcagctc atttagagca tctagtagca
 25 6961 tattcacccc gctattcatt ctttccttcc ttcttttctt ttcttttctt ttttcttttc
 7021 ttttctttca gactctggct ctgttgccca ggctggaggg cagttgcaca atctcagctc
 7081 actgcaacct cccctgtca ggttaaagtg attctcgtgc ctcagcctcc taagtagctg
 7141 ggattacagg cgcatgccac tacagcacag ctaatttttg tgttcttagt agagacggga
 7201 tttcgccacg ttagccaggc tggtcacgaa ctctggcct caagtgatcc accagcctgg
 30 7261 gcctcccaaa gtgctgggat tacaggcgtg agccaccatg cccaccgcc gctattttatt
 7321 cattcattca ttaataaata tttgttggct aacttccagg tgccaagtac ttaagaatct
 7381 tataacacat caggtccttg acagcatgcc cacatgaaga ttatagttta gctgagagat
 7441 ggagagtaga tgagcaagta aatatgccaa tagctatctc aggagaatgc ctacttacga
 7501 aggctaaaaa gactattagc ccatctcccc cagcaccac actggctggg gggagggtggc
 35 7561 atctcaagtg actgaggtct aagcctcctg ttgaggaggg tggagaagtg tgtgctaagt
 7621 ggtgtcaaaa aaagcagggt gtggatatgt atttgccatg ggggtgtggaa ggttgtgggt
 7681 gaagaatgtt ttggtagaaa aagtgttgaa gggccaggca cgggtggctca cgcctgtaat
 7741 ccgaatactt tgagaggccg aggtgggcga atcacttgag gccagggtgtt tgagaccagc
 7801 ctgaccaaca tggtgaaacc ccatctctac taaaaataca aaaattagtc aggtgtgatg
 40 7861 gcgtgtgcct gtagtccctg ctacttggga ggctgagaca cgagaattgc ttgaacctgg
 7921 gaggtggagg gtgcagtga ctgagatcgt tccactgcac tcaagcctgg gcgacagagg

7981 agactgtctc aaaaaaagaa agaaaaagtg ttcaagggat tttagggtca gctgaggggt
 8041 gaggagagca gcagtctagt tgactgcagt aggagttctg catctctctc tctctctctc
 8101 tctctctctc tctctctctg caggaacgga gctttgactg taagatctgt gggaagagct
 8161 tcaagaggtc atccacactg tccacacacc tgcttatcca ctacagacact cggccctacc
 5 8221 cctgtcagta ctgtggcaag aggttccacc agaagtcaga catgaagaaa cacactttca
 8281 tccacactgg tgagctaaaa aggcccttgg cttgtaggaa acaccctgag gccaacatta
 8341 ctcatcttct ctgatttctg gccccagtga gtggtggatg aggccttctt gatggagtta
 8401 ttctctgctc tgtgttaaag aaaacaaagg ggtgggttct ttggttcatt taccggcata
 8461 attctcccca gagccacctt gatttgggtg tgtgtctgaa aggcactca gcaggtcagc
 10 8521 tcacaggtac tctatacttg gaaagaacat tttcctttag gttagcagct gcttcccctg
 8581 ctgctgctc tgggtgaaat atgaagctcc agggtcctct tagagagttg ctctaaagct
 8641 tacctagaga ttgaggactt tccctaacca cctggccttt tgtgggaggg actcgtgtgg
 8701 actctccggc tgcattttca ggagtctgag agcttattct gattgaagag gaacaaataa
 8761 tggcaaatat gattaaactc tctgctaagc attttttaaa tgcattattt cattttatgc
 15 8821 tcacaacaac tctgagaggt agcgactact ccttctcccc atttttagaga tgaaaaaaat
 8881 gaggctaggt aatctgcccc gggtcacact gctagcaaat gacagggcca gagctcaaat
 8941 tcaggtctga cctctcaaat gttcactctt gaccactgtt tattgtattt tatgttcaga
 9001 gtcagtaggt tggtagacag aaagcttctg ttcacttatt gcccttttca aaatatctgc
 9061 aagttaatgc cataataatg ataattcctt acctattata atgctttata atttacaag
 20 9121 tacttttcaa tctatcattt catttgattc ttattgccac tcaagaaagt agaaggagct
 9181 gctcttacca tcctgaaact cagaaagagt gaatgattta tcagaagtaa gactgaatga
 9241 ataatgtagc catgtaatgc tctggctttt aatcctggac tgtttgtcta acactttatg
 9301 tgcgggtggg agtttttaag ccaagaacac tctaatagtc aaaagacatt tacatgagac
 9361 ccagaatttc tgaaaatttt attgcagaat atgaatactg atttagaaca aatcacagtg
 25 9421 tattctaaac acccaccctt gatgtttata aatatacttg ggtaatgtat atatttccat
 9481 tgaaaacca gaaaagtatt ctactttaat cattccctct tacctgaaat ttccatgtaa
 9541 ttcactcctt ataagtaagg tattcaggac acttatcaaa atgcaactag gatcttgact
 9601 gaataaaaca ttaagccctt atcaaacatt tacgttatac ctagaatttg ttttctcaga
 9661 tttgtttgac cctaaaggga tagaatacat tttgatgggt ggtttcttat caaggaaatc
 30 9721 tgaagcatga aaacagaaaa gagtttttag caaggaggac agaggggtcc tcaaaaacaa
 9781 acttcatcta ttttatactt tttccaaggc tgagccctga ctataatgcc atgctgggct
 9841 attggaaatt catgccattt acccaacaac acatgagatg gggaacaaga caaaccttc
 9901 ttgtgttctc ttattttatta atttgtggtg aagaattgct ggtatataaa gaatcatgtg
 9961 attaacccca taaaattaag gaaaaatcaa gacagtaaag tatcagctgc cttaatcctt
 35 10021 tgtggcccaa atgtggattt ttaaaataag atattgaaaa acgtatcctg cacatgtacc
 10081 ccggaactta gaaagaaaga gagagagaga gagaaataaa gaaagaaaga aagtccatgt
 10141 taagatgttt tttcagatat aatctgctgt ctttcaagaa caagaaagaa gacgggctca
 10201 ctgatccata caaactaaca cccacttgga aattcagatt tgaaaacttc ctctgaatta
 10261 gaacggagtc acacggtttt aggacagctt cccctcccc ttctgttga acatctgctc
 40 10321 tgagtgttca tggcttataa agtcagggga gtccctcccg ggtagattca gctggggagg
 10381 gcacgtggcc tttgctctgt ttccgtttag caggaaaccg tttgaggcct ttggctggga

10441 accccccttc agaaagtctc cctttcacct ggtgccccca tgggtgcttcc agggactcgc
 10501 attgcaggct gggagtcagt tcagggttgca acacgtcacc ctccaagttg cttgaaggcc
 10561 ttagactgtg gtgcaaccag ctgctgccaa gagcatgtgg gtcacagtgg gtccccctcta
 10621 gctttatcat agactcatac tttctcccct cccctccca tccccacagg tgagaagcct
 5 10681 cacaagtgcc aggtgtgcgg caaggcattc agccagagct ccaacctcat caccacagc
 10741 cgcaaacaca caggcttcaa gcccttcggc tgcgacctct gtgggaaggg tttccagagg
 10801 aagggtggacc tccgaaggca ccgggagacg cagcatgggc tcaaagtgc accctggctg
 10861 gctgcaagca gcagctacac aacactacag agggcagcct ccctgcttgc caccactctg
 10921 ctccctgctt gcctccactc ccttctgact ttccagaccc caggtccagt ctgcagatcc
 10 10981 taccagggtg ctccctcttc gccttacctc ctggagctgc cagaagaaat gaggtacctt
 11041 ttcaaagtgc agccgagagt gagaaccaag tgactctcta ggcttcggac acaaataaggc
 11101 tcctctacac ctgaagacaa aggcaaagtc aaatggggac cagaataaat cttagacccc
 11161 acagtccttc ccatttccag ccctaattcta cagacaggaa tgcccttcag gtttcttccc
 11221 tccccctct tgacctacc cagatatttg tgtggaagag gaggaatcac cattacaag
 15 11281 gtggacaaat gctaataatt ttatctagaa agaagagtga gtgttaactt ttatTTTTTT
 11341 ccttctgggg ggtctgttga ctcccttctt ttgggtgctg cctataaatc ttggaggaat
 11401 catttctcct cctcaaaaac tgattcagaa actgacttgg ggaaggaatt taatactttg
 11461 aagtcagtga atgcaccatc gaggtaccc ccaagaagaa gcagaagaga agttggaat
 11521 gagaggggat tagaggctct cccttcagta gggctgtgaa aacctcatca ctggaggtaa
 20 11581 aagcacaagc aatgcctgtg gacaagatgt cattcattca ctacagaaat gttcatggat
 11641 caccggctac caaggtagca ggccaccatgc taggtatttg ggaagagaga ctgaagtcac
 11701 aacccttgac tgctcctcaa aagctaacgg ttgcacctcc aagtggtggtg gtctgttctt
 11761 actcttgagg ggaattctga gaagacagca cagaattgta aaccttcctt tttgacctt
 11821 ttggatttta tcagggtgaa acaaaaagct gaacagttac ttcaaagata tgtgtgtata
 25 11881 ttcatgtttt tattgttaag ctgatatttt aaagatttct gagctagcag gcatgtggga
 11941 aggaaggctc tgtcttcaac tctttgacct tccatgtgta ccatagaggg gggaaagggtg
 12001 gtattttcac tttgatgagg ttggtaaatg ttttagatc ttctggtaag cattatgttt
 12061 gttaatacat atttattaga gtgatgtttt aagttaataa agtattaaga gtatta

30 *Neurotrophin 3 (NTF3)*

The NTF3 gene encodes the neurotrophin 3 protein, and has homology to sulfate transporters. NTF3 is expressed in inner hair cells and in surrounding supporting cells in the adult cochlea. NTF3 supports connectivity to spiral ganglia-like neurons (SGN). NTF3 induces synapse regeneration and SGN protection after damage (Wan et al. (2014) *Elife* 3; Budenz et al. 35 (2015) *Sci Rep* 5:8619; Suzuki et al. (2016) *Sci Rep* 6:24907).

The human NTF3 gene is located on chromosome 12p13. It contains 2 exons encompassing ~63 kilobases (kb) (NCBI Accession No. NG_050629.1). The full-length wildtype NTF3 protein expressed from the human NTF3 gene is 270 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype NTF3 protein is or includes the sequence of SEQ ID NO:

- 5 26. Non-limiting examples of a nucleic acid encoding a wildtype NTF3 protein is or includes SEQ ID NO: 29. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 29 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype NTF3 Protein (SEQ ID NO: 26)

10 MVTFATILQVNKVMISILFYVIFLAYLRGIQGNMMDQRSLPEDSLNSLI IKLIQADILKNKLSKQMVDVKENYQSTLP
KAEAPREPERGGPAKSAFQPVIAMDTELLRQRRYNSPRVLLSDSTPLEPPPLYLMEDYVGSPVVANRTSRRKRYAE
HKSHRGEYSVCDSESLWVTDKSSAIDIRGHQVTVLGEIKTGNSPVKQYFYETRCKEARPVKNGCRGIDDKHWSQCK
TSQTYVRALTSENKLVGWRWIRIDTSCVCALSRKIGRT

15 Mouse Full-length Wildtype NTF3 Protein (SEQ ID NO: 27)

MSILFYVIFLAYLRGIQGNMMDQRSLPEDSLNSLI IKLIQADILKNKLSKQMVDVKENYQSTLPKAEAPREPEQGEA
TRSEFQPMIATDTELLRQRRYNSPRVLLSDSTPLEPPPLYLMEDYVGNPVVANRTSPRRKRYAEHKSHRGEYSVCD
SESLWVTDKSSAIDIRGHQVTVLGEIKTGNSPVKQYFYETRCKEARPVKNGCRGIDDKHWSQCKTSQTYVRALTSE
NNKLVGWRWIRIDTSCVCALSRKIGRT

20

Rat Full-length Wildtype NTF3 Protein (SEQ ID NO: 28)

MSILFYVIFLAYLRGIQGNMMDQRSLPEDSLNSLI IKLIQADILKNKLSKQMVDVKENYQSTLPKAEAPREPEQGEA
TRSEFQPMIATDTELLRQRRYNSPRVLLSDSTPLEPPPLYLMEDYVGNPVVTNRTSPRRKRYAEHKSHRGEYSVCD
SESLWVTDKSSAIDIRGHQVTVLGEIKTGNSPVKQYFYETRCKEARPVKNGCRGIDDKHWSQCKTSQTYVRALTSE
25 NNKLVGWRWIRIDTSCVCALSRKIGRT

Human Wildtype NTF3 cDNA (SEQ ID NO: 29)

atggttacttttggccacgatcttacaggtgaacaaggtgatgtccatcttggtttatgtgatatttctcgcttatct
ccgtggcatccaaggtgaacaacatggatcaaaggagtttgccagaagactcgctcaattccctcattattaagctga
30 tccaggcagatatatttgaaaaacaagctctccaagcagatggtggacgttaaggaaaattaccagagcaccctgccc
aaagctgaggctccccgagagccggagcggggagggcccgccaagtcagcattccagccggtgattgcaatggacac
cgaactgctgcgacaacagagacgctacaactcaccgcggtcctgctgagcgacagcacccttgaggccccgc
ccttgatatctcatggaggattacgtgggcagccccgtggtggcgaacagaacatcacggcggaacggtacgcggag
cataagagtcaccgaggggagtgactcggtatgtgacagtgagagtcgtggtgaccgacaagtcacggccatcgac
35 cattcggggacaccaggtcacggtgctgggggagatcaaaacgggcaactctcccgtaaacaatatattttatgaaa

cgcgatgtaaggaagccaggccgggtcaaaaaacggttgtaggggtattgatgataaacactggaactctcagtgcaaaa
 acatcccaaacctacgtccgagcactgacttcagagaacaataaactcgtgggctggcggtggatacggatagacac
 gtcctgtgtgtgtgccttgtcgagaaaaatcggaagaacatga

5 A non-limiting example of a human wildtype NTF3 genomic DNA sequence is SEQ ID NO: 30. The exons in SEQ ID NO: 30 are: nucleotide positions 1-229 (exon 1) and nucleotide positions 62081-63186 (exon 2). The intron in SEQ ID NO: 30 is nucleotide positions 230-62080 (intron 1).

10 **Human Wildtype NTF3 Gene (SEQ ID NO: 30)**

	1	agt	tgaagct	cct	ctccctt	ccg	aacagct	ccg	cgcaccg	cccc	gcgacg	cag	ccccggcg
	61	caact	actttt	ctt	ctctctc	ctt	ctcttct	tcct	ctcctt	ttt	cccctgc	tgg	gtagtgg
	121	ctg	cggcggg	gtg	gggggaga	ctt	tgaatga	ccg	agctcgc	gtc	caccttt	ctt	ttcatgt
	181	cgac	gtccct	gga	aacggcc	aca	cggatgc	cat	ggttact	ttt	gccacgg	taa	gggggagg
15	241	cggc	ggggcac	ctt	gggtggg	cag	gtttggg	gat	gggggtc	cac	gtgggga	ggg	attttcc
	301	agt	ggactgg	tgc	ggggggc	ccc	agatccg	cat	cccggcc	cac	ccccatc	gcg	ccgcgct
	361	cact	cacttt	ccc	gggcttg	tgt	cttcccc	aa	gtttgcg	ctg	gggatctg	ctc	aggccga
	421	agc	gcaaccg	cag	ccacccc	gct	acacaca	cac	acacaca	cac	acacaca	cac	acacaca
	481	cac	acacaca	cac	agacacg	gac	acccttc	tcc	acctcct	ccc	ctcttgt	cc	tcggctg
20	541	ccc	aagaagc	ttc	ctcaat	ctg	gggaaagt	gat	cagggtt	aag	ggacctg	gatt	ggaaaag
	601	ggt	gggggca	ga	agagggga	aat	gggggacg	acg	aaaagagc	agg	aaaagaga	ttc	aacagaa
	661	tca	accaccc	acc	actccca	acc	gacccgc	ctg	ctcctcc	gag	aaagctc	ctag	cgcatc
	721	ctata	acaaaa	agg	gggtggc	ag	acagaact	ccg	ggcgggg	agg	tgcgcg	gcag	ctcccc
	781	tgc	acacgcc	ctg	cactctg	ccg	gcgcgtg	agc	ctgattc	tc	agctcgcc	cc	agcaccac
25	841	tct	ggcccgg	gcg	tgggctg	ggg	ggagggg	acg	gcgagct	cag	gacccgg	aac	ctcgcg
	901	tcc	agttttg	gg	agttggga	ctc	actgcca	cgc	ccgcgct	ac	ctgcgtt	gag	ttccccg
	961	aa	agggtttt	ttc	agaaaaag	ac	ctcgcgcc	ccg	gggtcct	ctt	ggccagc	gccc	acccgg
	1021	tgg	ccacccc	acc	ctggggc	ttt	gcgcaga	tgt	tggagct	ccg	tacgcag	ccc	gcacatc
	1081	tgg	gacccct	ccg	gggagcg	gcg	gggcaccc	ggg	cccggcc	at	cccagggg	at	ctcctgc
30	1141	ggt	atcgtcc	agc	ctgttct	cgg	actttga	gcg	gtggcgt	ggg	aggccgg	gag	acctggg
	1201	cacc	cgcgca	gcc	agccagg	tcc	gagttta	aag	gtcccac	gac	ggaccga	act	gtcccat
	1261	tg	ccccagag	ctt	actcag	tgg	tggatgc	tc	ctgatgaa	att	tgggacg	ctt	gggagtt
	1321	ga	aggttagg	gac	aggaggg	gcg	aggggccg	agg	gcatggg	atg	gggggag	agg	attctgc
	1381	ttg	tgtctct	ccg	cgggag	ggg	tgcgcgt	cc	aggaggcg	ctg	cttcttt	gcg	ggagttt
35	1441	ggc	tgtctgc	ttc	attcgct	gtc	tgcgctt	cag	atgcacg	gc	actgagac	cct	tgcgtcc
	1501	gac	ggtgtcg	ggg	ctgtgga	ct	agaaaagga	tcc	ctttttgc	tg	aatcgag	gct	gggggtg
	1561	gatt	gccgg	ggg	ggaacaa	ccg	aaaagat	cgt	ctggcct	cgg	cctctgg	cgg	cggggcg
	1621	cag	gttctga	gtc	cgaatgg	agg	ttgctcc	cgg	gagcgcc	ggg	ctcagag	ct	agagagct
	1681	cgg	gagactg	tgc	gcctgtg	gac	ttgttta	tgt	gtgtgaa	gag	gcggggg	cg	agggcctg
40	1741	ctg	agagggg	agg	ggagcct	gga	aggggtg	ggt	gtttctc	ctg	gagcctg	atg	tttgtaa
	1801	ctc	agctgat	tat	ggagtgc	act	gagcgac	ctg	cttttta	aata	aaagagg	tg	ccgctcc
	1861	tacc	ccgcaa	aac	agcgaac	gag	gagaaca	tg	gaagcgct	ctg	tcctaaa	cg	tcaggatg
	1921	gg	agaaagt	gta	acataga	gg	agactttt	cc	agaggtcc	tg	ttttcaca	ac	actcagaa
	1981	gtt	ctccagc	gt	actcagcc	tg	cctccgc	ctg	ccctcaa	ttc	ctttttg	ac	atgtcaca
45	2041	caa	agaaagc	tg	aaaggtag	aat	gtggagg	ata	aaactcca	ac	ccctctg	cct	tggggcg
	2101	aa	acacacag	act	taggtat	cgt	gttagga	gg	taagggg	ttg	gaaaata	atg	caggctc
	2161	cgg	tagacag	tgt	tgaagg	aga	tagaaaag	tct	gggggtat	tt	ccctagg	gag	agtgtgt
	2221	ggg	attttgt	gtt	gtgagg	act	ggagtgt	ag	ctggactt	aga	gggtttg	gtt	gtgtgtg
	2281	aat	gggatat	gct	tatctat	gg	agtgtgat	tgt	gaccatt	gag	tgagtgtg	tc	agggagg
50	2341	atg	gggatgt	ttt	tccaaag	tgt	gtgtgtg	tgt	gtgtgtg	tgt	gtgtgtg	tg	taggggaa

	2401	ggtggcaaaa	gagctggagt	cctatccctc	tttgggtgcct	gggaggtgtg	tgtttctggg
	2461	gagtctgtcg	gatgggcac	tgtgtgctta	gcgggagaca	agagcttgca	ggaacatagg
	2521	tgtagctaata	gcgaaatctg	tgcctatgaa	gggtgtcaga	gcaggaggag	aggtgcagcc
5	2581	tgttgaagac	agatggcttg	tgcattgcctg	cacaaacaca	tggagtgggt	gtggtgtagg
	2641	agggcggtcg	tgtgtgtgag	ggaaagggtg	atgcatgttg	aaggttgggc	gcctgaggat
	2701	taccaaagg	atgtgtgtg	gggggtgggg	aagattatct	gaggagcacg	gggtttgaac
	2761	ttcatgagg	tgcactggga	aattttggctg	gaatggcact	ggggcatgta	gggggtagaa
	2821	aagtgtgagc	tcagatccctc	tttatggaaa	gacaaagcat	ttgcagggaa	ggacaccggg
10	2881	ttgggtgttg	tgtttgcctc	ccccctcctt	cctccccaga	gagaagagaa	ttggaggagg
	2941	cttagaatga	gggatgtaca	tttaaggggga	aaagtccttg	tgcattggca	ctgggtggga
	3001	acagggaag	atgtgtgcat	aagactgaag	taagtagagg	aatgcactat	ctgtgcctgg
	3061	tgtcagggat	gcatgttgta	taatgttttg	taatgggttt	cttatgtaat	aggagaattc
	3121	actatcacca	gttcataattg	ataatcaatt	aaagctaata	gagtgccctac	tggatgctca
15	3181	gcccattgct	gggtactggg	gcatggcaca	ctgccctgga	aaaattcctt	ctccaggctc
	3241	agtaaaactg	atgcaattttt	gccagaaata	gacagagctc	gacgcttggt	gaggtttctt
	3301	caagggtctat	aattcagttt	ctagagatgt	ttatggatct	tggggcccaa	agggaagctg
	3361	acttgaaata	tatctctggt	tttgtaaaaa	gtggggcata	aagagaggga	tgatgtagac
	3421	ttttgtgtat	tgggaaggga	gctggacggt	gttgtgtgtt	cttgtgtggg	gtgtgggctg
20	3481	ggaattggag	ggtgctgcac	ttcaggtcag	aatgcaggtc	cagatccctc	taaggattga
	3541	gtttctaaga	gtacacagag	cctccttcat	agctggaatg	agacctgtta	tgtggagctg
	3601	gagaatggtg	aagtcattgat	caagggaag	atggactggg	agtgtttgtt	catctttgat
	3661	gttttctggg	ctacgagtgt	aaagctaacc	tgagtgtgtg	ttgggggggt	gggtatgcat
	3721	gcagggtcgt	gcacatgcac	tttggcctgg	agtaacgtaa	ctgggagaaa	actagcgagg
25	3781	tggataggag	gggtggatat	tctggctctt	gtgaaggggg	tgccagtatg	tgtttttgca
	3841	ggaattctgt	gtgtctgggg	aaggataaaa	attgtcttgg	atactcagat	gtgcatcttt
	3901	agcgaaaaaa	ggtgtacaca	aaagaagggc	cgtagacttt	ggggaatgag	acaaagggtg
	3961	cgttgatcgg	gtccataccc	tgagatggct	ggtgaggcaa	ctgggtagga	acagagtga
	4021	ggctgtgctt	ttgtgatgca	cacatctgcc	tgtctggagc	agacgatggg	gatccccatt
30	4081	tgattaggtt	ctgaccacg	ttgggcatat	gtttgggagg	actttatttg	actcatgaca
	4141	tttttgaatt	catgaagtta	cagctgtgcc	attttaaggg	tctaccttat	ttttgagagg
	4201	gaaactcatc	tgtgtaagag	atagtttata	tagggcttga	ctatttggtca	gaagatccgg
	4261	ggcagatatt	tctgagagca	gggtggtgaa	gtgcaagact	ggattgggaa	tgtcttagtt
	4321	ctagtccagt	tctaccactt	gggaaaagt	tgactttggg	gacattttta	atctctataa
35	4381	ctcttctgtt	tctctttttg	taaaggggca	atataatgac	tcaaggagaa	ggaaagttca
	4441	ctcctctttt	cccccatcaa	aacttgatgt	gttttagttgt	atacagccat	acaaagatga
	4501	aatgtaatta	aatacagggt	ctagcacgta	tgtagggtcat	tgtgaaaata	ggatctggac
	4561	ccctttctct	ttacaatgca	acttaatcag	taatcatgat	gccacagca	tctctactgc
	4621	tcactcattg	tgtaataaaa	actttttata	actgcctcag	tttactcatc	cagaagctgg
40	4681	aggcagccaa	acctagatta	tagatgatgg	aaatctttta	gatctttgag	ctattcagag
	4741	gaaaggtagt	ctaagaatca	atgtttattaa	tgatcagaga	cagacatttc	tttatttttt
	4801	aacatcagct	atgcagacca	cagcaccaaa	agaaatgtga	cgtggagaac	agcaaaaata
	4861	ataattgccc	ttgacatagg	caccatttga	ataaaaagaag	agcaaagtct	agcctgcctg
	4921	gtaagaggaa	agcctgggct	ttgggtcttt	gaaattcttt	gtgcccagtg	cagcaagttg
45	4981	ggtacagcca	ctttccccag	agatgaagat	ggagtcgtag	aggggcatgg	aaacacactg
	5041	aaacctccga	ttcctcccag	aagtcagtac	atgctgtcaa	ctcctggctc	agagtcaggc
	5101	tcccatcctt	gagttcagtg	gccataaata	gtcaggagct	aagctatttc	caagaggcac
	5161	taaactaatt	ctgttgcttt	tcttggcgag	attgggccaa	ttaagttgat	tggagagggt
	5221	cactgctgat	gggtctcaca	ttccccaggc	gggacaggcc	agcagagttt	aattacataa
50	5281	catttcccat	gtcttcactt	ccactcccag	atgtatcaga	tgttcaggcc	tctcctcttt
	5341	ccctacattc	aggatgcctg	acggagcagg	cggcttctgc	tacaagcttc	tccagacctt
	5401	ttcctgagca	gttacagcac	tcccacccca	actggaggga	atgcaagtct	gtgcattgcc
	5461	tggaaagtga	agaggggatg	caggctgaca	cagaagatct	ttagagccca	ccaccctgcc
	5521	tcaagcctca	cccgcggctg	ttttactctt	taacgaggca	ggcagtttct	ggcgttgggg
55	5581	atgttgaact	aatgaccgtt	gctagggttt	tagatttagt	tttattttaa	aatgagtttg
	5641	ttggaaaagg	accaaattctt	ctactggaga	caagcctaag	ccccgtgtgg	tgctgctgat
	5701	aaggctctgtg	gatgtacttt	gtgcatggac	aggcatgtgt	ggagtaagag	gagagcacag
	5761	gatgtgcgtg	cgtgcgtgct	gggggacacg	tgagcgtagg	agatgccttg	ggtagatggg

	5821	gcattttgtag	agcttatgtt	atgggtggca	ggatatgtga	tgcttggtc	agattggatt
	5881	gtcattggat	tgcttcttac	aggtgggaca	gcctgagaaa	aaagttgggg	acaccttct
	5941	tcagatcatt	tttatatggg	catgttgaac	cttgaagtat	gagccagtg	tagatattac
5	6001	ctcattcttg	cctggatgga	tttctaacaa	gattgaggct	ggacaaaaaa	aaaaaaagtc
	6061	tttcagagga	ctcatttttag	aacctgtgag	aacacatgta	gatggcccca	gagcgttaaa
	6121	gtattttttg	ttgttggttaa	tagttttcca	taactcagtc	ttttgtttta	cctaaaaata
	6181	ctgagaggct	agaggttgca	cacaccccca	catggtcaga	atgcatttgc	tgttcatgtc
	6241	cccattctgtg	ccatggtcct	tctcaaatag	gattaattgg	gaactacgga	gagcttgccg
10	6301	cagcttttgg	tgctctcaca	ggacaaacct	atgactgacc	gtgatgtcat	tgaaacagaa
	6361	cagcttttga	cgtgaaatgt	cgtgtggtct	gttgtgattg	ctcacttttt	agtaacatct
	6421	ccagagaaga	catttggtga	gccccatttt	tttttggtgg	tttggttcagt	cagagatttg
	6481	ttggaagatt	taaggttctc	caacattcaa	ggccagagat	tttgaggagaa	acagtagcta
	6541	acaagagccc	aggccagggt	gaccttcagt	agcctgaggc	agcaggtagc	tgctgactt
15	6601	agaagtgaag	gaccaggttc	agtgactggg	ctctgtcact	aattagctgc	tactgtgtga
	6661	ccctggacat	tctctgtagg	acatgacttt	gccatctgta	aatgaagggc	ttgaattaga
	6721	agatctccca	gtcccatctc	agccctgtca	ctgtaggatt	ccacagtagc	cacatcctcc
	6781	tacctaatca	gtgctaatta	cctcacttgg	agtgcaccct	ggcacgccct	taacgagctc
	6841	ctgtttttat	ctttggggaa	tggtgtgagc	aaaggcagaa	aagcgtgcaa	aaaatttgcc
20	6901	agagtgtttc	cttttgcca	agaatgctct	aaacattggc	ttagcctaca	cagggtgtgc
	6961	agagctgaaa	gggtgtggat	ttggctttct	ggtaggtctg	atgctaatac	ttcagcttcc
	7021	acgtgaatgt	ctttcaggtg	caatgtctct	gactaaggta	agcatccatc	tttaagatga
	7081	tggtgatga	tgatgatgac	agttaatat	agagtactta	ctgggtgcca	ggcattttta
	7141	cagctattac	atggcattat	ctcagtactc	aaaataatct	catttacaga	tggtgggtact
25	7201	gaggttatgt	gaggttaagg	ggtatgctga	tggtcttggc	tttccagcct	gaagtctgtg
	7261	gcattcccaag	cccattcacct	gctagtccga	ttctgttttt	tatgtctgtg	atgtcatcca
	7321	gatgggagct	gggtccatct	tcttctgcaa	agttggtcac	cccattaggt	aacctgctcc
	7381	ctgcccccg	aagcctcctt	gtcatgtgca	agacaaatgt	ttaggttttc	tagccacctg
	7441	ctggcaagca	tgtgtcagat	ctgtctcagc	ctccattggc	cctctccctt	cttctctgt
30	7501	tctgttgaca	agggaaacct	cccttacaga	atacggatat	tgatttaagc	cacaattgcc
	7561	tccttgtaaa	ctctataaaa	ttggagctgg	acatctgggt	ttgatccaca	cgcccacgag
	7621	aggtggtagg	tcccagcatt	cttgagtcct	ggaattgcac	agaatcaaag	tgaaaaagaa
	7681	tatgaatgga	gagcatgagg	tatctttaga	atttggaaga	tggtttctct	tgataggcta
	7741	gggaactcac	accatttaga	ctggggccaa	agttaaggta	ggtttcttgc	tggtgaggtg
35	7801	gggggcagg	gtcactgcta	aagactgagg	tccccaggct	gtgtttgatg	cagagcagag
	7861	cagctgtaga	gccctccagc	tgaggatttg	ccactgcaca	ccattctctg	ggcagtacca
	7921	gaaatagaga	agccaaacag	gttcttgggg	tggtggttgc	cctgtttgaa	gattgtgact
	7981	aaaggctcat	tccttcttcc	tccactgatt	tcctcaccat	cactaccaca	tatacatcat
	8041	caagacagtt	aaggacaggt	gactttgttc	tggtgcttaa	acataattgc	atgtacaaat
40	8101	aaatgttgaa	cactaaccta	atgcctatta	gaccataggt	actcaaaaat	tatgtattta
	8161	ctttgcttaa	aataaccttt	catgtgtagg	gtgataattt	ttttcttcat	atggagtata
	8221	tagcctttta	aaaactgata	gccaggttct	tgcaattttac	ttcccccttc	cctctgtttc
	8281	tagatctaag	accccagcat	ttaaaaatat	ccttctctgg	ctctaagtga	tccctcaggg
	8341	agtttggttc	aagttctcaa	tctgagatta	gtttcatccc	acatctagag	gacctgcttc
45	8401	taatggttca	taaggtactt	ccactgggtt	caagagggaag	acacagctct	agaaatcacc
	8461	tgctttccca	acaggtgtgt	cccattgggtg	ggaagaagcc	cagtcctcct	aacacccacc
	8521	tggttacctt	aaatagcttg	ccacagtctc	cctgcagaga	gcccagacta	gctgtgtgac
	8581	tttaagcaaa	tcactttccc	tccgtgggcc	tcttttcttg	gaaaaggag	gagttggact
	8641	cactagagg	tcacaaggac	cctaccagtg	tgacagatca	catcttccat	gactccgctg
50	8701	ctccctcgcc	cccacactgt	ggcccatgcc	tggtattttc	aggtaaccct	gctggcctag
	8761	aggacttaca	tagagtgatt	tggtgggggt	ggaggaggaa	attaagtatt	gtgatgccag
	8821	ataagtcatt	gaacatacgg	atgagtgcac	gtttatgaag	cactgaagaa	acaagcttgc
	8881	cagattcgtg	gcggatcttg	gattggaagt	tggtgagtatg	cgagaggttt	atgtgggatt
	8941	aaagttgtct	gtgggtagct	gaattcaata	agctaaatca	gagagatccc	aaactctgca
55	9001	gagagatcag	acccgcctc	taactccaac	agaggctaata	tacaattttg	ggaagtcctc
	9061	gaaccacact	aggtctcact	tcctcatcta	ttaagagcta	ctgactttga	ttatctctaa
	9121	ggtctttctt	gacttatcat	taattaaatt	gtgagaattg	ggagctagta	gaaaaatctg
	9181	gacaaatgcg	tggtatggaga	aggggatacc	tcaaatggcc	taaatataga	gggaaatggt

	9241	ctgtacaagt	gatacgttta	ttcttgaatc	ggagatagca	ttggaggata	tattctgtac
	9301	aaatatgtg	cttcaaattgc	ttacccctcc	atcatcatca	atgatattaa	aatcattgct
	9361	taacctggga	tatttgctta	acctgggaag	cctccctgca	tttcccaga	gagcctaagg
	9421	ctcccttccc	agcctcctga	tgccacttag	cagcctgtgt	ttaatttgca	ttcctagaat
5	9481	gtgacttcac	agagaaaagg	atctgttggt	ttcatttttg	tgtcccccac	cttgaatagc
	9541	gacagccaca	ggaatgtgtt	gagtgaatga	atgcgtgacc	agccattgtg	agctcagaat
	9601	gtgcagcagc	gatttgacag	tgccaggaac	ttggctcaga	gaggcagtgc	accttgccag
	9661	ggcttcaaaa	cagtgttacc	tttatccccc	ttttggcctg	tcccaaggct	aaggtcatga
	9721	atgctcagcg	cttggtggcc	agaaaaggct	gttcttactc	ccttttagag	acctttcttc
10	9781	cccacatctt	agcctgctta	gttctcctgc	aaattggggc	tcttggtcac	agggaaatcg
	9841	gtgtcattgc	atacgcat	ggataactcc	tcgttagaca	ctcggaaagc	aactcaaaac
	9901	acacgtgtac	agattattca	ttaagcagtc	tggtgtgtgt	cacgtgctaa	atgttgggga
	9961	agatgcagag	atgaatgtag	tgctgctggc	cttgggtgct	ggaaactggt	gaaggaggca
15	10021	gtcaggttac	aatcatagtg	ctccaggag	aacaaggga	gcaaggaga	tggaatacac
	10081	ggttcccact	tcaggtgttg	tttttttagta	caaaccatga	ttgagaatcc	tcttggttca
	10141	ggagcaagct	ggaggggtcca	ggaatggtct	ccgcttccac	atgacagtgg	ttttcttgtc
	10201	cctgaattcc	cagggtcagg	gtgatggatg	ctcagagacc	attgtgatgg	gagcccttgg
	10261	tccttgtggg	gagagtcggg	ggaacacgca	tggtgtacag	gtgttggggc	ttggggaggt
20	10321	gtagggtgtc	atgtctacca	gggtacttct	cagaagtatc	cctgtctctg	gcattggtgg
	10381	ggtaggtcca	cctcaggccc	ctaagctgtg	cttctacttc	ctttctaatt	atacagcatg
	10441	tattgtctct	gtgaccactt	ctgacacagg	aacagactgc	catctactgt	ccacttctgt
	10501	cctgagtccc	tcctgtctca	ggaggcattt	attttttctc	atcctgtttt	taatagatcc
	10561	ggacaccagg	agaaacagac	gtggagtcct	atttagtaat	tttttttagaa	tttagaaggg
25	10621	agccttcaaa	aaacatagca	catctgtcag	atggtttccg	tatctgttta	caaagtatta
	10681	ttgtaagggg	ccactctgct	gccactgagt	cttgctagcc	gctgtgagag	gtgccgtgtg
	10741	ggataccaag	agatacaaag	caaggcctta	ccttcaacag	ccttgacgtc	tacttgcgaa
	10801	tacagtgtag	atttgcagag	aagcagacta	caggatgtgg	cagtagatat	ttgtaagag
	10861	gcatctgggt	ggtaggagct	acagaggttt	tgggagtttg	gaggaaggag	gagcccttat
30	10921	aggctggggg	atttgggaaa	gaccaatatg	ggtctccttt	tagcaatccc	agtgttttag
	10981	aagaggcttc	taagagctgc	gcgatgttgt	gggacaggct	atgaactgga	aactcaggaa
	11041	cctggagggg	ctgagggtcta	gtccagcttc	tggcagtcct	gccctgtgac	ccagggcctt
	11101	gggttccgca	tcctcaaagc	aaggctgttg	gtcttcttgt	catgtaaggg	ccccccatc
	11161	cccgaccctg	catcatgtac	cactcttgga	tgagtgagg	ggtggtttcc	tggtttccag
35	11221	cttgtcaaga	gcaatagagt	caattggatc	ccatgcagga	aggataacct	gatgcagggc
	11281	cctgtgctga	tgaccccaca	tgacagcagga	gagagagaa	ttccccaggg	gagttttggg
	11341	tgctgttaaa	atacaaaggg	gggaaatgtt	ccatgcaacc	tctgcatgcc	actcctctcc
	11401	agagctttcc	tggactttcc	ttccaaaaat	atatgtgcat	ggtgctttgc	gtccaagata
	11461	gctaccccaa	aagaaattgt	attagatttc	tagcattgtt	catagctcaa	atcctaagatt
40	11521	ccttcacatt	agaattcatt	ccatgattgt	ttttatggag	gttttctttg	agtatttgag
	11581	actttttgta	gctgtggatt	aaaatgagga	aaagattgtt	ttctatagct	cttttactcc
	11641	ttgtctcttt	ttctctgtct	ttcttagcat	ttaaaccacc	accaaggaaa	cctccagggt
	11701	gtttattgca	tattcttgcc	ttttgatgtg	tggtgtgtga	tgagagagag	agaaaaaaa
	11761	agaaaggag	agaggagagg	cgagagagag	tacaagagag	aaagatattt	tagactgtgg
45	11821	tctattagca	tgtctaggaa	atcaagttga	atttggattg	ggttacaggt	tgagagacct
	11881	cagctcgtca	ctgtaatcct	ctgacatttc	ccttttactt	tttggtctct	gtctgaatca
	11941	agaatacatc	ttcccgttcc	ctcttccatg	tttacatctt	cttttggggc	agctgtataa
	12001	agttactgtc	tgtctacagc	aagtctcctt	aattcttttg	gggactgtc	gacagaggca
	12061	cagtcaagga	tagaaccatt	agagacggtc	agcttttgga	caaagacaaa	aaaatatgtt
50	12121	gaggaatagg	agaaatgttc	caattgtcga	ttatgtgttt	tcctgtggta	tgctgatgtc
	12181	tgtgactgtt	gtccacagaa	agagagcgtg	agcaggggct	ggggcagaga	taaaaggagg
	12241	attctcccca	aaccaacaaa	tgtggagaga	gagcaggagt	ccccagcaac	agatatggag
	12301	ccaaacaacc	ctgggaagta	actaggaaac	ttcctggtga	cctgagatgt	gttttgtaat
	12361	cactggtgaa	agtaggagga	gatggggtag	aggagagtta	caggtagaga	agttccgtgg
55	12421	gaggagacct	caaggtgaga	gaaagcaggg	acttaaagaa	ctgaaagaag	cccaggaggg
	12481	ggatgaacag	agaacatccc	atatatccca	tatccagtat	ggtggagagg	tgggcggggg
	12541	ccaactggga	gcttactttt	tttttcccag	gggcatggga	actatggaaa	actctgaggg
	12601	cacaattaag	cttttgtttt	taaaaagatc	cctccggctg	cagtgtggct	ctgaataaag

12661 agttgttcaa tgaagaaagg aatggctgat gggttgattg atggaagaat taatgaatga
 12721 gcaaacgaat gcatgagatg cagagggaac gcagcaagcc tgatttgatc tctggttact
 12781 ttgaccagta ctggctccag gggcgattat ctcagcctgg gaggccaggg agtaatgcat
 12841 tgattagaat gtctctggac acatggatta aaatatctga tatttttaggg tggttgatag
 5 12901 tggggaggac ttctgaaact ttttccctct tctatgcatt tccatcctaa tttgatttca
 12961 ttcaggatca aaaaagaaga tggcttttga aattacatcc cagaaaatct gaaactgtgg
 13021 cattgacttg ctccagagag ggctgcttgc atggaagacc ttttcatagg ctcatcgtgg
 13081 aatagggaca gatgataaag tttcttgggc atatgaaggg gtcccagatt tctggacgtc
 10 13141 agatccaccc atagatgatt ccttggatta aatgatgtgt gtgtgtatgt gtgtgtgtgt
 13201 gtgtgtgtgt gtgtgtgtgt gtgtttaaag ttttaaagtc tctgaaaat taccaagtgc
 13261 tactgaacat ttttttgtc agtagtggat atctggataa tttctttcaa ggccacattg
 13321 cttagcatgt ataaggaaag tgtgtgcggc agagaccag atggacagcc ggcatgccat
 13381 ccagttgctt ggggagtgtg aatactcctt ggcaaagcca aaggagacca aaggagacct
 13441 ctagtgtctt tgctctccca tgtcctgact attccaatct cactttgcat tttgagatct
 15 13501 ttttaactttt gctaacagtg attgctcagg atgtcattca ggccaattaa atttgaatgt
 13561 ctagggttgg gacagaagct cacgagggtga atgcaatgtg cagccacgtt gagaaccact
 13621 agctgacact gctaactctac tctgctctcc tccctgcctt ttggcggctt tgccatgata
 13681 tctgtcccat cctcacctca gtttgcctg aactcctcac ttagttaagg aagttcttct
 20 13741 aaggatgttc aactaatatg ggctaaggcc tccatccccc taaaaatcca gactttgcct
 13801 gagaaattgg acgctaggag gataagacag gttcttagca gattctgaag cactcacctg
 13861 ctctcatctg ctgaaggggtg tattgaggat gaatgatggg gaacagcagg gagaaagacc
 13921 agcatccata gggcagtaac aggggtgcaga cacctattta tatgtgtcat cgtagttccc
 13981 ctgacagctc ccagaaccag atattatacc aggcagatat accaggctta caggcaccta
 14041 tttatatgtg tcatctcctt agtccccca acaactcaca gaggcagata ttataccagg
 25 14101 cttgagggca caacaactca cagaggcaga tattatacca gacttgaggg cacagtttgc
 14161 ataatgcatt cataatgcaa agtggtttgt gatcttgaa tgatcagact tcctacactg
 14221 gttctcccca gtaatttgca tatcagccag ataccctcct ctagggttag ctatgatcag
 14281 ggctgctgat ttctggcact gatggctcaa tgggaaggaa agcgcctatc cctctgttc
 30 14341 tgtgggattc agcaacttg cctttatatt gagaaaatag tagttctctt ccttgccatt tccctgtagg
 14401 aaccaacagg cctttatatt gagaaaatag tagttctctt ccttgccatt tccctgtagg
 14461 gaagccatgg cttgcaggac agccagggtg tagaggttca gagaaggagg ctaaaaaaat
 14521 caagacctga aaaagttaag tttccaacca tatattccca aattgtagag tggaccatat
 14581 ggaatcattg caatccctgg gtttctcatt tacaactttt gcaaaagata ttactatgta
 35 14641 tgctgtcagg catgctggag ccggatgtct agcaccatt gctaagtga tgcttcggta
 14701 agagcttgg acttgggatt ctccagactg ctgagtgaac tacgtgtagt ttagcaacat
 14761 aatgcttgg atgagccacg ccgacctgat gcgagaaccg gatggtagcc gaagagagca
 14821 ggccaggtag tgagacaact atgatctgca gaggtgggca gcgggtgga gacaggcttc
 14881 ttctgcacca agtactgcta cctgaactgt acatcacaga gacaagccct gccgagagca
 40 14941 gtggaagcag gaaaacacag gcctatctct tgtgtgtgca tgtaagtaag tctctcttgg
 15001 ctgtgtgtac ctgccacat caccatcc caccacaaat atacagcatc accttcagct
 15061 atagcttttg gtttttgatt aaaggacagt aatatttcca ggaggagaa gagtgtgaa
 15121 cagaagcatt tatgtatgga cactgggcaa tttattgttt ttcaaagtct aatctagaac
 15181 atgtttgcaa gaaagcgtga attgagtaaa aagtttacta tcgtattaag gactgttagg
 45 15241 tacaatgact gaaggaggag ttaagagtga tccttcaata attcccctgt gggattatgc
 15301 acatttaagg aaaaatgttc ttcaggctat ttctgcctta gagctaggca tcattacagc
 15361 gaaatagaga actaaccttt aagcaagagg aaccaagttc tagttctatt tctgtcaatc
 15421 actgtacgca caacctctca aagccttagt ttctctttg caaagtggga taataaacc
 15481 tacctactta acacaatgtg gggtttcagg tgagatgat cataggatcg tgcttggaac
 50 15541 gctgtaaatc tgtaaatatt aaatatatat tatggtttca actggtacat tcctaagcga
 15601 atagcacatt gctctgttgg gaagacggct cttctccaag tcaggctggg ataagtgtcc
 15661 ctgacaagac actgccatac ctagggtgtt cccaaacatt gtctctggga accttgagga
 15721 agcaccataa gacatgggaa gaaatgttac agcgtggct tgaaagaata acaatgtatc
 15781 agtctactta tttctgataa tgtcatcttg ggataagaga ctcagggtag cttagttagg
 15841 gacatgggca tgcactgcac agtaaaaatg gtgtccagga aacctgggtt tatttcagta
 55 15901 tgggttgccc acacttctgc caaccagtc ccctacttcg tccccagctg ctcttgatga
 15961 actctctgca cacacttgca cctgtatctg tgaaacagag ctccctctct tacatgagaa
 16021 tggatctggg tgcaaatcta atagattccg ctaccacaat gtcccctgcc ttttttgttt

	16081	acttcatttta	tgaaaataacc	cttgaaacat	ccatagtccc	atthttgtaga	catgggtgctt
	16141	tatgtctttt	agattatttaa	atactcatgc	tccttttctga	ttgctgtttt	cacctcttct
	16201	ttaggcttgg	gcttttctat	tggtggaatt	tgctgttcct	tttcatgggtg	ctggctttcc
5	16261	tgagtgggt	ggttgctaata	tcttattgcc	agtttgtctt	ctgtgacaga	ttctccgaca
	16321	tgcttcggat	gtggcttccg	tgcttggtct	tagcttttta	tctgggtctct	cgtgtcctga
	16381	atattttacct	tctttcagga	atthttctaca	gcttctagcc	aatgaaaagtc	ccccttctta
	16441	ttgtcaaacc	cagacagtta	taattttatt	ttaaagtac	ttttgttac	attgttttgg
	16501	gagtagggct	ggaaggactt	gctaattggat	gagtgggtga	aagtgaagaga	aagaaagata
10	16561	tgatggaaga	aggagcaatc	aaagccccag	aaggaaagcta	tcgcattgtt	cttggattcc
	16621	taagcctgcc	agaaagagct	gacttacttt	acagtttctg	agagaactat	gtgtgctatt
	16681	actagaagca	caagagaaaag	aaaaatagga	tgttcagcat	gcttcatttta	tctaattgtga
	16741	aaaatgaact	ctgcccagtg	acttaattggg	cataaactct	gtttctaaaa	aagccactca
	16801	ttcggcaacg	cacttctgag	ttcttgatat	ggtaaggat	tgtgttctgt	gctggacagg
	16861	aatacaaaaa	tgacaggttc	ctcacctcca	agaacttata	gtacatgtag	ggaaataaga
15	16921	caacccttat	tgaatatcac	tcaagatgga	aatgactagg	gccatggaat	gtaacacaca
	16981	gagggtagct	ggagttccta	agacttctga	tggtatgattg	aggagagccc	tggtacaggt
	17041	aagcaatttg	aagggatagg	gatagcacag	acagcacagg	ggtggaaaca	agtgtggagt
	17101	gtcgagagct	tgaccagtag	gcctgaaagg	gagggagtgt	acacagagcg	ttaattaggag
20	17161	ctctgtctcg	aggcagcttc	cctcagcccc	tcccaggaca	tcgaggtttt	gggagaaaga
	17221	gcctattgct	cactctcacg	gctcttctcc	tttttctctg	ctttcagttt	gttctttgaa
	17281	cttttttgaa	acttcccctg	ttctttcttt	aacactgtgc	ttcatctctt	gggttcttac
	17341	gtttttgcagg	ttgtagtgtc	tgagatccag	ccttcccaaa	tgatttctct	gaatttagta
	17401	tttggtatgg	gttttgctat	tttgctgcca	tcccagccct	agcaaagaaa	cgacttacct
	17461	ggagtatgga	cagggcttca	gagaaaaccc	ctaacttcc	tgactcccga	ctttacagag
25	17521	ctctgccaaa	ccttgccctg	cgggagtaag	aaaagcgcta	acaagccatc	ctctttgggtg
	17581	tcaagtgcag	acaaatcact	tagccccctc	gaggtcctcc	aacagtaagc	tactggtttg
	17641	tgaaccccca	ggataatcca	tctgatttca	gtcctgcatt	tagtacttta	gaacattctc
	17701	gcacatgcat	ttgctggctc	atgtacatac	gaatatacac	atatccctct	ttgtgcccct
30	17761	tcttagcctc	tgatgatttc	ttctcctcca	ggaggcagga	atccaaggct	tataaacctat
	17821	gacttctggg	aagtthtttc	ttctgcttaa	ccagggtatc	attgtttttt	ctgccttccc
	17881	ctggagaatc	actggccact	gccctagtgg	ttggggcaag	gatcagaggt	agcttgcat
	17941	ctgggggtttg	tcccaaaagc	ctcgggtggga	ctctgcattg	gggtctgtag	cctggatcca
	18001	ctccagtacc	ttaactaatc	tcttgactcc	cagatgggtc	aaaatatgtg	gatttagaag
35	18061	agcaacagac	agctgttcct	ctgggcctct	ccaagaacac	ggtttgggtg	ctagaccacc
	18121	ttagagaaaac	atggcagagg	aaatcatggg	ggagcagcat	ggaaacaggt	gaaaccacga
	18181	cttagtacct	tgttaaattc	catcctggag	tgagataacc	agaggagcag	atattacctt
	18241	tattaactga	tagaaatggt	tggggatttc	tctgacttct	tgtagggttg	gataagcccc
	18301	aaagtgaaga	gaattttgct	ccttgthtta	gccattagga	aactcaagac	ctgctacag
40	18361	tgctattggt	ttaatttttc	cctatcacat	tgctctgca	acttctgaat	ggttgacgcc
	18421	atttcttaaa	atttccctgc	attgtcactc	agacaacaag	aatagatttg	gccttcttca
	18481	tctcaaaata	atggtcatga	ttaatagtta	ttggactggg	aacagtgtct	agccctctgt
	18541	acgtgatctc	aggaatcctc	acagtactca	atgaaatagc	aattttatta	tctcattttt
	18601	gcagacaaaag	caacggaaac	ttccacacat	tttctacatt	gcacctaaga	tcatctgaga
45	18661	aactatgctg	tacttgthttt	tctaattgat	gatctgattt	ttctattata	atgttaattc
	18721	tatgaggaca	gggctttctg	tggccttgct	tcattgctgt	atctccagca	cctggactac
	18781	tgcatggcac	ctggtagtta	cttagtaaaag	gtttttcaaa	tgactgagta	actcatccaa
	18841	gattaaatgt	ctaggaagtg	gtggcaccaa	gcttaggacg	actcttttct	gattccagag
	18901	tccagacagc	cctaaccact	atcccacact	accttcttgt	ttatttttaa	atcattttcc
50	18961	ttcccttcaa	tcccttccca	gtgccttaca	ccttcttgct	gtaatttgaa	gcatggccac
	19021	agtaagctac	ctcaagtttc	tcatctgtaa	aatggggata	atataatgaa	ctaccttatg
	19081	ggattgtacc	cctctgcatg	gtagcctcat	cctactgtgc	ctcctaacca	cggcctttaa
	19141	atcagcaggt	atagttaata	tatttagttc	ttttaatcta	atctgaaaca	caaagcattt
	19201	gcttccthta	ttcaagattt	ttggctttgc	ctagactaag	cttaaaacca	aagaagtact
	19261	gcagaactga	ctgaggctgc	cagaagtacc	acactcttgc	acccagccag	tggaagtggg
55	19321	aaagataaca	gctaagcctt	tggggatcct	tccagaagta	gtgatgacgt	acagcattct
	19381	ttctgattat	gaagtaaata	tctgttctaa	tgtatgttca	acatagagag	ttaagaaaat
	19441	ggggaaagaa	taaagagtaa	aacaatgacc	agaaatacct	tcaataccct	ttgacattct

19501	ttctctgtgt	gtgcatgtgt	ttgtgtgtct	ttgtttctgt	gtctgtatat	gtgtatttcc
19561	tttatttttg	tttttttact	ttaatgtaat	tttttagagac	aaggctttac	tttgtcttgt
19621	agactcgcgt	gtggtggcgt	ggcactcatg	gttctactgca	acctccaatt	cctgggctca
19681	agcgatcctc	ttacctcagc	ctcctgtgta	gttaggactg	caggcatgca	tcaccattcc
19741	tggctaattt	ttcaattttt	gtgaagacgg	gctctcacta	tgttactcag	gatgggtctca
19801	aactcctggc	ccgaagcaat	cctatcacct	tggcctccca	aagtgcctggg	attacaggcg
19861	tgagcgacca	tgcccagctc	cctttttataa	ataaggggct	caccatacaa	tataaccagt
19921	ttttacctgg	cattttccag	tcattattgc	attgtacgta	tctccccatg	tctttttctt
19981	ttcttttttt	tttttttttt	gttgacggag	tctcactctc	ttgccgggct	ggaaggcagt
20041	ggcgcaatct	cagctcactg	caacctccgc	ctcccgggtt	caagtgattc	tcctgctca
20101	gcctcccag	tagctgggac	tacaggcgcc	cgccaccacg	cccagcta	ttttgtattt
20161	ttagtagaga	cggggtttca	ccatgtcagc	ccaggaaggt	ctcgatctct	tgacctcgtg
20221	atccacccgc	ctcggcttcc	caaagtgtctg	gaattacagg	cgtgagccac	cgcgcttggc
20281	cctctccatg	tctttaatta	ttcttgcata	agatgacttt	tcactgcata	atattccatc
20341	acataccact	ctttaaccat	tttgcttctg	gggcacattt	tccttttggg	cacacttttt
20401	atactacagt	tgccatcctt	ttacatacat	ttgaatacac	atatctggct	attctctcag
20461	aatagatttc	cagacgttac	ctttccaagc	ttgaagatgt	taacatttta	aagaaagatg
20521	gttattcttg	aaagccctga	cagctctgag	tggggagccg	gggctgatgg	ttaccacagg
20581	atagcggaaa	ggcacactgg	ctggcctgtg	tgactcacg	catcccccca	cctaggcgag
20641	ccttgggaag	agcactcagg	attatgagaa	agactgtcgc	ctcccctttg	cttcattagc
20701	tgatcctcta	agcatatgtg	ctttcttggg	ctaattttcg	gattggctct	ctcctatatt
20761	ctcttcctac	tccccacccc	gaccttacag	ctaagtgcac	atctcatgta	gtgcagtggtg
20821	aaagaaccgt	aaggcagaag	ccgggctgac	ttggctgtga	atcccagctc	catcacttgc
20881	tggccagggtg	actgagtaag	atcgtttaca	catccatcat	cctcaagttt	ctcatctgta
20941	aaatggggat	aataatgtaa	ctgccttatg	gcattatata	aggattgtat	gactgaacac
21001	atgtagaatg	cttagaacia	tgcctggcat	atatgaagca	tttaatacat	ggtgtattaa
21061	attagttttg	aaaagaataa	attaataaca	atgatgaaca	tttttgatac	ctattttcct
21121	attgttttga	ctctcaaagc	cagttgcaag	catatttagc	actgtgatgt	atgtgtgact
21181	tactgcaaa	tcttttttcc	agtccttgat	accagctctc	tcttcacctc	cagtgtttcc
21241	taccctcct	gcctcccctt	ccctaagaat	attgctgttt	cacagagtgt	aggctttcct
21301	ctggcttcca	gatctgcccc	catatgcaca	cttctctttc	ccatccctgt	tggactcttt
21361	ctccttatca	gtttatttgt	tccagttggg	aagaactgga	acctggctcg	cagcttttcc
21421	agttggcttt	atctgtgcgc	tgcatgttaa	aactgttctc	tcttgcttag	aaatctcttt
21481	gatccatgtt	tagctgtatt	tattcttcca	acagatgttt	tgggtagtga	gaggattttc
21541	ttctcgcatt	tgccatgtct	catgctcctt	catgcttccc	acttgttcgg	gatctttttg
21601	ccagctgacc	acagacaggg	gccatctgtc	gtgaaggctc	ccctggccca	gcagaccagg
21661	aatggcccag	caaccaagac	tttctgaagg	gcttagtgaa	ggggaggagg	gaggaagatg
21721	ttggagaact	gtgtagggtg	gagtttgagt	ttccagaca	cattccagga	gtcttttga
21781	tccaaggtat	acatgatttg	gcttgtgctc	tgtggcagg	taacaaaaac	acaaccttcc
21841	attgtctcct	gtagacaaca	gagtgaaggc	cttgggcatg	gcaggtagcc	taagactacc
21901	cctgagagtt	gggaagtgtg	tgagtctcct	ggggctgccg	taacaaagca	ccacaaattg
21961	ggtggcttag	cacaacagaa	atgtattgcc	tcacagttct	ggaggccaga	agtccaagat
22021	caaggttgcg	gtcagggccg	tgtttcatct	gaaggcccag	ggaagcagct	gcccacgccc
22081	ttctcctagc	ctctggtagc	ctctggcatt	tcttggttta	tagatgcata	tgtcaaatcc
22141	tgtgtcttca	tatggctttc	tcctttgtct	cacactgtct	ttcctctgtg	catgtctgtg
22201	ttcagatgtc	cctttttata	aggatgtcaa	cccaattgga	ttaagttcta	ccctaattgat
22261	ctcattttta	cttggttacc	tctgtaaagt	ccttattttc	aaataaggtc	atgtgtctaa
22321	gtactaagg	gttaggactc	cagcatactc	tggtggtaga	cacaattcaa	cccataatgg
22381	gaaggaaa	gttgggcac	ctgtaactcc	tccaaacacc	cacagagtgc	agggtgagct
22441	gtgtgcta	acatagtcag	ttctctttgg	ggtgaggagg	cctaggggca	gggccccat
22501	gtgggtctc	tgtccacacc	agcaacaata	acaaccagg	aggaaagcat	ctcattttcc
22561	ttggctcagt	tcagcttttt	atgttttttag	cacaatgcct	gctttgtctc	tccaacaatt
22621	tgggaatctc	tgggagctgt	gcatggaaag	caaggaggac	agcggcgaga	aaaaggggga
22681	gtagatggag	ggtcttggaa	agcagagggc	ctaggcaggc	agagaggaca	ggaaagtata
22741	gcgagcagag	cggcaaattg	gtggggagg	gcagaaggct	gcttggcagc	caggagttct
22801	tgccctggcc	ctgccatgag	gctgcagtgc	tgtggcctag	gtattttacct	tctccaggcc
22861	tcagtttctc	tgtaggcaag	attgggagg	ggatgggtgc	tctctaggat	cccttctctg

	22921	ccagaataaac	attctcagca	ggagcctaac	gtgtggagca	aatgggagca	ctgggctccg
	22981	gcctcctgca	gtgagcacag	cccctgttct	tgtggaaaca	tcttccaata	gggctgccct
	23041	gcctacagg	tcatgcggca	tgcattctgt	gcctgcctgc	gctcttgaaa	cagcctccac
5	23101	tgctcccctc	ccagctcctg	tctctctgca	cacgcaagcg	tgctactcct	tttcatgata
	23161	cccattagta	ttcttttgacg	atggcataca	tctgtcttcg	atcggtgtca	gctctgggag
	23221	gcttatgcca	agcttcttga	gcgtaaccca	tgactgcctg	ggtagaggtg	tgtgagctgt
	23281	ccaggaggca	ggaggacgat	gcatgcaagt	cagggcctag	ggcagaagtg	cctgggcctg
	23341	gcctcccctt	ggactccagg	agtcctgtcc	taacagagcc	cacagcccc	tatccatctg
10	23401	gcctctgtaa	cccctcccca	acacacacac	acacacacac	acacacacac	acacacacac
	23461	acacacacac	atagcccctg	tgattgaggg	ggccccaatt	cctgttcata	tcctccagga
	23521	tagcccacct	gcacccctga	cagtgaagga	caaagtctta	ttccctgttt	agatgggtgc
	23581	tggggacaat	ggaaaggagg	tgtggctctg	agaagtctat	gtcttgctca	gggcacacag
	23641	cagctgatcg	ggaacatgtt	gctgactcca	agatgctgcc	ttgcaagaag	ctggctctat
15	23701	ccttctttttg	gctgaagtgc	ctttcatgga	tggtagggga	tgtgcagggga	gaagtgtcag
	23761	gagtgaaggt	cagtgggttag	aatcaggcag	tccacagagt	ctgagaaaagc	aagacattct
	23821	ctggcagctc	gggggtcatg	atcgcccacc	ccagcccaga	taaccctcac	agctgtgcgg
	23881	gccactagag	aaaaaggagg	gcatgtttgg	ggcaggagag	gcaaatgttt	gcttatctgt
	23941	gacttcttcc	tccaagcatg	tccggacctc	cagtcaatgg	tgggctgtca	gtcgtcagct
20	24001	gaggttgagc	tttccttagc	aggagcactg	gtcacttggg	ctgggtaggt	tcttagtggt
	24061	acaggatgca	ctgcaagctt	taaatgcgag	tggcatcatc	cccttccggg	caccatggca
	24121	accagaaaca	ccttgacaca	tttccaaatg	cccttttagta	gggcagtgc	agcccttttg
	24181	agaatcacat	agaatcgcat	tgattgatga	gtgaaaaata	aatggatggg	agcctccttt
	24241	tgtgattttt	gcagcggcct	ttagcttcct	ttactcacc	cagaaatcag	tgggaccctg
25	24301	ggagctgtgt	accctcaga	cccagttgga	accagccaa	gagtacttaa	tccatcccca
	24361	cttgtggggc	caacggcacc	taaccacctc	aggcacgggt	gacctgggtc	ctcagagagc
	24421	tctagggaca	gaggagagaa	agggtctgca	ttctgtttgc	agccctgata	gtgagctctg
	24481	ggggtcctct	tccaccccca	ccccacccc	cagccctgg	agcaggtact	cggggtcaga
	24541	gctctgtga	gggtctggct	ctgggagggg	aggtttgtgt	aagattccct	cccacgggtc
30	24601	agcacagatg	ggatgacaag	gaccaaattc	tgtttctggg	ctctgatatt	tgccaagatt
	24661	tttaccaggc	ttcctggaat	agacaggga	gcagagcaag	ctcccgtagg	tcaagtgatt
	24721	tggggcccgag	ttgaccacga	gtccctaata	gactgctgtg	tagctaccat	gagtgtgctg
	24781	agtggcccat	aggggcagg	atgagagagg	tgctgagggga	ggcagggggc	ccgcagaacg
	24841	gcctcccatc	tccactgccc	gtcccagggt	ccacaggctc	acagagcaga	cacgggtctgt
35	24901	gcctgggttt	gctcaccac	aagaggaaga	acataacatc	tcgctccttt	tgctgcacag
	24961	gataaaaacga	gaacagagag	gaaacaggaa	gtgctttgca	ttccagaaaag	agcagcaact
	25021	gtataaaagtc	atgcatatta	ggatttgagg	tatgcatggt	cagaagttag	aaactaacgg
	25081	aatcttgtca	ttgccaggaa	gtttcggggg	tctgtgactg	gtggccactg	atgttccctg
	25141	gttccctccat	tccagctcct	accttgactg	tgtcctcctc	ttcacacctc	acttctttag
40	25201	tgaaggctcc	atttctctat	ttcctgttca	atgcttaaac	cccttgcaat	ctggcttcta
	25261	ccatgcgctt	atcacagacc	cttctctggc	tttgcccttg	ctggcccttc	catgatgtct
	25321	tctttcttga	aacctctttt	ccttggtagg	ataccacggc	atcctgggtt	ttgtcctacc
	25381	tctgtggctg	tttctgctg	ctttccttag	ctgacttttg	ctcctttatc	tgacctgggc
	25441	tctcctctct	ccatacactc	tccatagcct	attctaagtg	tcccagggtc	cttatatctt
45	25501	atctctcaaa	tgacaatta	ctttgtgtta	gctacagacc	catatatcca	gcttccctat
	25561	agatacccca	aatgtctttg	taggctccta	aactcagtgt	atgctaagct	gaacaggggc
	25621	tcctttttgtg	ccccaaactt	tcccaactcc	agtgagtctt	ctccattgtc	tttcacctta
	25681	ataaatggaa	ccacctgcga	ctctagtgtg	tgggtccagag	acttgggaag	cacctcagct
	25741	tgtctctctg	tcacccgcag	gatcgggcag	cctccaagtc	ctcatcattc	taactctcat
50	25801	aatgcctctg	gagtttgtcc	agatctcctc	gtcaccactg	ccgctacgct	aatcaaaacc
	25861	accattgtct	cttgccatcc	tccatacttt	gcaaagttaa	ttgggtcatt	tttctacttg
	25921	aaatcttata	atggctctcc	agtgcctctg	agtccttggt	ttttcaacac	tgttcacact
	25981	ctccccacct	ctctctcact	catacccat	gcaccagcca	tctgggtttt	tgctgttttt
	26041	gtttcccgaga	atgcacaatg	caccttctgg	cctctgagcc	acagcacctg	ggtatttgct
55	26101	cacactactg	cagctctcct	cttccctgcc	accacgcctt	tcttgccctga	ctgttagtat
	26161	gcagcagtg	ccacttgagc	atgaccgcct	ctgggaggct	gtccctagtc	ctctgtcgca
	26221	ttctggggct	ccctatcaca	cactcccat	gcatgctgca	gcacctcag	caccagcat
	26281	tccttattgt	agttcctgat	tcaacacctt	tctcaggaga	ctctggactt	cttgaaggca

	26341	ggaacaattc	ccacttggtc	ctagtagcat	tccaaaccac	acgtgacagt	gtctggctta
	26401	taataagcag	ccaataaaaa	gttgatgaat	gaatgaataa	gtgaaaacag	aaggtgtttg
	26461	cctgcagaaa	tctggaataa	gatcaaagat	cagagctggg	attaaggaaa	aaacttcctt
5	26521	ggggtggcac	tatgaattcc	cagaacaggt	gactaaccct	cattcacttt	ggcaaatggt
	26581	tatcccatgc	cacgcaacca	ggcaaaaagt	tgaatgaggt	ttaatccctt	cccacacgga
	26641	gcttattcct	tctttggaag	tcctttaaac	aagctctgaa	atgatttttg	caggtagaca
	26701	aactggtcct	catttctctg	tgaccagtaa	gtagggaaag	caagcacaca	tacacacaca
	26761	cacacacacg	tgacacacga	cactgacaga	caaccttgct	cactcacatg	ggcatgcca
10	26821	aaccttcttt	ctattttata	ggatggtaac	tcactcttta	gttttagactc	ttgacgtgcc
	26881	atggaaaatc	ccactcgccc	tagaactggg	ggccgggcag	gtttgactgt	aacaacgaag
	26941	cctggagcct	actctttgct	gatttggttt	cctttctgtc	tccatttttc	ccctggtgag
	27001	cactgcagtt	gtgttcttcc	tcccaaaggt	aatgcctggt	ttggctcact	aaaacctggt
	27061	ctttctgtac	cgagagctca	tcttctcttc	ctcttctgga	ttctcaaagt	agatgacgtc
15	27121	agaggatgga	ggccaaccac	accttcctc	cttgaccctg	ataaagtttc	ttggaaccc
	27181	tatactcaga	ggcagccaat	tcttgccagt	ggaagagtga	aaagagggct	tggaagctc
	27241	aaggctcagt	gtctgtcccc	aggtccccc	gttaaagaca	catctgtcct	tcactctcaa
	27301	agatgttgcc	attgctcccc	tgctagagt	acacactgca	ctccctcctt	cccttcacac
	27361	cccagcaaga	ggctatttcc	caggggtctt	ataagcagat	ttcatcttct	cctgtgctgt
20	27421	tttcttattt	caattatctt	caggagaa	cgtgcatatt	gcgtcattgc	ctggctgtga
	27481	aatttcat	ccatttcttt	acacctgcag	ttgcaatacg	agagagaaaa	ggccagagct
	27541	tagcggatgt	cctagacgca	ggttatcaag	gtgctgtggc	tgtggtttcc	cggaaaaggc
	27601	cttgggtcca	gagcacattt	tatcagcagg	accttcgagg	ggctgcgttc	cttcaattgt
	27661	tttctctttg	gggtctctgg	tctccagttc	tttcttctct	agcatgtgag	atctgtgctt
25	27721	ttgattcatg	cctttaagtc	tgacattgaa	aaaatatctg	atttgccatt	ccagatgctc
	27781	gtcctcattt	gcaaattttc	ctaaagggcc	agattgtcct	ctggcctttt	cccttttctt
	27841	ggtcccacct	caccacccctc	ccactggggc	ttcacagagg	cagagctagt	ctcctttcat
	27901	tttttaaaat	taatagtctt	caatttttag	aacagtttta	agttcacaga	aaaattaacc
	27961	agctattaca	gagttaccgt	ataactcctc	cccctcactc	cccagttttc	tccattatta
30	28021	gcatgttgca	ttagtgtagt	acatttgttg	cagttaataa	gcaaataatta	gcccattcatt
	28081	atctaactca	gtctatagtt	tacattaagg	tgtattcttt	gtgttttaca	gttttatggg
	28141	ttatgacaaa	tgcataatgt	tatgtatcca	ccattatagc	atacagaata	gtttcactgc
	28201	cctaaaactc	tcctgtgtct	cacctgccca	tcctctctcc	ctcctctgcc	accaatccct
	28261	ggcagccacc	agtcttttga	ctgcctagag	tttcgccttt	ttcagaatat	catagtagtt
35	28321	ggaataatac	agtgcgtagt	cttttcagac	tggtctcctt	cacttagcaa	tatgcatttt
	28381	aagtttcttc	catggaaact	ttgctttcat	ccttttatca	ccacaaggcc	agtcattcaa
	28441	ggaatttctc	catctctgtc	tgttcctttc	tagttctatg	tgtgccactg	cttggcatag
	28501	aataggtatc	catttaatat	acattccctt	tcaccacctg	ggacaccttc	ccaggataaa
	28561	caaaaataaa	accagctagg	tcaatagcag	agccccatc	ccagttttta	ccttattctc
40	28621	ccctcttttc	acaataaact	ggatcagaac	cagcagctct	gtaagactgc	atttcttttc
	28681	cttaatacca	ggccccagag	agcatttgat	tccttggcag	agaggtgtag	gcttaattaa
	28741	tttttctcct	tttttctttg	aacatcttgg	aacacacaca	cacacattcg	catttatgca
	28801	caattgggtg	tatgagaatt	ttaatggcag	gtgggtgttag	cagttctttt	cctcctgata
	28861	cagatcaggg	tttttccatc	tgggcctttt	agcagggcct	atgaatattg	actttctaac
45	28921	cacttggaat	tgggtggagt	gtgcagagt	ggattggggg	ggaaggttca	agggagaact
	28981	atacttatgt	ataaatcaca	tgtgaaggga	gttttgaagt	cattatttgt	tcaggatgtg
	29041	cgaaccataa	ttattttttt	aggtcttgat	ttgccc aaag	agcattttcc	agggttgctg
	29101	ctccaagcat	gacgtctgtg	ctgtcaggag	gtgcagcata	gtctgattcg	agtttaacg
	29161	ctttaaagga	ggcctgggtg	aggatctggg	ctctagggtt	tcagctgtgg	tcagtcctcc
50	29221	atgcagcaaa	acatccagat	gacttagatg	attaagacag	cagacttaaa	gtgaagaaga
	29281	gatttttttc	cttattcttt	ccttttatta	ttagttttta	aatggttggc	tacatgggct
	29341	gttggctcatt	ctccatgttc	tctgtgctct	cctcagctct	ctgctcaaaa	caggctgcac
	29401	cggcctgcct	aaacctgaa	agcaacttct	cagctgccta	ctttctgcct	tttgaccccc
	29461	aagccaatcc	ccatctcctt	accacctctc	cgccatgtcc	tcatacacct	gcctctcctt
55	29521	gacttcattc	ttcatgctcc	atcagcaaca	gccctctgtc	aataatgatt	gtcccaggga
	29581	agtgtattca	agggtcacat	aaaatgtgcc	ctctctatgt	gttgagaagg	ttttctgtcc
	29641	ccaaaggagc	tctctggata	atgaggaagg	ttgaactggg	gcagcctaca	ggaagaagcc
	29701	cttagaaggg	aaacctgtgg	cataaacat	gctgatccac	gactcttatt	ttggaatagc

	29761	tattttaaaaa	gaaatatgaa	gaactcgtaa	gacttggaaa	agaaaactag	agaatgttga
	29821	aaatgtccaa	gggttatgtg	tatgatgtgt	atggggaaat	tttaaaaagaa	tggtgtagaa
	29881	aactgaattt	gtggtaaaa	gttgtcacag	gacggcctgt	tctttcattg	aattatgtct
	29941	tagtgcaggg	atgacaaaata	aatacagcaa	ctgtgctgcc	attctcacat	cttttcttac
5	30001	agtaggcac	actaatcaat	tacaacattc	ttttccactg	tcatgccttt	gcctacagca
	30061	gacatcacta	atcaatcaca	gcactctctc	tcattgagtt	tcaaagggtt	tttaatcctc
	30121	cacataggtc	tctaagtagc	tatataccaa	ttactttgat	ttagaacttg	gattaccaat
	30181	ttgccatctc	tagtttaatg	ggaagaacat	taggattaga	gtcagaatac	ctgaattcaa
	30241	gttactgacc	tactacttaa	taataattaa	ttttaataat	tctctaacct	cagctttctc
10	30301	ctctataata	agaaaataatg	cttaccacag	aggctcgagg	taaggattaa	gtaagataat
	30361	atataacatg	taagataaatt	tgtgtgatgg	ttcctagaat	agtacctggt	gcattgtaag
	30421	cacccagcaa	gtgatagcca	agtatgaatt	tgcagcggca	tggaacacac	atgccacacc
	30481	ccagatatgg	aacaatagtt	aggattccat	tcacctttgc	ttcttttaac	atcctcagtg
	30541	aaggcagatg	gagaacagct	ggaccctcta	attctacctg	actttaactt	cacattctta
15	30601	agatgcttat	taaatctctc	tttcctgaac	taatataatg	cttccttggt	acctggaag
	30661	gacagagtta	acacccatgt	attatgacat	gatggatggt	ctttttgctc	acactgcatg
	30721	catagccatg	tgccaaggcc	aggcctgcct	tagctttttg	gttctccaga	gcagcttcac
	30781	tgtgactgaa	gagagccagg	gaagacatct	tggtagagct	cttttatatg	ctcttttcag
	30841	aattgttctc	actgatggag	atgacagaaa	cttaggatga	tttgaggcca	ggaggacagg
20	30901	cttcttttga	aaaggtttca	ccatagatta	cctcaatcca	gggttggaag	actggaaagt
	30961	ggtctgcagg	gcatcccagg	ggatgccttc	ttcaagacaa	tattagacta	gcattggacc
	31021	ctgcctccag	actagtaaag	ggttattttc	agagcagcaa	caccagagaa	catgttttga
	31081	gcaaatcaaa	tccttaaaat	catgatttca	tacatcctga	gatactagtg	acttcaaaat
	31141	gccttttctc	aattatacat	atagtgcctt	ctgaatcaac	tgtctttttt	ctgtccaaca
25	31201	gtataaacat	aatctctgct	ctcacacagc	ttagaaactg	tcatgagagg	taaacaaaata
	31261	cacaaatgac	aaggtaaaat	agaagagcta	caaaagagat	catgaaaata	tccaggaggat
	31321	tcgaagggtg	cagataacaa	aaggcttcgt	ggaggaggta	gcctttgtgg	catttgagat
	31381	cagtcctgca	gcagggttat	aggtggacag	gcagagatgg	ttgggggtgg	gcggtaggag
	31441	ggacagacag	aaggaaccac	aagaatgagg	atatggagaa	aaagaccact	gaagggtgca
30	31501	aattatttaag	gatggattta	tgtaaaggat	gtggcctaag	aagaaggact	cttctttatg
	31561	atgaaattag	gttttagagg	cagaagtctg	gtgcttcccc	gtactaaaat	gacccagta
	31621	tggtgtctgc	agaatactta	gataactgca	tagatgggtc	agccttcttt	ccattatacc
	31681	acactactgg	gttaccaact	tgctgtgtga	ccttagataa	gcaactacct	ctctctgggc
	31741	ctcagttttc	atatctatga	aattaggagg	ttgaattatg	tcttggcata	ataacgaaat
35	31801	aaataagtga	tttattttta	tttctttgtc	attcctcttt	gaaatgggag	tggaataaaa
	31861	gtggtatttg	tttccacatg	aaaattaaag	ccagggggcc	aattctggac	ttggcttgag
	31921	ctgtgctttg	tctgtttttc	tgactgggcc	cagtaccctt	tacactggtg	taactaactc
	31981	cctggaaagg	gatgcaggta	gacagcgtga	ctgttccttc	ctatctgaga	ggcccaaca
	32041	gattctatat	tgaggacca	ccaacttgga	atttcagctc	agttcagcaa	acacttttga
40	32101	cttggttcgtg	tgctagacat	tgtgctgggc	ttatgtgggg	tagagatgcc	tgctaactta
	32161	aggctaacag	gaagacaggg	aagtggtttc	cagacctgat	gcttatcaga	gtcactgtgg
	32221	aagaactttt	tagaatgcag	acagcccca	acttaacagt	ggttggaact	tcaatttttc
	32281	accttataat	ggtgcaaaag	tgatgggcat	tcagtatgct	tctcaacata	caaagagggt
	32341	gtgtgaaata	aacctatcat	taattgaaga	tactgtcaat	caaaaatgta	cttttgactt
45	32401	acagtgtttt	caacttagga	ggggtcatag	gttgaggagc	atctccacag	attcctgagc
	32461	cccatccctt	gatgatgatg	atccagttag	cctgggtggg	gcctgagaat	ctgcatacta
	32521	actgctgaag	gtgagtgtga	tgacgggcca	agcttcagaa	cctctgatct	agagaagatt
	32581	ctgtgtcatc	acagctcagg	gtgacctagt	tctctcttac	ttgcacttac	atgaattcat
	32641	atgaattcag	tattacatga	gtaaattatg	attatatgaa	tttattgagt	atcctatgtg
50	32701	cttgatatag	gtgtttacat	cccagggttg	gggggtagag	atgaagaata	taagccggac
	32761	tatgttaaga	tcttcacttc	tcagcacctg	aacaactggc	agcctcatta	ggaaggcaag
	32821	tcattcacag	gtgaaatgac	atcatggctg	tcttcttcat	tcttctcttc	agcatccatt
	32881	taagactcac	attttatcac	cagagactga	aaagagccac	ctaaggcagg	caggtcaggt
	32941	ggtgttatct	ctttattttc	agatgtggag	gctgaggctc	agagagggtg	atccatattg
55	33001	ccaagctcac	atctcccgc	ctcagtcacg	ggcttctccc	cacttcatgg	gagaagcatc
	33061	ctcctcccca	gagcagtagg	ttctggagct	gggagaggcc	actgtgggct	ggattgttgg
	33121	ggacagcttc	agggagagcc	cgattcaagg	caatagagaa	ctttggctgc	aggccgttgc

	33181	ctagaatagg	gcagctgaca	cacctttgat	ctggaatgat	tcctgctgct	gagaatgagg
	33241	ttttttatat	ctggattctc	aggtagtaac	accacgacaa	cggtgtgttg	tgttatttca
	33301	tctgcacagc	attcacgtgt	agcagagaga	aggtagttta	ttccccagaa	gtttaccagt
5	33361	gggaaattga	ggtcaaaggt	aggactttcc	taaatgaaga	gaactactat	ttattggatg
	33421	cctcccatct	gccaaagcagt	gtgcacagca	ggcatatatc	atgaaatgcc	tcattttattc
	33481	ctcataactg	tagcctgatt	ttacattcgg	caaacctgag	gcttacagaa	atcatgtggc
	33541	tcgcactcag	actgctgatg	gccacgcaag	gctttcagct	tctctgattt	ccaggccccc
	33601	atccccacca	cactgtgctg	tccagctcac	ctgggtggaac	tggtatcctt	gagttccagg
10	33661	ccagggatcc	tactgagctc	ttccatcagg	gaaaccatag	cagtagcagc	tccaccaagg
	33721	acttggcatc	tatcttactg	cagcatccgt	gcctgtctaa	tggaaccatg	taccaggagt
	33781	agctacccaa	gaaacatttc	caccagaaac	ttctctttat	agctcccat	cgagcctcag
	33841	agagctgata	ggaattgctc	aagaccaccc	aacttgtaag	tggtagagct	gggactaaaa
	33901	ggcaagtccc	ccaactttca	gcctcactcg	tgcccaatgt	gtctcagcct	ccctgaagaa
15	33961	tatcagacca	agatggccag	aaaaggaacc	tggaagggac	gtgtgggct	catgcagccc
	34021	ttggcacagt	cttcaggctg	agctgcctcc	accttgtcat	ctcatcagag	ctttacatcc
	34081	atccttggga	gaggccagct	ccatgacctc	tcagtgtcat	ttaggatctt	cttcctaaac
	34141	ggctaaccce	agtctgtccg	ctgcataccta	cagtgaagtg	cactgtggcg	agagcagcta
	34201	ttggcattct	gcttacgtgc	tgtttccaga	ggtaaaactca	gtataaatgg	atctacagcc
20	34261	tgtcccatct	tgtacagctg	agttctaaaa	ccagcctgca	aggataatgc	tataaaaatg
	34321	tcctgccagc	cccaagggtg	tcttctacaa	caagttcttg	ttccctcata	attcttctga
	34381	caaattcttc	ttttgatcca	gactttccta	gccatcattt	cattccagaa	gtggcatgtg
	34441	tgtgcaaaac	ctcatgcttg	agggcgggac	aagcaaaggg	ataagagggg	aaatgggaac
	34501	tcgaaatctg	ctcaaatgtg	gtaaagaaat	atatccagaa	agtactgcct	actcaccaaa
25	34561	atactatttg	ctttttatat	tctttcctga	gtagaatttc	ctgttcaaac	ttgaaaatga
	34621	aaattcctcc	acttcaaaat	gaacaggcaa	gaaatgctgt	aggctgggtt	tcccggggag
	34681	gggggactga	cagccatctg	ccctgagact	gactgtcaaa	tctgaactct	gtgtacttgt
	34741	tagtggttgt	tatgggaggg	gtgagggag	ggaggaagca	acagggacct	gctaacccta
	34801	tgaattctcc	ctcataccct	taaaaagtcg	ggtgcttgcc	cgggtgcagt	cggtcacgcc
30	34861	tgttaattcca	acacttggtg	ggccaggggc	gttgatcac	aaggtcagga	gttcaagacc
	34921	agcctggcca	agatggtgaa	accccgctct	tactaaaaat	acaaaaatta	gccaggcgta
	34981	ctggcaggcg	ccggtaatcc	cagctactcg	ggaggctgag	gcagagaatt	gcttgaaccc
	35041	aggaagtga	ggttgcaatg	agccgagatc	atgccactgc	actccaacct	gtgcaacaga
	35101	gcgagactcc	gtctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	gtcgggggct
35	35161	tgccaagctc	gctttttggg	gcaatgggag	ggaaatttta	agagctgatt	tctgtgatct
	35221	ttatcagacc	tcttcttctg	tcctgcccac	aggctcagtga	tgggaaatgat	agtattgcat
	35281	taaaagagaa	gtgattcccc	agatgtggac	atctcagtgc	ctgggaggca	gacgtcttta
	35341	gagtgttgtg	tgttcctccc	catatctaac	ttacctgttc	ttaggacata	ttcccttgag
	35401	tgtcttttat	tatttcgttg	attgattaag	ggttacaaat	ttgttacaaa	agcagcagaa
40	35461	attttgaaag	ataaaaacag	tagaaaactgt	tatttaattg	gaggcacaaa	agccagttcc
	35521	agcttgagca	ttaactggct	gtgtggcttg	ggcagctccc	ctgcactttt	gggaccatgg
	35581	tttcttcagc	tctaaagtga	aagctgagag	ttcctctgaa	tgctaaaatc	ttgcttctct
	35641	gtgattattg	ttttttaaat	gccgccgtca	tttcaaaaca	catacatggg	gcagccctcc
	35701	tcagctgcct	gtcctgttct	cttccctcct	tccctttgtg	ccatcagcgt	ctcccctgga
45	35761	tttatcgtgc	tgtactcctt	ggacaccctt	gaaagtggag	agagaataat	tctcattccc
	35821	tttccccaaa	ctctgtcctg	aagcaccact	ccccacctcc	ctaccatgcc	tgcccctcca
	35881	ccccccacat	tttccctctg	atgagatttc	ctctctgagg	gaatattttg	gtcctgttct
	35941	ctgttcccaa	acaaattccc	caggtccttt	aaatctgggt	tcaatcaaca	taatcacagt
	36001	ctcttcgatg	ctatcaactc	tgcttgatt	tggttggtg	attcatgctg	tgctctccac
50	36061	ctcctcaggc	tgttttatat	atacacactt	aagtgtcat	atatatgtgt	atatattaca
	36121	caggttgtat	gtatatgatg	tattgaaaga	gagagatgtt	ttccaccgtg	cctgtaatta
	36181	ctttttatct	tcctccaaat	cagtatctta	atgcatatca	cctaaactgg	attcattttg
	36241	agaacactag	ggcctttttt	ttacttttca	agaacagcta	tctaattctg	ggccaaaaac
	36301	tagttatcaa	atgagggcag	gtaatggagt	tgtatccagc	gagggggcat	ttgcttctcc
	36361	atatgatata	ctcctcacca	gtaccaaggc	gttttctctt	cagtcttcat	gcctatgctg
55	36421	ttacgagtcc	tttctcctat	ttgaaagaaa	agatatgagg	cagcccaggg	aaagctctga
	36481	tggaggctat	aaagacaata	tggatgtaaa	ctaaaaaatg	gtagcttaga	gcttaatggt
	36541	agcctctcag	gacactcacc	acacctgaat	tctaccagcc	tgacgccaat	gctgctccgc

	36601	ttatgctggg	atgtagtgac	acaggacact	ctgctggggg	atagggatga	tctcccccca
	36661	ccaagtgggt	catcaacact	catcactcag	aggggagggtg	ggataccacc	ttgaacagag
	36721	aaagcggctg	ggcatgggtg	ctcacacctg	tagtcccagc	accttgggag	cccaggcgag
5	36781	gtggatcgct	ttagcccagg	agtaagagac	cagcctgggc	aacatggcaa	aaccccatct
	36841	tcattaaaaa	tagaaacatt	agcctggcct	ggggttgcac	acctatagtc	ccagctactg
	36901	gggaggctga	ggtgggagga	tcgcttgagc	ctaggggggtc	gaggctgcag	tgagccatga
	36961	ttgtgccatt	gaactcctgc	ctaggagaca	gagtgcagtc	ctgtctcaaa	aaaaaaaaaga
	37021	ggaagaagtt	gctgaatccg	tcttataaat	ctgtaacaga	aaattatggc	aagttgtggc
	37081	ttttaattgc	cagagctggg	cacttaagag	agaaaagggt	ttttgtgaat	tccaaaaatg
10	37141	taattgtatt	cattgaacac	tgaagtcagg	agaagtcaga	ccataatgac	atgggagcca
	37201	ttttcgcaac	cagcagacta	gaaggggagg	gtttggaagg	gtggcggcag	agcttagaca
	37261	cctgctgcga	agggagagaa	gtggggacag	aaagatggcc	tgtgtagagt	cccatgagaa
	37321	agacagaact	gacttgccag	gcactcttag	gggcccaggt	cacaatcatg	gggccggtgg
	37381	acagtctcca	gggcactgtg	attgtcacag	tgcacagcct	aatggggaaa	attgcacagc
15	37441	ttcacttaaa	tataggtgac	atatggacgt	aagaattcat	gattaatcta	gaactaacc
	37501	tgaccagcaa	ggccgaatga	agaaaaaatg	tagaaacact	gaccaaacc	tcctcaaagg
	37561	tcacagatct	tagggaatgt	gtctctcttt	cacatttcaa	aataacaaca	atttttgaaa
	37621	tatgtatttt	ttaacattta	attgttttat	ttggggataa	ttttgactc	atggcatcca
	37681	ctgttttttt	agttaccac	atatttgatt	cagtaagatt	ccctgtattg	atacaaaaaa
20	37741	agaagaggga	ccctgtttcc	ctttgacact	gttacttttt	cctgagctca	ttgtttgttt
	37801	gtattttgtt	ttgttttggg	ttggtttttg	agatgtagtc	ttgctctgtc	accagggcta
	37861	gagtgcgaat	gcgcgatctc	agctcactgc	aagctccgcc	tcccgggttc	actccattct
	37921	cctacctcag	cctcctgagt	agctgggact	acaggtgccc	gccaccatgc	ctggctaatt
	37981	ttttgtattt	ttagtagaga	tggggattca	cggtgttagc	catgatggtc	ttgatctcct
25	38041	gacctcgtga	tccaccggcc	tcggcctccc	aaagtgcctg	gattacaggc	gtgagccact
	38101	gcgctcggcc	tcctgagctc	attttaagag	agacttctgg	cctagagggt	tgaatgagaa
	38161	gaagtataac	agctgggatt	cttccctttc	tctgatatga	ggacaggagt	tctctctcat
	38221	ctccgccaag	agcaggaagc	tggagtaact	gccacaagct	ccaggaggga	gtgtctagaa
	38281	catccacggt	ttgcagcagg	aaaacacccc	ctcacgctga	agtttgattc	gtgaatcctg
30	38341	tgtcgcagtc	taaatgctga	ggcagaaggg	gacatccgtg	ttcctggggc	attccacttg
	38401	cagtcctggc	tgtaaccoga	gtgagccatc	cgtgtagtgc	ctgttgctaa	gtctcccctg
	38461	ccacctcttc	ttcccatggc	tgcagggcag	ggggccatgc	cctcctcttc	atttctgttc
	38521	cctgggtgag	cgtgccccct	gccttctccc	agatctctgc	tgtggcagct	tcacgtggga
	38581	ttcagcactg	tgtctccttc	ccctctgtct	ggcctgcccc	tacctgtcca	gcagagctgt
35	38641	aagaccagaa	gacagagcat	tccccctatc	tatgaagtca	aatgcatgtg	tggaaatgc
	38701	cacccagcct	gcagtctctc	tactataaaa	tactgctcat	aagacaaatg	tgtggcccag
	38761	atgatttctg	ataaagtcta	ttattttgaa	atacatatgt	atgtctctca	gccactgata
	38821	cacgcagaag	ctgcacatgt	tgtaaacacc	tgctttgagc	tcctttcctt	cccacccttc
	38881	cttttggaac	tgcaagtttc	cattcatttt	ctgcattact	ggctctcctc	cttctcccct
40	38941	actactagat	cttacaataa	acatttgaaa	tagtttattt	gtccacagtg	tgatttctgt
	39001	tgagaaacat	gggctcaccg	acttttgggt	ctcttctaac	actagaaatt	cctctgggtt
	39061	tttagactat	ttcaagggtc	attaatgtgg	aacagacggc	ctttagaaac	agcaatccac
	39121	agggggcccg	gagacctgga	ttcatatttc	tcctagcgct	aactagtggg	tgaccttaga
	39181	aatgtcattt	tccttgctcg	gctttagtgc	ccttatctgt	cacacagaag	cactgtgtga
45	39241	gtttgggaaa	ccaatactat	gttgaaatgt	acaaaataat	cttaaagcac	agatattctg
	39301	ttccctccaa	gaatacatca	aacaaaggaa	ctgacattgc	aagaagattt	tgaggagatg
	39361	gctggatgca	ctggagcagg	gattgctgag	ggaagccagg	ccctcacctg	gagcgtctca
	39421	ggagaggcag	cttcgggtgc	ggctgtttat	tgcaggcatc	tctttctgtg	tctgtgcgtt
	39481	taggggctct	tctttggaga	taagaaaagg	gttctggatg	gagggcagtg	aagaacagtg
50	39541	agaacttaac	atgaggatgt	ttgtatagag	gggaagactt	ctggacagtg	gcttgacttt
	39601	gctcactggg	catttctctc	tggatctctg	tagaagtcag	ggacagatct	cctcgtgcat
	39661	atctgtctcc	cagagacaga	tctcttccca	gcagaaagta	gaaagtgggc	ttcaggcatc
	39721	ctggaagttt	tctttcttgg	tgggtgataa	aagggcttgc	agagagagga	gaatcaaadc
	39781	tcccacatgt	gcacatgcgc	tgcgagtcct	atgcagagat	gtcttatgtt	caacatagaa
55	39841	agcaagcctg	gcagccccaa	gaccttcttc	tgcacaccgt	ccattttttac	ttggtttcat
	39901	tttgataaact	gtgcgggtctg	aggtcctggc	caagaaagca	tcacctggca	agaagtgtat
	39961	ttggccaatg	gtaaggttac	catctctgtg	taattaggct	ccgtaaaagct	ttgtttttaa

	40021	at tt t a t t a a t	g g g a a t g a t t	t g a c a t t c c t	a c a c a c t g a c	a t t a c c c t c a	t g g a a t g g a t
	40081	a a g a a t c t c a	a g g c t t g t t g	g g t g a a a g a a	g g g c a g t g t t	t g g t g t g a c g	g g a a g g g a a a
	40141	g t a t a a g c a g	g c a g c t c g t g	c g c a t g a g c a	t t t g g g a a a c	a g a a c a g a a a	t c a t a g a a t g
	40201	g c a g g c t t a a	t t c t a g c t c t	g t c a c c t a c t	g g c t g t t t g t	c a t t a g a a a a	a t t a t t t a c c
5	40261	c t t c g t g a g a	t t c a g t t t c c	t t a c a t t t a a	a a t a a a g a a a	a t a t t c g t c c	t c a t a t t g a a
	40321	a t g a a t t g g g	c t a t c a t g a t	c a a c t t t a a a	a t a c a a c g a a	c a g t a t a a a t	g t c a g g a a t t
	40381	a t a t g a c a t t	c g g g a c c t c c	a c t g c c a c c c	t c a c t t t t c t c	c c t c c a g t g g	t c a c t t a c t c
	40441	t c t g t c c c t c	t t t c t g g g t c	a g a g c t t c t g	t t a c t c c a g c	c t g g g c c t g c	c t t a a g t g g g
	40501	g a c a t g t g c t	g a t c c c t c a c	a a t g c c g g g t	g a c a a g g a g g	g t t t t c a a g g	c t g g c t t g a c
10	40561	t g c c a c t g c t	g g t c t c t c t t	c t c a t t t g c a	a c t g t c t t c t	c c c c g c t g g g	c t c a g t g t t c
	40621	c t g g g a g g g t	g a t g c t g a g g	g a g a a a g c t g	t g g c a g a g g g	a c g t g g c a g g	g t c a g a g a c c
	40681	a c t g a t t c g c	a g g a g c t g g c	c t c a g a g c t a	g c c t t t t t g c	a t t g a t c t a g	g g a a c c a g t g
	40741	a t c a t a g a t a	t c t a t g t t g a	c g c c t g t g t c	a a t t t a t c c c	t a g a g c c a t t	a t t c a g t g a a
	40801	t t t c c t a a g g	g g a a a a c a a t	t c t c c a g c a t	t a t t t t t c t t	t g g g c c a g g a	g a g c a c t t t c
15	40861	t c t g a g t t t t	a c t g g c a a g c	t a g a t a t a t t	c t t g a a a g g c	t c c a g c a g c a	g a g t t c c c g t
	40921	c t t g t t t a g t	t c a a a a a c a g	c t c c t g g c c c	g t c t c t a a a t	g g t c t t g c t a	a a a a c a t c t c
	40981	c t c c c a c c c a	t a g a c c t g a a	c t t a a g c c t c	a g a c t g c t a t	c c t c t c t t t c	t g c c a c t g t g
	41041	a g a g a c c t a t	g c c t c t t t t c	t c a t t g g c t t	t t g c c t g c c c	a g c c c t c c t c	a g a t c c t t g t
	41101	a c a t c c c t a g	g a a g t a c a t t	c t t t c c a c t c	c a c a c a c a t a	a t t g t g c a a c	t t g t g c t a g g
20	41161	a c c g c a t g a g	g c a c t g g a g c	t g c a g a t a g c	a a g g a a a t g t	g g g c c c t g c c	c t c a g g a a c t
	41221	t g a a a t c c g g	t a g t t a c c t a	g a g c t a c a c t	g a g t t c c c t g	a c g t g g t a g g	a a g c c c t c c c
	41281	a g a g c c t t g t	c c g t g c t t a g	a c g t t g c c t t	c a c a g a g g t g	g c t a a g g g g c	a t t t t g t c c c
	41341	t g c c c t a g t t	t t t a c a a g t c	c c c t g g a t g t	t a a c t c c t a c	t t g c t t t t a t	t t g c a g g t t c
	41401	t c c a g t c t t a	t g c a t t c t c t	c t t a t t c c t a	a a a a t t t c c a	a t c c a g t g c t	g a t a g t a t g t
25	41461	t a g a g c t g t a	g g g c c a g g a a	a c c c t g c t g g	g g g a a t c a t c	g t c t a g t a g g	t g g a g t g t g a
	41521	g a g a g a g g a g	a c t c a g g c c a	g a g g g g c t t c	t g a g t t c t g g	g c a a g t c c c t	t a c c a t c c t a
	41581	g t g c a g c t a t	c c t t c t c c a t	g t c c t a c t g a	g c t g c c t t c c	t t c t t g c t t c	t c a t t c c c a a
	41641	a g a g a g g c a a	c c a t g c c a t t	c t g g g c a g t g	g g c a a g g g c t	g c t g g a a t t g	a g a t t t t t a t
	41701	t t t t c t c t t t	g g c a t c t g a a	t c c t t c c t t t	c a g t t g c t g	a c a t g c a g t t	t t g t g t g a g a
30	41761	t c a c c c a t g t	c a t c c a t c a t	c t t c a g g a c a	c c c a g a a a c t	c c t c c t a c c t	c t c t c a g c c t
	41821	c a c a t g c g c a	t c t t c a g t c c	c c t t c t g a t t	c a c c c t a g t t	g g t c t a t t g c	t c t t g g c t c t
	41881	g a a t a a a t a c	c t t t a a a g g g	t a c a g g t a t c	c a t a t g g a a g	g c c t t c a g a g	a a g a g a g a g c
	41941	a t g g a a t t t a	t t t t c c a a a g	t g g g a c a c t c	g a g a g t g a a a	g g g t g a g c t g	a c a a t a a t g a
	42001	t g g c a g g a c a	a a a g c a t a t t	t c a g g c t t c t	t a a g a g c t t c	a t c c c t t t g a	t c c a g t g a c t
35	42061	t t a t t t c t a g	g a a t c t t t t a	t g g a g t a c c t	a c c a t g a g c c	a g g c a c c a t t	g c a g g t g c t a
	42121	a g g a t a t c a a	a a c a c a t c a a	a a g a g a g c a t	a c a g t t a a g t	g g a g c a g a c a	g g c a a g a a a a
	42181	a g g a t g g t t g	t a a a a g a a t g	c g g t a a g t g c	t g a g a a g t c t	t t c t c a g a g a	c g t g a t g t a t
	42241	g a g c a g g c t c	a g a c a g t c a a	a t a g t a g t t a	g c c a g g g a a g	g a a a g g a g a	a a t t c c t t c t
	42301	c t g g c a c a g c	g a g t t g t c a g	g g g a g c c c t g	g c a g a g g t g a	g a g t a t a c a t	g g t g t g t t t g
40	42361	g a a a a c c c a g	g t c c c t c g c t	g a g g c t g c t g	t g c a c a g g g a	g a g g a g g g c t	g g c a a g a g a t
	42421	g a g g c t g g c a	a g t g a g g c a c	c c a g g a g c c a	g a c g a t g a a g	g g c a g a c a a t	g c t g a g g a a t
	42481	t t t a g c t t g a	t c t t g g a g g t	t a t a g a g a a g	a a c a g a a a t t	t c t c c a g c t g	a g g a a t g a c c
	42541	c a a t t c c a t t	c c a g t g c g g g	g a a g a a t g a	t a g g g a a g g c	t g g g a g a g g g	c a a a a t a g g g
	42601	t c t c g g a a a c	c a t t t a a g a g	t c t a t t g a a g	t a a t c c a g g c	g a a a g a g g t a	g c a t g g a c t g
45	42661	t a g a g a a a a g	g a g g g t g g a t	t t t g a g a a a t	a g g a g g a g g c	a g a g t c a a c t	g g a a t t a t t a
	42721	a a g a t c a g t t	c a a t g t g g g g	g c a g g g a a g a	a a g g a a g a a a	g a g g a c t c a a	a c a t g a t t t c
	42781	t g g t g c t g c c	t t t t c c c a a g	a g a g a a a a t c	c g a g a g g c t g	g a g g a g g a g c	a g g g t t t g g g
	42841	g g a g g a t g g t	g a c t t c a g c t	t t g g a t a t a g	c a c c c t t g a g	a t t c c t c t g c	a g t t c a g g t g
	42901	g a g a t g t c t a	g t a g g c c a t g	g a t g t c c c a t	g g a t g t c t a t	g g g t t t g g a g	t a t a g g a c g g
50	42961	a g g t c t g t g c	c t g g a g a t g t	g a g c a t a g a c	g t a t c t a c a c	a t c a g c a g g g	t t g a t g a g a t
	43021	t g c a c a g t c a	t a t c c t c a g a	t g a g a a g g g a	a a g c a g a a t g	t t g g a g a a g a	a t a t g g a a a g
	43081	a a t g c a a g g g	a a g t g c t g g t	g g t g g t a g g a	g g a a g a c c a a	g a g a g g t g c c	g t c t g g a g g a
	43141	g g g a t c g g a g	t g t c c a g t g a	t t c a g a g a g c	a t g a g g g a g g	a g g g g c t g t g	g c a c a c c t g c
	43201	c t g a c t t t g a	a a t g a a g a a t	g g t a t t c t a g	g t g g g g c t g t	t t t t t t a g g g	a c c t g g t g c a
55	43261	g g t g g g g a t g	a t g g a c a t a g	g g a a t t c t a c	t c t c t g g g a g	a c a g c a t g t g	g c t g g g a t c a
	43321	g a t a g t t a c a	a a g t t a t a t g	a c a a c t a t t t	t a t t t c a t t t	c t c t a a t a c t	g g c a a c a t c t
	43381	g a a c a t t a t c	t a c a g a g g a a	g g a a a c c t t	t a t g g a g g g g	t a g g t t g a a a	a t a t g t c t g g

43441 gccaggcatg gtggctcaca cctgtaatcc cagcactttg ggaggctatg gtgggaggat
 43501 tgcttgagcc caggagtttg agaccagcct gtgcaacatg gcaaaacccc atctctacaa
 43561 aaaatacaaaa aaaatagctg ggcttggtgt gcattcctgt agtcctagct acttggaagg
 43621 ctgagggtggg aggcacacc tgagcccaag aggtcggggc tgcagtgaagc tgtgattatg
 5 43681 cccctgcact ctagactgag tgacagaatg agactctgcc tcaaaaaaga aaaaagaaga
 43741 aagagagaga gagagagaga gagaaggaaa gaaagaaaaga aaaagagaaa gaaagaaaaga
 43801 aagagagaaa gagagaaaaa gaaagagaaa agagaagaaa agacaagaca agaaaaagaaa
 43861 agagaaaaata cacctgggat ggaggccctg atggaagaag gtttctgagg aaatgggtgt
 43921 ggggtcccaga gtgtcagtat agggcctgtt ggagagcgag ,tatgataagg tcagccctgc
 10 43981 cactcagaag aaattctggt tatactcaca gtaatcacac ttctgtggtt acccagggca
 44041 gtcctaccgt gcaaccatta ccccaacaaa attattgctt acaccccttt tcaacttttga
 44101 aagtttaggt gataaagtat ctggtctcct cagagtatac atctaaggag attcgtagggt
 44161 ggagtttagac agctgaggga gctcctacct gaaggactct gtgcagcagg aggcaggatt
 44221 ctctgtatga tgttccaaga ggcttgagga gataggagca gctgtgaggg gaggggaagac
 15 44281 ctgaaatgag aaaaagagca aattgttgat ctacagtaag ggctgagcag agcagatcct
 44341 tgagaacatc aatttgctgg gcaccaagta gttcataggc cgcgtttccc tgtcttttagc
 44401 agcctgcgtg tctgctggga gaatagatgt ataaactgat ctgaggttgg agctgcacag
 44461 ggcagatgtg gccagagggt aaggaagtgg ctggagtggg ggatcatgag aatagcacac
 20 44521 ccgtagtaag ctctgcacat ggagatactt attgcagcac tgtttatggg agcagaggat
 44581 tcaaaatagt aaatgtccaa caagggggtg tcatccgta aattatggca aattaacgtg
 44641 atagaatatt atgtggccat caagattttg tttaggtaga attgtaatat aaggagagaa
 44701 agcctagatt gtaatagtaa gtgaaataag aaagtccag aatgatgtat attaaacat
 44761 cccgcctata atgcatagag atttttttat ttatttttaa aaggttggg acaaatattc
 44821 cataataata cctgtgggag ctgatagatg atatatattt gtatttttaa attttcttc
 25 44881 aataagttca ttatattact tttataataa ggaaaaaaca tctctctagg aaaaattatt
 44941 taaggaaaaa ataaacatga aaaaggggtg ctgtaatcac agtttaattt ggccttctca
 45001 gaaaaccatt gtagggagtt tttattctct agtttcccc ggttggggaa gatgagcagc
 45061 atggccctag tgtttgagaa cgtgggtttt gatatcagac atgcctgggt ttgaatccca
 30 45121 gctccgctaa gctgagctat agtattgcca tgggtgtgac tcaagcaggc tatccttatt
 45181 tatgaagtgg aataatacac atccccagtc tcttctttat aattccgaaa ccaaaaaagc
 45241 cctgaaaacc cagagtcgtt tcttaaaagt acagcaaatt catttggtta taaaacctaa
 45301 cctgaattga gctagtcttt atttcacttc agtgaatata attttgcagt ggaaatatta
 45361 acatgtttta tcattggatg ttccgccaga tctcacggag ggtatgagat aatacgcagg
 35 45421 gttcaggtaa gttgtgaact acaggatctg cgtcataggg ttattgtaga atgaagtcag
 45481 gttgtgattg caggggttctt agcacagttc ttgccagaga gtaagtactg acaaagttag
 45541 ttatcatcac tattgatgct ataattattc gtgagccttc actaagtatg gagtgcata
 45601 tttctgcttg cccctcttcc ctccatttcc ttctttattg ggatgttga gaagattaat
 45661 gtatcacata taatgagaat ttaggatggc attgtttagg acacaggact gcaaaaaagg
 40 45721 gaagtccag cacagatatt ttctgcctc ttttcttatt ccagatttct cactttctgg
 45781 ggaattagct gtagatata atgcacatta acaggacacc caaatacttg aagagcta
 45841 ggagaagtta cccttactcc ctgtgtgact gctttcgtct caaagcacat gtgttcatta
 45901 taaaatgtta atgtatttac accatagttg ctgtacattt aaaggtttac tgtgttttaa
 45961 gctaaaacta gctggatata aaatgtgctt ttaattgaaa agtggttatt tactgctttc
 45 46021 cataggaatc cacaataaga ttttttcatt gcaaaccctc ctagtatctt taaaatgcaa
 46081 ttcagttcac aaacattgat ttgcattcaa cgtttttctag aagcctaata agagtgcaa
 46141 atgagggtcca cctgcaagac tacagttatt acttctctct ttttcttgga agcatccatg
 46201 cattgtaaat tctgtttgtg tggcatcagc caactacatt attaacatca ccaggaggaa
 46261 ttaaaatgta tcatgttata ggatcaaaag tttcacattc tgaatcaatg ccagtataaa
 50 46321 aagaatgtcc agacccttcc gtctgactac agtcggcagg gatccccgcc ccagcagcca
 46381 tgtgactcac ttgcattgtt ggtgagggtc actcatggct catacggggc tgcggccgga
 46441 taggcctccc catggggcag catttagcct tttcaaaatc agcttaggat cacagacctt
 46501 caggctcttga aggaactttt tggagtttcc tctattaaac ccagttttaa tctggtcatt
 46561 tagactgcat gtctagaaac actacagcat ttttagtagt aaactgaaat aaatcttcaa
 46621 ataaacgact ttaacatttt gcacattggg ccagggggcag aagagctcgt ctgttgctt
 55 46681 ttaggagggc tctaattctt ttctgtttca gctcatcggg tcctcattgt cccctactgt
 46741 atgtgcccag cacaatgtca gggcacattt ctacacaatg ttggattcca gctgatctc
 46801 cgtgttaatc cattcgtagg tgccaaattg cagggttcct tgccaagccc gtcagcacia

	46861	tcagagcttc	tgaagtcagg	tacctcaatt	acattagttc	agttgtctca	ttagccccc
	46921	tgcattcatcc	aaactcatca	tcaaactgtgt	tttccttcat	tattgttgtc	atittcttaa
	46981	taccagtctt	gttgaaaaag	gattgtttta	gatacaattt	tgcctaactt	gtttcatctc
5	47041	cagaatacaa	tcctaaatag	agccaagatt	tgccggagag	aggagatagc	ggttgaggct
	47101	ggtgaagctt	ccctccggtt	agaattagag	agaagtggat	gcagaacttg	ggccagtaat
	47161	ggacccatcc	atctccctag	tcaagaggac	attcttgcag	ccagagggtg	ttccgcttgg
	47221	gaaatgatct	actttgtgta	aaccggtgca	tcacagatac	tgtagagta	tctcacagat
	47281	actgatacag	catggtctgc	attttgtaga	tatgagggtc	ctcccttagg	tttacagccg
10	47341	gataaaggag	ctccacttac	ctggtgatca	caggcatgtg	gtccaccctg	ctgggctcta
	47401	ggactttgaa	catagagaag	gctatctttt	aggaagactg	caggatcagg	agtcgggaag
	47461	cagggattct	gctttagatt	ttacatctag	cagtaacaac	ttttagccca	ttcctcacct
	47521	actctgtgcc	ctggatttct	tggctttaac	atagcaaaa	cctctcttta	tagttttcct
	47581	aactccatct	ctcccaaaaa	agctggagat	gtaagtttta	aattccagtc	cttggcacat
15	47641	agtaggtgct	caatctctca	tccttctgtc	cctcaagata	gccgatggct	cagccatata
	47701	tgttccagac	aggtctctct	gtgtcttttg	ggaaagcctt	gttgctctca	gactggcctc
	47761	taggcctcat	gttttccac	ttgctcttgc	cttcagtgtt	aggtcatggg	cttgccctcat
	47821	tacctgtttc	aaggaggcct	ctatttgatc	cctggagcat	ccctgggagt	ctgaggcctt
	47881	taaggatccc	tgtggtctca	aggaaaaga	actataagta	gttctcaggc	ttccttgatc
20	47941	aattcaaa	catcttcagc	tcttccagtg	cagcccctcc	agagctgttg	aacgccacct
	48001	ttttcctctt	tcgggaaacc	cacaaaacct	tgcatggggt	tgtgctctgt	atggttttcc
	48061	agcatcaaat	tcactcatca	aaagatgcat	caggagggtg	gaggggggtg	cgctgggggt
	48121	gaggggtcaa	caaggaaggc	cagaatgcct	gtttgcattc	gcaacgggaa	cattggaagt
	48181	tcgggtggaa	aaacaatccc	accggccttt	atttaccaca	gtgggtagcc	tcgggaggag
25	48241	ggggagctgg	gtgggaaagg	aagacaatag	aatgctgtgt	tgccctgtgt	taaaagcttg
	48301	ctctaagcaa	ataaagtggc	aagctcagtc	agagccctgc	tcaattagcc	ccagtgccta
	48361	aaggagggtc	tccttgcagc	tgttggcggc	gggcgggctg	gagcaaaggc	aggcgactc
	48421	tggggcactc	gggaggcgaa	ccggcaggaa	tcttgcattg	gagctgacct	gggagggaga
	48481	ccaaaggacc	ctccaacctg	atcccagtc	ctgcttcttg	aacagagggc	tacagaaggg
30	48541	gttgggtggg	ccactctggg	gagaggcagt	gtggggagac	cacggaggag	gtgaccagaa
	48601	aagttagtgc	aaacgtttca	tgccgagagt	gaccgacata	tggaacgtgt	tatccggaag
	48661	ggttacaggc	tggaggagct	cctagagttt	ttcagggggg	gggggtgggg	agggggaagc
	48721	cttcagaaac	tcaaggaggt	ttctaataca	aaggaatttg	agaagggttg	gagaaataat
	48781	gagaaatgca	aacttgacct	agagtttctc	cttccatctc	acggcttctc	ccctgcctgt
35	48841	cctcacatgt	ttccagttca	tggggaccca	ggaaggccac	tggagccctg	tgccctgactc
	48901	cacgtgcacc	tcactggggg	tgtgggtggc	ggggtaggga	ggacccgcag	agctggctcg
	48961	ctgctctgtg	ctgaaaggga	cccagagagc	gagagccctg	cctggcttta	gacccctgtg
	49021	gactagcagg	ctgctagcgc	cgggatggtc	tctggatgat	ttactccggt	tcctgcctt
	49081	gctggaggag	caatttgatg	ccggactggg	agtgaataac	agacttgccc	agggactcac
40	49141	agtggccaaa	ggggaggctg	ggagtagacg	caggctcttg	atcccttctg	ctttttggca
	49201	tctccccttc	tcagcatccc	ttgccagca	caccacgcac	acatgtacat	acacacgcag
	49261	acacacacac	caaacagatg	cacacagaat	aacaacacca	cacctacatg	cacacataca
	49321	cagacataca	tgacgcacag	acacacacag	aataacacca	cacatacatg	cacaccacag
	49381	atacacagaa	taccacacac	acatgcacac	acacagacac	attcacagaa	tacatatatg
45	49441	cacacacaca	catacatata	tgcaccacac	agaataaac	cacacacaca	tacagacaca
	49501	cagacataca	catacatgca	cacatacaga	tatactcaca	gaatacacac	atgcacacac
	49561	atgtatagac	atacaccaca	caagcagaat	aacaccacac	gcacacacca	cacacacaca
	49621	cagaataaca	ccacacacat	atgcacacca	tacagagaca	cactcacacc	tagacatata
	49681	caccacactc	atacagagac	acatgcacac	acatgcatac	tacacacatg	cacatgcaca
50	49741	cacaaaaaca	tacacagaca	tcacacatgc	atacatatac	acaacacaca	gacacaccca
	49801	cgggcttaca	cagacaccac	acacacgtag	atatgcacac	agcacacaca	cacacagaga
	49861	aacagacaca	ctcgcaggca	tacatgcatg	catacacaca	ctacacacag	acacactcgc
	49921	aggcatacat	gcatgcatac	acacactaca	cacagacaca	tctgcagtca	tacacaggca
	49981	ctacatacat	agagacatac	aatgcacact	cacacacaga	cacacataga	cacatgttgt
55	50041	gtgctcatgc	gtacacgagc	gcgcacacac	acacactgcc	tgacttgttt	cccacaagac
	50101	gggtactggc	ctgtcgttgg	tagccctcct	gtcccagcct	gtgttggcca	gggtgccagg
	50161	cactgccacc	cctcttggga	caaggtacag	gtggccagtg	tgatcacggc	ctgttctgag
	50221	agcctcttcc	tgagccagga	agcgtctgtg	gatactgagt	gcccgctgct	tcgtctttcc

	50281	catgggtgctg	ggtctttggcc	acctgctgca	ttgatagcac	ccgcatgttc	acttccctgg
	50341	cagtagaaaag	aaatgtaggt	tagtgcaggg	aggtcactgg	ctttggagat	ggtggcaggg
	50401	tggagcagtc	tctaagtga	atggaagtgc	acatgccctc	tggacctgca	gcaatgctgg
5	50461	cttcagaggg	gccttctctg	gtacctttcc	aattttcccc	aaccaggtag	aagagccact
	50521	gcccagtgctc	ttgggctcaa	ctgaaacccc	ctgtacaaga	aagagacccc	cctttccact
	50581	gtgtctctct	cctcctcccc	actcactgtc	tctctgtctc	cctcactccg	tgctctctct
	50641	ctctctttctg	tgtctctctg	cctctccctc	tctctgtttc	tctctgtatc	gctctttttt
	50701	ttcaatctct	gtctctatct	ctcctccttt	ctctttatct	ctctgtctct	ccctctgtct
	50761	ttctctctgt	gtgtgttttt	ttttctgtct	ttctgcatct	ctgtttctgt	ctccttctct
10	50821	cggttctgta	tctttgtccc	cctctctccc	tatctctgtt	tctctgtctc	agtccttttc
	50881	tgtctttatg	tctctgttta	tctctctcag	tctctgtatc	tctgtctctt	tctcttttta
	50941	tcctttctct	gtctctttgt	tctctatccc	tccatctctc	tctccctctc	cctcagctct
	51001	tctctgtctc	ttcctttctc	tgtctcagcc	tctctctgtc	tctgtatctc	tatctctttc
	51061	tctctctccc	agtctctctc	tttctatctc	tctctctctc	tctctctctg	gcactcactc
15	51121	accacttac	ttgaagtctc	catgagcagt	gggtggcttc	acctttctgt	ttcgccactt
	51181	tgcaagtcca	cccgtggggc	ttgcagaccc	tcttctggcg	cacaccttca	ggagaccaac
	51241	ggtgccaggg	cactcccgtg	ttcttcaggg	ttccagcccc	gagtagttgg	taaacaatcag
	51301	taatcgtcct	agagatccac	tgtagattcc	tcatccaggt	actgaatgag	gcccttctga
	51361	gcaaatttaa	tggaatgacc	ttggtgacat	tacaagatga	cggctcatct	cctgtagtct
20	51421	attattgttg	ggcatttagg	ttgatactat	gtcttcaacta	ttgtgagtag	tgctgcagtg
	51481	aacataatg	tgcatatgtg	tctttatgaa	agaatgattt	atagtctttt	gggtatatac
	51541	ccagtaaatg	gattgctggg	ttgagtggtg	tttctatttt	taggtctttg	aggaattgct
	51601	taggtctttg	aggaatcgca	tttgctgttt	caaacaatcat	ggagaacaca	ggtcttgagg
	51661	atgtgcagga	acatggagac	caaggaatga	aaaagccatc	agggtggaag	tgaagtcacc
25	51721	cagagtgtt	gcaggtttat	ggacagaaaag	gaaaatgtct	tgattcctgg	tgctcatgag
	51781	gaaggtggat	gccaatagat	ggtgctacac	agacacttgt	gaaaggaaga	agggaaagag
	51841	tctgcctcca	tgggctgagt	agggctgctg	gggcctcagg	cttcacacat	agtgctcagac
	51901	atgctgtgtt	ggtcagggtc	atctgggacc	cctgtgcctc	ccccaggaaa	ccctattggg
	51961	aggtgtgggc	aggcttccct	ccatgtgtcc	ttaaatacac	tgactgctct	gtgtgtgact
30	52021	gtggatgagc	agcgttaggc	tggtcatctg	gagggctggg	cttgggtcca	gctctgttac
	52081	tgtaggattg	attggttttg	ggcaagtcac	ttccccccat	aacctttcca	ttgtaccaca
	52141	atcctaatag	agttggatgc	cagtttctgg	caccctgggc	ttcacaggga	cacagagaaa
	52201	cagagaggac	cagaaaatta	ggctgataat	aatcatttct	tctctatttg	gtctggagaa
	52261	gaaatgactt	agggttggca	cagggttaata	tgtcaactag	aaagaacctt	tgaaaatctt
35	52321	tagttcaaga	attctaaacc	taaaatctgt	gggacactag	gatgattttt	atggactctt
	52381	tgaatcacc	tgaattttt	taaaatattt	tttctttatg	ggcacatgtg	tggtttttta
	52441	gcagaaggat	tccttagcat	tgattacatt	gtaaaagggc	caacaatcat	aaaacaatta
	52501	attccccagc	atcatgttat	gtgattcata	actgaggggt	ctccgcagat	ggcatttggg
	52561	aacgtgcggg	tccttttttt	cagttgtcag	catgactggg	gtcacgatgg	atatttagta
40	52621	ggtgggggcc	tgctgtgcat	gcgtggggga	ttaaacttaa	caataaaaaa	ctgtcctgtt
	52681	tgaaatccca	gtagcacc	tctggagaaa	cactgctcag	gaactgagcc	cccaagatgc
	52741	acagtgattc	tcccgtggcc	acaaagctca	ttagtggcag	agttggtttt	tattaacaga
	52801	tcaaaacagg	ggatgtgcca	agaacctaca	ttattttgtt	tctcatcatc	agctgcaaat
	52861	gtgctgcaat	ctgtgaaaac	aaaagaatca	gaaatctgtc	ccaccttcaa	cgctccagcc
45	52921	attttttaaaa	atgagtcttt	gggttcttgg	ttgtttttta	ccatcaaattg	agaaatgagg
	52981	aaaggaatat	tttacatttg	gagaaactaa	accataaaca	attgatccca	ctagccccat
	53041	atcacctaga	tgtgttcttc	agctactgtg	aattggtgac	gcaagcctta	gactggaaat
	53101	ttccccttta	tgtttcagag	gatctgtctt	gaatgtctct	tactctacaa	agaaagaaac
	53161	attaatatatac	cagcatgctg	cttgctacct	agtttactgc	aatgaagtgg	caggtgcctt
50	53221	agacttttga	gtgaaattga	gagatattcc	gcagtattag	ctaagagtta	ggcctgtggg
	53281	atcacagaga	caggggtttg	ttccctggct	ctgtctctca	attgatgtgt	cttcttaaga
	53341	aaaagtactt	aatccttctg	agcctcagtt	tccttggctg	aaaagtgggg	ataataatcg
	53401	tatccacaaa	gattaattga	gatcatccat	gtgaagtgtt	ccgcacagtc	tagcatatgg
	53461	ggaggctcaa	taaatgtgag	ctgttattac	aagcgattat	tatgacttgt	gtctgtgatt
55	53521	aaagacagcc	tgaggttcag	agtgtttgtg	ccaaaattgc	tttcagagga	caagctctgg
	53581	gttggttttt	cagtcttgcc	cagctgccc	gagtcactag	agtgtttggg	gctggagtca
	53641	cagtgtttat	taagtgccta	ccacacactc	cagttttgat	ggaagattgg	ttttcttctt

	53701	gtgtgcacat	ccccgcatca	ctacatgagt	gggtgagaag	agactcagaa	aagcaccaga
	53761	cacatgtttc	ttctgcctca	gttccctagt	ctgtaaaagt	actcaagagc	atgatggcgg
	53821	ccacagggaa	cagaggggag	agacacctgg	aggagcaaaa	gacagtccca	gctgtcagtc
5	53881	ttcagacacc	atggttttct	ggagaaggga	tccagaccag	accgccaatc	aatatagccg
	53941	cagaccgcca	gtcaatatac	ctcatccttg	tgaaaggcgg	ttgtctgtgt	ggcagcaggg
	54001	agacaggagg	ggccatattt	gacacaagcc	tgctggccaa	gtctaaagga	gtcagcctgc
	54061	catctgacca	cactttctgc	agccaagtcc	tctggggggc	aaatcagtat	ggtttgagtt
	54121	tatccccgac	ttgtctgctat	ggttgatgca	cacaaaaaag	gtttggccat	tgccaagcca
10	54181	gctgtgtgtc	tggtttgcaa	gaggcaacca	aagtaagaat	tgaagggaag	gcatgctttg
	54241	ggatttccct	tagccttcaa	ccctctagag	gaagccaact	tctttattct	gccattttga
	54301	gactcatgct	tgctctgggtg	gataagataa	atactttcat	cattatcatc	attattatta
	54361	attataatta	ttgaaaatca	attactcagc	agcccctacc	ttgtaatgg	atcacttaca
	54421	gttgtatagc	attttgcaga	caacgtatac	actagaccat	aaactcagct	cgaacataag
15	54481	cataaattta	aaacaaaaac	attgacttgg	gcagctatta	tctgaggtca	aagtatttct
	54541	ttaattccag	ctgtaccctc	agctgcctat	gtcgtctgaag	caacttcccc	ttgtgtgttt
	54601	tcacacttag	agattggagt	aatccagaca	cagaaaatga	tacaaaaatt	gaaatgtttc
	54661	tgaacgttaa	atgattcaat	aatttgaaaa	tagtttcaaa	tttttcaaaa	ttcaaaaaatt
	54721	tgaaaatagt	tcgactagag	cccatcctca	cctcaggtag	ttacgtgggtg	tgtgtgtgtg
20	54781	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtagg	cacacaaacc	catgcatgca	cacatgctatg
	54841	cacagtgggtg	gatgagtgtg	agttaccaaa	acaaatacca	ctaaatgcag	gataacacca
	54901	tgatggctag	tagctggatt	gactcagatt	ttctctgcag	atccttccta	gtccaccctc
	54961	cttccatgac	atgggagttt	gtgatatgca	aaggataaag	taaacctgag	tttctggggg
	55021	cttgtgtcgg	cataccaagc	tcctttaaga	tctctctttg	ccaagatggg	gctaagccag
25	55081	aggcttcctt	ggcctgggga	ataactgtcc	taccttcctc	cctggagcgg	ctggcgcaca
	55141	ctgttttgac	gtgtgccatg	ccttccatct	tttacggcct	cggggccaga	ggaggggatg
	55201	ctttgaacaa	tgtgaaattc	ctctcctaac	agctgtgcaa	aggaaactca	cagctttcta
	55261	tctccatctc	ggtccactta	gctttctttt	ggggtatagg	tttcttgtct	gttcaggggtg
	55321	ccctgttcct	cactttctgg	ccttcagaag	gaccagtttc	agtacttcct	ttcttaggga
30	55381	ggcaagtccag	tcttctgagg	ttctctgttt	tattttattt	tatcctagag	tatgccttca
	55441	ctgttgaccg	tgtcttgccct	cattttctgct	aagaagccag	catcttctga	aaattgagtg
	55501	tcttgtcctg	tttcagatct	tgggccgagg	attgaccatg	tatgttagtc	agtttcagag
	55561	ccctgagttt	taagtgcata	gttttaagga	gaggaacaac	agtagaatta	gatatcttaa
	55621	agaacaggat	aggaatccac	ttccctgccc	cttctctcac	ctggtttgac	actacataga
35	55681	tacagacctg	tcctgagttc	agggattgtc	taatgaggcc	tcaccaaata	ccaacagaga
	55741	gactagctgt	ttctcctcta	gctcactgac	tgcttctagt	tttccgtgac	tttagagcat
	55801	gtgcgtcagc	aggttttggg	gaccatctct	ggcctgcttt	ctgatcttgg	ctctgtccct
	55861	caatttatgt	cttttttgtg	ctccataaga	cagggaagtt	catggttgcc	actacccgag
	55921	aggtgtgcata	aagtgaatga	ggtggaggac	acagagctag	caggaagaa	ggtgccgtgg
40	55981	aaaccacac	agctttattc	atgtgagctc	atttctcatt	tactccttgt	agcaatgccca
	56041	agtcaccccc	tcccttcac	cctttccaag	catttccctc	ccttgttgg	tttggatcca
	56101	ttgtaacccc	agtgaatga	cctacatgtt	ttagtttgaa	tctaactcatg	ctatttccag
	56161	gttgccaacc	cccacgggtg	cttcagccta	attggtctgt	cctcactgg	gtctgcagca
	56221	ccttcccttt	gagcgtcatc	tgtgcatttt	aattaatctg	ccactttttc	ctccctcctg
45	56281	cttccggatc	attaataaag	agattaacca	aaaccagacc	aatccccagc	ggaggctcca
	56341	ccaggatctc	tctctcccag	ctctcagctt	tgtcatgggt	cagaacatca	tcatgtggcc
	56401	ccattgacta	attctgaggc	ctcaagacct	atgttcagag	caaagccagt	gtgaattaat
	56461	ttttcaatta	agatttgagg	aacaagcata	gccagtcttt	tatatatggc	cttagtttaa
	56521	tatttccacc	acattccttt	cccttttgta	attctacttt	aaatctatca	ctttcccacc
50	56581	tctctccctt	tccagtcctt	gaacaagaca	ttagggttat	ctggcaggat	gtctcctgag
	56641	cctgtagcat	tagtcagaga	tcatgtcata	gccgaagaaa	ttctccaaag	aatcaggccc
	56701	ttctggacac	agctggccag	cgtcctcttt	gaccccagcc	agagaatgca	tttgcttcca
	56761	gagaatatgg	ctacaccaga	aataatgaaa	tctcttgaat	ttgcatatta	aaaatttggc
	56821	aaattaaaa	ttgctttcac	ctctgtctct	atgactcctc	tctgaggaaa	gctaaaacag
55	56881	agcctttctg	ctttctttgt	aagtcaccca	gatttccagc	gctgggtctca	aaccagggat
	56941	gtgagatgg	gcactccgtg	ttctgtccac	tacaacacag	ggcctctgca	ggcacagggtg
	57001	acctgcagat	gggctgcact	ggagagcatc	aggcccagga	gggactaggg	tgaggggaatt
	57061	ggattgtgtt	tacttttgtg	gggatgctt	tcatagggaa	tacaggagg	aggacacagg

	57121	attgatggcc	tcaaattctct	gcaaactctg	cttagagctg	aggcttcttg	gggaaggaca
	57181	cagcctgctg	acagcagccg	tcctggctcc	aggcttggag	cttgggcttc	ccctggaagt
	57241	aaactctcta	gaatgtttga	gagcaaggac	agacggagag	gcagccgagt	gtcatcatta
5	57301	ataacatgag	ctctagagcc	agaagggtgca	aatcctgact	ctgttgcttg	ttagctgggt
	57361	gtggaacccc	tctttgcctt	ggtttcccca	ttcaaaatgg	gaaatataat	agtacgtact
	57421	tcaggatggg	tgttactagg	atttaagatt	tatgtaagcc	agtacctggc	actttgcatt
	57481	tttattcttg	gtgatctcac	ccctaagcag	tctaaaaatt	tacttcacac	atgtgcagca
	57541	agaggcgtct	ttaggtacac	ggtcaggagg	tgctcctgtg	tagggaagct	ggcctggag
10	57601	aggtgaattg	gaacctgaag	gctgcagccc	agtctctgca	cgtgacagg	cagacgcagg
	57661	ttgctggggg	ctcagattct	gtgccactgt	ctacctcagc	ttccttttcc	atgtcagctg
	57721	cttcttgaat	acccccaaac	cctctgtgga	tgctgtaaac	tgagccaaaa	aggaggaatc
	57781	agaaattgct	ctcttttgag	tattttgagg	gatggtgaaa	ccttgtagga	atgttgctcc
	57841	agggactgag	aaaacaccat	atgtgggaaa	gcgtttctaa	gtgaggaaaa	tcgcacatgc
15	57901	acgtcgggtca	ggatcttgga	aaactctgag	caggggaaat	gggcactggg	gagagtgtag
	57961	aaaggggggt	cactaagggtc	tttaacaggt	ctttttctaa	ttatagccaa	gaagagcctg
	58021	gggacctagg	ccgacttcag	agccaggggc	tgtattttct	ttgggaaaag	aaggcagagt
	58081	tgattggcct	ccaaaaccag	gcttggcaaa	agcccatgca	ggctctccag	ccagtaagca
	58141	gtgcctgggg	taggtctcga	tcgtagcagg	cgctcattca	gtaaatagac	gtcgaagtgc
20	58201	cgagcgtgt	ccagccctat	gctgggagtc	aaggatgtgt	ggatggaaa	gacatgggtcc
	58261	ctgtcgtcaa	agagaaggca	caccaacaat	aatagcacia	tatgacaatt	gcttaaatga
	58321	agctctctgc	agatgacagg	gacccctaga	gaaagagcgt	agacattccc	ctaaatagca
	58381	gaactcgtcg	ctgtgccttg	cctggatgga	gctggcactg	ccttctctgt	tccattcccc
	58441	actccaggga	actggccac	tgggggactc	ccaccccgga	catggagtga	gaaatgaagg
25	58501	tattggtggc	caaaaagtga	ccagagatca	gaatgcatct	attcttgccc	ccagcaaatt
	58561	cagaaacacc	acctctccct	tgcccgtgtt	tgctttcgtc	atttgtctgt	tctctgggct
	58621	tcgagttcat	aatattccta	tcttctccca	caaccagtgt	aactggctct	agatccaagc
	58681	agtcgggcac	catggcatct	gggggaaagc	cttgtgggga	ccgtgttggt	aaagatttgt
	58741	gtgagccaga	tgtagaggag	ggagctctcc	agactctggg	tcctctcgcc	gtgtcgctca
30	58801	agggcagggt	tgcaccatct	cccgtggcca	caggccattc	agcttcatgt	tctctaacat
	58861	ttctagtgtg	ctcactctgt	acctgatgct	atgctaggtg	cttccatgtg	gatttcattg
	58921	tttacccttc	acgataactt	tcattggtgca	caagccatgt	gatttttgtt	gcgcctacga
	58981	accaaagtcc	cccgctcactc	tctgccattg	tttagcagtc	tgaccttaga	cagctcactt
	59041	caccacaact	gcaggcctcc	gttcccttat	atgtataaat	gaggaggttg	tctagggtcc
35	59101	tttccagctg	caaacctctg	tgttttgggg	agttctgtta	gcaacattct	tggacttgct
	59161	ctctacgaga	aaagctagaa	gtcgattatt	aaaatggagt	ccacaggaag	cctgcagggc
	59221	tttaagacag	ccttctaaag	agtttagaaa	cactcaccag	cctgcagatg	acttactccc
	59281	actacttatt	acaaagtatc	atggcaagta	agaaagaaat	ggtcagccac	tgggagagca
	59341	ggctgaggtt	gcaaggaagt	ggcagggtgt	atcggtggga	tcttggggga	gcattgtaga
40	59401	aaggtcaggt	ttggaaccag	gcactgaaaa	acgggtgaaa	tttcaacagg	agaggcagtg
	59461	tcacatgggt	gtgatttgag	cagcagtgag	tagtgaatgg	ggttggagca	agaggcctgt
	59521	gagcccagtc	atcacatgga	aaagggaggt	gagggccaga	gtgcagagaa	cccctctcac
	59581	tttctctgtg	gtctctcctc	ttcttttttt	tttttttttt	ttttgaattg	gagtcttgct
	59641	ctgtcgccca	ggctggagtg	cagtggcggtg	atgttggtct	actgtaagct	ccgcctccc
45	59701	ggttcacacc	attctcctgc	ctcagcctcc	cgagtagctg	ggactacagg	ctgccaccgt
	59761	gccctgctaa	tttttgattt	tttttttaga	cggggtttca	ccttgtttagc	caggatgggtc
	59821	tcgatctcct	gacctcgtga	tccacccacc	tcggcctccc	aaagtgtctg	gattacaggc
	59881	ataagccacc	gtacccggcc	tttctcctct	tcttttaaact	gcatttaggt	tgtcctcatc
	59941	tataagatga	aggaataaga	ttagatggcc	tcttttatatg	gcttatctac	ttctaataaa
50	60001	ctttggttcc	atattccagc	caccacccat	aaagactctt	actttttcct	cctcaatcca
	60061	tcagcatcag	ccagcattct	ccctctcagt	tcactcctgt	cattacctca	ctaactctga
	60121	ccttgactct	tattctacct	gggacctccc	ctcactaacc	atacttgatt	tatttcttta
	60181	agagcggata	tcctgtgtat	tcttagctta	cagtggctga	gcttttggga	ggtttaacaa
	60241	attgtttttc	aatagactct	cagtcctaaa	tgattttatt	gcagctgtat	tattcttttt
	60301	gaggggagta	gagaatgtaa	tcctagttgt	atgccatatt	ttttgtattg	ttcaagttgc
55	60361	attttggttg	gcctgtataa	ttgaagaaaa	tgtcacttgt	atgcagaata	ggaggggagt
	60421	tccttatgct	gtagaggaac	cattttccta	ggatagtaca	attccttaga	gtcatctcta
	60481	agggtgagca	gcaaagctag	aatccctctc	ttggctttga	ccttgaagga	gtcagttttc

5 60541 ccaaagacta gttcccatgg aaagaagatg gtctccttat cacagcagca agaggaggca
 60601 gtatgagcag agtgccctcac tgttggtgac ctctagggac aagtgagccg gcagtatttc
 60661 agacaagctg ggaaagaggc tgtctgtgag tcctggggagc gagtgagcac tggcaggctc
 60721 agacataggt gctgggttagc aggactgctt ttctgtttct tgtgtcggct ttgtttat
 60781 cctctctatt ttccccctgg acttagtaaa gtctttccga aaataccaaa ggtgaaccag
 60841 gggaagagtt tttattttcc atgtttggac agaactttaa agaggaaatg atgtaccccc
 60901 ctgggagcca gtgaggtggc agcgatggtg attaaggagt gaatatctca aggaggtgga
 60961 cgaattcggg gatcactagc tcagctgccc ctctccacct ggagcatctc cttccagtgc
 10 61021 taccctcaga acatctgggc tttgctctag tgagggagag actagcaatg aagggtgtctt
 61081 gagatcagca ctgtaattcc accaggacgc caccggagtc cgggtgttaag cttctactat
 61141 ggcaacagaa tgagagcgtg gatgggttga aatgccattt caacaaggaa atagtagaat
 61201 tagaagaagc ataagcacta aaaggaacat tttgtagcag aaatgttaaa aatctaaagg
 61261 aggcaagtga atcaacaatg actcttctat tctgggcaag tccagcccat ttatgtaagg
 61321 tgggtattct gcatctctgt cttctgcaag tagtgctgtg gcagagctgc gttttgtgga
 15 61381 gagcgtcccc ggggatggag cagatcagtt ggtgatgctg atgtatcaga aagctcggca
 61441 gagcaccctg gaacgtaggc cctctcggcg agtgggtagt ggccctacat gttcatttcc
 61501 aagggcagga gaatagaccg ttccagctgc ggccctggcca gggatgacct cacatctgac
 61561 actgcaatat gggggcaact gaaccagtcc tcagcctcag tgtgttccag gggctgcagc
 61621 tggggagcag tcgaactctt tcttgagaca attacaaggc caccgctgct gctgtgtaag
 20 61681 ggaagttact ccatgtttac aattctcagg tttgaagttt tcatgctttg ccaaggtaga
 61741 gtgaaccatg cgtctttgca ggctcaaggg atgtttaaag aagcggtagg acatcgtcca
 61801 cccacaagca gagaccgcag gataaagcag acatccaatg taaatacaac ccgtgcaaaa
 61861 agcagagtcg gcagacctgg agtgcatctc cagtatctcc cgggggtggg ggaaagaaat
 61921 cacctcttca gaatgtccag aggggagttg ccttgcttac ctggggggcg gtaccctctc
 25 61981 tcgtgccctc acagggctac tcagcctcag gtagctggtg ccagaataac acagactcag
 62041 ctgccagagc ctgctcttaa cacctgtgtt tccttttcag atcttacagg tgaacaagg
 62101 gatgtccatc ttgttttatg tgatatttct cgcttatctc cgtggcatcc aaggttaaca
 62161 catggatcaa aggagtttgc cagaagactc gctcaattcc ctcatatta agctgatcca
 62221 ggcagatatt ttgaaaaaca agctctccaa gcagatggtg gacgttaagg aaaattacca
 30 62281 gagcaccctg ccaaaagctg aggtcctccg agagccggag cggggagggc ccgccaagtc
 62341 agcattccag ccggtgattg caatggacac cgaactgctg cgacaacaga gacgctacaa
 62401 ctaccgcggg gtccgtgctga gcgacagcac ccccttgagg ccccgccctc tgtatctcat
 62461 ggaggattac gtgggcagcc ccgtggtggc gaacagaaca tcacggcgga aacggtacgc
 62521 ggagcataag agtcaccgag gggagtactc ggtatgtgac agtgagagtc tgtgggtgac
 35 62581 cgacaagtca tcggccatcg acattcgggg acaccaggtc acggtgctgg gggagatcaa
 62641 aacgggcaac tctcccgtca aacaatattt ttatgaaacg cgatgtaagg aagccaggcc
 62701 ggtcaaaaac ggttgacagg gtattgatga taaacactgg aactctcagt gcaaaacatc
 62761 ccaaacctac gtccgagcac tgacttcaga gaacaataaa ctcggtgggt ggcggtggat
 40 62821 acggatagac acgtcctgtg tgtgtgcctt gtcgagaaaa atcggaagaa catgaattgg
 62881 catctctccc cataataaaa ttattacttt aaattatatg atatgcatgt agcatataaa
 62941 tgtttatatt gtttttatat attataagtt gacctttatt tattaactt cagcaaccct
 63001 acagtatata agcttttttc tcaataaaat cagtgtgctt gccttcctc aggcctctcc
 63061 catctgttaa aacttgtttt gtgatccggc tctcaggagt cactctgtaa aatctgtgta
 63121 caccagtatt ttgcattcag tattgtcaag gccatgactg ttgttttagt aaacttgta
 45 63181 aaatca

Brain-derived neurotrophic factor (BDNF)

The BDNF gene encodes the brain-derived neurotrophic factor protein. BDNF is
 expressed only in inner hair cells and outer hair cells during the neonatal stage. BDNF supports
 50 connectivity to SGN. BDNF induces synapse regeneration and SGN protection after damage
 (Takada et al. (2014) *Hear Res* 309:124-135; Budenz et al. (2015) *Sci Rep.* 5:8619).

The human BDNF gene is located on chromosome 11p14. It contains 2 exons encompassing ~67 kilobases (kb) (NCBI Accession No. NG_011794.1). The full-length wildtype BDNF protein expressed from the human BDNF gene is 255 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype BDNF protein is or includes the sequence of SEQ ID NO: 31. Non-limiting examples of a nucleic acid encoding a wildtype BDNF protein is or includes SEQ ID NO: 34. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 34 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype BDNF Protein (SEQ ID NO: 31)

MFHQVRRVMTILFLTMVISYFGCMKAAPMKEANIRGQGGLAYPGVIRTHGTLESVNGPKAGSRGLTSLADTFEHVIEE
LLDEDQKVRPNEENNKDADLYTSRVMSSQVPLEPPLLFLLEEYKNYLDAAANMSMRVRRHSDPARRGELSVCD
15 WVTAADKKTAVDMSGGTVTVLEKVPVSKGQLKQYFYETKCNPMGYTKEGCRGIDKRHWNSQCRTTQSYVRALTMDSK
KRIGWRFIRIDTSCVCTLTIKRGR

Mouse Full-length Wildtype BDNF Protein (SEQ ID NO: 32)

MTILFLTMVISYFGCMKAAPMKEVNVHGQGNLAYPGVIRTHGTLESVNGPRAGSRGLTTTSLADTFEHVIEELLDEDQ
20 KVRPNEENHKDADLYTSRVMSSQVPLEPPLLFLLEEYKNYLDAAANMSMRVRRHSDPARRGELSVCD
KKTAVDMSGGTVTVLEKVPVSKGQLKQYFYETKCNPMGYTKEGCRGIDKRHWNSQCRTTQSYVRALTMDSKKRIGW
FIRIDTSCVCTLTIKRGR

Rat Full-length Wildtype BDNF Protein (SEQ ID NO: 33)

25 MTILFLTMVISYFGCMKAAPMKEANVHGQGNLAYPAVIRTHGTLESVNGPRAGSRGLTTTSLADTFEHVIEELLDEDQ
KVRPNEENHKDADLYTSRVMSSQVPLEPPLLFLLEEYKNYLDAAANMSMRVRRHSDPARRGELSVCD
KKTAVDMSGGTVTVLEKVPVSKGQLKQYFYETKCNPMGYTKEGCRGIDKRHWNSQCRTTQSYVRALTMDSKKRIGW
FIRIDTSCVCTLTIKRGR

Human Wildtype BDNF cDNA (SEQ ID NO: 34)

atgaccatccttttccttactatgggttatcttcatactttggttgcatgaaggctgccccatgaaagaagcaaacat
ccgaggacaagggtggcttggcctaccaggtgtgcggacccatgggactctggagagcgtgaatgggccaaggcag
gttcaagaggccttgacatcattggctgacactttcgaacacgtgatagaagagctgttgatgaggaccagaaagt
cggcccaatgaagaaaacaataaggacgcagacttgtacacgtccagggtgatgctcagtagtcaagtgcctttgga

gcctcctcttctcttcttctgctggaggaatacaaaaaattacctagatgctgcaaacatgtccatgaggggtccggcgcc
 actctgaccctgcccgcgaggggagctgagcgtgtgtgacagtattagttagtgaggtaacggcgccagacaaaaag
 actgcagtggacatgtcgggcgggacgggtcacagtccttgaaaagggtccctgtatcaaaaggccaactgaagcaata
 cttctacgagaccaagtgcatacccatgggttacacaaaaagaaggctgcaggggcatagacaaaaggcattggaact
 5 cccagtgccgaactacccagtcgtacgtgcggggcccttaccatggatagcaaaaagagaattggctggcgattcata
 aggatagacacttcttgtgtatgtacattgaccattaaaagggaagatag

A non-limiting example of a human wildtype BDNF genomic DNA sequence is SEQ ID
 NO: 35. The exons in SEQ ID NO: 35 are: nucleotide positions 1-647 (exon 1) and nucleotide
 10 positions 63474-64238 (exon 2). The intron in SEQ ID NO: 35 is: nucleotide positions 648-
 63473 (intron 1).

Human Wildtype BDNF Gene (SEQ ID NO: 35)

15	1	attattaaag	cggtagtctg	ccggcgctga	taagcaacaa	gttccccagc	ggtcttcccg
	61	ccctagcctg	acaaggcgaa	ggttttctta	cctggcgaca	gggaaatctc	ccgagccgaa
	121	ttcagcttcg	ccggagcccc	aggtgtgacc	tgcgtagtgg	gcaagggagc	ggtgtgcagg
	181	ctgagttttt	ttttttacag	gggtaccctg	aaactcctca	ctttctctgg	gaactttcag
	241	tgccaggacc	cagtaacggg	cggttagaag	gcagccctag	gaaacacctg	ctacatagca
	301	gggcagttgg	gcaatcattg	gtaacctcgc	tcattcatta	gaatcacgta	agaactcaaa
20	361	aggaaacgtg	tctctcggag	tgagggcggt	tgcgtaaaatc	tataggtttt	tcgacatcga
	421	tgccagttgc	tttgtcttct	gtagtgcgca	aggtggttga	gagttaagc	ttgcggatat
	481	tgcaaagggg	tattagattc	ataagtcaca	ccaagtgggt	ggcgatccac	tgagcaaagc
	541	cgaacttctc	acatgatgac	ttcaaacaag	acacattacc	ttccagcatc	tggtggggag
	601	acgagatttt	aagacacttg	agtctccagg	acagcaaagg	cacaatgggt	agtagcaata
25	661	aaacctgcat	tataattgaa	aaatcttgac	atgttgctta	acaacgggca	tatcacggct
	721	cttcctagca	cttcacacgc	caaagaacag	cagctactca	ggccagggga	atcgggtttt
	781	tacacagtgc	aactttaatt	ggaatcattt	gagatttgac	acagctatgt	ggaactgcgt
	841	ggaacaaact	tggagctggg	tgggggggtg	tgtgttatat	tgggtgttca	aggctgatgc
	901	ttgtctctca	gcagtcttgc	attctattct	tttccttaat	gtgtatgggt	tatgatcata
30	961	ttctatgatt	tatatgtggg	catgtaattg	acatttgcaa	gggggttaat	ttccatctaa
	1021	aaacaataat	gctgttagag	gttgggggta	gggggtggag	tgggggtaag	ggtggggtaa
	1081	agactgggag	tttaggtgta	gatggggggg	gggggtgggg	ggagagaaat	aagtcagaag
	1141	tgcataatcac	cggtaatggg	taatcctctc	gtagaagaaa	aggttctcat	caacatgtga
	1201	tcaactatta	acaggatggc	tttgggaaaag	ccatccgcac	gtgacaaacc	gtaagggaag
35	1261	ggaagaaacc	gtctagagca	atatcaagta	tcacttaatt	agagattttt	aagccttttc
	1321	ctcctgctgt	gccgggtgtg	taatccgggc	gataggagtc	cattcagcac	cttggacaga
	1381	gccaacggat	ttgtccgagg	tggcggtacc	cccaggtagt	cttcttggcc	ccgctgtaaa
	1441	gccaaccctg	tgtcgcctt	aaaaagcgtc	ttttctgagg	ttcggtcac	actgagatcg
	1501	gggctggaga	gagagtcaga	ttttggagcg	gagcgttttg	aaagcgagcc	ccagtttggg
40	1561	cccctcattg	agctcgctga	agttggcttc	ctagcgggtg	aggctggaat	agactcttgg
	1621	caagctccgg	gttgggtatac	tgggttaact	ttgggaaatg	caagtgttta	tctccaggat
	1681	ctagccaccg	gggtgggtgta	agccgcaaag	aaggtaagca	ccagggcggg	gaccccttgc
	1741	atccccaatt	cttgagctat	tttgatactg	tcttccggag	aggacgctg	gtggagggga
	1801	ggaggtagag	ggagagcatg	agaggggggt	gtttcttggg	atttgccag	tttgaattgc
45	1861	cctaggtgag	aaccgtggg	caaggggaga	aagaaaaaaa	agaaactcag	tcttcttgcg
	1921	gatataatga	gttttagttaa	cttggacctg	caaagtctctg	attcaaatgt	aagattttatc
	1981	tctctttttc	tcctcttcac	ctccctcttt	tccgttctct	ttgctgggtg	gtgtgtgtgt
	2041	gtgtacagta	gattcattac	taattatgaa	gcttttgcaa	aacattcgaa	ttcctaaaat

	2101	ttgactttgt	agcattttaga	atcaggcggt	ggaggtggtg	tgcggtgggg	agaggaggtg
	2161	gaggttggga	agagggaagg	aggtaaagct	aaacctcaa	cacaaaaaaa	tgaatcaagg
	2221	taatttcagc	tcttctagtg	agaaggattc	attctctctg	tatccctccc	tccctctctt
5	2281	tccccctccc	tccctccttc	ccgccccct	tcttccaccc	cgccccctcc	tccagcctcc
	2341	atccctccct	cattctatct	cttctctcc	gtcgcctcg	ctcctcgctg	gatgcttctt
	2401	tctgggtttt	ctttttttt	tcccttctgt	cctccctccc	cgcgagtttc	gggcgctggc
	2461	ttagagggtt	cccgttttct	caagggaagg	ggagctgccg	agaccgcgct	ccgtcccca
	2521	gccggggccg	atgcctcact	gagcccaggt	ccgagtcagt	cggggtaact	cagggaaagg
10	2581	ggagcctccg	cctgggagta	gaaggtcctt	tccggaccga	agagccagag	agcgggccgg
	2641	gcgagggggc	ctggggcggt	ggaggcggtg	gagaagaaca	cttttagctc	cgtgcggcgg
	2701	ctggacagag	ccaccaatca	gctggacgcg	cagaccgccc	tgccaggggc	aggttgcgtc
	2761	cggaggcgcc	ggtggagggc	ggccggctag	tcgctgagcc	gccgcgcca	cccgggtggg
	2821	caggggactg	gcggtgggtg	gaggtgaggg	gcttggcggg	tgagatagaa	gcggcgcgga
	2881	gccgcccaga	cctgtgttct	acctctccc	ccccgcctg	cacccccggg	ggacagcgaa
15	2941	ctgccggaac	gcgcggctgc	gttatcctct	tgccactctt	cagggagctc	agggacttag
	3001	gcgccccctg	gcgggggcca	ccaggctctc	cacactccta	taaccctcac	ccccaccccc
	3061	ttctcaggcc	ttttgttccg	gccacagagc	caagcccggg	ggcagttttc	gcccagggt
	3121	ttggtcgttc	tggactcctc	ccccgatcag	ctcttcttaa	ttaaaaggca	agagccttgg
20	3181	gtagggttga	ggggaacgta	ggaaaaatct	gtttccgaaa	ctcaagacca	ctgttttaac
	3241	gaacgaaaga	aagaatccca	actctgcgca	ggtggattca	taggcgaagc	gaggatattg
	3301	tggaaattca	gaaggaaaa	ataaaaaaca	ggcgctagga	tcagatgacg	gtgataggct
	3361	gctcggcaca	caaagggagc	gtagggcagg	gtttacggag	caagcctgca	gcgaatgggg
	3421	cacagattgt	tccgagatcc	agtcgttttc	tcagtcagat	ctacgcgaag	ggaggggagg
25	3481	ggaggggagg	gcaggggagc	gtggcgggag	gggctgagct	tggggcgggg	gggatttctg
	3541	atcagtctga	tgcaattcca	agcgtgctgc	aaaggaactc	caaggcgccc	gcacaccat
	3601	cgccaccac	ccttcccaga	tgggtgctgt	ttaaatacgg	atctgcaggg	ctgaacgcag
	3661	aactgggaga	tttattgcaa	aatcccggga	ggggcggggg	gggtgtgtgt	gcggaacggg
	3721	gaatggagga	gcagaattta	aaggtgcaac	gcttgctttt	tccaatcagg	cggcaaccgg
30	3781	ccggaattat	tatttttttc	tttctgtctg	cttgtctctg	gatttctaatt	caccaagaaa
	3841	gaggtgtaaa	tattgtgaca	ttttgaggca	gcttgatgga	tgggaaagaa	atcatctgtc
	3901	actctaaatt	gcagagttcc	ctctccccgc	gccatccctt	gctagcgaat	actcgctgct
	3961	gcctaataca	gttgctaggg	cttcaaatga	atgcatcggt	aagggaatat	tatcctttta
	4021	gttgacttgc	caatttagtt	gacagttgaa	tcgagaaaat	tgtagatttc	gtgtctctgg
35	4081	gaggaaaaat	gcttaacagt	ctaagtcttg	taaccttgag	gtctttaaca	acttaataaa
	4141	acctcaaaa	tgacacgtca	tcctctacac	acacacacac	acacacacac	acacacactc
	4201	aacttgtaag	atgacatggt	ttcacctaaa	ctgttggtgga	aatgaatagc	actttaaaaa
	4261	tgggtgcacct	gatattcact	gtttatgtgt	atttacaagg	agctcttcag	catgaaggca
	4321	agacatttca	attgtcctgt	ttggaatcag	tcagaagact	agaagtgat	gaggaagaga
40	4381	agggaaagaa	gaggaagag	agagaatttt	aacctagatg	ctattaaatt	aacagtaacc
	4441	tagcctactt	ttatacccct	tggctctgca	tattaatatt	tctgtatgtg	agatttttagc
	4501	ttggtctagc	tcccccaatg	gagtatacca	gtattgattc	agatgagaat	atgagcatcc
	4561	tgccagtagc	tttttcagtg	tcattgatag	taagacctac	aacacagcaa	tttttgagg
	4621	atagaagaga	atatatataa	gggctttgca	aactgggaag	caggcactcc	ataaatggga
45	4681	ggtatcatta	tgacattctc	tttgacata	tcattttcat	ttcactgaac	cagagtacta
	4741	gttattttta	aacataatct	aatgtatatg	ctcaaggtag	taagtgggga	ttttaaaagc
	4801	aagtgattag	ttggcttata	aaatattatt	tttcaattgt	ctattaatgt	acattggaaa
	4861	gaaggctttt	aaagatctaa	aatcaacata	aataagcttc	ccctttcatt	tgccagactc
	4921	tttccctatc	agatttctga	tctaaaattct	taataagaag	agaagctggt	gaatttagtt
50	4981	tcttttcctt	tcctggcctg	tcctctaggg	gaagctttag	taagaaacaa	cattccaaaa
	5041	tcaggcagtg	agcgagagag	aaggcaaggg	actggatgac	cacaaaatag	ataatcagcc
	5101	aagaaacaga	aatgagggaa	aaccagcatt	aaagcatgac	ttacaaaggg	tttttatttt
	5161	gtaattctgt	aattttggga	ccaggctcaa	acttgctcaa	gtaacattca	ctcgatcata
	5221	ttgcttacia	tctgtcagtt	aaaatgatgt	ctaactgagc	atatttttta	ttaaatatac
	5281	ttctctcaaa	ggccagtaaa	gctactcttt	ggttttaatt	agacaaacta	gtctaaccac
55	5341	ttaaataact	ctaataaata	tgaactgata	tcatacagatt	taaaagctct	gctgaaaact
	5401	aaatttatct	tgaaaagcac	tgacttgcca	gaaaaaatatc	tatttttgca	gctttctttt
	5461	cactctatgg	ataatttaat	gagttgctta	ttttaatttt	acaactgcta	cctcagaagt

	5521	atctcaaaatt	atcttttcttt	ggctgggtgtc	tttctctgtct	gatctgctac	tgctgtgtgt
	5581	gtgtgtgtgt	gtgtgtgtgt	gtctgtgtgt	gtgtgtgtat	ggcggtgtgt	gtctctgtgt
	5641	gtgttttcta	gtgggaattt	aacaagcagt	gagtcctcta	aatttacatg	ccataatcta
5	5701	tgtcaagaac	attgctgact	acttagcaat	aaaaataaac	attagcatct	agtgaagct
	5761	taccatcatt	gagtgcctatg	gaaatagagg	tcttaaagaa	agattaaatt	tttcaacaaa
	5821	aaaatttttc	cccttttttg	cttaaagggtg	attataattt	caaaaatatg	acatctttcc
	5881	tcttttactt	tggaatgtag	agctgctgct	ttaacaagtg	tcttttgaga	aagatacacg
	5941	tgtttcataa	agattaatac	ccttaaaaca	ctatgggtgca	gagagggaag	gatgaattct
10	6001	ttaacccctgc	ctctaattctc	atttggcaat	ttttggagta	ttcattctga	ctttttaaaa
	6061	attcagggtg	attttttttc	tgcttttctt	ccaacattat	aaaacaatcc	tataagagat
	6121	ttttctgcta	tagtgcagac	tttatttgta	tttcctagta	ataacacttt	agattcatat
	6181	agtactttga	cagctctcta	taggtttcat	ttgatttcct	tatcagtcac	gtaaggtagg
	6241	aatcaccaat	caccttttac	agatgaggaa	agtaagggtc	agaattatct	aacactacac
15	6301	tgccagtaag	atgtaaagac	taactcagag	tctttcttca	aattttaagg	aaatttgtgt
	6361	ttgttcctct	ctagaccatg	ctgccttaaa	ctccactagg	gcacagagg	gagctgtagg
	6421	cattattttc	tcctattttg	atttatattaa	ttaaattaaa	aacatttttt	acaaatagtt
	6481	ttcaaaaatt	caggcctaata	ggaaagtgtt	aactagtctt	cctaagacag	tatttccctc
	6541	tcccacagtt	agacactcaa	agaaagcagg	actcttctct	tagttgacat	accatctaag
20	6601	tcatagttgc	taattcccca	aaaaacaaat	acaaagaata	agaagaacca	aagccaagat
	6661	gtacattacc	tttacttggt	aatcatagaa	tctgggttct	gggaagggtc	tcagagggtca
	6721	cttagcccag	ctcatatctg	atacatgaat	tcatgctggc	ttgttttggt	tcaatcctct
	6781	ggaatttcat	cccagtcttg	gacactgggt	acttggttat	cttggtagat	gttctgatac
	6841	ttatgtggat	catgagcttg	gctgatcatt	tcccattttg	cccaaaagg	atactttcta
25	6901	tagagactct	agcattcata	attttatttt	gtaagtatta	attggcaaca	cataatttgc
	6961	cattatgtgc	caggaaactgt	attacaatct	ggagattcaa	gaacaagcaa	gaccaaagtg
	7021	tccctatctt	ggtggatctt	gtggcctaga	gacaaagact	ggtctcttta	tttgctccaa
	7081	tccttagaag	gggacagacc	tctggactca	taagattcct	ttttcaaga	gaacacagtg
	7141	aataagatta	aaagcctgag	tttaggcttt	acttccaacc	gctgtcagat	acccaaaaaa
30	7201	gtcaatcagt	cactcttgat	ctctgcttcc	tcacacataa	aacaaaagag	atgttctcag
	7261	aataaactct	aaattccact	tccgttccaa	gtttgagtga	tataaaaagat	atagtctata
	7321	actatttctt	gcagtgtagg	ggaaattaag	gcctaagtta	ccctaacctt	tggtgattta
	7381	caattctggg	tgggtaccgc	aaatttttaa	cttggttaaga	agtatatcat	gaaaaaaatc
	7441	aaagtaacat	atttgaaccc	agaatagaag	gaatctaggt	gttgaaacct	ttcatttatg
35	7501	gaatgatggc	taggaaagtt	ttaatttaga	gaatgatctc	aaatttctgc	ggattatttt
	7561	taaaagcagg	tggtgcatg	gaagacctgc	tatcaaattt	actgcttatt	ttctttgtgc
	7621	aagcagaagt	tataattttt	cagggtcattt	ccttttaaaa	tcaaaaatat	tacatcctga
	7681	aattgcctgg	gtctcatgaa	taatgcatta	tatacacatg	ataatagata	attagatgga
	7741	caagccaaaa	gaaacatgaa	aggaagcagg	tagcccaagg	attgcagaag	gtgtgtgggc
40	7801	attttgacat	ccaggaaatg	ctatagatct	gtccttaact	aactcagcct	ggtgtgagata
	7861	attaagaaaa	aaaatgtggg	tgtagaaaaga	ctgcaagcca	ttccctggga	ttggctagat
	7921	tgctgcagta	gttcaaaaac	aattggcaca	gccacccaca	ctagacatga	ttgccctttg
	7981	atgaggcagc	tattgacttt	tataaagatg	tcatatttaa	aataacttct	gatgcactag
	8041	gcataacaga	catcattctt	gaattctatt	ttagacaaat	ggcaaaatgt	attacaaagt
45	8101	attaatttaa	aaataaaaaa	tctttaaagt	ctagtgtcta	aaaaccagca	gttttagtaac
	8161	atgcaacctc	tggattttaag	aattcagcct	gaagctggga	gaaagctgta	gcttgtatag
	8221	gacattttga	tccactctgg	gcatttccca	gaccactaca	ggaagtaaaa	tgtactttgt
	8281	caaagttttt	aaccttttag	tgaatgttaa	atccactcca	aaatcttcgc	aacctgggaa
	8341	aggtgatcca	acaattttcc	taaatagcgg	cagaaaaatgc	tctgagatct	ttgttcccag
50	8401	agtgaatgtt	ataatgttat	gctatctaga	aatttccttg	tagcaccatg	ctcatcagta
	8461	ccaaaaggag	ttagaattga	ttcctccgcg	ttcaaggaaa	tatatcaacc	acctcctgtc
	8521	tctaagtaac	aaggttactg	tggggaaaaa	atacacaat	taggtgattg	cagaaagggtg
	8581	tcacaaacat	ccaaagcctt	tgggataggg	cattgcagtg	tgagtgaata	gagaaaagaa
	8641	agagaatgtg	ggaaaaaatt	gagaaataaa	aagggaagtc	acagtggagt	tctaattata
55	8701	caggggctct	tgaattgact	gttctctacc	ttccatgctc	attgttggtc	tggctacttt
	8761	agtaggaaac	aatgattttc	tctgctttca	ccttccctct	ccgctaagga	cttcttactt
	8821	gccaataact	tccataatca	atgtttaaga	attgctctga	tgcccagtg	ggtggctaac
	8881	gcctgtaatc	ccaacacttt	gggaggccga	ggtgagtgga	tcacttgagg	tcaggagtgc

	8941	cagaccagcc	tggccaacat	ggcaaaaccc	tgtctctact	aaaaatacaa	aaaattagtg
	9001	gttcatgcct	gtaatctcag	ctacttggga	ggctgaggta	ggagactaac	ttgaaccatg
	9061	gagacagagg	ctgtagtaag	ccgagatcat	gccattgcac	tccagcctgg	caaaaaaaaaa
5	9121	aaaaaaaaaa	agaaagaaa	aaagaaaagaa	ttgctcactg	taatgacttt	catgccatgg
	9181	actcaactct	cttggcagtc	tggtaaagct	tatgtaaacc	cttctcataa	aaatgtctaa
	9241	atggggccag	gcgcgacggc	tcacacctgt	aatctcagca	ctttgggagg	ccaaggcggg
	9301	tggatcactt	gaggtcagga	gttccagacc	agcctggcca	acatggcaaa	accctgtcct
	9361	taccaaaaaa	aaaaaattag	ctgggcgtgg	tggcatgtac	ctttaatccc	agctacttgg
10	9421	aaggctgagg	cacgagaatc	acttgaatct	gggaggtgga	ggttgcaagt	agccgagatt
	9481	agccactgca	ctccagcctg	gatgacacag	tgagactttg	tctcaaaaaa	aaaaaaaaaat
	9541	tctaagttaa	tgaaataaat	gtataagatt	acaaaggaag	ccagtggcat	tgatgtacag
	9601	ttataaaaaa	atttataaat	atatattgtg	tgatataagta	atataatgc	tttttaatac
	9661	attaaataag	atctaacagc	aaggtaaata	ttataatttt	gaaataatga	taagtatcaa
15	9721	tgtattttga	aatatctata	aaactgacgt	gatatgaagg	tgtctgtgat	gtatactggt
	9781	gagaaaagcat	gcaagtacta	ctgtgttaaca	tttcccacac	atatttaaca	acagaacact
	9841	tgagaagcac	ttattaacac	agcatagatt	cagaaatatt	aatttagtaa	atgtcaacat
	9901	tagccattgt	tgtttccttc	ctggcaaaag	gaaatcagca	ttgggagaaa	acttttaaaa
	9961	ttcacatttg	ccattagaca	agctgtcaag	tggggaaaag	accaaatact	gagaaggcca
20	10021	gggtatggta	agcatgtttc	tattgactga	gcttgctatt	actctaactg	ttatctttag
	10081	catcaccagc	acaaccccat	taccctagca	atccatcact	ccattgaaaa	agataaaaaag
	10141	ttcagattct	ggtcattaac	tcagcattgc	ttaagatacc	tggtctgacc	tcactaaccc
	10201	aagagattac	tgaaactctt	cctgtttgtc	attactacac	catgggaaat	tataatgatg
	10261	tgggatgaca	tttactctgc	attcatccag	tgctgttatt	tgttttgtat	ttggcatata
25	10321	ttacttaact	cttaaagtaa	ctctcagaga	tagattaaga	aaactagagc	tcagagaatt
	10381	taagtaactt	gcccuaacta	acacagaaaa	tctgaagtgg	agaagctaaa	cttcaaaccc
	10441	aaagttttct	ggttccaaag	tccattatga	agttgtgcct	ccccatctta	tagctaccac
	10501	ccagatttaa	tctgggtctc	ccattatcag	atggtttaca	tacacatttt	cttacaagat
	10561	cttgaccaca	actctttgag	atggccatga	gtctcacaat	tcatttccag	gagtgtctact
30	10621	ttagaatcat	ttgatctttt	gctaaccgat	gagagatttt	caaataagcta	attgtcacct
	10681	accctttttg	aagcccagtt	ttcataatca	taaaatggaa	acagtattac	aatgttttgt
	10741	taggatcata	tacattaata	ataaaatcta	actttgttga	gctcactatg	gtgagcattc
	10801	tgcatttcct	tagttcattg	aatcctcaca	acaatctttc	taggctaaga	cgattatttt
	10861	tcttttaaa	ataaggaac	tgaggcatca	gtaattaatt	aactatctta	aattagcaga
35	10921	gccaataagt	ggcaaagctg	ggttcaaacc	taggtctgtc	taatgtcaaa	gccctttttt
	10981	taatcactaa	tctgcaaatc	actattcaat	cttagctttt	attattataa	ttatcatcac
	11041	acttaaaaca	ctatcaagat	acagaatgat	ccagacataa	gtatatagtc	actgaagaga
	11101	ttagaatctg	aaacttttca	cctgcatgtt	cttccttcca	ctttagttta	ttaaccaaat
	11161	ggatgatgtc	tgactccttt	cttaacttgt	ttagggcagt	tccaagttag	tgcactctg
40	11221	agagttattg	agtaagaat	gttataaatt	gtttggatta	ggatttagta	tgtttagaag
	11281	ctatttcata	agtttgcctt	tgcgaaactgt	tactggctat	aatactgcag	atgctgtgat
	11341	gaggaacacc	ctctccaaag	acacacagtg	gatgacaaac	ctccaaagct	aacatgttgt
	11401	ttacagatat	ggagaagaag	aggatggaca	agcacagtct	aaaacgtaat	tacaaggctt
	11461	atagtccctg	ttggggacta	gaatgtttat	tggctttcct	tgtgcaattc	aatgctcttc
45	11521	ctccaaagga	tccactccaa	acttggaaact	ttcctgaaaa	tagcatctca	tttgggagca
	11581	tgccaggaat	tgggtgtctg	gtcctttgtg	tctttgcacc	aactcagaac	tctggatact
	11641	agctctagaa	actaagctgg	gatataattct	gggtaaggga	gtagcatatc	tacttgggca
	11701	tcttcctgat	acattttatt	catccatctt	cctcctagag	agcacctcct	agaaagatgt
	11761	ggttttaaat	gagggattgg	atgcatactg	gtatgtctta	gcacacaagt	cagtggctct
50	11821	tgacagactg	ccaaaggcat	ataagtaatc	aaagatgagg	aagtctatga	agagacttca
	11881	tcccacctcc	actctgattt	attcagggaa	ggaccccatg	aacacataat	ggatttgata
	11941	cgtcccagag	ctctgaaagc	agcctagcaa	aaaaggataa	tcttgaagga	cattttgatg
	12001	tatgaaaaag	tccacctaata	gctttgtcag	agataactaa	gtaatatgat	ggctggtaga
	12061	ctgtaagtcc	ttaccttggc	tcaggaactg	tatatcattt	ggtaaactaa	acttgtcgtt
55	12121	caaattttaga	tagaaaaagt	accttacaaa	tgatctagtt	cactgattcc	cttcatgcat
	12181	tgaaatcacc	taaatcatct	cttctttctg	agataaggct	tgaatgtgtt	gccagcttta
	12241	gcaaactcag	tttgtagccc	actgacctca	tttgattgat	tgggcaactg	aggtgcacag
	12301	tggtagatct	ctcaatttat	tcaataaaca	attatatggc	ccttacgata	tctatctgaa

	12361	caatcttggg	ctagtgaagt	tgcttgccca	ggttacatgg	ccagaaaactg	acagtttttaa
	12421	attaggacca	aagttctttt	gactactatc	tgggccttaa	aataatatca	tatgacaaaag
	12481	atattttcttc	tgtttcctaa	tagtcacatc	aaaaggaaac	aatggacagt	ttgtgcaaga
5	12541	ttttagttac	tttaatgttc	aaaataaaat	taaaaacaga	ttattactaa	aacataagca
	12601	taacaacact	ttaatagcat	tctaatacaga	tattattaat	ttcaaaatgg	taggacaaaa
	12661	ctaattatac	tttatacttc	ttaaataatcc	tatagttact	ttatgactat	tgagacacta
	12721	gctaaaaactt	gaaacttcaa	gttttcattg	attcctatat	tattacttat	ttcagagtta
	12781	cttcattttg	ttctttttatc	tgagattgga	caacagcttt	atttgatttt	cagcgacaaa
10	12841	attctttttca	ctcctgatcc	tccaccccaa	gaaaacaaca	gctactaata	tattttccct
	12901	aaagtgatca	agaaataaaa	gaggaattct	agccaggcgc	ggtggctcat	gcctgtaatc
	12961	ccagctcttt	gggaggctga	ggcggttgga	tcacctgagg	tcaggagttc	gagaccagcc
	13021	tggccaacat	ggtgaaactc	cgtctctact	aaaacttaaa	aaatgagcca	agtgtggtgg
	13081	cgcagcctg	taatcccagc	tacttgggag	gctgaggcag	gagaattgct	tgaaccaggg
15	13141	aggcagaggt	tgcaagtgagc	caagattgcg	ccattgcact	ccagctctggg	tgacagagtg
	13201	agactctgta	tcaaaaaaaaa	aaaaaaaaaaa	aaagaggaat	tctaaaaatta	attatatcta
	13261	ttaatatccc	tactcttaaa	acgttagaaa	atgtttgctc	atttaaaaatt	tttattttta
	13321	aaaccacctt	atattccaac	taaatactct	ttggagcaat	ttctttgttc	ctcatataat
	13381	atccatacat	ataattctgc	ttttgtgatt	aacttttatt	actactcttc	taaaaattgtg
20	13441	ctcttataaa	catcagttaa	ttaagagtta	atctgatgtt	ttataaaattc	tttctagaaa
	13501	cagagagcaa	aatcatataa	ataacaatat	gaatttccaa	aagtacaata	ataaaaaaaaa
	13561	attagaaaaa	aattaatcta	ggaaaatagtc	aagaatatgt	caaacttgta	catacttttg
	13621	agataaattg	gcatcatgta	gattagcatg	attcttcttt	atggaattca	acttattttt
	13681	actcactttg	ctctaattag	tttttgtgtg	cggacaagat	ggaaggtaat	ggaaatttgg
25	13741	cttgcaaagt	agttctaaca	tgatctacat	ccacaatctg	gttataatgc	tataagaata
	13801	ttatgtggga	atagtagttc	aaatcagtat	ttagtatgaa	cataaaaggga	caaacaatgc
	13861	aaagctaact	taagttgttt	acacttggaa	cttatttaaa	ttaaaaaggc	cagtggatgg
	13921	tcatatgttt	ggctcattct	tctcaaggcc	ttcaggaaaa	catgcctatg	aaataaaaaga
	13981	tcctcaatat	taaacatttt	actgcatttg	ggggacacat	gaaatctggt	aataaaggaa
30	14041	gtgttggtct	tcatttttct	aattcagcat	ggaaaactatc	ttgaggaaaa	ctgactatgg
	14101	tcttagtttg	tgcttcagaa	atatattttag	tctgaatcat	ggcgtcgaca	tctgacttcc
	14161	aaaattggat	atctagccgt	atagtacctc	acctcccaca	cacaccaccc	cccattccca
	14221	ggtcatgact	actgtccaag	cagcaaaaaa	agaagtaatt	tcccagagta	catacatggc
	14281	agtgacaacc	aaccaaacaa	aaaacaatta	taggggctgg	aatttaaaat	aatggctgta
35	14341	ctctcaccaa	ttcattcccc	attccacccc	atctctctgt	cttcaacttt	tatgaaacat
	14401	tatatattgtc	ctattcttct	gtatcagcat	cagcctttcc	tatatccaac	tagacttata
	14461	acttcttgg	gcctctcact	ggctgactaa	ggtttcagaa	gtacctactt	acagcaaaca
	14521	cttgacagc	tctctttttg	gttacaaaag	ccctggacaa	tttctcaagg	cgatattatg
	14581	aagagggaag	aaacattctc	ctctgctacc	ccatttcttt	ttagagtgtc	aactttattc
40	14641	tatatctggt	ttaatgtctt	cttaggcca	ttggactgat	tttagagaca	ccatagaata
	14701	tctcctgagt	aatgggaaca	atatttctgc	tgatcccatg	atttggcttc	attgggttgt
	14761	taggccataa	tggagacata	cttgatgaat	ttatgaagac	ttgattctag	gtatcatgta
	14821	ggttagcata	attctctttt	actgaattca	acttagtttt	attcacttta	ttctaactgg
	14881	attttgtgtg	cagaccaa	gaaaagaaat	ggttcaattt	aggtgaaagg	taaagcttca
45	14941	aaagtagtgt	agtatttcat	agaccttacc	tttgagagaa	atttatatcag	tatataataa
	15001	gcacctgaga	atatgaaagc	acaaatccaa	tttaaagtgt	aaaggtctac	aacttgggat
	15061	tttaaagtga	gtacagaaaa	gccactgttt	cttaaacaa	tttgttgagg	gggaaaacag
	15121	tgaagctaa	atgttctatt	caagagttgt	ttcttttgaa	aataatgctt	catttaaaag
	15181	ctaaggacag	aagacgtagc	tttgttatga	aggctcatct	ttttattaaa	caaccactac
50	15241	tttgtctcca	agttgcaaag	ggaagatttg	tcaatctgat	tgaatcttcc	ctttagtttt
	15301	tcccaacagc	tgtgtccaga	taattcatga	ctcctgtgtt	tcctgagccc	tggataattt
	15361	cacacacatg	tctggttttg	ggctccacat	tttcagaaaa	atatagaaat	cttggaggag
	15421	gtccagggta	gaccaaggga	aatgattaat	gggttgaaag	ttggagtta	tgaagaaagg
	15481	ttgtgagatc	tgatcttttg	ctacgagaaa	agtctgagtg	gtgacttaat	aacataagga
55	15541	ggttagtaag	cagctgttct	ccatcttcac	taagggttgaa	tgaaatgaaa	taagatataa
	15601	attgcaacag	gaacaaaaat	gcattacaag	tgaggacttc	caagcaccag	cattgctgga
	15661	ttctagatag	ctcccaaaaa	gaaggatgtg	tagtctactt	ccctggtctg	caatgacagg
	15721	cctataaata	gtgagaaaaga	tgagataatc	tcttaagatc	ccttctggac	ctatctttta

	15781	taggtctatc	tatcatat	agaaaaatta	tttgctcaa	acaaaaatta	tctgatttcc
	15841	tccctctcac	cctatccact	ccttctcttt	tgtctacctt	ttgtaaaaca	ctgctaaccg
	15901	aaataactgg	ggactgatta	accgtgggtg	gccctcccc	gcctctaagt	gccactccag
5	15961	ctttgggagc	aagtttcttg	tccatcacta	ccacccctg	gccactaggg	gcatgtttac
	16021	catcatcttt	ctacacacca	aacctacggc	aaggggaaaat	aaaacaaaac	aaaacttcct
	16081	agacttaaca	aatttgcaag	tgtcaccatg	gattaaaaata	caactcttat	gtcctagaat
	16141	atgagcatgt	aaagggctaa	aatgtat	atgcatctgc	ctgtatcagc	ccatagaata
	16201	gcctcctgac	agatagtaga	tactcagcaa	tctttcatca	actgaatgac	tgtaactatg
10	16261	aagtgaagag	caactaaagt	tgagaaagtc	aggagtttcg	gatgtttcca	aatgattctg
	16321	tatgccagag	taatctaaag	cctaaccocat	tcttcacaac	catgcactat	taaggatttc
	16381	attctcacca	tgctgtgct	atctggaggt	agaaagaggg	ccagtgtcac	atcctgctca
	16441	agtccttggt	caaaaagacc	actaaagagt	gctttgtaga	ttcatgtatc	agaatcacat
	16501	gaaagtaggc	caaattctta	gtgtgtgttt	tfaaaataag	actttaggaa	gttcacttat
15	16561	ttttttctaa	attatttttg	catattcttc	tttttcattt	ttttcatgaa	gaatttaaaa
	16621	tttggtctgta	gaaaatctct	cactccaaac	atcacacagc	ctaaataggt	gagtctcaaa
	16681	aataagctaa	tgttcatctt	tcatctgatt	caatgtcctg	aaaccctttg	gtttaaat
	16741	gttaattctt	ctcatggctt	ttctcctagc	aaaaccaact	aataccacag	ctatttatta
	16801	ctgtcagctc	taacttatgc	ccacaatctc	acatcccttt	tgaccacgct	atagaacta
20	16861	ttacaacaag	taaaccaaat	ttattcttca	ttatttaattt	ttaaatgttc	tcagcacaaa
	16921	tctggtaact	tggagggcta	caagttgata	tttctcatat	gtttgggggt	ttagtctcaa
	16981	cagtttctta	atggttttcta	tgccgttttt	cttgatccaa	ctaaatatta	ttcccagatg
	17041	ggatcagctt	ttgaccctct	tggttctactc	tcctagtctt	ggcccttcta	aaagtcttct
	17101	gctgtggttc	ctttcttttg	tctgccacta	atggctatgc	ctggttacat	aactcctgta
25	17161	acaggtgttg	accaatttga	acacattttg	gtatggtatt	gagctattct	tatggttcat
	17221	aaaaagctta	gtgagaacgt	aacatctcat	gaatagggaa	attacttctc	ccttaagggt
	17281	tttctcagga	caggcctcat	acaagaat	caaggattgc	gagtgcata	gtttaacatt
	17341	ggaccaggcc	tttcaaatta	tccaggatga	gtttgaaaac	acctgtgcca	ctctgtctaa
	17401	cagcagagtt	ttctgtttac	taagtatttt	ccctatgcta	attacggaaa	gtttcaacag
30	17461	tttttttagg	ccaacttatt	tgtgcttaga	ctagacaact	tatttttttt	ttctgtcaag
	17521	gaatactgaa	ggtaggagta	actaggaagc	ttaaataaac	ataaatataa	aatgcttata
	17581	gtgatagaat	tgacctcagc	caattaaaaat	tattaataga	aaaaacatgt	caatgtcaag
	17641	cctactacct	ctgttctcac	ttgagtaatg	aggattagtt	tatat	gacaagaata
	17701	gatgggaatt	caaatttctt	cctgaccttt	gttccccctg	gaacattggg	ttaggatcat
35	17761	attagaacat	aaccaaaaag	aaataaagat	tcagacgaat	tcacaattaa	tttttaagcc
	17821	ccacaaaagt	gaaataggta	gcattat	ttcaagctgt	gaaactttcc	ctcattttag
	17881	taatagagaa	aatgttcaga	ttataaaactt	ggaaactttg	ctcctaacat	atcaattatg
	17941	ccagaggcca	atttttaaga	agaagagaaa	tgcatgctct	atattctcag	catcatcctt
	18001	gcccacata	gggaaataat	tttgtaaaat	gtttgatttt	agacctcaa	aattatctct
40	18061	atatgctacc	tgaattaaagc	aaataaaaaa	taatat	aattccatgc	aaggcactgg
	18121	tacaattttg	tttatcttgg	cttcattgtt	tttgaatgta	agatgtactt	ttaaggcaaa
	18181	taagtacatg	ttttaagctg	gtcgcataca	gtattggcaa	tgctataatc	acaaatcaga
	18241	aagt	atgcttaca	gtgtaagag	gtgtgattca	tcatggttat	ctgaattggc
	18301	atctgatctt	cttttctttc	taaatatccc	tgacatttct	gactcctctg	tcttttcctc
45	18361	agtaaaactg	caccacacac	tggaaagcga	agatacacac	atttatttat	ataatgtcaa
	18421	gggagagtag	gaataagaag	attggccata	gacccacca	atcagagtct	gggaaatgag
	18481	aacacttttt	ccttcagcag	aatgctgac	gtgccaatgt	gaatttagca	gaaaaaagat
	18541	ttgccataac	ttctaagtga	gcagccttca	gaatgctagc	ttagattcct	ggcattaact
	18601	tgccaggtat	tttttcagga	aggaaataaa	ttacaattga	gcttaaaaac	ctgagggtag
50	18661	aactcatttt	caagcaaatg	tgaagcatca	gtttgaagtt	aacaaagtta	aagt
	18721	tagggttccct	ccagtccttt	ataatgtagt	acaagtattt	tttttaaatg	tataacacta
	18781	gccttttaaa	ttgtattgtg	ctactaaaag	aaattgtgcc	tgcatcctc	ttacaacctg
	18841	ggaaccaacg	cagaggtctt	gtgggtagc	ggtatccagc	ttcatgccct	ctgtccttta
	18901	ttgctttctg	gttagcctgc	gtatttcaca	tacattaaat	attccacaat	aaactctgcc
55	18961	atctgtgctg	tagggtagtt	tgtattggtc	atgtgctctg	tcaagttgac	agaggtgcaa
	19021	agctaaatgt	gtgacactcg	aagaatatgc	atatatttga	ataatttgac	tatttagtcc
	19081	aacaatttgc	aaaggcgctc	tgaatgatca	cacattctga	taacacttcc	aaggaacaga
	19141	tagcttcact	taggggtg	gggagatgga	agcagggtta	tttctagcag	gaattcttga

19201 gttcactgaa gtcttgtccc tggacttca ctgtgtgaac gtgggtaaat tatttcctgg
 19261 cagaggatcg gattcttctt ttataaaacg ggtaaataat ttctgtcact agtctttaga
 19321 agttctaaaa tagctaagt tagtgaattc attttgctaa ctgtaaacc ttaggtaaat
 19381 tgaactgagt atgtaataat attatatatt cagttcaaca gcacattctt ggtaaccaca
 5 19441 agaggggtcca ggaaaggaaa ctgtttataa atctttccct ttagcaaaat taatgttgga
 19501 gtcttttaggg aaattcttac agcaatagtc ttcgcaatta ttaggtcaaa cccctttgag
 19561 attacagaaa aacgcacaca cacagaaagc tgcctgcaga atttgggtgt gggcttggtg
 19621 ggagattcct ctgataccca gtgtgtgacc cccaagagag tgtttctcaa agtgtgactt
 19681 cagattgtct gcattcgaat tgcttgtggt atttattaaa attataactc ctgggccctg
 10 19741 cccaccacct actaaatcac aatttcagga ggagggacct tcattttaac actcaccag
 19801 gtgattttta tgctccgagg aggtccaggg actccaagtt aagtacggta ctgctgtctt
 19861 attctttatt cttaaatttta aggtctgcac aaattggtt aactaatgag aagaaaattc
 19921 agctttaaag cagaaacaca ggtagacggt tgacagagtt catcaaatgg ataattgaaa
 19981 atgtcctctg gaccctagcc atataagttc tcttcaaggg tcttggtac aggcaaatga
 15 20041 gaaccgaaa ggctatttgc tcttttgctg cgggcagtgg tgggggtgga gggcggggga
 20101 ggattaactg agccagttct gccccaccc tcgaatcacc taccocccact ctggttaaag
 20161 cagaagactt tttatttatc ttggctgccc tggttcgta ttaaaagggt tagcttatac
 20221 gtgtgtttgc tggggctgga agtgaaaaca tctgcaaaag catgcaatgc cctggaacgg
 20281 aactcttcta ataaaagatg tatcatttta aatgcgctga attttgattc tggtaattcg
 20 20341 tgcactagag tgtctatttc gaggcagcgg aggtatcata tgacagcgca cgtcaaggca
 20401 ccgtggagcc ctctcgtgga ctcccaccca ctttccatt caccgcggag agggctgctc
 20461 tcgctgccgc tccccccggc gaactagcat gaaatctccc tgctctgcc gagatcaaat
 20521 ggagcttctc gctgatggg tgcgagtatt acctccgcca tgcaatttcc actatcaata
 20581 atttaacttc tttgctgcag aacagaagga gtacataccg ggcaccaag actcgcccc
 25 20641 cctcccccct ttaattaagc gaagggaacg tgaaaaaata atagagtgtg ggagtttgg
 20701 ggccgaagtc tttcccgag cagctgcctt gatggttact ttgacaagta gtgactgaaa
 20761 aggtggggtt gttttctttc tttctctttc cgtttttctg tttggtcggc tagaaagcgt
 20821 gtggctttag cgaggtctgt cattgcctgg gcttctctgg tgaacaagt aacttggtgt
 20881 aacgttatct gggggcggtc atcaataaaa aatgtctgta ttatctgtat ttaattccta
 30 20941 ttaggcaaac tctagagag tcagtcgcg aactctgtt aagccggcgt gtttaaggca
 21001 gcagagtaaa ccaatagccc ccatgctctg tgcgatttca ttgtgtgctc gcgttcgcaa
 21061 gctccgtagt gcaggaaggt gcgggaaggt gtgtctgtgg cccgggaac gcacgccctc
 21121 tcccagagaa cttgggtgct gggatgggga ggaaggggag agttgaaagc taggggagcg
 21181 agacctcggg gcgtgcgatt ctactcgct cctcccgcc ccagcgcca cagccgggt
 35 21241 ttctgcagag ggcgcgggac gcggggttcc ccggggctga ggctggggct ggaacacccc
 21301 tcgaagccgc gggcgctcctg tccaaggcgc ccaggaggg cgcaggactc gcagggcgat
 21361 gtcgcggggc cctaggggag gaggtgagga caggccccgg gggagcggg agttccgggc
 21421 gccctcgggt tccccgcgcg aggaaaagac gcggcgttcc cttaagcgg ccgctcgaa
 40 21481 cgggtatcgg tagcgcgggc gagcgggag cggggggcgg gggcgggcg gggcgggcg
 21541 gcgcggtttg accaatcgaa gctcaaccga agagctaaat aatgtctgac ccggcgcaa
 21601 ggcgcagcct ggagctccgg gtccccgacg ctgccgccgc cgcgccggg cgcacccgcc
 21661 cgctcgctgt cccgcgcacc ccgtagcgcc tcgggctccc gggccggaca gaggagccag
 21721 cccggtgcgc ccctccacct cctgctcggg gggctttaat gagacacca ccgctgctgt
 45 21781 ggggcccggc gggagcagca ccgcgacggg gaccggggct gggcgctgga gccagaatcg
 21841 gaaccacgat gtgactccgc cgccggggac ccgtgaggtt tgtgtggacc ccgaggtagg
 21901 caagcgtggt gaatggggct tgggtcagga gctgccgctc cgcgggagag agttgactgg
 21961 gggatcccc accccaaagt tgtgggacga ggccagtctc cttctttcct cccctccggt
 22021 agaagggacg atttgaggt actcttggg agttttctcc cccatcccac aaccagaag
 50 22081 gtcagccggc accaccagg aaaaaggac ccggggaagt cacgaagtag aggagggaag
 22141 gcctggagga gaccagagc tgcgtgatgg gagcaaagac ggcgaccgg ggatccctcg
 22201 cagccctccc ccagcccagg agtagtcgag agagacttag ggggcccagag ctgtcgaggg
 22261 tcctgactga ggggaggggt ctggggctag gctaggaatc cttccagggg gtgggtggtc
 22321 cccgcgccga cttgcgggg gagtgggag gaagcttgcg ccttcagccc gcatcccttc
 22381 cccggagctg cacacggcta cctgctcccc aggaattgag actgaagtgg acttacaagt
 55 22441 ccgaagccaa ttagcttgg aaaacttgg aggcggaatt cctaccgctg ggaactgaaa
 22501 gggctctgca cactctcgg caggccgaac ccacatctct acccatcctg cgccctctt
 22561 ctgaagcgcc ctccaggga gttaagagtt ttgactttcg gggagtgggt gggatgtacg

22621 tgggggattc ttgactcggg ttagtctctg gggatgcaga gccgggaaga ggaatgggtg
 22681 agtgagttac tcctggaaag aaatagctga ggattggggg ctctgtgcct gacgggcaag
 22741 aagaagggga gattacagac taggggcatc cctaagggaag aagcctcggg gctgcgaggg
 22801 tgaactggag gatgcagtgt ttgtgtgttg ggggtagagc ggggatgagg gaccgggggtg
 5 22861 gaggggagggc gaggaggagg aggggaccca gagaacgaag ctagggaagg tagaggggtgc
 22921 cctctgccgg ccatgctgcc aagagcagct actggggggcg ggaggctggg ggtggggaaag
 22981 tggtaaaggga aggttttgcg ggatccctta gagagctggt aggagggact tgttgaatgg
 23041 tgctgctgac tccagctcgg tggggcggtc gactcgtcgt cgggtggattt tgactcctcg
 10 23101 ttcttgtttg gcttctatgc aagttttcct cgcgctgggg gagctttgat aagcctcgat
 23161 tggcgggtgtg tttagggcttc ttggatctta ttttaggggtc ctctagtatt cctgcactta
 23221 ctcccttaatg tcagtagcaa ccaaaagaaca ttttccgaca agcacgcagg aatgttcttg
 23281 gccagaagca aagaaaggca tatttctgag tgtttattaa tcctcctagt aatcttttaa
 23341 agcaaagtaa tatgtaattg ggaacgttga ttttctaact gcatataaaa ggcgacatga
 23401 tattaatatga gacccctccc tactgactca atatcctgca aaatctctct ctccccttta
 15 23461 ttattatgga aaaatctatt tttatatgag tttgttgtaa ggtcaaaagc cattttggtc
 23521 ttacaatttg atatgtcttt acatttttaac ttattgaggc ataattacag atttaatttg
 23581 tatgaacgtg tgtgccttca atgtctatct catgcaacat aatttttagg ttggagattt
 23641 ctgatgttat ggcattgtagc gtttcaaggc attacacata ataggttaaca tagcatgtg
 23701 aaattacacc acaaaagtgtt gaccctggga acagcacctt ttaaaaacaa tcactaaact
 20 23761 cctgttcctg ttttctgatt ttgcaaatgc cttgcttaag actttttttt ttttttttt
 23821 tttttttttt ttgggaaatt tacctctggg ttagcaggag aggtaaaaaa aaggaaagag
 23881 acacttggtg aaatgtaacc ataaccttta ctggaattta aaacatgttg gtcaccatta
 23941 ctggaattcc agggccataa agtcgttggtc ttttttttct tctacttcat tttgtaaaat
 24001 gtgataaatg ttggtaaata tagaccagta gtaagtatta tgacactaaa agcattatgt
 25 24061 atgtggaact attttaagtt attacagaac attttctatt tataaatgat ataagcagaa
 24121 agaaatgatt tccagataaa caaggcttac gtacatgttt tgaagcatta gaacattgca
 24181 gacactctta gacatcacat tttttaaaagc aaaataacag taatttttca catacctttg
 24241 gagcctttca tagcccatc agagctgagt tagtagctgg aagtttcctt tattttaagg
 24301 tgatatttta aaaccattta acatgtatag taggtcaaca ttggtcattc cagaaaatga
 30 24361 agcatttagg aaatctgttt cagtgtcttt tcaatgtgtg taacttttac ttgcaaacca
 24421 atggaaccaa gaaagtcatc atttgcctaa aatgcagtca tcacctcaaa tgattcattt
 24481 atactatgtg agttaattgc cttcatctca ttaatggcca aggagggaag ggaggtcctg
 24541 gggatatttct tgttcatttt gactcaccag gagggaaaat cctgtaaaaa aagaaatgca
 24601 aattttctaaa atcctggctc aaagtccgtg ggtttcctgt ttaaaagggg cgccatgaaa
 35 24661 atgtaagcta ttcccttttt cctggaatct ttaagagtcc cagcttttca atagtcaaaa
 24721 tgtagatgat tgatatcatt tcttatatga atagcactgg tttgtagttc agcacgcaca
 24781 gtgagctggg cacgcccacc tgatagtata gcagagaact tgtttacatt ctttttacat
 24841 tcatcttcta aaacctgggg tgctctctct ctctctctct ctctctctct ctctgtgtgt
 40 24901 gtgtgtgtgt gtgtgcgtgc acgtgcgctg gtgtgtagag ggggagagag agagagagag
 24961 aactgtgaac tgtgaaatat aacacagcca gcagctttgg gtctcaatcg tagacttact
 25021 cttaaggaaa ttacagaat ggaaaggtca tgttcaagta gtttattaac attttgagat
 25081 gtaggaaatt aatcccggag tacagaagaa caatttcaga cttcctgaat aaaaacagac
 25141 agcatagaga gtggatgata gctaaactct gaatatcttt tgagaagaaa ggcactccca
 25201 tttcagggtgc ccataatatg gatttgattt tagtgattaa aacattaatt ttcaacttgc
 45 25261 atctccctgt gtggaagagt tcaatttgtg tgaggggtct cgcctatcca acaaaagtga
 25321 atatgtccct tttatagggg aattgctaac ttgtctcaac ttgttttcaa acaattgtta
 25381 tagagcactc agtttccact aattgcaaaa ttgttgctta attgaaggac tctcagccat
 25441 ctagtgcagc cattcagcca ctggcaggct ctgtgatctc aaactgtgaa ttgcatttta
 25501 aagaggaatc gaggagagaa ttctgtggaa ttctaggttt taagtgtctg ctgttgttca
 50 25561 atggaagagg aaatcatttg aacaagaatc gcatcaagtt gtgttgtgat aaattttctt
 25621 tattaggatg aataacatgc acagatgagc ttcaaaagtg aatgagcaaa cttactgggt
 25681 aactctgca tccatttact ctgttttagta tggagtaatg ttaggcaata aatgatgctg
 25741 gcaaatgaaa tccgtatgtt atttgcagtgt ggtattttaa cctaggaaac atagagtggc
 25801 tttgggtattt gtaggcttag tcatgtgtgt cctaaacgtc ctcttaaaact tctacttaag
 55 25861 gcatagaatt atttaatcct aaataatttt atacttaagt gcctcactgg atttccagaa
 25921 tatttacact gtaaagattt agaaaggtca tgaacccaat tattgactat atggaatcat
 25981 tattgatggc agatgcaaaa tggagctcac taatgtactg acattgaaaa ctttttgcat

5 26041 gggagaggag ggggagtggt aaatgtgtgt gttctttaag tggaacagga aggtattctc
 26101 ttttctgtag aaaaatttga gtatctgggtc agataagtgt ggaagctttc atttaaatta
 26161 agtattttaag ttcaagtaga agctctagggt cacttatcct cttgatgaga caaatcttat
 26221 caaatatact agatgctaag aagtggctca ttgccctgat gtctcattta tagattgatg
 26281 tttgaggatg ggttgcattha agtgagttag ggggctgagt gtgggacagg agaacgattg
 26341 gaaggaagca aagtaaaattt acaagcttta gtgacagcca taataaagta aaagtttatt
 26401 tccagagagc ctagagagta aggaacgtta tatagttttc cccaaaggtt cacttgaaag
 26461 aacttttcat tggttgtcat ggtagtaatg tcttgatttt gaaatctccc agaacctagt
 10 26521 agctcttaaa catgctttca tcttgggtcc tttgggtctga cggaaacttt atgacgaccc
 26581 tctgtgtttt tgacatgcct ctgcattttt ggagagagga ggtcaggcaa gggaggattt
 26641 cttaaaacta agacagtata gtaaggaaac ataaaaattat atgataaaaa atcactgaac
 26701 ttcaaattga cttactgaaa taaaacctag aaggcaacct gtcgtttaat tacaactagc
 26761 ttgtataaaa ttaaaattta taaaatggga attcaaagaa aataaacggg cagtccaag
 26821 taattttaagc aactcaccaa aaattgaagt aatagtgcc cctagagaac aaaatcacca
 15 26881 gctttactag ccaaatggct tatttccata tgaaccattt ttccaacgct acagttacta
 26941 ggatttcctt gttaccatat tcagatcttg tgagtgtgta tgggggtggg ggttgcatgt
 27001 ggaattacag atgaaaattt aaaacaagca gatccacaat ttgatataat cactaaatcc
 27061 ttttaacggt gtaatgtagc caaatgtaga atagcatgcc aggaatcaac ggctagcatc
 27121 ctttttaaca ttattttttt tcatggatat gtaccaaacc gaaccttga gtataaagg
 20 27181 tctgatttta tttatttgct acaggcaatt cattatactt tctgagatac aataacacca
 27241 aataatttga gtagagagac ctttaagaat gttttcgtat tatgatctac ctttaacttt
 27301 aatgtactca gaagatgtga gaataaaata aagtcaaata taagcaagat tttaaacaca
 27361 cacacaaaaa acaaacaaac aagaaaaagg aagaaaatta taaggattgc cttaacctta
 27421 gaatagatga aggtatacat ctgagccagc accaaaaaaa aaaaaaaaaa aagttatgga
 25 27481 accaggaacc aataattaca aattgactta aaattcttgg atgacaaaaa tctatattta
 27541 gttcattttt gcatgcgccc acaacagcat ccaaaacagt tctggggagg cactttgata
 27601 aatggttgctg aatgcactaa tagattgatt aatggctgct tcagattatc actagtgtatg
 27661 tagacagaaa cttcatgaaa atggtttgtc ttgtctggaag aaaggcagaa attggaggaa
 27721 aaggtttaat aatatttttc cccagtacct attataaaaag tcatttagtt ggcttagttc
 30 27781 tataatttct tatgtgtaat ttgattcact tatgaaattg tgaatatatg aaatgttaaa
 27841 gttgatttag acagcaacta taagcttggt gatttttctt taaatgtctt caaattttta
 27901 aatgccagtg gagatgccag cgactgtgct tcaggaggta gaatatagta tatcttaaat
 27961 ttgtgccaat ttctggtgaag cagagaaaaa attgcatgat aaccaaagaa agtcatattg
 35 28021 tttgtgcttt gtgttattca tggaagcaat caggtgcaga aaactttctt tttcagaaaa
 28081 aaaaaattac taaaataaag gtgctgtgtg gtgtatgcac atatatctaa agggagagag
 28141 ggagaaggaa acttactaaa taaaattttt gccacatggg atttagtcta atcagctctg
 28201 gtttttgagt tgctatcatc agtagttcca ttttgtgatt ctttctttct gccttcatgt
 28261 gcctttgaaa actgaaacta tgccaaattt aaaaacagtt tttctgtctt ttcattagtt
 40 28321 cacttatttc ttgaatgtgt ttttaaacac agacaaactt cttttacatc atgtagaatc
 28381 tgaagttcga gaaatttgca gtcattttgc tggagagaga tgcttggcgg agtcccaggc
 28441 cacattccta ggccaaactc tcgaagggtat tcctcttatg caacattggg aaaatacatc
 28501 cagcaccgac atgttggctg ataattgtgtc tgaaggcaca gacgatatgc ttatcatatg
 28561 aaacataaag ccagcagata ttgcagacat tctgttgaat gatagaatct ggatcattta
 45 28621 catttactta aatgtaaaat actatgatta agtacaaaaa aatcaattta ggagagaata
 28681 gagagttgct ggcacggttt taggggatga cttatcagca gattgtagaa aggaagcttg
 28741 aatgttttaa attaaactga agttcagtat aagccagtgg tgtgacaaga ggctgttatc
 28801 atagctactg aaattttggg ctgactgtct agaaatataa tactgaaatg gagaagctaa
 28861 taattcttca ctttttaaat agactgtatc tagaatatta tcatcagttc aaggaaatga
 50 28921 aataagttgt tttaggtaca tcatcgataa attagtgtac attcaaatca ctgtgaccag
 28981 gatgcatagg gaatttgaaa gcattgcatg tgagcaatgg ttgaggggac ttggaatgca
 29041 tgacttaggg acaagaaaac ttaggctgga atggcaagtg gtttttgaat gttgggttga
 29101 gaagaattct aaaactgtga aggattagta aaaataacat tcagattgct aatgcctact
 29161 gtggctggga gattagagt tcaacatgtg tgatgtatth ttgacatcct tattttgagg
 29221 atgggcttca aagatttgac gaactgtcat aagtgttaatt tgtgttgctt cagacagcag
 55 29281 ttctagaacc aatgatgtaa atttagatac tctacatggt agttagaaaa ctttccatta
 29341 atttaattta gcaaatattg aatgctcact gcatacagag cactttatta gaggaatata
 29401 taataaagaa aaaagaggtc tgggtgtggtg gctcatgcct gtaatcccag cactctggga

	29461	ggctgaggtg	ggaggatcac	ttgagcccag	gagatcatta	cgagtctgga	caacatagca
	29521	agaccccatc	tttacaaaag	acaaaaaaa	ttatccaagc	ctggtgacag	gcacttgtag
	29581	tcccagctac	ttgggaggct	gaggcagtag	gatcgtttga	gccaggagg	ttggggctgc
5	29641	agtgagccgt	gattgtccca	ctgctcttca	gcctgggtga	cagagtgaga	ccccgttga
	29701	gggaagggaag	ggaagggaag	aagaaaggta	ggaaggctga	aatgaaaccc	ttttagaaat
	29761	gacactaaaa	tgggaggttg	gagtaaggta	ttttctgaag	tgcctttgta	cttgtttttt
	29821	tctaatagcat	tggccataag	tctgctcctt	atttatagtc	cataaacaat	cctaatagaga
	29881	acagttatat	atttctgcct	ttgaatcatc	tacttgaagt	gtttagcatc	atgaattgag
10	29941	tatcagaaat	ccctcccatt	tctttgcaaa	gcgctgtatt	ttacttttcc	ttatttgtat
	30001	acagattctc	aaaattggct	attttctcct	tgggttagac	agaacagaat	gtctggaaaa
	30061	aaaagttctt	atcaaaattca	ggtgcccaaa	ttgcttaaga	aattaacttt	tgaggttata
	30121	tttttttagg	gttcagtagc	taaactaaga	aaacttctca	ccgttcacct	tcacttttgg
	30181	aaaccacaaa	atcttcagat	attacagttt	tccaaagagt	ttctcttttt	aaatataaac
15	30241	taaaaggaat	tgactcctcc	cccaactccc	tcaggcctca	gcatggatag	agttactttt
	30301	tttctttaat	aatttattta	taacttattt	tgctcttctg	tagaacagct	ggagattaag
	30361	caacatggcc	atgacataaa	atgcaagtta	gaccataaga	tgagcagccc	actccaagta
	30421	tgaatgagta	cttattcttt	gtgatctctc	atactgcttt	taggtattaa	tagtgtcagt
	30481	cagcaaaagca	aacagtttaa	tatttacatc	tcctttagga	tatcatatag	ttatagtttt
20	30541	gtatgtgttc	ttgcgtgtat	gttttcttct	gtttcaaatt	cttttttctt	aaagtaagaa
	30601	tggtatatgt	agcaaatggg	tctttcatta	attcatttgt	taattcaccg	tgcattaatt
	30661	gagtgtccag	tgagtgtcag	gcactgggct	acttgatttt	cttttccctg	tattgatcca
	30721	tatatcttca	ggtgctcctt	gatattggctc	tccattgact	ctctcatata	aggctttcat
	30781	tacctattca	catactccct	cccaaagagg	actggtccaa	aagtaaaatc	ttgagcaagt
25	30841	tctctgagtt	atttcagaac	ttcttgcccc	caaactgtat	ttttaattat	tgactggtag
	30901	cattttggaa	taacttacct	ctctttttta	aagattgaag	ttctttatcc	tctctgattt
	30961	tcagggtgtca	gtttgtgtac	aaattgagac	aataaaaaatg	tttgtagac	attttcttaa
	31021	agcattttgc	tatgtgagaa	tctttcatga	agaactcttt	ttaacaatga	ctctatagca
	31081	gaagccacag	tagagggaga	actactgaat	caaagatggg	gtttgagtct	atgattttat
30	31141	gatggatttt	tttttttttt	acatacaagg	atatgcatgg	gtcttttagt	gtctaggaaat
	31201	ctgttcttca	cttgacagta	tttataaatt	gtgtgtttcc	ccctaaaaaa	acttaaaattg
	31261	tgagaatgct	tccattttact	agaagttggg	taatgattat	gccaaataag	gaaaataaga
	31321	cagaaaaatc	agtttagtga	atactatttt	gcctttaaat	ttagtaattt	agtaacagta
	31381	tctctttggg	gtttactaga	aaccactttt	taatccaata	ggctcttttc	atttgtgaagt
35	31441	caggaggtga	ttttgcttaa	atgtgtagta	taggaatcta	tatgtgggtg	tcaaggatca
	31501	tgtaaatatg	ctgatataat	cggagcacag	tttggcatca	ttaactcaga	aatattttaa
	31561	ctcttgctat	acaacatgga	agcaaaactg	tgcatagttt	gtgtgtgtgt	gtgtgtgtga
	31621	atctcaaaaa	aagaaaaaaa	tcacaggatc	aggaagtcgg	aataggtccc	acttttcttc
	31681	tagtaccaaa	cctacagcca	tgttcctagc	cttctctttt	actccaagc	aagacagaca
40	31741	ggcaaatgac	catcctgctg	cccatttctg	tgtatattca	cttgcatgtg	agattgtatt
	31801	cacctgcttg	ttgagagtat	tcacaaatgg	tacctgataa	agtagatact	tctttaaaca
	31861	tgtgaatttt	tttgcattgt	ataatgttta	gaaataatca	tgtataaatg	gttgaatatt
	31921	aatacaggat	tgccttatca	agtattttat	taatcattaa	aatgtgggtg	cattaatata
	31981	atttattttta	agtgtctttc	ctaaaatacc	agattatttt	tctgattttc	acatccctga
45	32041	caatgacttt	cttaaaactg	gtagccagga	acagaaaacc	taacactgca	tgttctcact
	32101	cataagtggg	agctgaacag	agagaacaca	tggactcagg	gaggggaacc	acacacactg
	32161	gggcctggag	cagggggcag	gggaaggagg	agagtgcgtc	aggacaaaca	gacaaatagc
	32221	taatgcatgc	gagccttaat	acctagggtga	tgggttgata	ggtgcagcaa	accaccatga
	32281	cacatattta	cctatgtaac	aaacctgcac	attctgcaca	tgtatcccgg	aacttaaggt
50	32341	aaaaataaaa	ataataataa	aataaaaataa	acttagtagc	atctattgtt	ccagagcctg
	32401	taattgctct	tcaggcagtc	tcacataaaa	acctaggaga	accttactctg	tactgtttcc
	32461	atgagggtgt	aggaaaactt	gctctactgc	agtgtccccc	taggcattgg	tactgagacc
	32521	aaaattcagc	tggtttgttg	ttactacgat	tcctacgtga	tttacttgt	catgtagaca
	32581	agattgcaca	cttcaataat	aatcttgtcc	aaatgtgtgg	tattccatac	atttttaaaa
55	32641	tgcattcaca	tatctcattc	catttgatcc	tacaaaataac	tctataaaaa	agattggcag
	32701	acattatttc	tatataacag	aggaggaaac	tggagcttag	agaagctaaa	tagcaatcca
	32761	aaatgcacag	ctgtagaacc	agagcaagga	tgatagccca	gtgacttcac	ctaacctagt
	32821	ccccttacca	ccactccagc	tgtctataac	caaaacctgc	agtattcaag	taagaaacca

5 32881 tatcttgccc ttgatgcatt aatgtgagac ctggagcagg aacaggctga tattgtcacc
 32941 ctggcctact gtccaccttt gtctccagca gagactggta cccttctgtg tgccaaggaa
 33001 taaagtggta atgggaagat taaaaatgtt ttttccaagg agttttttaa ttttaattttt
 33061 ttaaaaaaga aaaaactctt agagggaaaa atgaatatat gacttttgat gtattgttcc
 33121 ttagtaactt agttataatt ttacttaaac ctgagactct tgctaagtga atgattagaa
 33181 atattagggtg gctggccaga tggcaaatag gaacagctgc agtctgcagc tcccagagag
 33241 atcaatgcag aaggtgggtg atttctgcat ttccaactga ggtaccgggc tcatctcatt
 33301 gggactggtt agacagtggg tgcagcccat ggagggtagg ccaaagcagg gtggagcatt
 10 33361 gcctcactca ggaagcgcaa ggggtcaggg gaactccctc ccctagccaa gggaaagccct
 33421 gagggactgt gccatgaggg agattgctat ctggcccaga tactacacat ttctacagt
 33481 ctttgcagcc ggcagaccag gagattccct tgggtgccta caccaccagg gccctgggtt
 33541 tcaagtacaa aactgggtgg ccatttgggc agacaccag ttagctgcag gagttttttc
 33601 tcatacccca gtggcacctg aaattccagt gagacagaac cattcactcc cccggaaaagg
 33661 ggctgaaggc caggcagcca agtgcctag ctcagcagat cccaccccca tggagcacgg
 15 33721 caagtaaga tctgctggtt tgaaattctc actgccagca cagctgcctg aagtcaacct
 33781 gggatgctcc agcttggctg ggggaggggc atccgccatt actgaggctt gagtaggctg
 33841 ttttcctctc acaatgtaaa caaagccact gggaagtttg aactgggtgg agcctaccac
 33901 agctcagcaa agcccctgta gccagattgc ctctctagat tctccctctc tgggcagggc
 20 33961 atctgggaaa gaaaggcagc agccccagtc aggggcttat agataaaact cccatctcat
 34021 gggacagagc acctgggaga ggggtggtg gtgggccag cttcagacaga cttaaatgtt
 34081 ctttgcctgt tggctgtgaa gagagcagtg gatctccag cacagcactt gagctctgct
 34141 aaggagacaga ctgccttctt aagcaggtcc ctgaccctcg tgattcctga gtgggagaca
 34201 cctcccagca ggggtcgaca gacacttcat acaggagagc tctggctggc atctgggtggg
 25 34261 tgcccctctg ggacaaacct tccagaggaa ggaacaggca gcagtctttg ctgttctgca
 34321 gcttctgctg gtgataccca ggcaaacagg gtctggagtg gacctccacc aaattccagc
 34381 agacctgcag cagaggggcc tgactgttag aaggaaaact aacaaacagg aatagcatca
 34441 acatcaacaa aaaggatgac cacacgaaaa ccccgtaaca aggtcgccaa catcaaagat
 34501 caaacataga taaatccaca aggatgagga aaatccagca caaaaaggct gaaaattcca
 30 34561 aaaaccagaa tgctcttct cctccaagg agcacaactc tttgcagca agggaaacaa
 34621 actgtaggga gaatgagttt gatgaattga cagaagtatg cttcagaaaag tgggtaaaaa
 34681 cagactcctc caagctaaag gagcatgctc taaccctaatg caaagaagct aagaaccttg
 34741 aaaaaagggtt ggaggaattg ctaactagaa taactagttt agagaagaac ataatgacc
 34801 taatggagct gaaaaacaca gcacgagaac tctgtgaagc atacacaagc ttcaataact
 34861 gaatcgataa agcagaggaa aggatatcag agattgaaga tcaatttaat gaaataaagc
 35 34921 atgaagacaa gattagagaa aaaaagaatg aaaaggaagg aacaaagcct ccaagaaatg
 34981 tgggactatg tgaaaagacc aaacctacat ttcatgtgtg tacctgaaag tggcggggac
 35041 aatggaacca agttggaaaa cactccttag gatattatcc aggagaactt ccccaaacta
 35101 gcaagacaag ccaacattca aattcaggaa atacagagaa caccacaaa ataccctca
 40 35161 agaagacaa acccaagaca tgtaattgtc agattcacca aggttgaatg gcagaaaacaa
 35221 aagtttaagg gagcgagaga gaaaggtcgg gttaccctaaa aagggaagcc catcagacta
 35281 acagtggatc tctcagcaga aacctacaa gcctacaagc cagaagagag tgggggcca
 35341 tattcaacat tcttaaagaa aagaattttc aaccacgaat ttcatatcca gccaaactaag
 35401 cttcagaagt gaagtagaaa taaaatcctt tacagacgag caaatgctga gagattttgt
 35461 caccaccagg catgccttac aagagctcct gaaggaagta ctaaaataag aaaggaaaaa
 45 35521 ccggtaccag ccaactgcaga aacataccaa attgtaaaga ccattgaaac tatgaagaaa
 35581 ctgcatcaac taatgggcaa aataaccagc taacatcata atgacaggat caaattcaca
 35641 cataacaata ttaaccttaa atataaatgg gctaaatgcc ccaattaaaa gaccacagac
 35701 tggcaaattg gataaagagt caagacccat cagtgtgctg tgttctggag acccatctca
 35761 catgcaaaga cacacatagg ctgaaaataa agggatggag gaagatctac caagcaaattg
 50 35821 gaaagcaaaa aaaaagcagg ggttgcaatc ctagtctctg ataaaaacaga ctttaaacca
 35881 acaaagatca aaagagacaa agaaggccat tacataatga taaagggatc aattcaacaa
 35941 gaagagctaa ctatcctaaa catatatgca cccaatacag gagcaccag attcataaag
 36001 caagttctta gagaccaca aagagaccaa gactcccaca caataatagt gtgagacttt
 36061 aacaccccaa tgtcaatatt aggtcaacga gacagaaaat taacaagcat attcaggatt
 55 36121 tgaactcagc tctggaccca gtggaactaa tagacatcta cagaactctc caccatcat
 36181 caacagaata tacattcttc tcagcaccac atcacactta ttctaaaatt gaccacataa
 36241 ttggaagtaa aacactcctc agcaaatgca aaagaatgga aatcataaca aacagtctct

	36301	cagaccaaaag	tgcaattaaa	ttagaactca	ggattaagaa	actaactcaa	aaccatacaa
	36361	ctacagtggga	aactgaacaa	cctgctcctg	aatgactact	gagtaaataa	caaaaagaag
	36421	gcagaaataa	atacattatt	tgagaccaat	gagaataaag	atacaacata	ccagaatctc
5	36481	tgggacacag	ctaaaacagt	gtttagggga	aattcatagc	aataaatgcc	cacaggagaa
	36541	agcaggaaaag	agctaaaatc	aacactctaa	catcacaatt	aaaggaaacta	gagaagcaag
	36601	agcaaacaca	ttcaaaagct	agcagaagac	aagaaataac	taagatcaga	gcagaactga
	36661	aggagattag	agacacaaaa	aacccttcaa	aaaaatcagt	gaatccagaa	gctggttttt
	36721	tgaaaagatt	aacaaaatag	atagaatgct	agccagattg	ataaagaaga	aaagagagaa
	36781	gaatcaaata	gacgcaataa	aagatgataa	agaggatatac	accactgatac	ccacaaaaat
10	36841	acaatctacc	atcagagaac	actataaaca	cctctatgca	aataaactag	aaaatctaga
	36901	agaaatgaat	aaattcctgg	acacatacac	cctcctaaga	ctaaagggaag	aagtcaaatt
	36961	cctgaataga	ccaataataa	gttctgaaat	cgaggcagta	attaacagcc	taccaaccaa
	37021	aaaaagccca	ggaccagacg	gattcacagc	tgaattctac	cagaagtaca	aagaagagct
15	37081	ggtaccattc	cttctgaaac	tattccaatc	aatagaaaag	gagggaaatcc	tccctaactc
	37141	attttatgag	tccggcatca	tcctgatata	aaaacctggc	agagacacag	caaaaaata
	37201	aaattgtagg	ccaatatccc	tgatgaacat	tgatgcaaaa	atcttcaata	aaaaactggg
	37261	aaactgaatc	cagcagcaca	tcaaaaagct	tatctaccat	gataatttgg	cttcatccct
	37321	gggatgcaag	gctgggtcaa	catatgcaaa	tcaataaaga	taatccatca	cataaagaga
20	37381	accaatgaca	aaaaccacat	gattatttca	atagatgcag	aaaaggcctt	tcataaaatt
	37441	caaccagccct	tcattgctaaa	aactctcaat	aaactaggtta	ttgatggaaac	atatctcaaa
	37501	ataataagag	ctattttatga	gaaaccacaca	gccaatatca	tactgaatgg	gcaaaagctg
	37561	gaagcattca	tttgaaaacc	ggcacaaaac	aaggatgccc	tctgtcacca	ctcctattca
	37621	acatagtatt	ggacgttcta	gccagggcaa	tcaggcaata	gaaagaaata	aagcatattc
25	37681	aaataggaag	agaggaagtc	aaattgtctc	tgtttgcaga	tgacatgatt	gtatatttag
	37741	aaaaccccat	catctcagcc	caaaatctcc	ttaagctgat	aagcaacttc	agcaaagtct
	37801	caggatacaa	aatcaatgtg	caaaaatcac	gagcattcct	atacaccaat	aatgacaaac
	37861	agccaagtca	tgagtgaact	cccattcaca	attgctacaa	agagaataaa	atgcctagga
	37921	atacaactta	caagggatgt	gaaggacctc	tttaaagaga	actacaaacc	actgctcaat
30	37981	gaaataagag	aggacacaaa	caaattgcaag	aacattccat	tctcatggat	aggaagaatc
	38041	aatatcgtga	aaatggccat	actgcccata	gtaatttata	gatccaatgc	tatccccatc
	38101	aagctaccat	tgacttttctt	catagaatta	gaaaaaacta	ctttaaattt	catatggaac
	38161	caaaaaacag	cccgtatagc	caagacaatc	ctaagcaaaa	tgaacaagct	ggaggcatca
	38221	tgctacctga	cttcaaaacta	tactacaagg	ctacagtaac	caaaacatca	tggtactggt
35	38281	acataaacag	atagatagac	caatggaaca	gaacagaggc	ctcagaaata	acgccacaca
	38341	tctacaacca	tctgatcttt	gacaaacatg	acaaaaacaa	gcaatgcaga	aaggattccc
	38401	tatttaataa	atggtgtcgg	gaaaactggc	tagccatttg	cagaaaactg	aaactggacc
	38461	ccttccttac	acgttatata	aaaattaact	caagatggat	taaagactta	aacataaaac
	38521	ataaaacctt	aaaaacccta	gaagaaaacc	taggcaatc	cattcaggac	ataggctagg
40	38581	caagacttct	atgactaaaa	taccaaaagc	aatggcaaca	aaagccaaaa	ttgacaaatg
	38641	ggatctaatt	aaactaaaga	gcttctgcac	agcaaaaagaa	actaacatca	gagtgaacag
	38701	gcaaccgaca	gaatgggtga	aattttttgc	aacgtatcca	tctgacaaaa	ggctaataatc
	38761	cagaatctac	aaggaacctt	aacaagttta	caagaaaaaa	aacaacccca	tcaaaaagtg
	38821	ggcgaagggt	atgaacagat	gcttctcaaa	agaagaaatt	tatgctgcca	acaaacatac
45	38881	gaagaaaagc	tcatcatcac	tggtcattag	agaaatgcaa	atcaaaaacca	cagtgcagata
	38941	ccatcttatg	ccagttagaa	tggcgatcat	taaaaagtca	ggaaacaaca	gatgcaggag
	39001	aggatgtaga	gaaataggaa	cactttttaca	ctgttggtgg	gagtgtaaat	tagttcaacc
	39061	attgtggaag	acagtgtggt	gattcctcaa	ggatctagaa	ccagaaatat	cttttgacct
	39121	agccatccca	ttactgggta	tatactcaaa	ggattataaa	tcattgtact	ataaagacac
50	39181	atgcacatgt	atgtttattg	tggcactatt	cacaatagca	aagacttggga	accaatccga
	39241	atgcccataca	atgatagact	ggataaagaa	aatgtgacac	acatacacca	tggaataacta
	39301	tgagccata	aaaaaggatg	agttcatgtc	ctttgcaggg	acatggatga	agctggaaac
	39361	catcattctt	ggcaaggtaa	cacaggaaca	gaaaacccaa	caccacatgt	tctcactcat
	39421	aagtgggagt	tgaacagtga	gaacacatgg	acactgggag	gagaacatca	cacactgggg
	39481	cctgtcaggg	tgtaggaggc	taggggaggg	atagcattag	gagaaatacc	taatgtagat
55	39541	gacaagttga	tgagtgcagc	aaaccacat	ggcatctgta	tacctaggta	acaaacctgc
	39601	acgttctgca	catgtacccc	agaacttaaa	agtattatta	ttattattat	aataataata
	39661	ataaaagaaa	tacaataaaa	tagaatgcag	catacagcag	tgatttctcaa	acacattcag

	39721	catcagaatt	acccttgaat	ctttaaaata	tatatacata	tgagatctta	gtctccaaga
	39781	tttghtaagt	tggtattggg	tccctggggc	tatgttgggt	ttagaaactt	ctacagatgg
	39841	tttggatgta	tgggacagtt	taagaatcgc	tgaactaaaa	tcaaataaac	tgaatatcct
5	39901	gtgattttaga	gagacttatc	gtttatttca	ctatccaagt	acttgcatta	gagcgtggct
	39961	agaagggatt	tgagcccttg	taaataatca	gaaattcaga	cattttgaga	tgagagaact
	40021	gctgaagatt	ttattctgac	ttgaaataaa	ttttctaatt	agaaacttcc	aggtgagagc
	40081	aaaggcctgg	aacaatatc	ctgagccaga	ggaggatcga	gtttgactcc	aggcctaaca
	40141	cttactaggt	ctatgacctt	gggtcagtaa	tttaaattct	ctgtatctca	acctctcaac
10	40201	agggtattgg	tagggattaa	atgtgttagt	gtctgtgaag	tgcttagagc	agtgttggc
	40261	atagtaaatg	cttaatgaat	ttcagccact	gtttttat	ttagtacttt	ccagctcccc
	40321	caaaaagata	ctttttttag	acttgtatta	agacaataaa	aagttaatc	agcatgcttc
	40381	atacctaaat	atgcttcact	ttatagcaaa	gtttacaaga	ctaaaactgt	ttgtgtgtaa
	40441	ttctctgagt	ctcatgtgtt	tattaatgat	ttttctgct	gtttattcat	ctgaattcta
15	40501	ctcattcttc	aagacctagc	tggaaatcctg	tttctagaaa	gactcttgcc	cataataata
	40561	aacctgccct	atctgagttc	ctagggtggtc	tgtacctcat	aatttggtaa	ttaattgtat
	40621	atgcacttat	ataacaaaac	attattgtgt	gtctttgctg	tatcagattc	taggctggaa
	40681	gttgtagata	tgatgttttt	gtctagaaaa	atgttctaga	atgtcctact	caggacagtc
	40741	tggtgacttt	aaagacacat	ttcctaataa	gacacttcat	gaggcagccc	gagcctgtac
20	40801	ctgtgttcct	ggacctgatg	atcaagtttg	atttaagcct	caccacttac	tagctctgtg
	40861	attttgggca	agttacttga	attctctgtg	tgtagataga	acaatgttga	gggaaatccc
	40921	tttcccccat	ccttgtgttt	ccacaaggga	acttgcctcc	taataagtaa	cactttcagg
	40981	ggaatattct	aggcccttct	cttatcccca	ttacttgctc	tttctgtgaa	aagaggagag
	41041	gttaatctga	tgatgaaat	ccttaatctt	tcattctctg	gactgtagag	cctgtgaacc
25	41101	aaagcaatgg	accacttgca	ctgaaattga	ggctgaccct	gtattttgat	tcttatttgg
	41161	caacttattt	ctattctgtt	cccaattcaa	aatcccaagg	ggagaaggaa	gataattgat
	41221	taccagaagt	atgtaatggt	ggtaggaagt	tgaataaatg	gtaacttttt	aaaagttgca
	41281	tgagatatag	tccttatccc	agagaagcta	agtttgcttt	tctttcctct	catgtatttt
	41341	agtattattt	ctacaattag	attgtaaacc	ctttaaagc	aagaatattt	ctacattttc
30	41401	ttactcctga	tagcacacag	tagactgtcg	ggcacatata	tagtaggttg	ctctgtaggt
	41461	acttgctaaa	tgattcaaca	tgtttttccc	tcattggaaa	gaaagatttc	agtattgttc
	41521	ttatcagcta	ggaaggcact	ctgaatagga	aatcagttct	aggcaggtat	ccataaatgg
	41581	gttatgattt	ccaacttact	tgccccagag	gctcgcta	gttgaactct	tcattgggtac
	41641	tttgtcttgc	ttcatgagct	atacatgcta	aggggttagc	agatcatata	atcttttgat
35	41701	ctacaaaata	tgatctttat	tgaacaaaaa	cttgggccaa	aggcctttct	cctttgccac
	41761	cttcctccct	cttttcattc	tcttttttgg	gaatgccctt	tgtgcatgtt	agttacagca
	41821	tgtaccacat	tgactgttat	tggttggttt	tgggtctaac	ccacccttaa	cactgcagtc
	41881	cccaagggca	gaaattcagt	ctcattcatt	ttgatgtcct	cagtgcctgt	gctcagagaa
	41941	tatctattat	ttgaaaaaat	agtgcaaaag	taaattttag	gagactacat	cacactcatc
40	42001	taaactgcaa	gtttgacaag	ttgacatcca	aaagaaaggc	tctcctaata	aacctcgcca
	42061	cagaaatttg	ggtgaccttt	gtagctctgg	agaaagcaga	ggcaaaaatg	aaacctaaaa
	42121	attatttgtg	ggttttttaa	aaatgttttc	tcattggagta	aaggctctaca	gctgagttct
	42181	tttcatatga	gggaatgaca	gaaacacagc	tggttctgac	tttcagcttc	aactgagcga
	42241	ccagagctct	gctggtgaaa	caggaacttg	tattgtgccc	ctgacgtgca	ccttgaagggt
45	42301	gtcagctcat	tgtccctttg	ttcacataaa	tagtttttta	agaattgttt	ttgatcttgt
	42361	gagcctctaa	ctaaatgatt	aacctatgca	agttggccat	ttggggtaat	actgaagcac
	42421	ttctcttgag	ggctattgac	aggtgggaat	gtgcccacct	ccttgggtct	ctggttttca
	42481	tgtcatactt	gcaaatcagt	gacagtttaa	acttggggca	atcacttagc	aagtctattg
	42541	agttaccaag	ttaattattc	ccactttgca	tgaagcaacc	ttgaaaatga	tttccctaaa
50	42601	gcaaagtaca	tccaaactca	gtaccttctt	aataaccttt	gctgaatgaa	taaatgacta
	42661	attcataaaa	aatgtaacat	atcttttaatt	cttacttacg	ggcagtttaa	gcctcttgtg
	42721	taagaggagg	cctcggttg	agataacata	ggatagtaag	cctcctagag	aaatttctat
	42781	atggaaacat	ggtctgctat	gaagctagaa	gtgagaggac	attatatttg	accattatat
	42841	ttggcttcag	agcttctcaa	catggggccc	aaagtcaagg	tcccttggtt	cattaagagg
55	42901	aggtccagga	gtgcatgaca	cccatcagac	tactgagacc	cagctggaac	taggcacctt
	42961	gcacaggggc	cttgccctaat	caaaaatagtt	cttatttttt	ctgagttcca	agtaactagt
	43021	ttcctaacc	agtgtctgga	tagtagtgcc	aagtgggagt	accttcaatg	aacttctcta
	43081	tgaggttatt	tctagcctat	tggaaatgtt	cgtttttagga	gggtgaggaa	gggaagtctt

	43141	gaattttttgt	gcttagtttta	atgtttgtgat	acagcttttga	ccatccgttt	aatgggagat
	43201	ctgtttttcca	gatgactata	catgtggaaa	ggagaagttt	tttgagtgtt	ttttttaacc
	43261	cctttttaaag	aatgggtttt	catttagtct	ctacatttgg	gggtaaaagg	tcctctaggg
5	43321	agactttttca	aaagtatttg	aagtttgcac	ctgatttcag	aggtgagttg	gaggcctatc
	43381	tgtgtatgac	agacacatgt	ctccaacaac	tatatgttca	caaggactaa	gagccatcct
	43441	tttgggtcca	tcattcaaca	ttgatctcac	attcgtgttc	gtatcagtat	ctttacagtg
	43501	cgctcccagt	tacatctccc	taatttccct	tagtaggctt	cacagaattt	gcagtgtatg
	43561	caatggcaga	tgaccacatg	tggagtcatt	taaccacatc	ttccactgca	agtcagccccg
10	43621	ctcttgatgt	ctgttttatgt	ttagattcca	tcttttggaa	gatttcattc	ctctgcaacta
	43681	tctcagtatc	tcagatgctt	ttgagactgg	gtccttttcc	cctcctatgt	ttggccatgg
	43741	ccacccccctc	aggggttgtg	tgtgtttcac	agctgctgtt	tgtagggttg	acctttacaa
	43801	tgtacaaagc	tctttcccat	atgttgacaa	tccctgggtg	gatgctgtga	gttaggcagg
	43861	gtgtgtatac	gtgtcctcat	catattacag	tggttaaggca	acagggtttt	tgaatttgat
15	43921	cacccatgaa	tttgtccta	ttgttggtaa	aaaatggtca	tgtatcagcc	gtttcacagg
	43981	gtcagcttaa	tagaaagtgg	gagttaggca	ggaccagaat	tcaggacttc	agccccgggt
	44041	cccaggggact	attctctata	cccaattgtc	ccaccttgaa	tcagtttctt	ctagggaagt
	44101	atctccaaaa	ctgagatggc	acccacagga	cttcttaatt	gtagtcatta	ccaggaaaaa
	44161	caagcaagg	aactgggtga	aatctctgtt	tttgggtgatt	gggtgagatt	tggagattgt
20	44221	cttgtgtcaa	aagtaaaagg	actagattaa	atgttttgtt	aataaattgg	ttatttttaa
	44281	tttaattatt	tgacagttaa	tttacattat	tcaaaaatca	aaataaaatt	taaaagaagt
	44341	ttactactgaa	aagtcttgcc	ccacttatac	cctgctcacc	tcagtatccc	ccaatacata
	44401	ccatctataa	ggtgatcatt	tgtattagtt	tcttgtgaat	ccttgatagt	gtgttttata
	44461	tagatacagg	taaatatgag	tatgtactat	tatttcccc	ccaccccacc	ctgttttttt
25	44521	tttgagacgg	agtctcgctc	tgctgcccag	gctggagtgc	agtggcacga	tctcggtcca
	44581	ccgcaagctc	caccattttt	ccccattttt	aaaacaaaag	gtagtagcca	tatatacact
	44641	attttacacc	ttgtttttat	acttactaat	atataccaga	gagctttcca	tcatttttga
	44701	catatgcacc	tatatctgtc	aattattccc	agaagtggaa	ttgctgggtc	agcaggaaaa
	44761	atcatgtata	attttgatga	gtattgccta	attgtcctgc	acagggttg	aattgtttgt
30	44821	actcccacct	ttagtgtatg	agaagacctg	tttctccata	gcctcatcaa	acagagtgtg
	44881	tgagattaga	tgagaaatag	gaggtgagca	gtctttttacc	ccatccgtag	tttgcagtgg
	44941	gaacactgca	cagttgcaag	agctggtgca	ggtatcagat	tagttccagt	ggaaacgctg
	45001	cctcaccatg	gccatgggct	tgccgacagc	ctagtacac	acacggaatg	gacccacgtt
	45061	gccacttgca	gaatttcctg	tagcagaaag	ttgaacatgc	attcattatt	catctaacta
35	45121	gccatgctgg	atctaaagag	cacaacagtg	tttttttagaa	ccaaaaagaa	aattgtttca
	45181	ctacaacaca	ctgtgtataa	ggctttcaat	gctcttttct	cagctattaa	cattattttc
	45241	aggactgagt	tcaagagatg	tatcccaa	cacagggatg	tcttgctaag	cttggaactt
	45301	tcatactcaa	gggatgcttt	tttgaggaa	gattttacac	ttactcaaca	tttgtaatta
	45361	aataataagt	actttataag	ataaatttaa	actgtccaag	tacaataata	acattgaact
40	45421	atgatgcatt	attgctagac	tttttcctta	aagttgccaa	gtggtttcct	gcattaggca
	45481	aataggggat	catataaaaa	tgccatgatt	tacggcctag	ataacatctc	caccatttga
	45541	gcagcatata	ttccagggtca	tccccacata	actccttacc	attctcatta	gaaaggttga
	45601	ttcttagtct	tattttttctc	tgaggacagc	aaaaaaaaaa	atccccttca	gttccactgc
	45661	atagaaaagt	gtggtaaagt	gagccgggca	cagtggttat	ttaatttaaa	tggaacaatat
45	45721	tttttataga	attttgacag	ggccactgta	taggggaaag	tcactcctct	tcccctttat
	45781	agaagagttg	cacctggaca	gttgcatgta	tgactgtatc	cagtctacac	aagaggtcat
	45841	tcctgggcat	agaatggac	tgccaaaatc	tagctgaaac	accattgaca	aatagacatt
	45901	ttcttttgtt	aataatacct	gtgaaggctt	tcataacaga	catttccagt	tttgttctca
	45961	ggctccttgc	agctgctcct	ctaaaagtgt	gctctcttcc	aagagctgac	aatggccaga
50	46021	agcaagggtg	tctgtctttt	gtgccatcat	catctaactt	gccacacaca	tttgggatgt
	46081	cagcctaggt	ataggttttt	tatccactca	gtatggcttg	tggtgtgttg	tgcttttgtt
	46141	attcatgctg	agggcctctg	ggcatcagtt	tggtgtgaga	gaacccattc	catgaccctc
	46201	cttccttttg	ctgttttgac	tcgatggctc	ttgttggcac	agtctgtgag	tgtctgatgc
	46261	tctatccatg	ccggaccatc	tggtctgctg	tctctgtggt	ctgaagtcgt	tttctgaact
55	46321	attccttgat	aataaatttg	agatgatctt	gttctacctt	tcttttcaag	tcacatctta
	46381	gccccttagc	cacattcccc	aagaacatga	caaatggatg	ggtcacaagt	cacgtagcat
	46441	aggggtgtcag	accacgaggc	tttgaaggga	ttctgttggg	tgctaaaaag	aaagattttg
	46501	tgtcaccacg	atttttttta	aaggcatggt	gacacttagg	ccttaattga	aagcgttctt

	46561	actcaagtag	agttgacaga	ggagatatttg	gtagtcgcgg	ttgctgggtct	gaagagcatg
	46621	tggttctggt	tcaatgccca	atgagatctt	ctcacgggaa	aatgttctga	catctcaaac
	46681	aaatgacctt	catgcatagt	tttgacaaaa	taccctatta	agtatgcata	tatggttggt
5	46741	accttggtgt	aataattcaa	tactggaaac	agagtagcaa	caaagaaaca	ttagggttat
	46801	atttaacctc	tgtggaatta	gtgtgtaaac	aaactgctta	tcagaaatgc	tcatatgggg
	46861	ctttgtttta	ataaataaga	aactggcata	tagggctctgc	aggatatttc	tgccaagtag
	46921	acctccctca	cattataaga	caccacatct	atgtctgacc	ccatatggaa	agaggcatag
	46981	caagccagca	ctgggtcata	ttccctctcc	accacataat	gggtatgtga	tcttagggaa
10	47041	tccaccgaaa	ctctctgggc	ctcagtttcc	tcagctataa	atgggtggata	atcaaattat
	47101	ttacctcacc	attaataaat	gttagctatt	atTTTTtattc	aagtttaata	caaagagaaa
	47161	cattttactt	atTTTTccag	ctatccagag	catcttccaa	aatcctatca	ccaacaaata
	47221	ctgtattgta	tttattatag	caactatgta	aaaaatggagt	ccctgtccta	tgcttagatg
	47281	aaatatgttg	gtatttgagt	ttgcatgtct	tctataggaa	tcagtgttta	gtgaaaacgg
	47341	gtggagataa	acagatgttt	tcacagtcct	gttggttcaca	gtaccgccaa	attgaatgtt
15	47401	tccatatagg	tgcatctctaa	tggcttaaat	gatgcagata	ttttctggcc	agccatatgg
	47461	atcttttgtc	atctaagatg	ttaatatattt	ccttatattt	tatagtagtt	ctggagtaca
	47521	gccagtttct	tgaatagggt	ccacatggct	cattatgcac	agggcctgga	aactgcctta
	47581	ctcgtgctgt	tgaatgaac	cgtgacactt	cagaagagct	gggagctggg	gtagagcagt
	47641	ggctaggaga	acataattcaa	ttatatattcc	tcctgcatta	agctacaaat	aatgagcact
20	47701	ttcctgtgct	ttacagttaa	gtaattaaaa	gaaattatag	agtgggatgc	aaaaataacc
	47761	cgaaggacaa	ctggatgtgt	ggagccacca	gttttctcca	tgagtgcaca	aggttaatcc
	47821	ttgttactac	tcagaatgct	gagtttctac	agaaaggggt	gcaggtccac	acatgttttg
	47881	gcgtctaccc	acacgcttct	gtatggcatg	actgtgcatc	ccagaagaag	ggctgtgctg
	47941	tgtacctcca	cgtttcagtg	gaatttaaca	aactgatccc	tgaaaatggg	ttcataaagg
25	48001	tgagtaacag	agagctaata	gccttctctt	gctaatttta	tctttccccc	aagatttctt
	48061	gataatagtt	tgaaaaggag	tgttattctt	tggtctctag	aggcaactta	cctttccagt
	48121	ttcttccatc	acctgttttc	atctctcttg	tttttttaaa	tttaatgctg	tatgtatttc
	48181	agaggatagg	atctaactta	gtgcggctcc	ttcatcaggt	gagaattatt	catctcattt
	48241	tcatttttagc	ccttctgaat	taatgacatt	gaagcccggc	agtttggtcc	taagatgggt
30	48301	ttaattatgt	acagatactc	tttctataat	ggaaattgct	cagataacta	attaaccaca
	48361	agaatacact	gtctatggaa	aatttcagga	gcaccgtctg	tggaataact	gggaagggca
	48421	tgctgtcacc	acagctctgg	ggtctattaa	aagtgtgggt	atgcagcact	ggtgtctagt
	48481	gggtgtgttg	ctctcaactg	ccagaattcc	catagcattt	catggcagaa	agtcaagggtg
	48541	tccagcaata	ctctgaaagt	gacctgttga	ttaaagtcgt	caattctgaa	gaaagagact
35	48601	gaaataagac	aaatgggtct	taactttttt	tctctttctc	tctcttgtaa	aaatgtgtga
	48661	ttgttctggc	atgttcccaa	tccccacata	atgccaacat	cttttcttaa	agggggattc
	48721	cctttatcct	tggtctctag	aattattgca	tggtctccct	ttagggacaa	tgaatgcagt
	48781	tgcatcacc	ttgttttttt	tttttttttt	gtacacagca	tgcttattct	tgtagtcagg
	48841	gacttgaaa	acaaagcccc	acctggcttt	cacaacatct	cctattagta	ggtgtgcctt
40	48901	gtgtgtaatt	tgaaggaggc	ggtcccttag	ctgtgtttac	actgtacttt	taaatgtggg
	48961	gctgaaggta	gaatcaacca	tacttaagat	gccacctggg	aaaatagggt	tctgtgtcat
	49021	ctcagcccca	cccatttgca	aatgacttaa	cagcagcact	attagggttc	ctagtgtgag
	49081	tcatttgcac	ttggactggg	gaacttggtg	acttcttggt	gtttggaac	aaacaacctt
	49141	tgcatctttt	cgtaaaaaagc	ctgaacagtg	gaccagtctc	cagttctact	tgcaaagctg
45	49201	cccccatcaa	atccctcata	atgttcaact	taaaaaatgt	tacacttttc	tctggaaatc
	49261	taaccttttt	tccttttttt	aaagccattt	taagtacttc	agtcttgaat	caaatgatcc
	49321	caaataattg	acaccaacct	agaaattggg	ttacctctg	ggaactttat	cgaagaagag
	49381	agatttttgt	tgagagggg	gttttgatgt	ttgatactta	tatttactat	tttaattatt
	49441	cattgttggt	gttgctgctg	ctgctgtatt	atTTTtgcgag	tttctgtttg	ttaaattttca
50	49501	tggtatttgg	taggagagag	ctggatctgt	tggtttcagg	acaagtctag	aaataagaaa
	49561	tctgccttga	gtgagttagt	tggttccctc	tggtgtctatt	tcaccattaa	ggacgaaagg
	49621	aactcacaag	gaccagagac	atctggctga	aagcaatact	agtgtgactg	gacatctact
	49681	acctgccata	gttggtcata	tcgtttccag	tatgattctg	attgagttag	tgatattagg
	49741	ctatgttcag	ggatcagggg	ggctaattat	gcttatattg	ccttgtagca	ttttggttaag
55	49801	aattaatgat	tgtgttagatg	tccagattta	ggtcagcaat	attctaaaag	ttctcattga
	49861	actaatcatg	tttataagta	gcctgtactt	tctatcataa	taacaatagt	ggaaaagcta
	49921	gttgacataa	aaggagccca	gatttttactt	aagtaaaaac	acaaaagcaa	agatattttc

	49981	ccacataaat	tacaaaagca	aagatatattt	cccacataaa	tgtccccata	aaacaagttg
	50041	aaccaaaagag	gaaagatgac	aggtaaccgt	atgacacgct	aagaaagtat	cataataactt
	50101	aagttaaactt	caacctttta	tttccttatc	ctaagcagcc	tcttttctct	ttatcattta
5	50161	gtcctgtgct	tctcaacttt	gataagtaaa	aaagtatttg	cactaaataa	atcttattga
	50221	aatgcaggat	ctgattgagt	gggtggggta	ggtggaatga	gggtggggaa	gttgagattc
	50281	tgcattttctt	agaagtttct	actttatgtt	aaaatggcta	atccatctca	acattgagaa
	50341	gtaaggtttc	acttaatttc	agcctgtgta	agtttatccc	atatgtacat	ttcctaaaaac
	50401	tctaattctca	ggccccagga	atttctcctt	tagttaaaat	atttttagga	ataaatttga
	50461	attgcattaa	tacacaattt	ataaatttaa	cacaaaaaat	tatttgaagt	ttgagacttt
10	50521	agggtgcatg	aaatcaattt	catacttgaa	aattttctat	aaattcaaaa	gtctgtgtat
	50581	ttaaatacaa	tttaaatacc	tgtgttacag	tgacatttgt	ttttctgtct	ctctctccac
	50641	catttccaga	gtcatcatcc	ctgtacagaa	aaatttttcc	cacatgattt	caccataaat
	50701	tcattaaata	tgatgcttac	ttgataattt	ctccagggtc	ttttttttt	taattatact
	50761	ttaagtctta	gggtacatgt	gcacaacctg	caggtttgtt	acatatgtat	acatgtgccca
15	50821	tgttggtgtg	ctgcacccat	taactcgtca	tttacattag	gtatatctcc	taatgctatc
	50881	cctccccct	accctactc	catgacaggt	cccagtgtgt	gatgttcccc	accctgtgtc
	50941	caagtgttct	cattgttcag	ttcccaccta	tgagtggagaa	catgcggtgt	ttgggtttct
	51001	gtccttgcca	tagtttgctc	agaattgatg	ccttgctcac	tgatggacat	ttggttggtc
	51061	ccaagtattt	gctattgtaa	atagtgccgc	aataaacata	cgtgtgcatg	tgtctttata
20	51121	gtagcatgat	ttataatcct	ctgggtatat	accagtaat	gggatggctg	gctcaaattg
	51181	tatttctagt	tctagatcct	agagggaatc	ccacactgtc	ttccacaatg	tttgaactag
	51241	tttacagtcc	catcaacagt	gtaaaagtgt	tcctatttct	ctacatcctt	tccagcacct
	51301	gttggttccg	gactttaatg	atcgccattc	taactggtgt	gagatgggat	ctcattgtgg
	51361	ttttgatttg	catttctctg	atggccagtg	atgatgagca	ttttttcatg	tgtcttttgg
25	51421	ctacataaat	gtcttctttt	gagaagtgtc	tgttcatatc	cttcacccac	tttttgatgg
	51481	ggtcatttga	ttttttcttg	taaatttggt	taagttcttt	tagattctgg	atattagccc
	51541	tttgtcagat	gggtagattg	taaaaatttt	ctccatttcc	gtaggtttcc	tattcactct
	51601	gatggtagtt	tcttttgctg	tgcagaagct	ctttagttta	attagatccc	atttgtcaat
	51661	tttggctttt	gttgccattg	cttttggtgt	tttagtcatg	aagtccttgt	ccatgcctat
30	51721	gtcctgaatg	gtattgccta	ggttttcttc	taggggtttt	atgggttttag	gtctaactgt
	51781	taagtcttta	attcatcttg	aattaatttt	tgtataaggt	gtaaagaagg	gatccagttt
	51841	cagctttcta	catatggcta	gccagttttc	ccagcaccat	ttattaaata	gagaatcctt
	51901	tccccatttc	ttgtttttgt	caggtttgct	aaagatcaga	tgggtgtaga	tgtgtgggat
	51961	tgtttctgag	ggctctgttc	tgttccattg	gtctatatct	ctgttttggt	accagtacca
35	52021	tgctgttttg	gttactgtag	ccttgtaata	tagtttgaag	tcaggtagcg	tgatgcctcc
	52081	agctttgttc	ttttggttta	ggattgtctt	ggcgatgcgg	gctctttttt	ggttccatat
	52141	gaactttaaa	gtagtttttt	ttcaaattctg	tgagaaaagt	cattggtagc	ttgatgggga
	52201	tggcatgaa	tctataaatt	accttgggta	gtatggccat	tttcatgata	ttgatcttct
	52261	ctacccatga	gaatggaatg	ttcttccatt	tgtttgctgc	ctcttttatt	tccttgagca
40	52321	gtgggttgta	gttctccttg	aagaggtctt	ccacatccct	tgtaagttgg	attcctaagt
	52381	attttattct	ctttgaaaca	attgtgaatg	ggagttcact	catgatttgg	ctctctgttt
	52441	gtctgttatt	gggtgatagg	aatgcttggt	atttttgcac	attgattttg	tatcctgaga
	52501	ctttgctgaa	gttgcttatc	agcttaagga	gattttgggc	tgagatgatg	gggttttcta
	52561	aatatacaat	catgtcatct	gcaaacagag	acaatttgac	ttcctctctt	cctatttgaa
45	52621	tatcctttat	ttctttctat	tgcttgattg	ccctggctag	aacgtccaat	actatgttga
	52681	ataggagtgg	tgacagagga	catccttggt	ttgtgccagt	tttcaaaggg	aatgcttcca
	52741	gcttttgccc	attcagtatg	acattggctg	tgggtttgtc	gtgaatagct	cttattattt
	52801	tgagatatgt	cccatcaata	cctagtttat	ttagagtttt	tagcacaaag	gctgttgaat
	52861	tttgtcaaag	gccttttctg	catctattga	gataatcatg	gtttttgtct	ttgattctgt
50	52921	ttatatgatg	gatttatatt	attgatttgc	atatgttgaa	ccagccttgc	atcccaggga
	52981	tgaagccaac	ttgatcatgg	tggataagct	ttttgatgtt	ctgctggatt	cgggttgcca
	53041	gtattttact	gaggattttt	ccatcgatct	tcatcaggga	tattggcctg	aaattctctt
	53101	tttttggtgt	gtctctgtca	ggctgtggta	tcaggatgat	gctggcctca	taaaatgagt
	53161	tagggaggat	tccctctttt	tctattgatt	agaatagttt	cagaatggta	ccagctcctc
55	53221	cttatacctc	tggtagaatt	cagctgtgaa	tccatctggt	cctgatggat	ttttttggtt
	53281	ggtaggctat	taattattgc	ctcaatttca	gagcctgtta	ttggtctatt	aagagattca
	53341	acttcttcc	ggtttagtcc	tgggagggtg	tgtgtgtcca	ggaatttata	aatttctttt

	53401	agggttttcta	gtttatttgc	atagaagtgt	ttatagtgtt	ctctgatggt	agtttgtatt
	53461	tctgtgggat	tgggtggtgat	atcccccttta	tcacctttta	ttgcatctat	ttgattcttt
	53521	tctcttttct	tctttattag	tcttgctagt	gatctatcaa	ttttgttgat	ctttttaaaa
5	53581	aaccagctcc	tgggttcatt	gattttttga	aggagttttt	ctgtctctat	ctccttcagt
	53641	tctactctga	tcttagttat	ttcttgtcct	ctgctagcct	ttgaatgtgt	ttgctcctgc
	53701	ttctctaaat	tgtgatgtta	gggtgtcaat	tttagatcct	tcttgctttc	tcttgtgggc
	53761	atttagtgct	ataaaattcc	ctctacacac	tgctttaaat	gtgtcccaga	gattctggta
	53821	tgttgtgtct	ttgttctcat	tggtttcaaa	gaacatcctt	atttctgcct	tcacttcggt
10	53881	aagtaccag	tagtcactca	ggagcagggt	gctcagtttc	catgtagttg	agtggttctg
	53941	agtgaagttc	ttaatcctga	gttctagttt	gaaagcactg	tagtctgaga	ggcagtttgt
	54001	tataatttct	gttcttttac	atttgctgag	gagtgcctta	cttccaaacta	tgtagtcaat
	54061	ttttggaata	agtgtgatgt	ggtgccgaga	agaatgtata	ttctgttgat	ttggagtgga
	54121	gagttctgta	gatgtctatt	agggtccgctt	ggtgcagagc	tgagttcaat	ttctggatat
	54181	ctttgttaat	tttctgtcct	gttgatctgt	ctaataattga	ccgtggggtg	ataaagtctc
15	54241	ccattattat	tgtgtgggag	tctaagtctc	ttttaggttc	tctaaggact	tgctttgtga
	54301	atctggtgct	cctgtattag	gtgcatatat	ttttaggata	gtagctcctt	cttgttgaat
	54361	tgatccccctt	atcattatgt	aatggcccttc	tttgcctcctt	ttgatctttg	ttggtttaaa
	54421	gtctgtttta	tcagagacta	ggattgcaac	tctgtcctttt	ttttgccttc	catttccttg
20	54481	gtagatcttc	ctccatccct	ttattttgag	cctatgtgcg	tctctgcaca	tgagatgggt
	54541	ctgctgaata	cagcacactg	atgggtccttg	actctttatc	caatttgcca	gtccatgtct
	54601	tttaactgga	gcatttagcc	cattttacatt	taagggttaat	attgttatgt	gtgaatttga
	54661	tcctgtcatt	atgatgttag	ctgggtatttt	tgctcgttag	ttgatgcagt	ttcttcctag
	54721	cctcaatgat	ctttacaatt	tggcatgttt	ttgcagtggc	tggtactggg	tgttcctttc
25	54781	catgttttagt	gcttccttca	ggagctcctg	taaggcaggc	ctgggtggta	caaaatctct
	54841	cagcatttgc	ttgtctgtaa	aggattttat	ttctccttca	cttatgaagc	ttagtttggc
	54901	tggatatgaa	attctggggt	gaaaattcctt	ttctttaaga	atgttgaata	ttggccccc
	54961	ctctcttctg	gctttagag	tttctgccga	aagatgctgt	tagtctgatg	gacttcctt
	55021	tgtgggtaac	ctgccctttc	tctctcgtg	cacttaatgt	ttttcccttc	atttcaactt
30	55081	tgtggaatct	gacaattatg	tgtcttttag	ttactcttct	tgaggagtat	ctttgcggca
	55141	ttctctgtat	ttcctgaatt	tgaatgctgg	cctgcctcac	tagattgggg	aagtctcct
	55201	ggataatatc	ctgcagagcg	ttttccaact	tgggtccatt	ctccccatca	ctttcaggta
	55261	caccaatcag	atgtagattt	ggtcttttca	catagtccca	tatttcttgg	aggctttgtt
	55321	catttctttt	tactcttttt	tctctaaact	tctcttcttg	cttcatttca	ttcatttgat
35	55381	cttcaatccc	tttcttccac	ttgattgaat	cagctactga	agcttgtgca	tgtgtcacat
	55441	agttctcgtg	ccatggtttt	cagctccatc	aggctcattta	aggctcttctc	tatgctgttt
	55501	tttctagtta	gccattcgtc	taatgttttt	tcaaggtttt	tagcttcttt	gctaaaagg
	55561	tcaaacatcc	tccttttagct	cggaggaggtt	tgttattact	gatcatctga	agccttcttc
	55621	tctcaacttg	tgaagtcac	tctctgtcca	gctttgttcc	attgctggcg	aggagctgca
40	55681	ttcctttgga	ggagaagacg	tgtctgtatt	ttagaattt	tcagcttctc	tgctctggtt
	55741	tctccccatc	ttattgggtt	tatctacctt	tgggtcttga	tgatgggtgac	gtacagatgg
	55801	ggttttgggtg	tggatgttct	ttctctttgt	tagtttttct	tctaacagtc	aggaccctca
	55861	gctgcaggctc	tgttggagtt	tgctggaggt	ccactccaga	ccctgtttgc	ttgggtatca
	55921	ccagcagagg	ctgcagaaca	gcaaataattg	cagaacggca	aatgttgctc	cctgattgtt
45	55981	cctctggaag	cttcgtctca	gaggggcacc	tggccgtatg	agggtgtcagt	cggcccctac
	56041	tgggaggtgc	ctcccagtta	ggctactcag	gggtcaggaa	cccacttgaa	gaggcagact
	56101	gtccattctc	agatatcata	ttccatgctg	ggaggacccc	tactcttttc	aaagctgtca
	56161	gacagggaca	tttaagtctg	cagaagtctc	tgctgtcttt	tggtcagctg	tgccctgccc
	56221	ctagaggtgg	agtctacaga	ggcaggcagg	cctccttgag	ctgcggtggg	ctccacccat
50	56281	ttcagagcttc	ctggctgctt	tgtttaccta	ctcaagtctc	agcaatgggtg	gacacccctc
	56341	ccccagcctc	gctgctgctt	tgcagttcga	tctcagactg	ctgtgctagc	agtgagccag
	56401	gctccgtggg	catgggaccc	tccgagccag	gcctgggaca	taatctcctg	gtgtgccgtt
	56461	tgctaagacc	attggaaaag	cacagtatta	gggtggggag	tgctctgatt	ttccaggtag
	56521	cgtcagtcac	ggcttccctt	ggctaggaaa	gggaattccc	caaccccttg	tgcttccctg
55	56581	gtgaggtgat	gccccaccct	gctttggctc	atgctccgtg	ggttgtaccc	actgtctgac
	56641	aagccccagt	gagatgaacc	cggtagctca	gttggaatg	cagaaatcac	ccgtcttctg
	56701	catcactcac	gctgggggct	gtagactgga	gctgttcata	tttgccatc	ttggaacctc
	56761	cctttccaag	ttctttatta	cagagtgggt	cactgaaact	tcatggaaca	aattggaat

	56821	tatcttctta	attaatgtca	ctgtctacca	tgtatgggaa	tttggtaa	attatatggt
	56881	ttcaataaca	tagtagatag	aacattgtca	aatctaaact	tcagtgaatt	gtaacagatc
	56941	ccacctgaaa	ttctaaagaa	aacagaattc	taattgaaga	ggttaaactt	ttacagggaa
5	57001	tgtcaactgc	catttgggtc	ctgtaaacaa	aaaactgttt	tttaaaaaag	taaactttta
	57061	aagtattttc	agatgacctc	atttgctatc	caagtggctt	gagtatgctt	gatgctaaga
	57121	cttctttgtt	acagactgga	gatgtgtgct	actggggcag	tgttgctctg	tgacaaggag
	57181	gcagaggatg	agggcaaggt	tcgatgtgac	tgtgaattct	gggtggctct	ggctatcggg
	57241	agccttcatt	gattacagca	aaacagttgc	tttcctaggg	caatagtgtc	tctgtcacc
10	57301	aggctggagt	tcagtggcat	gatcaatcgc	tcactgtagc	ctcaacttct	tagactcaag
	57361	taatcctccc	acctcagcct	cccaagttagc	tgaaactaca	gggtgtgcacc	accacaccta
	57421	atTTTTTTTaa	TTTTTaaagt	TTTgtagaga	catgggtctca	ctgtgttgcc	cagggtgatc
	57481	tcgaattcct	gggctctagt	gatcctcccg	cctcggcctc	ccaaagtgtt	gggattacaa
	57541	atgtgagcca	ctgcacctgg	ccctttgcaa	ccttcttgac	aatgcattcc	tttattccct
15	57601	aactggaagt	aacttctttc	tctttataaa	attgtatctg	taccttttct	gggtcatttc
	57661	tacctttata	ttctagttag	gtatgtccta	cctccctcct	aggaggagg	gtaagtaaga
	57721	ctggaaagta	gacttcatgt	gtgatgaatg	aatgaacaaa	aggagtcta	acatatggat
	57781	atagtcaact	ggatgcaa	taaaaaattt	taaatattga	tttgcaagat	ttcattaagg
	57841	tcaactctta	atagttagt	tcatatatgt	taggaaccaa	atattaataa	cttcttcagc
20	57901	attaccatta	tctttatagg	actgtctaaa	atgagcagcc	atatctttaa	actgtgtttt
	57961	ctctgattac	acgctcacag	gtaaaaccca	aaggggctgg	gaacaaacaa	gacttttttt
	58021	TTTTTctgta	tgctgaatt	atctgtactg	ttgcttgttt	tcccaccttt	ggccatagaa
	58081	acttagttct	aacatgctac	aatttttgca	gttctttctc	ttagaaaaag	accacattgt
	58141	ctgaaatttc	atccatttaa	gtaatcaagc	cctaaagtgt	aaggatcttg	gtcatgatta
25	58201	atctagacct	acaaagtagt	atcttaatgg	cactcctttt	agaaagttag	gttccaggac
	58261	acacatagct	gcagtgtcca	cattttgtaa	gctccttcgt	tgtcacagcc	actctcttct
	58321	ctgtggctga	tattctaaaa	ctggcaacac	atcctgatgg	taaaagcttg	gttcaggaga
	58381	caggtagacct	actagcttta	tggcatttga	caggttacct	aacctctctg	acgcataatt
	58441	gcctcatcta	tataatgggg	ataataatac	ccatcctgtc	tcttgtaaa	aatcaaatta
30	58501	gatgacgcct	gtgaatgttc	tatagtctct	tagacaaatg	taagttatga	ctacagcaag
	58561	agtaaaaagag	catgttggtta	tggcaattct	ttcagtga	tgtctaagac	ttgtgagtca
	58621	cacttaaaagc	taaaacttgat	atctacttca	ttgattttct	ttttagtctt	atgtactata
	58681	ttgaatttcc	tgacagtggg	gctatgaaag	ccttcctagc	attttataga	tgtgggtgaa
	58741	ttaatggctg	taagccttaa	agcagaatta	gacagcatca	atgaatttat	taagtataaa
35	58801	taaatatata	atctgcttag	caatattaca	cagcctcttt	atcttatgtg	tgataaagag
	58861	tcacccgaag	gttgaaaatg	aagaattgtc	ctggaagctc	ttacttaatc	ttttattatt
	58921	tcctaataca	gtatataaaa	ttactcattg	aaagcttagc	agaataagaa	acaagaagtt
	58981	aaaaggctga	aaactacaaa	ttttgctatt	attattgtta	ttacttccca	agtctcttat
	59041	tgatctgtta	gaaatagagc	tacacaggaa	attgtaggac	agttagtagt	tggtagtggt
40	59101	atctgctttt	taattattca	agtaaggttt	tattccatta	gaggaactca	ggaagtgtgt
	59161	catggctgat	aattgctatc	tgtcaaaattc	cttagagcag	ggatccgcaa	caccaggcc
	59221	atggattggt	accagtccct	ggcctgttag	gaaccaggct	gcacagtagg	aagtgagcgg
	59281	cgggtgagca	aacattgcc	tctgagctct	gtctcctttc	agatcagcag	cagcattaga
	59341	ttctcataga	agcatgagcc	ctgttgtaga	ctgcacatgc	aagggatcta	ggttggtg
45	59401	tccttatgag	aatccaattc	cttatgataa	tttaactgat	gatctaaggt	ggaaccattt
	59461	catcccaaaa	ccatccaccc	tgctactccc	agaccgtgga	aaaattgtct	tccacaaaac
	59521	tggttcctgc	tgccaaaaag	gttgggacca	ctgccttaga	gtttataatt	tgggggttagc
	59581	acagcctata	tttacctgag	aatttcaatg	ggttactga	tctttccaaa	tgaaaaggct
	59641	tcttacgaaa	attatatcca	aactgtcttt	tctcttagtt	taataaacct	atcagtaagt
50	59701	ttttactgag	tactgtatt	acatttttct	ctgttaagca	ttatgggggc	tcagacatga
	59761	tccattccct	caaagaactt	acctttcagc	tgaagactga	ctagaatgag	caaatacgg
	59821	taacaattaa	caagttagta	ggccagctcg	gccaacatgg	tgaaaccctg	tctctactaa
	59881	aaatacaaaa	attagccggg	catgggtgtg	ggcgctgta	atcccagcta	cctcctgctg
	59941	aggcaggaga	attgcttgaa	cccaggaggt	ggagattgca	gtgagccgag	attgcacat
55	60001	tgactccag	cctgggcgac	agagcaagac	tctgtcttgg	gaaaaaaaca	aaacaaaaca
	60061	agtgagaagg	gaatcaagta	ctacgtaaga	tgtaatgtgg	aatttttaggg	aaggaaggca
	60121	gtgtatgctg	gagtaattag	agaaaggggc	atgcatgatt	tgatgcttga	actggatcat
	60181	gaaggataag	caagattttg	gcagcaagtg	agagggagag	aggagtgtgt	cagagggaatg

5 60241 gaacaagtca ggaggcagta atgtgtacgg cactccaagg actctgtcct gatcggagca
 60301 gaagtgatga agcattgagt agtttgagaa agtttagctaa gaagggtgag ggcagatggt
 60361 ggagagagtt gaggatcaga gagaagacat tagatttgag cagatagaac aaaaatgcc
 60421 ttgccagttt ttgtgtaaga gaatagtatt aggggtggtta cccaggaagt ggtcatcaga
 60481 gttgaatgga gcagaaagag tgtctgatac agcatcggga cgttgtgggc acaaaatgag
 60541 gtgggtgaggg ctgagccaag atgggtggcag tgggatggat aaaaagggat ggccaggaca
 60601 aatatttttaa aggaaaaatt aacaggacat ctttactgac tggattggaa ggctatgcag
 60661 gaaatatatt gtcaaaacttg attccaggat ttctatccta tgcctgggtt gcccataa
 10 60721 tcagggaacc attgttagaa aaggtaggag ataccactgt tccaacaaa agtattgagt
 60781 ttgggtgttc acccacttaa cttcaaggcc ttacaagtga gtagacagtt agttagaatt
 60841 gcagaagtgc cactcagaga gcagggttg caatgtgggg ttggactttg tcaccattgt
 60901 gttaattcct aattctatgc agatgctcag cttgaggaat acccatgttt gggcttcaga
 60961 atgaaagcca agtaatatat actcagatgc caattttccc tctgaaatat ttgctcatgg
 15 61021 aactgagaga acaatatata aagcattaat tatttttctc ataaagtta taataaaaag
 61081 ataagatcag tgaaaggcag agtaaaactag aagccaagta tagaaaatgg tatcattcaa
 61141 agactcatta ctgtagtggt gaaaacaaaa caattttcca acagcttaag atgcctcagt
 61201 attttggacc atttttaagt agttagtgtg ggcacttagt aaatatgtat taaactatag
 61261 ttcattaatt cttttttttt tttttttgag atggagtttc actctgtgta cccaggctgc
 20 61321 atttttgctt tcttagtgat atataaaatg tccaggtttca caatgatggt atcttagatt
 61381 tgattaaata tggattataa aaatagctga tcacagaaag tctctaccag tgtgatgtag
 61441 atggctaaaag tattccacat ttgcaaaact ttattgacct aaataagagg tgccccttgg
 61501 gttgttttta tttggactgg gaatattagg agaaaagctt ttcattcagt gtgtaagtac
 61561 aatctaccag aaatagaaac cccatgggac gatctatttc tttgatggta caggactcag
 25 61621 aacattcaca aagatttagt tgtagcggga atagacatct gtattttatt caaaccaatt
 61681 ttcccttcct aatctgagaa cattgtgcaa tctaagcagt tctaagcatg tttgctattc
 61741 gtgcaaaagt agagtaaaatc taaaagaaat ttttttgtgt gtttagggat ggtaataaag
 61801 tctcttagtg gttgaaaatg ttattttcta caaaagtggga gaacatttgc ttttcaatac
 61861 cagagttttc agccatttct gcattctgac ctattgactg gaggtagggt gcctttgaat
 30 61921 tcagtaaaac ttcattgggca gaaacacagt tcttttctcct acttatttg atatcatgat
 61981 ggccattgca tgtatgtgtc tttttgtgta tccatgcctc agaactgaga agtaggaata
 62041 aaattagggt cagggtctgg gatgctactc tttgctgctg agaaacacaa tgcttcaggt
 62101 aagtgattct gaagtccttc accacctgac ggtaaccttg ggttggtcca taggtatggt
 62161 ttcattttgc ttgttcatcc attttaattg gcttcctaga gcatgcttgt agatgtagag
 35 62221 ccaaatttag agtagagcaa ccctctggca aacaggaaga gattaatttt gtggtatgct
 62281 tttaagggtgac ttcccaggaa acttcaaaag cagaaaaaga agcactagct gcctattcca
 62341 aaatgtgtaa aacaccactc agctttttta aagtaggata aactcagagc gcgcgcacac
 62401 gcgcgcgcgc acacacacac acacacacag agagaacatc tctagtaaaa agaaaagttg
 62461 agctttctta gctagatgtg tgtattagcc agaaaaagcc aaggagtga gggttttaga
 40 62521 gaactggagg agataaaagt gagtctgcat atgggaggca tttgaaatgg acttaaatgt
 62581 ctttttaatg ctgactttt cagttttctc cttaccagac acattgtttt catgacatta
 62641 gccccaggca tagacacatc attaaaatga acatgtcaaa aaatgatttc tgtttagaaa
 62701 taagcaaaac attttcagtt gtgaccaccc aggtgtagaa taaagaacag tgggaattggg
 62761 agccctgagt tctaacataa actttcttca tgacataagg caagtcttct atggcctttg
 45 62821 gtttccttac ctgtaaaaca ggatggctca atgaaattat ctttcttctt tgctataata
 62881 gagtatctct gtgggaagag gaaaaaaaga gtcaatttaa aggtcctta tagttcccca
 62941 actgctgttt tattgtgcta ttcattgccta gacatcacat agctagaaag gcccatcaga
 63001 cccctcaggc cactgctgtt cctgtcacac attcctgcaa aggaccatgt tgctaacttg
 63061 aaaaaaatta ctattaatta cacttgacgt tgttgcttag taacatttat gattttgtgt
 50 63121 ttctcgtgac agcatgagca gagatcatta aaaattaaac ttacaaagct gctaaaagtgg
 63181 gaagaaggag aacttgaagc cacaattttt gcacttgctt agaagccatc taatctcagg
 63241 tttatatgct agatcttggg ggaaacactg catgtctctg gtttatatta aaccacatac
 63301 agcacactac tgacactgat ttgtgtctgg tgcagctgga gtttatcacc aagacataaa
 63361 aaaaccttga ccctgcagaa tggcctggaa ttacaatcag atgggccaca tggcatcccg
 63421 gtgaaagaaa gccctaacca gttttctgtc ttgtttctgc tttctcccta cagttccacc
 55 63481 aggtgagaag agtgatgacc atccttttcc ttactatggt tatttcatac tttggttgca
 63541 tgaaggctgc ccccatgaaa gaagcaaaac tccgaggaca aggtggcttg gcctaccag
 63601 gtgtgaggac ccatgggact ctggagagcg tgaatgggc caaggcaggt tcaaggagct

	63661	tgacatcatt	ggctgacact	ttcgaacacg	tgatagaaga	gctgttggat	gaggaccaga
	63721	aagttcggcc	caatgaagaa	aacaataagg	acgcagactt	gtacacgtcc	agggatgatgc
	63781	tcagtagtca	agtgcctttg	gagcctcctc	ttctctttct	gctggaggaa	tacaaaaatt
	63841	acctagatgc	tgcaaacatg	tccatgaggg	tccggcgcca	ctctgaccct	gcccgcgag
5	63901	gggagctgag	cgtgtgtgac	agtattagtg	agtgggtaac	ggcggcagac	aaaaagactg
	63961	cagtggacat	gtcggggcggg	acggtcacag	tccttgaaaa	ggccctgtga	tcaaaaggcc
	64021	aactgaagca	atacttctac	gagaccaagt	gcaatcccat	gggttacaca	aaagaaggct
	64081	gcaggggcat	agacaaaagg	cattggaact	cccagtgccg	aactaccag	tcgtacgtgc
	64141	gggcccttac	catggatagc	aaaaagagaa	ttggctggcg	attcataagg	atagacactt
10	64201	cttgtgtatg	tacattgacc	attaaaaagg	gaagatagtg	gatttatgtt	gtatagatta
	64261	gattatattg	agacaaaaat	tatctatttg	tatatataca	taacagggtg	aattattcag
	64321	ttaagaaaaa	aataatttta	tgaactgcat	gtataaatga	agtttatata	gtacagtggg
	64381	tctacaatct	atattattgga	catgtccatg	accagaaggg	aaacagtcac	ttgcgcacaa
	64441	cttaaaaaagt	ctgcattaca	ttccttgata	atgttgtgtg	ttgttgccgt	tgccaagaac
15	64501	tgaaaaacata	aaaaggttaa	aaaaataata	aattgcatgc	tgctttaatt	gtgaattgat
	64561	aataaaactgt	cctctttcag	aaaacagaaa	aaaacacaca	cacacacaac	aaaaatttga
	64621	accaaaacat	tccgtttaca	tttttagacag	taagtatctt	cgttcttgtt	agtactatat
	64681	ctgttttact	gcttttaact	tctgatagcg	ttggaattaa	aacaatgtca	aggtgctgtt
	64741	gtcattgctt	tactggctta	ggggatgggg	gatggggggg	atatttttgt	ttgttttgtg
20	64801	tttttttttc	gtttgtttgt	ttgttttttt	agttcccaca	gggagtagag	atggggaaag
	64861	aattcctaca	atataatatt	tggctgataa	aagatacatt	tgtatgttgt	gaagatgttt
	64921	gcaatatcga	tcagatgact	agaaagtga	taaaaaattaa	ggcaactgaa	caaaaaaatg
	64981	ctcacactcc	acatcccgtg	atgcacctcc	caggccccgc	tcattctttg	ggcgttggtc
	65041	agagtaagct	gcttttgacg	gaaggacctg	tgtttgctca	gaacacattc	tttccccccc
25	65101	tccccctctg	gtctcctctt	tgttttgttt	taaggaagaa	aatcagttg	cggttcttga
	65161	aatattttac	cactgctgtg	aacaagtga	cacattgtgt	cacatcatga	cactcgtata
	65221	agcatggaga	acagtgattt	tttttttagaa	cagaaaacaa	caaaaaataa	ccccaaaatg
	65281	aagattattt	tttatgagga	gtgaacattt	gggtaaatca	tggctaagct	taaaaaaaac
	65341	tcattggtgag	gcttaacaat	gtcttgtaag	caaaaagtag	agccctgtat	caaccagaa
30	65401	acacctagat	cagaacagga	atccacattg	ccagtgacat	gagactgaac	agccaaatgg
	65461	aggctatgtg	gagttggcat	tgcattttacc	ggcagtgccg	gaggaatttc	tgagtggcca
	65521	tcccaaggct	taggtggagg	tggggcatgg	tatttgagac	attccaaaac	gaaggcctct
	65581	gaaggaccct	tcagaggtgg	ctctggaatg	acatgtgtca	agctgcttgg	acctcgtgct
	65641	ttaagtgcct	acattatcta	actgtgctca	agaggttctc	gactggagga	ccacactcaa
35	65701	gccgacttat	gcccaccatc	ccacctctgg	ataattttgc	ataaaattgg	attagcctgg
	65761	agcaggttgg	gagccaaatg	tggcatttgt	gatcatgaga	ttgatgcaat	gagatagaag
	65821	atgtttgcta	cctgaacact	tattgctttg	aaactagact	tgaggaaacc	aggttttacg
	65881	ttttgagaac	ttttggttaag	ggaaaaggga	acaggaaaag	aaaccccaaa	ctcaggccga
	65941	atgatcaagg	ggaccatag	gaaatcttgt	ccagagacaa	gacttcggga	aggtgtctgg
40	66001	acattcagaa	caccaagact	tgaagggtgcc	ttgctcaatg	gaagaggcca	ggacagagct
	66061	gacaaaattt	tgctccccag	tgaaggccac	agcaaccttc	tgcccatcct	gtctgttcat
	66121	ggagagggtc	cctgcctcac	ctctgccatt	ttgggttagg	agaagtcaag	ttgggagcct
	66181	gaaatagtgg	ttcttggaag	aatggatccc	cagtgaagaa	tagagctcta	agcccattca
	66241	gccattttca	cacctgaaaa	tgtagtgatg	caccacttgg	accagcatcc	ttaagtatca
45	66301	gaaagcccca	agcaattgct	gcattcttagt	agggtgaggg	ataagcaaaa	gaggatgttc
	66361	accataaccc	aggaatgaag	ataccatcag	caaagaattt	caatttgttc	agtctttcat
	66421	ttagagctag	tctttcacag	taccatctga	atacctcttt	gaaagaagga	agactttacg
	66481	tagttagatg	ttgttttgtg	ttgtttgaaa	atattatctt	tgtaattatt	tttaatatgt
	66541	aaggaatgct	tggaatatct	gctatatgtc	aactttatgc	agcttccttt	tgagggacaa
50	66601	attttaaaca	aacaaccccc	catcacaac	ttaaaggatt	gcaagggcca	gatctgttaa
	66661	gtggtttcat	aggagacaca	tccagcaatt	gtgtgggtcag	tggctctttt	acccaataag
	66721	atacatcaca	gtcacatgct	tgatgggtta	tggtgacctg	agattttatt	tgtaaaatc
	66781	tctctctgtt	gtgttcgttc	ttgttctgtt	ttgttttgtt	ttttaaaagc	ttgctgtggg
	66841	ctctttgtgg	cagaagtgtt	tcatgcatgg	cagcaggcct	gttgcttttt	tatggcgatt
55	66901	cccattgaaa	atgtaagtaa	atgtctgtgg	ccttggtctc	tctatggtaa	agatattatt
	66961	caccatgtaa	aacaaaaaac	aatttttatt	gtatttttagt	atatttatat	aattatgtta
	67021	ttgaaaaaaa	ttggcattaa	aacttaaccg	catcagaacc	tattgtaaat	acaagttcta

67081 ttttaagtgtgta ctaattaaca tataatatat gtttttaata tagaattttt aatgttttta
 67141 aatatatttt caaagtacat aaaa

Hair Cell Differentiation-Suppressing Gene

5 The term “hair cell differentiation-suppressing gene” refers to a gene encoding a protein (e.g., a transcription factor) that positively contributes (directly or indirectly) to the suppression of hair cell differentiation from supporting cells in a primate (e.g., a human). Non-limiting examples of hair cell differentiation-suppressing genes include: HES1, HES5, CDKN1B, and SOX2.

10 The term “mutation in a hair cell differentiation-suppressing gene” refers to a modification in a hair cell differentiation-suppressing gene that results in the production of a hair cell differentiation-suppressing protein having one or more of: one or more amino acid substitutions, and one or more amino acid insertions as compared to the wildtype hair cell differentiation-suppressing protein, and/or results in an increase in the expressed level of the
 15 encoded hair cell differentiation-suppressing protein in a primate cell as compared to the expressed level of the encoded hair cell differentiation-suppressing protein in a primate cell not having a mutation. In some embodiments, the mutation can result in the gain (or an increase in the level) of expression of a hair cell differentiation-suppressing mRNA or a hair cell differentiation-suppressing protein, or both the mRNA and protein. In some embodiments, the
 20 mutation can result in the production of an altered hair cell differentiation-suppressing protein having a gain or increase in one or more biological activities (functions) as compared to a wildtype hair cell differentiation-suppressing protein.

In some embodiments, the mutation is an insertion of one or more nucleotides into a hair cell differentiation-suppressing gene. In some embodiments, the mutation is in a regulatory
 25 sequence of the hair cell differentiation-suppressing gene, i.e., a portion of the gene that is not coding sequence. In some embodiments, a mutation in a regulatory sequence may be in a promoter or enhancer region and prevent or reduce the proper transcription of the hair cell differentiation-suppressing gene (e.g., a mutation in a regulatory sequence that increases the transcription of the hair cell differentiation-suppressing gene).

Hes family basic Helix-Loop-Helix (bHLH) transcription factor 1 (HES1)

The HES1 gene encodes hes family bHLH transcription factor 1, and acts as a transcriptional repressor. HES1 binds to the ATOH1 promoter to inhibit transcription in supporting cells and drives lateral inhibition (Abdolazimi et al. (2016) *Development* 143:841-850). Loss of HES1 results in supernumerary inner hair cells in early development. HES1 inhibition after damage induces hair cell regeneration (Du et al. (2018) *Mol. Ther.* 26(5):1313-1326).

The human HES1 gene is located on chromosome 3q29. It contains 4 exons encompassing ~15 kilobases (kb) (NCBI Accession No. NM_005524). The full-length wildtype HES1 protein expressed from the human HES1 gene is 280 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype HES1 protein is or includes the sequence of SEQ ID NO: 36. Non-limiting examples of a nucleic acid encoding a wildtype HES1 protein is or includes SEQ ID NO: 37.

Human Full-length Wildtype HES1 Protein (SEQ ID NO: 36)

MPADIMEKNSSSPVAATPASVNTTPDKPKTASEHRKSSKPIMEKRRRARINESLSQLKTLILDALKKDSRRHSKLEK
ADILEMTVKHLRNLQRAQMTAALSTDPSVLGKYRAGFSECMNEVTRFLSTCEGVNTEVTRLLGHLANCMTQINAMT
YPGQHPALQAPPPPPGPGGPQHAPFAPPPPLVPIPGGAAPPPGGAPCKLGSQAGEAAKVFGGFQVVPAPDGQFAF
LIPNGAFAHSGPVI PVYTSNSGTSVGPNAVSPSSGPSLTADSMWRPWRN

Human Wildtype HES1 cDNA (SEQ ID NO: 37)

atgccagctgatataatggagaaaaattcctcgtccccgggtggctgctaccccagccagtggtcaacacgacaccgga
taaaccaaagacagcatctgagcacagaaagtcataaagcctattatggagaaaagacgaagagcaagaataaatg
aaagtctgagccagctgaaaacactgattttggatgctctgaagaaagatagctcgcggcattccaagctggagaag
gcggaattctggaaatgacagtgaagcacctccggaacctgcagcgggcgagatgacggctgcgctgagcacaga
cccaagtgtgctggggaagtaccgagccggcttcagcgagtgcatgaacgaggtgacccgcttcctgtccacgtgcg
agggcggttaataccgaggtgcgcactcggctgctcgccacactggccaactgcatgaccagatcaatgccatgacc
taccgcgggcagccgcaccccgcttgcaggcgccgcccaccgccccaccgggacccggcgggccccagcacgcgcc
gttcgcgcgcgcgcgcactcgtgcccatccccggggcgcgggcgccccctcccgggcgcgccccctgcaagctgg
gcagccaggctggagaggcggttaaggtgtttggaggcttcagggtgtaccggctcccgatggccagtttgctttc

ctcattcccaacggggccttcgcgcacagcggccctgtcatccccgtctacaccagcaacagcggcacctccgtggg
 ccccaacgcagtgacaccttccagcggccctcgcttacggcggaactccatgtggaggccgtggcggaactga

A non-limiting example of a human wildtype HES1 genomic DNA sequence is SEQ ID

- 5 NO: 38. The exons in SEQ ID NO: 38 are: nucleotide positions 1-347 (exon 1), nucleotide positions 348-443 (exon 2), nucleotide positions 444-531 (exon 3), and nucleotide positions 532-1461 (exon 4).

Human Wildtype HES1 Gene (SEQ ID NO: 38)

10 1 gggatcacac aggatccgga gctggtgctg ataacagcgg aatcccccggt ctacctctct
 61 ccttggtcct ggaacagcgc tactgatcac caagtagcca caaaatataa taaaccctca
 121 gcacttgctc agtagttttg tgaaagtctc aagtaaaaga gacacaaaca aaaaattctt
 181 tttcgtgaag aactccaaaa ataaaattct ctagagataa aaaaaaaaaa aaaaggaaaa
 241 tgccagctga tataatggag aaaaattcct cgtccccggt ggctgctacc ccagccagtg
 15 301 tcaacacgac accggataaa ccaaagacag catctgagca cagaaagtca tcaaagccta
 361 ttatggagaa aagacgaaga gcaagaataa atgaaagtct gagccagctg aaaacactga
 421 ttttggatgc tctgaagaaa gatagctcgc ggcattccaa gctggagaag gcggacattc
 481 tggaaatgac agtgaagcac ctccggaacc tgcagcgggc gcagatgacg gctgcgctga
 541 gcacagaccc aagtgtgctg gggaaagtacc gagccggctt cagcgagtgc atgaacgagg
 20 601 tgaccgcgtt cctgtccacg tgcgagggcg ttaataccga ggtgcgcact cggctgctcg
 661 gccacctggc caactgcatg acccagatca atgccatgac ctaccccggt cagccgcacc
 721 ccgccttgca ggcgcgcga cgcgccccac cgggaccggg cggccccag cagcgcgcgt
 781 tcgcgcgcgc gccgcactc gtgccatcc ccgggggcgc ggcgccccct cccggcggcg
 841 cccctgcaa gctgggcagc caggctggag aggcggctaa ggtgtttgga ggcttccagg
 25 901 tggtaaccggc tcccgatggc cagtttgctt tcctcattcc caacggggcc ttcgcgcaca
 961 gcggccctgt catccccgtc tacaccagca acagcggcac ctccgtgggc cccaacgcag
 1021 tgtcaccttc cagcggcccc tcgcttacgg cgactccat gtggaggccg tggcggaact
 1081 gagggggctc aggccacccc tcctcctaaa ctcccaacc cacctctctt ccctccggac
 1141 tctaaacagg aacttgaata ctgggagaga agaggacttt tttgattaag tggttacttt
 30 1201 gtgttttttt aatttctaag aagttacttt ttgtagagag agctgtatta agtgactgac
 1261 catgcactat atttgtatat attttatatg ttcattattg attgcccctt tgtattataa
 1321 aagctcagat gacatttcgt tttttacacg agatttcttt tttatgtgat gccaaagatg
 1381 tttgaaaatg ctcttaaaat atcttccttt gggaagttt atttgagaaa atataataaa
 1441 agaaaaaagt aaaggctttt aaaaaaaaaa aaaaa

35

Non-limiting examples of siRNA targeting HES1 are described in, e.g., Zhang et al., *World J. Gastroenterol.* 24(29):3260-3272, 2018; Du et al., *Mol. Ther.* 26(5):1313-1326, 2018; Li et al., *Oncol. Lett.* 14(4):3989-3996, 2017; and Du et al., *Hear Res.* 304:91-110, 2013. Non-limiting examples of shRNA targeting HES1 are described in, e.g., Cenciarelli et al., *Oncotarget* 8(11):17873-17886, 2017, and Wang et al., *Oncotarget* 6(34):36713-36730, 2015.

Hes family bHLH transcription factor 5 (HES5)

The HES5 gene encodes hes family bHLH transcription 5, and acts as a transcriptional repressor. HES5 is a Notch-pathway activator, and binds the ATOH1 promoter to inhibit transcription in supporting cells. Loss of HES5 results in supernumerary outer hair cells in early development. HES5 inhibition in adult mouse utricle results in increased regeneration after aminoglycoside damage (Jung et al. (2013) *Mol. Ther.* 21(4):834-841; Abdolazimi et al. (2016) *Development* 143:841-850).

The human HES5 gene is located on chromosome 1p36. It contains 3 exons encompassing ~18 kilobases (kb) (NCBI Accession No. NM_001010926.3). The full-length wildtype HES5 protein expressed from the human HES5 gene is 166 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype HES5 protein is or includes the sequence of SEQ ID NO: 39. Non-limiting examples of a nucleic acid encoding a wildtype HES5 protein is or includes SEQ ID NO: 40.

Human Full-length Wildtype HES5 Protein (SEQ ID NO: 39)

MAPSTVAVELLSPKEKNRLRKPVVEKMRRDRINSSIEQLKLLLEQEFARHQPNSKLEKADILEMAVSYLKHSKAFVA
AAGPKSLHQDYSEGYSWCLQEAVQFLTLHAASDTQMKLLYHFQRPPAAPAPAKEPKAPGAAPPPALSAKATAAAAA
AHQPACGLWRPW

Human Wildtype HES5 cDNA (SEQ ID NO: 40)

atggccccccagcactgtggccgtggagctgctcagccccaaagagaaaaaccgactgcggaagccggtggtggagaa
gatgcgccgcgacccgcacatcaacagcagcatcgagcagctgaagctgctgctggagcaggagttcgcgcgccaccagc
ccaactccaagctggagaaggccgacatcctggagatggctgtcagctacctgaagcacagcaaagccttcgtcgcc
gccgcccggccccaagagcctgcaccaggactacagcgaaggctactcgtggtgcctgcaggaggccgctgcagttcct

gacgctccacgcccagccagcgacacgcagatgaagctgctgtaccacttccagcgcccccgccgccccgcccgcgc
ccgccaaggagcccaaggcgccggcgccgccccgccccgcgctctccgccaaggccaccgcccgcgcccgcgcgc
gcgcaccagccccgctgcgccctctggcgccctgggtga

- 5 A non-limiting example of a human wildtype HES5 genomic DNA sequence is SEQ ID NO: 41. The exons in SEQ ID NO: 41 are: nucleotide positions 1-135 (exon 1), nucleotide positions 136-301 (exon 2), and nucleotide positions 302-1306 (exon 3).

Human Wildtype HES5 Gene (SEQ ID NO: 41)

10 1 cgcgcttggc cttgcccgcg cccgctcgcc tcgtctcgcc cggcctcccc gcgtcgctc
61 gtcgcctgtt ccgcgccagg catggccccc agcactgtgg ccgtggagct gtcagcccc
121 aaagagaaaa accgactgcg gaagccggtg gtggagaaga tgcgcccga ccgcatcaac
181 agcagcatcg agcagctgaa gctgctgctg gagcaggagt tcgcgcgga ccagcccaac
241 tccaagctgg agaaggccga catcctggag atggctgtca gctacctgaa gcacagcaaa
15 301 gccttcgctg ccgcccgcg cccaagagc ctgcaccagg actacagcga aggctactcg
361 tgggtgctgc aggaggccgt gcagttcctg acgctccacg ccgccagcga cacgcagatg
421 aagctgctgt accacttcca gcggcccccg gccgcgcccg ccgcgcccgc caaggagccc
481 aaggcgccgg gcgcccgcgc cccgcccgcg ctctccgcca aggccaccgc cgccgcccgc
541 gccgcgcacc agcccgcctg cggcctctgg cggccctggt gaccggcgcg gacctgcggg
20 601 cgcgcgggccc gacgaccaga gggcgagcct gtcctctcg cctgtaggga agcgcttcc
661 cgccgtcgtc cgccccgggc ttggacgcgc cttctccgg aaggctctgg cccaagctg
721 gccggcccgc aggagcccca ttctcagaga atgtgtgtgc agagtccctg ccgttttagg
781 acaatcaggg cccatcttct gccaaagtgc tgaccccatg gggttgttct gtgtttgcat
841 ttaagcaagt gacttctggg aagtccccg ccgcccgggg ttctatgata tttgtagtgc
25 901 cggggctcgc aactgctgc cccagcctg tagaggactt tcttcagggc ccgtagctgc
961 tgggcgtacc cctggcaggc gggctgtgcc gcgggcacat ttgccttttg tgaaggccga
1021 actcgagctg taccctcata ggaaacagt atcaccccg acgggcgtcc aggaccctga
1081 gggccatggc caaaaggctc ctgagtgtgc ctggtggtct ggctggggct cacggtgggc
1141 tgtctgggga ggggtgggtg ctccactatg atccttaaag gattcctctg tgtgggtgga
30 1201 tgcgtgtggg cagactttg tactcagaaa ttgaactctc agtcacgtgg aagccacggg
1261 actgctccga agccgccata ataaaatctg attgttcagc ccccaaaaaa aaaaaaaaaa

- Non-limiting examples of siRNA targeting HES5 are described in, e.g., Gu et al., *Oncol. Rep.* 37(1):474-482, 2017; Zhu et al., *Exp. Mol. Pathol.* 99(3):474-484, 2015; Du et al., *Hear Res.* 304:91-110, 2013; Jung et al., *Mol. Ther.* 21(4):834-841, 2013; and Liu et al., *Int. J. Gynecol. Cancer* 20(7):1109-1116, 2010. Non-limiting examples of shRNA targeting HES5 are

described in, e.g., Lee et al., *J. Neurochem.* 100(6):1531-1542, 2007; and Osario et al., *Development* 140:1-12, 2013.

Cyclin dependent kinase inhibitor 1B (Cdkn1b) (p27^{kip1})

5 The CDKN1B gene encodes a cyclin-dependent kinase inhibitor (p27^{kip1}). CDKN1B is a cell cycle regulator and controls the cell cycle exit of supporting cells. For example, p27^{kip1} binds to and prevents activation of cyclin E (CDK2) and cyclin D (CDK4). Inhibition of CDKN1B promotes supporting cell proliferation and regeneration induction through its canonical pathway and a non-canonical pathway that involves Gata3 (Minoda et al. (2007) *Hear Res.* 232(1-2):44-51; Walters et al. (2014) *J. Neurosci* 34(47):15751-15763; Walters et al. (2017) 10 *Cell Rep* 19(2):307-320).

The human CDKN1B gene is located on chromosome 12p13. It contains 3 exons encompassing ~5 kilobases (kb) (NCBI Accession No. NG_016341.1). The full-length wildtype CDKN1B protein expressed from the human CDKN1B gene is 198 amino acids in length.

15 Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype CDKN1B (p27^{kip1}) protein is or includes the sequence of SEQ ID NO: 42. Non-limiting examples of a nucleic acid encoding a wildtype CDKN1B (p27) 20 protein is or includes SEQ ID NO: 43.

Human Full-length Wildtype CDKN1B (p27^{kip1}) Protein (SEQ ID NO: 42)

MSNVRVSNNGSPSLERMDARQAEHPKPSACRNLFPGVDHEELTRDLEKHCRDMEEASQRKWNFDFQNHKPLEGKYEWQ
EVEKGSLEPEFYRPPRPPKGACKVPAQESQDVSGSRPAAPLIGAPANSEDTLHVDPKTDPDSQTGLAEQCAGIRKR
25 PATDDSSSTQNKRANRTEENVSDGSPNAGSVEQTPKKPGLRRRQT

Human Wildtype CDKN1B (p27^{kip1}) cDNA (SEQ ID NO: 43)

atgtcaaacgtgctgagtgcttaacgggagccctagcctggagcggatggacgccaggcaggcggagcaccccaagcc
ctcggcctgcaggaacctcttcggcccggtggaccacgaagagttaaccgggacttggagaagcactgcagagaca
30 tgggaagaggcgagccagcgcaagtggaaatttcgattttcagaatcacaaccctagagggaagtacgagtggtgcaa
gaggtggagaagggcagcttgcccagatttactacagacccccgcggccccccaaaggtgcctgcaaggtgccggc
gcaggagagccaggatgtcagcgggagccgccccggcgcccttaattggggctccggctaactctgaggacacgc
atttggtggaccacaagactgatccgtcggacagccagacggggttagcggagcaatgcgcaggaataaggaagcga

cctgcaaccgacgattcttctactcaaaacaaaagagccaacagaaacagaagaaaatgtttcagacggttccccaaa
tgccggttctgtggagcagacgccaagaagcctggcctcagaagacgtcaaacgtaa

A non-limiting example of a human wildtype CDKN1B (p27^{kip1}) genomic DNA sequence
5 is SEQ ID NO: 44. The exons in SEQ ID NO: 44 are: nucleotide positions 1-1045 (exon 1),
nucleotide positions 1556-1685 (exon 2), and nucleotide positions 3767-5114 (exon 3). The
introns in SEQ ID NO: 44 are: nucleotide positions 1046-1555 (intron 1) and nucleotide
positions 1686-3766 (intron 2).

10 Human Wildtype CDKN1B (p27^{kip1}) Gene (SEQ ID NO: 44)

1 ttaaggccgc gctcgccagc ctcggcgggg cggctccgc cgcgcaacc aatggatctc 61
ctcctctgtt taaatagact cgccgtgtca atcattttct tcttcgtcag cctcccttcc 121
accgccatat tgggccacta aaaaaagggg gctcgtcttt tcggggtgtt tttctcccc 181
tcccctgtcc ccgcttgctc acggctctgc gactccgacg ccggcaagggt ttggagagcg 241
15 gctgggttcg cgggaccgcg gggcttgac ccgccagac tcggacgggc tttgccaccc 301
tctccgcttg cctggtcccc tctcctctcc gccctccgc tcgccagtcc atttgatcag 361
cggagactcg gcggccgggc cggggcttcc ccgcagcccc tcgcgcgtcc tagagctcgg 421
gccgtggctc gtccgggtct gtgtcttttg gctccgaggg cagtcgctgg gcttccgaga 481
ggggttcggg ctgcgtaggg gcgctttgtt ttgttcgggt ttgttttttt gagagtgcga 541
20 gagaggcggg cgtgcagacc cgggagaaa atgtcaaacg tgcgagtgtc taacgggagc 601
cctagcctgg agcggatgga cgccaggcag gcggagcacc ccaagccctc ggctgcagg 661
aacctcttcg gcccggtgga ccacgaagag ttaaccggg acttgagaa gcactgcaga 721
gacatggaag aggcgagcca gcgcaagtgg aatttcgatt ttcagaatca caaaccccta 781
gagggcaagt acgagtggca agagggtggg aagggcagct tgcccagatt ctactacaga 841
25 ccccgcggc ccccaaagg tgctgcaag gtgccggcg aggagagcca ggatgtcagc 901
gggagccgcc cggcggcgcc ttaattggg gctccggcta actctgagga cacgcatttg 961
gtggaccaa agactgatcc gtcggacagc cagacggggt tagcggagca atgcgcagga 1021
ataaggaagc gacctgcaac cgacggtaat gacccttcc caaccataga atgtgttttg 1081
ggccccgctt tgctgtctgg aggggtgtta ccttagcttg cttttcggcg tattctgatt 1141
30 tagctttggg agagctaact ttattggtct taggtgttca gtgctacctg gccactgct 1201
tgtctgtttg tgacttttaa gtcagaaact ggagatggta agatccgata atttccctaa 1261
cttaatacat cgcggtccct ctactagca actcctaggt atgtgacaaa gttgggatgt 1321
ttatcaacgg tccgcctcct ggctagggaa agagctctgg ggcgagaat gcactttctg 1381
ttttttgaaa acaacctcat tttgtgccct taaaagccac tggggatgac ggatccagga 1441
35 ttgtgggttg aggtagtggg tttttcatcc cctgactatg gggccaactt ctgccagcca 1501
ttgttttttc taataaagat tgtgtgttct ttttaaaat tcccctgcg cttagattct 1561

tctactcaaa acaaaaagagc caacagaaca gaagaaaatg tttcagacgg tttcccaaat 1621
 gccggttctg tggagcagac gccaagaag cctggcctca gaagacgtca aacgtaaaca 1681
 gctcggtggg ttgatcacta aaggagcacg cactggaacc cggggccttc agacctcacg 1741
 atacctgatc ttactggttg ctggcaaatt aaaagcttat ggggttttgt tttgtttata 1801
 5 cttcgtgagg tcaaaaaagt agcaatgggg aaggctgggg atacggtaat tcctcagagt 1861
 ttctatgcc agagatactt tctcttcaaa ctgttgacca gagcagctac ttgtaacca 1921
 ggccccatcg ggtaggaagg tcgtttccct gtgagtccca ctaaaacgtg ttgggagcaa 1981
 taggttcttt gcccatccga acaagaacta gggactccc tcagtccgaa ttaatgagaa 2041
 ttaatttcct agaggttcag cttgagtcgg taacagattt tgagccatac atggaaaaat 2101
 10 ggcaaataca tgattaagtt tcaattttga gggggaatgt ttggtagaaa ttgctcatct 2161
 ttggttatgc aagggattag agatgtgaat aggatggat gttgtgttct ttgacatttt 2221
 aataaactgt cactttccct gttgtctcct aagtttgag agagaaggaa ccagtatttg 2281
 caaaaaccaa atggaaagat aaaaaagtta ctaaagtctc tacagaattt ctggtaacac 2341
 tgaagttgca aagcagaagt taaattaact cttgtcagta agcaatccag gaacacgtca 2401
 15 gccagtgtat gctaattgtg ccgtaacagg gtgatttgga tattttagag ggaaatgggt 2461
 agtaaatatc aagactggtg accgtaggtc agcccagcac aaagggaagt gagatttttc 2521
 catgcacaag aatctgatca ctgtaaatac taatttcagta cccagataac 2581
 caacatgggt tggttattca taataaacta catattttaa tagtttatta gcttccttta 2641
 gaccaagact gtgacctctt tattttctaa agcacacacg tagtttagca tatgaggcga 2701
 20 taaaatattg atgttaactt tttaaatccc cagttataaa aattttaaaa taacagggat 2761
 taaggtgaga ttcaggtttg ttgtgtcttt aaattgtata tgtgacttca catatctttt 2821
 tcagcgctta tacaaaacgg cactatagaa cctccatttt acagcaccat atgaagtggg 2881
 aaaattaggt gaaaattttc ctgaagcaac cttacatgc gcagcccttg ttggtttgtg 2941
 acttggtggc tagctcatca gatgagccac gagaatcaga cctggatttt gatctggccc 3001
 25 tgttctgaca tgcaatgagg catttgtagg atttagtaat attgctagtt caaagaatac 3061
 tagaaatatt agtaagaacc tattcaaaag tattcatgag tattttctgc atatgaatca 3121
 ggaattagaa tattttgaaa atgatgttaa taaaattttc ctctggaagg cttttataat 3181
 ttttattccc aatcattttt caaatttaga aagtttaatc tgtcacagga gaaaaaaat 3241
 taaaaatttt caaaaattta gaaaattttt acccgtaagt attacagttt cctaattatc 3301
 30 ctattttattt ccacttgcc tttgacctag attatttaga gtagggtttc tcagcctctg 3361
 cactaatgac attttgggcc gaataattct gttgtaggag gctgtcctgt gtgttttaga 3421
 ttgtttgaa ttatccctgg cctctcacac tagatgccag cagtatctc ctccccagt 3481
 gtgacaacct aaaatgtctc cagacattgc caaatgttcc ctgtggggcg ggggcaacat 3541
 tgcctactgt taagaactac tgctctagac caaagaacac agcacagagg aaaggaaaaa 3601
 35 aaaatccagt taagagaatg ttaggtggag atgactatag tcatcaaact tttttcccca 3661
 tcaagtattt ccaagctaac atagtgacaa aataattcct gtactctact ggtaacgtta 3721
 atctagtgtt cttcctttta ttcttcggtt ttgttttctt ttgcagaatt aagaatatgt 3781
 ttccttggtt atcagataca tcaactgctg atgaagcaag gaagatatac atgaaaattt 3841

taaaaataca tatcgctgac ttcattggaat ggacatcctg tataagcact gaaaaacaac 3901
 aacacaataa cactaaaatt ttaggcactc ttaaatgac tgcctctaaa agcgttggat 3961
 gtagcattat gcaattaggt ttttccttat ttgcttcatt gtactacctg tgtatatagt 4021
 ttttaccttt tatgtagcac ataaactttg gggaaggag ggaggggtgg ggctgaggaa 4081
 5 ctgacgtgga gcgggggatg aagagcttgc tttgatttac agcaagtaga taaatatttg 4141
 acttgcatga agagaagcaa ttttggggaa gggtttgaat tgttttcttt aaagatgtaa 4201
 tgtccctttc agagacagct gatacttcat ttaaaaaaat cacaaaaatt tgaacactgg 4261
 ctaaagataa ttgctattta tttttacaag aagtttattc tcatttggga gatctggtga 4321
 tctccaagc tatctaaagt ttgtagata gctgcatgtg gcttttttaa aaaagcaaca 4381
 10 gaaacctatc ctactgccc tccccagtct ctcttaaagt tggaatttac cagttaatta 4441
 ctgagcagaa tggtagcac tccaggtagt ttggggcaaa aatccgaggt gcttgggagt 4501
 tttgaatgtt aagaattgac catctgcttt tattaatttt gttgacaaaa ttttctcatt 4561
 ttcttttcac ttcgggctgt gtaaacacag tcaaaataat tctaaatccc tcgatatttt 4621
 taaagatctg taagtaactt cacattaaaa aatgaaatat tttttaattt aaagcttact 4681
 15 ctgtccattt atccacagga aagtgttatt tttcaaggaa gggtcatgta gagaaaagca 4741
 cactttagg ataagtgaat tggatactac atctttaaac agtatttcat tgcctgtgta 4801
 tggaaaaacc atttgaagtg tacctgtgta cataactctg taaaaacact gaaaaattat 4861
 actaacttat ttatgttaaa agattttttt taatctagac aatatacaag ccaaagtggc 4921
 atgttttgtg catttgtaaa tgctgtgttg ggtagaatag gttttcccct cttttgttaa 4981
 20 ataatatggc tatgcttaaa aggttgcata ctgagccaag tataattttt tgtaatgtgt 5041
 gaaaaagatg ccaattattg ttacacatta agtaatcaat aaagaaaact tccatagcta 5101
 ttcattgagt caaa

Non-limiting examples of siRNA targeting CDKN1B (p27^{kip1}) are described in, e.g.,
 25 Galardi et al., *J. Biol. Chem.* 282:23716-23724, 2007; Liang et al., *Nature Cell Biol.* 9:218-224,
 2007; Tamamori-Adachi et al., *J. Biol. Chem.* 279:50429-50436, 2004; Akashiba et al., *Cell.*
Mol. Life Sci. 63:2397-2404, 2006; and Lee et al., *J. Mol. Med.* 83(4):296-307, 2005. Non-
 limiting examples of shRNA targeting CDKN1B (p27^{kip1}) are described in, e.g., Lin et al.,
Nature 464:374-379, 2010.

30

Sex determining region Y-box 2 (SOX2)

The SOX2 gene encodes the sex determining region Y- box 2 protein. SOX2 is a
 transcription factor that binds the ATOH1 3'- enhancer and activates initial hair cell
 differentiation. Low SOX2 expression levels are required for proper hair cell maturation.
 35 Haploinsufficiency of SOX2 results in a few extra inner hair cells. SOX2 also increases the

susceptibility to induce transdifferentiation in the presence of other contributing components, e.g., beta-catenin (Kempfle et al. (2016) *Sci Rep* 6:23293; Atkinson et al. (2018) *J Clin Invest* 128(4):1641-1656).

The human SOX2 gene is located on chromosome 3q26. It contains 1 exon encompassing ~3 kilobases (kb) (NCBI Accession No. NG_009080.1). The full-length wildtype SOX2 protein expressed from the human SOX2 gene is 317 amino acids in length.

Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype SOX2 protein is or includes the sequence of SEQ ID NO: 45. Non-limiting examples of a nucleic acid encoding a wildtype SOX2 protein is or includes SEQ ID NO: 46. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 46 can be codon-optimized to allow for optimal expression in a non-human primate.

Human Full-length Wildtype SOX2 Protein (SEQ ID NO: 45)

MYNMMETELKPPGPQQTSGGGGGNSTAAAAGGNQKNSPDRVKRPMNAFMVWSRGQRRKMAQENPKMHNSEISKRLGA
EWKLLSETEKRPFIDEAKRLRALHMKEHPDYKYRPRRKTKTLMKKDKYTLPGGLLAPGGNSMASGVGVGAGLGAGVN
QRMDSYAHMNGWSNGSYSMQDQLGYPQHPLNAHGAAQMMPHRYDVSALQYNSMTSSQTYMNGSPTYSMSYSQQG
TPGMALGSMGVSVKSEASSPPVVTSSSHSRAPCQAGDLRDMISMYLPGAIEVPEPAAPSRHMSQHYQSGPVPGTAI
NGTLPLSHM

Human Wildtype SOX2 cDNA (SEQ ID NO: 46)

atgtacaacatgatggagacggagctgaagccgcccggcccgagcaaaacttcggggggcgggcgggcggaactccac
cgcgggcgggcgccggcggaaccagaaaaacagcccggaccgcgtcaagcgggcccatgaatgccttcattggtgtggt
cccgcgggcagcggcgcaagatggcccaggagaaccccaagatgcacaactcggagatcagcaagcgctgggcgcc
gagtggaaaacttttgcggagacggagaagcgccgcttcattcgacgaggctaagcggtgcgagcgctgcacatgaa
ggagcaccggtattataaataaccggccccggcggaacccaagacgctcatgaagaaggataagtacacgctgccccg
gcgggctgctggcccccgggcaatagcatggcgagcggggtcggggtggcgccggcctgggcgcgggcggtgaac
cagcgcatggacagttacgcgcacatgaacggctggagcaacggcagctacagcatgatgcaggaccagctgggcta
cccgcagcaccgggctcaatgcgcacggcgagcgagatgcagcccatgcaccgctacgacgtgagcgccctgc
agtacaactccatgaccagctcgacagcctacatgaacggctcgccacctacagcatgtcctactcgacgagggc
acccctggcatggctcttggctccatgggttcgggtggtcaagtccgaggccagctccagccccctgtggttacctc
ttcctcccactccaggcgccctgccaggccggggacctccgggacatgatcagcatgtatctccccggcgccgagg

tgccggaacccgccgccccagcagacttcacatgtcccagcactaccagagcggcccggtgcccggcacggccatt
aacggcacactgcccctctcacacatgtga

A non-limiting example of a human wildtype SOX2 genomic DNA sequence is SEQ ID

5 NO: 47. The exon in SEQ ID NO: 47 is nucleotide positions 1-2520 (exon 1).

Human Wildtype SOX2 Gene (SEQ ID NO: 47)

1 ggatggttgt ctattaactt gttcaaaaaa gtatcaggag ttgtcaaggc agagaagaga
10 61 gtgtttgcaa aagggggaaa gtagtttgct gcctctttaa gactaggact gagagaaaga
121 agaggagaga gaaagaaagg gagagaagtt tgagccccag gcttaagcct ttccaaaaaa
181 taataataac aatcatcggc ggccggcagga tcggccagag gaggaggga gcgctttttt
241 tgatcctgat tccagtttgc ctctctcttt tttccccca aattattctt cgctgattt
301 tcctcgcgga gccctgcgct cccgacaccc ccgcccgcct cccctcctcc tctcccccg
15 361 ccgcggggcc ccccaaagtc ccggccgggc cgagggtcgg cggccgccgg cgggcccggc
421 ccgcgcacag cgcccgcag tacaacatga tggagacgga gctgaagccg ccgggcccgc
481 agcaaacttc ggggggccc ggccggcaact ccaccgccc ggccggccc ggcaaccaga
541 aaaacagccc ggaccgcgtc aagcggccca tgaatgcctt catggtgtgg tcccgccggc
601 agcggcgcaa gatggcccag gagaacccca agatgcacaa ctccggagatc agcaagcgcc
20 661 tgggcgccga gtggaaactt ttgtcggaga cggagaagcg gccgttcac gacgaggcta
721 agcggctgcg agcgtgcac atgaaggagc acccgatta taaataccgg ccccgccgga
781 aaaccaagac gctcatgaag aaggataagt acacgctgcc cggcgggctg ctggcccccg
841 gcggcaatag catggcgagc ggggtcggg tgggcgccgg cctgggcgcg gggtgaacc
901 agcgcagga cagttacgc cacaatgaac gctggagcaa cggcagctac agcatgatgc
25 961 aggaccagct gggctacccg cagcaccg gcctcaatgc gcacggcgca gcgcagatgc
1021 agcccatgca ccgctacgac gtgagcgccc tgcatgacaa ctccatgacc agctcgcaga
1081 cctacatgaa cggctcggcc acctacagca tgtcctactc gcagcagggc acccctggca
1141 tggctcttgg ctccatgggt tcggtggtca agtccgaggc cagctccagc ccccctgtgg
1201 ttacctcttc ctccactcc agggcgccct gccaggccgg ggacctccgg gacatgatca
30 1261 gcatgtatct ccccgccg ccgggtgccg aaccgcccgc cccagcaga cttcacatgt
1321 ccagcacta ccagagcggc ccgggtcccg gcacggccat taacggcaca ctgcccctct
1381 cacacatgtg agggccggac agcgaactgg aggggggaga aattttcaaa gaaaaacgag
1441 ggaaatggga ggggtgcaaa agaggagagt aagaaacagc atggagaaaa cccggtacgc
1501 tcaaaaagaa aaaggaaaaa aaaaaatccc atcaccaca gcaaatgaca gctgcaaaag
35 1561 agaacaccaa tcccatccac actcacgcaa aaaccgcat gccgacaaga aaacttttat
1621 gagagagatc ctggacttct ttttggggga ctatttttgt acagagaaaa cctggggagg
1681 gtggggaggg cgggggaatg gaccttgtat agatctggag gaaagaaagc tacgaaaaac
1741 tttttaaaag ttctagtgg acggtaggag ctttgcagga agtttgcaaa agtctttacc

1801 aataatat ttt agagctagtc tccaagcgac gaaaaaaatg ttttaatat tgcaagcaac
 1861 ttttgtacag tatttatcga gataaacatg gcaatcaaaa tgtccattgt ttataagctg
 1921 agaatttgcc aatat ttttc aaggagaggc ttcttgctga attttgattc tgcagctgaa
 1981 atttaggaca gttgcaaacg tgaaaagaag aaaattattc aaatttggaac attttaattg
 5 2041 ttttaaaaatt gtacaaaagg aaaaaattag aataagtact ggcgaaccat ctctgtggtc
 2101 ttgttttaaaa agggcaaaag ttttagactg tactaaattt tataacttac tgttaaaagc
 2161 aaaaatggcc atgcagggtg acaccgttg taatttataa tagcttttgc tcgatcccaa
 2221 ctttccattt tggtcagata aaaaaacca tgaaattact gtgtttgaaa tattttctta
 2281 tggtttgtaa tattttctga aatttattgt gatattttta ggttttcccc cttttatttt
 10 2341 ccgtagttgt attttaaaag attcggctct gtattatttg aatcagtctg ccgagaatcc
 2401 atgtatataat ttgaactaat atcatcctta taacagggtac attttcaact taagttttta
 2461 ctccattatg cacagtttga gataaataaa tttttgaaat atggacactg aaa

15 Non-limiting examples of siRNA targeting SOX2 are described in, e.g., Kondo et al.,
Genes Develop. 18:2963-2972, 2004; Tani et al., *J. Cancer Res. Clin. Oncol.* 133(4):263-269,
 2007; Chen et al., *J. Biol. Chem.* 283:17969-17978, 2008; and Card et al., *Mol. Cell. Biol.*
 28(20):6426-6438, 2008. Non-limiting examples of shRNA targeting SOX2 are described in,
 e.g., Rudin et al., *Nature Genetics* 44:1111-1116, 2012; Basu-Roy et al., *Oncogene* 31:2270-
 20 2282, 2012; and Marques-Torrejon et al., *Cell Stem Cell* 12(1):88-100, 2013.

Vectors

Some of the compositions provided herein can include at least two (e.g., two, three, four,
 five, or six) AAV vectors, where: each of the at least two different AAV vectors includes a
 25 coding sequence that encodes a different portion of a hair cell differentiation protein, each of the
 encoded portions being at least 30 amino acids (e.g., about 30 amino acids to about 800 amino
 acids, about 30 amino acids to about 780 amino acids, about 30 amino acids to about 760 amino
 acids, about 30 amino acids to about 750 amino acids, about 30 amino acids to about 740 amino
 acids, about 30 amino acids to about 720 amino acids, about 30 amino acids to about 710 amino
 30 acids, about 30 amino acids to about 700 amino acids, about 30 amino acids to about 690 amino
 acids, about 30 amino acids to about 680 amino acids, about 30 amino acids to about 670 amino
 acids, about 30 amino acids to about 660 amino acids, about 30 amino acids to about 650 amino
 acids, about 30 amino acids to about 640 amino acids, about 30 amino acids to about 630 amino
 acids, about 30 amino acids to about 620 amino acids, about 30 amino acids to about 610 amino

127

128

5
10
15
20
25
30

130

131

[illegible]

133

[illegible]

[illegible]

[illegible]

about 710 amino acids, about 650 amino acids to about 700 amino acids, about 700 amino acids to about 800 amino acids, about 700 amino acids to about 790 amino acids, about 700 amino acids to about 780 amino acids, about 700 amino acids to about 770 amino acids, about 700 amino acids to about 760 amino acids, about 700 amino acids to about 750 amino acids, or about 750 amino acids to about 800 amino acids), where the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes the hair cell differentiation protein (e.g., a full-length hair cell differentiation protein (e.g., a full-length wildtype hair cell differentiation protein)); and, when introduced into a primate cell (e.g., a hair cell or a supporting cell of the inner ear), the at least two different AAV vectors undergo homologous recombination with each other, where the recombined nucleic acid encodes a hair cell differentiation protein (e.g., a full-length hair cell differentiation protein).

In some embodiments of the compositions that include at least two AAV vectors, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of hair cell differentiation genomic DNA, and lacks the intronic sequence that naturally occurs between the two neighboring exons.

In some embodiments of the compositions that include at least two AAV vectors, the amino acid sequence of none of the encoded portions overlaps even in part with the amino acid sequence of a different one of the encoded portions. In some embodiments of the compositions that include at least two AAV vectors, the amino acid sequence of one or more of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments of the compositions that include at least AAV vectors, the amino acid sequence of each of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions.

In some embodiments of the compositions that include at least two AAV vectors, the overlapping amino acid sequence is between about 30 amino acid residues to about 800 amino acids (e.g., or any of the subranges of this range described herein) in length.

In some examples, the compositions include two different AAV vectors, each of which comprises a different segment of an intron, where the intron includes the nucleotide sequence of an intron that is present in a hair cell differentiation genomic DNA, and where the two different segments overlap in sequence by at least 100 nucleotides (e.g., about 100 nucleotides to about

3,000 nucleotides, about 100 nucleotides to about 2,500 nucleotides, about 100 nucleotides to about 2,000 nucleotides, about 100 nucleotides to about 1,500 nucleotides, about 100 nucleotides to about 1,000 nucleotides, about 100 nucleotides to about 800 nucleotides, about 100 nucleotides to about 600 nucleotides, about 100 nucleotides to about 400 nucleotides, about 100 nucleotides to about 200 nucleotides, about 200 nucleotides to about 3,000 nucleotides, about 200 nucleotides to about 2,500 nucleotides, about 200 nucleotides to about 2,000 nucleotides, about 200 nucleotides to about 1,500 nucleotides, about 200 nucleotides to about 1,000 nucleotides, about 200 nucleotides to about 800 nucleotides, about 200 nucleotides to about 600 nucleotides, about 200 nucleotides to about 400 nucleotides about 400 nucleotides to about 3,000 nucleotides, about 400 nucleotides to about 2,500 nucleotides, about 400 nucleotides to about 2,000 nucleotides, about 400 nucleotides to about 1,500 nucleotides, about 400 nucleotides to about 1,000 nucleotides, about 400 nucleotides to about 800 nucleotides, about 400 nucleotides to about 600 nucleotides, about 600 nucleotides to about 3,000 nucleotides, about 600 nucleotides to about 2,500 nucleotides, about 600 nucleotides to about 2,000 nucleotides, about 600 nucleotides to about 1,500 nucleotides, about 600 nucleotides to about 1,000 nucleotides, about 600 nucleotides to about 800 nucleotides, about 800 nucleotides to about 3,000 nucleotides, about 800 nucleotides to about 2,500 nucleotides, about 800 nucleotides to about 2,000 nucleotides, about 800 nucleotides to about 1,500 nucleotides, about 800 nucleotides to about 1,000 nucleotides, about 1,000 nucleotides to about 3,000 nucleotides, about 1,000 nucleotides to about 2,500 nucleotides, about 1,000 nucleotides to about 2,000 nucleotides, about 1,000 nucleotides to about 1,500 nucleotides, about 1,500 nucleotides to about 3,000 nucleotides, about 1,500 nucleotides to about 2,500 nucleotides, about 1,500 nucleotides to about 2,000 nucleotides, about 2,000 nucleotides to about 3,000 nucleotides, about 2,000 nucleotides to about 2,500 nucleotides, or about 2,500 nucleotides to about 3,000 nucleotides), in length.

The overlapping nucleotide sequence in any two of the different AAV vectors can include part or all of one or more exons of a hair cell differentiation gene.

In some embodiments, the number of different AAV vectors in the composition is two, three, four, or five. In compositions where the number of different AAV vectors in the composition is two, the first of the two different vectors can include a coding sequence that encodes an N-terminal portion of the hair cell differentiation protein. In some embodiments, the N-terminal portion can include a portion having about 30 amino acids to about 800 amino acids

(or any of the subranges of this range described herein). In some examples, the N-terminal portion encoded by one of the two vectors can include a portion comprising amino acid position 1 to about amino acid position 800, about amino acid position 790, about amino acid position 780, about amino acid position 770, about amino acid position 760, about amino acid position 750, about amino acid position 740, about amino acid position 730, about amino acid position 720, about amino acid position 710, about amino acid position 700, about amino acid position 690, about amino acid position 680, about amino acid position 670, about amino acid position 660, about amino acid position 650, about amino acid position 640, about amino acid position 630, about amino acid position 620, about amino acid position 610, about amino acid position 600, about amino acid position 590, about amino acid position 580, about amino acid position 570, about amino acid position 560, about amino acid position 550, about amino acid position 540, about amino acid position 530, about amino acid position 520, about amino acid position 510, about amino acid position 500, about amino acid position 490, about amino acid position 480, about amino acid position 470, about amino acid position 460, about amino acid position 450, about amino acid position 440, about amino acid position 430, about amino acid position 420, about amino acid position 410, about amino acid position 400, about amino acid position 390, about amino acid position 380, about amino acid position 370, about amino acid position 360, about amino acid position 350, about amino acid position 340, about amino acid position 330, about amino acid position 320, about amino acid position 310, about amino acid position 300, about amino acid position 290, about amino acid position 280, about amino acid position 270, about amino acid position 260, about amino acid position 250, about amino acid position 240, about amino acid position 230, about amino acid position 220, about amino acid position 210, about amino acid position 200, about amino acid position 190, about amino acid position 180, about amino acid position 170, about amino acid position 160, about amino acid position 150, about amino acid position 140, about amino acid position 130, about amino acid position 120, about amino acid position 110, about amino acid position 100, about amino acid position 90, about amino acid position 80, about amino acid position 70, about amino acid position 60, about amino acid position 50, or about amino acid position 40 of a wildtype hair cell differentiation protein.

In compositions where the number of different AAV vectors in the composition is two, the second of the two different vectors can include a coding sequence that encodes a C-terminal

portion of the hair cell differentiation protein. In some embodiments, the C-terminal portion can include a portion having about 30 amino acids to about 800 amino acids (or any of the subranges of this range described herein).

As used herein, the term “vector” means a composition including a polynucleotide capable of carrying at least one exogenous nucleic acid fragment, e.g., an adeno-associated virus (AAV) vector. A vector can, e.g., include sufficient cis-acting elements for expression; other elements for expression can be supplied by the host primate cell or in an *in vitro* expression system. The term “vector” includes any genetic element (e.g., a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector, etc.) that is capable of replicating when associated with the proper control elements.

“Recombinant AAV vectors” or “rAAVs” are typically composed of, at a minimum, a transgene or a portion thereof and a regulatory sequence, and optionally 5' and 3' AAV inverted terminal repeats (ITRs). Such a recombinant AAV vector is packaged into a capsid and delivered to a selected target cell (e.g., an inner or outer hair cell, or a supporting cell of the inner ear).

The AAV sequences of the vector typically comprise the cis-acting 5' and 3' ITR sequences (See, e.g., B. J. Carter, in "Handbook of Parvoviruses", ed., P. Tijsser, CRC Press, pp. 155-168, 1990). Typical AAV ITR sequences are about 145 nucleotides in length. In some embodiments, at least 75% of a typical ITR sequence (e.g., at least 80%, at least 85%, at least 90%, or at least 95%) is incorporated into the AAV vector. The ability to modify these ITR sequences is within the skill of the art. (See, e.g., texts such as Sambrook et al., "Molecular Cloning. A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory, New York, 1989; and K. Fisher et al., *J Virol.* 70:520-532, 1996). In some embodiments, any of the coding sequences described herein are flanked by 5' and 3' AAV ITR sequences in the AAV vectors. The AAV ITR sequences may be obtained from any known AAV, including presently identified AAV types. In some examples of any of the vectors described herein, the vector includes a 5' ITR sequence

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGGCGAC
CTTTGGTCTGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACT
CCATCACTAGGGGTTTCCT (SEQ ID NO: 51) and/or a 3' ITR sequence
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTG

AGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTG
AGCGAGCGAGCGCGCAGCTGCCTGCAGG (SEQ ID NO: 57).

AAV vectors as described herein may include any of the regulatory elements described herein (e.g., one or more of a promoter, a polyA sequence, and an IRES).

5 In some embodiments, the AAV vector is selected from the group consisting of: an AAV1 vector, an AAV2 vector, an AAV3 vector, an AAV4 vector, an AAV5 vector, an AAV6 vector, an AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV2.7m8 vector, an AAV8BP2 vector, and an AAV293 vector. Additional exemplary AAV vectors that can be used herein are known in the art. See, e.g., Kanaan et al., *Mol. Ther. Nucleic Acids* 8:184-197, 2017; 10 Li et al., *Mol. Ther.* 16(7): 1252-1260; Adachi et al., *Nat. Commun.* 5: 3075, 2014; Isgrig et al., *Nat. Commun.* 10(1): 427, 2019; and Gao et al., *J. Virol.* 78(12): 6381-6388.

In some embodiments, an AAV vector provided herein includes or consists of a sequence that is at least 80% identical (e.g., at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 15 99%, or 100% identical) to SEQ ID NO: 50, 58, 60, 64, 66, 68, 78, 79, 81, 82, 83 or 94.

The AAV vectors provided herein can be of different sizes. In some embodiments, the AAV vector(s) can include a total number of nucleotides of up to 5 kb. In some embodiments, the AAV vector(s) can include a total number of nucleotides in the range of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 2 kb to 20 about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, or about 4 kb to about 5 kb.

In some embodiments of any of the compositions, kits, and methods provided herein, the at least two different AAV vectors can be substantially the same type of vector and may differ in size. In some embodiments, the at least two different AAV vectors can be different types of 25 AAV vector, and may have substantially the same size or have different sizes.

In some embodiments, any of the at least two AAV vectors can have a total number of nucleotides in the range of about 500 nucleotides to about 10,000 nucleotides, about 500 nucleotides to about 9,500 nucleotides, about 500 nucleotides to about 9,000 nucleotides, about 500 nucleotides to about 8,500 nucleotides, about 500 nucleotides to about 8,000 nucleotides, 30 about 500 nucleotides to about 7,800 nucleotides, about 500 nucleotides to about 7,600 nucleotides, about 500 nucleotides to about 7,400 nucleotides, about 500 nucleotides to about

5
10
15
20
25
30

5
10
15
20
25
30

5
10
15
20
25
30

5
10
15
20
25
30

[illegible]

2,400 nucleotides, about 2,000 nucleotides to about 2,200 nucleotides, about 2,200 nucleotides to about 10,000 nucleotides, about 9,500 nucleotides, about 9,000 nucleotides, about 8,500 nucleotides, about 8,000 nucleotides, about 7,800 nucleotides, about 7,600 nucleotides, about 7,400 nucleotides, about 7,200 nucleotides, about 7,000 nucleotides, about 6,800 nucleotides, about 6,600 nucleotides, about 6,400 nucleotides, about 6,200 nucleotides, about 6,000 nucleotides, about 5,800 nucleotides, about 5,600 nucleotides, about 5,400 nucleotides, about 5,200 nucleotides, about 5,000 nucleotides, about 4,800 nucleotides, about 4,600 nucleotides, about 4,400 nucleotides, about 4,200 nucleotides, about 4,000 nucleotides, about 3,800 nucleotides, about 3,600 nucleotides, about 3,400 nucleotides, about 3,200 nucleotides, about 3,000 nucleotides, about 2,800 nucleotides, about 2,600 nucleotides, about 2,400 nucleotides, about 2,400 nucleotides to about 10,000 nucleotides, about 2,400 nucleotides to about 9,500 nucleotides, about 2,400 nucleotides to about 9,000 nucleotides, about 2,400 nucleotides to about 8,500 nucleotides, about 2,400 nucleotides to about 8,000 nucleotides, about 2,400 nucleotides to about 7,800 nucleotides, about 2,400 nucleotides to about 7,600 nucleotides, about 2,400 nucleotides to about 7,400 nucleotides, about 2,400 nucleotides to about 7,200 nucleotides, about 2,400 nucleotides to about 7,000 nucleotides, about 2,400 nucleotides to about 6,800 nucleotides, about 2,400 nucleotides to about 6,600 nucleotides, about 2,400 nucleotides to about 6,400 nucleotides, about 2,400 nucleotides to about 6,200 nucleotides, about 2,400 nucleotides to about 6,000 nucleotides, about 2,400 nucleotides to about 5,800 nucleotides, about 2,400 nucleotides to about 5,600 nucleotides, about 2,400 nucleotides to about 5,400 nucleotides, about 2,400 nucleotides to about 5,200 nucleotides, about 2,400 nucleotides to about 5,000 nucleotides, about 2,400 nucleotides to about 4,800 nucleotides, about 2,400 nucleotides to about 4,600 nucleotides, about 2,400 nucleotides to about 4,400 nucleotides, about 2,400 nucleotides to about 4,200 nucleotides, about 2,400 nucleotides to about 4,000 nucleotides, about 2,400 nucleotides to about 3,800 nucleotides, about 2,400 nucleotides to about 3,600 nucleotides, about 2,400 nucleotides to about 3,400 nucleotides, about 2,400 nucleotides to about 3,200 nucleotides, about 2,400 nucleotides to about 3,000 nucleotides, about 2,400 nucleotides to about 2,800 nucleotides, about 2,400 nucleotides to about 2,600 nucleotides, about 2,600 nucleotides to about 10,000 nucleotides, about 2,600 nucleotides to about 9,500 nucleotides, about 2,600 nucleotides to about 9,000 nucleotides, about 2,600 nucleotides to about 8,500 nucleotides, about 2,600 nucleotides to about 8,000 nucleotides, about 2,600 nucleotides to about 7,800 nucleotides, about 2,600

[illegible]

5
10
15
20
25
30

5
10
15
20
25
30

5
10
15
20
25
30

[illegible]

[illegible]

[illegible]

5
10
15
20
25
30

nucleotides, about 7,200 nucleotides to about 9,500 nucleotides, about 7,200 nucleotides to about 9,000 nucleotides, about 7,200 nucleotides to about 8,500 nucleotides, about 7,200 nucleotides to about 8,000 nucleotides, about 7,200 nucleotides to about 7,800 nucleotides, about 7,200 nucleotides to about 7,600 nucleotides, about 7,200 nucleotides to about 7,400 nucleotides, about 7,400 nucleotides to about 10,000 nucleotides, about 7,400 nucleotides to about 9,500 nucleotides, about 7,400 nucleotides to about 9,000 nucleotides, about 7,400 nucleotides to about 8,500 nucleotides, about 7,400 nucleotides to about 8,000 nucleotides, about 7,400 nucleotides to about 7,800 nucleotides, about 7,400 nucleotides to about 7,600 nucleotides, about 7,600 nucleotides to about 10,000 nucleotides, about 7,600 nucleotides to about 9,500 nucleotides, about 7,600 nucleotides to about 9,000 nucleotides, about 7,600 nucleotides to about 8,500 nucleotides, about 7,600 nucleotides to about 8,000 nucleotides, about 7,600 nucleotides to about 7,800 nucleotides, about 7,800 nucleotides to about 10,000 nucleotides, about 7,800 nucleotides to about 9,500 nucleotides, about 7,800 nucleotides to about 9,000 nucleotides, about 7,800 nucleotides to about 8,500 nucleotides, about 7,800 nucleotides to about 8,000 nucleotides, about 8,000 nucleotides to about 10,000 nucleotides, about 8,000 nucleotides to about 9,500 nucleotides, about 8,000 nucleotides to about 9,000 nucleotides, about 8,000 nucleotides to about 8,500 nucleotides, about 8,500 nucleotides to about 10,000 nucleotides, about 8,500 nucleotides to about 9,500 nucleotides, about 8,500 nucleotides to about 9,000 nucleotides, about 9,000 nucleotides to about 10,000 nucleotides, about 9,000 nucleotides to about 9,500 nucleotides, or about 9,500 nucleotides to about 10,000 nucleotides (inclusive).

Figures 4A-D, Figures 7A-B and Figures 11A-B provide schematic representations of exemplary nucleic acid vectors that can be included in any of the compositions and methods described herein.

In some embodiments of any of the compositions described herein, the vector comprises or consists of pITR-CMV-mScarlet (SEQ ID NO: 50). In some embodiments of any of the compositions described herein, the vector comprises a sequence that has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity to SEQ ID NO: 50.

pITR-CMV-mScarlet (SEQ ID NO: 50)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggtcgcccggcctcagt
 gagcgagcgagcgcgagagagggagtgcccaactccatcactaggggttcctctagatcccatatatggagttccg
 cgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgcccattgacgtcaataatgacgt
 5 atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaaactgccacttg
 gcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcatta
 tgcccagttacatgaccttatgggactttcctacttggcagttacatctacgtatttagtcatcgctattaccatgggtga
 tgccgtttttggcagttacatcaatgggctggatagcgggtttgactcacggggatttccaagtctccaccccattgac
 gtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgca
 10 aatgggcggttaggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgctggagac
 gcatgcctaagaagaagcggaaagtccggctccggcggtgagcaagggcgaggcagtgatcaaggagttcatgcggttc
 aaggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggcgccccctacgaggg
 caccagacccgccaagctgaaggtgaccaaggtggccccctgcccttctcctgggacatcctgtccccctcagttca
 tgtacggctccagggccttcatcaagcaccgcccgcacatccccgactactataagcagtccttccccgagggcttc
 15 aagtgggagcgctgatgaacttcgaggacggcgccgctgaccgtgacccaggacacctccctggaggacggcac
 cctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccccgtaatgcagaagaagacaatgggct
 gggaagcgctccaccgagcggttgtaccccgaggacggcggtgctgaagggcgacattaagatggccctgcgcctgaag
 gacggcgccgctacctggcggaacttcaagaccacctacaaggccaagaagcccgctgcagatgcccggcgccctacaa
 cgctcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtggaacagttacgaacgctccgagggcc
 20 gccactccaccggcgcatggacgagctgtacaagtaagctgatcagcctcgactgtgccttctagttgccagccat
 ctgttgtttgcccctcccccgctgccttcccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgag
 gaaattgcacgcattgtctgagtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggagga
 ttgggaagacaatagcaggcatgctggggatgcggtgggctctatggaggaacccctagtgtgagttggccactc
 cctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccggcgggcc
 25 tcagtgcgagcgagcgcgagctgcctgcagg

5' ITR cDNA sequence (SEQ ID NO: 51)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggtcgcccggcctcagt
 gagcgagcgagcgcgagagagggagtgcccaactccatcactaggggttcct
 30

CMV_enhancer cDNA sequence (SEQ ID NO: 52)

ctagatcccatatatggagttccgcgttacataacttacggtaaatggcccgctggctgaccgcccacgaccccc
 gccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggag
 tatttacggtaaaactgccacttggcagttacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatga
 35 cggtaaatggcccgctggcattatgccagttacatgaccttatgggactttcctacttggcagttacatctacgtat
 tagtcatcgctattaccatg

CMV_promoter cDNA sequence (SEQ ID NO: 53)

gtgatgCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTCCACCCCAT
 TGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCGCCCCATTGA
 CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGG
 5 agacgc

SV40-NLS cDNA sequence (SEQ ID NO: 54)

atgcctaagaagaagcggaaagtcggctccggc

10 mScarlet cDNA sequence (SEQ ID NO: 55)

gtgagcaagggcgaggcagtgatcaaggagttcatgCGGTTCAAGGTGCACATGGAGGGCTCCATGAACGGCCACGA
 GTTCGAGATCGAGGGCGAGGGCGAGGGCGGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTG
 GCCCCCTGCCCTTCTCCTGGGACATCCTGTCCCCTCAGTTTCATGTACGGCTCCAGGGCCTTCATCAAGCACCCCGCC
 GACATCCCCGACTACTATAAGCAGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGTGAACCTCGAGGACGGCGG
 15 CGCCGTGACCGTGACCCAGGACACCTCCCTGGAGGACGGCACCCCTGATCTACAAGGTGAAGCTCCGCGGCACCAACT
 TCCCTCCTGACGGCCCCGTAATGCAGAAGAAGACAATGGGCTGGGAAGCGTCCACCGAGCGGTTGTACCCCGAGGAC
 GCGGTGTGAAGGGCGACATTAAGATGGCCCTGCGCCTGAAGGACGGCGGCGCTACCTGGCGGACTTCAAGACCAC
 CTACAAGGCCAAGAAGCCCGTGCAGATGCCCGGCGCCTACAACGTCGACCGCAAGTTGGACATCACCTCCCACAACG
 AGGACTACACCGTGGTGGAACAGTACGAACGCTCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAG
 20 taa

BGHpA cDNA sequence (SEQ ID NO: 56)

gctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgTgccttcttgaccctg
 gaaggTgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctat
 25 tctggggggTgggTggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgCGGTGG
 gctctatgg

3' ITR cDNA sequence (SEQ ID NO: 57)

aggaacccttagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgggCGACCAAAGGTC
 30 gcccgacgcccgggctttgcccgggCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

In some embodiments of any of the compositions described herein, the vector comprises or consists of pITR-CMV-mScarlet-DD (SEQ ID NO: 58). In some embodiments of any of the compositions described herein, the vector comprises a sequence that has at least 75% (e.g., at
 35 least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at

least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity to SEQ ID NO: 58.

pITR-CMV-mScarlet-DD (SEQ ID NO: 58)

5 cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggctcgcccggcctcagt
gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctctagatcccatatatggagttccg
cgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgccattgacgtcaataatgacgt
atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttg
gcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaattggcccgctggcatta
10 tggccagtacatgaccttatgggactttcctacttggcagtagatctacgtatttagtcatcgctattaccatgggtga
tgccgtttttggcagtagatcaatggcggtgtagcggttgactcacggggattttccaagtctccacccccattgac
gtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgca
aatggcggttaggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaaccgctcagatcgctggagac
gcattgcctaagaagaagcggaaagtccgctccggcgtagcaaggcgaggcagtgatcaaggagttcatgcggttc
15 aagggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggcgccctacgaggg
caccagaccgccaagctgaaggtgaccaagggtggccccctgcccttctcctgggacatcctgtccctcagttca
tgtacggctccagggccttcatcaagcaccgcccgcacatccccgactactataagcagtccttccccgagggcttc
aagtgggagcgctgatgaacttcgaggacggcgggcgccgtgaccgtgaccagggacacctccctggaggacggcac
cctgatctacaaggtgaagctccgcgccaccaacttccctcctgacggccccgtaatgcagaagaagacaatgggct
20 gggaagcgctccaccgagcggttgtaacccgaggacggcggtgctgaaggcgacattaagatggccctgcgcctgaag
gacggcgccgctacctggcggaacttcaagaccacctacaaggccaagaagcccgtagcagatgcccggcgcctacaa
cgctcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtggaacagtagcaacgctccgagggcc
gccactccaccggcgcatggacgagctgtacaagggtaccatcagtcctgattgcggcggttagcggtagattacgtt
atcggcattggaaaacgccatgccgtggaacctgcctgccgatctcgctggtttaaacgcaacaccttaataaaacc
25 cgtgattatgggcccgcatacctgggaatcaatcggtcgctccgttgccaggacgcaaaaatattatcctcagcagtc
aaccgagtagcgacgatcggtaacgtgggtgaagtcggtggatgaagccatcgcgcggtgtggtgacgtaccagaa
atcatggtgattggcgcggtcggttattgaacagttccttgccaaaagcgcaaaaactgtatctgacgcataatcga
cgcagaagtgggaaggcgacaccatttcccggattacgagccggatgactgggaatcggtattcagcgaattccacg
atgctgatgcgcagaactctcacagctattgctttgagattctggagcggcgataagctgatcagcctcgactgtgc
30 cttctagttgccagccatctgttgtttgcccctcccccgtagccttccctgaccctggaagggtgacctccactgtc
ctttcctaataaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctgggggggtgggggtggggca
ggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcgggtgggctctatggaggaaccctag
tgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccggcgaccaaaggtcgcccgacgccc
ggctttgcccggcgccctcagtgagcgagcgagcgcgagctgcctgcagg
35

DHFR-DD cDNA sequence (SEQ ID NO: 59)

ggtaccatcagtcgtgattgcggcggttagcggtagattacgttatcggcatggaaaacgccatgccgtggaacctgcc
 tgccgatctcgctgggtttaaacgcaacaccttaataaaccgctgattatgggcccgcataacctgggaatcaatcg
 gtctgctcgttgccaggacgcaaaaatattatcctcagcagtcacccgagtagcgacgatcgcgtaacgtgggtgaag
 5 tccgtggatgaagccatcgccggcggtgtggtgacgtaccagaaatcatggtgattggcgggcggtcgcggtattgaaca
 gttcttgccaaaagcgcaaaaactgtatctgacgcataatcgacgcagaagtggaggcgacacccatttcccgatt
 acgagccggatgactgggaatcgggtattcagcgaattccacgatgctgatgcgcagaactctcacagctattgcttt
 gagattctggagcggcgataa

- 10 In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-hPou4f3-T2A-mScarlet-DD (SEQ ID NO: 60). In some embodiments
 of any of the compositions described herein, the vector comprises a sequence that has at least
 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at
 least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%)
 15 sequence identity to SEQ ID NO: 60.

pITR-CMV-hPou4f3-T2A-mScarlet-DD 60

- cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcggcgacaccttggtcgcccggcctcagt
 gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctctagatcccatatatggagttccg
 20 cgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgccattgacgtcaataatgacgt
 atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttg
 gcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcatta
 tgcccagtagatgaccttatgggactttcctacttggcagtacatctacgtatttagtcacgtctattaccatgggtga
 tgccgtttttggcagtacatcaatggcggtggatagcgggttgactcacggggatttccaagtctccaccccatgac
 25 gtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgttaacaactccgccccattgacgca
 aatggcggttaggcgtgtacgggtgggaggtctatataagcagagctcggttagtgaaacctcagatcgccgtggagac
 gcatgatggccatgaactccaagcagcctttcggcatgcacccggtgctgcaagaacccaaattctccagtctgcac
 tctggctccgaggccatgcgccgagctctgtctcccagccccgcagctgcagggtaatatatttgggaagctttgatga
 gagcctgctggcacgcgcgaagctctggcggtggatcgtctccacggcaagaacctccgttcaagccccg
 30 acgccacctaccataccatgagcagcgtgcctgcacgtccacttcgtccaccgtgcccattctcccacccagctgcg
 ctacacctcacacctcaccacgcccgtgcaccagggcctcgaaggcgacctgctggagcacatctcgccacgctgag
 tgtgagcggcctggcgctccggaacactcgggtgatgccgcacagatccatccacaccacctggcgccatgggccc
 acctgcaccaggccatgggcatgagtcacccgcacaccgtggccccctcatagcgccatgcctgcatgcctcagcgac
 gtggagtcagacccgcgcgagctggaagccttcgccgagcgttcaagcagcggcgcatcaagctgggggtgaccca
 35 ggcggacgtggcgcggtcttggttaattctcaagatccccggcggtgggctcgctgagccaaagcaccatctgcaggt
 tcgagctctctcactctctcgcaacaacatgatcgctctcaagccggtgctccaggcctggttggaggaggccgag
 gccgcctaccgagagaagaacagcaagccagagctcttcaacggcgagcgaacggaagcgcaaacgcacgtccatcgc

ggcgccggagaagcggttcactcgaggcctatttcgctatccagccacgtccttcatctgagaagatcgcgcccatcg
 ctgagaaactggaccttaaaaagaacgtggtgagagtctggttctgcaaccagagacagaaacagaaacgaatgaag
 tattcggctgtccacgttaacgattacaaggatgacgacgataaggactataaggacgatgatgacaaggactacaa
 agatgatgacgataaaggatccggcgagggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcc
 5 caatgcctaagaagaagcggaaagtcggctccggcgtgagcaagggcgaggcagtgatcaaggagttcatgcggttc
 aaggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgaggg
 caccagaccgccaagctgaaggtgaccaaggtggccccctgcccttctcctgggacatcctgtcccctcagttca
 tgtacggctccagggccttcatcaagcaccgccgacatccccgactactataagcagtccttccccgagggcttc
 aagtgggagcgcgatgaacttcgaggacggcgggcgcggtgaccgtgaccaggacacctccctggaggacggcac
 10 cctgatctacaaggtgaagctccggcgaccaaacttccctcctgacggccccgtaatgcagaagaagacaatgggct
 gggaagcggtccaccgagcggttgtaacccgaggacggcggtgctgaagggcgacattaagatggccctgcgctgaag
 gacggcgggcgctacctggcggaacttcaagaccacctacaaggccaagaagcccgatgcagatgcccggcgctacaa
 cgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtggaacagtagcaacgctccgagggcc
 gccactccaccggcggtatggacgagctgtacaaggtaccatcagtcgtgattgcggcggttagcggtagattacgtt
 15 atcggcatggaaaacgccatgccgtggaacctgcctgccgatctcgcttggtttaacgcaacaccttaataaacc
 cgtgattatgggcccatacctgggaatcaatcggtcgctccgttgccaggacgcaaaaatattatcctcagcagtc
 aaccgagtagcgacgatcgcgtaacgtgggtgaagtcggtggatgaagccatcgcgcggtgtggtgacgtaccagaa
 atcatggtgattggcgcggtcgcggttattgaacagttcttgccaaaagcgcaaaaactgtatctgacgcatatcga
 cgcagaagtggaaaggcgacaccatttcccggattacgagccggatgactgggaatcggtattcagcgaattccacg
 20 atgctgatgcgcagaactctcacagctattgctttgagattctggagcgcgataagctgatcagcctcgactgtgc
 cttctagttgccagccatctgttgtttgcccctcccccgctgccttccctgaccctggaaggtgccactcccactgtc
 ctttcctaataaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctggggggtgggggtggggca
 ggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggaggaacccctag
 tgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccggcgaccaaaggtcgcccgacgcccg
 25 ggctttgcccggcgcgctcagtgagcgagcgagcgcgagctgcctgcagg

hPou4f3 cDNA sequence (SEQ ID NO: 61)

atgatggccatgaactccaagcagcctttcggcatgcacccggtgctgcaagaacccaaattctccagtctgcactc
 tggctccgaggccatgcgccgagtcgtctcccagccccgcagctgcagggtaatatatttggaaagctttgatgaga
 30 gcctgctggcacgcgccgaagctctggcgcggtggatatcgtctcccacggcaagaaccatccgttcaagcccgac
 gccacctaccataccatgagcagcgtgccctgcacgtccacttcgtccaccgtgcccatctcccaccagctgcgct
 cacctcacaccctcaccacgcccgtgcaccagggcctcgaaggcgacgtgctggagcacatctcgccacgctgagtg
 tgagcggcctggcgctccggaacactcggtgatgccgcacagatccatccacaccacctggcgccatggggccac
 ctgcaccaggccatgggcatgagtcacccgcacaccgtggccccctcatagcgccatgcctgcatgcctcagcgacgt
 35 ggagtgcagaccgcgcgagctggaagccttcgccgagcgcttcaagcagcgcgcatcaagctgggggtgaccacgg
 cggacgtggcgcggtctggctaattctcaagatccccggcggtgggctcgctgagccaaagcaccatctgcaggttc
 gagtctctcactctctcgacaacaacatgatcgctctcaagccggtgctccaggcctggttgaggaggccgaggg
 cgccatccgagagaagaacagcaagccagagctcttcaacggcagcgaacggaagcgcaaacgcacgtccatcgcg

cgccggagaagcggttcactcgaggcctatttcgctatccagccacgtccttcattctgagaagatcgcgcccatcgct
gagaaactggaccttaaaaagaacgtggtgagagtctggttctgcaaccagagacagaaacagaaacgaatgaagta
ttcggctgtccacgttaac

5 3x FLAG cDNA sequence (SEQ ID NO: 62)

gattacaaggatgacgacgataaggactataaggacgatgatgacaaggactacaaagatgatgacgataaaggatc
cggc

T2A cDNA sequence (SEQ ID NO: 63)

10 gagggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggccca

T2A cDNA sequence (SEQ ID NO: 89)

GCGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCA

15 In some embodiments of any of the compositions described herein, the vector comprises
or consists of pITR-CMV-hGFI1-T2A-mScarlet-DD (SEQ ID NO: 64). In some embodiments
of any of the compositions described herein, the vector comprises a sequence that has at least
75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at
least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%)
20 sequence identity to SEQ ID NO: 64.

pITR-CMV-hGFI1-T2A-mScarlet-DD (SEQ ID NO: 64)

cctgcaggcagctgcgcgctcgctcgctcactgagggccgcccggcgctcgggcgacctttggctcgcccggcctcagt
gagcgagcgagcgcgagagagggagtggccaactccatcactaggggttcctctagatcccatatatggagttccg
25 cgttacataacttacggtaaatggcccgctggctgaccgccaacgacccccgcccattgacgtcaataatgacgt
atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttg
gcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaattggcccgctggcatta
tgcccagtagatgaccttatgggactttcctacttggcagtagatctacgtatttagtcatcgctattaccatgggtga
tgcggttttggcagtagatcaatggcggtgtagcggtttgactcacggggatttccaagtctccacccattgac
30 gtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgca
aatgggcggtaggcgtgtacgggtgggaggtctatataagcagagctcggttagtgtaaccgtcagatcgctggagac
gcatgccgcgctcatttctcgtaaaaagcaagaaggctcacagctaccaccagccgcgctccccaggaccagactat
tcctccggttagagaatgtaccggcgcttagccgagcagacagcacttcaaattgcaggcggggcgaaggcggagcc
ccgggaccgtttgtccccgaatcgagctgaccgaagccccagacagagcctccgcatccccagacagctgcgaag
35 gcagcgtctgcgaacggagctcgaggtttgaggacttctggaggcccccgctaccctccgctctccagcctcgag
aagtcaatgtgcccatcgctggacgaagcccagcccttccccctgcctttcaaaccgtactcatggagcggcctggc

gggttctgacctgcggcacctggtgcagagctaccgacctgtggggccctggagcgtggcgctggcctgggcctct
 tctgcgaaccgcggcgagcctggccaccggcgcgctgtacggcccgaagcgggctgcggcgggcgggggcc
 gggcgccaggagctgcagcgcaggggcccgtgccaccgctggccctggcctagggtctacggcgacttcgggtc
 tggcgagccgggctgtatgagaggccacggcagcggcggttgcgtgtaccccgagcgtggccacgggctgcacg
 5 cagacaagggcgctggcgctcaaggtggagtcggagctgctgtgcacccgctgctgctggcgggcggtcctacaag
 tgcataaagtgcagcaaggtgttctccacgcgcacgggctcgaggtgcacgtgcgcaggtcccacagcggcaccag
 accctttgcctgcgagatgtgcggcaagaccttcgggcacgcgggtgagcctggagcagcacaagccgtgcactcgc
 aggaacggagctttgactgtaagatctgtgggaagagcttcaagaggtcatccacactgtccacacacctgcttatac
 cactcagacactcggccctaccctgtcagtagctgtggcaagaggttccaccagaagtgcagacatgaagaaacacac
 10 tttcatccacactggtgagaagcctcacaagtgccaggtgtgcggcaaggcattcagccagagctccaacctcatca
 cccacagccgcaaacacacaggttcaagcccttcggctgcgacctctgtgggaagggtttccagaggaaggtggac
 ctccgaaggcaccgggagacgcagcatgggctcaaagttaacgattacaaggatgacgacgataaggactataagga
 cgatgatgacaaggactacaaagatgatgacgataaaggatccggcgagggcagaggaagtctgctaacatgcggtg
 acgtcgaggagaatcctggcccaatgcctaagaagaagcggaaagtgcggtccggcgtagcaagggcgaggcagt
 15 atcaaggagttcatgcggttcaaggtgcacatggagggtccatgaacggccacgagttcgagatcgagggcgaggg
 cgagggccgccccctacgagggcaccagaccgccaagctgaagggtgaccaagggtggccccctgcccttctcctggg
 acatcctgtccccctcagttcatgtacgggtccagggccttcatcaagcaccggcgacatccccgactactataag
 cagtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcgggcgcggtgacctgacccagga
 cacctccctggaggacggcaccctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccccgtaa
 20 tgcagaagaagacaatgggctgggaagcgtccaccgagcgggtgtaccccgaggacggcggtgctgaaggggcgacatt
 aagatggccctgcgcctgaaggacggcgccgctacctggcggaacttcaagaccacctacaaggccaagaagccggt
 gcagatgcccggcgcttacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtggaac
 agtacgaacgctccgagggccgcccactccaccggcggtatggacgagctgtacaagggtaccatcagttctgattgcg
 gcgttagcggttagattacgttatcggcattgaaaaacgcatgcccgtggaacctgcctgccgatctcgctggtttaa
 25 acgcaacaccttaataaaccgtgattatgggcccgcatacctgggaatcaatcggtcgtccgttgccaggacgca
 aaaatattatcctcagcagtcacaccgagtagcgacgatcgcgtaacgtgggtgaagtgcgtggatgaagccatcgcg
 gcgtgtggtgacgtaccagaaatcatggtgattggcgggcggtcgcggttattgaacagttcttgccaaaagcgcaaaa
 actgtatctgacgcataatcgacgcagaagtgggaaggcgacaccatttcccggattacgagccggatgactgggaat
 cggtattcagcgaattccacgatgctgatgcgcagaactctcacagctattgctttgagattctggagcggcgataa
 30 gctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgccctcccccgctgccttcttgaccctg
 gaagggtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctat
 tctgggggggtgggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgg
 gctctatggaggaacccttagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccggggcga
 ccaaaggtcgcccgacgcccgggctttgcccggggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

35

hGFI1 cDNA sequence (SEQ ID NO: 65)

ATGCCGCGCTCATTTCTCGTCAAAAGCAAGAAGGCTCACAGCTACCACCAGCCGCGCTCCCCAGGACCAGACTATTC
 CCTCCGTTTAGAGAATGTACCGGCGCCTAGCCGAGCAGACAGCACTTCAAATGCAGGCGGGGCGAAGGCGGAGCCCC

GGGACCGTTTGTCCCCGAATCGCAGCTGACCGAAGCCCCAGACAGAGCCTCCGCATCCCCAGACAGCTGCGAAGGC
 AGCGTCTGCGAACGGAGCTCGGAGTTTGAGGACTTCTGGAGGCCCCCGTCACCCTCCGCGTCTCCAGCCTCGGAGAA
 GTCAATGTGCCCATCGCTGGACGAAGCCCAGCCCTTCCCCCTGCCTTTCAAACCGTACTCATGGAGCGGCCTGGCGG
 GTTCTGACCTGCGGCACCTGGTGAGAGCTACCGACCGTGTGGGGCCCTGGAGCGTGGCGCTGGCCTGGGCCTCTTC
 5 TCGAACCCGCCCCGGAGCCTGGCCACCCGGCCGCGCTGTACGGCCCCGAAGCGGGCTGCCGGCGGCGGGGGCCGG
 GGCGCCAGGGAGCTGCAGCGCAGGGGCCGGTGCCACCGCTGGCCCTAGGGCTCTACGGCGACTTCGGGTCTG
 CGGCAGCCGGGCTGTATGAGAGGCCACGGCAGCGGGGCTTGTGTACCCCCAGCGTGGCCACGGGCTGCACGCA
 GACAAGGGCGCTGGCGTCAAGGTGGAGTCGGAGCTGCTGTGCACCCGCTGCTGCTGGGCGGCGGCTCCTACAAGTG
 CATCAAGTGACGAAGGTGTTCTCCACGCCGCACGGGCTCGAGGTGCACGTGCGCAGGTCCACAGCGGCACACAGAC
 10 CCTTTCCTGCGAGATGTGCGGCAAGACCTTCGGGCACGCGGTGAGCCTGGAGCAGCACAAGCCGTGCACTCGCAG
 GAACGGAGCTTTGACTGTAAGATCTGTGGGAAGAGCTTCAAGAGGTATCCACACTGTCCACACACCTGCTTATCCA
 CTCAGACACTCGGCCCTACCCCTGTCACTGTGGCAAGAGGTTCCACCAGAAGTCAGACATGAAGAAACACACTT
 TCATCCACACTGGTGAGAAGCCTCACAAGTGCCAGGTGTGCGGCAAGGCATTTCAGCCAGAGCTCCAACCTCATCACC
 CACAGCCGCAAACACACAGGCTTCAAGCCCTTCGGCTGCGACCTCTGTGGGAAGGGTTTCCAGAGGAAGGTGGACCT
 15 CCGAAGGCACCGGGAGACGCAGCATGGGCTCAAAGTTAAC

In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-hATOH1-T2A-mScarlet-DD (SEQ ID NO: 66). In some
 embodiments of any of the compositions described herein, the vector comprises a sequence that
 20 has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least
 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least
 99%) sequence identity to SEQ ID NO: 66.

pITR-CMV-hATOH1-T2A-mScarlet-DD (SEQ ID NO: 66)
 25 cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcggcgacctttggtcgcccgccctcagt
 gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctctagatcccatatatggagttccg
 cgttacataacttacggtaaatggccgcctggctgaccgcccacgacccccgcccattgacgtcaataatgacgt
 atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttg
 gcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggccgcctggcatta
 30 tgcccagttacatgaccttatgggactttcctacttggcagttacatctacgtattagtcacgtattaccatgggtga
 tgcggtttttggcagttacatcaatggcggtgtagcggtttgactcacggggatttccaagtctccacccattgac
 gtcaatgggagtttgttttggcaccaaaatcaacgggacttttccaaaatgtcgtacaaactccgccccattgacgca
 aatgggcggtaggcggtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgccctggagac
 gcatgtcccgctgctgcatgcagaagagtggtgaagtgaaggagttgggagaccaccatcgccagccccagccg
 35 catcatctcccgcaaccgcccgcgcccgcagccacctgcaactttgcaggcgagagagcatcccgtctaccgcg
 tgagctgtccctcctggacagcaccgaccacgcgctggctggctccactttgcagggcacgtgcacggcacgcg
 ccgcccagtatattgctacattccccggagctgggtgcctcagaggccgctgcgccccgggacgaggtggacggccg

ggggagctggtaaggaggagcagcggcggtgccagcagcagcaagagccccgggcccgggtgaaagtgcgggaacagct
 gtgcaagctgaaaggcgggggtggtggtagacgagctgggctgcagccgccaacggggccccttccagcaaacaggtga
 atgggggtgcagaagcagagacggctagcagccaacgccaggagcggcgaggatgcatgggctgaaccacgccttc
 gaccagctgcgcaatgttatcccgtcgttcaacaacgacaagaagctgtccaaatatgagaccctgcagatggcca
 5 aatctacatcaacgccttgtccgagctgctacaaacgcccagcggagggggaacagccaccgcgcctccagcctcct
 gcaaaagcgaccaccaccaccttcgcaccgcggcctcctatgaagggggcggggcaacgcgaccgcagctggggct
 cagcaggcttccggaggagccagcggccgaccccgccgggagttgcccggactcgttctcagccccagcttctgc
 gggagggtactcgggtgcagctggacgctctgcacttctgcacttctgcaggacagcgcctgacagcgatgatggcg
 aaaagaatttgtctccttctctccccgggagcatcttgcagccagtgcaggaggaaaacagcaaaacttcgcctcgg
 10 tcccacagaagcgacggggaattttccccccattcccattacagtgactcggatgaggcaagtgttaacgattacaa
 ggatgacgacgataaggactataaggacgatgatgacaaggactacaaagatgatgacgataaaggatccggcgagg
 gcagaggaagtctgctaaccatgcggtgacgtcgaggagaatcctggcccaatgcctaagaagaagcggaaagtccgc
 tccggcgtgagcaagggcgaggcagtgatcaaggagttcatgcggttcaagggtgcacatggagggtccatgaacgg
 ccacgagttcgagatcgaggggcgaggggcgaggggccgcccctacgagggcacccagaccgccaagctgaagggtacca
 15 aggggtggccccctgccttctcctgggacatcctgtcccctcagttcatgtacggctccagggccttcatcaagcac
 cccgcgcacatccccgactactataagcagtccttccccgagggttcaagtgggagcgcgtgatgaacttcgagga
 cggcgggcgccgtgaccgtgacccaggacacctccctggaggacggcaccctgatctacaaggtgaagctccgcggca
 ccaacttccctcctgacggccccgtaatgcagaagaagacaatgggctgggaagcgtccaccgagcgggtgtacccc
 gaggacggcgtgctgaaggggcgacattaagatggccctgcgcctgaaggacggcgccgctacctggcggaacttcaa
 20 gaccacctacaaggccaagaagcccgtgcagatgcccggcgccctacaacgtcgaccgcaagttggacatcacctccc
 acaacgaggactacaccgtggtggaacagtagcaacgctccgaggggccgcccactccaccggcgccatggacgagctg
 tacaagggtaccatcagctctgattgcggcggttagcggttagattacgttatcgccatggaaaacgccatgccgtggaa
 cctgcctgccgatctcgcttggtttaaacgcaacaccttaaataaaccggtgattatgggcccatacctgggaat
 caatcggtcgtccgttgccaggacgcaaaaatattatcctcagcagtcacccgagtagcgacgatcgcgtaacgtgg
 25 gtgaagtcgggtggatgaagccatcgcggcgtgtggtgacgtaccagaaaatcatggtgattggcgggcggtcgcggtat
 tgaacagttcttgccaaaagcgcaaaaactgtatctgacgcataatcgacgcagaagtggaaaggcgacacccatttcc
 cggattacgagccggatgactgggaatcggtattcagcgaattccacgatgctgatgcgacagaactctcacagctat
 tgctttgagattctggagcggcgataagctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgc
 ccctcccccggtgccttcccttgaccctggaaggtgccactcccactgtcctttcctaataaaaatgaggaaattgcac
 30 gcattgtctgagtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgggaagaca
 atagcaggcatgctggggatgcggtgggctctatggaggaacccctagtgatggagttggccactccctctctgcgc
 gctcgcctcgtcactgaggccggggcgaccaaaggtcgcccgcgcccgggctttgcccggggcgccctcagtgcgcga
 gcgagcgcgcagctgcctgcagg

35 hATOH1 cDNA sequence (SEQ ID NO: 67)

atgtccgcctgctgcatgcagaagagtggtgtaagtgaaggagttgggagaccaccatcgccagccccagccgca
 tcatctcccgcaaccgcgcgcgcgcgcgcagccacctgcaactttgcaggcgagagagcatcccgtctaccgcctg
 agctgtccctcctggacagcaccgacccacgcgcctggctggctcccactttgcagggcacgtgcacggcacgcgc

gccagtatattgctacattccccggagctgggtgcctcagaggccgctgcgccccgggacgaggtggacggccggg
 ggagctggtaaggaggagcagcggcggtgccagcagcagcaagagccccgggcccgggtgaaagtgcgggaacagctgt
 gcaagctgaaaggcgggggtggtggtagacgagctgggtgcagccgccaacgggccccttcagcaaacaggtgaat
 ggggtgcagaagcagagacggctagcagccaacgccaggaggagcggcgaggatgcatgggtgaaccacgccttga
 5 ccagctgcgcaatgttatcccgtcggttcaacaacgacaagaagctgtccaaatatgagaccctgcagatggccaaa
 tctacatcaacgccttgtccgagctgctacaaacgccagcggagggaacagccaccgcgcctccagcctcctgc
 aaaagcgaccaccaccaccttcgcaccgcggcctcctatgaagggggcggggcaacgcgacgcagctggggctca
 gcaggcttcggaggaggagccagcggccgacccccggggagttgcgggactcgcttctcagccccagcttctgcgg
 gagggtaactcggtgcagctggacgctctgcacttctcgactttcgaggacagcgccctgacagcgatgatggcgcaa
 10 aagaatttgtctccttctctccccgggagcatcttgacagccagtgcaggaggaaaacagcaaaacttcgcctcggtc
 ccacagaagcgacggggaattttccccccattcccattacagtgactcggtgagggaagtgttaac

In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-Luc-T2A-mScarlet-U6-Hes1 –S3 (SEQ ID NO: 68). In some
 15 embodiments of any of the compositions described herein, the vector comprises a sequence that
 has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least
 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least
 99%) sequence identity to SEQ ID NO: 68.

20 **pITR-CMV-Luc2-T2A-mScarlet-U6-Hes1-S3 (SEQ ID NO: 68)**
 cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggctcgggcgacctttgggtcgcccgccctcagt
 gagcgagcgagcgcgcagagaggagtgcccaactccatcactaggggttcctctagatcccatatatggagtccg
 cgttacataacttacggtaaatggcccgctgggtgaccgccaacgacccccgcccattgacgtcaataatgacgt
 atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaaactgccacttg
 25 gcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggccgcctggcatta
 tggccagtagacgttatgggactttcctacttggcagtagacatctacgtattagtcacgctattaccatgggtga
 tgcggttttggcagtagacatcaatggcggtgtagcggtttgactcacggggatttccaagtctccacccattgac
 gtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgca
 aatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgccctggagac
 30 gcatggaagatgcaaaaaacattaagaagggccagcgccattctaccactcgaagacgggaccgcccggcgagcag
 ctgcacaaagccatgaagcgtacgcccctggtgccgggcaccatcgcccttaccgacgcacatatcgaggtggacat
 tacctacgccgagtagtctcgagatgagcggttcgggtggcagaagctatgaagcgctatgggtgaatacaaaccatc
 ggatcggtggtgtgcagcgagaatagcttgagttcttcatgcccgtgttgggtgccctgttcacggtgtgggtgtg
 gccccagctaacgacatctacaacgagcgcgagctgctgaacagcatgggcatcagccagcccaccgtcgatttcgt
 35 gagcaagaaagggtgcaaaagatcctcaacgtgcaaaagaagctaccgatcataaaaagatcatcatcatggata
 gcaagaccgactaccagggttccaaagcatgtacaccttcgtgacttccatttgccaccgggttcaacgagtagc
 gacttcgtgcccagagacttcgaccgggacaaaaccatcgccctgatcatgaacagtagtggcagtagcggattgcc

caagggcgtagccctaccgcaccgcaccgcttggtgctcgattcagtcacgcccgcgaccccatcttcggcaaccaga
 tcatccccgacaccgctatcctcagcgtggtgccatttcaccacggcttcggcatggtcaccacgctgggctacttg
 atctgcggcttttcgggtcgtgctcatgtaccgcttcgaggaggagctattcttgcgagcttgcaagactataagat
 tcaatctgccctgctggtgccacactatcttagcttcttcgctaagagcactctcatcgacaagtagcacctaagca
 5 acttgacgagatcgccagcggcgggcgccgctcagcaaggaggtaggtgaggccgtggccaaacgcttccaccta
 ccaggcatccgcccagggtacggcctgacagaaacaaccagcgccattctgatcaccocgaaggggacgacaagcc
 tggcgagtaggcaaggtggtgcccttcttcgaggctaaggtggtggacttgacacaggtaagacactgggtgtga
 accagcgcgagctgtgcgtccgtggccccatgatcatgagcggtacgttaacaacccgagggtacaaacgct
 ctcatcgacaaggacggctggctgcacagcgggacatcgccactactgggacgaggacgagcacttcttcacgtgga
 10 ccggctgaagagcctgatcaaatacaagggtaccaggtagccccagccgaactggagagcatcctgctgcaacacc
 ccaacatcttcgacgcccgggtgcgcggcctgccgcagcagcatgccggcgagctgccgcgcgagctcgtcgtgctg
 gaacacggtaaaaccatgaccgagaaggagatcgtggactatgtggccagccaggttacaaccgccaagaagctgcg
 cgggtggtgtgtgttcgtggacgaggtgcctaaaggactgaccggcaagttggacgcccgaagatccgcgagattc
 tcatgaaggccaagaagggcggaagatcgccgtgggctccggagagggcgaggaagctcgtacatgaggtgac
 15 gtcgaggagaatcctggcccaatggtgagcaaggcgaggcagtgatcaaggagttcatgagggttcaaggtgcacat
 ggagggtccatgaacggccacgagttcgagatcgaggcgaggcgaggcgccctacgagggcacccagaccg
 ccaagctgaaggtgaccaagggtggccccctgcccttctcctgggacatcctgtcccctcagttcatgtacggctcc
 agggccttcatcaagcaccocgcgacatccccgactactataagcagtccttccccgagggttcaagtgggagcg
 cgtgatgaacttcgaggacggcgccgtgaccgtgaccaggaacacctccctggaggacggcacccctgatctaca
 20 aggtgaagctccgcggcaccaacttccctcctgacggccccgtaatgcagaagaagacaatgggctgggaagcgtcc
 accgagcgggtgtaccccgaggacggcggtgctgaaggcgacattaagatggccctgcgcctgaaggacggcgccg
 ctacctggcgacttcaagaccacctacaaggccaagaagcccgtgcagatgccggcgccctacaacgtcgaccgca
 agttggacatcacctcccacaacgaggactacaccgtggtggaacagtagcaacgctccgaggggccgccaactccacc
 ggcgcatggacgagctgtacaagtaagctgatcagcctcgataagatacattgatgagtttgacaaaaccacaact
 25 agaatgcagtgaaaaaatgctttatgtgaaatgtgtgatgctattgctttatgttaaccattataagctgcaa
 taaacaagtttaaggtcgggcaggaagaggcctatttcccatgattccttcatatttgcatatacgatacaaggctg
 ttagagagataattagaattaatttgactgtaaacacaaagatattagtaaaaatacgtgacgtagaaagtaataa
 tttcttggttagtttgacgttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtattt
 cgatttcttggttttatatatcttgtggaaaggacgaaacaccgaaagtcataaagcctatcgaaataggctttga
 30 tgactttcttttttaggaaccctagtgtgaggttgccactccctctcgcgcgctcgtcgtcactgaggccg
 ggcgaccaaaggtcgcccgacgcccgggctttgccggggcgccctcagtgagcgagcgagcgcgagctgc

In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-Luc-T2A-GFP-U6-Hes1 –S5 (SEQ ID NO: 78). In some
 35 embodiments of any of the compositions described herein, the vector comprises a sequence that
 has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least

88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity to SEQ ID NO: 78.

pITR-CMV-Luc2-T2A-GFP-U6-Hes1 –S5 (SEQ ID NO: 78)

5 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGT
GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTCTAGATCCCATATATGGAGTTCCG
CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTTGACGTCAATAATGACGT
ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAACTGCCCCACTTG
10 GCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
TGCCCAGTACATGACCTTATGGGACTTTCTACTTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA
TGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC
GTCAATGGGAGTTTGTGGTGGCACAAAATCAACGGGACTTTCCAAAATGTGTAACAACCTCCGCCCCATTGACGCA
AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGAC
15 GCATGGAAGATGCCAAAAACATTAAGAAGGGCCAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAG
CTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACAT
TACCTACGCCGAGTACTTCGAGATGAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATC
GGATCGTGGTGTGCAGCGAGAATAGCTTGCACTTCTTCATGCCCCGTGTTGGGTGCCCTGTTTCATCGGTGTGGCTGTG
GCCCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCTG
20 GAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAGATCATCATCATGGATA
GCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTGCCCACCCGGCTTCAACGAGTAC
GACTTCGTGCCCCGAGAGCTTCGACCGGGACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCC
CAAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTTCAGTCATGCCCCGACCCCATCTTCGGCAACCAGA
TCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTACCCACGGCTTCGGCATGTTACCCACGCTGGGCTACTTG
25 ATCTGCGGCTTTTCGGGTCTGTGCTCATGTACCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGAT
TCAATCTGCCCTGCTGGTGCCCACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGCA
ACTTGACAGAGATCGCCAGCGGGCGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCCAAACGCTTCCACCTA
CCAGGCATCCGCCAGGGCTACGGCCTGACAGAAACAACAGCGCCATTCTGATCACCCCGAAGGGGACGACAAGCC
TGGCGCAGTAGGCAAGGTGGTGGCCCTTCTTCGAGGCTAAGGTGGTGGACTTGGACACAGGTAAGACACTGGGTGTGA
ACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATGATCATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCT
30 CTCATCGACAAGGACGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGTGGA
CCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAGTGGAGAGCATCCTGCTGCAACACC
CCAACATCTTCGACGCCGGGGTCGCCGGCTGCCCGACGACGATGCCGGCGAGCTGCCCGCCGAGTCGTCGTGCTG
GAACACGGTAAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAAGAAGCTGCG
CGGTGGTGTGTGTTCTGTTGGACGAGGTGCCTAAAGGACTGACCGGCAAGTTGGACGCCCCGAAGATCCGCGAGATTC
35 TCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGGGCTCCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGAC
GTCGAGGAGAATCCTGGCCCAATGGTGAGCAAGGGCGAGGCAGTGATCAAGGAGTTCATGCGGTTCAAGGTGCACAT
GGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCG
CCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCTCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCC

AGGGCCTTCATCAAGCACCCCGCCGACATCCCCGACTACTATAAGCAGTCCTTCCCCGAGGGCTTCAAGTGGGAGCG
 CGTGATGAACTTCGAGGACGGCGGCCCGTGACCGTGACCCAGGACACCTCCCTGGAGGACGGCACCCCTGATCTACA
 AGGTGAAGCTCCGCGGCACCAACTTCCCTCCTGACGGCCCCGTAATGCAGAAGAAGACAATGGGCTGGGAAGCGTCC
 ACCGAGCGGTTGTACCCCGAGGACGGCGTGCTGAAGGGCGACATTAAGATGGCCCTGCGCCTGAAGGACGGCGGCCG
 5 CTACCTGGCGGACTTCAAGACCACCTACAAGGCCAAGAAGCCCCGTGCAGATGCCCCGGCGCCTACAACGTCGACCGCA
 AGTTGGACATCACCTCCCACAACGAGGACTACACCGTGGTGGAACAGTACGAACGCTCCGAGGGCCGCCACTCCACC
 GGCGGCATGGACGAGCTGTACAAGTAAGCTGATCAGCCTCGATAAGATACATTGATGAGTTTGGACAAACCACAAC
 AGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAA
 TAAACAAGTTAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCCTTCATATTTGCATATACGATACAAGGCTG
 10 TTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAA
 TTTCTTGGGTAGTTTGCAGTTTTAAATATGTTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTT
 CGATTTCTTGGCTTTATATATCTTGTGGAAGGACGAAACACCACTGCATGACCCAGATCAAcgaaTTGATCTGGGT
 CATGCAGTTTTTTTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCG
 GGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGC

15 In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-Luc-T2A-GFP-U6-Hes1 –KOP (SEQ ID NO: 79). In some
 embodiments of any of the compositions described herein, the vector comprises a sequence that
 has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least
 20 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least
 99%) sequence identity to SEQ ID NO: 79.

pITR-CMV-Luc2-T2A-GFP-U6-Hes1 –KOP (SEQ ID NO: 79)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcggcgacacctttggtcgccggcctcagt
 25 gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctctagatcccatatatggagttccg
 cgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgcccattgacgtcaataatgacgt
 atgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtaatttacggtaaactgcccacttg
 gcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggcccgccctggcatta
 tgcccagttacatgaccttatgggactttcctacttggcagttacatctacgtatttagtcacgtattaccatgggtga
 30 tgcggtttttggcagttacatcaatggcgctggatagcggtttgactcacggggatttccaagtctccacccattgac
 gtcaatgggagttttgtttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacgca
 aatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgccctggagac
 gcatggaagatgcaaaaaacattaagaaggggccagcgccattctacccactcgaagacgggaccgcccggcgagcag
 ctgcacaaagccatgaagcgctacgcccctgggtcccggcaccatcgccctttaccgacgcacatatcgaggtggacat
 35 tacctacgcccagttacttcgagatgagcggttcggctggcagaagctatgaagcgctatgggctgaatacaaacctac
 ggatcggtgtgtgcagcgagaatagcttgagttcttcatgcccgtgttggtgcccgtgttcacgtgtgtggctgtg
 gccccagctaacgacatctacaacgagcgcgagctgctgaacagcatgggcatcagccagcccaccgtcgatttcgt

gagcaagaaagggctgcaaaagatcctcaacgtgcaaaagaagctaccgatcatataaaaagatcatcatcatggata
 gcaagaccgactaccagggcttccaaagcatgtacaccttcgtgacttcccatttgccacccggcttcaacgagtac
 gacttcgtgcccagagagcttcgacccgggacaaaaccatcgccctgatcatgaacagtagtggcagtagccgattgcc
 caagggcgtagccctaccgcaccgcaccgcttggtggtccgattcagtcatgcccgcgaccccatcttcggcaaccaga
 5 tcatccccgacaccgctatcctcagcgtggtgccatttcaccacggcttcggcatgttcaccacgctgggctacttg
 atctgcggttttcgggtcgtgctcatgtaccgcttcgaggaggagctattcttgcgagcttgcaagactataagat
 tcaatctgcccgtgctggtgcccacactatttagcttcttcgctaagagcactctcatcgacaagtagcacctaagca
 acttgacagagatcgccagcggcgggcgccgctcagcaaggaggtaggtgaggccgtggccaaacgcttccaccta
 ccaggcatccgcccagggtacggcctgacagaaacaaccagcgccattctgatccccccgaaggggacgacaagcc
 10 tggcgcagtaggcaaggtggtgcccttcttcgaggctaaggtggtggacttggacacaggtaagacactgggtgtga
 accagcgcggcgagctgtgctgctggtgccccatgatcatgagcggctacgttaacaaccccgaggctacaaacgct
 ctcatcgacaaggacggctggtgacagcggcgacatcgccctactgggacgaggacgagcacttcttcatcgtgga
 ccggtgaagagcctgatcaaatacaagggtaccaggtagccccagccgaactggagagcatcctgctgcaacacc
 ccaacatcttcgacgcccgggtcgccggcctgcccgcagcagatgcccggcgagctgcccgcgcagctcgtcgtgctg
 15 gaacacggtaaaaccatgaccgagaaggagatcgtggactatgtggccagccaggttacaaccgccaagaagctgcg
 cgggtggtgtgtgcttcgtggacgaggtgcctaaaggactgaccggcaagttggacgcccgaagatccgcgagattc
 tcattaaggccaagaagggcggaagatcgccgtgggtccggagagggcagaggaagtctgctaacatgcggtgac
 gtcgaggagaatcctggcccaatggtgagcaagggcgaggcagtgatcaaggagttcatgcggttcaagggtcacat
 ggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgagggcaccagaccg
 20 ccaagctgaagggtgaccaaggggtggccccctgcccttctcctgggacatcctgtcccctcagttcatgtacggctcc
 agggccttcatcaagcaccgccgacatccccgactactataagcagtccttccccgagggttcaagtgggagcg
 cgtgatgaacttcgaggacggcgccgctgaccgtgaccacaggacacctccctggaggacggcacctgatctaca
 aggtgaagctccgcggcaccaacttccctcctgacggccccgtaatgcagaagaagacaatgggctgggaagcgtcc
 accgagcgggtgtaccccgaggacggcgctgtaagggcgacattaagatggccctgcccctgaaggacggcgggccg
 25 ctacctggcggacttcaagaccactacaaggccaagaagcccgtgcagatgcccggcgccctacaacgtcgaccgca
 agttggacatcacctcccacaacgaggactacaccgtggtggaacagtagaacgctccgagggccgcccactccacc
 ggcgcatggacgagctgtacaagtaagctgatcagcctcgataagatacattgatgagtttgacaaaaccacaact
 agaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgaaccattataagctgcaa
 30 taaacaagttaaggctcgggcaggaagagggcctatttcccattgatccttcatatttgcataacgatacaaggctg
 ttagagagataattagaattaatttgactgtaaacacaaagatatttagtataaaaatacgtgacgtagaaagtaataa
 tttcttgggtagtttgagtttttaaattatgttttaaaatggactatcatatgcttacggtaacttgaaagtattt
 cgatttcttggctttatatatcttgtggaaaggacgaaacaccgcagctgatataatggagaacgaattctccatta
 tatcagctgttttttaggaacccttagtgatggagttggccactccctctctgcgcgctcgtcgtcactgaggcc
 gggcgaccaaaaggtcgcccgacggcggtttgcccggcgccctcagtgagcgagcgagcgcgagctgc
 35

Hes1 –KOP (SEQ ID NO: 80)
 gcagctgatataatggagaa

SV40pA cDNA sequence (SEQ ID NO: 70)

gctgatcagcctcgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatt
tgtgaaatttgtgatgctattgctttatttgaaccattataagctgcaataaacaagtt

5 U6 cDNA sequence (SEQ ID NO: 71)

aaggctcgggcaggaagagggcctatttcccatgattccttcataatttgcataacgatacaaggctgtagagagat
aattagaattaatttgaactgtaaacacaaagatattagtagacaaatacgtgacgtagaaagtaataatttcttgggt
agtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttg
gctttatatacttctgtggaaaggacgaaacacc

10

U6 cDNA sequence (SEQ ID NO: 84)

CGGTGTTTCGTCCTTTCCACAAGATATATAAAGCCAAGAAATCGAAATACTTTCAAGTTACGGTAAGCATATGATAG
TCCATTTTAAACATAATTTTAAAACTGCAAACTACCCAAGAAATTATTACTTTCTACGTCACGTATTTGTACTAA
TATCTTTGTGTTTACAGTCAAATTAATTCTAATTATCTCTTAACAGCCTTGTATCGTATATGCAAATATGAAGGAA
15 TCATGGGAAATAGGCCCTCTTCCTGCCCGACC

siRNA cDNA sequence (SEQ ID NO: 72)

(N)₂₀CGAA(N)₂₀TTTTTT

20 In some embodiments of any of the compositions described herein, the vector comprises
or consists of pITR-CMV-mScarlet-bGHpA (SEQ ID NO: 76). In some embodiments of any of
the compositions described herein, the vector comprises a sequence that has at least 75% (e.g., at
least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at
least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity
25 to SEQ ID NO: 76.

pITR-CMV-mScarlet (SEQ ID NO: 76)

cctgcaggcagctgcgcgctcgctcgctcactgagggccgcccggcgctcggcgacctttggtcgccggcctcagt
gagcgagcgagcgcgagagagggagtggtgccaactccatcactaggggttcctgtgatgcggttttggcagtagatc
30 aatgggctgtagatagcggtttgactcacggggatttccaagtctccacccattgacgtcaatgggagtttgttttg
gcacaaaaatcaacgggactttccaaaatgtcgtaacaaactccgccccattgacgcaaatgggcggtaggcgtgtac
ggtgggaggtctatataagcagagctcggttagtgaaccgtcagatcgcttgagacgctgagcaagggcgaggca
gtgatcaaggagttcatgcggttcaagggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcga
ggcgagggccgcccctacgagggcaccagaccgccaagctgaagtgaccaaggggtggccccctgcccttctcct
35 gggacatcctgtcccctcagttcatgtacgggtccagggccttcatcaagcaccgcccgcacatccccgactactat
aagcagtccttccccgagggcttcaagtgaggcgctgatgaacttcgaggacggcgccgctgaccgtgaccca
ggacacctccctggaggacggcaccctgatctacaaggtgaagctccgcgccaccaacttccctcctgacggccccg
taatgcagaagaagacaatgggctgggaagcgtccaccgagcggttgtagcccgaggacggcgctgtagggcgac

attaagatggccctgcgctgaaggacggcgccgctacctggcggacttcaagaccacctacaaggccaagaagcc
 cgtgcagatgccccggcgctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtgg
 aacagtacgaacgctccgagggccgcccactccaccggcgcatggacgagctgtacaagtaagctgatcagcctcga
 ctgtgccttctagttgccagccatctgttgtttgcccctccccctgcttcttgaccctggaaggtgccactccc
 5 actgtcctttcctaataaaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctggggggtggggt
 ggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggaggaac
 ccctagtgtgaggttggccactccctctctgcgcgctcgctcgctcactgaggccggcgaccaaaggtcgcccgga
 cgcccggttttggccggcgccctcagtgagcgagcgagcgcgagctgctgcagg

10 In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-mScarlet-DD-bGHpA (SEQ ID NO: 77). In some embodiments of
 any of the compositions described herein, the vector comprises a sequence that has at least 75%
 (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least
 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence
 15 identity to SEQ ID NO: 77.

pITR-CMV-mScarlet (SEQ ID NO: 77)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgccccggcgctcgggcgacctttggtcgcccgccctcagt
 gagcgagcgagcgcgagagaggagtggccaactccatcactaggggttctgtgatgcgggttttggcagtacatc
 20 aatgggcgtggatagcggtttgactcacgggatttccaagtctccacccattgacgtcaatgggagtttgttttg
 gcacaaaatcaacgggactttccaaaatgtcgtaacaaactccgccccattgacgcaaatgggcggtaggcgtgtac
 ggtgggaggtctatataagcagagctcggttagtgaaccgtcagatcgctggagacgcgtgagcaagggcgaggca
 gtgatcaaggagttcatgcggttcaaggtgcacatggagggctccatgaacggccacgagttcgagatcgagggcga
 gggcgagggcgccctacgagggcaccagaccgccaagctgaaggtgaccaaggtggccccctgcccttctcct
 25 gggacatcctgtcccctcagttcatgtacggctccagggccttcatcaagcaccgcccagacatccccgactactat
 aagcagtccttccccgagggcttcaagtgggagcgcggtgatgaacttcgaggacggcgcgccgtgaccgtgacca
 ggacacctccctggaggacggcaccctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccccg
 taatgcagaagaagacaatgggctgggaagcgtccaccgagcggttgtaccccgaggacggcggtgctgaagggcgac
 attaagatggccctgcgctgaaggacggcgccgctacctggcggacttcaagaccacctacaaggccaagaagcc
 30 cgtgcagatgccccggcgctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtgg
 aacagtacgaacgctccgagggccgcccactccaccggcgcatggacgagctgtacaagtaaatcagttctgattgcg
 gcgttagcggttagattacgttatcgccatggaaaacgcatgccgtggaacctgctgccgatctcgcttggttaa
 acgcaacaccttaataaaacccgtgattatgggcccgcatacctgggaatcaatcggtcgcttgccaggacgca
 aaaatattatcctcagcagtcacccgagtagcgacgatcgcgtaacgtgggtgaagtcggtggatgaagccatcgcg
 35 gcgtgtggtgacgtaccagaaatcatggtgattggcgccggtcgcggttattgaacagttcttgccaaaagcgcaaaa
 actgtatctgacgcataatcgacgcagaagtgaaggcgacacccatttcccgattacgagccggatgactgggaat

cggatttcagcgaattccacgatgctgatgcgacagaactctcacagctattgctttgagattctggagcggcgagct
 gatcagcctcgactgtgccttctagttgccagccatctgttggtttgcccctcccccgctgccttccttgaccctggaa
 ggtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattct
 ggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggct
 5 ctatggaggaaccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccggggcgacca
 aaggtcgcccgacgcccgggctttgcccggggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-CMV-mScarlet (SEQ ID NO: 81). In some embodiments of any of the
 10 compositions described herein, the vector comprises a sequence that has at least 75% (e.g., at
 least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at
 least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity
 to SEQ ID NO: 81.

15 pITR-CMV-mScarlet (SEQ ID NO: 81)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggggcgctcgggcgacctttggctcgcccggccctcagt
 gagcgagcgagcgcgagagaggagtgcccaactccatcactaggggttcctgcggccgcacgcgtctagatccca
 tatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccacgacccccgcccattgac
 gtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacgggt
 20 aaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatgg
 cccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcacgc
 tattaccatgggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtc
 tccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactcc
 gccccattgacgcaaattggcggttaggcgtgtacgggtgggaggtctatataagcagagctcggttagtgaaccgtca
 25 gatcgcttgagacgcaccggtgccaccatgcctaagaagaagcgaaagtcgggctccggcgtagcaagggcgagg
 cagtgatcaaggagttcatgcggttcaaggtgcacatggagggtccatgaacggccacgagttcgagatcgagggc
 gagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaaggggtggccccctgccccttctc
 ctgggacatcctgtcccctcagttcatgtacgggtccagggtccatcaagcaccggcgacatccccgactact
 ataagcagtccttccccgagggcttcaagtgaggcgcggtgatgaacttcgaggacggcgcgccgtgaccgtgacc
 30 caggacacctccctggaggacggcaccctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccc
 cgtaatgcagaagaagacaatgggctgggaagcggtccaccgagcggttgtagcccgaggacggcggtgtagggcg
 acattaagatggccctgcgctgaaggacggcgccgctacctggcggaacttcaagaccacctacaaggccaagaag
 cccgtgcagatgcccggcgctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggt
 ggaacagtacgaacgctccgagggccgacctccaccggcgcatggacgagctgtacaagtaataagagctcgctg
 35 atcagcctcgactgtgccttctagttgccagccatctgttggtttgcccctcccccgctgccttccttgaccctggaag
 gtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctg

gggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctc
tatggaagcttgaattcagctgacgtgcctcggaccgctaggaacccctagtgatggagttggccactccctctctg
cgcgctcgcctcgcctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggaggcctcagtgag
cgagcgagcgcgagctgcctgcaggggagcctgatgcggtatcttctccttacgcattctgtgcggtatctcacacc
5 gcatacgtcaaagcaaccatagtagcgccctgtagcggcgcatataagcgcgggcggtgtggtggttacgcgcagcg
tgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcccttctcgccacgttcgcgggc
tttcccgctcaagctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaa
acttgatttgggtgatggttcacgtagtgggccatcgccctgatagacgggttttgcgctttgacgttggagtcca
cggttctttaatagtggactcttggttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataa
10 gggattttgcccatttccgctattgggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaaat
attaacgtttacaattttatgggtgcactctcagtacaatctgctctgatgcgcgatagttaagccagccccgacacc
cgccaacacccgctgacgcgcccctgacgggcttgctctgctcccgcatccgcttacagacaagctgtgaccgtctcc
gggagctgcatgtgtcagaggttttcacgctcatcaccgaaacgcgcgagacgaaagggcctcgatgacgcctatt
tttatagggttaatgtcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgttatgagccatatt
15 caacgggaaacgtcgaggccgagattaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataa
tgtcgggcaatcaggtgcgacaatctatcgcttgatgggaagcccgatgcgccagagttgtttctgaaacatggca
aaggtagcgttgccaatgatgttacagatgagatgggtcagactaaactggctgacggaatttatgcctcttccgacc
atcaagcattttatccgtactcctgatgatgcattactcaccactgcgatccccgaaaaacagcattccaggt
attagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtggtcctgcgccggttgcatctgattc
20 ctgtttgtaattgtccttttaacagcgatcgcgattttcgtctcgctcaggcgcaatcacgaatgaataacgggtttg
gttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaaacagctctggaaagaaatgcataaactttt
gccattctcaccggattcagtcgctcactcatggtgattttctcacttgataaccttatttttgacgaggggaaattaa
taggttgattgatgttgagcagagtcggaatcgacagccgataaccaggatcttgccatcctatggaactgcctcggt
gagttttctccttcattacagaaacggctttttcaaaaatatgggtattgataatcctgatatgaataaattgcagtt
25 tcatattgatgctcgatgagttttctaatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgtca
gaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaa
accaccgctaccagcggtgggtttgtttgcccggatcaagagctaccaactctttttccgaaggtaactggcttcagca
gagcgagataccaaataactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcct
acatacctcgctctgctaatacctgttaccagtggtgctgctgacagtgagcagataagtcgtgtcttacggggttgactc
30 aagacgatagttaccggataaggcgagcggtcggtgacgggggttcgtgcacacagcccagcttgagcgaa
cgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggagaaaggcgagac
aggtatccggtaagcggcagggctcggaacaggagagcgacgagggagcttccagggggaaacgccttggtatcttta
tagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatgga
aaaacgcccagcaacgcggcctttttacgggttccctggccttttgctggccttttgctcacatgt

35

In some embodiments of any of the compositions described herein, the vector comprises or consists of pITR-CMV-mScarlet-DD (SEQ ID NO: 82). In some embodiments of any of the

compositions described herein, the vector comprises a sequence that has at least 75% (e.g., at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity to SEQ ID NO: 82.

5

pITR-CMV-mScarlet-DD (SEQ ID NO: 82)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggctcgcccggcctcagt
gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctgcgggcgacgcgctctagatccca
tataatggagttccgcttacataacttacggtaaattggcccgccctggctgacgccccaacgacccccgcccattgac
10 gtcaataatgacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagtatattacggg
aaactgcccacttggcagtagcatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatgg
ccgcctggcattatgcccagtagcatgaccttatgggactttcctacttggcagtagcatctacgtatttagtcatcgc
tattaccatggtagtgcggttttggcagtagcatcaatgggcgtggatagcggtttgactcacgggggatttccaagtc
tccaccccatgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactcc
15 gccccattgacgcaaattgggcgttaggcgtgtacgggtgggaggtctatataagcagagctcggttagtgaaaccgtca
gatcgccctggagacgcaccggtgccaccatgcctaagaagaagcgaaagtcggctccggcgtagcaaggggcgagg
cagtgatcaaggagttcatgcggttcaagggtgcacatggagggctccatgaacggccacgagttcgagatcgagggc
gagggcgagggccgcccctacgagggcaccagaccgccaagctgaagggtgaccaagggtggccccctgcccttctc
ctgggacatcctgtcccctcagttcatgtacggctccagggccttcatcaagcaccgcccgcacatccccgactact
20 ataagcagtccttccccgagggttcaagtgggagcgcgtagtgaacttcgaggacggcgccgtagcaggtgacc
caggacacctccctggaggacggcaccctgatctacaagggtgaagctccgcgccaccaacttccctcctgacggccc
cgtaatgcagaagaagacaatgggctgggaagcggtccaccgagcggttgtagccccgaggacggcggtgctgaaggggcg
acattaagatggccctgcgcctgaaggacggcgccgctacctggcggaacttcaagaccacctacaaggccaagaag
ccggtgcagatgcccggcgccctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggt
25 ggaacagtagcaacgctccgagggccgcccactccaccggcgcatggacgagctgtacaagggtagcatcagttcga
ttgcggcgttagcggttagattacgttatcggcgtggaacacgcatgcccgtggaacctgcctgccgatctcgccctgg
tttaaacgcaacaccttaataaaacccgtgattatgggcccgcatacctgggaatcaatcggtcgctccgttgccagg
acgcaaaaatattatcctcagcagtagcaaccgagtagcgacgatcgcgtaacgtgggtgaagtcgggtggatgaagcca
tcgcggcgtgtggtgacgtaccagaaatcatggtgattggcgggcggtcgcggttattgaacagttcttgccaaaagcg
30 caaaaactgtatctgacgcataatcgacgcagaagtgggaaggcgacaccatttcccggattacgagccggatgactg
ggaatcggtattcagcgaattccacgatgctgatgcgcagaactctcacagctattgctttgagattctggagcggc
gataataagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccctgac
cttccttgaccctggaagggtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcgcatgtctgagtg
aggtgtcattctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgc
35 tggggatgcggtgggctctatggaagcttgaaatcagctgacgtgcctcgaccgctaggaacccctagtgtgagag
ttggccactccctctctgcgcgctcgctcgctcactgaggccggggcgaccaaaggctcgcccgacgcccgggctttgc
ccggggcgccctcagtgagcgagcgagcgcgagctgcctgcagggggcgccctgatgcggtattttctccttacgcac

tgtgcgggtatttcacaccgcatacgtcaaagcaaccatagtagcgccctgtagcggcgcatataagcgcggcggtg
 tgggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttccttt
 ctgcgcacggttcgcgggctttcccggtcaagctctaaatcgggggctccctttaggggtccgatttagtgctttacg
 gcacctcgacccccaaaaaacttgatttggtgatggttcacgtagtgggccatcgccctgatagacgggtttttcgcc
 5 ctttgacggttgaggtccacggttctttaatagtggaactcttggttccaaactggaacaacactcaaccctatctcgggc
 tattcttttgatttataagggattttgccgatttcggcctattgggttaaaaaatgagctgatttaacaaaaatttaa
 cgcgaattttaacaaaatattaacggtttacaattttatgggtgactctcagtacaatctgctctgatgccgcatagt
 taagccagccccgacaccgccaacaccgctgacgcgcctgacgggcttgctctgctcccgccatccgcttacaga
 caagctgtgaccgtctccgggagctgcatgtgtcagagggttttcaccgctcatcaccgaaacgcgcgagacgaaaggg
 10 cctcgtgatacgcctatttttataggttaatgtcatgaacaataaaactgtctgcttacataaacagtaatacaagg
 ggtgttatgagccatattcaacgggaaacgtcgaggcccgattaaattccaacatggatgctgatttatatgggta
 taaatgggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgcttgatgggaagcccgatgcgccagagt
 tgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacggaa
 tttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatgggttactcaccactgcgatccccgg
 15 aaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattggtgatgcgctggcagtggttctgc
 gccggttgcatcgcattcctgtttgtaattgtccttttaacagcgatcgcgatatttcgtctcgctcaggcgcaatca
 cgaatgaataacgggttggttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaaacaagtctggaa
 agaaatgcataaacttttgccattctcaccggattcagtcgctcactcatgggtgattttctcacttgataaccttattt
 ttgacgaggggaaattaataggttgattgatgttgacgagtcggaatcgacagccgataccaggatcttgccatc
 20 ctatggaaactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctga
 tatgaataaaattgcagttttcatttgatgctcgatgagtttttctaattctcatgacccaaaatcccttaacgtgagttt
 tcggtccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaatctg
 ctgcttgcaacaaaaaaaccaccgctaccagcggtggtttggttgccggatcaagagctaccaactctttttccga
 aggttaactggcttcagcagagcgcagataccaaataactgtccttctagtgtagccgtagttaggccaccacttcaag
 25 aactctgtagcaccgcctacatacctcgtctgctaatacctgttaccagtggtgctgccagtgggcgataagtcgtg
 tcttaccgggttgactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcacac
 agcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttccc
 gaaggagaaaaggcggacaggtatccggtaagcggcagggctcggaacaggagagcgcacgaggagcttccaggggg
 aaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcag
 30 gggggcggagcctatggaaaaacgcagcaacgcggcctttttacgggttcttgcccttttgctggccttttgctcac
 atgt

In some embodiments of any of the compositions described herein, the vector comprises
 or consists of pITR-U6-shHES1-S5-CMV-3xFLAG-hATOH1-DD-T2A-hPOU4F3-U6-shHES1-
 35 S3 (SEQ ID NO: 83). In some embodiments of any of the compositions described herein, the
 vector comprises a sequence that has at least 75% (e.g., at least 80%, at least 82%, at least 84%,

at least 85%, at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%) sequence identity to SEQ ID NO: 83.

pITR-U6-shHES1-S5-CMV-3xFLAG-hATOH1-DD-T2A-hPOU4F3-U6-shHES1-S3 (SEQ ID NO: 83)

5 cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggctcgcccggcctcagt
 gagcgagcgagcgcgagagagggagtgcccaactccatcactaggggttcctgcgggccgcaaaaaaactgcatgac
 ccagatcaattcggtgatctgggtcatgcagtcggtgtttcgtcctttccacaagatatataaagccaagaaatcga
 aatactttcaagttacggtaagcatatgatagtcatttttaaaacataattttaaaactgcaaactacccaagaaat
 10 tattactttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaattctaattatctcttaac
 agccttgatatcgatatatgaaatatgaaggaatcatgggaaataggccctcttctgcccgaccacgcgtctagatc
 ccataatggagttccgcgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgcccatt
 gacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatTTTAC
 ggtaaaactgcccacttggcagttacatcaagtgtatcatatgccagtagcggccctattgacgtcaatgacggtaaa
 15 tggcccgctggcattatgccagtagatgaccttatgggactttcctacttggcagtagatctacgtattagtcac
 cgctattaccatgggtgatgcggttttggcagtagatcaatggcggtggatagcggtttgactcacggggattttccaa
 gtctccacccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgttaacaac
 tccgccccattgacgcaaatggcggttaggcgtgtacggtgggaggtctatataagcagagctaccggtgccaccat
 ggattacaaggatgacgacgataaggactataaggacgatgatgacaaggactacaaagatgatgacgataaagtta
 20 actcccgctgctgcatgcagaagagtggtgctgaagtgaaggagttgggagaccaccatcgccagccccagccgcat
 catctcccgcaaccgcccgcgcccgcagccacctgcaactttgcaggcgagagagcatcccgtctaccgcctga
 gctgtccctcctggacagcaccgaccacgcgcctggctggctcccactttgcaggggcatctgcacggcacgcgccc
 ccagtatTTTGctacattccccggagctgggtgcctcagaggccgctgcgccccgggacgaggtggacggccggggg
 gagctggtaaggaggagcagcgcggtgccagcagcagcaagagccccgggcccgtgaaagtgcggggaacagctgtg
 25 caagctgaaaggcgggggtggtgtagacgagctgggctgcagccgccaacggggcccttccagcaaacaggtgaatg
 ggggtgcagaagcagagacggctagcagccaacgccaggagcgggcgaggtgcatgggctgaaccacgccttcgac
 cagctgcgcaatgttatcccgctgttcaacaacgacaagaagctgtccaaatatgagacctgcagatggcccaaat
 ctacatcaacgccttgtccgagctgctacaaacgcccagcggaggggaacagccaccgcccgcctccagcctcctgca
 aaagcgaccaccaccaccttcgacccgcccctcctatgaagggggcggggcaacgcgaccgcagctggggctcag
 30 caggcttccggagggagccagcgccgacccccggggagttgccggactcgcttctcagccccagcttctgcggg
 aggtactcggtgcagctggacgtctgcacttctcagctttcgaggacagcgccctgacagcgatgatggcgcaaa
 agaatttgtctccttctctccccgggagcatcttgcagccagtgcaggaggaaaacagcaaaacttcgcctcggtcc
 cacagaagcgacggggaattttccccccattcccattacagtgactcggtatgaggcaagtgttaacatcagttgat
 tgcggcgttagcggttagattacgttatcggtatggaaaacgcatgcccgtggaacctgcctgccgatctcgctgggt
 35 ttaaacgcaacaccttaataaaaccggtgattatgggcccatacctgggaatcaatcggtcgctccgttgccagga
 cgcaaaaatattatcctcagcagtaaccgagtagcgacgatcgcgtaacgtgggtgaagtgcgtggatgaagccat
 cgcgcggtgtggtgacgtaccagaaatcatggtgattggcgggcggtcgcggtattgaacagttcttgccaaaagcgc

aaaaactgtatctgacgcataatcgacgcagaagtggaaggcgacacccatttcccggattacgagccggatgactgg
 gaatcgggtattcagcgaattccacgatgctgatgcgagaactctcacagctattgctttgagattctggagcggcg
 aggatccggcgagggcgagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcccaatgatggccatga
 actccaagcagcctttcggcattgcacccgggtgctgcaagaacccaaattctccagtctgcaactctggctccgagggc
 5 atgcgccgagctctgtctcccagccccgcagctgcagggtaatatatttggaagctttgatgagagcctgctggcacg
 cgccgaagctctggcggcggtggatatcgtctcccacggcaagaaccatccgttcaagccccgacgccacctaccata
 ccatgagcagcgtgccttgacgtccacttcgtccacgtgcccatctcccacccagctgcgtcacctcacaccct
 caccacgcgtgcaccagggcctcgaaggcgacctgctggagcacatctcgcccacgtgagtgtagcggcctggg
 cgctccggaacactcggtagtcccgcacagatccatccacaccacctgggcgccatgggccacctgcaccaggcca
 10 tgggcatgagtcacccgcacaccgtggccccctcatagcgccatgcctgcatgcctcagcgacgtggagtcagacccg
 cgcgagctggaagccttcgcccagcgttcaagcagcggcgcatcaagctgggggtgacccaggcggacgtgggcgc
 ggctctggctaattctcaagatccccggcgtgggctcgctgagccaaagcaccatctgcaggttcgagttctctactc
 tctcgcaacaacatgatcgctctcaagccggtgctccaggcctgggtggaggaggccgagggccgctaccgagag
 aagaacagcaagccagagctcttcaacggcagcgaacggaagcgcaaacgcacgtccatcgcggcgcgggagaagcg
 15 ttactcagggcctatttcgctatccagccacgtccttcatctgagaagatcgcgccatcgctgagaaactggacc
 ttaaaaagaacgtggtgagagctctggttctgcaaccagagacagaaacagaaacgaatgaagtattcggctgtccac
 taataataaaaatatctttattttcattacatctgtgtgttggtttttgtgtgttaataaaaaaagaaagtcac
 caaagcctatttcgataggtttgatgactttcgggtgtttcgtcctttccacaagatatataaagccaagaaatcga
 aatactttcaagttacggtaagcatatgatagtcattttaaaacataattttaaaactgcaaactacccaagaaat
 20 tattactttctacgtcacgtattttgtactaataatctttgtgtttacagtcaaattaatttctaattatctcttaac
 agccttgatatgtatatgcaaatatgaaggaatcatgggaaataggccctcttctcgccgacccggaccgctagga
 acccctagtgtgaggttgccactccctctctgcgcgtcgctcgctcactgaggccggcgaccaaaggctcgccc
 gacgcccgggctttgcccggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

25 U6 cDNA sequence (SEQ ID NO: 91)

Ggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagt
 ccattttaaaacataattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaat
 atctttgtgtttacagtcaaattaatttctaattatctcttaacagccttgatcgatatgcaaatatgaaggaat
 catgggaaataggccctcttctcgccgacc

30

shHES1-1 (SEQ ID NO: 85)

cgggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatag
 tccattttaaaacataattttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaa
 tatctttgtgtttacagtcaaattaatttctaattatctcttaacagccttgatcgatatgcaaatatgaaggaa
 35 tcatgggaaataggccctcttctcgccgacc

atggattacaaggatgacgacgataaggactataaggacgatgatgacaaggactacaaagatgatgacgataaa

5 gttaactccgcctgctgcatgcagaagagtgggctgaagtgaaggagttgggagaccaccatcgccagccccagcc
gcatcatctcccgaaccgccgcgcgcgcagccaccttgcagggcagagagcatcccgctctaccgc
ctgagctgtccctcctggacagcacccagcgcgcctggctggctcccactttgcagggcatctgcacggcacgc
gccgcccagtatattgctacattccccggagctgggtgcctcagagggcgctgcgccccgggacgaggtggacggccg
gggggagctggtaaggaggagcagcggcggtgccagcagcagcaagagccccgggcccgggtgaaagtgcgggaacagc
10 tgtgcaagctgaaaggcggggtggtggtagacgagctgggctgcagccgccaacgggccccttcagcaaacaggtg
aatggggtgcagaagcagagacggctagcagccaacgccaggggagcggcgaggatgcatgggctgaaccacgcctt
cgaccagctgcgcaatgttatcccgtcgttcaacaacgacaagaagctgtccaaatatgagaccctgcagatggccc
aaatctacatcaacgccttgtccgagctgctacaaacgcccagcggaggggaacagccaccgcgcctccagcctcc
tgcaaaagcgaccaccaccaccttcgcaccgcggcctcctatgaagggggcgcggggaacgcgaccgcagctggggc
15 tcagcaggcttccggaggggagccagcggccgaccccgcggggagttgcgggaactcgcttctcagccccagcttctg
cgggaggggtactcgggtgcagctggacgctctgcacttctcgactttcgagggacagcgccttgacagcgatgatggcg
caaaagaatttgtctccttctctccccgggagcatcttgagcagcagtgaggaggaaaacagcaaaacttcgcctcg
gtccacagaagcgacggggaattttccccccattcccattacagtgactcggatgaggcaagt

ataataaaatatctttatttcattacatctgtgtgttggtttttgtgtg

25 ggtgtttcgtcctttccacaagatatataaagccaagaaatcgaaatactttcaagttacggttaagcatatgatagt
ccattttaaaacataatttttaaaactgcaaactacccaagaaattattactttctacgtcacgtattttgtactaat
atctttgtgtttacagtcaaattaattctaattatctctctaacagoccttgatatcgatatgcaaatatgaaggaat
catgggaaataggccctcttctgcccgaacc

35

**pITR-U6-shHES1-S5, hATOHessps-3xFLAG-hATOH1-T2A-hPOU4F3-US-shHES1-S3
(SEQ ID NO: 93)**

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggcgctcgggcgacctttggctcgccggcctcagt
gagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctgcgggcgcaaaaaaactgcatgac
5 ccagatcaattcggttgatctgggtcatgcagtcggtgtttcgtcctttccacaagatatataaagccaagaaatcga
aatactttcaagttacggtaagcatatgatagtcattttaaaacataattttaaaactgcaaactaccaagaat
tattactttctacgtcacgtattttgtactaatatctttgtgtttacagtcaaattaattctaattatctcttaac
agccttgatcgtatatgcaaataatgaaggaaatcatgggaaataggccctcttctgcccagaccacgcgtctatgga
gtttgcataacaaacggtttggcagctcgctctcttacctccattaacaagctgtaacatatagctgcaggttgcta
10 taatctcattaataattttggaaacttgaatattgagatatttctgagtgctcattccccatatgccagccacttctgc
catgctgactgggttcctttctctccattattagcaattagcttcttaccttccaaagtcagatccaaggtatccaag
atactagcaaaggaatcaactatgtgtgcaagttaagcatgcttaatatcacccaaacaaacaaagaggcagcattt
cttaaagtaatgaagatagataaatcgggttagtcctttgcgacactgctggtgctttctagagttttatatatattt
aagcagcttgctttatattctgtctttgcctccccaccaccagcacttttatttggagggttttggctcgccac
15 actttgggaaacttatttgatttcacggagagctgaaggaagatcatttttggcaacagacaagtttaaacacgatt
tctatgggacattgctaactggggcccctaaggagaaaggggaaactgagcggagaatgggttaaactccttggaagc
aggggagaggcaggggaggagagaagtcggaggagtataaagaaaaggacaggaaccaagaagcgtgggggtggttt
gccgtaatgtgagtggtttcttaattagagaacggttgacaatagaggggtctggcagaggtcctggccgcggtgcgg
agcgtctggagcggagcacgcgtgtcagctggtgagcgcactctcctttcaggcagctccccggggagctgtgcgg
20 ccacatttaacaccatcatcaccctccccggcctcctcaacctcggcctcctcctcgctcgacagccttccttggcc
cccaccagcagagctcacagtagcgagcgtctctcgccgtctcccgcaactcggccggggcctctctcctccccagc
tgcgagcggggagccgcaactgccactgcacctccagcaaccagcccagcacgcaaagaagctgcgcaaagttaa
agccaagcaatgccaaggggaggggaagctggaggcgggctttgagtggtcttgggcgcctggcgggtccagaatc
gccagagccgcccgcggtcggtgcacatctgacccgagtcagcttgggcaccagccgagagccggtccgcaccgct
25 cccgcaccccagccgcccgggtggtgacacacaccggagtcgaattacagccctgcaattaacatatgaatctgacg
aatttaaaagaaggaaaaaaaaaaaaaaaaacctgagcaggcttgggagtcctctgcacacaagaacttttctcggggt
gtaaaaactctttgattggctgctgcacgcgcctgcccgcgcctccattggctgagaagacacgcgaccggcgcg
aggagggggttgggagaggagcgggggagactgagtggcgctgcccgtttttaaaggggagcgccttcagca
accggagaagcatagttgcacgcgacctggtgtgtgatctccgagtggtggggagggtcgaggagggaacaaaaa
30 ataagacgttgagaagagacccggaaggccttttttttgggtgagctggtgtcccagtgctgcctccgatcctg
agcctccgagcctttgcagtgcaaccggtgccaccatggattacaaggatgacgacgataaggactataaggacgat
gatgacaaggactacaaagatgatgacgataaagttaacatgtcccgctgctgcatgcagaagagtgggctgaagt
gaaggagttgggagaccaccatcgccagcccagccgcatcatctcccgcaaccgcccgcgcccgcagccacctg
caactttgcaggcgagagagcatcccgctctaccgcctgagctgtccctcctggacagcaccgacccacgcgcctgg
35 ctggctcccactttgcagggcacatctgcacggcacgcgcccagctatttgctacattccccggagctgggtgcctc
agaggccgctgcgccccgggagcaggtggacggccgggggagctggtaaggaggagcagcggcggtgccagcagca
gcaagagccccgggcccgtgaaagtgcgggaacagctgtgcaagctgaaaggcgggggtggtggtagacgagctgggc
tgagccgccaacggggccccttcagcaaacaggtgaatggggtgcagaagcagagacggctagcagccaacgccag

ggagcggcgagcagatgcatgggctgaaccacgccttcgaccagctgcgcaatgttatcccgctcggttcaacaacgaca
 agaagctgtccaaatatgagaccctgcagatggcccaaattctacatcaacgccttggtccgagctgctacaaacgccc
 agcggaggggaacagccaccgcccgcctccagcctcctgcaaaagcgaccaccaccaccttcgaccgcgccctccta
 tgaagggggcgcggggaacgcgaccgcagctggggctcagcaggcttccggaggagccagcgccgaccccgccg
 5 ggagttgccggactcgcttctcagccccagcttctgcgggagggtactcggtgcagctggacgctctgcacttctcg
 actttcgaggacagcgccctgacagcgatgatggcgcaaaagaatttgctctccttctctccccgggagcatcttgca
 gccagtgcaggaggaaaacagcaaaacttcgcctcggtcccacagaagcgacggggaattttccccccattccatt
 acagtgactcggatgaggcaagtgttaacgagggcagaggaaagtctgctaacaatgcggtgacgtcgaggagaatcct
 ggcccaatgatggccatgaactccaagcagcctttcgcatgcacccggtgctgcaagaacccaaattctccagtct
 10 gcactctggctccgaggccatgcgccgagctctgtctcccagccccgcagctgcagggtaatatatttggaagctttg
 atgagagcctgctggcacgcgccaagctctggcggggtggatatcgctctccacggcaagaacccatccgttcaag
 cccgacgccacctaccataccatgagcagcgctgcctgcacgtccacttcgtccaccgtgcccatctcccaccacgc
 tgcgctcacctcacaccctcaccacgcgctgcaccaggcgctcgaaggcgacctgctggagcacatctcgccacgc
 tgagtgtgagcggcctggcgctccggaacactcggtgatgcccgcacagatccatccacaccacctggcgccatg
 15 ggccacctgcaccaggccatgggcatgagtcacccgcacaccgtggccctcatagcgccatgcctgcatgcctcag
 cgacgtggagtcagacccgcgcgagctggaagccttcgcccagcgcttcaagcagcggcgcatcaagctgggggtga
 cccaggcggacgtggcgcggtcttggttaattctcaagatccccggcggtgggtcgctgagccaaagcaccatctgc
 aggttcgagttctctcactctctcgcaacaacatgatcgctctcaagccggtgctccaggcctggttgaggaggc
 cgaggccgctaccgagagaagaacagcaagccagagctcttcaacggcagcgaacggaagcgcaaacgcacgtcca
 20 tcgcgggcgccggagaagcggttactcgaggcctatttcgctatccagccacgtccttcatctgagaagatcgcgcc
 atcgctgagaaaactggaccttaaaaagaacgtgggtgagagctctggttctgcaaccagagacagaaaacagaacgaat
 gaagtattcggtgtccactaaataataaaaatatctttattttcattacatctgtgtgttggtttttgtgtgttaa
 ttaaaaaaaagaaagtcataaagcctatttcgataggctttgatgactttcggtgtttcgctcctttccacaagata
 tataaagccaagaaatcgaaatactttcaagttacggtaagcatatgatagtcattttaaaacataattttaaaac
 25 tgcaaaactaccaagaaattattactttctacgtcacgtattttgtactaataatctttgtgtttacagtcaaattaa
 ttctaattatctcttaacagccttgatcgatatgcaaatatgaaggaaatcatgggaaataggccctcttcctgc
 ccgacccggaccgctaggaacccttagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggcc
 gggcgaccaaaggctcgcccgacgcccgggctttgcccggcgccctcagtgagcgagcgagcgcgagctgcctgca
 gg

30

Human POU4F3 sequence (SEQ ID NO: 95)

atgatggccatgaactccaagcagcctttcgcatgcacccggtgctgcaagaacccaaattctccagtctgcactc
 tggctccgaggccatgcgccgagctctgtctcccagccccgcagctgcagggtaatatatttggaagctttgatgaga
 gcctgctggcacgcgccaagctctggcgggcggtggatatcgctctcccacggcaagaaccatccgttcaagcccgac
 35 gccacctaccataccatgagcagcgtgccttcacgtccacttcgtccaccgtgcccattctcccaccagctgcgct
 cacctcacaccctcaccacgcgctgcaccaggcgctcgaaggcgacctgctggagcacatctcgccacgctgagtg
 tgagcggcctggcgctccggaacactcggtgatgcccgcacagatccatccacaccacctggcgccatgggcccac
 ctgcaccaggccatgggcatgagtcacccgcacaccgtggccctcatagcgccatgcctgcatgcctcagcgacgt

ggagtcagacccgcgcgagctggaagccttcgccgagcgcttcaagcagcgccgcacatcaagctgggggtgacccagg
 cggacgtggcgcggtctctggctaattctcaagatccccggcggtgggtcgctgagccaaagcaccatctgcaggttc
 gagtctctcactctctcgcacaacaacatgatcgctctcaagccggtgctccaggcctgggtggaggaggccgagggc
 cgctaccgagagaagaacagcaagccagagctcttcaacggcagcgaacggaagcgaaacgcacgtccatcgcg
 5 cgccggagaagcggttcaactcgaggcctatttcgctatccagccacgtccttcatctgagaagatcgcgccatcgct
 gagaaactggaccttaaaaagaacgtggtgagagtcgtggttctgcaaccagagacagaaacagaaacgaatgaagta
 ttcggtgtccactaa

A variety of different methods known in the art can be used to introduce any of the AAV
 10 vectors disclosed herein into a primate cell (e.g., a supporting cell or a hair cell (e.g., an inner or
 outer cochlear hair cell)). Non-limiting examples of methods for introducing an AAV vector
 into a primate cell include: lipofection, transfection (e.g., calcium phosphate transfection,
 transfection using highly branched organic compounds, transfection using cationic polymers,
 dendrimer-based transfection, optical transfection, particle-based transfection (e.g., nanoparticle
 15 transfection), or transfection using liposomes (e.g., cationic liposomes)), microinjection,
 electroporation, cell squeezing, sonoporation, protoplast fusion, impalefection, hydrodynamic
 delivery, gene gun, magnetofection, viral transfection, and nucleofection.

Skilled practitioners will appreciate that any of the AAV vectors described herein can be
 introduced into a primate cell (e.g., a hair cell or a supporting cell of the inner ear) by, for
 20 example, lipofection.

Various molecular biology techniques that can be used to correct a mutation(s) in an
 endogenous gene are also known in the art. Non-limiting examples of such techniques include
 site-directed mutagenesis, CRISPR (e.g., CRISPR/Cas9-induced knock-in mutations and
 CRISPR/Cas9-induced knock-out mutations), and TALENs. These methods can be used to
 25 correct the sequence of a defective endogenous gene present in a chromosome of a target cell
 (e.g., any of the exemplary cells described herein).

Any of the AAV vectors described herein can further include a control sequence, e.g., a
 control sequence selected from the group of a transcription initiation sequence, a transcription
 termination sequence, a promoter sequence, an enhancer sequence, an RNA splicing sequence, a
 30 polyadenylation (polyA) sequence, a Kozak consensus sequence, and a destabilizing domain
 sequence. Non-limiting examples of these control sequences are described herein. In some
 embodiments, a promoter can be a native promoter, a constitutive promoter, an inducible

promoter, and/or a tissue-specific promoter.

Some embodiments of any of the compositions and kits described herein can include any combination of the AAV vectors described herein. Some embodiments of any of the methods described herein can include the use of any combination of the AAV vectors described herein.

5

Promoters

The term “promoter” means a DNA sequence recognized by enzymes/proteins in a primate cell required to initiate the transcription of a specific gene (e.g., a hair cell differentiation gene). A promoter typically refers to, e.g., a nucleotide sequence to which an RNA polymerase and/or any associated factor binds and at which transcription is initiated. Non-limiting examples of promoters are described herein. Additional examples of promoters are known in the art.

In some embodiments, an AAV vector encoding an N-terminal portion of a hair cell differentiation protein (e.g., a human hair cell differentiation protein) can include a promoter and/or an enhancer. The AAV vector encoding the N-terminal portion of the hair cell differentiation protein can include any of the promoters and/or enhancers described herein or known in the art.

In some embodiments, the promoter is an inducible promoter, a constitutive promoter, a primate cell promoter, a viral promoter, a chimeric promoter, an engineered promoter, a tissue-specific promoter, or any other type of promoter known in the art. In some embodiments, the promoter is a RNA polymerase II promoter, such as a primate RNA polymerase II promoter. In some embodiments, the promoter is a RNA polymerase III promoter, including, but not limited to, a H1 promoter, a human U6 promoter, a mouse U6 promoter, or a swine U6 promoter. The promoter will generally be one that is able to promote transcription in cochlear cells such as hair cells or supporting cells. In some examples, the promoter is a cochlea-specific promoter or a cochlea-oriented promoter.

A variety of promoters are known in the art that can be used herein. Non-limiting examples of promoters that can be used herein include: human EF1a, human cytomegalovirus (CMV)

(GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA CTCACGGGG
 ATTTCCAAGTCTCCACCCCAT TGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCA
 ACGGGACTTTCCAAAATGTCGTAACA ACTCCGCCCCATTGACGCAAATGGGCGGTA

GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATC
GCCTGGAGACGC; SEQ ID NO: 53; US Patent No. 5,168,062), human ubiquitin C (UBC),

mouse phosphoglycerate kinase 1, polyoma adenovirus, simian virus 40 (SV40), β -globin, β -actin, α -fetoprotein, γ -globin, β -interferon, γ -glutamyl transferase, mouse mammary tumor virus

(MMTV), Rous sarcoma virus, rat insulin, glyceraldehyde-3-phosphate dehydrogenase,
metallothionein II (MT II), amylase, cathepsin, MI muscarinic receptor, retroviral LTR (e.g.

human T-cell leukemia virus HTLV), AAV ITR, interleukin-2, collagenase, platelet-derived
growth factor, adenovirus 5 E2, stromelysin, murine MX gene, glucose regulated proteins
(GRP78 and GRP94), α -2-macroglobulin, vimentin, MHC class I gene H-2k b, HSP70,

proliferin, tumor necrosis factor, thyroid stimulating hormone α gene, immunoglobulin light
chain, T-cell receptor, HLA DQ α and DQ β , interleukin-2 receptor, MHC class II, MHC class II
HLA-DR α , muscle creatine kinase, prealbumin (transthyretin), elastase I, albumin gene, c-fos, c-

HA-ras, neural cell adhesion molecule (NCAM), H2B (TH2B) histone, rat growth hormone,
human serum amyloid (SAA), troponin I (TN I), duchenne muscular dystrophy, human

immunodeficiency virus, and Gibbon Ape Leukemia Virus (GALV) promoters. Additional
examples of promoters are known in the art. See, e.g., Lodish, Molecular Cell Biology, Freeman
and Company, New York 2007. In some embodiments, the promoter is the CMV immediate
early promoter. In some embodiments, the promoter is a CAG promoter or a CAG/CBA
promoter.

The term “constitutive” promoter refers to a nucleotide sequence that, when operably
linked with a nucleic acid encoding a protein (e.g., a hair cell differentiation protein), causes
RNA to be transcribed from the nucleic acid in a primate cell (e.g., a hair cell or a supporting cell
of the inner ear) under most or all physiological conditions.

Examples of constitutive promoters include, without limitation, the retroviral Rous
sarcoma virus (RSV) LTR promoter, the cytomegalovirus (CMV) promoter (see, e.g., Boshart et
al, *Cell* 41:521-530, 1985), the SV40 promoter, the dihydrofolate reductase promoter, the beta-
actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1-alpha promoter
(Invitrogen).

Inducible promoters allow regulation of gene expression and can be regulated by
exogenously supplied compounds, environmental factors such as temperature, or the presence of
a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in

replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech, and Ariad. Additional examples of inducible promoters are known in the art.

Examples of inducible promoters regulated by exogenously supplied compounds include
 5 the zinc-inducible sheep metallothioneine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al, *Proc. Natl. Acad. Sci. U.S.A.* 93:3346-3351, 1996), the tetracycline-repressible system (Gossen et al, *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551, 1992), the tetracycline-inducible system (Gossen et al, *Science* 268:1766-1769, 1995, see
 10 also Harvey et al, *Curr. Opin. Chem. Biol.* 2:512-518, 1998), the RU486-inducible system (Wang et al, *Nat. Biotech.* 15:239-243, 1997) and Wang et al, *Gene Ther.* 4:432-441, 1997), and the rapamycin-inducible system (Magari et al. *J. Clin. Invest.* 100:2865-2872, 1997).

The term “tissue-specific” promoter refers to a promoter that is active only in certain specific cell types and/or tissues (e.g., transcription of a specific gene occurs only within cells
 15 expressing transcription regulatory proteins that bind to the tissue-specific promoter).

In some embodiments, the regulatory sequences impart tissue-specific gene expression capabilities. In some cases, the tissue-specific regulatory sequences bind tissue-specific transcription factors that induce transcription in a tissue-specific manner.

Exemplary tissue-specific promoters include but are not limited to the following: a liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a creatine kinase (MCK) promoter, a primate desmin (DES) promoter, an alpha-myosin heavy chain (a-MHC) promoter, and a cardiac Troponin T (cTnT) promoter. Additional exemplary promoters include Beta-actin promoter, hepatitis B virus core promoter (Sandig et al., *Gene*
 20 *Ther.* 3:1002-1009, 1996), alpha-fetoprotein (AFP) promoter (Arbuthnot et al., *Hum. Gene Ther.* 7:1503-1514, 1996), bone osteocalcin promoter (Stein et al., *Mol. Biol. Rep.* 24:185-196, 1997); bone sialoprotein promoter (Chen et al., *J. Bone Miner. Res.* 11:654-664, 1996), CD2 promoter (Hansal et al., *J. Immunol.* 161:1063-1068, 1998); immunoglobulin heavy chain promoter; T cell receptor alpha-chain promoter, neuronal such as neuron-specific enolase (NSE) promoter
 25 (Andersen et al., *Cell. Mol. Neurobiol.* 13:503-515, 1993), neurofilament light-chain gene
 30

promoter (Piccioli et al., *Proc. Natl. Acad. Sci. U.S.A.* 88:5611-5615, 1991), and the neuron-specific vgf gene promoter (Piccioli et al., *Neuron* 15:373-384, 1995).

In some embodiments, the tissue-specific promoter is a cochlea-specific promoter. In some embodiments, the tissue-specific promoter is a cochlear hair cell-specific promoter. Non-limiting examples of cochlear hair cell-specific promoters include but are not limited to: a
 5 ATOH1 promoter, a ATOH1 3'-enhancer, a POU4F3 promoter, a LHX3 promoter, a MYO7A promoter, a MYO6 promoter, a CHRNA9 promoter, and a CHRNA10 promoter. In some embodiments, the promoter is an outer hair cell-specific promoter such as a SLC26A5 promoter or an OCM promoter. See, e.g., Zheng et al., *Nature* 405:149-155, 2000; Tian et al. *Dev. Dyn.*
 10 231:199-203, 2004; and Ryan et al., *Adv. Otorhinolaryngol.* 66: 99-115, 2009.

In some embodiments of any of the AAV vectors described herein, the AAV vector includes a human ATOH1 enhancer-promoter (SEQ ID NO: 94).

Human ATOH1 enhancer-promoter (SEQ ID NO: 94)

15 ctatggaggtttgcataacaaacgtttggcagctcgctctcttacctccattaacaagctgtaacatatagctgcag
 gttgctataatctcattaatatatttggaaacttgaatattgagtatttctgagtgctcattccccatagccagcca
 cttctgccatgctgactgggttcctttctctccattattagcaattagcttcttaccttccaaagtcagatccaaggt
 atccaagatactagcaaaggaatcaactatgtgtgcaagttaagcatgcttaatatcacccaaacaaacaaagaggc
 agcatttcttaaagtaataagatagataaatcgggtagtcctttgcgacactgctggtgctttctagagttttat
 20 atattttaagcagcttgctttatattctgtctttgcctcccaccccaccagcacttttatttggagggttttggc
 tcgccacactttgggaaacttatttgatttcacggagagctgaaggaagatcatttttggcaacagacaagtttaaa
 cacgatttctatgggacattgctaactggggccctaaaggagaaaggggaaactgagcggagaatgggttaaactcct
 tggagcaggggagaggcaggggaggagagaagtcggaggagtataaagaaaaggacaggaaccaagaagcgtgggg
 gtggtttgccgtaatgtgagtgtttcttaattagagaacggttgacaatagagggtctggcagaggctcctggccgc
 25 ggtgcggagcgtctggagcggagcacgcgctgtcagctggtgagcgcactctcctttcaggcagctccccggggagc
 tgtgcggccacatttaacaccatcatcacccctccccggcctcctcaacctcggcctcctcctcgtcgacagccttc
 cttggccccaccagcagagctcacagtagcagcgtctctcgccgtctcccgactcggcggggcctctctcctc
 cccagctgcgcagcgggagccgcccactgcccactgcacctcccagcaaccagcccagcacgcaaagaagctgcgca
 aagttaaagccaagcaatgccaaggggaggggaagctggaggcgggctttgagtggcttctgggcgccctggcgggtc
 30 cagaatcgcccagagccgcccgcggtcgtgcacatctgacccgagtcagcttgggcaccagccgagagccggtccg
 caccgctcccgcaccccagccgcccgggtggtgacacacaccggagtcgaattacagccctgcaattaacatatgaa
 tctgacgaatttaaaagaagaaaaaaaaaaaaaaaaacctgagcaggcttgggagtcctctgcacacaagaacttttc
 tcggggtgtaaaaactctttgattggctgctcgcacgcgcctgcccgcgcctccattggctgagaagacacgcgac
 cggcgcgaggagggggttgggagaggagcggggggagactgagtggcgctgccgcttttttaaggggcgagcgcc
 35 ttcagcaaccggagaagcatagttgcacgcgacctggtgtgtgatctccgagtgggtgggggaggggtcaggaggga

aaaaaaaaataagacgttgcagaagagacccggaaagggccttttttttggttgagctggtgtcccagtgctgcctcc
gacctgagcctccgagcctttgcagtgcaa

Enhancers and 5' Cap

5 In some instances, an AAV vector can include a promoter sequence and/or an enhancer sequence. The term “enhancer” refers to a nucleotide sequence that can increase the level of transcription of a nucleic acid encoding a protein of interest (e.g., a hair cell differentiation protein). Enhancer sequences (50-1500 basepairs in length) generally increase the level of transcription by providing additional binding sites for transcription-associated proteins (e.g.,
10 transcription factors). In some embodiments, an enhancer sequence is found within an intronic sequence. Unlike promoter sequences, enhancer sequences can act at much larger distance away from the transcription start site (e.g., as compared to a promoter). Non-limiting examples of enhancers include a RSV enhancer, a CMV enhancer
(CTAGATCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGG
15 CTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAA
TGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG; SEQ ID NO: 52), and a
20 SV40 enhancer.

In some embodiments of any of the AAV vectors described herein, the AAV vector includes a CMV enhancer-promoter sequence (SEQ ID NO: 96)

CMV enhancer-promoter sequence (SEQ ID NO: 96)

25 CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCC
ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG
ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTA
TCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCA
TTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
30 GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAG
CGGTTTGACTCACGGGGATTTC AAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG

TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTG
ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT

Poly(A) Sequences

5 In some embodiments, any of the AAV vectors provided herein can include a poly(A) sequence. Most nascent eukaryotic mRNAs possess a poly(A) tail at their 3' end which is added during a complex process that includes cleavage of the primary transcript and a coupled polyadenylation reaction (see, e.g., Proudfoot et al., *Cell* 108:501-512, 2002). The poly(A) tail confers mRNA stability and transferability (Molecular Biology of the Cell, Third Edition by B. Alberts et al., Garland Publishing, 1994). In some embodiments, the poly(A) sequence is positioned 3' to the nucleic acid sequence encoding the C-terminus of the hair cell differentiation protein or a protein of interest (e.g., a Cas9 endonuclease, e.g., a SaCas9 endonuclease (e.g., any of the SaCas9 endonucleases described herein), a reporter protein (e.g., a GFP protein, a mScarlet protein)).

15 As used herein, "polyadenylation" refers to the covalent linkage of a polyadenylyl moiety, or its modified variant, to a messenger RNA molecule. In eukaryotic organisms, most messenger RNA (mRNA) molecules are polyadenylated at the 3' end. The 3' poly(A) tail is a long sequence of adenine nucleotides (e.g., 50, 60, 70, 100, 200, 500, 1000, 2000, 3000, 4000, or 5000) added to the pre-mRNA through the action of an enzyme, polyadenylate polymerase. In higher eukaryotes, the poly(A) tail is added onto transcripts that contain a specific sequence, the polyadenylation signal or "poly(A) sequence." The poly(A) tail and the protein bound to it aid in protecting mRNA from degradation by exonucleases. Polyadenylation is also important for transcription termination, export of the mRNA from the nucleus, and translation. Polyadenylation occurs in the nucleus immediately after transcription of DNA into RNA, but additionally can also occur later in the cytoplasm. After transcription has been terminated, the mRNA chain is cleaved through the action of an endonuclease complex associated with RNA polymerase. The cleavage site is usually characterized by the presence of the base sequence AAUAAA near the cleavage site. After the mRNA has been cleaved, adenosine residues are added to the free 3' end at the cleavage site.

25

As used herein, a “poly(A) sequence” is a sequence that triggers the endonuclease cleavage of an mRNA and the additional of a series of adenosines to the 3' end of the cleaved mRNA.

There are several poly(A) sequences that can be used, including those derived from
 5 bovine growth hormone (bgh) (Woychik et al., *Proc. Natl. Acad. Sci. U.S.A.* 81(13):3944-3948, 1984; U.S. Patent No. 5,122,458), mouse- β -globin, mouse- α -globin (Orkin et al., *EMBO J.* 4(2):453-456, 1985; Thein et al., *Blood* 71(2):313-319, 1988), human collagen, polyoma virus (Batt et al., *Mol. Cell Biol.* 15(9):4783-4790, 1995), the Herpes simplex virus thymidine kinase gene (HSV TK), IgG heavy-chain gene polyadenylation signal (US 2006/0040354), human
 10 growth hormone (hGH) (Szymanski et al., *Mol. Therapy* 15(7):1340-1347, 2007), the group of SV40 poly(A) sites, such as the SV40 late and early poly(A) site (Schek et al., *Mol. Cell Biol.* 12(12):5386-5393, 1992).

The poly(A) sequence can be a sequence of AATAAA. The AATAAA sequence may be substituted with other hexanucleotide sequences with homology to AATAAA which are capable
 15 of signaling polyadenylation, including ATTAAA, AGTAAA, CATAAA, TATAAA, GATAAA, ACTAAA, AATATA, AAGAAA, AATAAT, AAAAAA, AATGAA, AATCAA, AACAAA, AATCAA, AATAAC, AATAGA, AATTAA, or AATAAG (see, e.g., WO 06/12414).

In some embodiments, the poly(A) sequence can be a synthetic polyadenylation site (see,
 20 e.g., the pCl-neo expression vector of Promega which is based on Levitt et al., *Genes Dev.* 3(7):1019-1025, 1989). In some embodiments, the poly(A) sequence is the polyadenylation signal of soluble neuropilin-1 (sNRP) (AAATAAAATACGAAATG) (see, e.g., WO 05/073384). Additional examples of poly(A) sequences are known in the art.

In some embodiments, the poly(A) sequence is a bGHpA sequence
 25 (GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC
 CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACCTGTCCTTTCCTAATAAAATG
 AGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG
 GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCTCTATGG; SEQ ID NO: 56).

30

Internal Ribosome Entry Site (IRES)

In some embodiments, an AAV vector encoding the C-terminus of the hair cell differentiation protein can include a polynucleotide internal ribosome entry site (IRES). An IRES sequence is used to produce more than one polypeptide from a single gene transcript. An IRES forms a complex secondary structure that allows translation initiation to occur from any position with an mRNA immediately downstream from where the IRES is located (see, e.g., Pelletier and Sonenberg, *Mol. Cell. Biol.* 8(3):1103-1112, 1988).

There are several IRES sequences known to those skilled in the art, including those from, e.g., foot and mouth disease virus (FMDV), encephalomyocarditis virus (EMCV), human rhinovirus (HRV), cricket paralysis virus, human immunodeficiency virus (HIV), hepatitis A virus (HAV), hepatitis C virus (HCV), and poliovirus (PV). See e.g., Alberts, *Molecular Biology of the Cell*, Garland Science, 2002; and Hellen et al., *Genes Dev.* 15(13):1593-612, 2001.

In some embodiments, the IRES sequence that is incorporated into the vector that encodes the C-terminus of a hair cell differentiation protein is the foot and mouth disease virus (FMDV) 2A sequence. In some embodiments, the IRES sequence that is incorporated into the vector that encodes the C-terminal portion of a protein of interest (e.g., a Cas9 endonuclease, e.g., a SaCas9 endonuclease (e.g., any of the SaCas9 endonucleases described herein)) is the FMDV 2A sequence. The Foot and Mouth Disease Virus 2A sequence is a small peptide (approximately 18 amino acids in length) that has been shown to mediate the cleavage of polyproteins (Ryan, M D et al., *EMBO* 4:928-933, 1994; Mattion et al., *J. Virology* 70:8124-8127, 1996; Furler et al., *Gene Therapy* 8:864-873, 2001; and Halpin et al., *Plant Journal* 4:453-459, 1999). The cleavage activity of the 2A sequence has previously been demonstrated in artificial systems including plasmids and gene therapy vectors (AAV and retroviruses) (Ryan et al., *EMBO* 4:928-933, 1994; Mattion et al., *J. Virology* 70:8124-8127, 1996; Furler et al., *Gene Therapy* 8:864-873, 2001; and Halpin et al., *Plant Journal* 4:453-459, 1999; de Felipe et al., *Gene Therapy* 6:198-208, 1999; de Felipe et al., *Human Gene Therapy* 11:1921-1931, 2000; and Klump et al., *Gene Therapy* 8:811-817, 2001).

Destabilizing Domain (DD)

Any of the AAV vectors provided herein can optionally include a sequence encoding a destabilizing domain (“a destabilizing sequence”) for temporal control of protein expression.

Non-limiting examples of destabilizing sequences include sequences encoding: a FK506

5 sequence, a dihydrofolate reductase (DHFR) sequence. An exemplary DHFR destabilizing sequence is:

MISLIAALAVDYVIGMENAMPWNL PADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRK
NIILSSQPSTDDRVTWVKSVD EAIACGDVPEIMVIGGGRVIEQFLPKAQKLYLTHIDAEV
EGDTHFPDYEPDDWESVFSEFHDADAQNSHSYC FEILERR (SEQ ID NO: 48). An

10 exemplary DHFR destabilizing domain sequence is

GGTACCATCAGTCTGATTGCGGCGTTAGCGGTAGATTACGTTATCGGCATGGAAAAC
GCCATGCCGTGGAACCTGCCTGCCGATCTCGCCTGGTTTAAACGCAACACCTTAAAT
AAACCCGTGATTATGGGCCGCCATACCTGGGAATCAATCGGTTCGTCCGTTGCCAGGA
CGCAAAAATATTATCCTCAGCAGTCAACCGAGTACGGACGATCGCGTAACGTGGGT
15 GAAGTCGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGA
TTGGCGGCGGTTCGCGTTATTGAACAGTTCTTGCCAAAAGCGCAAAAACCTGTATCTGA
CGCATATCGACGCAGAAAGTGAAGGCGACACCCATTTCCCGGATTACGAGCCGGAT
GACTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGC
TATTGCTTTGAGATTCTGGAGCGGCGATAA (SEQ ID NO: 59).

20

In some embodiments of any of the AAV vectors described herein, the AAV vector includes a destabilizing domain (SEQ ID NO: 88).

Destabilizing domain (SEQ ID NO: 88)

25 Atcagtctgattgcggcgtagcggtagattacggttatcggcattggaacacgcatgccgtggaacctgcctgccga
tctcgcttggtttaaacgcaacaccttaaataaaccgctgattatgggcccgcatacctgggaatcaatcggtcgtc
cggtgccaggacgcaaaaatattatcctcagcagtcacccgagtagcgacgatcgcgtaacgtgggtgaagtcgggtg
gatgaagccatcgcggtgtgtgtgacgtaccagaaatcatggtgattggcggcggtcgcggttattgaacagttctt
gccaaaagcgcaaaaactgtatctgacgcataatcgacgcagaagtggaaggcgacacccatttcccggttacgagc
30 cggtgactgggaatcggtattcagcgaattccacgatgctgatgcgcaactctcacagctattgctttgagatt
ctggagcgcgca

Additional examples of destabilizing sequences are known in the art. In some embodiments, the destabilizing sequence is a FK506- and rapamycin-binding protein (FKBP12) sequence, and the stabilizing ligand is Shield-1 (Shld1) (Banaszynski et al. (2012) *Cell* 126(5): 995-1004). An exemplary FKBP12 destabilizing sequence is:

5 MGVEKQVIRPGNGPKPAPGQTVTVHCTGFG
 KDGDLSQKFWSTKDEGQKPFSFQIGKGA VIKGWDEGVIGMQIGEVARLRCSSDYAYGA
 GGFP AWGIQPNSVLD FEIEVLSVQ (SEQ ID NO: 49). In some embodiments, the destabilizing sequence is a DHFR sequence, and the stabilizing ligand is trimethoprim (TMP) (Iwamoto et al. (2010) *Chem Biol* 17:981-988).

10 In the absence of a stabilizing ligand, the protein sequence operatively linked to the destabilizing sequence is degraded by ubiquitination. In contrast, in the presence of a stabilizing ligand, protein degradation is inhibited, thereby allowing the protein sequence operatively linked to the destabilizing sequence to be actively expressed. As a positive control for stabilization of protein expression, protein expression can be detected by conventional means, including
 15 enzymatic, radiographic, colorimetric, fluorescence, or other spectrographic assays; fluorescent activating cell sorting (FACS) assays; immunological assays (e.g., enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and immunohistochemistry).

In some embodiments, the destabilizing sequence is a FKBP12 sequence, and the presence of an AAV vector carrying the FKBP12 gene in a primate cell (e.g., a supporting
 20 cochlear outer hair cell) is detected by western blotting. In some embodiments, the destabilizing sequence can be used to verify the temporally-specific activity of any of the AAV vectors described herein.

In some embodiments of any of the AAV vectors described herein, the AAV vector comprising the C-terminal portion of a hair cell differentiation gene, the vector further includes a
 25 destabilizing sequence 3' of the C-terminal portion of the hair cell differentiation gene. In some embodiments of the AAV vector including a sequence encoding the C-terminal portion of an ATOH1 protein, the vector further comprises a sequence encoding a destabilizing domain (DD) (e.g., any of the destabilizing domain described herein).

Reporter Sequences/Detectable Marker Genes

Any of the AAV vectors provided herein can optionally include a sequence encoding a reporter protein or a detectable marker ("a reporter sequence" or "a detectable marker gene"). Non-limiting examples of reporter sequences or detectable marker genes include DNA sequences encoding: a beta-lactamase, a beta-galactosidase (LacZ), an alkaline phosphatase, a thymidine kinase, a green fluorescent protein (GFP), a red fluorescent protein, an mCherry fluorescent protein, a yellow fluorescent protein, a chloramphenicol acetyltransferase (CAT), and a luciferase. Additional examples of reporter sequences or detectable markers are known in the art. When associated with regulatory elements which drive their expression, the reporter sequence or detectable marker gene can provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence, or other spectrographic assays; fluorescent activating cell sorting (FACS) assays; immunological assays (e.g., enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and immunohistochemistry).

In some embodiments, the reporter sequence or detectable marker gene is a 3x Flag sequence

(GATTACAAGGATGACGACGATAAGGACTATAAGGACGATGATGACAAGGACTACA AAGATGATGACGATAAAGGATCCGGC; SEQ ID NO: 62). In some embodiments, the reporter sequence or detectable marker gene is a luciferase sequence

(ATGGAAGATGCCAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAG ACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCC GGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTAC TTCGAGATGAGCGTTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAA CCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCCGTGTTGGG TGCCCTGTTTCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGA GCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAG GGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATAAAAAGATCATC ATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACT TCCCATTGTCACCCGGCTTCAACGAGTACGACTTCGTGCCCCGAGAGCTTCGACCGG GACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGG CGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTTCAGTCATGCCCGCGACCCCAT CTTCGGCAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTCACCA

CGGCTTCGGCATGTTCAACACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTC
 ATGTACCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAA
 TCTGCCCTGCTGGTGCCCACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACA
 AGTACGACCTAAGCAACTTGCACGAGATCGCCAGCGGGCGGGGCGCCGCTCAGCAAG
 5 GAGGTAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTA
 CGGCCTGACAGAAACAACCAGCGCCATTCTGATCACCCCCGAAGGGGACGACAAGC
 CTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGGACTTGAC
 ACAGGTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCAT
 GATCATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAAGG
 10 ACGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTC
 ATCGTGGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGC
 CGAACTGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGGTCCGCCG
 GCCTGCCCCGACGACGATGCCGGCGAGCTGCCCCGCCGACGTCGTCTGTGCTGGAACAC
 GGTAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAAC
 15 CGCCAAGAAGCTGCGCGGTGGTGTGTTGTCGTGGACGAGGTGCCTAAAGGACTGA
 CCGGCAAGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGC
 GGCAAGATCGCCGTGGGCTCCGGA; SEQ ID NO: 69).

In some embodiments, the reporter sequence or detectable marker gene is the LacZ gene, and the presence of a vector carrying the LacZ gene in a primate cell (e.g., a supporting cochlear
 20 outer hair cell) is detected by assays for beta-galactosidase activity. In other embodiments, the reporter sequence or detectable marker gene is a fluorescent protein (e.g., green fluorescent protein) or luciferase, the presence of a vector carrying the fluorescent protein or luciferase in a primate cell (e.g., a supporting cochlear outer hair cell) may be measured by fluorescent
 25 techniques (e.g., fluorescent microscopy or FACS) or light production in a luminometer (e.g., a spectrophotometer or an IVIS imaging instrument). In some embodiments, the reporter sequence or detectable marker gene can be used to verify the tissue-specific targeting capabilities and tissue-specific promoter regulatory activity of any of the vectors described herein.

Flanking Regions Untranslated Regions (UTRs)

30 In some embodiments, any of the AAV vectors described herein (e.g., any of the at least two different vectors) can include an untranslated region. In some embodiments, an AAV vector

can includes a 5' UTR or a 3' UTR.

Untranslated regions (UTRs) of a gene are transcribed but not translated. The 5' UTR starts at the transcription start site and continues to the start codon but does not include the start codon. The 3' UTR starts immediately following the stop codon and continues until the transcriptional termination signal. There is growing body of evidence about the regulatory roles played by the UTRs in terms of stability of the nucleic acid molecule and translation. The regulatory features of a UTR can be incorporated into any of the vectors, compositions, kits, or methods as described herein to enhance the stability of a hair cell differentiation protein or of a protein of interest (e.g., a Cas9 endonuclease, e.g., a SaCas9 endonuclease (e.g., any of the SaCas9 endonucleases described herein), a reporter protein (e.g., a GFP protein, a mScarlet protein).

Natural 5' UTRs include a sequence that plays a role in translation initiation. They harbor signatures like Kozak sequences, which are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus sequence CCR(A/G)CCAUGG, where R is a purine (A or G) three bases upstream of the start codon (AUG), which is followed by another "G". The 5' UTR have also been known, e.g., to form secondary structures that are involved in elongation factor binding.

For example, in some embodiments, a 5' UTR is included in any of the AAV vectors described herein. Non-limiting examples of 5' UTRs including those from the following genes: albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, and Factor VIII, can be used to enhance expression of a nucleic acid molecule, such as a mRNA.

In some embodiments, a 5' UTR from a mRNA that is transcribed by a cell in the cochlea can be included in any of the vectors, compositions, kits, and methods described herein.

3' UTRs are known to have stretches of adenosines and uridines embedded in them. These AU-rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU-rich elements (AREs) can be separated into three classes (Chen et al., *Mol. Cell. Biol.* 15:5777-5788, 1995; Chen et al., *Mol. Cell Biol.* 15:2010-2018, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. For example, c-Myc and MyoD mRNAs contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A) (U/A) nonamers. GM-CSF and TNF-alpha mRNAs are examples that contain class II AREs. Class III AREs are less well

defined. These U-rich regions do not contain an AUUUA motif. Two well-studied examples of this class are c-Jun and myogenin mRNAs.

Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3' UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message *in vivo*.

In some embodiments, the introduction, removal, or modification of 3' UTR AREs can be used to modulate the stability of an mRNA encoding a hair cell differentiation protein. In other embodiments, AREs can be removed or mutated to increase the intracellular stability and thus increase translation and production of a hair cell differentiation protein.

In other embodiments, non-UTR sequences may be incorporated into the 5' or 3' UTRs. In some embodiments, introns or portions of intron sequences may be incorporated into the flanking regions of the polynucleotides in any of the vectors, compositions, kits, and methods provided herein. Incorporation of intronic sequences may increase protein production as well as mRNA levels.

Inhibitory Nucleic Acids

Some embodiments of the compositions provided herein include a single AAV vector that encodes an inhibitory nucleic acid that decreases the expression of a hair cell differentiation-suppressing protein in a primate cell (e.g., a hair cell or a supporting cell of the inner ear). Inhibitory nucleic acids include, e.g., siRNA, shRNA, antisense nucleic acids, and ribozymes.

Non-limiting examples of siRNAs that can decrease the expression of a hair cell differentiation-suppressing protein in a primate cell (e.g., a hair cell or a supporting cell of the inner ear) are described herein. An inhibitory nucleic acid can be, e.g., a chemically-modified siRNAs or a vector-driven expression of short hairpin RNA (shRNA) that are then cleaved to siRNA. In some examples, an inhibitory nucleic acid can be a dsRNA (e.g., siRNA) including 16-30 nucleotides, e.g., 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in each strand, where one of the strands is substantially identical, e.g., at least 80% (or more, e.g., 85%, 90%, 95%, or 100%) identical, e.g., having 3, 2, 1, or 0 mismatched nucleotide(s), to a target region in the hair cell differentiation-suppressing mRNA, and the other strand is

complementary to the first strand. dsRNA molecules can be designed using methods known in the art, e.g., Dharmacon.com (see, siDESIGN CENTER) or “The siRNA User Guide,” available on the Internet at mpibpc.gwdg.de/abteilungen/100/105/sirna.html website.

Several methods for expressing siRNA duplexes within cells from a vector to achieve
 5 long-term target gene suppression in cells are known in the art, e.g., including vectors that use a mammalian Pol III promoter system (e.g., H1 or U6/snRNA promoter systems (Tuschl, *Nature Biotechnol.*, 20:440-448, 2002) to express functional double-stranded siRNAs; (Bagella et al., *J. Cell. Physiol.*, 177:206-213, 1998; Lee et al., *Nature Biotechnol.*, 20:500-505, 2002; Paul et al., *Nature Biotechnol.*, 20:505-508, 2002; Yu et al., *Proc. Natl. Acad. Sci. U.S.A.*, 99(9):6047-6052,
 10 2002; Sui et al., *Proc. Natl. Acad. Sci. U.S.A.* 99(6):5515-5520, 2002). Transcriptional termination by RNA Pol III occurs at runs of four consecutive T residues in the DNA template, and can be used to provide a mechanism to end the siRNA transcript at a specific sequence. The siRNA is complementary to the sequence of the target gene in 5'-3' and 3'-5' orientations, and the two strands of the siRNA can be expressed in the same construct or in separate constructs.
 15 Hairpin siRNAs, driven by H1 or U6 snRNA promoter and expressed in cells, can inhibit target gene expression (Bagella et al., 1998, supra; Lee et al., 2002, supra; Paul et al., 2002, supra; Yu et al., 2002, supra; Sui et al., 2002, supra).

Animal cells express a range of noncoding RNAs of approximately 22 nucleotides termed micro RNA (miRNAs) and can regulate gene expression at the post transcriptional or
 20 translational level during animal development. miRNAs are excised from an approximately 70 nucleotide precursor RNA stem-loop. By substituting the stem sequences of the miRNA precursor with miRNA sequence complementary to the target mRNA, a vector construct that expresses the novel miRNA can be used to produce siRNAs to initiate RNAi against specific mRNA targets in mammalian cells (Zeng, *Mol. Cell*, 9:1327-1333, 2002). When expressed by
 25 DNA vectors containing polymerase III promoters, micro-RNA designed hairpins can silence gene expression (McManus, *RNA* 8:842-850, 2002).

In some examples, an inhibitory nucleic acid can be an antisense nucleic acid molecules, i.e., nucleic acid molecules whose nucleotide sequence is complementary to all or part of an mRNA encoding a hair cell differentiation-suppressing protein. An antisense nucleic acid
 30 molecule can be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a hair cell differentiation-suppressing protein. The non-coding

regions ("5' and 3' untranslated regions") are the 5' and 3' sequences that flank the coding region and are not translated into amino acids. Based upon the sequences disclosed herein, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules to target a hair cell differentiation-suppressing gene described herein. For example, a "gene walk" comprising a series of oligonucleotides of 15-30 nucleotides spanning the length of a nucleic acid (e.g., a hair cell differentiation-suppressing mRNA) can be prepared, followed by testing for inhibition of expression of the gene. Optionally, gaps of 5-10 nucleotides can be left between the oligonucleotides to reduce the number of oligonucleotides synthesized and tested.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 nucleotides or more in length.

In some embodiments, the inhibitory nucleic acid can be a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach, *Nature*, 334:585-591, 1988)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. Methods of designing and producing ribozymes are known in the art (see, e.g., Scanlon, 1999, *Therapeutic Applications of Ribozymes*, Humana Press). A ribozyme having specificity for a hair cell differentiation-suppressing mRNA can be designed based upon the nucleotide sequence of a hair cell differentiation-suppressing cDNA (e.g., any of the exemplary cDNA sequences described herein). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a hair cell differentiation-suppressing mRNA (Cech et al. U.S. Patent No. 4,987,071; and Cech et al., U.S. Patent No. 5,116,742). Alternatively, an mRNA encoding a hair cell differentiation-suppressing protein can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (See, e.g., Bartel and Szostak, *Science*, 261:1411-1418, 1993).

In some embodiments, the administration of the single AAV vector including a sequence that encodes an inhibitory nucleic acid results in at least a 1% to about 99% decrease (e.g., a 1% decrease to about a 99% decrease, a 1% decrease to about a 95% decrease, a 1% decrease to about a 90% decrease, a 1% decrease to about a 85% decrease, a 1% decrease to about a 80% decrease, a 1% decrease to about a 75% decrease, a 1% decrease to about a 70% decrease, a 1%

decrease to about a 65% decrease, a 1% decrease to about a 60% decrease, a 1% decrease to about a 55% decrease, a 1% decrease to about a 50% decrease, a 1% decrease to about a 45% decrease, a 1% decrease to about a 40% decrease, a 1% decrease to about a 35% decrease, a 1% decrease to about a 30% decrease, a 1% decrease to about a 25% decrease, a 1% decrease to about a 20% decrease, a 1% decrease to about a 15% decrease, a 1% decrease to about a 10% decrease, about a 20% decrease to about a 99% decrease, about a 20% decrease to about a 95% decrease, about a 20% decrease to about a 90% decrease, about a 20% decrease to about a 85% decrease, about a 20% decrease to about a 80% decrease, about a 20% decrease to about a 75% decrease, about a 20% decrease to about a 70% decrease, about a 20% decrease to about a 65% decrease, about a 20% decrease to about a 60% decrease, about a 20% decrease to about a 55% decrease, about a 20% decrease to about a 50% decrease, about a 20% decrease to about a 45% decrease, about a 20% decrease to about a 40% decrease, about a 20% decrease to about a 35% decrease, about a 20% decrease to about a 30% decrease, about a 50% decrease to about a 99% decrease, about a 50% decrease to about a 95% decrease, about a 50% decrease to about a 90% decrease, about a 50% decrease to about a 85% decrease, about a 50% decrease to about a 80% decrease, about a 50% decrease to about a 75% decrease, about a 50% decrease to about a 70% decrease, about a 50% decrease to about a 65% decrease, about a 50% decrease to about a 60% decrease, about a 50% decrease to about a 55% decrease, about a 70% decrease to about a 99% decrease, about a 70% decrease to about a 95% decrease, about a 70% decrease to about a 90% decrease, about a 70% decrease to about a 85% decrease, about a 70% decrease to about a 80% decrease, about a 70% decrease to about a 75% decrease, about a 80% decrease to about a 99% decrease, about a 80% decrease to about a 95% decrease, about a 80% decrease to about a 90% decrease, about a 80% decrease to about a 85% decrease, about a 90% decrease to about a 99% decrease, or about a 90% decrease to about a 95% decrease) in the level of expression of the hair cell differentiation-suppressing mRNA or protein in a primate cell (e.g., as compared to the level of expression before administration of the single AAV vector that encodes the inhibitory nucleic acid that targets the hair cell differentiation-suppressing mRNA).

Primate Cells

Also provided herein is a cell (e.g., a primate cell, e.g., a hair cell or a supporting cell of the inner ear) that includes any of the nucleic acids, vectors (e.g., at least two different vectors

described herein), or compositions described herein. In some embodiments, the primate cell is a human cell (e.g., a human supporting cell or a human hair cell of the inner ear). In other embodiments, the primate is a non-human primate (e.g., simian cell (e.g., a monkey cell (e.g., a marmoset cell, a baboon cell, a macaque cell), or an ape cell (e.g., a gorilla cell, a gibbon cell, an orangutan cell, a chimpanzee cell). Skilled practitioners will appreciate that the AAV vectors described herein can be introduced into any primate cell (e.g., a primate supporting cell or a primate hair cell of the inner ear). Non-limiting examples of AAV vectors and methods for introducing AAV vectors into primate cells are described herein.

In some embodiments, the primate cell can be a supporting hair cell of the inner ear of a mammal. For example, a supporting cell can be Hensen's cells, Deiters' cells, inner pillar cells, outer pillar cells, Claudius cells, inner border cells, inner phalangeal cells, or cells of the stria vascularis.

In some embodiments, the primate cell is a specialized cell of the cochlea. In some embodiments, the primate cell is a hair cell. In some embodiments, the primate cell is a cochlear inner hair cell or a cochlear outer hair cell. In some embodiments, the primate cell is a cochlear inner hair cell. In some embodiments, the primate cell is a cochlear outer hair cell.

In some embodiments, the primate cell is *in vitro*. In some embodiments, the primate cell is present in a primate. In some embodiments, the primate cell is autologous cell obtained from a primate and cultured *ex vivo*.

Methods

Also provided herein are methods of promoting differentiation of a supporting cell of an inner ear of a primate into a hair cell that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering promotes differentiation of the supporting cell of the inner ear of the primate into a hair cell. Differentiation of a supporting cell of the inner ear into a hair cell can be determined using, e.g., indirect functional assays (e.g., hearing testing, e.g., pure tone audiometry).

Also provided herein are methods of increasing the expression level of a hair cell differentiation protein in a hair cell or a supporting cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in an increase (e.g., a 1% to

500% increase, a 1% to 450% increase, a 1% to 400% increase, a 1% to 350% increase, a 1% to 300% increase, a 1% to 250% increase, a 1% to 200% increase, a 1% to 150% increase, a 1% to 100% increase, a 1% to 50% increase, a 50% to 500% increase, a 50% to 450% increase, a 50% to 400% increase, a 50% to 350% increase, a 50% to 300% increase, a 50% to 250% increase, a 50% to 200% increase, a 50% to 150% increase, or a 50% to 100% increase) in the expression level of the hair cell differentiation protein in the hair cell or the supporting cell of the inner ear of the primate (e.g., as compared to the level of expression of the hair cell differentiation protein in the hair cell or the supporting cell of the inner ear of the primate before administration of the composition).

Also provided herein are methods of decreasing the expression level of a hair cell differentiation-suppressing protein in a hair cell or a supporting cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein, where the administering results in a decrease (e.g., a 1% decrease to 99% decrease, or any of the subranges of this range described herein) in the expression level of the hair cell differentiation-suppressing protein in the hair cell or the supporting cell of the inner ear of the primate (e.g., as compared to the level of expression of the hair cell differentiation-suppressing protein in the hair cell or the supporting cell of the inner ear of the primate before administration of the composition).

Also provided herein are methods of increasing (e.g., a 1% to 500% increase, or any of the subranges of this range described herein) the number of functional hair cells in a primate in need thereof (e.g., as compared to the number of functional hair cells in a primate prior to the administration of the composition) that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of improving hearing in a primate in need thereof, the method comprising administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein. In some embodiments, the administering improves hearing in a primate following environmental damage (e.g., noise, chemotherapeutic treatment (e.g., cisplatin treatment) or aminoglycoside treatment).

Also provided herein are methods of repairing a hair cell toxicity-inducing mutation in an endogenous hair cell differentiation gene locus in a hair cell or a supporting cell of an inner ear of a primate that include: administering to the inner ear of the primate a therapeutically effective

amount of any of the compositions described herein, where the administering results in repair of the hair cell toxicity-inducing mutation in the endogenous hair cell differentiation gene locus in the hair cell or the supporting cell of the inner ear of the primate.

Also provided herein are methods of decreasing the risk of hearing loss due to hair cell loss or dysfunction in a primate in need thereof that include: administering to the inner ear of the primate a therapeutically effective amount of any of the compositions described herein.

In some embodiments of any of these methods, the primate has been previously identified as having a defective hair cell differentiation gene (e.g., a hair cell differentiation gene having a mutation that results in a decrease in the expression and/or activity of a hair cell differentiation protein encoded by the gene). In some embodiments of any of these methods, the primate has been previously identified as having a defective hair cell differentiation-suppressing gene (e.g., a hair cell differentiation-suppressing gene having a mutation that results in an increase in the expression and/or activity of a hair cell differentiation-suppressing protein encoded by the gene). Some embodiments of any of these methods further include, prior to the introducing or administering step, determining that the primate has a defective hair cell differentiation gene and/or a defective hair cell differentiation-suppressing gene. Some embodiments of any of these methods can further include detecting a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene in a primate. Some embodiments of any of the methods can further include identifying or diagnosing a primate as having non-syndromic sensorineural hearing loss. Some embodiments of any of the methods can further include identifying or diagnosing a primate as having syndromic sensorineural hearing loss.

In some embodiments of any of these methods, two or more doses of any of the compositions described herein are introduced or administered into the cochlea of the primate. Some embodiments of any of these methods can include introducing or administering a first dose of the composition into the cochlea of the primate, assessing hearing function of the primate following the introducing or the administering of the first dose, and administering an additional dose of the composition into the cochlea of the primate found not to have a hearing function within a normal range (e.g., as determined using any test for hearing known in the art).

In some embodiments of any of the methods described herein, the composition can be formulated for intra-cochlear administration. In some embodiments of any of the methods described herein, the compositions described herein can be administered via intra-cochlear

administration or local administration. In some embodiments of any of the methods described herein, the compositions are administered through the use of a medical device (e.g., any of the exemplary medical devices described herein).

In some embodiments, intra-cochlear administration can be performed using any of the methods described herein or known in the art. For example, a composition can be administered or introduced into the cochlea using the following surgical technique: first using visualization with a 0 degree, 2.5-mm rigid endoscope, the external auditory canal is cleared and a round knife is used to sharply delineate an approximately 5-mm tympanomeatal flap. The tympanomeatal flap is then elevated and the middle ear is entered posteriorly. The chorda tympani nerve is identified and divided, and a curette is used to remove the scutal bone, exposing the round window membrane. To enhance apical distribution of the administered or introduced composition, a surgical laser may be used to make a small 2-mm fenestration in the oval window to allow for perilymph displacement during trans-round window membrane infusion of the composition. The microinfusion device is then primed and brought into the surgical field. The device is maneuvered to the round window, and the tip is seated within the bony round window overhang to allow for penetration of the membrane by the microneedle(s). The footpedal is engaged to allow for a measured, steady infusion of the composition. The device is then withdrawn and the round window and stapes foot plate are sealed with a gelfoam patch.

In some embodiments of any of the methods described herein, the primate has or is at risk of developing non-syndromic sensorineural hearing loss. In some embodiments of any of the methods described herein, the primate has been previously identified as having a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene. In some embodiments of any of the methods described herein, the primate has any of the mutations in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene that are described herein or are known in the art to be associated with non-syndromic sensorineural hearing loss or syndromic sensorineural hearing loss.

In some embodiments of any of the methods described herein, the primate has been identified as being a carrier of a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene (e.g., via genetic testing). In some embodiments of any of the methods described herein, the primate has been identified as having a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene and has been diagnosed

with non-syndromic sensorineural hearing loss. In some embodiments of any of the methods described herein, the primate has been identified as having a mutation in a hair cell differentiation gene and/or a hair cell differentiation-suppressing gene and has been diagnosed with syndromic sensorineural hearing loss. In some embodiments of any of the methods described herein, the primate has been identified as having non-syndromic sensorineural hearing loss. In some embodiments of any of the methods described herein, the primate has been identified as having syndromic sensorineural hearing loss.

In some embodiments, successful treatment of non-syndromic sensorineural hearing loss, or syndromic sensorineural hearing loss, can be determined in a primate using any of the conventional functional hearing tests known in the art. Non-limiting examples of functional hearing tests are various types of audiometric assays (e.g., pure-tone testing, speech testing, test of the middle ear, auditory brainstem response, and otoacoustic emissions).

In some embodiments of these methods, the primate cell is *in vitro*. In some embodiments of these methods, the primate cell is originally obtained from a primate and is cultured *ex vivo*. In some embodiments, the primate cell has previously been determined to have a defective hair cell differentiation protein and/or a defective hair cell differentiation-suppressing protein.

Methods for introducing any of the compositions described herein into a primate cell are known in the art (e.g., via lipofection or through the use of a viral vector, e.g., any of the viral vectors described herein).

An increase in expression of an active hair cell differentiation protein and/or an active hair cell differentiation-suppressing protein (e.g., a full-length hair cell differentiation protein and/or a full-length hair cell differentiation-suppressing protein) as described herein is, e.g., as compared to a control or to the level of expression of an active hair cell differentiation protein and/or a hair cell differentiation-suppressing protein (e.g., a full-length hair cell differentiation protein and/or a full-length hair cell differentiation-suppressing protein) prior to the introduction of the vector(s).

Methods of detecting expression and/or activity of a hair cell differentiation protein and/or a hair cell differentiation-suppressing protein are known in the art. In some embodiments, the level of expression of a hair cell differentiation protein and/or a hair cell differentiation-suppressing protein can be detected directly (e.g., detecting hair cell differentiation protein

and/or a hair cell differentiation-suppressing protein or detecting hair cell differentiation mRNA and/or a hair cell differentiation-suppressing mRNA). Non-limiting examples of techniques that can be used to detect expression and/or activity of hair cell differentiation proteins and/or hair cell differentiation-suppressing proteins directly include: real-time PCR, Western blotting, immunoprecipitation, immunohistochemistry, or immunofluorescence. In some embodiments, expression of a hair cell differentiation protein and/or a hair cell differentiation-suppressing protein can be detected indirectly (e.g., through functional hearing tests).

Pharmaceutical Compositions and Kits

In some embodiments, any of the compositions described herein can further include one or more agents that promote the entry of any of the AAV vectors described herein into a primate cell (e.g., a liposome or cationic lipid).

In some embodiments, any of the AAV vectors described herein can be formulated using natural and/or synthetic polymers. Non-limiting examples of polymers that may be included in any of the compositions described herein can include, but are not limited to, DYNAMIC POLYCONJUGATE® (Arrowhead Research Corp., Pasadena, Calif.), formulations from Mirus Bio (Madison, Wis.) and Roche Madison (Madison, Wis.), PhaseRX polymer formulations such as, without limitation, SMARTT POLYMER TECHNOLOGY® (PhaseRX, Seattle, Wash.), DMRI/DOPE, poloxamer, VAXFECTIN® adjuvant from Vical (San Diego, Calif.), chitosan, cyclodextrin from Calando Pharmaceuticals (Pasadena, Calif.), dendrimers and poly (lactic-co-glycolic acid) (PLGA) polymers, RONDEL™ (RNAi/Oligonucleotide Nanoparticle Delivery) polymers (Arrowhead Research Corporation, Pasadena, Calif.), and pH responsive co-block polymers, such as, but not limited to, those produced by PhaseRX (Seattle, Wash.). Many of these polymers have demonstrated efficacy in delivering nucleic acid *in vivo* into a primate cell (see, e.g., deFougerolles, *Human Gene Ther.* 19:125-132, 2008; Rozema et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:12982-12887, 2007; Rozema et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:12982-12887, 2007; Hu-Lieskovan et al., *Cancer Res.* 65:8984-8982, 2005; Heidel et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:5715-5721, 2007).

Any of the compositions described herein can be, e.g., a pharmaceutical composition. A pharmaceutical composition can include any of the compositions described herein and one or more pharmaceutically or physiologically acceptable carriers, diluents, or excipients. Such

compositions may comprise one or more buffers, such as neutral-buffered saline, phosphate-buffered saline, and the like; one or more carbohydrates, such as glucose, mannose, sucrose, and dextran; mannitol; one or more proteins, polypeptides, or amino acids, such as glycine; one or more antioxidants; one or more chelating agents, such as EDTA or glutathione; and/or one or more preservatives.

In some embodiments, the composition includes a pharmaceutically acceptable carrier (e.g., phosphate buffered saline, saline, or bacteriostatic water). Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, injectable gels, drug-release capsules, and the like.

As used herein, the term “pharmaceutically acceptable carrier” includes solvents, dispersion media, coatings, antibacterial agents, antifungal agents, and the like that are compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into any of the compositions described herein.

In some embodiments, a single dose of any of the compositions described herein can include a total amount (e.g., total sum amount of the at least two different AAV vectors, or the total amount of the single AAV vector) of at least 1 ng, at least 2 ng, at least 4 ng, about 6 ng, about 8 ng, at least 10 ng, at least 20 ng, at least 30 ng, at least 40 ng, at least 50 ng, at least 60 ng, at least 70 ng, at least 80 ng, at least 90 ng, at least 100 ng, at least 200 ng, at least 300 ng, at least 400 ng, at least 500 ng, at least 1 μ g, at least 2 μ g, at least 4 μ g, at least 6 μ g, at least 8 μ g, at least 10 μ g, at least 12 μ g, at least 14 μ g, at least 16 μ g, at least 18 μ g, at least 20 μ g, at least 22 μ g, at least 24 μ g, at least 26 μ g, at least 28 μ g, at least 30 μ g at least 32 μ g, at least 34 μ g, at least 36 μ g, at least 38 μ g, at least 40 μ g, at least 42 μ g, at least 44 μ g, at least 46 μ g, at least 48 μ g, at least 50 μ g, at least 52 μ g, at least 54 μ g, at least 56 μ g, at least 58 μ g, at least 60 μ g, at least 62 μ g, at least 64 μ g, at least 66 μ g, at least 68 μ g, at least 70 μ g, at least 72 μ g, at least 74 μ g, at least 76 μ g, at least 78 μ g, at least 80 μ g, at least 82 μ g, at least 84 μ g, at least 86 μ g, at least 88 μ g, at least 90 μ g, at least 92 μ g, at least 94 μ g, at least 96 μ g, at least 98 μ g, at least 100 μ g, at least 102 μ g, at least 104 μ g, at least 106 μ g, at least 108 μ g, at least 110 μ g, at least 112 μ g, at least 114 μ g, at least 116 μ g, at least 118 μ g, at least 120 μ g, at least 122 μ g, at least 124 μ g, at least 126 μ g, at least 128 μ g, at least 130 μ g at least 132 μ g, at least 134 μ g, at least 136 μ g, at least 138 μ g, at least 140 μ g, at least 142 μ g, at least 144 μ g, at least 146 μ g, at least 148

μg, at least 150 μg, at least 152 μg, at least 154 μg, at least 156 μg, at least 158 μg, at least 160 μg, at least 162 μg, at least 164 μg, at least 166 μg, at least 168 μg, at least 170 μg, at least 172 μg, at least 174 μg, at least 176 μg, at least 178 μg, at least 180 μg, at least 182 μg, at least 184 μg, at least 186 μg, at least 188 μg, at least 190 μg, at least 192 μg, at least 194 μg, at least 196 μg, at least 198 μg, or at least 200 μg, e.g., in a buffered solution.

The compositions provided herein can be, e.g., formulated to be compatible with their intended route of administration. A non-limiting example of an intended route of administration is local administration (e.g., intra-cochlear administration). In some embodiments, the therapeutic compositions are formulated to include a lipid nanoparticle. In some embodiments, the therapeutic compositions are formulated to include a polymeric nanoparticle. In some embodiments, the therapeutic compositions are formulated to comprise a synthetic perilymph solution. An exemplary synthetic perilymph solution includes 20-200 mM NaCl; 1-5 mM KCl; 0.1-10 mM CaCl₂; 1-10 mM glucose; 2-50 mM HEPES, having a pH of between about 6 and about 9.

Also provided are kits including any of the compositions described herein. In some embodiments, a kit can include a solid composition (e.g., a lyophilized composition including the single AAV vector or the at least two different vectors described herein) and a liquid for solubilizing the lyophilized composition. In some embodiments, a kit can include a pre-loaded syringe including any of the compositions described herein.

In some embodiments, the kit includes a vial comprising any of the compositions described herein (e.g., formulated as an aqueous composition, e.g., an aqueous pharmaceutical composition).

In some embodiments, the kits can include instructions for performing any of the methods described herein.

Devices and Surgical Methods

Provided herein are therapeutic delivery systems for treating non-syndromic sensorineural hearing loss, or syndromic sensorineural hearing loss. In one aspect, the therapeutic delivery systems include i) a medical device capable of creating one or a plurality of incisions in a round window membrane of an inner ear of a primate in need thereof, and ii) an effective dose of a composition (e.g., any of the compositions described herein). In some

embodiments, the medical device includes a plurality of micro-needles.

Also provided herein are surgical methods for treatment of hearing loss (e.g., non-syndromic sensorineural hearing loss, syndromic sensorineural hearing loss). In some embodiments, the methods include the steps of: introducing into a cochlea of a primate first
5 incision at a first incision point; and administering intra-cochlearly a therapeutically effective amount of any of the compositions provided herein. In some embodiments, the composition is administered to the primate at the first incision point. In some embodiments, the composition is administered to the primate into or through the first incision.

In some embodiments of any of the methods described herein, any of the compositions
10 described herein is administered to the primate into or through the cochlea oval window membrane. In some embodiments of any of the methods described herein, any of the compositions described herein is administered to the primate into or through the cochlea round window membrane. In some embodiments of any of the methods described herein, the composition is administered using a medical device capable of creating a plurality of incisions in
15 the round window membrane. In some embodiments, the medical device includes a plurality of micro-needles. In some embodiments, the medical device includes a plurality of micro-needles including a generally circular first aspect, where each micro-needle has a diameter of at least about 10 microns. In some embodiments, the medical device includes a base and/or a reservoir capable of holding the composition. In some embodiments, the medical device includes a
20 plurality of hollow micro-needles individually including a lumen capable of transferring the composition. In some embodiments, the medical device includes a means for generating at least a partial vacuum.

The invention is further described in detail by reference to the following experimental examples. These examples are provided for purposes of illustration only, and are not intended to
25 be limiting unless otherwise specified. Thus, the invention should in no way be construed as being limited to the following examples, but rather should be construed to encompass any and all variations that become evident as a result of the teaching provided herein.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the
30 compounds of the present invention and practice the claimed methods. The following working

examples specifically point out various aspects of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

Example 1. AAV single vector injection into the inner ear

Immunofluorescent staining was performed on cochlear tissue of a cynomolgus macaque (non-human primate) following administration of a single Anc80-GFP AAV vector directly into the inner ear through the round window.

The cochlear tissue from the treated macaque was processed for immunofluorescence analysis using Myo7a as a marker for hair cells and Iba-1 as a marker for macrophages. The middle turn is representative of the entire sensory epithelium. The data in Figures 1A-1C show clear GFP expression in both the hair cells and the supporting cells, including the following supporting cell subtypes: Hensen's cells (HC), Claudius cells (CC), Dieter cells (DC), inner and outer pillar cells (OPC/IPC), inner border cells, and inner phalangeal cells (IPHC/IBC). These data demonstrate successful Anc80-GFP AAV vector transduction into different cell types of the inner ear sensory epithelium, and the resulting expression of the encoded reporter gene (GFP) in these different cell types. These data indicate the present claimed compositions including a single AAV vector or two or more AAV vectors can be used to express a gene in hairs cells and supporting cells, and can be used to repair a mutation in a gene in hair cells and supporting cells. Figures 2A and 2B are representative images of Anc80-GFP immunofluorescent staining of the cochlear tissue. As shown in Figure 2B, expression is detected in inner hair cells.

Example 2. Exemplary vectors for promoting differentiation of a supporting cell of an inner ear of a primate into a hair cell

As shown in Figure 3, progenitor cells differentiate into either supporting cells or hair cells. Expression of Notch 1 and Hes1/5 in progenitor cells leads to the generation of supporting cells, whereas expression of Atoh1 and Wnt in progenitor cells leads to the generation of hair cells. Figures 4A-4C are exemplary vectors that can be used to promote differentiation of a supporting cell. Figure 4D is an exemplary vector that encodes a shRNA that decreases the expression of a hair cell differentiation-suppressing protein in a primate cell. The data in Figure

5A shows the relative mRNA expression levels of Hes1 in HEK293 cells that were transfected with a vector encoding S3 (SEQ ID NO: 68), a vector encoding S5 (SEQ ID NO: XX), a vector encoding Kop (SEQ ID NO: 75), vectors encoding S3 plus S5, vectors encoding S3 plus Kop and vectors encoding S5 and Kop. Relative expression was determined using RTqPCR. Cells
5 transfected with the dual vectors show increased reduction in Hes1 mRNA levels. The data in Figure 5B shows reduced Hes1 protein levels in these same cells as determined by Western blotting. Taken together, the data in Figures 5A and 5B confirms the ability of vectors to decrease target mRNA and protein levels.

The data in Figures 6A and 6B show overexpression of ATOH1, POU4F3 and GFI-1 in
10 HEK293FT cells that were transfected with the vectors of Figures 4A-4D. As shown in Figure 6A, overexpression of POU4F3 in HEK293FT cells also led to an increase in ATOH1 and GFI-1 mRNA levels. Figure 6B shows overexpression of ATOH1, GFI-1 and POU4F3 in HEK293FT cells, respectively.

HEK293FT cells were transfected with mScarlet and mScarlet-DD vectors (Figures 7A-
15 7B). The data in Figures 8A and 8B show the functionality and reversibility of the destabilizing domain (DD) using fluorescence microscopy and flow cytometry, respectively. As shown in Figure 8A, the percentage of mScarlet positive cells increased proportionately with increasing concentration of TMP in mScarlet-DD transfected HEK293FT cells, whereas the percentage of mScarlet positive cells remained constant regardless of TMP concentration in mScarlet
20 transfected HEK293FT cells. As shown in Figures 9A and 9B, mScarlet expression was seen in all HEK293FT cells transfected with mScarlet, whereas mScarlet expression was primarily seen in mScarlet-DD transfected cells in the presence of TMP. Figure 10 displays the same response in cochlear explants, where transduction and subsequent expression of mScarlet is seen in hair cells and supporting cells, whereas expression of mScarlet-DD is only seen in the presence of
25 TMP.

Figures 11A and 11B are exemplary combined vectors that can be used to promoter differentiation of a supporting cells. The vectors are combined from vectors of Figures 4A-C.

The data in Figures 12A and 12B show overexpression of ATOH1 and POU4F3 and reduction in HES1 mRNA and protein respectively, after transfection with the vectors of Figures
30 11A and 11B.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, section headings, the materials, methods, and examples are illustrative only and not intended to be limiting.

What is claimed is:

1. A composition comprising at least two different nucleic acid vectors, wherein:
each of the at least two different adeno-associated virus (AAV) vectors comprises a coding sequence that encodes a different portion of a hair cell differentiation protein, each of the encoded portions being at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions;
no single vector of the at least two different vectors encodes the full-length hair cell differentiation protein;
at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons; and
when introduced into a primate cell the at least two different vectors undergo concatamerization or homologous recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length hair cell differentiation protein that is expressed in the primate cell.
2. The composition of claim 1, wherein the amino acid sequence of none of the encoded portions overlaps with the amino acid sequence of a different one of the encoded portions.
3. The composition of claim 1, wherein the amino acid sequence of each of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions.
4. The composition of claim 3, wherein the overlapping amino acid sequence is between 30 amino acid residues to about 390 amino acid residues in length.
5. The composition of any one of claim 1-4, wherein the vectors include two different vectors, each of which comprises a different segment of an intron, wherein the intron comprises the nucleotide sequence of an intron that is present in a hair cell differentiation genomic DNA, and wherein the two different segments overlap in sequence by at least 100 nucleotides.

6. The composition of claim 5, wherein the two different intron segments overlap in sequence by about 100 nucleotides to about 800 nucleotides.

7. The composition of any one of claims 1-6, wherein the entire nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 10,000 nucleotides in length.

8. The composition of claim 7, wherein the entire nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 5,000 nucleotides in length.

9. The composition of any one of claims 1-8, wherein the number of different vectors in the composition is two.

10. The composition of claim 9, wherein a first of the two different vectors comprises a coding sequence that encodes an N-terminal portion of the hair cell differentiation protein.

11. The composition of claim 10, wherein the N-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 750 amino acids in length.

12. The composition of claim 11, wherein the N-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 320 amino acids in length.

13. The composition of any one of claims 10-12, wherein the first vector further comprises one or both of a promoter and a Kozak sequence.

14. The composition of claim 13, wherein the first vector comprises a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.

15. The composition of any one of claim 10-14, wherein the second of the two different vectors comprises a coding sequence that encodes a C-terminal portion of the hair cell differentiation protein.

16. The composition of claim 15, wherein the C-terminal portion of the hair cell differentiation protein is between about 30 amino acids to about 750 amino acids in length.

17. The composition of claim 16, wherein the C-terminal portion of the hair cell differentiation portion is between about 30 amino acids to about 320 amino acids in length.

18. The composition of any one of claim 15-17, wherein the second vector further comprises a poly(dA) sequence.

19. The composition of any one of claims 15-18, wherein the second vector further comprises a destabilizing sequence.

20. The composition of claim 19, wherein the second vector further comprises a FKB12 destabilizing sequence.

21. A composition comprising two different nucleic acid vectors, wherein:

a first nucleic acid vector of the two different nucleic acid vectors comprises a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, and a splicing donor signal sequence positioned at the 3' end of the first coding sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequences of the encoded portions do not overlap, wherein no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell, to produce RNA transcripts, splicing occurs between the splicing donor signal sequence on one transcript and the splicing acceptor signal sequence on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

22. The composition of claim 21, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons.

23. A composition comprising:

a first nucleic acid vector comprising a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' of the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence; and

a second nucleic acid vector, different from the first nucleic acid vector, comprising a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the respective amino acid sequences of the encoded portions do not overlap with each other, wherein no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell to produce RNA transcripts, splicing occurs between the splicing donor signal on one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

24. The composition of claim 23, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the neighboring exons.

25. The composition of claim 23, wherein the first or second detectable marker gene is alkaline phosphatase.

26. The composition of claim 23 or 25, wherein the first and second detectable marker genes are the same.

27. A composition comprising:

a first nucleic acid vector comprising a promoter, a first coding sequence that encodes an N-terminal portion of a hair cell differentiation protein positioned 3' to the promoter, a splicing

donor signal sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splicing donor signal sequence; and

a second nucleic acid vector, different from the first nucleic acid vector, comprising a second F1 phage recombinogenic region, a splicing acceptor signal sequence positioned 3' of the second F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a hair cell differentiation protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the respective amino acid sequences of the encoded portions do not overlap with each other, wherein no single vector of the two different vectors encodes the full-length hair cell differentiation protein, and, when the coding sequences are transcribed in a primate cell to produce RNA transcripts, splicing occurs between the splicing donor signal one transcript and the splicing acceptor signal on the other transcript, thereby forming a recombined RNA molecule that encodes a full-length hair cell differentiation protein.

28. The composition of claim 27, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of a hair cell differentiation genomic DNA, and lacks an intronic sequence between the two neighboring exons.

29. A composition comprising a single adeno-associated virus (AAV) vector, where the single AAV vector comprises a nucleic acid sequence that encodes a hair cell differentiation protein; and

when introduced into a primate cell, a nucleic acid encoding the hair cell differentiation protein is generated at the locus of the hair cell differentiation gene and the primate cell expresses the hair cell differentiation protein.

30. The composition of any one of claims 1-29, wherein the hair cell differentiation gene is selected from the group consisting of: atonal bHLH transcription factor 1 (ATOH1), POU Class 4 Homeobox 3 (POU4F3), catenin beta 1 (CTNNB1), Noggin (NOG), growth factor

independent 1 transcriptional repressor (GFI-1), neurotrophin 3 (NTF3) and brain-derived neurotrophic factor (BDNF).

31. A composition comprising two different nucleic acid vectors, wherein:

a first nucleic acid vector comprises a first nucleic acid sequence that encodes a first hair cell differentiation protein; and

a second nucleic acid vector comprises a second nucleic acid sequence that encodes a second hair cell differentiation protein, wherein the first hair cell differentiation protein and the second hair cell differentiation protein are different proteins, and

when introduced into a primate cell, the first nucleic acid and the second nucleic acid encoding the first hair cell differentiation protein and the second hair cell differentiation protein are generated at the locus of the hair cell differentiation gene and the primate cell expresses the first hair cell differentiation protein and the second hair cell differentiation protein.

32. The composition of claim 31, wherein the first and the second hair cell differentiation proteins are selected from the group consisting of: atonal bHLH transcription factor 1 (ATOH1), POU Class 4 Homeobox 3 (POU4F3), catenin beta 1 (CTNNB1), Noggin (NOG), growth factor independent 1 transcriptional repressor (GFI-1), neurotrophin 3 (NTF3) and brain-derived neurotrophic factor (BDNF).

33. The composition of any one of claims 21-32, wherein the second nucleic acid vector further comprises a destabilizing sequence.

34. The composition of claim 33, wherein the second nucleic acid vector further comprises a FKB12 destabilizing sequence.

35. A composition comprising at least one adeno-associated virus (AAV) vector that encodes an inhibitory nucleic acid that decreases the expression of a hair cell differentiation-suppressing protein in a primate cell.

36. The composition of claim 35, wherein the inhibitory nucleic acid is a short interfering RNA (siRNA), a short hairpin RNA (shRNA), an antisense oligonucleotide, or a ribozyme.

37. The composition of claim 30 or 32, wherein the hair cell differentiation-suppressing gene is hes family bHLH transcription factor 1 (HES1), hes family bHLH transcription factor 5 (HES5), sex determining region Y-box 2 (SOX2), and p27kip (CDKN1B).

38. The composition of any one of claims 1-37, further comprising a pharmaceutically acceptable excipient.

39. A kit comprising a composition of any one of claims 1-38.

40. A kit of claim 39, further comprising a pre-loaded syringe comprising the composition.

41. A method of promoting differentiation of a supporting cell of an inner ear of a primate into a hair cell, the method comprising:

administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-38,

wherein the administering promotes differentiation of the supporting cell of the inner ear of the primate into a hair cell.

42. A method of increasing the expression level of a hair cell differentiation protein in a supporting cell or a hair cell of an inner ear of a primate, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-32,

wherein the administering results in an increase in the expression level of the hair cell differentiation protein in the supporting cell or the hair cell of the inner ear of the primate.

43. The method of claim 42, wherein the hair cell differentiation protein is selected from the group consisting of: Atoh1, Pou4f3, β -Catenin, Noggin, GFI-1, NTF3, and BDNF.

44. The method of claim 42, wherein the primate has previously been determined to have a defective hair cell differentiation gene.

45. A method of decreasing the expression level of a hair cell differentiation-suppressing protein in a supporting cell or a hair cell of an inner ear of a primate, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 35-37,

wherein the administering results in a decrease in the expression level of the hair cell differentiation-suppressing protein in the supporting cell or the hair cell of the inner ear of the primate.

46. A method of increasing the number of functional hair cells in a primate in need thereof, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-38.

47. A method of improving hearing in a primate in need thereof, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-38.

48. The method of claim 47, further comprising, prior to the administering step, determining that the primate has a defective hair cell differentiation gene.

49. A method of repairing a hair cell toxicity-inducing mutation in an endogenous hair cell differentiation gene locus in a supporting cell or a hair cell of an inner ear of a primate, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-32,

wherein the administering results in repair of the hair cell toxicity-inducing mutation in the endogenous hair cell differentiation gene locus in the supporting cell or the hair cell of the inner ear of the primate.

50. A method of decreasing the risk of hearing loss due to hair cell loss or dysfunction in a primate in need thereof, the method comprising administering to the inner ear of the primate a therapeutically effective amount of a composition of any one of claims 1-38.

51. The method of claim 50, wherein the primate has been previously identified as having a defective hair cell differentiation gene.

1/10

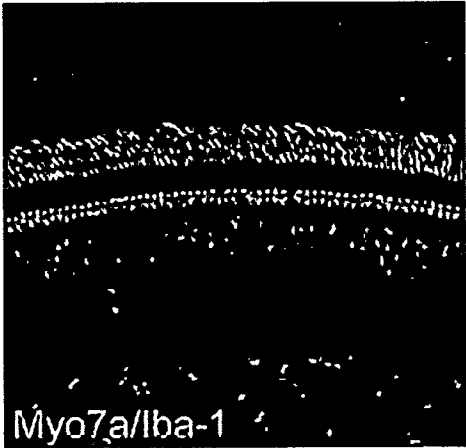


Figure 1A

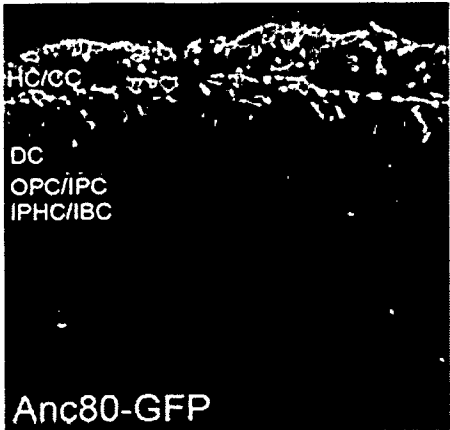


Figure 1B

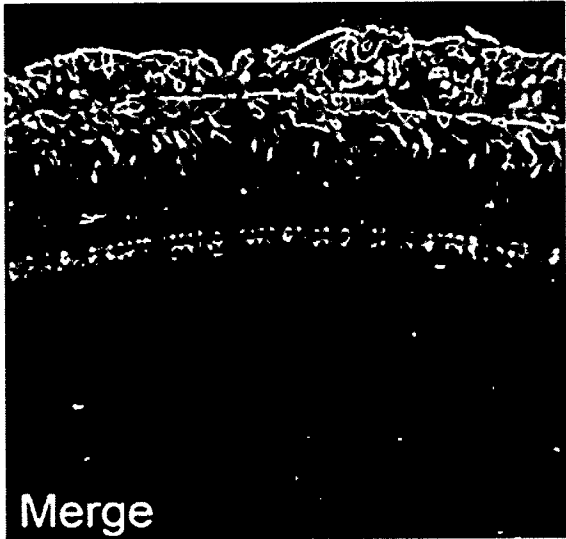


Figure 1C

2/10

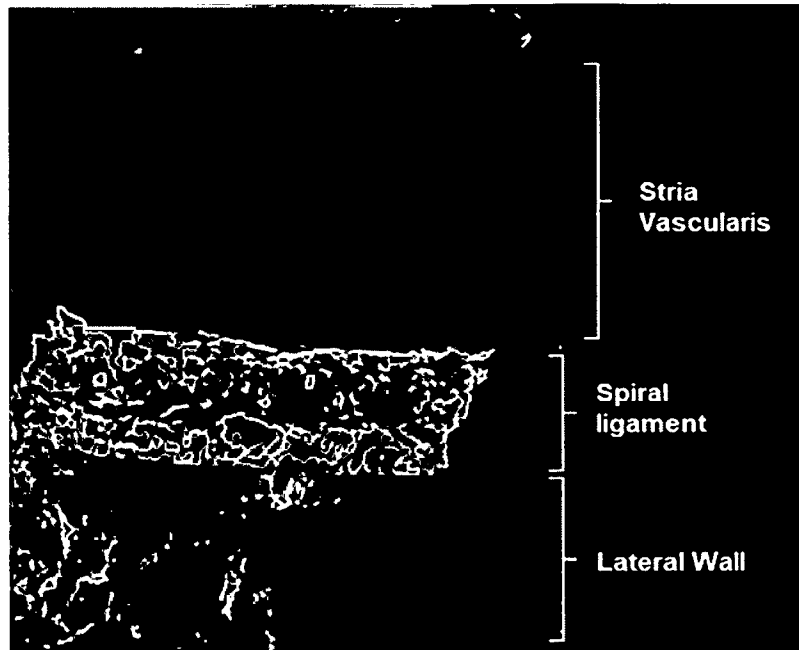


Figure 2A

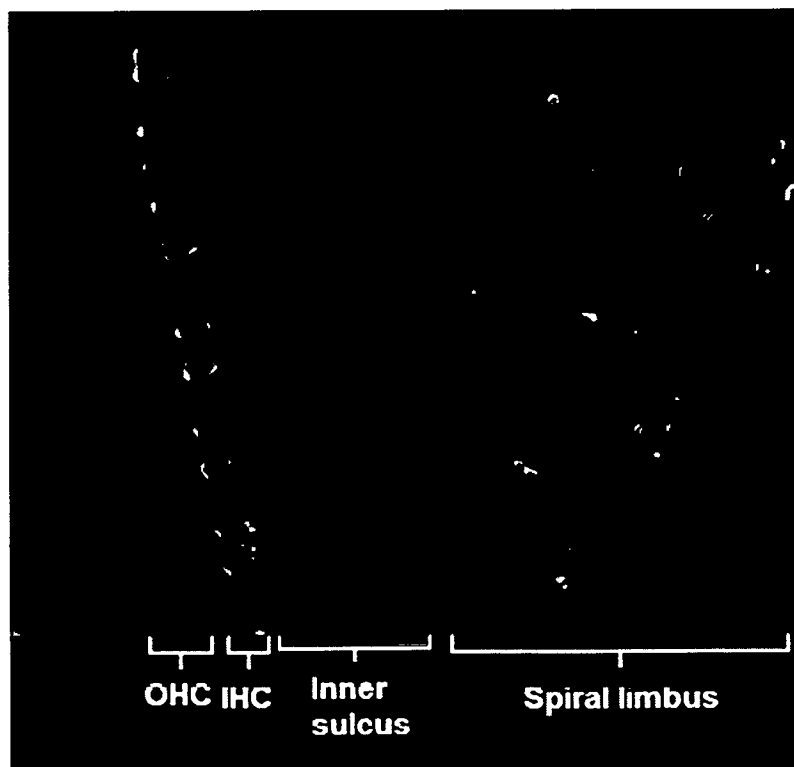


Figure 2B

3/10

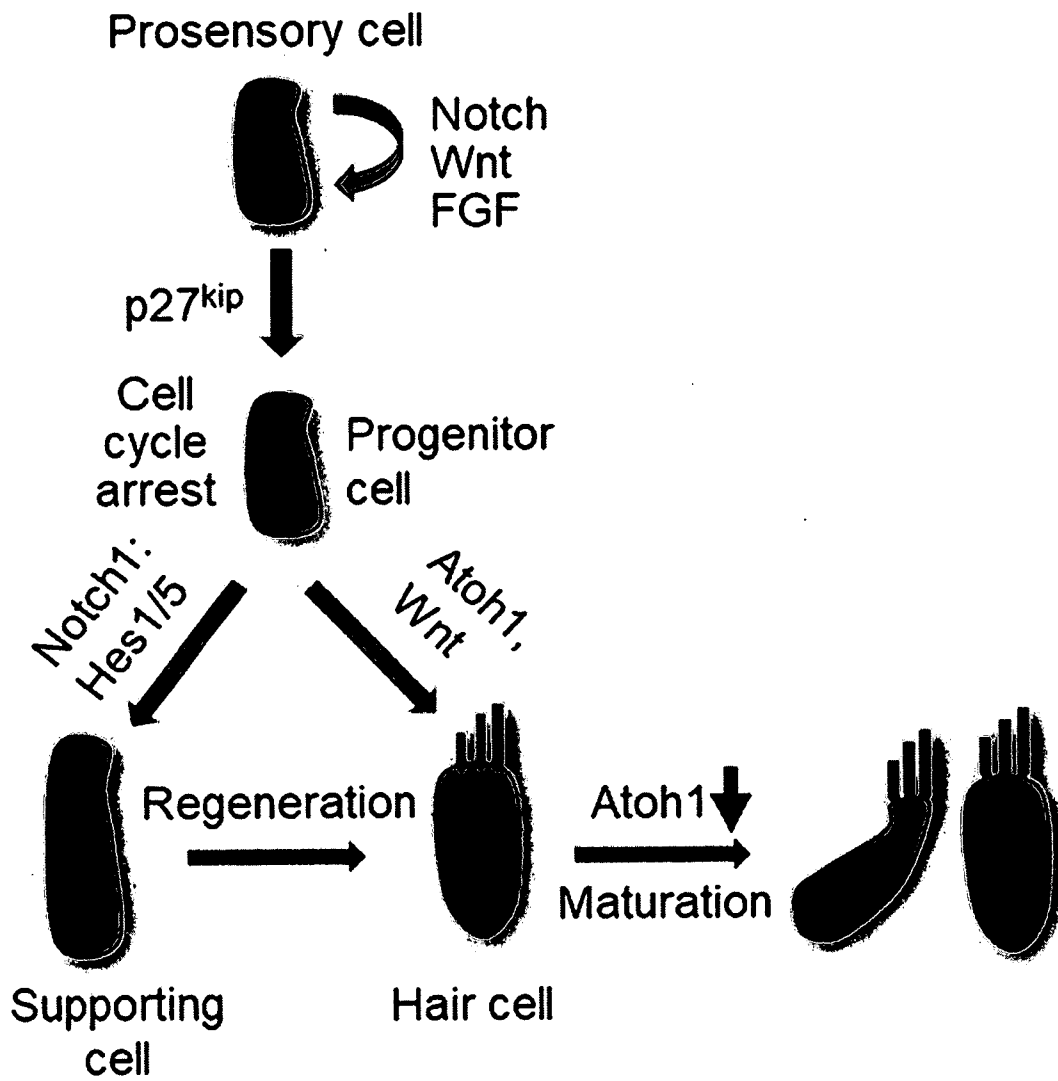


Figure 3

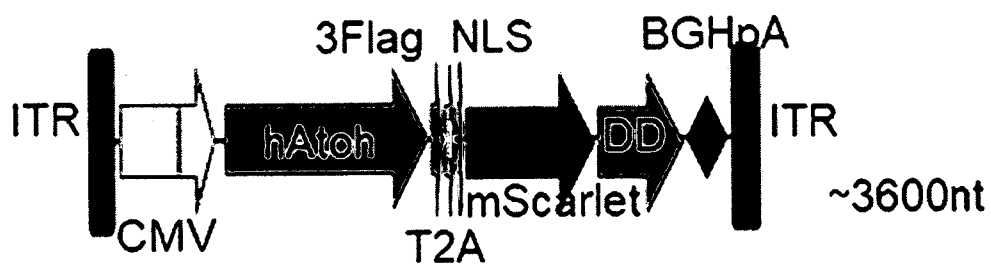


Figure 4A

4/10

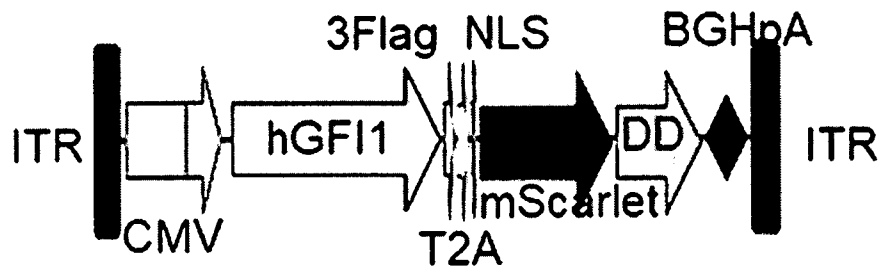


Figure 4B

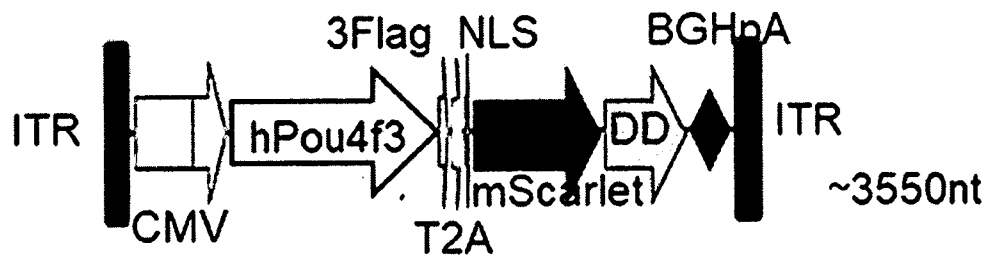


Figure 4C

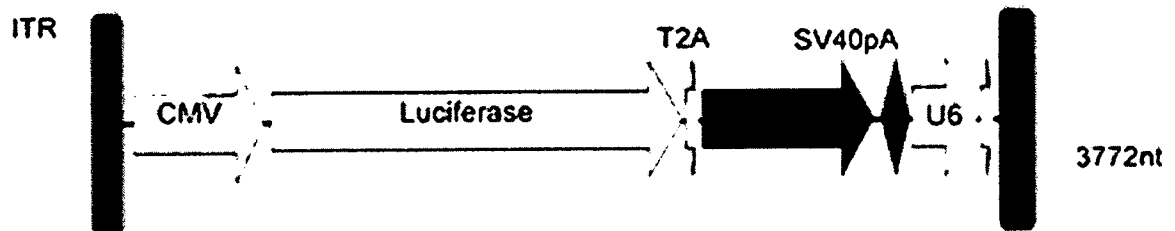


Figure 4D

5/10

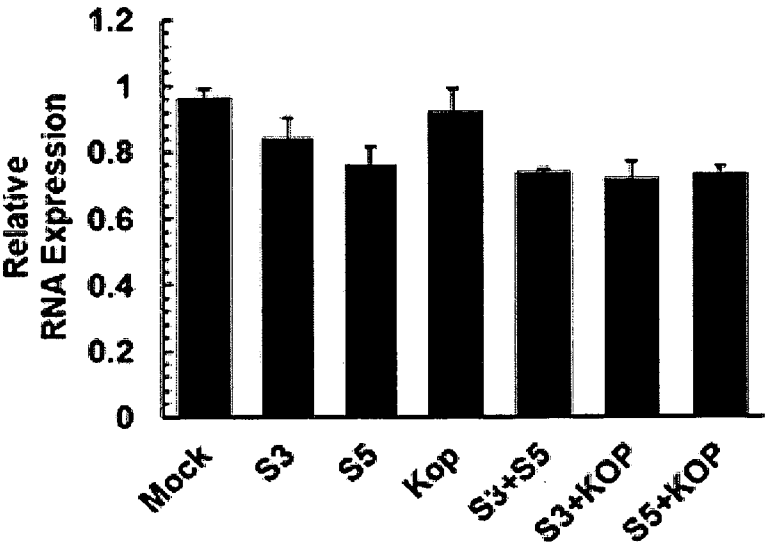


Figure 5A

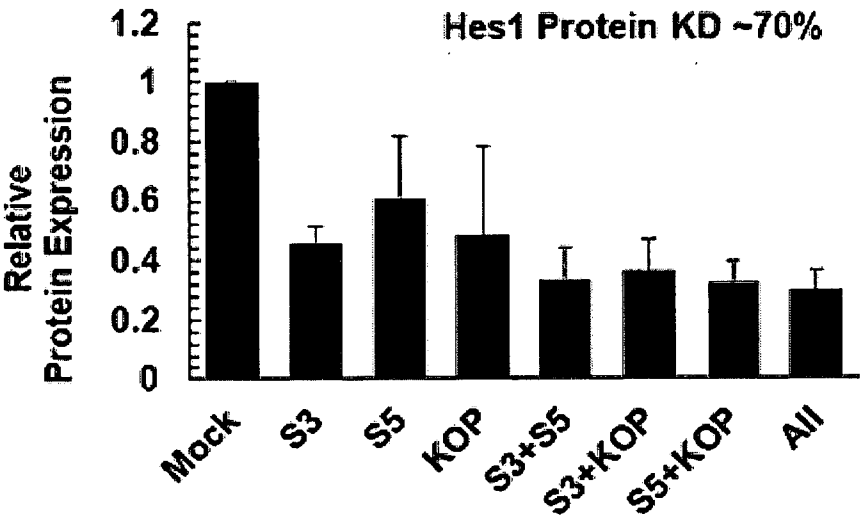


Figure 5B

6/10

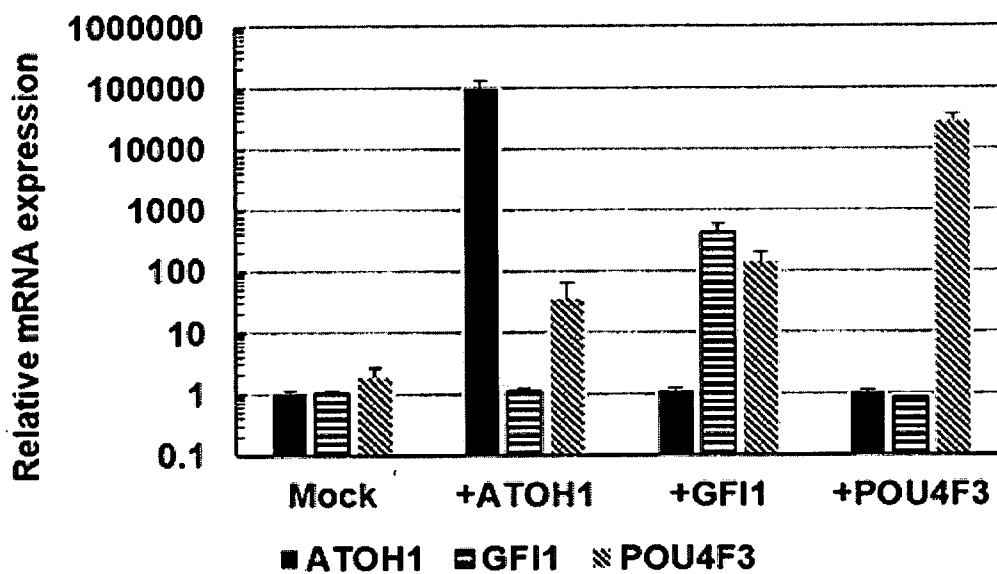
ATOH1, POU4F3 and GFI-1 overexpression in HEK293ft

Figure 6A

Mock +A +P +G +AP +AG +APG

GFI1
ATOH1
POU4F3

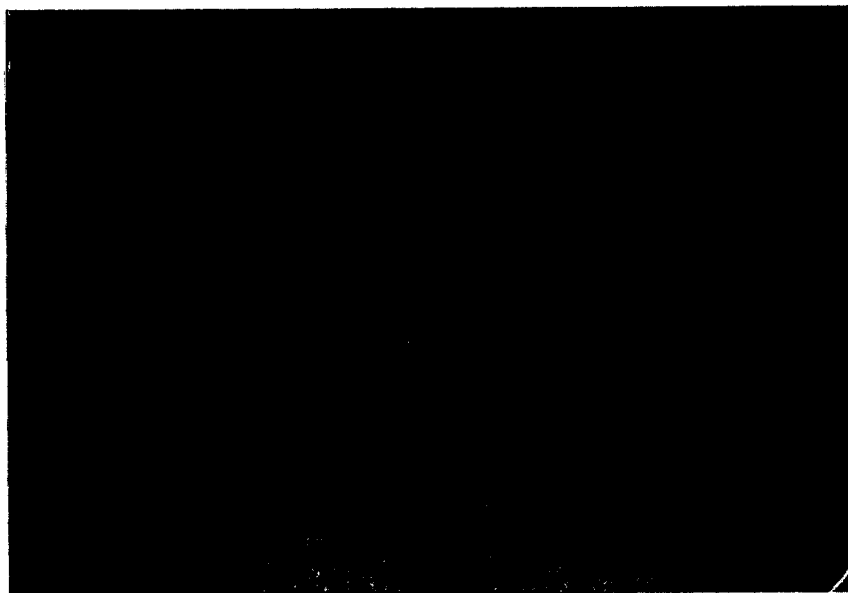


Figure 6B

7/10

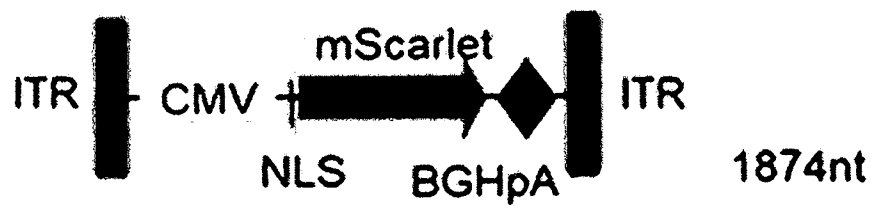


Figure 7A

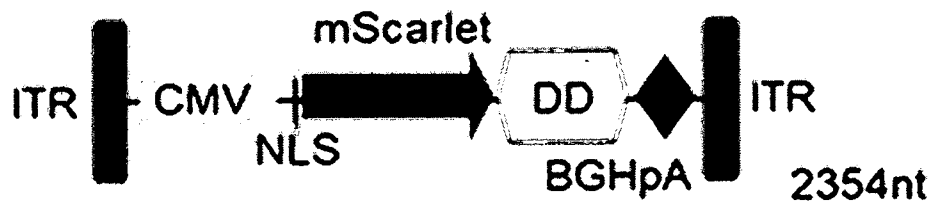


Figure 7B

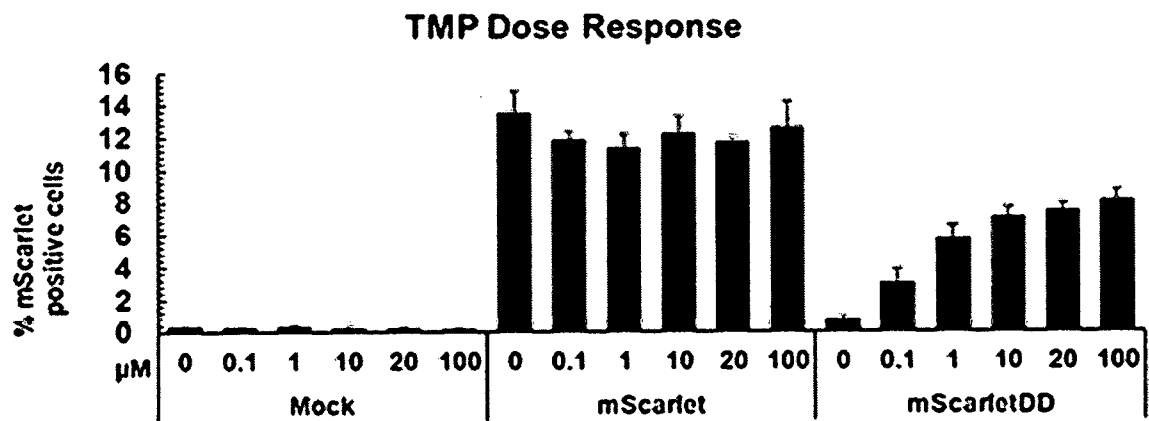


Figure 8A

8/10

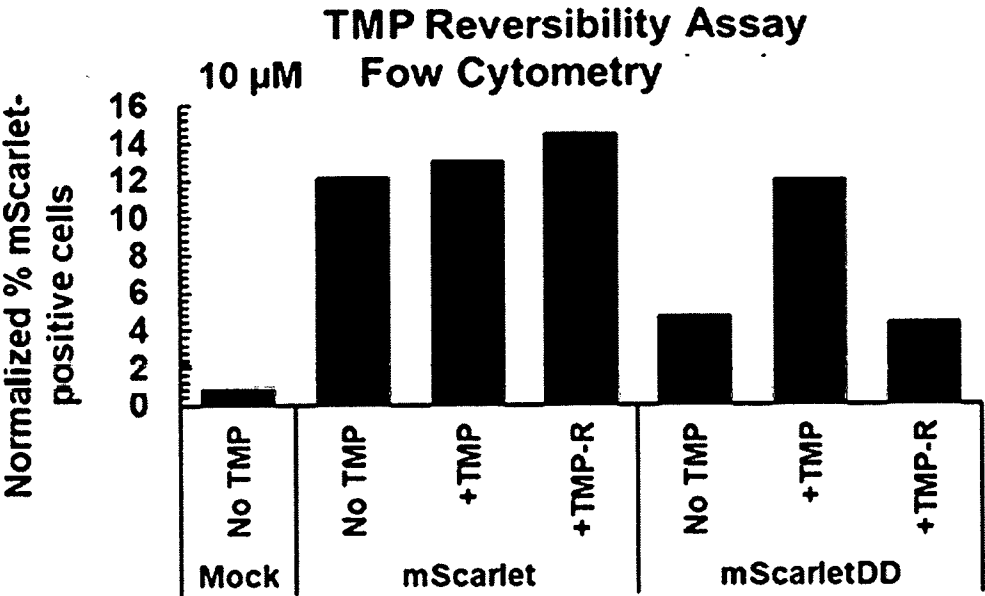


Figure 8B

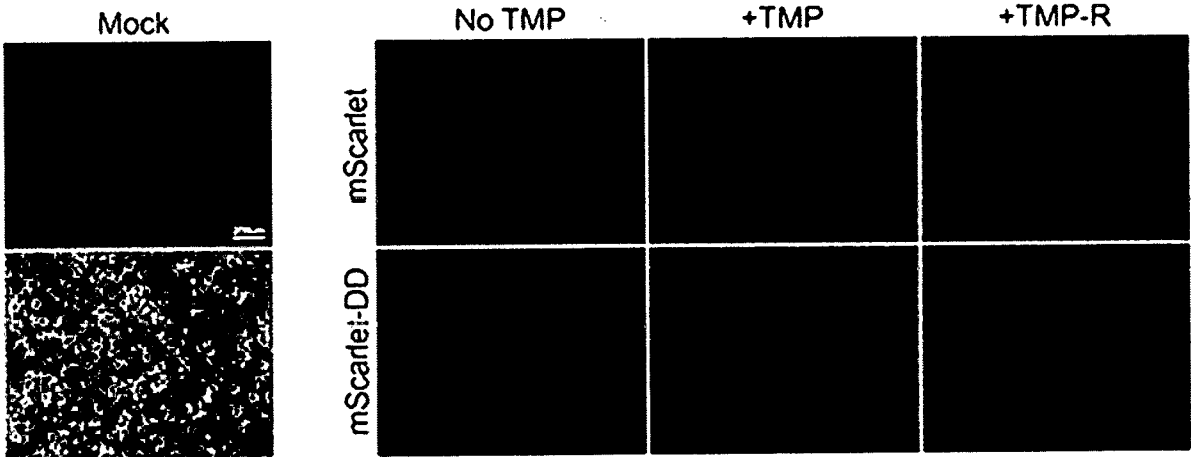


Figure 9A

Figure 9B

9/10

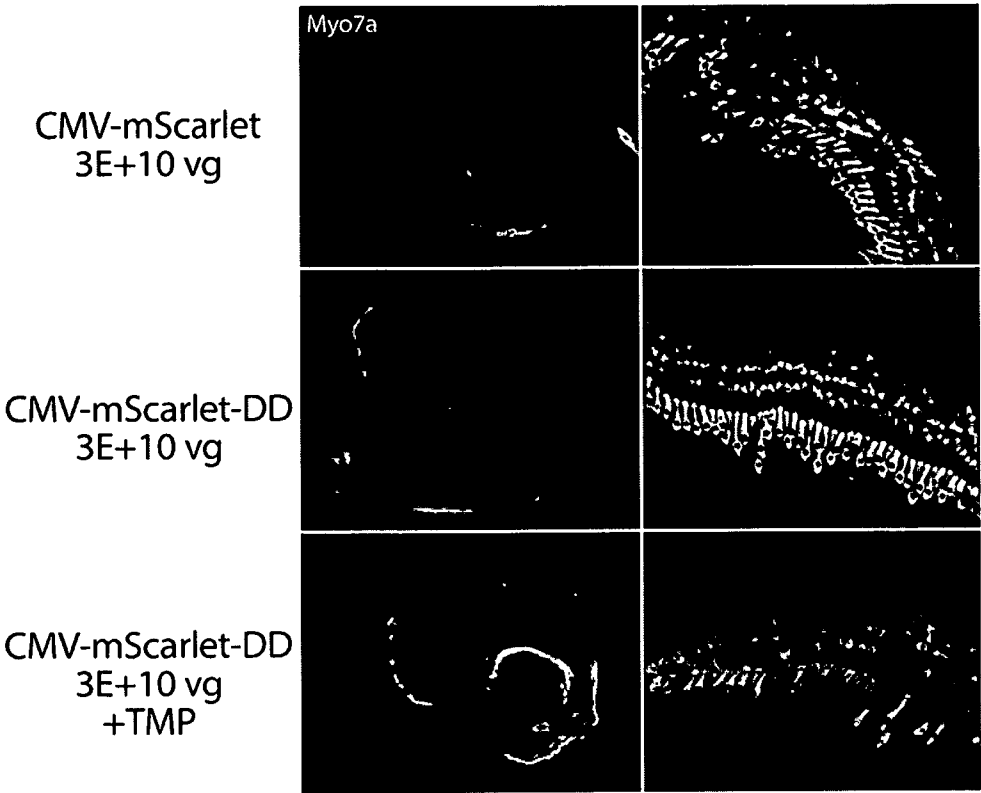


Figure 10

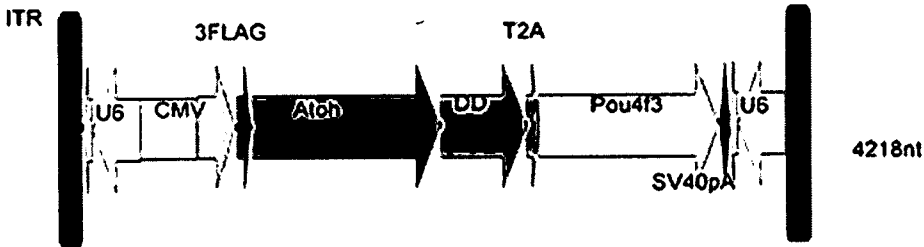


Figure 11A

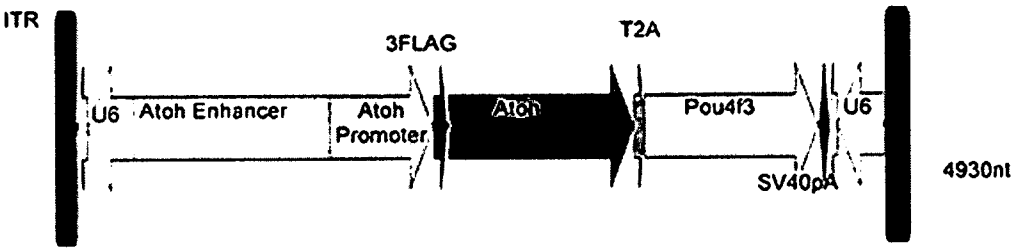


Figure 11B

10/10

Transfection of combined plasmid in HEK

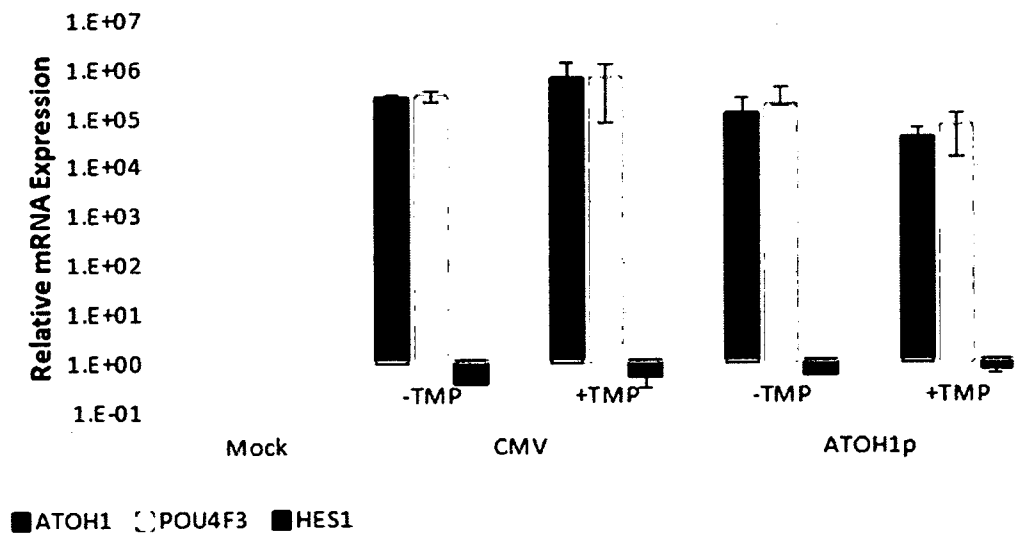


Figure 12A

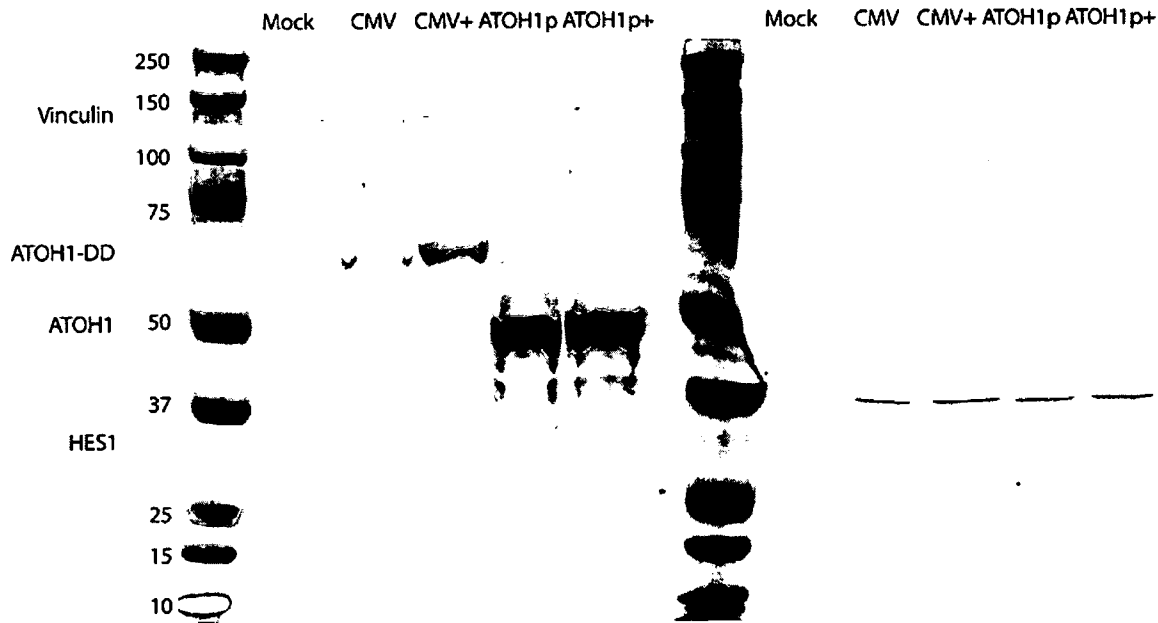


Figure 12B

SEQUENCE LISTING

<110> Akouos, Inc.

<120> COMPOSITIONS AND METHODS OF INDUCING DIFFERENTIATION OF A HAIR CELL

<130> 44628-0028W01

<150> US 62/888,105

<151> 2019-08-16

<150> US 62/756,910

<151> 2018-11-07

<160> 96

<170> PatentIn version 3.5

<210> 1

<211> 354

<212> PRT

<213> Homo sapiens

<400> 1

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1 5 10 15

Gly Asp His His Arg Gln Pro Gln Pro His His Leu Pro Gln Pro Pro
20 25 30

Pro Pro Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Glu His Pro Val
35 40 45

Tyr Pro Pro Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp
50 55 60

Leu Ala Pro Thr Leu Gln Gly Ile Cys Thr Ala Arg Ala Ala Gln Tyr
65 70 75 80

Leu Leu His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Arg
85 90 95

Asp Glu Val Asp Gly Arg Gly Glu Leu Val Arg Arg Ser Ser Gly Gly
100 105 110

Ala Ser Ser Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gln Leu
115 120 125

Cys Lys Leu Lys Gly Gly Val Val Val Asp Glu Leu Gly Cys Ser Arg
130 135 140

Gln Arg Ala Pro Ser Ser Lys Gln Val Asn Gly Val Gln Lys Gln Arg
145 150 155 160

Arg Leu Ala Ala Asn Ala Arg Glu Arg Arg Arg Met His Gly Leu Asn
165 170 175

His Ala Phe Asp Gln Leu Arg Asn Val Ile Pro Ser Phe Asn Asn Asp
180 185 190

Lys Lys Leu Ser Lys Tyr Glu Thr Leu Gln Met Ala Gln Ile Tyr Ile
195 200 205

Asn Ala Leu Ser Glu Leu Leu Gln Thr Pro Ser Gly Gly Glu Gln Pro
210 215 220

Pro Pro Pro Pro Ala Ser Cys Lys Ser Asp His His His Leu Arg Thr
225 230 235 240

Ala Ala Ser Tyr Glu Gly Gly Ala Gly Asn Ala Thr Ala Ala Gly Ala
245 250 255

Gln Gln Ala Ser Gly Gly Ser Gln Arg Pro Thr Pro Pro Gly Ser Cys
260 265 270

Arg Thr Arg Phe Ser Ala Pro Ala Ser Ala Gly Gly Tyr Ser Val Gln
275 280 285

Leu Asp Ala Leu His Phe Ser Thr Phe Glu Asp Ser Ala Leu Thr Ala
290 295 300

Met Met Ala Gln Lys Asn Leu Ser Pro Ser Leu Pro Gly Ser Ile Leu
305 310 315 320

Gln Pro Val Gln Glu Glu Asn Ser Lys Thr Ser Pro Arg Ser His Arg
325 330 335

Ser Asp Gly Glu Phe Ser Pro His Ser His Tyr Ser Asp Ser Asp Glu
340 345 350

Ala Ser

<210> 2
<211> 351
<212> PRT
<213> Mus musculus

<400> 2

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1 5 10 15

Gly Asp His His Arg His Pro Gln Pro His His Val Pro Pro Leu Thr
20 25 30

Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Asp Leu Pro Val Tyr Pro
35 40 45

Ala Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp Leu Thr
50 55 60

Pro Thr Leu Gln Gly Leu Cys Thr Ala Arg Ala Ala Gln Tyr Leu Leu
65 70 75 80

His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Arg Asp Glu
85 90 95

Ala Asp Ser Gln Gly Glu Leu Val Arg Arg Ser Gly Cys Gly Gly Leu
100 105 110

Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gln Leu Cys Lys Leu
115 120 125

Lys Gly Gly Val Val Val Asp Glu Leu Gly Cys Ser Arg Gln Arg Ala
130 135 140

Pro Ser Ser Lys Gln Val Asn Gly Val Gln Lys Gln Arg Arg Leu Ala

145		150		155		160									
Ala	Asn	Ala	Arg	Glu	Arg	Arg	Arg	Met	His	Gly	Leu	Asn	His	Ala	Phe
				165					170					175	
Asp	Gln	Leu	Arg	Asn	Val	Ile	Pro	Ser	Phe	Asn	Asn	Asp	Lys	Lys	Leu
			180					185					190		
Ser	Lys	Tyr	Glu	Thr	Leu	Gln	Met	Ala	Gln	Ile	Tyr	Ile	Asn	Ala	Leu
		195					200					205			
Ser	Glu	Leu	Leu	Gln	Thr	Pro	Asn	Val	Gly	Glu	Gln	Pro	Pro	Pro	Pro
	210					215					220				
Thr	Ala	Ser	Cys	Lys	Asn	Asp	His	His	His	Leu	Arg	Thr	Ala	Ser	Ser
225					230					235					240
Tyr	Glu	Gly	Gly	Ala	Gly	Ala	Ser	Ala	Val	Ala	Gly	Ala	Gln	Pro	Ala
				245					250					255	
Pro	Gly	Gly	Gly	Pro	Arg	Pro	Thr	Pro	Pro	Gly	Pro	Cys	Arg	Thr	Arg
			260					265					270		
Phe	Ser	Gly	Pro	Ala	Ser	Ser	Gly	Gly	Tyr	Ser	Val	Gln	Leu	Asp	Ala
		275					280					285			
Leu	His	Phe	Pro	Ala	Phe	Glu	Asp	Arg	Ala	Leu	Thr	Ala	Met	Met	Ala
	290					295					300				
Gln	Lys	Asp	Leu	Ser	Pro	Ser	Leu	Pro	Gly	Gly	Ile	Leu	Gln	Pro	Val
305					310					315					320
Gln	Glu	Asp	Asn	Ser	Lys	Thr	Ser	Pro	Arg	Ser	His	Arg	Ser	Asp	Gly
			325						330					335	
Glu	Phe	Ser	Pro	His	Ser	His	Tyr	Ser	Asp	Ser	Asp	Glu	Ala	Ser	
			340					345					350		

<210> 3
 <211> 351
 <212> PRT

<213> Rattus norvegicus

<400> 3

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1 5 10 15

Gly Asp His His Arg His Pro Gln Pro His His Ile Pro Gln Leu Thr
20 25 30

Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Asp His Pro Val Tyr Pro
35 40 45

Ala Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp Leu Thr
50 55 60

Pro Thr Leu Gln Gly Leu Cys Thr Ala Arg Ala Ala Gln Tyr Leu Leu
65 70 75 80

His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Gly Asp Glu
85 90 95

Ala Asp Gly Gln Gly Glu Leu Val Arg Arg Ser Gly Cys Gly Gly Leu
100 105 110

Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gln Leu Cys Lys Leu
115 120 125

Lys Gly Gly Val Val Val Asp Glu Leu Gly Cys Ser Arg Gln Arg Ala
130 135 140

Pro Ser Ser Lys Gln Val Asn Gly Val Gln Lys Gln Arg Arg Leu Ala
145 150 155 160

Ala Asn Ala Arg Glu Arg Arg Arg Met His Gly Leu Asn His Ala Phe
165 170 175

Asp Gln Leu Arg Asn Val Ile Pro Ser Phe Asn Asn Asp Lys Lys Leu
180 185 190

Ser Lys Tyr Glu Thr Leu Gln Met Ala Gln Ile Tyr Ile Asn Ala Leu
195 200 205

Ser Glu Leu Leu Gln Thr Pro Ser Val Gly Glu Gln Pro Pro Pro Pro
210 215 220

Ala Ala Ser Cys Lys Asn Asp His His His Leu Arg Ala Ala Ala Ser
225 230 235 240

Tyr Glu Gly Gly Ala Gly Ala Ser Ala Val Ala Gly Ala Gln Pro Ala
245 250 255

Pro Gly Gly Gly Pro Arg Pro Thr Pro Pro Gly Ala Cys Arg Thr Arg
260 265 270

Phe Ser Thr Pro Ala Ser Ser Gly Gly Tyr Ser Val Gln Leu Asp Ala
275 280 285

Leu His Phe Pro Ala Phe Glu Asp Arg Ala Leu Thr Ala Met Met Ala
290 295 300

Gln Lys Asp Leu Ser Pro Ser Leu Pro Gly Gly Ile Leu Gln Pro Val
305 310 315 320

Pro Glu Asp Ser Ser Lys Thr Ser Pro Arg Ser His Arg Ser Asp Gly
325 330 335

Glu Phe Ser Pro His Ser His Tyr Ser Asp Ser Asp Glu Ala Ser
340 345 350

<210> 4
<211> 1065
<212> DNA
<213> Homo sapiens

<400> 4
atgtcccgcc tgctgcatgc agaagagtgg gctgaagtga aggagttggg agaccaccat 60
cgccagcccc agccgcatca tctcccgcaa ccgccgccgc cgccgcagcc acctgcaact 120
ttgcaggcga gagagcatcc cgtctacccg cctgagctgt ccctcctgga cagcaccgac 180
ccacgcgcct ggctggctcc cactttgcag ggcatctgca cggcacgcgc cgcccagtat 240
ttgctacatt ccccgagct gggtgcctca gaggccgctg cgccccggga cgaggtggac 300

ggccgggggg agctggtaag gaggagcagc ggcggtgcca gcagcagcaa gagccccggg	360
ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg gggtaggtgt agacgagctg	420
ggctgcagcc gccaacgggc cccttcagc aaacaggtga atggggtgca gaagcagaga	480
cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac	540
cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagacc	600
ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gccagcgga	660
ggggaacagc caccgccgc tccagcctcc tgcaaaagcg accaccacca ctttcgcacc	720
gcggcctcct atgaaggggg cgcgggcaac gcgaccgcag ctggggctca gcaggcttcc	780
ggagggagcc agcggccgac ccgcccggg agttgccgga ctgccttctc agccccagct	840
tctgcgggag ggtactcggg gcagctggac gctctgcact tctcgacttt cgaggacagc	900
gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg	960
cagccagtgc aggaggaaaa cagcaaaact tcgcctcggg cccacagaag cgacggggaa	1020
ttttcccc attccatta cagtgactcg gatgaggcaa gttag	1065

<210> 5
 <211> 1065
 <212> DNA
 <213> Homo sapiens

<400> 5	
atgtcccgcc tgctgcatgc agaagagtgg gctgaagtga aggagtggg agaccacat	60
cgcagcccc agccgcatca tctcccga cgcgcgcgc cgcgcagcc acctgcaact	120
ttgaggcga gagagcatcc cgtctacccg cctgagctgt ccctcctgga cagcaccgac	180
ccacgcgcct ggctggctcc cactttgcag ggcattctgca cggcacgcgc cgcagtat	240
ttgctacatt ccccgagct ggggtgcctca gaggccgctg cgcggggga cgagggtggac	300
ggccgggggg agctggtaag gaggagcagc ggcggtgcca gcagcagcaa gagccccggg	360
ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg gggtaggtgt agacgagctg	420
ggctgcagcc gccaacgggc cccttcagc aaacaggtga atggggtgca gaagcagaga	480
cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac	540
cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagacc	600
ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gccagcgga	660

ggggaacagc caccgccgcc tccagcctcc tgcaaaagcg accaccacca ccttcgcacc	720
gcggcctcct atgaaggggg cgcgggcaac gcgaccgcag ctgggggtca gcaggcttcc	780
ggagggagcc agcggccgac cccgcccggg agttgccgga ctcgcttctc agccccagct	840
tctgcgggag ggtactcggg gcagctggac gctctgcact tctcgacttt cgaggacagc	900
gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg	960
cagccagtgc aggaggaaaa cagcaaaact tcgcctcggg cccacagaag cgacggggaa	1020
ttttcccccc attccatta cagtgactcg gatgaggcaa gttag	1065

<210> 6
 <211> 338
 <212> PRT
 <213> Homo sapiens

 <400> 6

Met Met Ala Met Asn Ser Lys Gln Pro Phe Gly Met His Pro Val Leu	
1 5 10 15	
Gln Glu Pro Lys Phe Ser Ser Leu His Ser Gly Ser Glu Ala Met Arg	
20 25 30	
Arg Val Cys Leu Pro Ala Pro Gln Leu Gln Gly Asn Ile Phe Gly Ser	
35 40 45	
Phe Asp Glu Ser Leu Leu Ala Arg Ala Glu Ala Leu Ala Ala Val Asp	
50 55 60	
Ile Val Ser His Gly Lys Asn His Pro Phe Lys Pro Asp Ala Thr Tyr	
65 70 75 80	
His Thr Met Ser Ser Val Pro Cys Thr Ser Thr Ser Ser Thr Val Pro	
85 90 95	
Ile Ser His Pro Ala Ala Leu Thr Ser His Pro His His Ala Val His	
100 105 110	
Gln Gly Leu Glu Gly Asp Leu Leu Glu His Ile Ser Pro Thr Leu Ser	
115 120 125	

Val Ser Gly Leu Gly Ala Pro Glu His Ser Val Met Pro Ala Gln Ile
130 135 140

His Pro His His Leu Gly Ala Met Gly His Leu His Gln Ala Met Gly
145 150 155 160

Met Ser His Pro His Thr Val Ala Pro His Ser Ala Met Pro Ala Cys
165 170 175

Leu Ser Asp Val Glu Ser Asp Pro Arg Glu Leu Glu Ala Phe Ala Glu
180 185 190

Arg Phe Lys Gln Arg Arg Ile Lys Leu Gly Val Thr Gln Ala Asp Val
195 200 205

Gly Ala Ala Leu Ala Asn Leu Lys Ile Pro Gly Val Gly Ser Leu Ser
210 215 220

Gln Ser Thr Ile Cys Arg Phe Glu Ser Leu Thr Leu Ser His Asn Asn
225 230 235 240

Met Ile Ala Leu Lys Pro Val Leu Gln Ala Trp Leu Glu Glu Ala Glu
245 250 255

Ala Ala Tyr Arg Glu Lys Asn Ser Lys Pro Glu Leu Phe Asn Gly Ser
260 265 270

Glu Arg Lys Arg Lys Arg Thr Ser Ile Ala Ala Pro Glu Lys Arg Ser
275 280 285

Leu Glu Ala Tyr Phe Ala Ile Gln Pro Arg Pro Ser Ser Glu Lys Ile
290 295 300

Ala Ala Ile Ala Glu Lys Leu Asp Leu Lys Lys Asn Val Val Arg Val
305 310 315 320

Trp Phe Cys Asn Gln Arg Gln Lys Gln Lys Arg Met Lys Tyr Ser Ala
325 330 335

Val His

<210> 7
<211> 338
<212> PRT
<213> Mus musculus

<400> 7

Met Met Ala Met Asn Ala Lys Gln Pro Phe Gly Met His Pro Val Leu
1 5 10 15

Gln Glu Pro Lys Phe Ser Ser Leu His Ser Gly Ser Glu Ala Met Arg
20 25 30

Arg Val Cys Leu Pro Ala Pro Gln Leu Gln Gly Asn Ile Phe Gly Ser
35 40 45

Phe Asp Glu Ser Leu Leu Ala Arg Ala Glu Ala Leu Ala Ala Val Asp
50 55 60

Ile Val Ser His Gly Lys Asn His Pro Phe Lys Pro Asp Ala Thr Tyr
65 70 75 80

His Thr Met Ser Ser Val Pro Cys Thr Ser Thr Ser Pro Thr Val Pro
85 90 95

Ile Ser His Pro Ala Ala Leu Thr Ser His Pro His His Ala Val His
100 105 110

Gln Gly Leu Glu Gly Asp Leu Leu Glu His Ile Ser Pro Thr Leu Ser
115 120 125

Val Ser Gly Leu Gly Ala Pro Glu His Ser Val Met Pro Ala Gln Ile
130 135 140

His Pro His His Leu Gly Ala Met Gly His Leu His Gln Ala Met Gly
145 150 155 160

Met Ser His Pro His Ala Val Ala Pro His Ser Ala Met Pro Ala Cys
165 170 175

Leu Ser Asp Val Glu Ser Asp Pro Arg Glu Leu Glu Ala Phe Ala Glu
180 185 190

Arg Phe Lys Gln Arg Arg Ile Lys Leu Gly Val Thr Gln Ala Asp Val
195 200 205

Gly Ala Ala Leu Ala Asn Leu Lys Ile Pro Gly Val Gly Ser Leu Ser
210 215 220

Gln Ser Thr Ile Cys Arg Phe Glu Ser Leu Thr Leu Ser His Asn Asn
225 230 235 240

Met Ile Ala Leu Lys Pro Val Leu Gln Ala Trp Leu Glu Glu Ala Glu
245 250 255

Ala Ala Tyr Arg Glu Lys Asn Ser Lys Pro Glu Leu Phe Asn Gly Ser
260 265 270

Glu Arg Lys Arg Lys Arg Thr Ser Ile Ala Ala Pro Glu Lys Arg Ser
275 280 285

Leu Glu Ala Tyr Phe Ala Ile Gln Pro Arg Pro Ser Ser Glu Lys Ile
290 295 300

Ala Ala Ile Ala Glu Lys Leu Asp Leu Lys Lys Asn Val Val Arg Val
305 310 315 320

Trp Phe Cys Asn Gln Arg Gln Lys Gln Lys Arg Met Lys Tyr Ser Ala
325 330 335

Val His

<210> 8

<211> 338

<212> PRT

<213> Rattus norvegicus

<400> 8

Met Met Ala Met Asn Ala Lys Gln Pro Phe Gly Met His Pro Val Leu
1 5 10 15

Gln Glu Pro Lys Phe Ser Ser Leu His Ser Gly Ser Glu Ala Met Arg
20 25 30

Arg Val Cys Leu Pro Ala Pro Gln Leu Gln Gly Asn Ile Phe Gly Ser
35 40 45

Phe Asp Glu Ser Leu Leu Ala Arg Ala Glu Ala Leu Ala Ala Val Asp
50 55 60

Ile Val Ser His Gly Lys Asn His Pro Phe Lys Pro Asp Ala Thr Tyr
65 70 75 80

His Thr Met Ser Ser Val Pro Cys Thr Ser Thr Ser Pro Thr Val Pro
85 90 95

Ile Ser His Pro Ala Ala Leu Thr Ser His Pro His His Pro Val His
100 105 110

Gln Gly Leu Glu Gly Asp Leu Leu Glu His Ile Ser Pro Thr Leu Ser
115 120 125

Val Ser Gly Leu Gly Ala Pro Glu His Ser Val Met Pro Ala Gln Ile
130 135 140

His Pro His His Leu Gly Ala Met Gly His Leu His Gln Ala Met Gly
145 150 155 160

Met Ser His Pro His Ala Val Ala Pro His Ser Ala Met Pro Ala Cys
165 170 175

Leu Ser Asp Val Glu Ser Asp Pro Arg Glu Leu Glu Ala Phe Ala Glu
180 185 190

Arg Phe Lys Gln Arg Arg Ile Lys Leu Gly Val Thr Gln Ala Asp Val
195 200 205

Gly Ala Ala Leu Ala Asn Leu Lys Ile Pro Gly Val Gly Ser Leu Ser
210 215 220

Gln Ser Thr Ile Cys Arg Phe Glu Ser Leu Thr Leu Ser His Asn Asn

225		230		235		240									
Met	Ile	Ala	Leu	Lys	Pro	Val	Leu	Gln	Ala	Trp	Leu	Glu	Glu	Ala	Glu
			245					250						255	
Ala	Ala	Tyr	Arg	Glu	Lys	Asn	Ser	Lys	Pro	Glu	Leu	Phe	Asn	Gly	Ser
			260					265					270		
Glu	Arg	Lys	Arg	Lys	Arg	Thr	Ser	Ile	Ala	Ala	Pro	Glu	Lys	Arg	Ser
		275					280					285			
Leu	Glu	Ala	Tyr	Phe	Ala	Ile	Gln	Pro	Arg	Pro	Ser	Ser	Glu	Lys	Ile
	290					295					300				
Ala	Ala	Ile	Ala	Glu	Lys	Leu	Asp	Leu	Lys	Lys	Asn	Val	Val	Arg	Val
305					310					315					320
Trp	Phe	Cys	Asn	Gln	Arg	Gln	Lys	Gln	Lys	Arg	Met	Lys	Tyr	Ser	Ala
			325					330						335	

Val His

<210> 9
 <211> 1017
 <212> DNA
 <213> Homo sapiens

<400> 9	
atgatggcca tgaactccaa gcagcctttc ggcatgcacc cggctgctgca agaaccctaaa	60
ttctccagtc tgcactctgg ctccgaggcc atgcgccgag tctgtctccc agccccgcag	120
ctgcagggtgta atatatttgg aagctttgat gagagcctgc tggcacgcgc cgaagctctg	180
gcggcggtgg atatcgtctc ccacggcaag aaccatccgt tcaagcccga cgccacctac	240
cataccatga gcagcgtgcc ctgcacgtcc acttcgtcca ccgtgcccac ctccccacca	300
gctgcgctca cctcacaccc tcaccacgcc gtgcaccagg gcctcgaagg cgacctgctg	360
gagcacatct cgcccacgct gagtgtgagc ggcctggggc ctccggaaca ctcggtgatg	420
cccgcacaga tccatccaca ccacctgggc gccatggggc acctgcacca ggccatgggc	480
atgagtcacc cgcacaccgt ggcccctcat agcgccatgc ctgcatgcct cagcgacgtg	540

gagtcagacc cgcgcgagct ggaagccttc gccgagcgct tcaagcagcg gcgcatcaag	600
ctgggggtga cccaggcgga cgtgggcgcg gctctggcta atctcaagat ccccggcgtg	660
ggctcgctga gccaaagcac catctgcagg ttcgagtctc tcaactctctc gcacaacaac	720
atgatcgctc tcaagccggt gctccaggcc tggttggagg aggccgaggc cgcctaccga	780
gagaagaaca gcaagccaga gctcttcaac ggcagcgaac ggaagcgcaa acgcacgtcc	840
atcgcggcgc cggagaagcg ttcactcgag gcctatttcg ctatccagcc acgtccttca	900
tctgagaaga tcgcggccat cgctgagaaa ctggacctta aaaagaacgt ggtgagagtc	960
tggttctgca accagagaca gaaacagaaa cgaatgaagt attcggctgt ccactga	1017

<210> 10
 <211> 1497
 <212> DNA
 <213> Homo sapiens

<400> 10	
cgctgagcag cgctcacttg gagagcggca agcaagctag acaagcctga ttccatgtca	60
cccgctgcca ccctgccagg agcgcgaaga tgatggccat gaactccaag cagcctttcg	120
gcatgcaccc ggtgctgcaa gaacccaaat tctccagtct gcaactctggc tccgaggcca	180
tgcgccgagt ctgtctccca gccccgcagg tacgtagtgg agcataatta ccgctctaag	240
gcacatTTTT tgacaggcac tagcttcatg tttttttcat gtcgcccaga acaatcgccg	300
ctgtctgaac ccctctcctt gtctcccccg cgttctctcc cggcgcgctc tctctctcat	360
tcatgtctct gatccacacg tctgttccag cagagccgct gcctccgtat taatTTTTat	420
gacctgggct ttgaggagag gcatctcggt tgcttgaaaa tgtgttttaa tcctgtgttg	480
acagtattcc ctactgaccg tgctgtgcgc cttctcgctt gcagctgcag ggtaatatat	540
ttggaagctt tgatgagagc ctgctggcac gcgccgaagc tctggcggcg gtggatatcg	600
tctcccacgg caagaacat ccgttcaagc ccgacgccac ctaccatacc atgagcagcg	660
tgccctgcac gtccacttcg tccaccgtgc ccatctccca cccagctgcg ctcacctcac	720
accctcacca cgccgtgcac cagggcctcg aaggcgacct gctggagcac atctcgccca	780
cgctgagtgt gagcggcctg ggcgctccgg aacactcggt gatgcccga cagatccatc	840
cacaccacct gggcgccatg ggccacctgc accaggccat gggcatgagt caccgcaca	900

```

ccgtggcccc tcatagcgcc atgcctgcat gcctcagcga cgtggagtca gacccgcgcg      960
agctggaagc cttcgccgag cgcttcaagc agcggcgcat caagctgggg gtgacccagg      1020
cggacgtggg cgcggtcttg gctaattctca agatccccgg cgtgggctcg ctgagccaaa      1080
gcaccatctg caggttcgag tctctcactc tctcgacaaa caacatgatc gctctcaagc      1140
cggtgctcca ggcctggttg gaggaggccg aggccgccta ccgagagaag aacagcaagc      1200
cagagctctt caacggcagc gaacggaagc gcaaacgcac gtccatcgcg gcgccggaga      1260
agcgttcact cgaggcctat ttcgctatcc agccacgtcc ttcacatgag aagatcgcg      1320
ccatcgctga gaaactggac cttaaaaaga acgtggtgag agtctggttc tgcaaccaga      1380
gacagaaaca gaaacgaatg aagtattcgg ctgtccactg attgcggcag ggcgcagcgt      1440
cgggagccgg gagagcctag tgctcatccc tcccgggttc gggggatggt tatcggg      1497

```

```

<210> 11
<211> 781
<212> PRT
<213> Homo sapiens

```

```
<400> 11
```

```

Met Ala Thr Gln Ala Asp Leu Met Glu Leu Asp Met Ala Met Glu Pro
1           5           10          15

```

```

Asp Arg Lys Ala Ala Val Ser His Trp Gln Gln Gln Ser Tyr Leu Asp
          20          25          30

```

```

Ser Gly Ile His Ser Gly Ala Thr Thr Thr Ala Pro Ser Leu Ser Gly
          35          40          45

```

```

Lys Gly Asn Pro Glu Glu Glu Asp Val Asp Thr Ser Gln Val Leu Tyr
          50          55          60

```

```

Glu Trp Glu Gln Gly Phe Ser Gln Ser Phe Thr Gln Glu Gln Val Ala
          65          70          75          80

```

```

Asp Ile Asp Gly Gln Tyr Ala Met Thr Arg Ala Gln Arg Val Arg Ala
          85          90          95

```

```

Ala Met Phe Pro Glu Thr Leu Asp Glu Gly Met Gln Ile Pro Ser Thr
          100         105         110

```

Gln Phe Asp Ala Ala His Pro Thr Asn Val Gln Arg Leu Ala Glu Pro
115 120 125

Ser Gln Met Leu Lys His Ala Val Val Asn Leu Ile Asn Tyr Gln Asp
130 135 140

Asp Ala Glu Leu Ala Thr Arg Ala Ile Pro Glu Leu Thr Lys Leu Leu
145 150 155 160

Asn Asp Glu Asp Gln Val Val Val Asn Lys Ala Ala Val Met Val His
165 170 175

Gln Leu Ser Lys Lys Glu Ala Ser Arg His Ala Ile Met Arg Ser Pro
180 185 190

Gln Met Val Ser Ala Ile Val Arg Thr Met Gln Asn Thr Asn Asp Val
195 200 205

Glu Thr Ala Arg Cys Thr Ala Gly Thr Leu His Asn Leu Ser His His
210 215 220

Arg Glu Gly Leu Leu Ala Ile Phe Lys Ser Gly Gly Ile Pro Ala Leu
225 230 235 240

Val Lys Met Leu Gly Ser Pro Val Asp Ser Val Leu Phe Tyr Ala Ile
245 250 255

Thr Thr Leu His Asn Leu Leu Leu His Gln Glu Gly Ala Lys Met Ala
260 265 270

Val Arg Leu Ala Gly Gly Leu Gln Lys Met Val Ala Leu Leu Asn Lys
275 280 285

Thr Asn Val Lys Phe Leu Ala Ile Thr Thr Asp Cys Leu Gln Ile Leu
290 295 300

Ala Tyr Gly Asn Gln Glu Ser Lys Leu Ile Ile Leu Ala Ser Gly Gly
305 310 315 320

Pro Gln Ala Leu Val Asn Ile Met Arg Thr Tyr Thr Tyr Glu Lys Leu
325 330 335

Leu Trp Thr Thr Ser Arg Val Leu Lys Val Leu Ser Val Cys Ser Ser
340 345 350

Asn Lys Pro Ala Ile Val Glu Ala Gly Gly Met Gln Ala Leu Gly Leu
355 360 365

His Leu Thr Asp Pro Ser Gln Arg Leu Val Gln Asn Cys Leu Trp Thr
370 375 380

Leu Arg Asn Leu Ser Asp Ala Ala Thr Lys Gln Glu Gly Met Glu Gly
385 390 395 400

Leu Leu Gly Thr Leu Val Gln Leu Leu Gly Ser Asp Asp Ile Asn Val
405 410 415

Val Thr Cys Ala Ala Gly Ile Leu Ser Asn Leu Thr Cys Asn Asn Tyr
420 425 430

Lys Asn Lys Met Met Val Cys Gln Val Gly Gly Ile Glu Ala Leu Val
435 440 445

Arg Thr Val Leu Arg Ala Gly Asp Arg Glu Asp Ile Thr Glu Pro Ala
450 455 460

Ile Cys Ala Leu Arg His Leu Thr Ser Arg His Gln Glu Ala Glu Met
465 470 475 480

Ala Gln Asn Ala Val Arg Leu His Tyr Gly Leu Pro Val Val Val Lys
485 490 495

Leu Leu His Pro Pro Ser His Trp Pro Leu Ile Lys Ala Thr Val Gly
500 505 510

Leu Ile Arg Asn Leu Ala Leu Cys Pro Ala Asn His Ala Pro Leu Arg
515 520 525

Glu Gln Gly Ala Ile Pro Arg Leu Val Gln Leu Leu Val Arg Ala His
530 535 540

Gln Asp Thr Gln Arg Arg Thr Ser Met Gly Gly Thr Gln Gln Gln Phe
545 550 555 560

Val Glu Gly Val Arg Met Glu Glu Ile Val Glu Gly Cys Thr Gly Ala
565 570 575

Leu His Ile Leu Ala Arg Asp Val His Asn Arg Ile Val Ile Arg Gly
580 585 590

Leu Asn Thr Ile Pro Leu Phe Val Gln Leu Leu Tyr Ser Pro Ile Glu
595 600 605

Asn Ile Gln Arg Val Ala Ala Gly Val Leu Cys Glu Leu Ala Gln Asp
610 615 620

Lys Glu Ala Ala Glu Ala Ile Glu Ala Glu Gly Ala Thr Ala Pro Leu
625 630 635 640

Thr Glu Leu Leu His Ser Arg Asn Glu Gly Val Ala Thr Tyr Ala Ala
645 650 655

Ala Val Leu Phe Arg Met Ser Glu Asp Lys Pro Gln Asp Tyr Lys Lys
660 665 670

Arg Leu Ser Val Glu Leu Thr Ser Ser Leu Phe Arg Thr Glu Pro Met
675 680 685

Ala Trp Asn Glu Thr Ala Asp Leu Gly Leu Asp Ile Gly Ala Gln Gly
690 695 700

Glu Pro Leu Gly Tyr Arg Gln Asp Asp Pro Ser Tyr Arg Ser Phe His
705 710 715 720

Ser Gly Gly Tyr Gly Gln Asp Ala Leu Gly Met Asp Pro Met Met Glu
725 730 735

His Glu Met Gly Gly His His Pro Gly Ala Asp Tyr Pro Val Asp Gly
740 745 750

Leu Pro Asp Leu Gly His Ala Gln Asp Leu Met Asp Gly Leu Pro Pro
755 760 765

Gly Asp Ser Asn Gln Leu Ala Trp Phe Asp Thr Asp Leu
770 775 780

<210> 12
<211> 781
<212> PRT
<213> Mus musculus

<400> 12

Met Ala Thr Gln Ala Asp Leu Met Glu Leu Asp Met Ala Met Glu Pro
1 5 10 15

Asp Arg Lys Ala Ala Val Ser His Trp Gln Gln Gln Ser Tyr Leu Asp
20 25 30

Ser Gly Ile His Ser Gly Ala Thr Thr Thr Ala Pro Ser Leu Ser Gly
35 40 45

Lys Gly Asn Pro Glu Glu Glu Asp Val Asp Thr Ser Gln Val Leu Tyr
50 55 60

Glu Trp Glu Gln Gly Phe Ser Gln Ser Phe Thr Gln Glu Gln Val Ala
65 70 75 80

Asp Ile Asp Gly Gln Tyr Ala Met Thr Arg Ala Gln Arg Val Arg Ala
85 90 95

Ala Met Phe Pro Glu Thr Leu Asp Glu Gly Met Gln Ile Pro Ser Thr
100 105 110

Gln Phe Asp Ala Ala His Pro Thr Asn Val Gln Arg Leu Ala Glu Pro
115 120 125

Ser Gln Met Leu Lys His Ala Val Val Asn Leu Ile Asn Tyr Gln Asp
130 135 140

Asp Ala Glu Leu Ala Thr Arg Ala Ile Pro Glu Leu Thr Lys Leu Leu
145 150 155 160

Asn Asp Glu Asp Gln Val Val Val Asn Lys Ala Ala Val Met Val His
165 170 175

Gln Leu Ser Lys Lys Glu Ala Ser Arg His Ala Ile Met Arg Ser Pro
180 185 190

Gln Met Val Ser Ala Ile Val Arg Thr Met Gln Asn Thr Asn Asp Val
195 200 205

Glu Thr Ala Arg Cys Thr Ala Gly Thr Leu His Asn Leu Ser His His
210 215 220

Arg Glu Gly Leu Leu Ala Ile Phe Lys Ser Gly Gly Ile Pro Ala Leu
225 230 235 240

Val Lys Met Leu Gly Ser Pro Val Asp Ser Val Leu Phe Tyr Ala Ile
245 250 255

Thr Thr Leu His Asn Leu Leu Leu His Gln Glu Gly Ala Lys Met Ala
260 265 270

Val Arg Leu Ala Gly Gly Leu Gln Lys Met Val Ala Leu Leu Asn Lys
275 280 285

Thr Asn Val Lys Phe Leu Ala Ile Thr Thr Asp Cys Leu Gln Ile Leu
290 295 300

Ala Tyr Gly Asn Gln Glu Ser Lys Leu Ile Ile Leu Ala Ser Gly Gly
305 310 315 320

Pro Gln Ala Leu Val Asn Ile Met Arg Thr Tyr Thr Tyr Glu Lys Leu
325 330 335

Leu Trp Thr Thr Ser Arg Val Leu Lys Val Leu Ser Val Cys Ser Ser
340 345 350

Asn Lys Pro Ala Ile Val Glu Ala Gly Gly Met Gln Ala Leu Gly Leu
355 360 365

His Leu Thr Asp Pro Ser Gln Arg Leu Val Gln Asn Cys Leu Trp Thr

370

375

380

Leu Arg Asn Leu Ser Asp Ala Ala Thr Lys Gln Glu Gly Met Glu Gly
 385 390 395 400

Leu Leu Gly Thr Leu Val Gln Leu Leu Gly Ser Asp Asp Ile Asn Val
 405 410 415

Val Thr Cys Ala Ala Gly Ile Leu Ser Asn Leu Thr Cys Asn Asn Tyr
 420 425 430

Lys Asn Lys Met Met Val Cys Gln Val Gly Gly Ile Glu Ala Leu Val
 435 440 445

Arg Thr Val Leu Arg Ala Gly Asp Arg Glu Asp Ile Thr Glu Pro Ala
 450 455 460

Ile Cys Ala Leu Arg His Leu Thr Ser Arg His Gln Glu Ala Glu Met
 465 470 475 480

Ala Gln Asn Ala Val Arg Leu His Tyr Gly Leu Pro Val Val Val Lys
 485 490 495

Leu Leu His Pro Pro Ser His Trp Pro Leu Ile Lys Ala Thr Val Gly
 500 505 510

Leu Ile Arg Asn Leu Ala Leu Cys Pro Ala Asn His Ala Pro Leu Arg
 515 520 525

Glu Gln Gly Ala Ile Pro Arg Leu Val Gln Leu Leu Val Arg Ala His
 530 535 540

Gln Asp Thr Gln Arg Arg Thr Ser Met Gly Gly Thr Gln Gln Gln Phe
 545 550 555 560

Val Glu Gly Val Arg Met Glu Glu Ile Val Glu Gly Cys Thr Gly Ala
 565 570 575

Leu His Ile Leu Ala Arg Asp Val His Asn Arg Ile Val Ile Arg Gly
 580 585 590

Leu Asn Thr Ile Pro Leu Phe Val Gln Leu Leu Tyr Ser Pro Ile Glu
595 600 605

Asn Ile Gln Arg Val Ala Ala Gly Val Leu Cys Glu Leu Ala Gln Asp
610 615 620

Lys Glu Ala Ala Glu Ala Ile Glu Ala Glu Gly Ala Thr Ala Pro Leu
625 630 635 640

Thr Glu Leu Leu His Ser Arg Asn Glu Gly Val Ala Thr Tyr Ala Ala
645 650 655

Ala Val Leu Phe Arg Met Ser Glu Asp Lys Pro Gln Asp Tyr Lys Lys
660 665 670

Arg Leu Ser Val Glu Leu Thr Ser Ser Leu Phe Arg Thr Glu Pro Met
675 680 685

Ala Trp Asn Glu Thr Ala Asp Leu Gly Leu Asp Ile Gly Ala Gln Gly
690 695 700

Glu Ala Leu Gly Tyr Arg Gln Asp Asp Pro Ser Tyr Arg Ser Phe His
705 710 715 720

Ser Gly Gly Tyr Gly Gln Asp Ala Leu Gly Met Asp Pro Met Met Glu
725 730 735

His Glu Met Gly Gly His His Pro Gly Ala Asp Tyr Pro Val Asp Gly
740 745 750

Leu Pro Asp Leu Gly His Ala Gln Asp Leu Met Asp Gly Leu Pro Pro
755 760 765

Gly Asp Ser Asn Gln Leu Ala Trp Phe Asp Thr Asp Leu
770 775 780

<210> 13
<211> 781
<212> PRT
<213> Rattus norvegicus

<400> 13

Met Ala Thr Gln Ala Asp Leu Met Glu Leu Asp Met Ala Met Glu Pro
1 5 10 15

Asp Arg Lys Ala Ala Val Ser His Trp Gln Gln Gln Ser Tyr Leu Asp
20 25 30

Ser Gly Ile His Ser Gly Ala Thr Thr Thr Ala Pro Ser Leu Ser Gly
35 40 45

Lys Gly Asn Pro Glu Glu Glu Asp Val Asp Thr Ser Gln Val Leu Tyr
50 55 60

Glu Trp Glu Gln Gly Phe Ser Gln Ser Phe Thr Gln Glu Gln Val Ala
65 70 75 80

Asp Ile Asp Gly Gln Tyr Ala Met Thr Arg Ala Gln Arg Val Arg Ala
85 90 95

Ala Met Phe Pro Glu Thr Leu Asp Glu Gly Met Gln Ile Pro Ser Thr
100 105 110

Gln Phe Asp Ala Ala His Pro Thr Asn Val Gln Arg Leu Ala Glu Pro
115 120 125

Ser Gln Met Leu Lys His Ala Val Val Asn Leu Ile Asn Tyr Gln Asp
130 135 140

Asp Ala Glu Leu Ala Thr Arg Ala Ile Pro Glu Leu Thr Lys Leu Leu
145 150 155 160

Asn Asp Glu Asp Gln Val Val Val Asn Lys Ala Ala Val Met Val His
165 170 175

Gln Leu Ser Lys Lys Glu Ala Ser Arg His Ala Ile Met Arg Ser Pro
180 185 190

Gln Met Val Ser Ala Ile Val Arg Thr Met Gln Asn Thr Asn Asp Val
195 200 205

Glu Thr Ala Arg Cys Thr Ala Gly Thr Leu His Asn Leu Ser His His
210 215 220

Arg Glu Gly Leu Leu Ala Ile Phe Lys Ser Gly Gly Ile Pro Ala Leu
225 230 235 240

Val Lys Met Leu Gly Ser Pro Val Asp Ser Val Leu Phe Tyr Ala Ile
245 250 255

Thr Thr Leu His Asn Leu Leu Leu His Gln Glu Gly Ala Lys Met Ala
260 265 270

Val Arg Leu Ala Gly Gly Leu Gln Lys Met Val Ala Leu Leu Asn Lys
275 280 285

Thr Asn Val Lys Phe Leu Ala Ile Thr Thr Asp Cys Leu Gln Ile Leu
290 295 300

Ala Tyr Gly Asn Gln Glu Ser Lys Leu Ile Ile Leu Ala Ser Gly Gly
305 310 315 320

Pro Gln Ala Leu Val Asn Ile Met Arg Thr Tyr Thr Tyr Glu Lys Leu
325 330 335

Leu Trp Thr Thr Ser Arg Val Leu Lys Val Leu Ser Val Cys Ser Ser
340 345 350

Asn Lys Pro Ala Ile Val Glu Ala Gly Gly Met Gln Ala Leu Gly Pro
355 360 365

His Leu Thr Asp Pro Ser Gln Arg Leu Val Gln Asn Cys Leu Trp Thr
370 375 380

Leu Arg Asn Leu Ser Asp Ala Ala Thr Lys Gln Glu Gly Met Glu Gly
385 390 395 400

Leu Leu Gly Thr Leu Val Gln Leu Leu Gly Ser Asp Asp Ile Asn Val
405 410 415

Val Thr Cys Ala Ala Gly Ile Leu Ser Asn Leu Thr Cys Asn Asn Tyr
420 425 430

Lys Asn Lys Met Met Val Cys Gln Val Gly Gly Ile Glu Ala Leu Val
435 440 445

Arg Thr Val Leu Arg Ala Gly Asp Arg Glu Asp Ile Thr Glu Pro Ala
450 455 460

Ile Cys Ala Leu Arg His Leu Thr Ser Arg His Gln Glu Ala Glu Met
465 470 475 480

Ala Gln Asn Ala Val Arg Leu His Tyr Gly Leu Pro Val Val Val Lys
485 490 495

Leu Leu His Pro Pro Ser His Trp Pro Leu Ile Lys Ala Thr Val Gly
500 505 510

Leu Ile Arg Asn Leu Ala Leu Cys Pro Ala Asn His Ala Pro Leu Arg
515 520 525

Glu Gln Gly Ala Ile Pro Arg Leu Val Gln Leu Leu Val Arg Ala His
530 535 540

Gln Asp Thr Gln Arg Arg Thr Ser Met Gly Gly Thr Gln Gln Gln Phe
545 550 555 560

Val Glu Gly Val Arg Met Glu Glu Ile Val Glu Gly Cys Thr Gly Ala
565 570 575

Leu His Ile Leu Ala Arg Asp Val His Asn Arg Ile Val Ile Arg Gly
580 585 590

Leu Asn Thr Ile Pro Leu Phe Val Gln Leu Leu Tyr Ser Pro Ile Glu
595 600 605

Asn Ile Gln Arg Val Ala Ala Gly Val Leu Cys Glu Leu Ala Gln Asp
610 615 620

Lys Glu Ala Ala Glu Ala Ile Glu Ala Glu Gly Ala Thr Ala Pro Leu
625 630 635 640

Thr Glu Leu Leu His Ser Arg Asn Glu Gly Val Ala Thr Tyr Ala Ala
645 650 655

Ala Val Leu Phe Arg Met Ser Glu Asp Lys Pro Gln Asp Tyr Lys Lys
660 665 670

Arg Leu Ser Val Glu Leu Thr Ser Ser Leu Phe Arg Thr Glu Pro Met
675 680 685

Ala Trp Asn Glu Thr Ala Asp Leu Gly Leu Asp Ile Gly Ala Gln Gly
690 695 700

Glu Ala Leu Gly Tyr Arg Gln Asp Asp Pro Ser Tyr Arg Ser Phe His
705 710 715 720

Ser Gly Gly Tyr Gly Gln Asp Ala Leu Gly Met Asp Pro Met Met Glu
725 730 735

His Glu Met Gly Gly His His Pro Gly Ala Asp Tyr Pro Val Asp Gly
740 745 750

Leu Pro Asp Leu Gly His Ala Gln Asp Leu Met Asp Gly Leu Pro Pro
755 760 765

Gly Asp Ser Asn Gln Leu Ala Trp Phe Asp Thr Asp Leu
770 775 780

<210> 14
<211> 2346
<212> DNA
<213> Homo sapiens

<400> 14
atggctactc aagctgattt gatggagttg gacatggcca tggaaccaga cagaaaagcg 60
gctgttagtc actggcagca acagtcttac ctggactctg gaatccattc tggtgccact 120
accacagctc cttctctgag tggtaaaggc aatcctgagg aagaggatgt ggatacctcc 180
caagtcctgt atgagtggga acagggattt tctcagtcct tcactcaaga acaagtagct 240
gatattgatg gacagtatgc aatgactcga gctcagaggg tacgagctgc tatgttcct 300
gagacattag atgagggcat gcagatccca tctacacagt ttgatgctgc tcatcccact 360

aatgtccagc gtttggctga accatcacag atgctgaaac atgcagttgt aaacttgatt	420
aactatcaag atgatgcaga acttgccaca cgtgcaatcc ctgaactgac aaaactgcta	480
aatgacgagg accagggtggt ggttaataag gctgcagtta tgggtccatca gcttttctaaa	540
aaggaagctt ccagacacgc tatcatgcgt tctcctcaga tgggtgtctgc tattgtacgt	600
accatgcaga atacaaatga tgtagaaaca gctcgttgta ccgctgggac cttgcataac	660
ctttcccatc atcgtgaggg cttactggcc atctttaagt ctggaggcat tcctgccctg	720
gtgaaaatgc ttggttcacc agtggattct gtgttgtttt atgccattac aactctccac	780
aaccttttat tacatcaaga aggagctaaa atggcagtgct gtttagctgg tgggctgcag	840
aaaatggttg ctttgctcaa caaaacaaat gttaaattct tggctattac gacagactgc	900
cttcaaattt tagcttatgg caaccaagaa agcaagctca tcatactggc tagtggtgga	960
ccccaagctt tagtaaatat aatgaggacc tatacttacg aaaaactact gtggaccaca	1020
agcagagtgc tgaaggtgct atctgtctgc tctagtaata agccggctat tgtagaagct	1080
ggtggaatgc aagctttagg acttcacctg acagatccaa gtcaacgtct tgttcagaac	1140
tgtctttgga ctctcaggaa tctttcagat gctgcaacta aacaggaagg gatggaaggt	1200
ctccttgga ctcttgttca gcttctgggt tcagatgata taaatgtggt cacctgtgca	1260
gctggaattc tttctaacct cacttgcaat aattataaga acaagatgat ggtctgccaa	1320
gtgggtggta tagaggctct tgtgcgtact gtccttcggg ctggtgacag ggaagacatc	1380
actgagcctg ccatctgtgc tcttcgtcat ctgaccagcc gacaccaaga agcagagatg	1440
gcccagaatg cagttcgcct tcactatgga ctaccagttg tggttaagct cttacacca	1500
ccatcccact ggcctctgat aaaggctact gttggattga ttcgaaatct tgccctttgt	1560
cccgcaaadc atgcaccttt gcgtgagcag ggtgccattc cacgactagt tcagttgctt	1620
gttcgtgcac atcaggatac ccagcgccgt acgtccatgg gtgggacaca gcagcaattt	1680
gtggaggggg tccgcatgga agaaatagtt gaaggttgta ccggagccct tcacatccta	1740
gctcgggatg ttcacaaccg aattgttatc agaggactaa ataccattcc attgtttgtg	1800
cagctgcttt attctcccat tgaaaacatc caaagagtag ctgcaggggt cctctgtgaa	1860
cttgctcagg acaaggaagc tgcagaagct attgaagctg agggagccac agctcctctg	1920
acagagttac ttcactctag gaatgaaggt gtggcgacat atgcagctgc tgttttgttc	1980

cgaatgtctg aggacaagcc acaagattac aagaaacggc tttcagttga gctgaccagc	2040
tctctcttca gaacagagcc aatggcttgg aatgagactg ctgatcttgg acttgatatt	2100
ggtgcccagg gagaaccctt tggatatcgc caggatgatc ctagctatcg ttcttttcac	2160
tctggtggat atggccagga tgccttgggt atggaccca tgatggaaca tgagatgggt	2220
ggccaccacc ctggtgctga ctatccagtt gatgggctgc cagatctggg gcatgcccag	2280
gacctcatgg atgggctgcc tccaggtgac agcaatcagc tggcctgggt tgatactgac	2340
ctgtaa	2346

<210> 15
 <211> 40998
 <212> DNA
 <213> Homo sapiens

<400> 15	
aggatacagc ggcttctgcg cgacttataa gagctccttg tgcggcgcca ttttaagcct	60
ctcggctctgt ggcagcagcg ttggcccggc cccgggagcg gagagcgagg ggaggcggag	120
acggaggaag gtctgaggag cagcttcagt ccccgccgag ccgccaccgc aggtcgagga	180
cggtcggact cccgcggcgg gaggagcctg ttcccctgag gtgcttgggc gtcctttcc	240
ttatccttcc ggggctgctc ccgcttcctc tcggagccaa acttcgtagc aggcgcgcgg	300
tccgggcggc gggctgggcg cagccgggag gcctgggggtt gggagcgggg agctcagggtg	360
ggggacggtg aggggtgggc gcgcccgggg cgcgaggggc ggcggccggg cccgggttcc	420
ggtcgcgctg cctctctggg gccctggggg catcgcttgc ggggaggggg cgccgcgggg	480
gcgcgtacag gagcccggat ggcaggcggg gtgggggtgg ggggtgggggt ctgtggtttc	540
cgtccggggc tctggccttg gccgagtttg ggggagggac ccggtgcctc gggatgcgcc	600
gggccctggg tggggggcgg ggtggggacg gggggctccg ctttctcagc tcttgcggcg	660
agttgggggt cgggcgctga ggcagagacg ccaccctaag tcccatcagt cctggggatc	720
ggaccagtgg actttctctt aagatttcct ctttcattct taagaataga agtggttatta	780
ttttttttaa tgccctggct atgtgagttt gaatcgaagc aactttaaac cttagagcaa	840
ctaaactcta agtgcagcgg gtgcgatgcg tcagtagggg gagcacataa aaaatccatg	900
tcttgcacct gtatttttagc gtactatgca ggtgagtga agcagtggat aatgtactgg	960
gagtcttatg gatttatggg agtgggtatg agaccctggg gaaataaggg ggtggaggaa	1020

ggcgaaggtg atggcttact gtttcttacc aagtgaactg caggattcag cctctgactc	1080
agaccgcttc gagaattttg ttcgtagaaa taattttaa attttgaatg	1140
cagctaaaa atgaattatag agcacgtttt cttttcagcg gaggtaattt ttccttcgct	1200
ccaaagctgg ccaaatggaa ttcaagcatt gcaacttctt tcagtgtttt gtctggagag	1260
aggactttga accgagactt ttcgaagtta agttcctata gcctgcttct gaatctgcca	1320
agcttgaaag ctttggcagt tgggtgtatg tagttgttgc cttcgttctc ttcccttttg	1380
gaggagcgt tgtctctac tttgtatctt ccagacatct gtggtcttcc cccaccct	1440
cgagtttg agtggtgaat gaagaaagac taggctgctg gtatgcagag gtcggcaaaa	1500
ggaaatcgag gagggtttt agtgaaatga gagctttgta tcatgaataa tgggtggctta	1560
ggctagacat caacttgaag agacggcagc atttcctttc ataaagtcta ggctaattgt	1620
tttcagatcg ctaagttgta gtttgtctgg aatttaggaa gccatttcag tatttgtcac	1680
ttggtgaacg aacattcaat accttcagat gtcttcgtgt tgacttgat tcatcctaag	1740
aaatagtaaa tatagtctca agtggtttt atgttatact gctggtttat tctctgctta	1800
aattattgac ataaatttct actttggagg cttttcgttt gaactaaggc tgtgcggaat	1860
ttattttact tttatattta aatctttgaa aaatctctga ttaaaaaaaaa agtaccctta	1920
aaggtttgag gatgtccttt cacaccagac aaaatttggg taatttgcgc ccaatattca	1980
ttactttgac ctaacctttg ttctgaaggc cgtgtacaag gacaaggccc tgagattatt	2040
gcaacagtaa cttgaaaaac tttcagaagt ctattctgta ggattaaagg aatgctgaga	2100
ctattcaagt ttgaagtcct ggggggtgggg aaaaataaaa aacctgtgct agaaagctta	2160
gtatagcatg taactttaga gtcctgtgga gtcctgagtc tcccacagac cagaacagtc	2220
atttaaaagt tttcaggaaa aaccaactta aaaaaaata aggtggctaa ttaaaaaaaaa	2280
atgaagcatt taacagtgtt caggtttcag agtatggaag aggggttttt taaactgtta	2340
tctgattatt tcttttacca acatgatata gaaaagtgtt tttccagtat taaaatttat	2400
cagactgagc ttactgttcc tgttaatgac tggaataaaa attggcataa atgagggtct	2460
gtatgcttgt ttttaataca ccaccaccaa gatagaaaac gaggaggcaa gtttctcaa	2520
gggtattttg aaatgtgtta gcaaaactat tgcagatact cgtttttgtt atagggtgag	2580
gtggggagag gcgcatgcta agtattgttg aaactaggga tgtagagaat taaaagtttg	2640

aatataatta	ttttgtagtt	ataagtagca	gtgaaattaa	atctcctgca	atagactata	2700
gaagtatatt	tagccaaatg	aaacttcagt	gttattgaaa	tgaaataata	catctgtcct	2760
gttacaagat	tatTTTTtatt	tctcttgtgg	tttcctagct	tctgataatc	aataattgta	2820
gatgagtagg	tggtaaagttt	taagtttgta	ctttgagctt	agtcggaagc	atgcttgact	2880
gccaacccgg	ggcacaaaagg	atgaaggctt	ttagaactgg	acaaacttct	aacaaaaggt	2940
atttgcaact	cttttgtagt	gtgtcatgtt	gatttgtgac	attgtttttg	aaaatatgtg	3000
ttaacttagt	tttcttgtag	ccctcttttt	attggaactg	tggtatctat	tgttgaaact	3060
gcttgactga	gaacatTTTT	ataccataaa	agtaaatagt	aaacatagcc	caggagcggc	3120
ttctggtttg	tccatcgtat	gtagccattg	cctccttgta	ctctcattga	gaagatactg	3180
atttgcagat	tcagttgtcc	ttctctaaca	gactatTTtT	gtaatatTgc	agttgtgatt	3240
gtgataggta	agtggaaccag	tcggttaaaa	taaatactca	ggtttcacaa	aaggaaaata	3300
atatgatttg	tgttgatcta	aatgagtata	ggagttaact	cctatagttt	ttcatcactt	3360
aaactcaggg	gaaagtTctt	tatttcctct	gtttacttaa	gaatgctgct	tttgtgtttc	3420
atgcaagact	gagcttgact	cagtttgaaa	cctaggctca	tctgttgagg	cctgaaccct	3480
gctgtccttg	aagtatgcat	ataatttgct	tccttcctaa	ggaaaaataa	gctcttgaaa	3540
gataaagtca	atcacattag	gaaccatttt	ttagggttta	gccactTTTT	TTTTTTTTTT	3600
TTTTtaactc	atgggcatct	cttctgttaa	gagacattcc	ccactctcca	agtttccctc	3660
aagcctgaag	cagcagagtg	agtagtgttg	gagcatgttt	tcattgcatg	cttgggtcat	3720
gttgagtgcc	ctccagtgga	tatagtataa	tgcttgatgat	TTTTTTTTTT	ttaattccaa	3780
acaagtttat	gtgggatata	tttaggaata	gttctgatga	gggagaatca	actaagaaac	3840
ctttgatttc	taaaataatt	aatatcatta	ctgctaatta	aaatacaggc	ttgagaaaat	3900
gtcttctcag	ccaatatttg	cagtagaaaa	gtcgggaggt	TTTTtaaggt	cactttgagt	3960
aggcagttct	gcttaaatat	atcataatga	taaaccagaa	tctcagtata	gtactttagg	4020
aggtaaaaga	tcataatatt	cagtttatatt	gatgaattac	agcaactgaa	attctcagaa	4080
aaaaattaat	gaaaatgtga	attgtcaatt	tgtctaaaaat	cattcacaga	gtaaaacata	4140
agtgctcaac	ttgattatat	taggaaatag	atagaaataa	aggtaattga	gccagtgtat	4200
gtgacctaaa	atataatgcc	cttagtgacc	atagggttgg	tctcatttgt	acatagtggg	4260

gggcatgat gaactgtgtt ttgccctttg aatttttcct taaaaagctt tctctaggct	4320
cctatgttca tggtttttct gttagtaata ttattttctg aaaatccatg tttcaaatca	4380
gaatctaatt agcaacagga atgaagctta ttctaaatta gtttttggaa gttaaacggt	4440
cagcatatgg aaatttttca gggtttagat ttttaaaaat ttgtttttca gaatatgttg	4500
ctggaatgaa aacgttagcg tagggacgga aaatgacact taccagtgat tgctttactt	4560
tgcctgtgga attcagtgtg attttgtgga aacattggta tatgattttt tactacttaa	4620
gaaatgtatt gctatagtta gggttttttt ttttttaaag gcaagaatgc ctcaagtgc	4680
ttatgtgaat gattatttca ggatggatta aatattcctc catcaaggac catacttgta	4740
aatcagtgat ttccaagttg gtgcttagta tttacagcat ttactgtcta taagcttctg	4800
ttctgatttt tcaagagttt tctgagaaat gagagtaggc ttaaaagtgc tttgaaaaat	4860
tatgtacata caacttactg aaaaaaattg ctaccgggga cttaatattgt ctcttgaaat	4920
gggctacttg ctttcattaa tgtagcatac tacaatttga tgttcaagat atgttactaa	4980
gaataagatc gctttcagaa gccttatata ggattggctt tactacattg tagtgggaat	5040
ggctactcaa atgtctccag ggccagttag gtattgggta aatgggacca tgcagactat	5100
taaaaattga agtgcacatg aagcagccag tcataagcag ctccagccac tgtgtgggaa	5160
tatagtttat gttgccagat catctgattt ctttccccta agtgggaaat ccagatcaat	5220
gtacatctct tgatttgcaa gtgttggtga acaaaattca tattttaaga tgctgtattc	5280
agcaciaaatt aaatacactt atttgctgaa tactgccagt ttgtccctct gcagtagtac	5340
catttgaagt acagtgtttt cataatgatt ctgtgaaatg actgggttctg tgaatgtaca	5400
taatttagca gataacattg ttaaattatt aggtttgtat ttatttaggc acttgggaaa	5460
tgccttgtgt caattgatta tagattagga gcttaaaagc aagatttata ttatcaactt	5520
atttgtgaag actgggaaac ccacattttt aaagttagga attaagatgg ccaggttcaa	5580
ggaaaagggg gagaagtaac tttcttatta ctcaaccatc ttaaatagag ttctttaagt	5640
gtatttttaa gaggtctcaa aacttaattc gaagggacgt caaatgctgg acaaattctg	5700
tgtatacaac tcaagtcagc ccccaatttt actggtcttt aaatcatgtc ctttttacca	5760
gaagtttgca tttctaagct aaactattac tgtttagacta gatccaaaac ttaaaaacag	5820
tttaggtaat taaaaattaa ttgaatataa acgttttact taaattaatg gcaaatggct	5880

ttttggccaa ttttaagttta tgtaggcagt taaatcgatt ttggttaaat cttttgctgc	5940
taacaaggta tttccagatt ttgaaaagtg gggtaggcctg gtgcctgtag taccagcact	6000
ttgggaggct ggggagggtg gatcacctga ggtagcaggagt tcgagactag cctggccgac	6060
gtggtgaata caaaaattag ccaggcatgg tggcagggtgc ctgtgatccc agctgcttgg	6120
aagtctgaag catgagaatt gcttgaacct gggaagcgga ggtagcagtg agctgagatc	6180
acgccactgc actccagctg gggcaacaga gcgagactcc atctcaagaa agaaaagtgg	6240
ggtagtttagt cttcaaaactc cgtgtttaag tgactggagt gaaaatgtaa atcataggcc	6300
ggtagttggtt taaaaagcat catctgaaaa taatgctgta gtctgcaatt atttttatta	6360
cgatacgatg gtgtaaaata caagcagatc agtgaacctat tcatgaaaca ttaatcctaa	6420
aggcgtctca cccaagtct atcccacaat ctccatgaga cttcgtggaa ccactgtaaa	6480
gtttcttgtg taatatccca gaagtttcct acctctggta tcttttgaac ttgttgaaaa	6540
ggcttttcca cccctcttt atgatggttt gaagagtgtg aacatctgaa tgatgctggg	6600
gtgaaactgc ttcataacac ttccattttc tcccctattt atttccatat ttttattttt	6660
tcactaatat cccacgggtt ttacttctgt tttagtaatt cacatgttgc tggactaatt	6720
ctttttaact gacttgtaac agatatgtta aaccgtttta aacttggggg gtatttttaa	6780
cctactttta gttagttcaa gttaatcagt ctacatggca tataaacctt atgattaata	6840
aatcttaaat gctggtagct gagttggaag ccaaagacgt acaaaaaagc tgaagtgtta	6900
ggtttagtgt gataagcttc tcttactaac agggttttgt aatagcagaa atagatatat	6960
gcatatatat gtgcatatat atagcatacc ttattggatg tccatataaa aatgtgtaag	7020
aagttaaatt tactgcaaaa tttcttggga gtgcaatttg aagatgatct taagtggatg	7080
tagtagtttg ctacactggg ggatagttgt tgcaaactgc tcctaatttt cttttactgt	7140
gaagtaaact gaacagctgt aatagggatt aggaactgta ctccctctct ctctttttta	7200
agtataatta agtggttttg gggtaagggt gtagggagtg agtgtctttg aagttttgca	7260
tatactagat gaatgccaca tgtataaggg aggaacaagg gattcttgga aatatttttc	7320
aatccaagta actttggagg cttccaagtg gagttcattc ccctgtgtag gaaagtgctg	7380
gggtagacct ttaaattcct ttctgagcca ttgaaagaat gtcctcaaac ttcgcttata	7440
ctttatagtt catttagata caaaagttac aaactgaatg ctatttagga aacgtaatac	7500

actgacatac cgctctttaa atagattata aatttagtat atcaattttc tggcattttg	7560
ctgaatttta ttgttttagtt ttcaagccca actatcttgt tactttgtat atcgtagttg	7620
tccccggtg atcactgttt cctgcttaat tgtgctgtcg tttttcctgg gtcctgattc	7680
agagtgtcag cattctgttc cccatagaat aagaagaggc tagaaagttt acagatgaga	7740
tatctaggaa tgccagaaga tcaggggtca ccgttgaggc agagtaatta attatggtta	7800
aaatgggtgtt gctgataagt ggggtgctggg aaataattaa aatttgattt tttagaagaa	7860
tactttcat gcttgaagag cgccctcatt atatgctaaa gggcctcagg tttttcctta	7920
ttgccattat gctgcagatt ctattacatt tgtctgaaaa gatctaagac agaagggtg	7980
tttaatacct tcccttttct cctgaacttc ccctctcctc tccccatca ggagctaagt	8040
aggaaccct tcacctgtt accatcagat ttcacatg gtctgtcttt acaatgaagg	8100
aagtagtact gcattctggg cagaggccag tcctgaggca tgccttttca aggacattgt	8160
tacttttagtt aactggctc ttctgtttta actcttatcc cccagactct aatcctgttg	8220
cttttttgg tccccatctc ccacctttca tcatctgaaa tccattcatt gtaacttctg	8280
gaactcagtc gtagaaaaat cttttatatt ctcaatcttg tgaatgttcc tttctttctt	8340
attccagctg taacctagcc ttctcccaa gaatgctact tcccttgag ctctctcaag	8400
tggtgaattt ttcccttctt gcacacctta taacactgaa ctaggagggtg tgtggactaa	8460
atgtctgctt ttgttcctta ttgtcacttc ttgacctta ttttccaaaa cttcaagctt	8520
tgactttcat gtgatcaaat tataccacc actgcctgtc tttatttcaa gcacctgcaa	8580
accttcctgg gtcattcaca tccttctttg ttcacttcat tagctcttgg ctcatgtca	8640
ctgtctctta tttctgtcat aattcttggg gacatcagta tctatgtaga gcaatactag	8700
tgaagatgtg gtctggtaac tgttacctgt atgaattaag ataaggagtt atgccagaat	8760
ataagtcacc tgtgtcacta agtttactgt ttagcttact tttttttag caagattttg	8820
atgaaggacg caatatgttg atttacagtc tgggtacaaat tttgatgtag aagatgcttc	8880
caatatcctg gtctcttagt tccttgattt cttctccagt gatcttattt tctaccctaa	8940
ctcaactaca tattcccatt gtcataatcct agaataattt gtcttttatc tgtaactctg	9000
ctctcttccc ccaatctcat ttcaagcatc ccactttcta attcctctag taaatacgtc	9060
agttccaaca gcccatcaat cccattggga cctacagttt atctatccaa gcttttcct	9120

gttcctcacc ctcacttcta tacagctgaa gtttcatact gaattataat cactttctcg	9180
tatacacgtt taacaatctt gtccctccct ggcttcatgc ccagtgatct cttgtatcta	9240
tgaccatgtc ctttatcttc tcctctgtca ctggatgaac tgtagccttc caagataagg	9300
ccactcagtt catttgtaca gcagattcca tcccctcttg ctctcaagaa tattactgtg	9360
gtatctctct tttcttgtct ctactggctc tttccatgag caaacatggg attatcccat	9420
tacaaaaaaaa attttttctc cgtctctcct tccactcacc acctcagtct ctgcttctct	9480
ttcccgcaaa ataaccttga aaaattgctt tatgtactcc cgttttcttt tgaaccctg	9540
ccagtgacca ccacgttata aattttagtg tgtcatctca cttaatctgt tagtagtatt	9600
tggcaccatt gctacagttg cttgaaatgc cttttcattg gtttccaggc caccatgtct	9660
gtagcagct tttcctctta cttcactagc atttccttct ttgttttttc tgttatcttt	9720
ctgacctctg ttggagtggc tgaaggttta gtccttgaat ctttttttgt tgtgcatatt	9780
tactccagta tcatagcttt atacagatgg tatttacatc tgtttgctaa cgatttccaa	9840
attggtatcc ttaaactggg atccagctat tttttggta gcattttgga tgtctaagaa	9900
gcttctcaaa ctaaactgac ctcccggttt tcccaaagc tgcacttag tcttttccga	9960
aatgcaattc tgtctttcca gttacctagc ttaaaagctt gcagttcttg actcatcttt	10020
ctctcatacc acgtatctga attctctctg caaaaaattg tctgttctcc cttcagaata	10080
aagtcacgtg tcattttatg atggggatac attcagaaat gcgtcattag gagataatca	10140
tggttgtgtg aacatcagag tatacataga caaacctaga tggatatagcc tactacacat	10200
ctaggctata tgggtgtggc aattactatg atgaatactg taggtaattg taacataaag	10260
gtaggtatth ttatctaaac gtattgaaac atagaaaaag tacagtaaaa aatatggtat	10320
caaaaataaa aaatggtaca actgtataag gcagttgtga tgaatggagc ttgcaggata	10380
tgttgctctg ggtgagtcag cgagtgcga ttgagggac gtgaaagatg tgggacatca	10440
ctgtacacta ctgtagactt tataaacact gtacacttgg gctacactac atttttgtaa	10500
ggttttaaaa gacttttttc tataataaac cttaaattac tgtcactttt ttactttatg	10560
aattcttaat tttttaaacg ttttcactct tgtaataaca cgtagcttaa aacatacatt	10620
gtacagctgt acaaaaattt tctttatatc tttataagct tttttatatt tttaaaatta	10680
ctttttacct tttagctttt ttgttgaaaa actaagacat gggccaggcg cgggtggctca	10740

cgctgtaat	cccagcactt	tgggaggctg	aggcaggcgg	atcacgaggt	caggagataa	10800
gagaccatcc	tggctaacat	ggtgaaaccc	cgtctctact	aaaaatacaa	aaaattagcc	10860
gggcgtggtg	gcgggcacct	gtagtccgag	ctacttggga	ggctgaggca	ggagaatggc	10920
gtgaaccag	gaggcggagt	ttgcagtgag	ccgagatagc	gccactgcac	tccagtctgg	10980
gcgacagagc	ggaaactccg	tctcaaaaaa	aaacaaacaa	aaaactaaga	catgaacaca	11040
ttagcctagg	cctacagagg	gtcaggatca	tcagtatcac	tgtattttcca	tctccacatc	11100
ttgtccttct	ggaatgtctt	cagaggcagt	aaacataaat	ggagctgcca	cctcctgtga	11160
taacagtgcc	ttctggaata	cctcttgaag	gacctacctg	tggctgtttt	atagttaact	11220
tttttttttt	aagaagtaac	agaaggagta	cactctaata	ataaaaagta	tagtaagtac	11280
ataaacctgt	aacaatcatt	atcattatca	agtgtcatgt	actggacata	actgtatatg	11340
ctatactttt	tttttttgag	atggcatctc	actctgtcac	ccaggctgga	gtgcagtggg	11400
gcgaggatag	ctcactgtaa	cctcagactc	ctgggctcaa	gtgatcctcc	tacctcagcc	11460
tccaagtag	ctgggactac	accaggcacc	ccaccatgcc	tggctaatta	aaaaaaattt	11520
ttttagaga	cagggtctca	ctctgttgcc	agggctggcc	ttgaattcct	ggcatcaagt	11580
aatcctcca	ctttggcctc	acaaagtgcg	aggattacag	gtaagagcca	ccatgtctgg	11640
cccactgtac	ttttatacaa	ctgaagcaca	gtaaacctac	tgtggtttcg	tttacaccag	11700
catcaccaca	aacaccatga	gtagaacatt	gtgctgcgac	gttaacgatg	gctacaacat	11760
cactaggtga	taggaatttt	tcagctccat	tataatctta	tgagaccact	gttgtatgtg	11820
cagttcatca	tccactgaaa	tgtccttatg	tgatgcatgt	cttcatatcc	aaaaatatta	11880
atcatttctc	actgaagcca	tgccatgcca	tgccatcttt	tgctgtatt	attatttttc	11940
agcttttatt	ttagattcag	ggtgtacatg	tgccaggttg	ttagaaagag	tatatcgtat	12000
gatgctgaag	tttgggatac	agttgaacca	gtcaccagg	tagtgagcat	agtactcaat	12060
agataacgtt	ctaacattac	tcctccttcc	ctccctgttc	ttgtctctgt	ctattgtatc	12120
tttatgtcca	tgtgtaccaa	atgttttagct	cattcttgtg	agaacatgtg	gcatttgatt	12180
ttgtttctgt	gttaatttgc	ttacaataaa	tagtctccag	ctgcatccac	attgtacaaa	12240
aggacatgat	tttgttcttt	tttataggct	gcacatatt	ccatggtgta	taggtaccac	12300
atcttcttga	tccagtctac	cgttcatggg	catttgggtt	gattgtatct	ttgctattat	12360

ggatggcttt tgcctatatt attggaaagg ctttctaact ggtgtccctg cttacaccgt	12420
tttccccctt aaatgtgttt tcaacatggt agccagagta acccttttta taacaataaa	12480
tcgtgtaact tttttgttca gaaacttaca gggcttacca tttcattcag taaaagctca	12540
agctcctgta tagtcagacc atatccttca tcacctgtta cttttctcct ctgactcttc	12600
agcctttttg tttttcctca aactgatgaa gccttcatgg ctgatgtcag atgttttgcc	12660
cattgagatc ttccttgttg actcagttgc acttggtcat atgattttca tttatttggg	12720
gtatctaadc ataacttgaa agttggctac ttatttttac ccctttgagg gtccttgccc	12780
tgtttttgta tccctgatag cgggacagcc agatatctgg aacttacagg tgttcaataa	12840
agttttgttg aatgaatatt ctggaatcac ccaacctttt ttttccccct cacttatttt	12900
tcttctccct ttcacggcct gaaagatgtc ctatgtatat ggttccactt atcactctca	12960
tcccagtttg tgatatacta ttccattata ttactattat taatacaatt ccattgaact	13020
tgctcttgct gacttcacca ctggacctac atgttggcca aatggatact ttataatttt	13080
agtcttgacc cctgcctttg gcacatttct tacctctagc acagcactgt ccagtaatcc	13140
acactttctg agacagtgga aatgttcagt atctgtgctg ttcagttggt agcaaccagc	13200
tacccatgcc tattaacat ttgaaatgtg gctgtgtgac tagtggcaat tatgttggag	13260
agtacagttt tagaaactcc tgtttttctt acatggcact acatttagta tcacaatcta	13320
attgtgcaag ccagataggt aggagtcac tttattcctg ttatttaatt tttctcatct	13380
actatatcca gttcatcaca tcaacagcgc ctgttggttc tacctcctaa atatttcttt	13440
agtctaacta ctacttgccc ctagtgccac caccatctat cagctggaat attgctatag	13500
ctgccttaca ggtttccctt ctttctgtt ctcttctagt tttttgaatt ttagtcagca	13560
cgagatttta aaaactcaaa taagattgtg ttattcacct gcttaaaacc tttcatgact	13620
ttcagtgta cgtagaacag aaaacacttt tcttaccaaa ggctagagag ctctacgtga	13680
tctggctatt tttaacgttt cattgcactc acccttttcc tctataatca aactactctg	13740
atctcaaggg ttagttcttg aaagatgac atgttcttta atgactttag gtttttgtgt	13800
gttattttct atttctggga tgtttattct ctgttcctta catgctggcc cttttgcac	13860
cttcttcagg tctcagctta catgttacct tcaagaagcc tttgaccact ctaagtgggc	13920
ccttccttcc acttctgctg tgtaatccca ctcccttctc ccacttgta attagttaca	13980

tacttttttg taattgttta tttggttgct gtctccctct caagaatgca gggaccatgt	14040
ctgcattctg cagtaatcac tactgcacac ccagaatcta ttacagatcc tggcatgtag	14100
ctgatgcata aatatttggt gaatgaaagt ctgtacattg tattttatgct attggtattg	14160
ctatgacctg aaactaaaag gagttgtgga aaagatttct tatggaacag aaatatccct	14220
tttgattaat atcacaatct cgtaaattga gaaaacaaaa aaatatatac tactggagca	14280
ttcatgtata gttggagatt atgactcatt tattggtgtg tttttggact cagaacaaag	14340
atgagggaat attccttaaa gctctgtatt gaaataacga aaagcagtca cattttaata	14400
atagaagctt cctagcttac tctttctgta atcttctttt cctaaatgta agagagcctc	14460
ataattatga ggcttattac tagagtaagg ctgtcaaagg cagcaaaatg tctttctgtt	14520
tggaagaata acataaactt gacatgtatg gtggggggaca gaaggtttca aaagttaaag	14580
aatctgtgtt gtcttaacaa atagatgctt ctcaaggagc ttacgctagt ggttactctg	14640
tccagtcagg gttttttctt ctttaacttg ggttcatttc ctgatggcac acatgaagtt	14700
tggatcatat ggtttgactt tagctatggt ccttagctat ggggagcagc atcagcgacc	14760
tgtgacatgt aaattaaaaa tacaatgcca gggcccttcc ccagcccctc tgatagagaa	14820
cctcttggcc atctgtattt ttagatgttc caggttagtc tgattaacac ccttggttaa	14880
gaaccattgg gaggatctga ttgccagttt aaggggacct tcaagcctgt aggtctttat	14940
agttaaaaaa aaaaaaagat tttaaaaatc atgcatatgt tgtggctgaa ttctggttta	15000
gcacatactg cttttaatgg cctgaaatgt ttttcccaaa taaattgtct tgttatagct	15060
ttcatgtgtg atttggtcca gcttcttggt ttgaagatac ttacgggggg gaacactttg	15120
tgatttctct tagtaacata ttaaccact taaaaacctt ttctattaca ggtcttcaca	15180
tttaggctta atgtgcttaa ttcaaagtga aaaatacacc tgcctttgtt ctcagtga	15240
gtatgtaata aataaatgag gggttggcaa actactgccc accatctgtt tttttatggc	15300
ctatgaacta agaatcgttt tggatagcta aaaaaaaaaa tcaaaaggat aattattttg	15360
tgacgtgaaa attatatgaa attcaaattt cagtttctgt gaatgaagtt ttaatggaac	15420
acagccatcc atgcttatgt aagtgtgcat attctctggc tgttttcact gcaatagcag	15480
agttgagtag ttgtgacaaa gagtttatgg ccacaaaaac ctaaaatatt tactttctga	15540
tgctttacag aaaaagtttc ctgaacctta ttctagctat atgttggttca taaatgaatc	15600

tttcgtgggtt	ctgaaggcat	ttaagaatct	cttaggttat	aaattggctg	ggcgcagtgg	15660
ctcacgcctg	taatcccagc	actttgggag	gccgaggctg	gtggatcacg	agggcaggag	15720
ttcaagatca	gcctagccaa	gatggtgaaa	ccctgtctcc	attaaaaaaaa	aaaaaaaaaa	15780
aaaaaaaaata	gctgggggtt	gtggtgggca	gtaatcccag	ctactcggga	ggctgaggca	15840
gagaattgct	taaaccagg	aggcggagga	tgcagtgagc	caagatcgcg	ccactgcact	15900
ccagcctggg	caacagagt	aaacaccatc	tcaaaaaaaaa	aaaaaaaaaa	aaaacactct	15960
taggttataa	ataattgttg	ttagctctcc	aagcctccat	attacatttt	gtgtgttctc	16020
ctgttcacat	tttgagcatt	ttatTTTTTA	ttagcacatt	cagttcatca	ggtattttaag	16080
agcttaatat	atgccaaagc	atatattaag	cgagaagctg	tttctaaatg	tactgtctca	16140
gccctcacag	agttcacttc	attaggctct	ttaaattttc	tttctttaaa	aggtcagcgt	16200
gctggtatag	tggggaaggg	aaactcttac	aacacgtcga	gtagaggaag	gttatcatta	16260
tgggatataa	tttgaagtc	attgagtacc	tgccattaat	tctgcctgta	gtctgaatgt	16320
agagattaac	atgtagaaac	TTTTTTgaaa	taaaatcttc	aatttctttg	gcatatctag	16380
tactgtctag	ctaggcatat	agtcaaagta	tggtgtatat	ttcaagtatt	aaaagttttt	16440
ttgggctgta	gtcactgttg	aaaggatata	gttctttact	attacatgtg	atacctttat	16500
ataaaattgg	ctaaccctg	tctttcattt	atctgcaaca	ctgactgtta	ccagttgtct	16560
ctaactttgg	tatgggggggt	ggaaatatga	ttagattgaa	aggggtacatg	actgagccac	16620
aagcagacct	ggatttgaat	tttaactgaa	cggtttatta	gctattctta	cattaatact	16680
gctaatacagt	tttcttgtga	tatgaggaat	gatgtcttct	ttatgagggt	gctaggaaga	16740
ttcaatgaga	taacatacta	ggctcagaac	tgaagttgct	aggaatttaa	ttatgctacc	16800
ttgttaaagt	atgtcaaagg	cagaattcag	tgtttagctg	ataccacaag	gcagtatcct	16860
aaaattatgc	tgtaaaagat	ataaagatgc	tgtaagtgac	tcagaaacct	agtgactttg	16920
taatgcagtt	gattcttaga	atactgtcac	tttaacagaa	taggagctag	gaatgaagaa	16980
atagttatta	aattactaaa	atagaaaatt	tattgacaca	tgtaaagtga	catttgctta	17040
aatattgaaa	aattttagt	actatttcct	tgcttttagaa	aacattgggt	accacttttt	17100
ttatttatag	cagtttgttt	ttgccttgag	gcaagatggg	tgactgagta	gttgccacat	17160
ttcttttgta	caaagtccat	ttcataggcc	atctagcttt	tatgcttaga	aacatttcct	17220

taacgttata	tttcagtatt	tggttaacct	atataggggt	aaattatata	ggctaacttc	17280
tcggacagat	atctctaata	atctatgtat	ttggttctgc	aatgtatgc	aaaaatatat	17340
gtacaaaggt	atgcagatgc	cttgcatact	tgatatatgt	taaatTTTTT	ttaatgtaga	17400
cctttttcgt	tctctttaat	gactatatgg	tattccacca	ttccccgctc	acctggacaa	17460
ctacagtaac	ctcctaaatg	gtgtttctac	tttgctattg	ccccttattg	tcttttttcc	17520
cctttatagc	tgctggagtg	aattttagaa	agcctaagtc	atacatcaca	ttgcttcatg	17580
ggcatcccag	tacactttgg	atcttatttt	acatccttac	tgatctgatt	ctcatctctg	17640
tctcttcatg	gttctctgcc	ttctagttac	actggtgacc	tttcaaaacc	tttaccacat	17700
tgagttcatt	ccttactttt	cactctttct	ctgcctggag	tgttctgccc	catctttacg	17760
tggccagctg	ctcctcctct	gatgaaatgt	ctcttcctca	caggccttcc	ctgaccaccc	17820
actagagtag	cacatcttct	acctcataaa	cttgttttatt	agtatttctt	actctaaatt	17880
ttcttttaaa	ttgcttaatt	ccctaacagt	agaatataag	cttcactgta	tgtatgatct	17940
tgttgactct	cttactcatt	gttattgtaa	taccagtaac	aaagggtgtt	taaaatttgt	18000
tcagtgggtg	aatatatgtt	ccatttaatg	gataaattat	tttttattca	gtctcctgtt	18060
gatggacatt	tgaataatct	ccatcttttt	ctctatgaat	gcctcacttg	gcatgcttct	18120
gacagtattg	ccacagaata	catttctgtt	ataaaaaattg	aattttttaag	tcaaagggta	18180
gttacacttt	aatggatagt	ggcagcttac	tatcaaaagt	ttctgctagt	ttcaccatat	18240
ccttattagc	agtagatatt	atcaatcttt	tcaatctttg	ccaatctgat	aagcaaaaag	18300
taaatgggtt	taaacatcct	ttgtatatat	tcattgctca	ctttatgttt	ttcctttgaa	18360
atgttatttc	ttgttctttc	cctgcagtat	gattctttct	ttttttgact	tgttcccagt	18420
tttttgtgta	ctatggatat	tagcctttta	ttatgttacg	gatgttctag	tatgttattt	18480
tttgaattac	ttcaaagtgt	atttgttgct	cagattttta	aaactacata	cacaaattat	18540
ctcatgtttc	cctttttggg	ttcaatttcg	actcatgctt	aatcagttca	tcgattgggc	18600
atgggttttat	tcttaatata	taccctgatt	ttatctcatt	ttattttttt	acgtgtaaat	18660
atttggtgaa	tataggttta	atcttaatgt	aaaataagga	tgaaaaatga	tagttggaat	18720
tacaagccca	tttctcctaa	tacttttaat	caagtaatcc	actaattgaa	atattacctt	18780
cttcatttat	gaaattgcc	cattatatct	gggtgttttt	ctgcctacta	cagtctctta	18840

cccatttctt	tcctaataat	acaatacttg	aattgctgtg	gttggtgatt	tataatgtta	18900
tcttaatgat	aacattataa	atgtgatgga	actggttcct	ccttatagtt	cttcttaaat	18960
caagaacaag	acatatcttc	ccattttactc	tcgtatgtat	ctcattttac	tgttatgaat	19020
gaaatctgtc	ctatttgtgt	ataggaaaat	agtttttgta	tgtaattgtg	atatggccag	19080
ttttattaaa	aatttggtta	aactaagagt	tgttttctgt	tcagccttat	catactataa	19140
aatccacata	aaatgggtat	aaaagtgtcg	caggacactg	ggctcagatg	attctcccac	19200
ctcagcttcc	caagtagctg	ggactacagc	ggcatatgcc	accacaccca	gccaattttt	19260
aaataagttt	taaaaatagt	attttttagta	gagacagggt	ttcaccatgt	tgcccaggct	19320
ggtcttgaac	tcctggactc	agacaatcca	cctgccttgg	cttcccaaag	tgttgggatt	19380
acaggtgtga	gccaccacac	cttgccgaat	tgcagccata	tttaatactt	ttttccatcc	19440
tattcccttt	gctgccccca	ggcctcctgt	attgatagcc	cgctattaag	aagctagtgt	19500
atattctttg	catactttta	cttcataaac	tatatgaagc	attgttctgt	tttttaactt	19560
aattggtata	aaattatatt	ttggaaattc	agtatatctt	gtgaaaatta	tttagaaaat	19620
gtgcctctga	gataaagcct	attcaggatg	tatcttaaag	gagatagctg	tgctttaaca	19680
ttatcagtct	ttttggctgc	ttatgttaat	ataagttgga	gaaaaacagt	ctgctttttg	19740
tgataatatg	ttcttggaga	tggagtgaag	gattgtttta	aaacattgtc	ttttttttcc	19800
cctgaagtac	cagtatttat	tttaggatta	tgttactgat	caaagatgct	gtgtggagtt	19860
actcattggt	gagactaaca	ataaatcaca	catgcaaagg	atgttaccat	aatctaatta	19920
ttttaaacag	taaaattata	ttctaagaca	tccagttggc	ctatatgtgc	tatatcaatg	19980
actatcaagg	ggctttttat	gtatactgta	tacatgtact	tcacaaaaat	ataaaaggat	20040
gacatcaaaa	atctggcaag	ccaaaagcct	acattacatg	tagcaaataa	ataagcatat	20100
gaacttattg	gaatttaaaa	ccctgtagga	tgggcgggtg	atggtatgta	tgtagatgt	20160
gtggacatat	ctattaaaag	ttgtgtcaga	taacagctgg	tgctgacaag	cccttggtta	20220
gatggcagca	tgttcaatat	gttctgtgaa	aattatctca	gtttatgatc	tgtcagtatt	20280
gtggagctat	gcatgaaagg	acttaaaatt	cttaccctta	aactcagtaa	cagtgtttct	20340
agaacttctg	gtgatatggg	aaattaagag	aattatttat	atgcaaagg	gtttattgca	20400
gcattgttgg	aataatagac	aaaatgggga	agaacaagct	cagaatggag	gaggtagctt	20460

atagtataga	catacgatac	aatccagatg	ataatatattt	ataatagtct	tcacaaggaa	20520
ttttatatatt	ttattttttaa	aaatacatag	cagtgagttt	aatataccaa	acataccaaa	20580
atgtcatcat	ttactgtgtg	gtggactcat	atgatggaga	tgataaataa	aaatattaat	20640
ttatttgagg	catatatatta	tggctgagga	aggaagacag	ttatgaagaa	cagctcattc	20700
tggaaacata	ctaatttttc	ccagccataa	agagatttcc	tatttctttt	ttttttccat	20760
ttaccttctg	tttcttacct	gagaagattt	catacttcta	ataaccattt	gtgtacctat	20820
ttaaagacag	taccaaaggc	atacatttta	gtgtttggag	gaccaagggt	catttgatgt	20880
ttgatgctta	ttgactattc	gaggatgaca	agacaccttg	agaacacaca	cacccacacc	20940
cacaccacaca	ccctcaccca	cccaccccac	ccccctcccc	gaagaaagct	gtgaaggaag	21000
aaagcagaaa	agaacctgga	gtgagttgta	acttaaaatg	ttagtgttgc	atgaagtgtg	21060
ttaaaacagg	aagatttgag	gaaattgcat	acattttcta	gatggcaaag	tattactggt	21120
gacagttaat	gaaaatgcat	atgcatgtgt	ttttagattt	acaaatttta	ctaagaactt	21180
tttaaaaatc	cctgaagggtg	tatcaaaagt	ttatcatgct	tatgaaatag	agtagcactt	21240
tctaacttta	aaacggggaa	taattctttg	gatcttgatt	attggaaaag	tgaattatga	21300
attgctagta	taaaactgtg	gttttaaaat	atgtctgctt	tatatatttta	tgtagcagat	21360
ttactcctag	ttaataatac	tcaaacttac	tgaaaactaa	ggtaattaag	ataattctgt	21420
cctgatggga	agaggaaaaa	taacttcagt	gtgaaatcta	ttatatatta	gttgtggcaa	21480
gatttctccc	attgactttg	actggagaca	tttatagggt	taaaatcgga	aatagcacgg	21540
tgaattttga	agtatccttg	tagttggaaa	gagtattatg	ttcatattgc	caaaaaaaaaag	21600
atgcatggat	gcattagact	ggatggaaaa	tacatgagaa	gttggttagc	cccctctttg	21660
tcaaaacatc	acttggtggt	gataaagctg	ttggaaaaca	cagcattcta	atgtagtctg	21720
tagtttaatg	ataatctgtg	tcttgaaaca	tttagcgtag	tacttatata	aacctagatg	21780
gcatagtgtg	ctgcatgcct	agcctatata	gtatagcctg	ttgcttctag	ggtgtaaagc	21840
tgtatagcgt	gttactatag	gcagttgaaa	cagtgggtatt	tatgtatcct	tttttttttt	21900
tttaaatctt	tttaagagac	agggtcttgc	tctgttgccc	aggctggatg	cattgggtgtg	21960
atcatagctc	actataacct	tgaactccta	agtgatcctc	tttgcctcag	cctccccagt	22020
ggctaggact	acaggcacat	actaccacac	ctggctaatt	tttaacattt	ttttgtagag	22080

atggaatttc gctgtgttgt ccaggctggt cttggaactc ttgtgctgca gcaatccacc	22140
cgcctcccaa agtgtttagaa ttacaagcca cttcgcctgg cttgtttacc taaacataga	22200
aaagatccag taaaaataca gaattaaaat cttgtggggc cactgtagca tatgtagtcc	22260
atcttgactg aaatgtcctt atgcagtgc tgaattgtact tcataatttt taagcactcc	22320
tccctcttga ttggtactta gtggatttta tcatttttgt ttcttcataa ttctttctga	22380
aatgtctact ggttggacct ttgatctcct gaattgatcg tgatttcttc tgttgatatt	22440
tttgtctttg tcattttttt gtactctagg cagttttctc aatttttagtt tctattcaac	22500
tttttgtttt tattttattct ctccagtatt tatggagata ctaaattgaa gtgttctgtt	22560
tctctctcca ccctatccct agtttcaagt tttatctcag tttctatgga gtcagttttt	22620
tcgttgcttt aaaaaaaaaat tttcctgaag tgattggtaa gttttggcta attgggagca	22680
ctagaattgg gcccttaatg gttggcaggg tgtgggtggag gagagacagc ccttagtcca	22740
aaggctcagg ccagaaaaag aaagaggaag gctttccttt tcctttccgg agcagggttc	22800
tgccctaggt cttgcttggc agtctatttg atttctttag cagttaatgc tcagtttttt	22860
ggcatatgtg gatctgcctc cagagcaggt acaaggtagag tgagtctatg ctgttaccta	22920
attagatccc catttctacc ctttgttttt acttctctat ctactgatag gtttttacc	22980
tccttcacct catagggttg cagtgaagag caagatgaat ttttatttat gttgcataaa	23040
ttttaaaagc taaaaaatat atatgtaatg ttgggaagtc ccagtgtaca aatggctatt	23100
gtaaatttgg aacatgaact tgcttttttc cattgtaaaa atgaaatcat tataaattgc	23160
ggccaagtta ctaggtcagc ccacacagag tttaccagc aatatgcgta aatgttttgc	23220
ctttgcatca acaacaagga aaaacagtac tataaaaaaa tgttcctgga agccggatgt	23280
atcaaagcac ttctgaaata gctatatagc ctatagacat gaccagttgg tttctgagtc	23340
tgttgacatt ggccaaagga gaagctcagt gtagaacatg tttggagtct ctttttgcag	23400
aaatacattg gaggctggag tggggaacca atttttcaga aagggtggtga agtagttaca	23460
tagccactct tttaaagaca gtcaaaagat agaaactaag gccagggtgt ggctcacatc	23520
tgagatagga aaatcacttg aacctgggag gcggagggtg cagtgagccc agtatgcacc	23580
tctgcactcc agcctggttt ggcaagagac caaaactctg tctcaaaaaa aaacaaaaca	23640
tagttcacac ttaaatattt tattccatat ctttacatac ccaatatgtt aatttatagt	23700

tcaagatgaa	cttgttttggg	acagattttg	taataaagga	aatcgtgtta	ttagaaatat	23760
ctagaggcca	tgagccctta	aactgttcta	atttgcaagt	agttccctgt	gtgatgcagt	23820
ttttttcaat	attgcacaat	aaaggcaaaa	tacggacaaa	ttagatgata	agatttatat	23880
aaatttttaa	aatattgatc	aaaatatgta	tccatattgg	taatatttgt	atttataata	23940
aatcattgct	gtaaatttga	acttagaaaa	attttactaa	taaagggtgct	tttgtgttgc	24000
aaactttcat	ttgaaaagta	atttttcttt	gtaccaaaaa	atctaaaatt	cgctatttcta	24060
gtcaccaaaa	tttgctttat	gaaaaataat	ttttgatggc	actatatcag	aaaacaactt	24120
gttaaagaaa	atgtggagtt	tttaaaatcc	cactgtacct	ctgttatcca	aaggggatct	24180
gtgaattttt	ctgtgaaagg	ttaaaaaagg	agagaccttt	aggaattcag	agagcagctg	24240
atttttgaat	agtgttttcc	cctccctggc	ttttattatt	acaactctgt	gctttttcat	24300
caccatcctg	aatatctata	attaatat	atactattaa	taaaaagaca	tttttggtaa	24360
ggaggagttt	tcactgaagt	tcagcagtga	tggagctgtg	gttgagggtg	ctggaggaga	24420
ccatgaggtc	tgcgtttcac	taacctggta	aaagaggata	tgggtttttt	ttgtgggtgt	24480
aatagtgaca	tttaacaggt	atcccagtga	cttaggagta	ttaatcaagc	taaatttaaa	24540
tcctaatagac	ttttgattaa	cttttttttag	ggtatttgaa	gtataccata	caactgtttt	24600
gaaaatccag	cgtaggacaat	ggctactcaa	ggtttgtgtc	attaaatctt	tagttactga	24660
attggggctc	tgcttcgttg	ccattaagcc	agtctggctg	agatccccct	gctttcctct	24720
ctccctgctt	acttgtcagg	ctaccttttg	ctccattttc	tgctcactcc	tcctaattggc	24780
ttggtgaaat	agcaaacaag	ccaccagcag	gaatctagtc	tggatgactg	cttctggagc	24840
ctggatgcag	taccattctt	ccactgattc	agtgagtaac	tgtaggtgg	ttccctaagg	24900
gattaggtat	ttcatcactg	agctaaccct	ggctatcatt	ctgcttttct	tggctgtctt	24960
tcagatttga	ctttatttct	aaaaatattt	caatgggtca	tatcacagat	tctttttttt	25020
taaatttaaag	taacatttcc	aatctactaa	tgctaatact	gtttcgtatt	tatagctgat	25080
ttgatggagt	tggacatggc	catggaacca	gacagaaaag	cggctgttag	tcactggcag	25140
caacagtctt	acctggactc	tggaatccat	tctggtgcca	ctaccacagc	tccttctctg	25200
agtggtaaag	gcaatcctga	ggaagaggat	gtggatacct	cccaagtcct	gtatgagtgg	25260
gaacagggat	tttctcagtc	cttcactcaa	gaacaagtag	ctggtaagag	tattat	25320

cattgcctta	ctgaaagtca	gaatgcagtt	ttgagaacta	aaaagttagt	gtataatagt	25380
ttaaataaaa	tgttgtgggtg	aagaaaagag	agtaatagca	atgtcacttt	taccatttag	25440
gatagcaaat	acttaggtaa	atgctgaact	gtggatagtg	agtgttgaat	taaccttttc	25500
cagatattga	tggacagtat	gcaatgactc	gagctcagag	ggtagcagct	gctatgttcc	25560
ctgagacatt	agatgagggc	atgcagatcc	catctacaca	gtttgatgct	gctcatccca	25620
ctaattgtcca	gcgtttggct	gaaccatcac	agatgctgaa	acatgcagtt	gtaaacttga	25680
ttaactatca	agatgatgca	gaacttgcca	cacgtgcaat	ccctgaactg	acaaaactgc	25740
taaatgacga	ggaccaggta	agcaatgaca	tagctagctt	tttagtctgc	tttgaagtaa	25800
atgctcaagg	ggagtagttt	cagaatgtct	acccaatacc	agtacttgaa	aactaacgat	25860
gtttctgaat	tcctgtatta	caggtgggtg	ttaataaggc	tgcagttatg	gtccatcagc	25920
tttctaaaaa	ggaagcttcc	agacacgcta	tcatgcgttc	tcctcagatg	gtgtctgcta	25980
ttgtacgtac	catgcagaat	acaaatgatg	tagaaacagc	tcgttgtagc	gctgggacct	26040
tgcataacct	ttcccatcat	cgtgagggtc	tactggccat	ctttaagtct	ggaggcattc	26100
ctgccctggg	gaaaatgctt	gggtaagaaa	acatgtcaga	atgcttgaag	ctaaaaagta	26160
gaagagtata	ctcacaatat	ttctgatgag	gcttttttct	tcttcccagt	tcaccagtgg	26220
attctgtgtt	gttttatgcc	attacaactc	tccacaacct	tttattacat	caagaaggag	26280
ctaaaatggc	agtgcgttta	gctgggtggc	tgcagaaaat	ggttgccttg	ctcaacaaaa	26340
caaatgttaa	attcttggct	attacgacag	actgccttca	aatttttagct	tatggcaacc	26400
aagaaagcaa	ggtaagagaa	ttattcttta	tgtggttttc	atggagcatt	ggacacctcc	26460
agtgtcatgt	cattccatgc	agtgttccta	accttttttg	caccagggac	cagtttcgtg	26520
gaaaacagtt	tttccatgaa	tgggttgtgg	gaatggtttc	tggatgacac	cattccacct	26580
cagataatca	ggcattagat	tctcataggg	agcgtgcagc	ctagatccct	cgcatgtgca	26640
gtccacacta	gggtttctac	tcctatgaga	ctctcatggg	gcagttgatc	tgacaggagg	26700
tagagctcaa	gccaggtaat	gctcgtcac	ctgccactta	cctcctgctg	tgcagcccag	26760
ttcatttctg	ttcttttaaa	tttttgagtt	tccatatgta	aagcactatg	cgaagtagta	26820
gggatatggg	aggcaagctt	ctcttcacac	ttttgttctt	aggtgggatg	tagatgttgg	26880
gaataataac	ctaataattta	atttgtgtag	tgggaagaag	tggggctatg	agggcacata	26940

acacaagttg aaactgactc tttttgaggg ttcaaggaga cctcttggag gaagtgatag	27000
ttgagttcag tgttcaagga tgagaaggga ttcactaggt gaaggtagg tgagaaaaca	27060
acatctttga aacgaaggaa ggagatggaa agttttggga atttaagaaa tactaatagt	27120
aaggaggaag aaaggtttga ggtgaggcta ttgagataga cttagcagat ctcatagggc	27180
ttttagagc atgtttaaaa gcacaatggg aaatttcagc agaagcctga aatgatgaaa	27240
tttgttttta gaaaattggg gcagtgttga aagggaagat atacaggga tgaaaggaca	27300
agcatgaatg atcattttat ggtatctgtt ttttaagggtg atataattag gaaaattaa	27360
gggccaaatg atgaggagtt aagtgccagt tctggttcaa attttcagtg aatcagtttt	27420
gatataactt tcatcttagg gcattactct tgcctacca catagtttct aaattttttt	27480
cttttggtgt gatcactgtg ggaagaagga aattgggccc aaactgatac attgtttgga	27540
ggactgggat gtctgaattt gagtggaatg ctttaaaagg acaagttgga tagggcccca	27600
gtatgggggt ctgagtgatg ggggtccagga atacatttag gtccaatggc aagctggctg	27660
aaattcttgt ataataaaat aggttggtaa tatggctctt ctcagacatg tgatcaagat	27720
tccttgacta acaagatata tatatatatc tttctagctc atcatactgg ctagtggttg	27780
acccaagct ttagtaaata taatgaggac ctatacttac gaaaaactac tgtggaccac	27840
aagcagagtg ctgaagggtc tatctgtctg ctctagtaat aagccggcta ttgtagaagc	27900
tggtaaagtat atgtatctat tctgagtctt gtgtatagca tctgcagttc taattagatt	27960
acttttctta ggaaaagggt gtagaacttt aactactgaa aataaatggt cctattcagt	28020
ttgcagccaa gatttacatt cagagtacct gtcacttgga ttgtagctaa atatttaagg	28080
ctagtttagg tagagttctt attatccatc aaaaatgatg gcatatgttt tgcttaataa	28140
aatttgtttg taatttcagt tttgagtaaa cctaagattt gctaacagag ctgtgaattt	28200
ataggagaaa agacaaattc taatatagta cagttttatg taaagtgatt gctttattag	28260
tagatgctca tgagcagttt ttgttttggt ttaactttta gggtccgggt aatgtgcagg	28320
cttgttatat aggtaaattg catgtcacag gggtttcgtg tgcagattat tttgtcacc	28380
aggcagtaag tattgtaccc aataggtagt ttttcagttc ttacctccc acccgtaagt	28440
aggccccagt gtctgttggt cccttctttg tgcccgtgtg tactcagtggt ttacctcca	28500
cttataagtg agaacatgtg gtatttggtt ttctattcct atgttagttt gcttaggata	28560

atggcctcca gctccatcca tgttgctgag gaagacatct tggatattttt ttatggctgc	28620
ttagtattcc atagtatata tgtaccacat tttctttatc tagtctacca ttgatgggca	28680
tttaggttaa ttccatatct ttgctattgt gaataatgct gcagtgaaca tatgcatgca	28740
tgtgtcttta tggtaaaaag atttcttttt ctttgggcat atacctaata ataggattgc	28800
tggattgaat ggtaattctg tcagggttttt tgagaaatca ccaaattgct ttccacaatg	28860
gctgaactaa tttactttcc caccagcagt gtataagcat tctcttttct cagcaacctc	28920
accagcatct gtcatttttt gactttttat tagtagccat tctaactggg gtgagacggg	28980
atctcattgt ggttttgatt tgcattttctc taatgatcag tgatgtcgag cttttcttca	29040
tatgtttctt ggccacttgt atgtcttctt ttgaaaagtg tctgttcatg tcctttgccc	29100
actttttaat ggggttggtc ttttttgctt gttaatttaa gtttattgta aactctggat	29160
attagacctt tgtcagatgc atagtttgcc agtactttct cccatgccag tactttctcc	29220
cattctgtag gttgtctgtt tactctgttg atttcttttg ctgcgagaa gctctttata	29280
ctgtcccatt tgtcagtttt tgtttttgtt gcaacttctc ttggcatctt cgtcatgaaa	29340
tctttgccag gtcttatgtc cagaatggta tttcctagggt tatcttgcag agtttttaca	29400
gttttaagtt ttatatataa gtctttaatc cattctgagt tgatttttgt acatcatgta	29460
aggatggggg gcagtttcaa tcttggaagt ggctagccag ttatcccagc accattttatt	29520
gaatagggag tcctttcccc attgcttggt tttgtttact tgttaggtgt gcggcctaac	29580
ttctgggctt tcttttctgt tccattgggtc tctgtgtctg tttgtatacc agtaccatgc	29640
tgtgattgta accttgtatt aacagtatag cttgaagttg ggtaaagtga ttcctccagt	29700
tttgttcttt ttgcttagga ttgccttggc tattcagggt cttttttggg ttcatatgaa	29760
tttttaaata gttttttttt aattatgtga agaatgccat tggtagtttg gtaggaatag	29820
cattgaatct gtgaattgct ttgggctgta tggccatttt aacaatattg attcttcctg	29880
ccatgaaata gaatgttttt tcatttggtg gtgtcatctc tgatttcttt gagcagtgtt	29940
ttttgtaatt ctcatgtag agatctttca cctccctggg tagttgtatt cctaggtatt	30000
ttattctttt tgtggctttg gtaaattggga ttgcattctt gat ttggctt gcagcttgga	30060
tgttggttgg gtctagaaat gcttctgact tttgtacatt gat tttttata tcctgaaact	30120
ttgctgaagt ttattggatc aaggagcttt tgggcagaga ttatgggggt ttctaggtat	30180

agaatcatat	tgtttgcaaa	cagacttcct	atttggatgc	atcttctttc	tcttgctga	30240
ttatgagcag	tgttttgccc	tgatattctg	tattctcagt	gaatagatgt	cgtctaagta	30300
tgagaaacaa	tttttttcta	ttctgagtat	ttttaagaag	gcaacttata	tgtgggtactt	30360
tgtatattgt	gtatgttggc	aattggggaa	aagaatagat	ggtttgtact	agggcctctt	30420
gggttctgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtcat	gaaaacagtt	acttttttagc	30480
taccaagcat	tttttctcct	ttcagtaacc	cacctaacaa	catttactca	gaatttcaaa	30540
gcaagcttca	aatcagtatt	gaaagaagga	aaaatataaa	ggcattttaat	ggaagaaaat	30600
gttggaata	aagtataggg	ctggcaacac	ttacttttct	cacttattga	gagtaatttt	30660
acttgggaat	ttatgagaga	gaaagacatt	atgattgctc	caggtaacta	ctggcagagg	30720
aaccatagtc	ttggggatag	acaaatgtgg	ctgagttcat	atagaatgag	gggatgggat	30780
gtaaattctg	tcagctgttc	cagcagtaac	ctgtaatgta	ggctaataaat	acagattttg	30840
agatttattt	aatcagaatc	cctggagtgt	taatttttat	atcaagatct	catagtgttt	30900
tatttgaagt	gacagggagg	tctgtagata	gctggacatg	tatgggactg	gaagcttagg	30960
aatctttaag	ttcttccagg	ttattcttat	gttcatttgt	ttattctgaa	aatagcatct	31020
aatgtatttt	aagaaatgga	ataggcacat	agtatacatt	gggtaacaca	acagataggg	31080
tccccgtgct	taattcttag	tcttgtgaag	gtgacaaaaa	tacttaaaaa	tatgtgatcc	31140
taaattagaa	tgagtgttat	gggagaaatg	acagcaaata	gtgatgagaa	ttaatgggga	31200
ggggaattgt	ctagatgaga	gggaaaaggt	ctccttgaaa	aggggatggt	aagtgggact	31260
gcaggatgag	agggaaccgt	ctcttgtcta	tatgagaagt	gagggttaaa	cgttttccag	31320
gtagagaaaa	ggaacaccat	gtgctatgtc	ttagaaccag	ggatatccag	tcttttggct	31380
tccctgggcc	acattggaag	aagaataatt	gtcttgggct	acacacaaaa	tacactaatg	31440
atagctgatg	agctaaaaca	aaaaaaaaatt	gcaaaagaat	ctcataatgt	ttaagaaagt	31500
ttacgaattt	gtgttgggct	acattcagag	ctgtcctagg	ccatgtggcc	catgggctgc	31560
aggttggaca	agcttgcctt	agaaggaaaag	agattgggtca	ggcacgggtg	ctcacgcctg	31620
taattccagc	actttgggag	gctgaggtgg	gcggatcatg	aggtcaggag	atcgagacca	31680
gcctggctaa	cacagtgaag	ccccatctct	actaaaaata	caaaaagtta	gccgggcgtg	31740
gtggcaggcg	cctgtagtcc	cagctacttg	ggaggctgag	gcaggagaat	ggtgtgaacc	31800

cgggaggcgg agcttgcagt gagctgagat agcgccactg cacttcagcc tgggcgacag	31860
agtgagactc tatctcaaaa aaaaaaaaaa gggaaagaga ttgtggagat ccagggtgctg	31920
aagagaaggt ctgcataaac agaacttagt aatgaggtgg atggcctggt atgaggttga	31980
ggttaggtaa gcagagccat aacatgcagg actttctagg ttcctataag atagttacta	32040
ctcatggagt ttattcatgc tttattccag ctttggagcc atagatacag aatactttgg	32100
tcagtttga aggctaggtg ggatccaaat tctaaacggg tcctcagggt tataactaaag	32160
tatttctatt atcttaaaag gatgctgaga cactttcgat gggtgtttat caatagcaaa	32220
gcatcacagt ggtgtgttta aaatattaat aatagcattg tatagattaa cagtttgaat	32280
gaccaaagc tagaagacca gactactgag atgttacagg cttttaggaa tgaaatagtt	32340
tgcttttaga actcaatagc aaagggcaga tgtctgagat gcctgaaaga atcatagaat	32400
gtaataatat aggagctaag ggagcaacca aaaacggttt gtggagggga caacattggt	32460
accatgaaga taaatggaac cctcagaagg catccttaat ttttgaacat aataatttaa	32520
gaagctgact taaagtgact taaaaggtca gtaggtagct ggaaatgtat gatactagaa	32580
tgcaagagag gcaggctaga gatttggaag tttccctctt agtatatagg ggtaagggca	32640
gcaggggaagg ggaggtagag gtgccacaga gtcacttgta tgggactttt ttttttacc	32700
tagaactgct gaatcagaat gtgtgtgttt taaagtctct gtaggccatt ctgatggaca	32760
tctggggtta aaatccattc tcttagagtt aatagttatg taaagggagg gaatgaagtc	32820
ttaaagaggg gaaagaaggt agtcatttca caaatactga gcatcctgat catcagtctt	32880
acgcagatca ttctattagt agctggagct actatgaaaa aggaacccaa cagaggtgat	32940
ctttgtcttg tagggaaagt ggagtaactt acactatgaa ggagaagtgc agggtagcat	33000
aagaattaca gcagatagac ctcatctgag gaaataaaac agacccgaaa gatgaaggag	33060
acaaggaaaa gtatctctta ctgcattcag aagtgattta agttgaagat ggatgagcga	33120
agttaatcta ctatgtgggc attgggcttc catttatact cctttgccag agtaaatgtc	33180
ccccatttaa gggtcctaaa ggatggaaga ttgtaaacct tggaacacat gttttgtagt	33240
cagtgaattg tataaagtcc ctgacagtaa gtgttttcat gccgtctttc tggattgttc	33300
ttaccccagg aatttaccta gcttctttag gtcttttagtc agatgtcacc ttcacagtga	33360
ggtgacctaa ttatctattt aaaatcgcag cccactcca ttatttttct ccatagccct	33420

ttaatatcat	ctgacatact	glatggTTTT	agtttattgt	atatTTTTct	gcctcttcca	33480
actagatcat	aaattctgag	ggtaggaact	tctgaatatt	tttgttcact	ggtctatctg	33540
cagctcagaa	caggacctgg	tactgaataa	atatTTTTga	aatgattgaa	tggatgaaaa	33600
gaaatgagta	ataagaatat	tacctaaggg	ggacagtgga	gataacaaag	gctttttcgg	33660
cttaggaaag	gaacagtagc	tatttgagag	tttgtcacta	gtgaggtgaa	ctggcaaagt	33720
gaaggaaact	gagcaacatt	ctagaaaatg	agaggaaatc	aaatacttag	gtgaaaggaa	33780
gtaaactctg	gaaatacaga	aggacacctc	ctaaggctag	aacagatatt	taggattgat	33840
aggcacttct	agctaattgac	tagggcctta	tatcctTTTT	aattttctag	gtggaatgca	33900
agcttttagga	cttcacctga	cagatccaag	tcaacgtctt	gttcagaact	gtctttggac	33960
tctcaggaat	ctttcagatg	ctgcaactaa	acaggtaaata	tctgagtaaa	ctggtgccat	34020
gggaatagag	tcaagatgag	tatgtgcttg	tactgaccat	ctgtttttat	ctccatagga	34080
agggatggaa	ggtctccttg	ggactcttgt	tcagcttctg	ggttcagatg	atataaatgt	34140
ggtcacctgt	gcagctggaa	ttcttttctaa	cctcacttgc	aataattata	agaacaagat	34200
gatggctctgc	caagtgggtg	gtatagaggc	tcttgtgcgt	actgtccttc	gggctgggtga	34260
cagggaagac	atcactgagc	ctgccatctg	tgctcttcgt	catctgacca	gccgacacca	34320
agaagcagag	atggcccaga	atgcagttcg	ccttcactat	ggactaccag	ttgtgggttaa	34380
gctcttacac	ccaccatccc	actggcctct	gataaaggta	aattgtcaaa	gtagaattta	34440
cctttgttgc	agaattgaaa	atgaagcatc	tctagctggt	ggatggctgt	ctaagcatag	34500
tgatcaataa	gtaggaattg	tattccttag	taagtaggaa	gtatggctgc	gataggggta	34560
agattctgaa	atgttttgtgt	agtcagaact	acttttagtt	gataccaata	gatttagtgt	34620
ggtgggaatt	ttagggtaag	aaaatgattt	tgttgagttg	tatgccagtt	cttccttctg	34680
tttttcaggc	tactgttgga	ttgattcgaa	atcttgccct	ttgtcccgca	aatcatgcac	34740
ctttgcgtga	gcagggtgcc	attccacgac	tagttcagtt	gcttgttcgt	gcacatcagg	34800
ataccagcg	ccgtacgtcc	atgggtggga	cacagcagca	atttgtggta	ggtaaattct	34860
tacagtgata	cctggctatc	taaaaggaat	gcataaatcc	aaaggatcct	gaacttcttt	34920
ctttggtcac	tggttcccc	catccgtctt	cctgaagagc	taatgacaaa	gtaaataaat	34980
aaataattac	acatttctat	ggctgcagag	aaaataaggc	atagtgtggc	cccagtgata	35040

tttccttgga	cacgtccttc	acatggtcag	tcttacaaag	gttgggtag	gtgtttcata	35100
aagtgttctc	atttaattta	cacaaaggcc	cacttcctta	ggaagaggta	gagtcataat	35160
ttgagatcaa	atctgtgtaa	tttcagagcc	tcttaccctt	gcctcatcat	gcattttgac	35220
tataaatatt	tagcagtccg	ttttattatc	ttttctgtga	gttaaacttt	tttcatggac	35280
ctaagaatat	tcagaaataa	gtagtagcat	ttctgtactc	ttaaccacaa	aaatctcaac	35340
ctgaagcttt	gatacaaagt	ttgtgtctta	aaagtagctt	cattaaaagt	atagtctaata	35400
gacatttctg	atttctcaga	ctttaagacc	ttattagggt	agtttagaaa	acaaagatgg	35460
agcctaccag	aacagatggt	aggaatctca	ttttgctggg	tgctttgtgt	atgtactcat	35520
attggggctt	tggttttctt	catttattac	tgttgggtatt	ggcccatctc	catgagggtga	35580
cttaatagaa	cggtgagggc	accttttatt	ttaaatctct	tttctaggaa	gaagagagtt	35640
tttgtgtcct	tgtaagaatc	aagttattta	taaaagctgc	taaatgtagc	agaataataa	35700
ccccttttaa	aactcaaadc	cagaaacagg	agaaacagat	ggtacttaca	tattgcaaaa	35760
gctatcttcc	ttctatacat	gaggctgtca	gctgaatagt	cttggaagag	tgaggagtga	35820
atthtttctgc	tggtcaactcg	gttagtttta	gcagttgggtg	ctaaaacttg	gcaaagttht	35880
caccaaatac	atggaagata	tacaaaaata	gagggggcat	gtaaaagaaa	aacgttgaca	35940
tagtctgagc	attactttct	catcttctct	ttttatatac	cttttaccba	gaatgattgg	36000
tgcccttact	gtaggaaagt	tgtctttggg	attcagcgct	gtatggaagc	tctgttgcac	36060
tgtgtatggg	ggaggggtgc	tgctttgaat	tagtgctgcc	aggaggcctc	ttttcagtga	36120
cattcaagtt	aatggaatcc	ttcttccttc	ctgaactaat	tgcaagttac	ggggaacttc	36180
gggtatataa	tgtaaataat	tacagtctaa	taattgttcc	tcaaacttta	cagaggagaa	36240
tgccctgttt	gttaaccatg	tttcttttgg	caggaggggg	tccgcatgga	agaaatagtt	36300
gaaggttgta	ccggagccct	tcacatccta	gctcgggatg	ttcacaaccg	aattgttatc	36360
agaggactaa	ataccattcc	attgtttgtg	cagggtatgtt	ttaagtgaag	tgthctaggt	36420
tttatgtcca	taaaatttcc	agattgtaat	gactaataac	atttcagaaa	attagggacc	36480
ataatagggt	taccaacatt	taattttatg	aaaattccct	acattttttg	gtcagtaaga	36540
gaaacattga	gacttgagaa	gagggaggag	atttcacatt	tcacttttat	gggtgcctag	36600
aggggagagc	tgacctgggc	tgccagaggc	agggcataga	cccccaacca	attctgggtt	36660

ttccaaatct tagatcagtt agagctgcct ctgaagaaag ggtttatagc taaaaaatat	36720
tatggaaatc cagtgcctca gagcattaaa caccccaaga cataaaattc agagaatatt	36780
atttactaca gtgtgaatgc ctcttgact ctgaattggg aatgtttgca ccacagtggg	36840
gggcttgcca tgtttttagct ttagatttaa ttaggttttg tttgtgtttt ctccttagct	36900
gctttattct cccattgaaa acatccaaag agtagctgca ggggtcctct gtgaacttgc	36960
tcaggacaag gaagctgcag aagctattga agctgaggga gccacagctc ctctgacaga	37020
gttacttcac tctaggaatg aaggtgtggg taagtaaaaa ggaaccaaag cctttagcag	37080
atgtgtacat tgaagtctca gtttttcctc aagggccttt ttctccttgt ctcttagcga	37140
catatgcagc tgctgttttg ttccgaatgt ctgaggacaa gccacaagat tacaagaaac	37200
ggctttcagt tgagctgacc agctctctct tcagaacaga gccaatggct tggaatgagg	37260
tagggaaatg tgagcagtta tttatctggt agtttcctag agcaggatat gcagcttggt	37320
ctttcctctc aaaacactta gtacacattc atttgcattg atgtttccct ggcttgagta	37380
tttcttcttt atgctgtcta gcaactgctc tgaggaagaa ctataatata agctttaaag	37440
agtctgttca gaatcattac aaataagttg tgttatttaa aattataatt cataagggag	37500
aaagatgaaa aatgtttacca gattaaagaa gatttttcaa aaggatgtaa ggaaagaggc	37560
agtgttaaac actgttaaga ggacagttta tcagtathtt ttactaaact ttaataaaac	37620
ttttctatht gaatttctgc tatgaathtt tcttcagcat ttgtcctcag tacaggtggt	37680
tccttgaaac attgtttcta ataaaactag aacatcctga tatttttatcc attctataga	37740
gatcattgat ggtacacaga catacagtgg attatgtttg ttgagtgaat ggaaagagag	37800
attgttaggt ttacaacgat gcagctcttg agaccggagt ttaagatcag cctgggcaac	37860
atagtgaaac cccatcttta gctgggcatg gagatggatg cctatagtcc tagctactgg	37920
ggagacgggg gcaggaggat tgcttgaacc caggagttaa cagactgcac tcagtgcag	37980
agccagactc caacacaaaa aaaaaaaaaa aaaaaaagca aattaccagt gagtagtgtg	38040
ttacttgggt ttttaatagg catcttatta acatgttcca acttgagccc ttaactttct	38100
ccacctacc ccttcacaa acctgttttc actgtcttct ctgtcttagt taatgtcagc	38160
tttgtctgtc cagctgctca ggctaaaact tttctttcat ataacacatc ctatcagcag	38220
ctcctgtttg tgggtaggca ttttgccttt tttttttttt ttttttttaa actgctatat	38280

ctctagcatg tagaacagtg cctggcagca cataataggt gcttaatata atatttggtg	38340
aaagaacaag tcagttagta tttttaatgt gaggtgcaaa gagaaaaaaa aatgtatctt	38400
tgaggtgtgg agttttgaag aacttccatt ttctaagcat ttgtgtaatg ttggagttac	38460
ttgttccttt tgtaatctga aagtatgctt taaaaaaaat tagtgtactt ttgagaattt	38520
tcattttgct ttctattctt ccttgctttg tgcatgttta tctagactgc tgatcttgga	38580
cttgatattg gtgcccaggg agaaccctt ggatatcgcc aggatggat gtgtctcata	38640
tttctcgatt aactccagat caagctaaag ttctaaaact tttatcagaa gagccggtt	38700
gctcatctgg gaaaccagtg ttggcagaaa agtagtggct tcaattaaaa gcagttctta	38760
aattccagtc agcaacagta tctttaatgg agcacaggga attcagagcc acacaatgag	38820
tagcagtagg attacaccac caacaaatac atgctactgc taggcctctg cagtgcagga	38880
tgttacaatt tacctggctt tttattctct ttttggccag aggactcata atacctttgt	38940
ctacaagcta cccaaggaag ataggaaaac tcctgtttct aggctcagat ctcgggtggg	39000
tttttacata gttgcattat catcagggtt ttcttgaaaa gctaatttaa atctgggtaa	39060
tgaacatgga ggatggcata gaccactaac aattataact gtcttacatt tataaccgca	39120
tctgcttcta cctaattatg aaaccactaa agcgcagatt cttactgtga gaaataacat	39180
gtcaacccta agataaaaata tgttgaggtt tcatggaaat agtgcctttc cttagtactt	39240
ttgtgggtgt cacttggcct ttttgtcaag atagattaca cctgccagac ctcattattg	39300
tcttaatcct ctttcccatg acttctcact gcctaggtgg tcacacagta gattcctgct	39360
tcttctcctc gggaaccca agtctcttga caggggtaaa tgcagagtgt tcagggttag	39420
actaatgatg tgactaggcc ctgctggtgt gcctgtctga tggaaataga tgttatttgt	39480
gtagtctcat ggggtggcctg gcactgagta attacttggc taaagaaagc tggaggttga	39540
agaggctaga aagcgttggtt ttctgacaag tttgctgctg aactttggat gccctaacct	39600
cagtgttaac gtctatgtct gcttctctcc tctctctttt gccttccttc ttgcctattt	39660
tgttgacacc ctgactcttc tagatcctag ctatcgttct tttcactctg gtggatatgg	39720
ccaggatgcc ttgggtatgg accccatgat ggaacatgag atgggtggcc accaccctgg	39780
tgctgactat ccagttgatg ggctgccaga tctggggcat gcccaggacc tcatggatgg	39840
gctgcctcca ggtgacagca atcagctggc ctggtttgat actgacctgt aaatcatcct	39900

ttaggtaaga agttttaaaa agccagtttg ggtaaaatac ttttactctg cctacagaac	39960
ttcagaaaga cttggttggg aggggtgggag tggtttaggc tatttgtaaa tctgccacaa	40020
aaacaggtat atactttgaa aggagatgtc ttggaacatt ggaatgttct cagatttctg	40080
gttgttatgt gatcatgtgt ggaagttatt aactttaatg ttttttgcca cagcttttgc	40140
aacttaatac tcaaatgagt aacatttgct gttttaaaca ttaatagcag cctttctctc	40200
tttatacagc tgtattgtct gaacttgcac tgtgattggc ctgtagagtt gctgagaggg	40260
ctcgaggggt gggctgggat ctgagaaagt gcctgacaca ctaaccaagc tgagtttcct	40320
atgggaacaa ttgaagtaaa ctttttgttc tggtcctttt tggtcgagga gtaacaatac	40380
aaatggattt tgggagtgac tcaagaagtg aagaatgcac aagaatggat cacaagatgg	40440
aatttatcaa accctagcct tgcttggtta attttttttt tttttttttt aagaatatct	40500
gtaatggtac tgactttgct tgctttgaag tagctctttt tttttttttt tttttttttt	40560
tgtagtaact gttttttaag tctctcgtag tgtttaagtta tagtgaatac tgctacagca	40620
atttctaatt tttagaatt gagtaatggg gtagaacact aattcataat cactctaatt	40680
aattgtaatc tgaataaagt gtaacaattg tgtagccttt ttgtataaaa tagacaaata	40740
gaaaatggtc caattagttt cttttttaat atgcttaaaa taagcagggt gatctatttc	40800
atgtttttga tcaaaaacta tttgggatat gtatgggtag ggtaaatcag taagaggtgt	40860
tatttggaac cttgttttgg acagtttacc agttgccttt tatcccaaag ttgttgtaac	40920
ctgctgtgat acgatgcttc aagagaaaat gcggttataa aaaatgggtc agaattaaac	40980
ttttaattca ttcgattg	40998

<210> 16
 <211> 232
 <212> PRT
 <213> Homo sapiens

<400> 16

Met	Glu	Arg	Cys	Pro	Ser	Leu	Gly	Val	Thr	Leu	Tyr	Ala	Leu	Val	Val
1				5					10					15	

Val	Leu	Gly	Leu	Arg	Ala	Thr	Pro	Ala	Gly	Gly	Gln	His	Tyr	Leu	His
			20					25					30		

Ile Arg Pro Ala Pro Ser Asp Asn Leu Pro Leu Val Asp Leu Ile Glu
35 40 45

His Pro Asp Pro Ile Phe Asp Pro Lys Glu Lys Asp Leu Asn Glu Thr
50 55 60

Leu Leu Arg Ser Leu Leu Gly Gly His Tyr Asp Pro Gly Phe Met Ala
65 70 75 80

Thr Ser Pro Pro Glu Asp Arg Pro Gly Gly Gly Gly Gly Ala Ala Gly
85 90 95

Gly Ala Glu Asp Leu Ala Glu Leu Asp Gln Leu Leu Arg Gln Arg Pro
100 105 110

Ser Gly Ala Met Pro Ser Glu Ile Lys Gly Leu Glu Phe Ser Glu Gly
115 120 125

Leu Ala Gln Gly Lys Lys Gln Arg Leu Ser Lys Lys Leu Arg Arg Lys
130 135 140

Leu Gln Met Trp Leu Trp Ser Gln Thr Phe Cys Pro Val Leu Tyr Ala
145 150 155 160

Trp Asn Asp Leu Gly Ser Arg Phe Trp Pro Arg Tyr Val Lys Val Gly
165 170 175

Ser Cys Phe Ser Lys Arg Ser Cys Ser Val Pro Glu Gly Met Val Cys
180 185 190

Lys Pro Ser Lys Ser Val His Leu Thr Val Leu Arg Trp Arg Cys Gln
195 200 205

Arg Arg Gly Gly Gln Arg Cys Gly Trp Ile Pro Ile Gln Tyr Pro Ile
210 215 220

Ile Ser Glu Cys Lys Cys Ser Cys
225 230

<210> 17

<211> 232
<212> PRT
<213> Mus musculus

<400> 17

Met Glu Arg Cys Pro Ser Leu Gly Val Thr Leu Tyr Ala Leu Val Val
1 5 10 15

Val Leu Gly Leu Arg Ala Ala Pro Ala Gly Gly Gln His Tyr Leu His
20 25 30

Ile Arg Pro Ala Pro Ser Asp Asn Leu Pro Leu Val Asp Leu Ile Glu
35 40 45

His Pro Asp Pro Ile Phe Asp Pro Lys Glu Lys Asp Leu Asn Glu Thr
50 55 60

Leu Leu Arg Ser Leu Leu Gly Gly His Tyr Asp Pro Gly Phe Met Ala
65 70 75 80

Thr Ser Pro Pro Glu Asp Arg Pro Gly Gly Gly Gly Gly Pro Ala Gly
85 90 95

Gly Ala Glu Asp Leu Ala Glu Leu Asp Gln Leu Leu Arg Gln Arg Pro
100 105 110

Ser Gly Ala Met Pro Ser Glu Ile Lys Gly Leu Glu Phe Ser Glu Gly
115 120 125

Leu Ala Gln Gly Lys Lys Gln Arg Leu Ser Lys Lys Leu Arg Arg Lys
130 135 140

Leu Gln Met Trp Leu Trp Ser Gln Thr Phe Cys Pro Val Leu Tyr Ala
145 150 155 160

Trp Asn Asp Leu Gly Ser Arg Phe Trp Pro Arg Tyr Val Lys Val Gly
165 170 175

Ser Cys Phe Ser Lys Arg Ser Cys Ser Val Pro Glu Gly Met Val Cys
180 185 190

Lys Pro Ser Lys Ser Val His Leu Thr Val Leu Arg Trp Arg Cys Gln
195 200 205

Arg Arg Gly Gly Gln Arg Cys Gly Trp Ile Pro Ile Gln Tyr Pro Ile
210 215 220

Ile Ser Glu Cys Lys Cys Ser Cys
225 230

<210> 18
<211> 232
<212> PRT
<213> Rattus norvegicus

<400> 18

Met Glu Arg Cys Pro Ser Leu Gly Val Thr Leu Tyr Ala Leu Val Val
1 5 10 15

Val Leu Gly Leu Arg Ala Ala Pro Ala Gly Gly Gln His Tyr Leu His
20 25 30

Ile Arg Pro Ala Pro Ser Asp Asn Leu Pro Leu Val Asp Leu Ile Glu
35 40 45

His Pro Asp Pro Ile Phe Asp Pro Lys Glu Lys Asp Leu Asn Glu Thr
50 55 60

Leu Leu Arg Ser Leu Leu Gly Gly His Tyr Asp Pro Gly Phe Met Ala
65 70 75 80

Thr Ser Pro Pro Glu Asp Arg Pro Gly Gly Gly Gly Gly Pro Ala Gly
85 90 95

Gly Ala Glu Asp Leu Ala Glu Leu Asp Gln Leu Leu Arg Gln Arg Pro
100 105 110

Ser Gly Ala Met Pro Ser Glu Ile Lys Gly Leu Glu Phe Ser Glu Gly
115 120 125

Leu Ala Gln Gly Lys Lys Gln Arg Leu Ser Lys Lys Leu Arg Arg Lys
130 135 140

Leu Gln Met Trp Leu Trp Ser Gln Thr Phe Cys Pro Val Leu Tyr Ala
145 150 155 160

Trp Asn Asp Leu Gly Ser Arg Phe Trp Pro Arg Tyr Val Lys Val Gly
165 170 175

Ser Cys Phe Ser Lys Arg Ser Cys Ser Val Pro Glu Gly Met Val Cys
180 185 190

Lys Pro Ser Lys Ser Val His Leu Thr Val Leu Arg Trp Arg Cys Gln
195 200 205

Arg Arg Gly Gly Gln Arg Cys Gly Trp Ile Pro Ile Gln Tyr Pro Ile
210 215 220

Ile Ser Glu Cys Lys Cys Ser Cys
225 230

<210> 19
<211> 699
<212> DNA
<213> Homo sapiens

<400> 19
atggagcgct gccccagcct aggggtcacc ctctacgcc ttggtggtggt cctggggctg 60
cgggcgacac cggccggcgg ccagcactat ctccacatcc gcccggcacc cagcgacaac 120
ctgcccctgg tggacctcat cgaacacca gaccctatct ttgaccccaa ggaaaaggat 180
ctgaacgaga cgctgctgcg ctcgctgctc gggggccact acgaccagg cttcatggcc 240
acctcgcccc ccgaggaccg gcccggcggg ggcgggggtg cagctggggg cgcggaggac 300
ctggcggagc tggaccagct gctgcggcag cggccgtcgg gggccatgcc gagcgagatc 360
aaagggctag agttctccga gggcttggcc cagggaaga agcagcgcct aagcaagaag 420
ctgcggagga agttacagat gtggctgtgg tcgcagacat tctgccccgt gctgtacgcg 480
tggaacgacc tgggcagccg cttttggccg cgctacgtga aggtgggcag ctgcttcagt 540
aagcgctcgt gctccgtgcc cgagggcatt gtgtgcaagc cgtccaagtc cgtgcacctc 600
acggtgctgc ggtggcgctg tcagcggcgc gggggccagc gctgcggctg gattcccatc 660
cagtaccca tcatttccga gtgcaagtgc tcgtgctag 699

<210> 20
<211> 1892
<212> DNA
<213> Homo sapiens

<400> 20
aaaccggtgc caacgtgcgc ggacgccgcc gccgccgccg ccgctggagt ccgccgggca 60
gagccggccg cggagcccgg agcaggcggga gggaagtgcc cctagaacca gctcagccag 120
cggcgcttgc acagagcggc cggacgaaga gcagcgagag gaggagggga gagcggctcg 180
tccacgcgcc ctgcgccgcc gccggcccgg gaaggcagcg aggagccggc gcctcccgcg 240
ccccgcggtc gccctggagt aatttcggat gccagccgc ggccgccttc ccagtagac 300
ccgggagagg agttgcggcc aacttggtgt cttttcttcc gccccggtgg gagccggcgc 360
tgcgcgaagg gctctcccg cggctcatgc tgccggccct gcgcctgccc agcctcgggt 420
gagccgcctc cggagagacg ggggagcgcg gcggcgccgc gggctcggcg tgctctcctc 480
cggggacgcg ggacgaagca gcagccccgg gcgcgcgcca gaggcattga gcgctgcccc 540
agcctagggg tcaccctcta cgccctggtg gtggtccttg ggctgcgggc gacaccggcc 600
ggcggccagc actatctcca catccgcccg gcacccagcg acaacctgcc cctggtggac 660
ctcatcgaac acccagaccc tatctttgac cccaaggaaa aggatctgaa cgagacgctg 720
ctgcgctcgc tgctcggggg ccactacgac ccaggcttca tggccacctc gccccccgag 780
gaccggcccc gcggggggcg ggggtgcagct gggggcgcgagg agcacctggc ggagctggac 840
cagctgctgc ggcagcggcc gtcggggggc atgccgagcg agatcaaagg gctagagttc 900
tccgagggct tggcccaggg caagaagcag cgcctaagca agaagctgcg gaggaagtta 960
cagatgtggc tgtggtcgca gacattctgc cccgtgctgt acgcgtggaa cgacctgggc 1020
agccgctttt ggccgcgcta cgtgaagggt ggagctgct tcagtaagcg ctcgtgctcc 1080
gtgcccagg gcatggtgtg caagccgtcc aagtccgtgc acctcacggg gctgcgggtg 1140
cgctgtcagc ggcgcggggg ccagcgctgc ggctggattc ccatccagta ccccatcatt 1200
tccgagtgca agtgctcgtg ctagaactcg ggggccccct gcccgcaccc ggacacttga 1260
tcgatcccca ccgacgcccc ctgcaccgcc tccaaccagt tccaccaccc tctagcgagg 1320
gttttcaatg aacttttttt tttttttttt tttttttttc tgggctacag agacctagct 1380

ttctggttcc tgtaatgcac tgtttaactg tgtaggaatg tatatgtgtg tgtatatacg	1440
gtcccagttt taatttactt attaaaaggt cagtattata cgttaaaagt taccggcttc	1500
tactgtattt ttaaaaaaaaa gtaagcaaaa gaaaaaaaaa agaacagaga aaagagagac	1560
ttattctggt tgttgctaata atgtttaacc tgctatttat attccagtgc ccttcgcatg	1620
gcgaagcagg ggggaaaagt tttttttttc ttgaagtaca aagagacggg ggaacttttg	1680
tagaggactt tttaaaagct attttccatt cttcggaag tgttttgggt ttccttggac	1740
ctcgaagaag ctatagagtt caatgttatt ttacagttat tgtaaataata gagaacaaat	1800
ggaatgacta atcattgtaa attaagagta tctgctattt attctttata atatcccggtg	1860
tagtaaata gaaagaagt cagagcagga tt	1892

<210> 21
 <211> 422
 <212> PRT
 <213> Homo sapiens

<400> 21

Met	Pro	Arg	Ser	Phe	Leu	Val	Lys	Ser	Lys	Lys	Ala	His	Ser	Tyr	His
1				5					10					15	

Gln	Pro	Arg	Ser	Pro	Gly	Pro	Asp	Tyr	Ser	Leu	Arg	Leu	Glu	Asn	Val
			20					25					30		

Pro	Ala	Pro	Ser	Arg	Ala	Asp	Ser	Thr	Ser	Asn	Ala	Gly	Gly	Ala	Lys
		35					40					45			

Ala	Glu	Pro	Arg	Asp	Arg	Leu	Ser	Pro	Glu	Ser	Gln	Leu	Thr	Glu	Ala
	50					55					60				

Pro	Asp	Arg	Ala	Ser	Ala	Ser	Pro	Asp	Ser	Cys	Glu	Gly	Ser	Val	Cys
65					70					75					80

Glu	Arg	Ser	Ser	Glu	Phe	Glu	Asp	Phe	Trp	Arg	Pro	Pro	Ser	Pro	Ser
				85					90					95	

Ala	Ser	Pro	Ala	Ser	Glu	Lys	Ser	Met	Cys	Pro	Ser	Leu	Asp	Glu	Ala
			100					105						110	

Gln Pro Phe Pro Leu Pro Phe Lys Pro Tyr Ser Trp Ser Gly Leu Ala
115 120 125

Gly Ser Asp Leu Arg His Leu Val Gln Ser Tyr Arg Pro Cys Gly Ala
130 135 140

Leu Glu Arg Gly Ala Gly Leu Gly Leu Phe Cys Glu Pro Ala Pro Glu
145 150 155 160

Pro Gly His Pro Ala Ala Leu Tyr Gly Pro Lys Arg Ala Ala Gly Gly
165 170 175

Ala Gly Ala Gly Ala Pro Gly Ser Cys Ser Ala Gly Ala Gly Ala Thr
180 185 190

Ala Gly Pro Gly Leu Gly Leu Tyr Gly Asp Phe Gly Ser Ala Ala Ala
195 200 205

Gly Leu Tyr Glu Arg Pro Thr Ala Ala Ala Gly Leu Leu Tyr Pro Glu
210 215 220

Arg Gly His Gly Leu His Ala Asp Lys Gly Ala Gly Val Lys Val Glu
225 230 235 240

Ser Glu Leu Leu Cys Thr Arg Leu Leu Leu Gly Gly Gly Ser Tyr Lys
245 250 255

Cys Ile Lys Cys Ser Lys Val Phe Ser Thr Pro His Gly Leu Glu Val
260 265 270

His Val Arg Arg Ser His Ser Gly Thr Arg Pro Phe Ala Cys Glu Met
275 280 285

Cys Gly Lys Thr Phe Gly His Ala Val Ser Leu Glu Gln His Lys Ala
290 295 300

Val His Ser Gln Glu Arg Ser Phe Asp Cys Lys Ile Cys Gly Lys Ser
305 310 315 320

Phe Lys Arg Ser Ser Thr Leu Ser Thr His Leu Leu Ile His Ser Asp
325 330 335

Thr Arg Pro Tyr Pro Cys Gln Tyr Cys Gly Lys Arg Phe His Gln Lys
340 345 350

Ser Asp Met Lys Lys His Thr Phe Ile His Thr Gly Glu Lys Pro His
355 360 365

Lys Cys Gln Val Cys Gly Lys Ala Phe Ser Gln Ser Ser Asn Leu Ile
370 375 380

Thr His Ser Arg Lys His Thr Gly Phe Lys Pro Phe Gly Cys Asp Leu
385 390 395 400

Cys Gly Lys Gly Phe Gln Arg Lys Val Asp Leu Arg Arg His Arg Glu
405 410 415

Thr Gln His Gly Leu Lys
420

<210> 22
<211> 423
<212> PRT
<213> Mus musculus

<400> 22

Met Pro Arg Ser Phe Leu Val Lys Ser Lys Lys Ala His Ser Tyr His
1 5 10 15

Gln Pro Arg Ser Pro Gly Pro Asp Tyr Ser Leu Arg Leu Glu Thr Val
20 25 30

Pro Ala Pro Gly Arg Ala Glu Gly Gly Ala Val Ser Ala Gly Glu Ser
35 40 45

Lys Met Glu Pro Arg Glu Arg Leu Ser Pro Asp Ser Gln Leu Thr Glu
50 55 60

Ala Pro Asp Arg Ala Ser Ala Ser Pro Asn Ser Cys Glu Gly Ser Val
65 70 75 80

Cys Asp Pro Cys Ser Glu Phe Glu Asp Phe Trp Arg Pro Pro Ser Pro

85

90

95

Ser Val Ser Pro Ala Ser Glu Lys Ser Leu Cys Arg Ser Leu Asp Glu
 100 105 110

Ala Gln Pro Tyr Thr Leu Pro Phe Lys Pro Tyr Ala Trp Ser Gly Leu
 115 120 125

Ala Gly Ser Asp Leu Arg His Leu Val Gln Ser Tyr Arg Gln Cys Ser
 130 135 140

Ala Leu Glu Arg Ser Ala Gly Leu Ser Leu Phe Cys Glu Arg Gly Ser
 145 150 155 160

Glu Pro Gly Arg Pro Ala Ala Arg Tyr Gly Pro Glu Gln Ala Ala Gly
 165 170 175

Gly Ala Gly Ala Gly Gln Pro Gly Arg Cys Gly Val Ala Gly Gly Ala
 180 185 190

Thr Ser Ala Ala Gly Leu Gly Leu Tyr Gly Asp Phe Ala Pro Ala Ala
 195 200 205

Ala Gly Leu Tyr Glu Arg Pro Ser Thr Ala Ala Gly Arg Leu Tyr Gln
 210 215 220

Asp His Gly His Glu Leu His Ala Asp Lys Ser Val Gly Val Lys Val
 225 230 235 240

Glu Ser Glu Leu Leu Cys Thr Arg Leu Leu Leu Gly Gly Gly Ser Tyr
 245 250 255

Lys Cys Ile Lys Cys Ser Lys Val Phe Ser Thr Pro His Gly Leu Glu
 260 265 270

Val His Val Arg Arg Ser His Ser Gly Thr Arg Pro Phe Ala Cys Glu
 275 280 285

Met Cys Gly Lys Thr Phe Gly His Ala Val Ser Leu Glu Gln His Lys
 290 295 300

Ala Val His Ser Gln Glu Arg Ser Phe Asp Cys Lys Ile Cys Gly Lys
305 310 315 320

Ser Phe Lys Arg Ser Ser Thr Leu Ser Thr His Leu Leu Ile His Ser
325 330 335

Asp Thr Arg Pro Tyr Pro Cys Gln Tyr Cys Gly Lys Arg Phe His Gln
340 345 350

Lys Ser Asp Met Lys Lys His Thr Phe Ile His Thr Gly Glu Lys Pro
355 360 365

His Lys Cys Gln Val Cys Gly Lys Ala Phe Ser Gln Ser Ser Asn Leu
370 375 380

Ile Thr His Ser Arg Lys His Thr Gly Phe Lys Pro Phe Gly Cys Asp
385 390 395 400

Leu Cys Gly Lys Gly Phe Gln Arg Lys Val Asp Leu Arg Arg His Arg
405 410 415

Glu Thr Gln His Gly Leu Lys
420

<210> 23
<211> 423
<212> PRT
<213> Rattus norvegicus

<400> 23

Met Pro Arg Ser Phe Leu Val Lys Ser Lys Lys Ala His Ser Tyr His
1 5 10 15

Gln Pro Arg Ser Pro Gly Pro Asp Tyr Ser Leu Arg Leu Glu Thr Val
20 25 30

Pro Val Pro Gly Arg Ala Asp Gly Gly Ala Val Ser Ala Gly Glu Ser
35 40 45

Lys Met Glu Pro Arg Glu Arg Leu Ser Pro Glu Ser Gln Leu Thr Glu
50 55 60

Ala Pro Asp Arg Ala Ser Ala Ser Pro Asn Ser Cys Glu Gly Ser Val
65 70 75 80

Cys Asp Pro Ser Ser Glu Phe Glu Asp Tyr Trp Arg Pro Pro Ser Pro
85 90 95

Ser Val Ser Pro Ala Ser Glu Lys Ser Leu Cys Arg Ser Leu Asp Glu
100 105 110

Ala Gln Pro Tyr Thr Leu Pro Phe Lys Pro Tyr Ala Trp Ser Gly Leu
115 120 125

Ala Gly Ser Asp Leu Arg His Leu Val Gln Ser Tyr Arg Gln Cys Ser
130 135 140

Ala Leu Glu Arg Ser Ala Gly Leu Ser Leu Phe Cys Glu Arg Gly Ala
145 150 155 160

Glu Ser Gly Arg Pro Ala Ala Arg Tyr Gly Ser Glu Gln Ala Ala Gly
165 170 175

Gly Ala Gly Ala Gly Gln Pro Gly Ser Cys Gly Ala Ala Ser Gly Ala
180 185 190

Thr Ser Ala Gly Gly Leu Gly Leu Tyr Gly Asp Phe Ala Pro Ala Ala
195 200 205

Ala Gly Leu Phe Glu Arg Pro Ser Thr Ala Ala Gly Arg Leu Tyr Gln
210 215 220

Asp Arg Gly His Glu Leu His Ala Asp Lys Ser Val Gly Val Lys Val
225 230 235 240

Glu Ser Glu Leu Leu Cys Thr Arg Leu Leu Leu Gly Gly Gly Ser Tyr
245 250 255

Lys Cys Ile Lys Cys Ser Lys Val Phe Ser Thr Pro His Gly Leu Glu
260 265 270

Val His Val Arg Arg Ser His Ser Gly Thr Arg Pro Phe Ala Cys Glu
 275 280 285

Met Cys Gly Lys Thr Phe Gly His Ala Val Ser Leu Glu Gln His Lys
 290 295 300

Ala Val His Ser Gln Glu Arg Ser Phe Asp Cys Lys Ile Cys Gly Lys
 305 310 315 320

Ser Phe Lys Arg Ser Ser Thr Leu Ser Thr His Leu Leu Ile His Ser
 325 330 335

Asp Thr Arg Pro Tyr Pro Cys Gln Tyr Cys Gly Lys Arg Phe His Gln
 340 345 350

Lys Ser Asp Met Lys Lys His Thr Phe Ile His Thr Gly Glu Lys Pro
 355 360 365

His Lys Cys Gln Val Cys Gly Lys Ala Phe Ser Gln Ser Ser Asn Leu
 370 375 380

Ile Thr His Ser Arg Lys His Thr Gly Phe Lys Pro Phe Gly Cys Asp
 385 390 395 400

Leu Cys Gly Lys Gly Phe Gln Arg Lys Val Asp Leu Arg Arg His Arg
 405 410 415

Glu Thr Gln His Gly Leu Lys
 420

<210> 24
 <211> 1269
 <212> DNA
 <213> Homo sapiens

<400> 24
 atgccgcgct catttctcgt caaaagcaag aaggctcaca gctaccacca gccgcgctcc 60
 ccaggaccag actattccct ccgttttagag aatgtaccgg cgcctagccg agcagacagc 120
 acttcaaattg caggcggggc gaaggcggag ccccgggacc gtttgtcccc cgaatcgag 180
 ctgaccgaag cccagacag agcctccgca tccccagaca gctgcgaagg cagcgtctgc 240

gaacggagct cggagtttga ggactttctgg aggccccgt caccctccgc gtctccagcc	300
tcggagaagt caatgtgccc atcgctggac gaagcccagc ctttccccct gcctttcaaa	360
ccgtactcat ggagcggcct ggcggttct gacctgcggc acctggtgca gagctaccga	420
ccgtgtgggg ccctggagcg tggcgctggc ctgggcctct tctgcgaacc cgccccggag	480
cctggccacc cggccgcgct gtacggcccc aagcgggctg ccggcggcgc gggggccggg	540
gcgccagggga gctgcagcgc aggggcccgt gccaccgctg gccctggcct agggctctac	600
ggcgacttcg ggtctgcggc agccgggctg tatgagaggc ccacggcagc ggcgggcttg	660
ctgtaccccg agcgtggcca cgggctgcac gcagacaagg gcgctggcgt caagggtggag	720
tcggagctgc tgtgcacccg cctgctgctg ggcgggcggt cctacaagt catcaagtgc	780
agcaaggtgt tctccacgcc gcacgggctc gaggtgcacg tgcgcaggtc ccacagcggg	840
accagaccct ttgcctgcga gatgtgcggc aagaccttcg ggcacgcggg gagcctggag	900
cagcaciaag ccgtgcactc gcaggaacgg agctttgact gtaagatctg tgggaagagc	960
ttcaagaggt catccacact gtccacacac ctgcttatcc actcagacac tcggccctac	1020
ccctgtcagt actgtggcaa gaggttccac cagaagtcag acatgaagaa acacactttc	1080
atccacactg gtgagaagcc tcacaagtgc cagggtgtgcg gcaaggcatt cagccagagc	1140
tccaacctca tcacccacag ccgcaaacac acaggcttca agcccttcgg ctgcgacctc	1200
tgtgggaagg gtttcagag gaaggtggac ctccgaaggc accgggagac gcagcatggg	1260
ctcaaatga	1269

<210> 25
 <211> 12116
 <212> DNA
 <213> Homo sapiens

<400> 25	
gagggtgcgc ccaccgtcc cgccgggcgc ccgcgggacg cgccgccagg gccctctccg	60
ccgggggctc ggcgctcgcc cacctcttcc aaatttaacc attacctaaa tccgaaggga	120
aatgagcaaa cctctcggat tgggtgtcaa ggtattttca gcctcgttgg gcgtatttat	180
ccccaaagtgt ttccacaaca agctatttcg gggcctgcgg ggcaggtttc gctctgcgga	240
cgccgtggcc actcgccggg ctccaggccg gcggcaccgc gggccggtga ttcacggtcc	300
cgacccgggg gtggtgcagc ctaggaggc ggcggggtcg ggggtggggg gggcgggtga	360

ccgaggcctg aggggcgggg agggtcctcg gagcggggcg ccccccaccc ctctctcgcc	420
agtcaatctg tgtcctcaat ctgtggcttc tctcgctgcg gaagtctccc tggagccaag	480
aatagttcat tttctttcaa gtcatttcta gtgcctaagt gtccggacct ccaatttccc	540
ccatcccctg ccgaccacaca gggagagaac tgggaggact actaaggggc gcgcgggacg	600
ggctggaaag gccaggcccc ccaccgcctg gccacttgcg caaaggagcg cggccggccg	660
cccgacgggg gttgggagca ggtctgggag ggctatgcga gcgactcagt aacgctcagg	720
aagtgaagct tgtggttttg ggggctgagc tcggaaggag aatttttttt ttttttaagt	780
cagagagata gagcgggtctc tcccgaagc aagattccgt ttgaaaactc tcctagcgcg	840
gtgcccgcgc cgtgactccg caggtaggtc cgccgagcct gttctgcgcc tcctgccctg	900
gtggggggcgg ccgcggggac tcgcagagca ctggcactgc gggggcgatc agagggcgcg	960
ggcggtttcc cacctgctgc ggaccgccgt gcgggagccc agagagctcc ggagctggg	1020
ttagggcgcg acccgcgag tgtgactgga aatctggagc tggggcgcg cagcaggcgg	1080
tctggtggtt cggcagggga gccaaatcca ccagggaagg aaacatctgg tggggaggcg	1140
gcggcagctg cgctcgggag gacgccccct tagcgccctc ggctccctcc ttcctgggcc	1200
cggacggtga ggagaggcct gagcgcgcg aggggcccgc ccacctccg cgccagctgc	1260
agcgcggggt tgccttccca cgcgctcggc ccaggccccg gggcccctat ctctccaaa	1320
ctctgtcgct cccacagga accagcaaag cgggccgggg tgcgagagag gcgtgggctt	1380
acagcccggg tggggaggcg gcctccgcgt ccgcctggtc tctggtggcg ccagcccaga	1440
cccagctccg gcgctgacag ttacccgcc cccatctgtc ccgctcccag ccaacgtggg	1500
tccaagctgc agcgggacct tcgggacaac gccactccgt ttttctttt ttccgagttt	1560
cgtggctgtt taaagaattg ggtttgggt ttgtggcgtc taattgtacg gacgagaagt	1620
gcaggaagcg acaaagctct agccctagga gccaccccg agggaggcgg atggcactct	1680
cacccctagg ggcattctgg cgcttgggta gcgggaaact tcgcgggagc ccgcgacac	1740
gtcccaggcg tcttttctcc caggtctatt cccattcctc cggagaaggg gacacaatgg	1800
ggctggggat ctggagcagg gggcctgcac cctacaggga ccaaggcctg taggactcgt	1860
ttgagctgag agcgccaacg gacagacgta gactgtgtgg cctgcatctt gcctaggaag	1920
ccgaggggct cctagtccgg cagtggaaac agcgcgaagc cggaggactg caggtcctgc	1980

cccgggcccag agttcccagc accctcgttt ctgaaccagc cgaggccacg gagaactgct	2040
gtactgcagc tcacgtgtga acccggtcac catcgccttc accccgggag gaaggcagat	2100
tcgtttactc cagaccacct cgactgtggg gtaccgcccc cggagccggc tggagcctag	2160
cggcaggcgc agccacgccc tcccgtgcg ctcagatttc gacctggtat taggtgaact	2220
gattgggggt taatgagagc gacgccccgg gcagctagtt ccctcccggg cccgggcccg	2280
acccccgctc tctgctaata cagcctgcgc gctctggcgt cctgtctttt ttgtctgcta	2340
aattgtcggg gcactaccga ctcgggacac ctagcatttc ccagtcaacg ttcgtggatc	2400
gggctccacc tccctaggac aagatTTTTT ggtgagcaga acggaaagtg cTTTTTccc	2460
gggacctgat tcccagggtt aggtctccat ggtctgggat ggctcgccgc agcctcgacc	2520
ggtgcccgcc gcagccggga gtccaagggc aaagtttctc ctacgtgggg cactagttag	2580
gggcgagtgg gatcacccag atgcgagttt ctcctggggc gggggtttgt cgtctgttgt	2640
tcccctcact ttctgtccc ttgtctctcc cacctccttt ctctggcctc tgctgtcccc	2700
aatccctctg ctgctgtcct cccgccgcc cacagtTTTA tcttggtgtc tgttcccccc	2760
tcccccggt ctttctactc cagttggtag ctggctcttg gaggtcttgg ctccttgccc	2820
cttccgggtc ctcgaccact ggcatcccc ggcccctaaa ccgatccgcg tgtccccgcc	2880
ctccctcgcc agccgtaaag cacagccagg caggcgatga gtagctgagt tggggtaacc	2940
cacccgatgg gaactacagc tctccaggga gtttgattgc cggagcgagc ttcgctagga	3000
aaggggagga gctggggggc gtgggcaggg aggaggaaag gggcctgaga cagggccccg	3060
ggacaggttt taccgctgag ctgtgtcagt ggcggcggcg gcaacgacgg cgggttcgcg	3120
ccacctgtcc aagtgccacc tggtaagcgc ggcgagcag ggtcaagccc ctctcccgt	3180
gggccctctg cgcgcctccc tggcccgcgc tctccctccg cctgggtgcc cagtccgccg	3240
caccggagag tactgacca cgtctccacc tggTTTTctt tccctctcag gtctctccg	3300
ggctggggct gagcaagccc tcggagtgc cgtgggtgac agcggctcca gggactcttg	3360
gggcgcagtg gggaaagtgc cggaccacca tgccgcgctc atttctcgtc aaaagcaaga	3420
aggctcacag ctaccaccag ccgcgctccc caggaccaga ctattccctc cgtttagaga	3480
atgtaccggc gcctagccga gcaggtgcga ggcgcgcgcg ggccaggcgg ggctgctccc	3540
ccggatgcct actgcacctc ggcacaccat tagtccggag ctgggagggg ctgccccaac	3600

gtcccttttg ctgctgtttt tgtttcctac tgtcctgggt cctccgggtt tgtctcctag	3660
gtgccatggc ctctctgcgc ctgccctcgg atccgagagg gttcccggcc ggggtctggg	3720
tggagagggg aagacgctcg gctgccctgg tcgggggatt gggggagcct tcagcacctt	3780
cagactcaac cggtcccagc ctgagcccct cacctgcctc ctctttgcag acagcacttc	3840
aaatgcaggc ggggcgaagg cggagccccg ggaccgtttg tccccgaat cgcagctgac	3900
cgaagcccca gacagagcct ccgcatcccc agacagctgc gaaggcagcg tctgcgaacg	3960
gagctcggag tttgaggact tctggaggcc cccgtcacc cccgcgtctc caggtaggaa	4020
cccactggga acctcttggg cgggagctgc agggaccgg cagtgtctgg ggggaattgg	4080
cgcgaccttg ggcgtagaaa tgctaacggg gagttggaga gtctttccgg gagaaggag	4140
ctgattcgta ggggaaggag gcatccggct tctctgggac ttggacagct tgcccgtgg	4200
ggctgtgcc tccatcccag gcggcaggac cctagctgct tgtcgcttag attcgtttgc	4260
gcggagctgg ccagtgcgg aaaacaaacc agtcgtttcg actggcggca acgctgacct	4320
ttcattttct gaccagacct gactgtttta taaagtgcta ggatcctgca atctagacct	4380
ccaaacctca aacagagaac agggcagaac gggccaggca gaggagctag gcgctgggcg	4440
gcagggaggg ggcaggacga aaatctcagc ccgcggcttg gtcttcacag gcgcagattg	4500
ggggcctgtt tcatttttctg ttttgccggg ttaacctagc ctcaggggcc tgctctctgg	4560
gtttcatttc cagcgagcaa tccagcttca ggcaaactaa gtgaccacac gttgggtggg	4620
ggcgtctcga gtcccggccg ggggaaggaa tgagcagacc agccggattc tgtcaagggc	4680
cggttatatc cagaatatgt ttgctagttt tagaagatac caccacccgt ccacaaatca	4740
gtgagttgac ttggcgaaaa ccatagctcc agcaagtgtg tctgggagcc ggccggcgga	4800
ggattcttcc tgccagggcg tcaagtggcc agacaaggat tgggcgcgcc ccgaaccctt	4860
ccgaacgaaa ctccgggtac agcctctcac tgaagtggcc agcctgaact ggagtgttgt	4920
gcgcacacac acacacacac atttgtaaat gccgtatgca ctcatatgcg ttggggtcac	4980
tagtttttagc aaaattcacg tgggtggggg cgtagcaggc cgagaattca gagctgtctc	5040
cttgaggtg gtggctaaac cttatgagtt atagttattc tctgagaaat tcaggttccc	5100
cgctccatc aaactgtaac aggaatgggg agtatttggc tgtcaattta agccaaaag	5160
cccctttcct gctgctcctt tgctacgtac actgggcact taacttcgtg aaatcttaat	5220

accttcgggt ttattcagac agcagccttt cgggtagttc ggggccgcat ttatggacct	5280
tctccctcct tcctcttgga ttctgggaag aaaaagaatt gaatgggaac atgtaggggc	5340
tgggagagtg cctgcgctgg tggctggacc cttccgccct tgagtgctgt gaggggccga	5400
acggccgcca ctttctcctt cttaacagct caactacggg catttataga tgcgcccttc	5460
cctgtaggat ctccaggtgc gcgggtccag ccagaaaaga tcctcggaac gccgagcgcc	5520
tccgctgcac tcgcacagaa tttagacact cctctcccga ggtcttttca atgatctgtt	5580
tactgttctg cctcctatag tggcctgcga ggccccaggg cccgggccac gttttaccct	5640
ggggcgagcc tggcacctgg cgcacgcagt gttctacaag cgctgggtgc cccgcagtcc	5700
gcgaacacgc cacgctcgca gccgcagccc ggcggcctcc gctctgccgt ctgaagcctg	5760
accggacgct ccccttgtgc ctccacagcc tcggagaagt caatgtgccc atcgctggac	5820
gaagcccagc ctttccccct gcctttcaaa ccgtactcat ggagcggcct ggcgggttct	5880
gacctgcggc acctggtgca gagctaccga ccgtgtgggg ccctggagcg tggcgctggc	5940
ctgggcctct tctgcgaacc cgccccggag cctggccacc cggccgcgct gtacggccccg	6000
aagcgggctg ccggcggcgc gggggccggg gcgccaggga gctgcagcgc aggggccggt	6060
gccaccgctg gccctggcct agggctctac ggcgacttcg ggtctgcggc agccgggctg	6120
tatgagaggc ccacggcagc ggcgggcttg ctgtaccccg agcgtggcca cgggctgcac	6180
gcagacaagg gcgctggcgt caaggtggag tcggagctgc tgtgcacccg cctgctgctg	6240
ggcggcggct cctacaagtg catcaagtgc agcaaggtga ggctcccag ctcaccacct	6300
cgctgccgt gcgcccgtt cccctacccg cgctctgcct gcgccccgcg gccctctca	6360
gcggccttct ctctggcccc accgcctta ggtgttctcc acgccgcacg ggctcgaggt	6420
gcacgtgcgc aggtcccaca gcggtaccag accctttgcc tgcgagatgt gcggcaagac	6480
cttcgggcac gcggtgagcc tggagcagca caaagccgtg cactcgcagg taagcgcggg	6540
gcgcaccgcc gcgcgcggcc ctgctcgggg atcttctgca tctcctcggg gcagcaccag	6600
ccactctctg cctggaagtt ttctcctcga cttccccag tttcctcccc caagccctcc	6660
gctgcgtccc cttgccctgg tgcaggtgtg tagggaaagg aggattgtgg ccggctcagg	6720
ccttgaggca gccctggatt ttggtgtcac accactgtga gcctcgagag tgtgatctc	6780
attgttactt tgggcttgag gtaggtttgt atgcactgat tcgtgctgct gatatatcag	6840

acttactagc tctgtttctt tgtgcctatt cttttcacca aatggttgtc acttaatttg	6900
cattgacccc tctcgactga aaaggcagga atctcagctc atttagagca tctagtagca	6960
tattcacccc gctattcatt ctttccttcc ttcctttctt ttcttttctt ttttcttttc	7020
ttttctttca gagtctggct ctgttgccca ggctggaggg cagttgcaca atctcagctc	7080
actgcaacct cccctgtca ggttaaagtg attctcgtgc ctcagcctcc taagtagctg	7140
ggattacagg cgcatgccac tacagcacag ctaatTTTTg tgttcttagt agagacggga	7200
tttcgccacg ttagccaggc tggtcacgaa ctcttgccct caagtgatcc accagcctgg	7260
gcctcccaaa gtgctgggat tacaggcgtg agccaccatg ccccacgcc gctatttatt	7320
cattcattca ttaataaata ttgtttggct aacttccagg tgccaagtac ttaagaatct	7380
tataacacat caggtccttg acagcatgcc cacatgaaga ttatagttta gctgagagat	7440
ggagagtaga tgagcaagta aatatgccaa tagctatctc aggagaatgc ctacttacga	7500
aggctaaaaa gagtattagc ccatctcccc cagcaccac actggctggg gggaggtggc	7560
atctcaagtg actgaggtct aagcctcctg ttgaggaggg tggagaagtg tgtgctaattg	7620
ggtgtcaaaa aaagcagggt gtggatatgt atttgccatg ggggtgtggaa ggttgtgggt	7680
gaagaatgtt ttggtagaaa aagtgttgaa gggccaggca cgggtggctca cgcctgtaat	7740
ccaataactt tgagaggccg aggtgggcga atcacttgag gccaggtgtt tgagaccagc	7800
ctgaccaaca tggtgaaacc ccatctctac taaaaataca aaaattagtc aggtgtgatg	7860
gcgtgtgcct gtagtccctg ctacttggga ggctgagaca cgagaattgc ttgaacctgg	7920
gaggtggagg gtgcagtga ctgagatcgt tccactgcac tcaagcctgg gcgacagagg	7980
agactgtctc aaaaaaagaa agaaaaagtg ttcaagggat tttagggtca gctgaggggt	8040
gaggagagca gcagtctagt tgactgcagt aggagtcttg catctctctc tctctctctc	8100
tctctctctc tctctctctg caggaacgga gctttgactg taagatctgt gggaagagct	8160
tcaagaggtc atccacactg tccacacacc tgcttatcca ctcagacact cggccctacc	8220
cctgtcagta ctgtggcaag aggttccacc agaagtcaga catgaagaaa cacactttca	8280
tccacactgg tgagctaaaa aggcccttgg cttgtaggaa acaccctgag gccaacatta	8340
ctcatcttct ctgatttctg gccccagtga gtgggtggatg aggcctttct gatggagtta	8400
ttctctgctc tgtgttaaag aaaacaaagg ggtgggttct ttggttcatt taccggcata	8460

attctcccca gagccacctt gatttggggt tgtgtctgaa aggccactca gcaggtcagc	8520
tcacaggtac tctatacttg gaaagaacat tttccttttag gttagcagct gcttcccctg	8580
ctgcctgctc tgggtgaaat atgaagctcc agggctcctct tagagagttg ctctaaagct	8640
tacctagaga ttgaggactt tccctaacca cctggccttt tgtgggaggg actcgtgtgg	8700
actctccggc tgcattttca ggagtctgag agcttattct gattgaagag gaacaaataa	8760
tggcaaatat gattaaactc tctgctaagc attttttaaa tgcattattht cattttatgc	8820
tcacaacaac tctgagaggt agcgactact ctttctcccc atttttagaga tgaaaaaaat	8880
gaggctaggt aatctgcccc gggtcacact gctagcaaat gacagggcca gagctcaaat	8940
tcaggtctga cctctcaaat gttcactctt gaccactgtt tattgtattht tatgttcaga	9000
gtcatgaggt tggtagacag aaagcttctg ttcacttatt gcccttttca aaatatctgc	9060
aagttaatgc cataataatg ataattcctt acctattata atgctttata atttaciaag	9120
tactttcaaa tctatcattt catttgattc ttattgccac tcaagaaagt agaaggagct	9180
gctcttacca tcctgaaact cagaaagagt gaatgattta tcagaagtaa gactgaatga	9240
ataatgtagc catgtaatgc tctggctttt aatcctggac tgtttgtcta acactttatg	9300
tgcgggtggg agttttaatg ccaagaacac tctaatagtc aaaagacatt tacatgagac	9360
ccagaatttc tgaaaatttt attgcagaat atgaatactg atttagaaca aatcacagt	9420
tattctaaac acccaccctt gatgtttata aatatacttg ggtaatgtat atatttccat	9480
tgaaaacca gaaaagtatt ctactttaat cattccctct tacctgaaat ttccatgtaa	9540
ttcactcctt ataagtaagg tattcaggac acttatcaaa atgcaactag gatcttgact	9600
gaataaaaca ttaagccctt atcaaacatt tacgttatac ctagaatttg ttttctcaga	9660
tttgtttgac cctaaaggga tagaatacat tttgatgggt ggtttcttat caaggaaatc	9720
tgaagcatga aaacagaaaa gagtttttag caaggaggac agagggttcc tcaaaaacaa	9780
acttcatcta ttttatactt tttccaaggc tgagccctga ctataatgcc atgctgggct	9840
attggaaatt catgccattt acccaacaac acatgagatg gggaacaaga caaaccttc	9900
ttgtgttctc ttattttatta atttgtggtg aagaattgct ggtatataaa gaatcatgtg	9960
attaaccca taaaattaag gaaaaatcaa gacagtaaag tatcagctgc cttaatcctt	10020
tgtggcccaa atgtggattt ttaaaataag atattgaaaa acgtatcctg cacatgtacc	10080

ccggaactta gaaagaaaga gagagagaga gagaaataaa gaaagaaaga aagtccatgt	10140
taagatgttt tttcagatat aatctgctgt ctttcaagaa caagaaagaa gacgggctca	10200
ctgatccata caaactaaca cccacttgga aattcagatt tgaaaacttc ctctgaatta	10260
gaacggagtc acacggtttt aggacagctt cccctcccc ttcctgttga acatctgctc	10320
tgagtgttca tggcttataa agtcaggga gtcctcccg gtagattca gctggggagg	10380
gcacgtggcc tttgctctgt ttccgttttag caggaaaccg tttgaggcct ttggctggga	10440
acccccctt agaaagtctc cttttcacct ggtgccccca tgggtgcttcc agggactcgc	10500
attgcaggct gggagtcagt tcaggttgca acacgtcacc ctccaagttg cttgaaggcc	10560
ttagactgtg gtgcaaccag ctgctgcaa gagcatgtgg gtcacagtgg gtcccctcta	10620
gctttatcat agactcatac tttctcccct cccctccca tccccacagg tgagaagcct	10680
cacaagtgcc aggtgtgcgg caaggcattc agccagagct ccaacctcat caccacagc	10740
cgaaacaca caggcttcaa gcccttcggc tgcgacctt gtgggaaggg tttccagagg	10800
aaggtggacc tccgaaggca cggggagacg cagcatgggc tcaaatgagc accctggctg	10860
gtgcaagca gcagctacac aacactacag agggcagcct ccctgcttgc caccactctg	10920
ctccctgctt gcctccactc ctttctgact ttccagacc caggtccagt ctgcagatcc	10980
taccaggttg ctctccttc gccttacctc ctggagctgc cagaagaaat gaggtacctt	11040
ttcaaagtgc agccgagagt gagaaccaag tgactctcta ggcttcggac acaaataggc	11100
tcctctacac ctgaagacaa aggcaaagtc aaatggggac cagaataaat ctagacccc	11160
acagtccttc ccatttccag ccctaatac cagacaggaa tgcccttcag gtttcttccc	11220
tccccctct tgacctacc cagatatattg tgtggaagag gaggaatcac catttacaag	11280
gtggacaaat gctaataattt ttatctagaa agaagagtga gtgttaactt ttatTTTTTT	11340
ccttctgggg ggtctgttga ctcttttctt ttgggtgctg cctataaatc ttggaggaat	11400
catttctcct cctcaaaaac tgattcagaa actgacttgg ggaaggaatt taatactttg	11460
aagtcatgag atgcaccatc gaggctaccc ccaagaagaa gcagaagaga agttggtaat	11520
gagaggggat tagaggctct cccttcagta gggctgtgaa aacctcatca ctggaggtaa	11580
aagcacaagc aatgcctgtg gacaagatgt cattcattca ctacagaaat gttcatggat	11640
caccggctac caaggtacca ggcaccatgc taggtattgg ggaagagaga ctgaagtcac	11700

aacccctgac tgctcctcaa aagctaacgg ttgcacctcc aagtggtctgg gtctgttctt	11760
actcttggag ggaattctga gaagacagca cagaattgta aaccttcctt tttgaccctt	11820
ttggatttta tcaggtgtaa acaaaaagct gaacagttac ttcaaagata tgtgtgtata	11880
ttcagttttt tattgttaag ctgatatttt aaagatttct gagctagcag gcatgtggga	11940
aggaaggctc tgtcttcaac tctttgaccc tccatgtgta ccatagaggg gggaaagggtg	12000
gtattttcac tttgatgagg ttggtaaattg tttttagatc ttctggtaag cattatgttt	12060
gttaatacat atttattaga gtgatgtttt aagttaataa agtattaaga gtatta	12116

<210> 26
 <211> 270
 <212> PRT
 <213> Homo sapiens

<400> 26

Met	Val	Thr	Phe	Ala	Thr	Ile	Leu	Gln	Val	Asn	Lys	Val	Met	Ser	Ile
1				5					10					15	

Leu	Phe	Tyr	Val	Ile	Phe	Leu	Ala	Tyr	Leu	Arg	Gly	Ile	Gln	Gly	Asn
			20					25					30		

Asn	Met	Asp	Gln	Arg	Ser	Leu	Pro	Glu	Asp	Ser	Leu	Asn	Ser	Leu	Ile
		35					40					45			

Ile	Lys	Leu	Ile	Gln	Ala	Asp	Ile	Leu	Lys	Asn	Lys	Leu	Ser	Lys	Gln
	50					55					60				

Met	Val	Asp	Val	Lys	Glu	Asn	Tyr	Gln	Ser	Thr	Leu	Pro	Lys	Ala	Glu
65					70					75					80

Ala	Pro	Arg	Glu	Pro	Glu	Arg	Gly	Gly	Pro	Ala	Lys	Ser	Ala	Phe	Gln
				85					90					95	

Pro	Val	Ile	Ala	Met	Asp	Thr	Glu	Leu	Leu	Arg	Gln	Gln	Arg	Arg	Tyr
			100					105					110		

Asn	Ser	Pro	Arg	Val	Leu	Leu	Ser	Asp	Ser	Thr	Pro	Leu	Glu	Pro	Pro
		115					120					125			

Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Ser Pro Val Val Ala Asn
130 135 140

Arg Thr Ser Arg Arg Lys Arg Tyr Ala Glu His Lys Ser His Arg Gly
145 150 155 160

Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val Thr Asp Lys Ser
165 170 175

Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val Leu Gly Glu Ile
180 185 190

Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr Glu Thr Arg Cys
195 200 205

Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly Ile Asp Asp Lys
210 215 220

His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg Ala Leu
225 230 235 240

Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp Ile Arg Ile Asp
245 250 255

Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly Arg Thr
260 265 270

<210> 27
<211> 258
<212> PRT
<213> Mus musculus

<400> 27

Met Ser Ile Leu Phe Tyr Val Ile Phe Leu Ala Tyr Leu Arg Gly Ile
1 5 10 15

Gln Gly Asn Ser Met Asp Gln Arg Ser Leu Pro Glu Asp Ser Leu Asn
20 25 30

Ser Leu Ile Ile Lys Leu Ile Gln Ala Asp Ile Leu Lys Asn Lys Leu
35 40 45

Ser Lys Gln Met Val Asp Val Lys Glu Asn Tyr Gln Ser Thr Leu Pro
50 55 60

Lys Ala Glu Ala Pro Arg Glu Pro Glu Gln Gly Glu Ala Thr Arg Ser
65 70 75 80

Glu Phe Gln Pro Met Ile Ala Thr Asp Thr Glu Leu Leu Arg Gln Gln
85 90 95

Arg Arg Tyr Asn Ser Pro Arg Val Leu Leu Ser Asp Ser Thr Pro Leu
100 105 110

Glu Pro Pro Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Asn Pro Val
115 120 125

Val Ala Asn Arg Thr Ser Pro Arg Arg Lys Arg Tyr Ala Glu His Lys
130 135 140

Ser His Arg Gly Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val
145 150 155 160

Thr Asp Lys Ser Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val
165 170 175

Leu Gly Glu Ile Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr
180 185 190

Glu Thr Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly
195 200 205

Ile Asp Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr
210 215 220

Val Arg Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp
225 230 235 240

Ile Arg Ile Asp Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly
245 250 255

Arg Thr

<210> 28
<211> 258
<212> PRT
<213> Rattus norvegicus

<400> 28

Met Ser Ile Leu Phe Tyr Val Ile Phe Leu Ala Tyr Leu Arg Gly Ile
1 5 10 15

Gln Gly Asn Asn Met Asp Gln Arg Ser Leu Pro Glu Asp Ser Leu Asn
20 25 30

Ser Leu Ile Ile Lys Leu Ile Gln Ala Asp Ile Leu Lys Asn Lys Leu
35 40 45

Ser Lys Gln Met Val Asp Val Lys Glu Asn Tyr Gln Ser Thr Leu Pro
50 55 60

Lys Ala Glu Ala Pro Arg Glu Pro Glu Gln Gly Glu Ala Thr Arg Ser
65 70 75 80

Glu Phe Gln Pro Met Ile Ala Thr Asp Thr Glu Leu Leu Arg Gln Gln
85 90 95

Arg Arg Tyr Asn Ser Pro Arg Val Leu Leu Ser Asp Ser Thr Pro Leu
100 105 110

Glu Pro Pro Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Asn Pro Val
115 120 125

Val Thr Asn Arg Thr Ser Pro Arg Arg Lys Arg Tyr Ala Glu His Lys
130 135 140

Ser His Arg Gly Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val
145 150 155 160

Thr Asp Lys Ser Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val
165 170 175

Leu Gly Glu Ile Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr
180 185 190

Glu Thr Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly
195 200 205

Ile Asp Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr
210 215 220

Val Arg Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp
225 230 235 240

Ile Arg Ile Asp Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly
245 250 255

Arg Thr

<210> 29
<211> 813
<212> DNA
<213> Homo sapiens

<400> 29
atggttactt ttgccacgat cttacaggtg aacaaggtga tgtccatctt gttttatgtg 60
atattttctcg cttatctccg tggcatccaa ggtaacaaca tggatcaaag gagtttgcca 120
gaagactcgc tcaattccct cattattaag ctgatccagg cagatatattt gaaaaacaag 180
ctctccaagc agatggtgga cgtaaaggaa aattaccaga gcaccctgcc caaagctgag 240
gctccccgag agccggagcg gggagggccc gccaaagtcag cattccagcc ggtgattgca 300
atggacaccg aactgctgcg acaacagaga cgctacaact caccgcgggt cctgctgagc 360
gacagcacc ctttgagacc cccgcccttg tatctcatgg aggattacgt gggcagcccc 420
gtggtggcga acagaacatc acggcggaaa cggtagcgcg agcataagag tcaccgaggg 480
gagtactcgg tatgtgacag tgagagtctg tgggtgaccg acaagtcac ggccatcgac 540
attcggggac accaggtcac ggtgctgggg gagatcaaaa cgggcaactc tcccgtaaaa 600
caatattttt atgaaacgcg atgtaaggaa gccaggccgg tcaaaaacgg ttgcaggggt 660
attgatgata aacactggaa ctctcagtcg aaaacatccc aaacctacgt ccgagcactg 720

acttcagaga acaataaaact cgtgggctgg cggtaggatac ggatagacac gtcctgtgtg	780
tgtgccttgt cgagaaaaat cggaagaaca tga	813

<210> 30
 <211> 63186
 <212> DNA
 <213> Homo sapiens

<400> 30	
agttgaagct cctctccctt ccgaacagct ccgcgcaccg ccccgcgacg cagcccggcg	60
caactacttt cttctctctc ctttctttct tcctctcctt tttcccctgc tgggtagtgg	120
ctgcggcggg gtgggggaga ctttgaatga ccgagctcgc gtccaccttt ctcttcatgt	180
cgacgtccct ggaaacggcc acacggatgc catggttact tttgccacgg taaggggagg	240
cggcgggcac cttgggtggg caggtttggg gatgggggtc cacgtgggga gggattttcc	300
agtggactgg tgcggggggc cccagatccg catcccgccc ccccccatc gcgccgcgct	360
cactcacttt cccgggcttg tgtcttcccc aaagtttgcg ctgggatctg ctcaggccga	420
agcgcaaccg cagccacccc gctacacaca cacacacaca cacacacaca cacacacaca	480
cacacacaca cacagacacg gacacccttc tcacctcct cccctcttgt ccctcggctg	540
cccaagaagc ttccctcaat ctgggaaagt gatcaggttt aagggaacctg gattggaaag	600
ggtggggggca gaagagggga aatggggacg acgaaagagc aggaaagaga ttcaacagaa	660
tcaaccaccc accactccca accgaccgc ctgctcctcc gagaaagctc ctagcgcac	720
ctataacaaa aggggggtggc agacagaact ccgggcgggg aggtgccgcg gcagctcccc	780
tgcacacgcc ctgcactctg ccggccgctg agcctgattc tcagctcgcc ccagcaccac	840
tctggcccgg gcgtgggctg gggggagggg acgcgcagct caggacccgg aacctcgct	900
tccagttttg ggagttggga ctactgcca cgcgccgct acctgcgttg gagttccccg	960
aaagggtttt ttcagaaaag acctcgcgcc ccgggctcct cttggccagc gcccaccgg	1020
tggccacccc accctgggcc tttgcgcaga tggttgagct ccgtacgcag cccgcacatc	1080
tgggaccct ccggggagcg gcgggcaccc gggcccggcc atcccagggg atctccttgc	1140
ggtatcgtcc agcctgttct cggactttga gcggtggcgt gggaggccgg gagacctggg	1200
caccgcgca gccagccagg tcggagtta aagggtccac gacggaccga actgtcccat	1260

tgccccagag	ctttactcag	tggtggatgc	tcctgatgaa	atttgggacg	cttgggagtt	1320
gaaggtagg	gacaggaggg	gcgagggccg	agggcattggg	atgggggagt	aggattctgc	1380
ttgttgctct	ccgcgggagt	gggtgcgcgt	ccaggaggcg	ctgcttcttt	gcgggagttt	1440
ggctgctgcg	ttcattcgtc	gtctgcgctt	cagatgcacg	gcactgagac	ccttgctgcc	1500
gacggtgtcg	gggctgtgga	ctagaaagga	tcccttttgc	tggaatcgag	gctgggggtgg	1560
gattgccggt	gggggaaaca	ccgaaaagat	cgtctggcct	cggcctctgg	cggcgggcgg	1620
caggttctga	gtccgaatgg	aggttgctcc	cgggagcgcc	gggctcagag	ctagagagct	1680
cgggagactg	tgcgcctgtg	gacttgttta	tgtgtgtgaa	gaggcggggg	cgagggcctg	1740
ctgagagggg	aggggagcct	ggaaggggtg	ggtgtttctc	ctggagcctg	atgtttgtaa	1800
ctcagctgat	tatggagtgc	actgagcgac	ctgcttttta	aataaagagg	tgcccgtccc	1860
taccccgcaa	aacagcgaac	gaggagaaca	tggaagcgct	ctgtcctaaa	cgtcaggatg	1920
ggagaaaagt	gtaacataga	ggagactttt	ccagaggctc	tgttttcaca	acactcagaa	1980
gttctccagc	gtactcagcc	tgcctccgc	ctgccctcaa	ttcctttttg	acatgtcaca	2040
caaagaaagc	tgaaaggtag	aatgtggagg	ataaactcca	acccctctg	ccttgggcgc	2100
aaacacacag	acttaggtat	cgtgttagga	ggtaaggggg	ttggaaaata	atgcaggctc	2160
cggtagacag	tgttgaaggg	agatagaaag	tctggggtat	ttcccctagg	gagagtgtgt	2220
gggattttgt	gttggtgagg	actggagtgt	agctggactt	agagggtttg	gttgtgtgtg	2280
aatgggatat	gcttatctat	ggagtgagat	tgtgaccatt	gagtgagtgg	tcagggaggg	2340
atggggatgt	ttttccaaag	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	tgtaggggaa	2400
ggtggcaaaa	gagctggagt	cctatccctc	tttgggtgcct	gggagggtgtg	tgtttctggg	2460
gagtctgtcg	gatgggcata	tgtgtgctta	gcgggagaca	agagcttgca	ggaacatagg	2520
tgtagctaata	gcgaaatctg	tgcctatgaa	gggtgtcaga	gcaggaggag	agggtgcagcc	2580
tgttgaagac	agatggcttg	tgcatacctg	cacaaacaca	tggagtgggt	gtggtgtagg	2640
agggcggctg	tgtgtgtgag	ggaaagggtg	atgcatgtgg	aaggttgggc	gcctgaggat	2700
taccaaaggg	atgtgtgtgg	ggggtggggg	aagattatct	gaggagcacg	gggtttgaac	2760
ttcatgaggg	tgactggga	aatttggtg	gaatggcact	ggggcatgta	gggggtagaa	2820
aagtgtgagc	tcagatcctc	tttatggaaa	gacaaagcat	ttgcagggaa	ggacaccggg	2880

ttggtgttgg	tgtttgcctc	ccccctcctt	cctccccaga	gagaagagaa	ttggaggagg	2940
cttagaatga	gggatgtaca	tttaagggga	aaagtccttg	tgcataggca	ctgggtggga	3000
acaggaagag	atgtgtgcat	aagactgaag	taagtagagg	aatgcactat	ctgtgcctgg	3060
tgtcagggat	gcatgttgta	taatgttttg	taatgggttt	cttatgtaat	aggagaattc	3120
actatcacca	gttcatattg	ataatcaatt	aaagctaattg	gagtgacctac	tggatgctca	3180
gcccattgct	gggtactggg	gcatggcaca	ctgccctgga	aaaattcctt	ctccaggctc	3240
agtaaaactg	atgcaatfff	gccagaaata	gacagagctc	gacgcttggt	gaggtttctt	3300
caaggctctat	aattcagttt	ctagagatgt	ttatggatct	tgggggcca	agggaagctg	3360
acttgaaata	tatctctggg	tttgtaaaaa	gtggggcata	aagagaggga	tgatgtagac	3420
ttttgtgtat	tgggaaggga	gctggacggg	gttggtgttt	cttggtggg	gtgtgggctg	3480
ggaattggag	ggtgctgcac	ttcaggtcag	aatgcaggctc	cagatccctc	taaggattga	3540
gtttctaaga	gtacacagag	cctccttcat	agctggaatg	agacctgtta	tgtggagctg	3600
gagaatggtg	aagtcattgat	caaggaagag	atggactggg	agtgtttgtt	catctttgat	3660
gttttctggg	ctacgagtgt	aaagctaacc	tgagtgtgtg	ttgggggggt	gggtatgcat	3720
gcagggtcgt	gcacatgcac	tttggcctgg	agtaacgtaa	ctgggagaaa	actagcgagg	3780
tggataggag	gggtggatat	tctggctctt	gtgaaggggg	tgccagtatg	tgtttttgca	3840
ggaattctgt	gtgtctgggg	aaggataaaa	attgtcttgg	atactcagat	gtgcatcttt	3900
agcgaaaaaa	ggtgtacaca	aaagaagggc	cgtagacttt	ggggaatgag	acaaagggtg	3960
cgttgatcgg	gtccataccc	tgagatggct	ggtgaggcaa	ctgggtagga	acagagtga	4020
ggctgtgctt	ttgtgatgca	cacatctgcc	tgtctggagc	agacgatggg	gatccccatt	4080
tgattagggt	ctgaccaccg	ttgggcatat	gtttgggagg	actttatttg	actcatgaca	4140
tttttgaatt	catgaagtta	cagctgtgcc	attttaaggg	tctaccttat	ttttgagagg	4200
gaaactcatc	tgtgtaagag	atagttttata	tagggcttga	ctattgggtca	gaagatccgg	4260
ggcagatatt	tctgagagca	gggtgggtgaa	gtgcaagact	ggattgggaa	tgtcttagtt	4320
ctagtccagt	tctaccactt	gggaaaagtg	tgactttggg	gacattttta	atctctataa	4380
ctcttctgtt	tctctttttg	ttaaaggggca	atataatgac	tcaaggagaa	ggaaagttca	4440
ctcctctttt	ccccatcaa	aacttgatgt	gttttagttgt	atacagccat	acaaagatga	4500

aatgtaatta aatacaggtg ctagcacgta tgtaggtcat tgtgaaaata ggatctggac	4560
ccctttctct ttacaatgca acttaatcag taatcatgat gcccacagca tctctactgc	4620
tcactcattg tgtaaataaa actttttata actgcctcag tttactcatc cagaagctgg	4680
aggcagccaa acctagatta tagatgatgg aaatctttaa gatctttgag ctattcagag	4740
gaaaggtagt ctaagaatca atgttattaa tgatcagaga cagacatttc tttatTTTTT	4800
aacatcagct atgcagacca cagcaccaaa agaaatgtga cgtggagaac agcaaaaata	4860
ataattgccc ttgacatagg caccatttga ataaaagaag agcaaagtct agcctgcctg	4920
gtaagaggaa agcctgggct ttgggtcttt gaaattcttt gtgcccagtg cagcaagttg	4980
ggtacagcca ctttccccag agatgaagat ggagtcgtag aggggcatgg aaacacactg	5040
aaacctccga ttcctcccag aagtcagtac atgctgtcaa ctctgggctc agagtcaggc	5100
tcccatcctt gagttcagtg gccataaata gtcaggagct aagctatttc caagaggcac	5160
taaactaatt ctgttgcttt tcttggcgag attgggcca ttaagttgat tggaagaggt	5220
cactgctgat gggcttcaca ttccccaggc gggacaggcc agcagagttt aattacataa	5280
catttcccat gtcttcactt ccactcccag atgtatcaga tgttcaggcc tctcctcttt	5340
ccctacattc aggatgcctg acggagcagg cggcttctgc tacaagcttc tccagacctt	5400
ttcctgagca gttacagcac tcccaccca actggaggga atgcaagtct gtgcattgcc	5460
tggaagtgag agaggggatg caggctgaca cagaagatct ttagagccca ccaccctgcc	5520
tcaagcctca cccgcggtg ttttactctt taacgaggca ggcagtttct ggcgttgggg	5580
atgttgaact aatgaccgtt gctagggttt tagatttagt tttatttaaa aatgagtttg	5640
ttggaaaagg accaaatctt ctactggaga caagcctaag ccccggtgtg tgctgctgat	5700
aaggtctgtg gatgtacttt gtgcatggac aggcatgtgt ggagtaagag gagagcacag	5760
gatgtgcgtg cgtgcgtgct gggggacacg tgagcgtagg agatgccctg ggtagatggg	5820
gcatttgtag agcttatgtt atgggtggca ggtatctgga tgcttggctc agattggatt	5880
gtcattggat tgcttcttac aggtgggaca gcctgagaaa aaagttgggg acaccttcct	5940
tcagatcatt tttatatggg catgttgaac cttgaagtat gaggccagtg tagatattac	6000
ctcattcttg cctggatgga tttctaaca gattgaggct ggacaaaaaa aaaaaaagtc	6060
tttcagagga ctcatcttag aacctgtgag aacacatgta gatggcccca gagcgtaaaa	6120

gtatTTTTtg ttgttgTtaa tagTTTTcca taactcagtc ttttgTTTTa cctaaaaata	6180
ctgagaggct agaggTTgca cacacCCCCa catggTcaga atgcattTgc tgttcatgtc	6240
cccatctgtg ccatggTcct tctcaaatag gattaattgg gaactacgga gagcttgccg	6300
cagctTTTggT tgtcctcaca ggacaaacCT atgactgacc gtgatgtcat tgaaacagaa	6360
cagatTTTga cgtgaaatTT cgtgtggTct gttgtgattg ctcaCTTTTt agtaacatct	6420
ccagagaaga catttgtTga gCCCCatTTT tTTTtgTggg tttgtTcagt cagagattTg	6480
ttggaagatt taaggTtctc caacattcaa ggccagagat tttggaggaa acagtagcta	6540
acaagagccc aggccagggt gacCTTcagt agcctgaggc agcaggtagc tgcctgactt	6600
agaagtgaag gaccaggTtc agtgactggg ctctgtcact aattagctgc tactgtgtga	6660
ccctggacat tctctgtagg acatgactTT gccatctgta aatgaagggc ttgaattaga	6720
agatctCCA gtccatttcc agccctgtca ctgtaggatt ccacagtagc cacatcctcc	6780
tacctaatca gtgctaatta cctcactTgg agtgcaccct ggcacgccct taacgagctc	6840
ctgtTTTTat cTTTggggaa tggctgagtc aaaggcagaa aagcgtgcaa aaaattTgcc	6900
agagtgtTtc cTTTtgGcca agaatgctct aaacattggc ttagcctaca cagggtgtgc	6960
agagctgaaa ggggtgggat ttggctTTtct ggtaggtctg atgctaataca ttcagcttcc	7020
acgtgaatgt cTTTcagtgg caatgtctct gactaaggta agcatccatc tTTaagatga	7080
tggatgatga tgatgatgac agTTaatatt agagtactta ctgggtgcca ggcattTTta	7140
cagctattac atggcattat ctcagtactc aaaataatct catttacaga tgggggtact	7200
gaggTtatgt gaggTTaagg ggtatgctga tgttctTggc tttccagcct gaagtctgtg	7260
gcatCCcaag cccatcacct gctagTcaga ttctgtTTTT tatgtctgtg atgtcatcca	7320
gatgggagct gggTccatct tcttctgcaa agttggTcac cccatagggt aacctgctcc	7380
ctgCCCCcg aagcctcctt gtcatgtgca agacaaatgt ttaggtTTTt tagccacctg	7440
ctggcaagca tgtgtcagat ctgtctcagc ctccattggc cctctccctt tacttctTgt	7500
tctgttgaca agggaaacTg ccTTtacaga atacggatat tgattTaaGc cacaattgcc	7560
tccttgtaaa ctctataaaa ttggagctgg acatctggTt ttgatccaca cggccacgag	7620
aggtggtagg tcccagcatt cttgagTcct ggaattgcac agaatcaaag tgaaaaagaa	7680
tatgaatgga gagcatgagg tatctTTtaga atttggaaga tggctTTtctt tgataggcta	7740

gggaactcac accatttaga ctggggccaa agttaaggta ggtttcttgc tgggaggttg	7800
gggggcaggg gtcactgcta aagactgagg tccccaggct gtgtttgatg cagagcagag	7860
cagctgtaga gccctccagc tgaggatttg cactgcaca ccattctctg ggcagtagca	7920
gaaatagaga agccaaacag gttcttgggg tggatgttgc cctgtttgaa gattgtgact	7980
aaaggctcat tccttcttcc tccactgatt tcctcaccat cactaccaca tatacatcat	8040
caagacagtt aaggacaggt gactttgttc tgttgcttaa acatatttgc atgtacaaat	8100
aatgttgaa cactaaccta atgcctatta gaccataggt actcaaaaat tatgtattta	8160
ctttgcttaa aataaccttt catgtgtagg gtgataattt ttttcttcat atggagtata	8220
tagcctttta aaaactgata gccaggttct tgcaatttac ttcccccttc cctctgtttc	8280
tagatctaag accccagcat ttaaaaatat ctttctctgg ctctaagtga tccctcaggg	8340
agtttgtttc aagttctcaa tctgagatta gtttcatccc acatctagag gacctgcttc	8400
taatggttca taaggtactt ccactgggtt caagaggaag acacagctct agaaatcacc	8460
tgctttcca acaggtgtgt cccatgggtg ggaagaagcc cagtcctcct aacaccacc	8520
tggctacctt aaatagcttg ccacagtctc cctgcagaga gcccagacta gctgtgtgac	8580
tttaagcaaa tcactttccc tccgtgggcc tcttttcttg gaaaaggag gagttggact	8640
cactagaggg tcacaaggac cctaccagtg tgacagatca catcttccat gactccgctg	8700
ctccctcggc cccacactgt ggcccatgcc tgggattttc aggtaacctt gctggcctag	8760
aggacttaca tagagtgatt tggagggggg ggaggaggaa attaagttat gtgatgccag	8820
ataagtcatt gaacatacgg atgagtgcac gtttatgaag cactgaagaa acaagcttgc	8880
cagattcgtg gcggatcttg gattggaagt tgggagtatg cgagagggtt atgtgggatt	8940
aaagttgtct gtgggtagct gaattcaata agctaaatca gagagatccc aaactctgca	9000
gagagatcag accccgcctc taactccaac agaggctaata tacaattttg ggaagtcctc	9060
gaaccacact aggtctcact tcctcatcta ttaagagcta ctgactttga ttatctctaa	9120
ggtctttctt gacttatcat taattaaatt gtgagaattg ggagctagta gaaaaatctg	9180
gacaaatgcg tggatggaga aggggatacc tcaaatggcc taaatataga gggaaatgtt	9240
ctgtacaagt gatacgttta ttcttgaatc ggagatagca ttggaggata tattctgtac	9300
aaatattgtg cttcaaatgc ttaccctctc atcatcatca atgatattaa aatcattgct	9360

taacctggga tatttgctta acctgggaag cctccctgca tttccccaga gaggcctaagg	9420
ctcccttccc agcctcctga tgccacttag cagcctgtgt ttaatttgca ttcctagaat	9480
gtgacttcac agagaaaggg atctgttggt ttcatTTTTg tgtccccac cttgaatagc	9540
gacagccaca ggaatgtgtt gagtgaatga atgcgtgacc agccattgtg agctcagaat	9600
gtgcagcagc gatttgacag tgccaggaac ttggctcaga gaggcagtgc accttgccag	9660
ggcttcaaaa cagtgttacc tttatcccc ttttggcctg tccaaggct aaggctcatga	9720
atgctcagcg cttgggtggc agaaaaggct gttcttactc ctttttagag acctttcttc	9780
cccacatctt agcctgctta gttctcctgc aaattggggc tcttggtcac agggaaatcgg	9840
gtgtcattgc atacgcatta ggataactcc tcgttagaca ctcggaaagc aactcaaaac	9900
acacgtgtac agattattca ttaagcagtc tgttggtggt cacgtgctaa atgttgggga	9960
agatgcagag atgaatgtag tgctgctggc cttgggggtct ggaaactggt gaaggaggca	10020
gtcaggttac aatcatagtg ctccaggagg aacaaggga gcaaggaga tggataacag	10080
ggttcccact tcagggtgtg ttttttagta caaacatga ttgagaatcc tcttggttca	10140
ggagcaagct ggagggtcca ggaatggtct ccgcttcac atgacagtgg ttttcttgtc	10200
cctgaattcc cagggttagg gtgatggatg ctgagagacc attgtgatgg gagcccttgg	10260
tccttgtggg gagagtcggg ggaacacgca tgttgtacag gtgttggggc ttggggagggt	10320
gtaggtgtgc atgctcacca gggtacttct cagaagtatc cctgctcgtg gcatgggtggg	10380
ggtgagtcca cctcaggccc ctaagctgtg cttctacttc ctttctaatt atacagcatg	10440
tattgtctct gtgaccactt ctgacacagg aacagactgc catctactgt ccacttctgt	10500
cctgagtccc tcctgctcaa ggaggcattt attttttctc atcctgtttt taatagatcc	10560
ggacaccagg agaaacagac gtggagtcct atttagtaat ttttttagaa tttagaaggg	10620
agccttcaaa aaacatagca catctgtcag atggtttccg tatctgttta caaagtatta	10680
ttgtaagggg ccactctgct gccactgagt cttgctagcc gctgtgcgag gtgccgtgtg	10740
ggataccaag agatacaaag caaggcctta cttcaacag cttgcagtc tacttgcgaa	10800
tacagtgtag atttgcagag aagcagacta caggatgtgg cagtagatat ttggtaagag	10860
gcatctgggt ggtagggact acagaggttt tgggagtttg gaggaaggag gagcccttat	10920
aggctggggg atttgggaaa gaccaatatg ggtctccttt tagcaatccc agtgtttagg	10980

aagaggcttc taagagctgc gcgatgttgt gggacaggct atgaactgga aactcaggaa	11040
cctggagggc ctgaggtcta gtccagcttc tggcagtcct gccctgtgac ccagggcctt	11100
gggttccgca tcctcaaagc aaggctgttg gtcttcttgt catgtaaggg ccccccatc	11160
cccgaccctg catcatgtac cactcttgga tgcagtggga ggtggtttcc tggtttccag	11220
cttgtcaaga gcaatagagt caattggatc ccatgcagga aggatacctg gatgcagggc	11280
cctgtgctga tgaccccaca tgcagcagga gagagagaat ttccccaggg gagttttggg	11340
tgctgttaaa atacaaaggg gggaaatgtt ccatgcaacc tctgcatgcc actcctctcc	11400
agagctttcc tggactttcc ttccaaaaat atatgtgcat ggtgctttgc gtccaagata	11460
gctaccccaa aagaaattgt attagtattc tagcattgtt catagctcaa atctaagatt	11520
ccttcacatt agaattcatt ccatgattgt ttttatggag gttttctttg agtatttgag	11580
acttttggtg gctgtggatt aaaatgagga aaagattgtt ttctatagct cttttactcc	11640
ttgtctcttt ttctctgtct ttcttagcat ttaaaccacc accaaggaaa cctccagggt	11700
gtttattgca tattcttgcc ttttgatgtg tgtgtgtgca tgagagagag agaaaaaaaa	11760
agaaaggagg agaggagagg cgagagagag tacaagagag aaagatattt tagactgtgg	11820
tctattagca tgtctaggaa atcaagtga atttggattg ggttacagggt tgagagacct	11880
cagctcgtca ctgtaatcct ctgacatttc ctttttactt tttggctctt gtctgaatca	11940
agaatacatc ttcccgttcc ctcttccatg ttacatctt cttttggggc agctgtataa	12000
agttactgtc tgtctacagc aagtctcctt aattcttttg gggactgctc gacagaggca	12060
cagtcaagga tagaaccatt agagacggtc agcttttgga caaagacaaa aaaatatgtt	12120
gaggaatagg agaaatgttc caattgtcga ttatgtgttt tcctgtggta tgctgatgtc	12180
tgtgactgtt gtccacagaa agagagcgtg agcaggggct ggggcagaga taaaaggagg	12240
attctcccca aaccaacaaa tgtggagaga gagcaggagt cccagcaac agatatggag	12300
ccaaacaacc ctgggaagta actaggaaac ttcttggtga cctgagatgt gttttgtaat	12360
cactggtgaa agtaggagga gatggggtag aggagagtta caggtagaga agttccgtgg	12420
gaggagacc caaggtgaga gaaagcaggg acttaaagaa ctgaaagaag cccagggagg	12480
ggatgaacag agaacatccc atatatcca tatccagtat ggtggagagg tgggcgggag	12540
ccaactggga gcttactttt tttttcccag gggcatggga actatggaaa actctgaggg	12600

cacaattaag	cttttgtttt	taaaaagatc	cctccggctg	cagtgtggct	ctgaataaag	12660
agttgttcaa	tgaagaaagg	aatggctgat	ggttggattg	atggaagaat	taatgaatga	12720
gcaaacgaat	gcatgagatg	cagaggggaac	gcagcaagcc	tgatttgatc	tctggttact	12780
ttgaccagta	ctggctccag	gggcgattat	ctcagcctgg	gaggccaggg	agtaatgcat	12840
tgattagaat	gtctctggac	acatggatta	aaatatctga	tatttttaggg	tggttgatag	12900
tggggaggac	ttctgaaact	ttttccctct	tctatgcatt	tccatcctaa	tttgatttca	12960
ttcaggatca	aaaaagaaga	tggcttttga	aattacatcc	cagaaaatct	gaaactgtgg	13020
cattgacttg	ctccagagag	ggctgcttgc	atggaagacc	ttttcatagg	ctcatcgtgg	13080
aatagggaca	gatgataaag	tttcttgggc	atatgaaggg	gtcccagatt	tctggacgtc	13140
agatccaccc	atagatgatt	ccttggatta	aatgatgtgt	gtgtgtatgt	gtgtgtgtgt	13200
gtgtgtgtgt	gtgtgtgtgt	gtgtttaaag	ttttaaaagtc	tcctgaaaat	taccaagtgc	13260
tactgaacat	tttttttgtc	agtagtggat	atctggataa	tttctttcaa	ggccacattg	13320
cttagcatgt	ataaggaaag	tgtgtgcggc	agagaccag	atggacagcc	ggcatgccat	13380
ccagttgctt	ggggagtgtg	aatactcctt	ggcaaagcca	aagggagcca	aagaggacct	13440
ctagtgtctt	tgctctccca	tgtcctgact	attccaatct	cactttgcat	tttgagatct	13500
tttaactttt	gctaacagtg	attgctcagg	atgtcattca	ggccaattaa	atttgaatgt	13560
ctagggttgg	gacagaagct	cacgaggtga	atgcaatgtg	cagccacgtt	gagaaccact	13620
agctgacact	gctaactctac	tctgctctcc	tcctgcctt	ttggcggctt	tgccatgata	13680
tctgtcccat	cctcacctca	gtttgctgag	aactcctcac	ttagttaagg	aagttcttct	13740
aaggatgttc	aactaatatg	ggctaaggcc	tcctatcccc	taaaaatcca	gcatttgcct	13800
gagaaattgg	acgctaggag	gataagacag	gttcttagca	gattctgaag	cactcaccgt	13860
ctctcatctg	ctgaagggtg	tattgaggat	gaatgatggg	gaacagcagg	gagaaagacc	13920
agcatccata	gggcagtaac	agggtgcaga	cacctattta	tatgtgtcat	cgtagtcccc	13980
ctgacagctc	ccagaaccag	atattatacc	aggcagatat	accaggctta	caggcaccta	14040
tttatatgtg	tcatctcctt	agttccccca	acaactcaca	gaggcagata	ttataaccagg	14100
cttgagggca	caacaactca	cagaggcaga	tattatacca	gacttgaggg	cacagtttgc	14160
ataatgcatt	cataatgcaa	agtggtttgt	gatcttggaa	tgatcagact	tcctacactg	14220

gttctcccca gtaatttgca tatcagccag atacctcct ctagggtag ctatgatcag	14280
ggctgctgat ttctggcact gatggctcaa tgggaaggaa agcgccatc ccctctgttc	14340
tgtgggattc agcaatcttg ggctggtcct tccaggaggc ccaacctgag gacatgcttt	14400
aaccaacagg cttttatatt gagaaaatag tagttctctt ccttgccatt tccctgtag	14460
gaagccatgg cttgcaggac agccaggagg tagaggttca gagaaggagg ctaaaaaaat	14520
caagacctga aaaagttaag tttccaacca tatattccca aattgtagag tggaccatat	14580
ggaatcattg caatccctgg gtttctcatt tacaactttt gcaaaagata ttactatgta	14640
tgctgtcagg catgctggag ccggatgtct agcaccatt gctaagtga tgcttcggtg	14700
agagctttgg acttgggatt ctccagactg ctgagtgacc tacgtgtagt ttagcaacat	14760
aatgcttgga atgagccacg ccgacctgat gcgagaaccg gatggtagcc gaagagagca	14820
ggccaggtag tgagacaact atgatctgca gaggtgggca gcgggggtgga gacaggcttc	14880
ttctgcacca agtactgcta cctgaactgt acatcacaga gacaagccct gccgagagca	14940
gtggaagcag gaaaacacag gcctatctct gtgtctgtca tgtaagtaag tctctcttgg	15000
ctgtgtgtac ctgccacat caccatcc caccctaat atacagcatc accttcagct	15060
atagcttttg gtttttgatt aaaggacagt aatatttcca ggaggagaa gagtgatgaa	15120
cagaagcatt tatgtatgga cactgggcaa tttattgttt ttcaaagtct aatctagaac	15180
atgtttgcaa gaaagcgtga attgagtaaa aagtttacta tcgtattaag gactgttagg	15240
tacaatgact gaaggaggag ttaagagtga tccttcaata attcccctgt gggattatgc	15300
acatttaagg aaaaatgttc ttcaggctat ttctgcctta gagctaggca tcattacagc	15360
gaaatagaga actaaccttt aagcaaggag aaccaagttc tagttctatt tctgtcaatc	15420
actgtacgca caacctctca aagccttagt ttcctctttg caaagtggga taataaacc	15480
tacctactta acacaatgtg gggtttcagg tgagatgatg cataggatcg tgcttgga	15540
gctgtaaadc tgtaaattac aaatatatat tatggtttca actggtacat tcctaagcga	15600
atagcacatt gctctgttgg gaagacggct cttctccaag tcaggctggg ataatgttcc	15660
ctgacaagac actgccatac ctaggtgttc cccaaacatt gtctctggga accttgagga	15720
agcaccataa gacatgggaa gaaatgttac agcgctggct tgaaagaata acaatgtatc	15780
agtctactta tttctgataa tgtcatcttg ggataagaga ctcaggtag cttagtggg	15840

gacatgggca tgcactgcac agtaaaaatg gtgtccagga aacctggggtt tatttcagta	15900
tgggttgccc acacttctgc caacccagtc ccctacttcg tccccagctg ctcttgatga	15960
actctctgca cacacttgca cctgtatctg tgaaacagag ctcctcctct tacatgagaa	16020
tggatctggg tgcaaatacta atagattccg ctaccacaat gtcccctgcc ttttttgttt	16080
acttcattta tgaaaatacc cttgaaacat ccatagtccc attttgtaga catgggtgctt	16140
tatgtctttg agattattaa atactcatgc tcctttctga ttgctgtttt cacctcttct	16200
ttaggcttgg gcttttctat tgggtggaatt tgctgttcct tttcatgggtg ctggccttcc	16260
tgagttgggt ggttgctaata tcttattgcc agtttgtcct ctgtgacaga ttctccgaca	16320
tgccctggat gtggcttccg tgcttggctt tagcttttta tctgggctct cgtgtcctga	16380
atatttacct tctttcagga atttctcaca gcttctagcc aatgaaagtc ccccttctta	16440
ttgtcaaacc cagacagtta taattttatt ttaaatgtac tttttgttac attgttttgg	16500
gagtagggct ggaaggactt gctaattgat gagtggtgga aagtgagaga aagaaagata	16560
tgatggaaga aggagcaatc aaagccccag aaggaagcta tcgcattgtt cttggattcc	16620
taagcctgcc agaaagagct gacttacttt acagtttctg agagaactat gtgtgctatt	16680
actagaagca caagagaaag aaaaatagga tgttcagcat gcttcattta tctaattgta	16740
aaaatgaact ctgcccagtg acttaatggg cataaactct gtttctaataa aagccactca	16800
ttcggcaacg cacttctgag ttcttgatat ggtaaggat tgtgttctgt gctggacagg	16860
aatacaaaaa tgcacggttc ctcacctcca agaacttata gtacatgtag ggaaataaga	16920
caacccttat tgaatatcac tcaagatgga aatgactagg gccatggaat gtaacacaca	16980
gagggtagct ggagttccta agacttctga tggatgattg aggagagccc tggatcaggt	17040
aagcaatttg aagggatagg gatagcacag acagcacagg ggtggaaaca agtgtggagt	17100
gtcgagagct tgaccagtac gcctgaaagg gagggagtgt acacagagcg ttaataggag	17160
ctctgtctcg aggcagcttc cctcagcccc tcccaggaca tcgaggtttt gggagaaaga	17220
gcctattgct cactctcacg gctcttctcc tttttctctg ctttcagttt gttctttgaa	17280
ctttttggaa acttcccctg ttctttcttt aacactgtgc ttcacatctt ggggttctac	17340
gttttgagg ttgtagtgct tgagatccag ctttcccaaa tgatttctct gaatttagta	17400
tttggtatgg gttttgctat tttgctgcca tcccagccct agcaaagaaa cgacttacc	17460

ggagtatgga cagggcttca gagaaaaccc ctaacattcc tgactcccga ctttacagag	17520
ctctgccaaa ccttgcccttg cgggagtaag aaaagcgcta acaagccatc ctctttggtg	17580
tcaagtgcag acaaactact tagcccctct gaggtcctcc aacagtaagc tactggtttg	17640
tgaaacccca ggataatcca tctgatttca gtcctgcatt tagtcactta gaacattctc	17700
gcacatgcat ttgctggctc atgtacatac gaatatacac atatccctct ttgtgccctt	17760
tcttagcctc tgatgatttc ttctcctcca ggaggcagga atccaaggct tataaaccat	17820
gacttctggg aagttttttc ttctgcttaa ccagggtatc attgtttttt ctgccttccc	17880
ctggagaatc actggccact gccctagtgg ttggggcaag gatcagaggt agcttgcat	17940
ctgggggttg tcccaaagc ctcggtggga ctctgcattg gggctctgtag cctggatcca	18000
ctccagtacc ttaactaatc tcttgactcc cagatggtcc aaaatatgtg gatttagaag	18060
agcaacagac agctgttcct ctgggcctct ccaagaacac ggtttggtgt ctagaccacc	18120
ttagagaaac atggcagagg aaatcatggt ggagcagcat ggaaacaggt gaaaccaga	18180
cttagtacct tgttaaattc catcctggag tggagatacc agaggagcag atattacctt	18240
tattaactga tagaaatgtt tggggatttc tctgacttct tgtaggggtg gataagcccc	18300
aaagtgaaga gaattttgct ccttgtttta gccattagga aactcaagac cctgctacag	18360
tgctattggt ttaatttttc cctatcacat tgcctctgca acttctgaat gggtgcagcc	18420
atttcttaaa atttccctgc attgtcactc agacaacaag aatagatttg gccttcttca	18480
tctcaaaata atgggtcatga ttaatagtta ttggactggg aacagtgctc agccctctgt	18540
acgtgatctc aggaatcctc acagtactca atgaaatagc aattttatta tctcattttt	18600
gcagacaaag caacggaaac ttccacacat tttctacatt gcacctaaga tcatctgaga	18660
aactatgctg tacttgtttt tctaattgtat gatctgattt ttctattata atgttaattc	18720
tatgaggaca gggctttctg tggccttgct tcattgctgt atctccagca cctggactac	18780
tgcatggcac ctggtagtta cttagtaaag gtttttcaaa tgactgagta actcatccaa	18840
gattaaatgt ctaggaagtg gtggcaccaa gcttaggacg actcttttct gattccagag	18900
tccagacagc cctaaccact atcccacact accttcttgt ttatttttaa atcattttcc	18960
ttcccttcaa tccctctcca gtgccttaca ccttcttgct gtaatttgaa gcatggccac	19020
agtaagctac ctcaagtttc tcatctgtaa aatggggata atataatgaa ctaccttatg	19080

ggattgtacc cctctgcatg gtagcctcat cctactgtgc ctcctaacca cggccttttaa	19140
atcagcaggt atagttaata tatttagttc ttttaatatc atctgaaaca caaagcattt	19200
gcttccttaa ttcaagattt ttggctttgc ctagactaag cttaaaacca aagaagtact	19260
gcagaactga ctgaggctgc cagaagtacc acactcttgc acccagccag tgggaagtgg	19320
aaagataaca gctaagcctt tggggatcct tccagaagta gtgatgacgt acagcattct	19380
ttctgattat gaagtaaata tctgttctaa tgtatgttca acatagagag ttaagaaaat	19440
ggggaaagaa taaagagtaa aacaatgacc agaaatacct tcaataccct ttgacattct	19500
ttctctgtgt gtgcatgtgt ttgtgtgtct ttgtttctgt gtctgtatat gtgtatttcc	19560
tttatttttg tttttttact ttaatgtaat ttttagagac aaggctttac tttgtcttgt	19620
agactcgcgt gtgggtggcgt ggcactcatg gttcactgca acctccaatt cctgggctca	19680
agcgatcctc ttacctcagc ctctgtgta gttaggactg caggcatgca tcaccattcc	19740
tggctaattt ttcaattttt gtgaagacgg gctctcacta tgttactcag gatggtctca	19800
aactcctggc ccgaagcaat cctatcacct tggcctccca aagtgtggg attacaggcg	19860
tgagcgacca tgcccagctc ctttttataa ataaggggct caccatacaa tataaccagt	19920
ttttacctgg cattttccag tcattattgc attgtacgta tctccccatg tctttttctt	19980
ttcttttttt tttttttttt gttgacggag tctcactctc ttgccgggct ggaaggcagt	20040
ggcgcaatct cagctcactg caacctccgc ctcccgggtt caagtgattc tcctgcctca	20100
gcctcccag tagctgggac tacaggcgcc cgccaccacg cccagctaatt ttttgtattt	20160
ttagtagaga cggggtttca ccatgtcagc ccaggaaggc ctcgatctct tgacctcgtg	20220
atccacccgc ctcggcttcc caaagtgtg gaattacagg cgtgagccac cgcgcctggc	20280
cctctccatg tctttaatta ttcttgcata agatgacttt tctactgcata atattccatc	20340
acataccact ctttaaccat tttgcttctg gggcacattt tccttttggt cacacttttt	20400
atactacagt tgccatcctt ttacatacat ttgaatacac atatctggct attctctcag	20460
aatagatttc cagacgttac ctttccaagc ttgaagatgt taacatttta aagaaagatg	20520
gttattcttg aaagccctga cagctctgag tggggagccg gggctgatgg ttaccacagg	20580
atagcggaaa ggcacactgg ctggcctgtg tgtactcacg catccccca cctagggcag	20640
ccttgggaag agcactcagg attatgagaa agactgtcgc ctccccttg cttcattagc	20700

tgatcctcta agcatatgtg ctttcttggt ctaattttcg gattggctctt ctcctatatt	20760
ctcttcctac tccccacccc gaccttacag ctaagtgcac atctcatgta gtgcagtggg	20820
aaagaaccgt aaggcagaag cggggtgac ttggctgtga atcccagctc catcacttgc	20880
tggccaggtg actgagtaag atcgtttaca catccatcat cctcaagttt ctcatctgta	20940
aatggggat aataatgtaa ctgccttatg gcattatata aggattgtat gactgaacac	21000
atgtagaatg cttagaaca tgcctggcat atatgaagca ttttaatacat ggtgtattaa	21060
attagttttg aaaagaataa attaataaca atgatgaaca tttttgatac ctattttcct	21120
attgttttga ctctcaaagc cagttgcaag catatttagc actgtgatgt atgtgtgact	21180
tactgcaaag tcttttttcc agtccctgat accagctctc tcttcacctt cagtgtttcc	21240
taccctcct gcctcccctt ccctaagaat attgctgttt cacagagtgt aggctttcct	21300
ctggcttcca gatctgcca catatgcaca cttctctttc ccatccctgt tggactcttt	21360
ctccttatca gtttatttgt tccagttggg aagaactgga acctggtcgg cagcttttcc	21420
agttggcttt atctgtgctg tgcattgtaa aactgttctc tcttgcttag aaatctcttt	21480
gatccatgtt tagctgtatt tattcttcca acagatgttt tgggtagtga gaggattttc	21540
ttctcgcat tgcctagtct catgctcctt catgcttccc acttgttcgg gatctttttg	21600
ccagctgacc acagacaggg gccatctgtc gtgaaggctc ccctggcca gcagaccagg	21660
aatggcccag caaccaagac tttctgaagg gcttagtgaa ggggaggagg gaggaagatg	21720
ttggagaact gtgtagggta gagtttgagt ttcccagaca cattccagga gctcttttga	21780
tccaaggtat acatgatttg gcttgtgctc tgtggcaggt taacaaaaac acaaccttcc	21840
attgtctcct gtagacaaca gagtgaggcc cttgggcatg gcaggtagcc taagactacc	21900
cctgagagtt gggaagtgtg tgagtctcct ggggctgccg taacaaagca ccacaaattg	21960
ggtggcttag cacaacagaa atgtattgcc tcacagttct ggaggccaga agtccaagat	22020
caaggttgcg gtcagggccg tgtttcatct gaaggcccag ggaagcagct gccccacgcc	22080
ttctcctagc ctctggtagc ctctggcatt tcttggtta tagatgcatc tgtcaaattc	22140
tgtgtcttca tatggctttc tcctttgtct cacactgtct ttcctctgtg catgtctgtg	22200
ttcagatgtc cttttttata aggatgtcaa cccaattgga ttaagttcta ccctaattgat	22260
ctcattttaa cttggttacc tctgtaaagt cttattttcc aaataaggctc atgtgctcaa	22320

gtactaaggg gttaggactc cagcatatct tggtaggtaga cacaattcaa cccataatgg	22380
gaaggaaaga tggtgggcac ctgtaactcc tccaaacacc cacagagtgc agggtagagct	22440
gtgtgctaac acatagtcag ttctcttttg ggtgaggagg cctaggggca gggcccccatt	22500
gtggggtctc tgtccacacc agcaacaata acaaccaggg aggaaagcat ctcatTTTTcc	22560
ttggctcagt tcagctTTTT atgtTTTTtag cacaatgcct gctttgctct tccaacaatt	22620
tgggaatctc tgggagctgt gcatggaaag caaggaggac agcggcgaga aaaaggggga	22680
gtagatggag ggtcttgga agcagagggc ctaggcaggc agagaggaca ggaaagtata	22740
gcgagcagag cggcaaattg gtggggaggt gcagaaggct gcttggcagc caggagtctt	22800
tgccctggcc ctgccatgag gctgcatgtc tgtggcctag gtattttacct tctccaggcc	22860
tcagtttctc tgtaggcaag attgggaggt ggatgggtgc tctctaggat cccttcctgg	22920
ccagaataac attctcagca ggagcctaac gtgtggagca aatgggagca ctgggctccg	22980
gcctcctgca gtgagcacag cccctgttct tgtggaaaca tcttccaata gggctgccct	23040
gcctacaggg tcatgcggca tgcattctgt gcctgcctgc gctcttgaaa cagcctccac	23100
tgctcccctc ccagctcctg tctctctgca cacgcaagcg tgctactcct tttcatgatc	23160
cccattagta ttctttgacg atggcataca tctgtcttcg atcgttgtca gctctgggag	23220
gcttatgcca agcttcttga gcgtaacca tgactgcctg ggtaggtgt tgtgagctgt	23280
ccaggaggca ggaggacgat gcatgcaagt cagggttag ggcaagaagt cctgggcctg	23340
gcctcccctt ggactccagg agtcctgtcc taacagagcc cacagcccc tatccatctg	23400
gcctctgtaa cccctcccc acacacacac acacacacac acacacacac acacacacac	23460
acacacacac atagcccctg tgattgaggg ggcccccaatt cctgttcata tcctccagga	23520
tagcccacct gcaccctcga cagttagaga caaagttcta ttccctgttt agatgggtgc	23580
tggggacaat ggaaaggagg tgtggctctg agaagttcat gtcttgctca gggcacacag	23640
cagctgatcg ggaacatgtt gctgactcca agatgctgcc ttgcaagaag ctggctctat	23700
ctttcttttg gctgaagtgc ctttcatgga tggtagggga tgtgcaggga gaagtgtcag	23760
gagtgaggggt cagtgggttag aatcaggcag tccacagagt ctgagaaagc aagacattct	23820
ctggcagtct gggggtcatg atcgcccacc ccagcccaga taaccctcac agctgtgcgg	23880
gccactagag aaaaaggagg gcatgttttg ggcaggagag gcaaattgtt gcttatctgt	23940

gacttcttcc tccaagcatg tccggacctc cagtcaatgg tgggctgtca gtcgtcagct	24000
gaggttgagc tttccttagc aggagcactg gtcacttggg ctgggatggg tcttagtggt	24060
acaggatgca ctgcaagctt taaatgcgag tggcatcatc cccttccggt caccatggca	24120
accagaaaca ccttgacaca tttccaaatg ccctttagta gggcagtgac agcccttttg	24180
agaatcacat agaatcgcat tgattgatga gtgaaaaata aatggatggg agcctccttt	24240
tgtgattttt gcagcggcct ttagcttcct ttactcacc cagaaatcag tgggaccctg	24300
ggagctgtgt acccctcaga ccagttgga acccagccaa gagtacttaa tccatcccca	24360
cttgtggggc caacggcacc taaccacctc aggcacgggtg gacctggctc ctgagagagc	24420
tctagggaca gaggagagaa agggctctgca ttctgtttgc agccctgac gtgagctctg	24480
ggggtcctct tccaccccca ccccccaccc cagcccctgg agcaggtact cggggtcaga	24540
gctctgctga gggctctggct ctgggagggg aggtttgtgt aagattccct cccacgggtc	24600
agcacagatg ggatgacaag gaccaaattc tgtttctggg ctctgatatt tgccaagatt	24660
tttaccaggc ttcctggaat agacaggga gacagagcaag ctcccgtagg tcaagtgatt	24720
tgggcccag ttgaccaga gtccctaaat gactgctgtg tagctaccat gagtgtgctg	24780
agtggcccat aggggcaggt atgagagagg tgctgaggga ggcagggggc ccgcagaacg	24840
gcctcccatc tccactgccc gtccccaggt ccacaggctc acagagcaga cacggctctgt	24900
gcctggggtt gctcaccac aagaggaaga acataacatc tcgctccttt tgctgcacag	24960
gataaaacga gaacagagag gaaacaggaa gtgctttgca ttccagaaag agcagcaact	25020
gtataaagtc atgcatatta ggatttgagg tatgcatggg cagaagttag aaactaaccg	25080
aatcttgtca ttgccaggaa gtttcggggg tctgtgactg gtggccactg atgttcctgt	25140
gttcctccat tccagctcct accttgactg tgtcctcctc ttcacaccta acttctttag	25200
tgaaggctcc atttcctcat ttcctgttca atgcttaaac cccttgcaat ctggcttcta	25260
ccatgcctt atcacagacc cttctctggc ttgacctgct ctggcccttc catgatgtct	25320
tctttcttga aaccctcttt ccttggtagg ataccacggc atcctgggtt ttgtcctacc	25380
tctgtggctg tttctgcgtg ctttccttag ctgacttttg ctcttttatc tgacctgggc	25440
tctcctctct ccatacactc tccatagcct attctaagtg tcccaggctc cttatatctt	25500
atctctcaaa tgcacaatta ctttgtgtta gctacagacc catatatcca gcttcctat	25560

agatacccca aatgtctttg taggctccta aactcagtgt atgctaagct gaacaggggc	25620
tccttttgtg ccccaaactt tcccaactcc agtgagtctt ctccattgtc tttcacctta	25680
ataaatggaa ccacctgcga ctctagtgtg tgggccagag acttggaagt cacctcagct	25740
tgtctctctg tcatccgcag gatcgggcag cctccaagtc ctcatcattc taactctcat	25800
aatgcctctg gagtttgtcc agatctcctc gtcaccactg ccgctacgct aatcaaaacc	25860
accattgtct cttgccatcc tccatacttt gcaaagttaa ttgggtcatt tttctacttg	25920
aaatcttata atggctctcc agtgcctctg agtccttggt ttttcaacac tgttcacact	25980
ctccccacct ctctctcact catacccat gcaccagcca tcctgggttt tgctgttttt	26040
gtttcccaga atgcacaatg caccttctgg cctctgagcc acagcacctg ggtatttgct	26100
cacactactg cagctctcct cttccctgcc accacgcctt tcttgccctga ctgttagtat	26160
gcagcagtgg ccacttgagc atgaccgcct ctgggaggct gtccctagtc ctctgtcgca	26220
ttctggggct ccctatcaca cactcccatt gcatgctgca gcatcctcag caccagcat	26280
tccttattgt agttcctgat tcaacacctt tctcaggaga ctctggactt cttgaaggca	26340
ggaacaattc ccacttgttc ctagtagcat tccaaaccac acgtgacagt gtctggctta	26400
taataagcag ccaataaaaa gttgatgaat gaatgaataa gtgaaaacag aagggtgttg	26460
cctgcagaaa tctggaataa gatcaaagat cagagctggg attaaggaaa aaacttcctt	26520
ggggtggcac tatgaattcc cagaacaggt gactaaccct cattcacttt ggcaaagtgt	26580
tatcccatgc cacgcaacca ggcaaaaagt tgaatgaggt ttaatccctt cccacacgga	26640
gcttattcct tctttggaag tccttttaac aagctctgaa atgattttgg caggtagaca	26700
aactggtcct catttctctg tgaccagtaa gtagggaaag caagcacaca tacacacaca	26760
cacacacacg tgcacacgca cactgacaga caaccttgct cactcacatg ggcatgcccc	26820
aaccttctt ctattttata ggatggtaac tcactcttta gtttagactc ttgacgtgcc	26880
atggaaaatc ccactcgccc tagaactggg ggccgggcag gtttgactgt aacaacgaag	26940
cctggagctt actctttgct gattggcttt ctttctgtc tccatttttc ccctggtgag	27000
cactgcagtt gtgttcttcc tcccaaaggt aatgcctggt ttggctcact aaaacctgtt	27060
ctttctgtac cgagagctca tcttctcttc ctcttctgga ttctcaaag agatgacgtc	27120
agaggatgga ggccaaccac accttcctc cttgaccctg ataaagtctt ttggaaacct	27180

tatactcaga ggcagccaat tcttgccagt ggaagagtga aaagagggtc tgggaagctc	27240
aaggctcagt gtctgtcccc aggtccccca gttaaagaca catctgtcct tcactctcaa	27300
agatgttgcc attgctcccc tgctagagtg acacactgca ctccctcctt cccttcacac	27360
cccagcaaga ggctattttcc caggggtctt ataagcagat ttcattcttct cttgtgctgt	27420
tttcttattt caattatctt cagggaggaa cgtgcatatt gcgtcattgc ctggctgtga	27480
aatttcattt ccatttcttt acacctgcag ttgcaatacg agagagaaaa ggccagagct	27540
tagcggatgt cctagacgca ggttatcaag gtgctgtggc tgtgggtttcc cggaaaaggc	27600
cttgggtcca gagcacattt tatcagcagg accttcgagg ggctgcgttc cttcaattgt	27660
tttctctttg gggctctctgg tctccagttc tttcttctct agcatgtgag atctgtgctt	27720
ttgattcatg cctttaagtc tgacattgaa aaaatatctg atttgccatt ccagatgctc	27780
gtcctcattt gcaaattttc ctaaagggcc agattgtcct ctggcctttt cccttttcct	27840
gggtcccacct caccaccctc ccactggggc ttcacagagg cagagctagt ctcttttcat	27900
tttttaaaat taatagtctt caatttttag aacagtttta agttcacaga aaaattaacc	27960
agctattaca gagttaccgt ataactcctc cccctcactc cccagttttc tccattatta	28020
gcatgttgca ttagtgcagt acatttgttg cagttaataa gcaaataatta gcccatcatt	28080
attaactcaa gtctatagtt tacattaagg tgtattcttt gtgttttaca gttttatggg	28140
ttatgacaaa tgcataatgt tatgtatcca ccattatagc atacagaata gtttcactgc	28200
cctaaaactc tcctgtgctc cacctgcca tccctcctcc ctctcttgcc accaatccct	28260
ggcagccacc agtcttttga ctgcctagag tttcgccttt ttcagaatat catagtagtt	28320
ggaataatac agtgcgtagt cttttcagac tggctccttt cacttagcaa tatgcatttt	28380
aagtttcttc catggaaact ttgctttcat ctttttatca ccacaaggcc agtcatccaa	28440
ggaatttctc catctctgtc tgttcctttc tagttctatg tgtgccactg cttggcatag	28500
aataggtatc catttaatga acattccctt tcaccacctg ggacaccttc ccagggataa	28560
caaaaataaa accagctagg tcaatagcag agccccatc ccagttttaa cctcattctc	28620
ccctctttcc acaataaact ggatcagaac cagcagctct gtaagactgc atttctttcc	28680
cttaatacca ggccccagag agcatttgat tccttggcag agagggtgtag gcttaattaa	28740
ttttctcct ttttctttg aacatcttgg aacacacaca cacacattcg catttatgca	28800

caattgggtg tatgagaatt ttaatggcag gtggtgttag cagttctttt cctcctgata	28860
cagatcaggg tttttccatc tgggcctttt agcagggcct atgaatattg actttctaac	28920
cacttggatt tgggtggagt gtgcagagtg ggattggggg ggaagggtca agggagaact	28980
atacttatgt ataaatcaca tgtgaaggga gttttgaagt cattattgct tcaggatgtg	29040
cgaaccataa ttatTTTTTT aggtcttgat ttgccccaaag agcatttccc agggttgctg	29100
ctccaagcat gacgtctgtg ctgtcaggag gtgcagcata gtctgattcg agtttaatcg	29160
ctttaaagga ggccctgggt aggatctggt ctctaggttc tcagctgtgg tcagtcctcc	29220
atgcagcaaa acatccagat gacttagatg attaagacag cagacttaaa gtgaagaaga	29280
gattttttcc cttattcttt ctttttatta ttagttttta aatggttggc tacatgggct	29340
gttggtcatt ctccatgttc tctgtgctct cctcagctct ctgctcaaaa caggctgcac	29400
cggcctgcct aaaccctgaa agcaacttct cagctgccta ctttctgcct ttgaccccc	29460
aagccaatcc ccatctcctt accaccctcc cgccatgtcc tcatacacct gcctctcctt	29520
gacttcattc ttcatgctcc atcagcaaca gccctctgtc aataatgatt gtcccaggga	29580
agtgtattca agggtcacat aaaatgtgcc ctctctatgt gttgagaagg ttttctgtcc	29640
ccaaaggagc tctctggata atgaggaagg ttgaactggg gcagcctaca ggaagaagcc	29700
cttagaaggg aaacctgtgg cataaaccat gctgatccac gactcttatt ttggaatagc	29760
tattttaaaaa gaaatatgaa gaactcgtaa gacttggaaa agaaaaactag agaatgttga	29820
aaatgtccaa gggttatgtg tatgatgtgt atgggggaaat tttaaaagaa tgtggtagaa	29880
aactgaattt gtggtaaaat gttgtcacag gacggcctgt tctttcattg aattatgtct	29940
tagtgcaggg atgacaaata aatacagcaa ctgtgctgcc attctcacat cttttcctac	30000
agtaggcac actaatcaat tacaacattc ttttccactg tcatgccttt gcctacagca	30060
gacatcacta atcaatcaca gcactctctc tcattgagtt tcaaagggtt ttaaatcctc	30120
cacataggtc tctaagtagc tatataccaa ttactttgat ttagaacttg gattaccaat	30180
ttgccatctc tagtttaatg ggaagaacat taggattaga gtcagaatac ctgaattcaa	30240
gttactgacc tactacttaa taataattaa ttttaataat tctctaacct cagctttctc	30300
ctctataata agaaataatg cttaccacag aggctcgagg taaggattaa gtaagataat	30360
atataacatg taagataatt tgtgtgatgg ttcttagaat agtacctggt gcattgtaag	30420

cacccagcaa	gtgatagcca	agtatgaatt	tgcagcggca	tggacacatc	atgccacacc	30480
ccagatatgg	aacaatagtt	aggattccat	tcacctttgc	ttcttttaac	atcctcagtg	30540
aaggcagatg	gagaacagct	ggaccctcta	attctacctg	actttaactt	cacattctta	30600
agatgcttat	taaatctctc	tttcctgaac	taatataatg	cttccttggt	accctggaag	30660
gacagagtta	acacccatgt	attatgacat	gatggatggt	ctttttgcgc	acactgcatg	30720
catagccatg	tgccaaggcc	aggcctgcct	tagctttttg	gttctccaga	gcagcttcac	30780
tgtgactgaa	gagagccagg	gaagacatct	tggtagagct	ctttatatga	ctcttttcag	30840
aatgtttctc	actgatggag	atgacagaaa	gctaggatga	tttgaggca	ggaggacagg	30900
cttctttgga	aaaggtttca	ccatagatta	cctcaatcca	gggttggaag	actggaaagt	30960
ggtctgcagg	gcatcccagg	ggatgccttc	ttcaagacaa	tattagacta	gcattggacc	31020
ctgcctccag	actagtaaag	ggttattttc	agagcagcaa	caccagagaa	catgttttga	31080
gcaaatcaaa	tccttaaaat	catgatttca	tacatcctga	gatactagtg	acttcaaaat	31140
gcctttctca	aattatacat	atagtgcctt	ctgaatcaac	tgtctttttt	ctgtccaaca	31200
gtataaacat	aatctctgcc	ctcacacagc	ttagaaactg	tcatgagagg	taaacaaata	31260
cacaaatgac	aaggtaaaat	agaagagcta	caaaagagat	catgaaaata	tccagggagt	31320
tcgaagggtg	cagataacaa	aaggcttcgt	ggaggaggta	gcctttgtgg	catttgagat	31380
cagtcctgca	gcagggttat	aggtggacag	gcagagatgg	ttgggggtgg	gcggtaggag	31440
ggacagacag	aaggaaccac	aagaatgagg	atatggagaa	aaagaccact	gaagggtgca	31500
aattattaag	gatggattta	tgtaaaggat	gtggcctaag	aagaaggact	cttctttatg	31560
atgaaattag	gttttagagg	cagaagtctg	gtgcttcccg	gtactaaaat	gaccccagta	31620
tggtgtctgc	agaatactta	gataactgca	tagatggttc	agccttcttt	ccattatacc	31680
acactactgg	gttaccaact	tgctgtgtga	ccttagataa	gcaactacct	ctctctgggc	31740
ctcagttttc	atatctatga	aattaggagg	ttgaattatg	tcttggcata	ataacgaaat	31800
aaataagtga	tttattttta	tttctttgtc	attcctcttt	gaaatgggag	tggggaataaa	31860
gtggtatattg	tttccacatg	aaaattaaag	ccagggggcc	aattctggac	ttggcttgag	31920
ctgtgctttg	tctgtttttc	tgcactggcc	cagtaccctt	tacactgggtg	taactaactc	31980
cctggaaagg	gatgcaggta	gacagcgtga	ctgttccttc	ctatctgaga	ggccccaaca	32040

gattctatat tgcaggacca ccaacttggga atttcagctc agttcagcaa acacttttga	32100
ctggttcgtg tgctagacat tgtgctgggc ttatgtgggg tagagatgcc tgctaatcta	32160
aggctaacag gaagacaggg aagtggtttc cagacctgat gcttatcaga gtcactgtgg	32220
aagaactttt tagaatgcag acagcccca acttaacagt ggttggactt tcaatttttc	32280
accttataat ggtgcaaaag tgatgggcat tcagtatgct tctcaacata caaagaggtt	32340
gtgtgaaata aacccatcat taattgaaga tactgtcaat caaaaatgta cttttgactt	32400
acagtgtttt caacttagga ggggtcatag gttgaggagc atctccacag attcctgagc	32460
cccatccctt gatgatgatg atccagtgag cctgggtggg gcctgagaat ctgcatacta	32520
actgctgaag gtgagtgtga tgcagggcca agcttcagaa cctctgatct agagaagatt	32580
ctgtgtcatc acagctcagg gtgaccatgt tctctcttac ttgcacttac atgaattcat	32640
atgaattcag tattacatga gtaaattatg attatatgaa tttattgagt atcctatgtg	32700
cttgatatag gtgtttacat cccagggttg gggggtagag atgaagaata taagccggac	32760
tatgttaaga tcttcacttc tcagcacctg aacaactggc agcctcatta ggaaggcaag	32820
tcattcacag gtgaaatgac atcatggctg tcttcttcat tcttctcttc agcatccatt	32880
taagactcac attttatcac cagagactga aaagagccac ctaaggcagg caggtcaggt	32940
ggtgttatct ctttatttcc agatgtggag gctgaggctc agagagggtga atccatatgt	33000
ccaagctcac atctcccgcc ctcagtccag ggcttctccc cacttcatgg gagaagcatc	33060
ctcctcccca gagcagtagg ttctggagct gggagaggcc actgtgggct ggattgttgg	33120
ggacagcttc agggagagcc cgattcaagg caatagagaa ctttggctgc aggccgttgc	33180
ctagaatagg gcagctgaca cacctttgat ctggaatgat tcctgctgct gagaatgagg	33240
ttttttatat ctggattctc aggtagtaac accacgacaa cgtgtgtttg tgttatttca	33300
tctgcacagc attcacgtgt agcagagaga aggtagttta ttccccagaa gtttaccagt	33360
gggaaattga ggtcaaagggt aggactttcc taaatgaaga gaactactat ttattggatg	33420
cctcccatct gccaaagcagt gtgcacagca ggcatatatc atgaaatgcc tcattttattc	33480
ctcataactg tagcctgatt ttacattcgg caaacctgag gcttacagaa atcatgtggc	33540
tcgcactcag actgctgatg gccacgcaag gctttcagct tctctgattt ccaggcccc	33600
atccccacca cactgtgctg tccagctcac ctggtggaac tggatcctt gagttccagg	33660

ccagggatcc tactgagctc ttccatcagg gaaaccatag cagtagcagc tccaccaagg	33720
acttggcatc tatcttactg cagcatccgt gcctgtctaa tggaaccatg taccaggagt	33780
agctacccaa gaaacatttc caccagaaac ttctctttat agctcccat cgagcctcag	33840
agagctgata ggaattgctc aagaccaccc aacttgtaag tggtagagct gggactaaaa	33900
ggcaagtccc ccaactttca gcctcactcg tgcccaatgt gtctcagcct ccctgaagaa	33960
tatcagacca agatggccag aaaaggaacc tggaaggac gtgtgggcgt catgcagccc	34020
ttggcacagt cttcaggctg agctgcctcc acctgtgcat ctcatcagag ctttacatcc	34080
atccttggga gaggccagct ccatgacctc tcagtgtcat ttaggatctt cttcctaaac	34140
ggctaacca agtctgtccg ctgcatccta cagtgagatg cactgtggcg agagcagcta	34200
ttggcattct gcttacgtgc tgtttccaga ggtaaactca gtataaatgg atctacagcc	34260
tgtcccattt tgtacagtcg agttctaaaa ccagcctgca aggataatgc tataaaaatg	34320
tcctgccagc cccaagggtg tcttctacaa caagttcttg ttccctcata attcttctga	34380
caaattcttc ttttgatcca gactttccta gccatcattt cattccagaa gtggcatgtg	34440
tgtgcaaaca ctcatgcttg agggcgggac aagcaaaggg ataagagggg aaatgggaac	34500
tcgaaatctg ctcaaagtgt gtaaagaaat atatccagaa agtactgcct actcacaaa	34560
atactatttg ctttttatat tctttcctga gtagaatttc ctgttcaaac ttgaaaatga	34620
aaattcctcc acttcaaaat gaacaggcaa gaaatgctgt aggctgggtt tcccggggag	34680
gggggactga cagccatctg ccctgagact gactgtcaaa tctgaactct gtgtacttgt	34740
tagtgttgtt tatgggaggg gtgagggaag ggaggaagca acagggacct gctaacccta	34800
tgaattctcc ctcataccct taaaaagtcg ggtgcttggc cgggtgcagt ggctcacgcc	34860
tgtaattcca acacttggga ggccagggcg gttggatcac aaggtcagga gttcaagacc	34920
agcctggcca agatggtgaa accccgtctc tactaaaaat acaaaaatta gccaggcgta	34980
ctggcaggcg ccggtaatcc cagctactcg ggaggctgag gcagagaatt gcttgaaccc	35040
aggaagtga ggttgcaatg agccgagatc atgccactgc actccaacct gtgcaacaga	35100
gcgagactcc gtctcaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gtcgggggct	35160
tgccaagctc gctttttggg gcaatgggag ggaaatttta agagctgatt tctgtgatct	35220
ttatcagacc tcttcttctg tcctgcccac aggtcagtga tgggaatgat agtattgcat	35280

taaaagagaa	gtgattcccc	agatgtggac	atctcagtgc	ctgggaggca	gacgtcttta	35340
gagtgccttg	tgttcctccc	catatctaac	ttacctgttc	ttaggacata	ttcccttgag	35400
tgtcttttat	tatttcgttg	attgattaag	ggttacaaat	ttgttacaaa	agcagcagaa	35460
attttgaaag	ataaaacagg	tagaaactgt	tatttaattg	gaggcacaaa	agccagttcc	35520
agcttgagca	ttaactggct	gtgtggcttg	ggcagctccc	ctgcactttt	gggaccatgg	35580
tttcttcagc	tctaaagtga	aagctgagag	ttcctctgaa	tgctaaaatc	ttgcttctct	35640
gtgattattg	ttttttaaat	gccgccgtca	tttcaaaaca	catacatggg	gcagccctcc	35700
tcagctgcct	gtcctgttct	cttcctcctc	tccctttgtg	ccatcagcgt	ctcccctgga	35760
tttatcgtgc	tgtactcctt	ggacaccctt	gaaagtggag	agagaataat	tctcattccc	35820
tttcccaaaa	ctctgtcctg	aagcaccact	ccccacctcc	ctaccatgcc	tgcccctcca	35880
ccccccacat	tttccctctg	atgagatttc	ctctctgagg	gaatattttg	gtcctgttct	35940
ctgttcccaa	acaaattccc	caggtccttt	aaatctgggt	tcaatcaaca	taatcacagt	36000
ctcttcgatg	ctatcaactc	tgccttgatt	tggttgtgtg	attcatgctg	tgtcctccac	36060
ctcctcaggc	tgttttatat	atacacactt	aagtgtcat	atatatgtgt	atatattaca	36120
caggttgtat	gtatatgatg	tattgaaaga	gagagatgtt	ttccaccgtg	cctgtaatta	36180
ctttttatct	tcctccaaat	cagtatctta	atgcatatca	cctaaactgg	attcattttg	36240
agaacactag	ggcctttttt	ttacttttca	agaacagcta	tctaattctg	ggccaaaaac	36300
tagttatcaa	atgagggcag	gtaatggagt	tgtatccagc	gagggggcat	ttgcttctcc	36360
atatgataca	ctcctcacca	gtaccaaggc	gttttctctt	cagtcttcat	gcctatgctg	36420
ttacgagtcc	tttctcctat	ttgaaagaaa	agatatgagg	cagcccaggg	aaagctctga	36480
tggaggctat	aaagacaata	tggatgtaaa	ctaaaaaatg	gtagcttaga	gcttaatggt	36540
agcctctcag	gacactcacc	acacctgaat	tctaccagcc	tgacgccaat	gctgctccgc	36600
ttatgctggg	atgtagtgac	acaggacact	ctgctggggg	atagggatga	tctcccccca	36660
ccaagtgggt	catcaacact	catcactcag	aggggaggtg	ggataccacc	ttgaacagag	36720
aaagcggctg	ggcatgggtg	ctcacacctg	tagtcccagc	accttgggag	cccgaggcag	36780
gtggatcgct	ttagcccagg	agtaagagac	cagcctgggc	aacatggcaa	aaccccatct	36840
tcattaataaa	tagaaacatt	agcctggctt	ggggttgcac	acctatagtc	ccagctactg	36900

gggaggctga ggtgggagga tgccttgagc ctaggggggtc gaggctgcag tgagccatga	36960
ttgtgccatt gaactcctgc ctaggagaca gaggtagatc ctgtctcaaa aaaaaaaga	37020
ggaagaagtt gctgaatccg tcttataaat ctgtaacaga aaattatggc aagttgtggc	37080
ttttaattgc cagagctggg cacttaagag agaaaagggt ttttgtgaat tccaaaaatg	37140
taattgtatt cattgaacac tgaagtcagg agaagtcaga ccataatgac atgggagcca	37200
ttttcgcaac cagcagacta gaaggggagg gtttggaagg gtggcggcag agcttagaca	37260
cctgctgcga agggagagaa gtggggacag aaagatggcc tgtgtagagt cccatgagaa	37320
agacagaact gcactggcag gcatcttttag gggcccagggt cacaatcatg gggccgggtg	37380
acagtctcca gggcactgtg attgtcacag tgcacagcct aatggggaaa attgcacagc	37440
ttcacttaaa tataggtgac atatggacgt aagaattcat gattaatcta gaactaacc	37500
tgaccagcaa ggccgaatga agaaaaaatg tagaaacact gaccaaacct tcctcaaagg	37560
tcacagatct tagggaatgt gtctctcttt cacatttcaa aataacaaca atttttgaaa	37620
tatgtatfff ttaacattta attgttttat ttggggataa ttttagactc atggcatcca	37680
ctgttttttt agttaccac atatttgatt cagtaagatt ccctgtattg atacaaaaaa	37740
agaagaggga ccctgtttcc ctttgacact gttacttttt cctgagctca ttgtttgttt	37800
gtatfffgtt ttgttttgggt ttggtttttg agatgtagtc ttgctctgtc acccaggcta	37860
gagtgcaatg gcgcgatctc agctcactgc aagctccgcc tcccgggttc actccattct	37920
cctacctcag cctcctgagt agctgggact acaggtgccc gccaccatgc ctggctaatt	37980
ttttgtatff ttagtagaga tggggattca cgggtgtagc catgatggtc ttgatctcct	38040
gacctcgtga tccacccgcc tcggcctccc aaagtgtctg gattacaggc gtgagccact	38100
gcgctcggcc tcctgagctc atfftaagag agacttctgg cctagagggt tgaatgagaa	38160
gaagttatac agctgggatt cttccctttc tctgatatga ggacaggagt tctctctcat	38220
ctccgccaag agcaggaagc tggagtaact gccacaagct ccaggaggga gtgtctagaa	38280
catccacgtt ttgcagcagg aaaacacccc ctcacgtga agtttgattc ctgaatcctg	38340
tgtcgcagtc taaatgctga ggcagaaggg gacatccgtg ttcctggggc attccacttg	38400
cagtcctggc tgtaaccga gtgagccatc cgtgtagttc ctgttgctaa gtctcccctg	38460
ccacctcttc ttcccatggc tgcagggcag ggggccatgc cctcctcttc atttcctgtc	38520

cctgggtgag	cgtgccccct	gccttctccc	agatctctgc	tgtggcagct	tcacgtggga	38580
ttcagcactg	tgtctccttc	ccctctgctc	ggcctgcccc	tacctgtcca	gcagagctgt	38640
aagaccagaa	gacagagcat	tccccttata	tatgaagtca	aatgcatgtg	tggaacatgc	38700
caccagcct	gcagtctctc	tactataaaa	tactgctcat	aagacaaatg	tgtggcccag	38760
atgatttctg	ataaagtcta	ttattttgaa	atacatatgt	atgtctctca	gccactgata	38820
cacgcagaag	ctgcacatgt	tgtcaacacc	tgctttgagc	tcctttcctt	cccacccttc	38880
cttttggcaa	tgcaagtttc	cattcatttt	ctgcattact	ggctctcctt	cttctcccct	38940
actactagat	cttacaataa	acatttgaaa	tagttttattt	gtccacagtg	tgatttctgt	39000
tgagaaacat	gggctcaccg	acttttgggt	ctcttctaac	actagaaatt	cctctggttt	39060
tttagactat	ttcaagggt	attaatgtgg	aacagacggc	ctttagaaac	agcaatccac	39120
agggggcccg	gagacctgga	ttcatatttc	tcctagcgct	aactagtggg	tgaccttaga	39180
aatgtcattt	tccttgctcg	gcttttagtt	ccttatctgt	cacacagaag	cactgtgtga	39240
gtttgggaaa	ccaatactat	gttgaaatgt	acaaaataat	cttaaagcac	agatattctg	39300
ttccctccaa	gaatacatca	aacaaaggaa	ctgacattgc	aagaagattt	tgaggagatg	39360
gctggatgca	ctggagcagg	gattgctgag	ggaagccagg	ccctcacctg	gagcgtctca	39420
ggagaggcag	cttcggtgct	ggctgtttat	tgcaaggcatc	tctttctgtg	tctgtgcgtt	39480
taggggctct	tctttggaga	taagaaaagg	gttctggatg	gagggcagtg	aagaacagtg	39540
agaacttaac	atgaggatgt	ttgtatagag	gggaagactt	ctggacagtg	gcttgacttt	39600
gctcactggg	catttccttc	tggatctctg	tagaagtcag	ggacagatct	cctcgtgcat	39660
atctgtctcc	cagagacaga	tctcttccta	gcagaaagta	gaaagtgggc	ttcaggcatc	39720
ctggaagttt	tctttcttgg	tgggtgataa	aagggttgct	agagagagga	gaatcaaadc	39780
tcccacatgt	gcatcatgcc	tgcgagtctc	atgcagagat	gtcttatgtt	caacatagaa	39840
agcaagcctg	gcagcccca	gaccttcctc	tgcacaccgt	ccatttttac	ttggtttcat	39900
tttgataact	gtgcggtctg	aggtcctggc	caagaaagca	tcacctggca	agaagtgtat	39960
ttggccaatg	gtaaggttac	catctctgtg	taattaggct	ccgtaaagct	ttgtttttaa	40020
atttattaat	gggaatgatt	tgacattcct	acacactgac	attaccctca	tggaatggat	40080
aagaatctca	aggcttgttg	ggtgaaagaa	gggcagtggt	tggtgtgacg	ggaagggaaa	40140

gtataagcag gcagctcgtg cgcatgagca tttgggaaac agaacagaaa tcatagaatg	40200
gcaggcttaa ttctagctct gtcacctact ggctgtttgt cattagaaaa attatattacc	40260
cttcgtgaga ttcagtttcc ttacatttaa aataaagaaa atattcgtcc tcatattgaa	40320
atgaattggg ctatcatgat caactttaaa atacaacgaa cagtataaat gtcaggaatt	40380
atatgacatt cgggacctcc actgccaccc tcactttctc cctccagtgg tcacttactc	40440
tctgtccctc tttctgggtc agagcttctg ttactccagc ctgggcctgc cttaaagtggg	40500
gacatgtgct gatccctcac aatgccgggt gacaaggagg gttttcaagg ctggcttgac	40560
tgccactgct ggtctctctt ctcatattgca actgtcttct ccccgctggg ctcagtgttc	40620
ctgggagggt gatgctgagg gagaaagctg tggcagaggg acgtggcagg gtcagagacc	40680
actgattcgc aggagctggc ctcagagcta gcctttttgc attgatctag ggaaccagtg	40740
atcatagata tctatgttga cgctgtgtc aatttatccc tagagccatt attcagtga	40800
tttcctaagg ggaaaacaat tctccagcat tatttttctt tgggccagga gagcactttc	40860
tctgagtttt actggcaagc tagatatatt cttgaaaggc tccagcagca gagttcccgt	40920
cttgtttagt tcaaaaacag ctcttgccc gtctctaaat ggtcttgcta aaaacatctc	40980
ctcccacca tagacctgaa cttaaagctc agactgctat cctctctttc tgccactgtg	41040
agagacctat gcctcttttc tcattggctt ttgcctgccc agccctctc agatccttgt	41100
acatccctag gaagtacatt ctttccactc cacacacata attgtgcaac ttgtgctagg	41160
accgcatgag gcaactggagc tgcagatagc aaggaaatgt gggccctgcc ctcaggaact	41220
tgaaatccgg tagttaccta gagctacact gagttccctg acgtggtagg aagccctccc	41280
agagccttgt ccgtgcttag acgttgctt cacagagggtg gctaaggggc attttgtccc	41340
tgccctagtt tttacaagtc ccctggatgt taactcctac ttgcttttat ttgcaggttc	41400
tccagtctta tgcattctct cttattccta aaaatttcca atccagtgtg gatagtatgt	41460
tagagctgta gggccaggaa accctgctgg gggaatcatc gtctagtagg tggagtgtga	41520
gagagaggag actcaggcca gaggggcttc tgagttctgg gcaagtccct taccatccta	41580
gtgcagctat ctttctccat gtcctactga gctgccttcc ttcttgcttc tcattcccaa	41640
agagaggcaa ccatgccatt ctgggcagtg ggcaagggt gctggaattg agatttttat	41700
ttttctcttt ggcattctgaa tccttccttt gcagttgctg acatgcagtt ttgtgtgaga	41760

tcacccatgt catccatcat cttcaggaca cccagaaact cctcctacct ctctcagcct	41820
cacatgcgca tcttcagtcc ctttctgatt caccctagtt ggtctattgc tcttggctct	41880
gaataaatac ctttaaaggg tacaggtatc catatggaag gccttcagag aagagagagc	41940
atggaattta ttttccaaag tgggacactc gagagtgaag ggggtgagctg acaataatga	42000
tggcaggaca aaagcatatt tcaggcttct taagagcttc atccccttga tccagtgact	42060
ttatttctag gaatctttta tggagtacct accatgagcc aggcaccatt gcaggtgcta	42120
aggatatcaa aacacatcaa aagagagcat acagttaagt ggagcagaca ggcaagaaaa	42180
aggatgggtg taaaagaatg cggtaatgac tgagaagtct ttctcagaga cgtgatgtat	42240
gagcaggctc agacagtcaa atagtagtta gccagggaag gaaagggaga aattgccttc	42300
ctggcacagg gagttgtcag gggagccctg gcagaggtga gagtatacat ggtgtgtttg	42360
gaaaaccag gtccctcgct gaggctgctg tgcacaggga gaggagggct ggcaagagat	42420
gaggctggca agtgaggcac ccaggagcca gacgatgaag ggagacaaat gctgaggaat	42480
tttagcttga tcttggaggt tatagagaag aacagaaatt tctccagctg aggaatgacc	42540
caattccatt ccagtgcggg gaagaaatga tagggaaggc tgggagaggg caaaataggg	42600
tctcggaac catttaagag tctattgaag taatccaggc gaaagaggta gcatggactg	42660
tagagaaaag gagggtagat tttagaaaat aggaggaggc agagtcaact ggaattatta	42720
aagatcagtt caatgtgggg gcagggaaga aaggaagaaa gaggactcaa acatgatttc	42780
tgggtgctgcc ttttcccaag agagaaaatc cgagaggctg gaggaggagc agggtttggg	42840
ggaggatggt gacttcagct ttggatatag cacccttgag attcctctgc agttcaggtg	42900
gagatgtcta gtaggcatg gatgtcccat ggatgtctat gggtttggag tataggacgg	42960
aggtctgtgc ctggagatgt gagcatagac gtatctacac atcagcaggg ttgatgagat	43020
tgcacagtca ttcctcaga tgagaaggga aagcagaatg ttggagaaga atatggaaag	43080
aatgcaaggg aagtgtggt ggtggtagga ggaagaccaa gagagggtgc gtctggagga	43140
gggatcggag tgtccagtga ttcagagagc atgaggaggg aggggctgtg gcacacctgc	43200
ctgactttga aatgaagaat ggtattctag gtggggctgt ttttttaggg acctggtgca	43260
ggtggggatg atggacatag ggaattctac tctctgggag acagcatgtg gctgggatca	43320
gatagttaca aagttatatg acaactatct tatttcattt ctctaatact ggcaacatct	43380

gaacattatc tacagaggaa ggaaaacctt tatggagggg taggttgaaa atatgtctgg	43440
gccaggcatg gtggctcaca cctgtaatcc cagcactttg ggaggctatg gtgggaggat	43500
tgcttgagcc caggagtttg agaccagcct gtgcaacatg gcaaaacccc atctctacaa	43560
aaaatacaaa aaaatagctg ggcttggtgt gcattcctgt agtcctagct acttggaagg	43620
ctgaggtggg aggcatcacc tgagcccaag aggtcggggc tgcagtgagc tgtgattatg	43680
cccctgcact ctagactgag tgacagaatg agactctgcc tcaaaaaaga aaaaagaaga	43740
aagagagaga gagagagaga gagaaggaaa gaaagaaaga aaaagagaaa gaaagaaaga	43800
aagagagaaa gagagaaaaa gaaagagaaa agagaagaaa agacaagaca agaaaagaaa	43860
agagaaaata cacctgggat ggaggccctg atggaagaag gtttctgagg aaatgggtgt	43920
gggtcccaga gtgtcagtat agggcctgtt ggagagcgag tatgataagg tcagccctgc	43980
cactcagaag aaattctggt tatactcaca gtaatcacac ttcctgggtg acccagggca	44040
gtcctaccgt gcaaccatta cccaacaaa attattgctt acacccttt tcacttttga	44100
aagtttaggt gataaagtat ctggtctcct cagagtatac atctaaggag attcgtaggt	44160
ggagttagac agctgagggg gctcctacct gaaggactct gtgcagcagg aggcaggatt	44220
ctctgtatga tgttccaaga ggcttgagga gataggagca gctgtgaggg gagggaagac	44280
ctgaaatgag aaaaagagca aattgttgat ctacagtaag ggctgagcag agcagatcct	44340
tgagaacatc aatttgctgg gcaccaagta gttcataggc cgcgtttccc tgtcttttagc	44400
agcctgcgtg tctgctggga gaatagatgt ataaactgat ctgaggttgg agctgcacag	44460
ggcagatgtg gccagagggt aaggaagtgg ctggagtggg ggatcatgag aatagcacac	44520
ccgtagtaag ctctgcacat ggagatactt attgcagcac tgtttatggg agcagaggat	44580
tcaaaatagt aaatgtccaa caagggggtg tcatccggtg aattatggca aattaacgtg	44640
atagaatatt atgtggccat caagattttg tttaggtaga attgtaatat aaggagagaa	44700
agcctagatt gtaatagtaa gtgaaataag aaagttccag aatgatgtat attaaaccat	44760
cccgctata atgcatagag atttttttat ttatttttta aaggttgggc acaaatttc	44820
cataataata cctgtgggag ctgatagatg atatatTTTT gtatttttaa atttttcttc	44880
aataagttca ttatattact ttataataa ggaaaaaaca tctctctagg aaaaattatt	44940
taaggaaaaa ataaacatga aaaagggtgt ctgtaatcac agtttaattt ggccttctca	45000

gaaaaccatt gtagggagtt tttatttctct agtttcccca ggggtggggaa gatgagcagc	45060
atggccctag tgtttgagaa cgtggggtttt gatatcagac atgcctgggt ttgaatccca	45120
gctccgctaa gctgagctat agtattgcca tgggtgtgac tcaagcaggc tatccttatt	45180
tatgaagtgg aataatacac atccccagtc tcttctttat aattccgaaa ccaaaaaagc	45240
cctgaaaacc cagagtcgtt tcttaaagtg acagcaaatt catttggtta taaaacctaa	45300
cctgaattga gctagtcttt atttcacttc agtgaatata attttgtagt ggaaatatta	45360
acatgtttta tcattggatg ttcgcccaga tctcacggag ggtatgagat aatacgagg	45420
gttcaggtaa gttgtgaact acaggatctg cgtcataggg ttattgtaga atgaagtcag	45480
gttgtgattg cagggttctt agcacagttc ttgccagaga gtaagtactg acaaagttag	45540
ttatcatcac tattgatgct ataattattc gtgagccttc actaagtatg gtagtcatga	45600
tttctgcttg cccctcttcc ctccatttcc ttctttattg ggatgtttga gaagattaat	45660
gtatcacata taatgagaat ttaggatggc attgtttagg acacaggact gcaaaaaagg	45720
gaagttccag cacagatatt ttcctgcctc ttttcttatt ccagatttct cactttctgg	45780
ggaattagct gtaggatata atgcacatta acaggacacc caaatacttg aagagcta	45840
ggagaagtaa cccttactcc cttggtgact gctttcgtct caaagcacat gtgttcatta	45900
taaaatgta atgtatttac accatagtgt ctgtacattt aaagggttac tgtgctttta	45960
gctaaaacta gctggatata aaatgtgctt ttaattgaaa agtgggttatt tactgctttc	46020
cataggaatc cacaataaga ttttttcatt gcaaaccctc ctagtatctt taaaatgcaa	46080
ttcagttcac aaacattgat ttgcattcaa cgtttttcag aagcctaaat agagtgcata	46140
atgaggtcca cctgcaagac tacagttatt acttcctcct ttttcttgga agcatccatg	46200
cattgtaaat tctgtttgtg tggcatcagc caactacatt attaacatca ccaggaggaa	46260
ttaaaatgta tcatgttata ggatcaaaag tttcacattc tgaatcaatg ccagtataaa	46320
aagaatgtcc agacccttcc gtctgactac agtcggcagg gatccccgcc ccagcagcca	46380
tgtgactcac ttcgcattgt ggtgagggtc actcatggct catacgggcc tgcggcccga	46440
taggcctccc catggggcag catttagcct tttcaaaatc agcttaggat cacagaccct	46500
caggcttga aggaactttt tggagtttcc tctattaaac ccagttttta tctggtcatt	46560
tagactgcat gtctagaaac actacagcat ttttagtagt aaactgaaat aaatcttcaa	46620

ataaacgact ttaacatttt gcacattggt ccaggggcag aagagctcgt ctgttgccctt	46680
ttaggagggc tctaattcttt ttctgtttca gctcatcggg tcctcattgt cccctactgt	46740
atgtgcccag cacaatgtca gggcacattt ctacacaatg ttggattcca gctgatcctc	46800
cgtgttaatc cattcgtttag tgccaaattg cagggttcct tgccaagccc gtcagcacia	46860
tcagagcttc tgaagtcagg tacctcaatt acattagttc agttgtctca ttagccccc	46920
tgcacatcc aaactcatca tcaaactgtt tttccttcat tattgttgctc attttcttaa	46980
taccagtctt gttgaaaaag gattgtttta gatacaattt tgcctaactt gtttcatctc	47040
cagaatacaa tcctaaatag agccaagatt tgccggagag aggagatagc ggttgaggct	47100
ggtgaagctt ccctccggtt agaattagag agaagtggat gcagaacttg ggccagtaat	47160
ggacccatcc atctccctag tcaagaggac attcttgtag ccagagggtg ttccgctttg	47220
gaaatgatct actttgtgta aaccggtgca tcacagatac tgtttagagta tctcacagat	47280
actgatacag catggtctgc atttttaga tatgagggtc ctcccttagg tttacagccg	47340
gataaaggag ctccacttac ctggtgatca caggcatgtg gtccaccctg ctgggctcta	47400
ggactttgaa catagagaag gctatctttt aggaagactg caggatcagg agtcgggaag	47460
cagggtattc gcttgtagtt ttacatctag cagtaacaac ttttagccca tccctcacct	47520
actctgtgcc ctggatttct tggctttaac atagcaaaag cctctcttta tagttttcct	47580
aactccatct ctccccaaaa agctggagat gtaagtttta aattccagtc cttggcacat	47640
agtaggtgct caatctctca tccttctgtc cctcaagata gccgatggct cagccatata	47700
tgttcagac aggtctctct gtgtcttttg ggaaagcctt gttgctctca gactggcctc	47760
taggcctcat gttttccac ttgctcttgc cttcagtgtt aggtcatggg cttgcctcat	47820
tacctgtttc aaggaggcct ctatttgatc cctggagcat ccctgggagt ctgaggcctt	47880
taaggactcc tgtggtctca aggaaaaaga actataagta gttctcaggc tcccttgatc	47940
aattcaaagt catcttcagc tcttcagtg cagccccctc agagctgttg aacgccacct	48000
ttttcctctt tcgggaaacc cacaaaacct tgcattgggt tgtgctctgt atggttttcc	48060
agcatcaaat tcactcatca aaagatgcat caggagggtg gaggggggtg cgcctggggg	48120
gagggtcaaa caaggaaggc cagaatgcct gtttgcattc gcaacgggaa cattggaagt	48180
tcgggtggaa aaacaatccc accggccttt atttaccaca gtgggtagcc tcgggaggag	48240

ggggagctgg	gtgggaaagg	aagacaatag	aatgctgtgt	tgcccttgctg	taaaagcttg	48300
ctctaagcaa	ataaagtggc	aagctcagtc	agagccctgc	tcaattagcc	ccagtgccta	48360
aaggagggtc	tccctgcagc	tgttggcggc	gggcgggctg	gagcaaaggc	aggcgcactc	48420
tggggcactc	gggaggcgaa	ccggcaggaa	tcttgcatgg	gagctgaccc	gggagggaga	48480
ccaaaggacc	ctccaacctg	atcccagtc	ctgcttcttg	aacagagggc	tacagaaggg	48540
gttggtgggg	ccactctggg	gagaggcagt	gtggggagac	cacggaggag	gtgaccagaa	48600
aagtgagtgc	aaacgtttca	tgccgagagt	gaccgacata	tggaacgtgt	tatccggaag	48660
ggttacaggc	tggaggagct	cctagagttt	ttcagggggg	gggggtggggg	agggggaagc	48720
cttcagaaac	tcaaggaggt	ttctaataca	aaggaatttg	agaagggttg	gagaaataat	48780
gagaaatgca	aacttgaccc	agagtttctc	cttccatctc	acggcttctc	ccctgcctgt	48840
cctcacatgt	ttccagttca	tggggaccca	ggaaggccac	tggagccctg	tgccctgactc	48900
cacgtgcacc	tcactggggg	tgtgggtggc	ggggtaggga	ggacccgcag	agctggctcg	48960
ctgctctgtg	ctgaaaggga	cccagagagc	gagagccctg	cctggcttta	gacccctgtg	49020
gactagcagg	ctgctagcgc	cgggatggtc	tctggatgat	ttactccggt	tcctgccctt	49080
gctggaggag	caatttgatg	ccggactggg	agtgaataac	agacttgccc	agggactcac	49140
agtggccaaa	ggggaggctg	ggagtagacg	caggctcttg	atcccttctg	ctttttggca	49200
tctccccttc	tcagcatccc	ttgcccagca	caccacgcac	acatgtacat	acacacgcag	49260
acacacacac	caaacagatg	cacacagaat	aacaacacca	cacctacatg	cacacataca	49320
cagacataca	tgacgcacag	acacacacag	aataacacca	cacatacatg	cacaccacag	49380
atacacagaa	taccacacac	acatgcacac	acacagacac	attcacagaa	tacatatatg	49440
cacacacaca	catacataca	tgaccacac	agaataacac	cacacacaca	tacagacaca	49500
cagacataca	catacatgca	cacatacaga	tatactcaca	gaatacacac	atgcacacac	49560
atgtatagac	atacaccaca	caagcagaat	aacaccacac	gcacacacca	cacacacaca	49620
cagaataaca	ccacacacat	atgcacacca	tacagagaca	cactcacacc	tagacatata	49680
caccacactc	atacagagac	acatgcacac	acatgcatac	tacacacatg	cacatgcaca	49740
cacaaaaaca	tacacagaca	tcacacatgc	atacatatac	acaacacaca	gacacacca	49800
cgggcttaca	cagacaccac	acacacgtag	atatgcacac	agcacacaca	cacacagaga	49860

aacagacaca	ctcgcaggca	tacatgcatg	catacacaca	ctacacacag	acacactcgc	49920
aggcatacat	gcatgcatac	acacactaca	cacagacaca	tctgcagtca	tacacaggca	49980
ctacatacat	agagacatac	aatgcacact	cacacacaga	cacacataga	cacatgttgt	50040
gtgctcatgc	gtacacgagc	gcgcacacac	acacactgcc	tgacttgttt	cccacaagac	50100
gggtactggc	ctgtcgcttg	tagccctcct	gtcccagcct	gtgttggcca	gggtgccagg	50160
cactgccacc	cctcttggga	caaggtacag	gtggccagtg	tgatcacggc	ctgttctgag	50220
agcctcttcc	tgagccagga	agcgctgtgt	gatactgagt	gcccgtgcct	tcgtctttcc	50280
catggtgctg	ggtcttggcc	acctgctgca	ttgatagcac	ccgcatgttc	acttccctgg	50340
cagtagaaag	aaatgtaggt	tagtgcaggg	aggtcactgg	ctttggagat	ggtggcaggg	50400
tggagcagtc	tctaattgta	atggaagtgc	acatgccctc	tggacctgca	gcaatgctgg	50460
cttcagaggg	gccttctctg	gtacctttcc	aattttcccc	aaccaggtag	aagagccact	50520
gcccagtgtc	ttgggctcaa	ctgaaacccc	ctgtacaaga	aagagacccc	cctttccact	50580
gtgtctctct	cctcctcccc	actcactgtc	tctctgtctc	cctcactccg	tgtctctctc	50640
ctctcttctg	tgtctctctg	cctctccctc	tctctgtttc	tctctgtatc	gctctttttt	50700
ttcaatctct	gtctctatct	ctcctccttt	ctctttatct	ctctgtctct	ccctctgtct	50760
ttctctctgt	gtgtgttttt	ttttctgtct	ttctgcatct	ctgtttctgt	ctccttctct	50820
cggttctgta	tctttgtccc	cctctctccc	tatctctgtt	tctctgtctc	agtccctttc	50880
tgtctttatg	tctctgttta	tctctctcag	tctctgtatc	tctgtctctt	tctcttttta	50940
tcctttctct	gtctctttgt	tctctatccc	tccatctctc	tctccctctc	cctcagtctc	51000
tctctgtctc	ttcctttctc	tgtctcagcc	tctctctgtc	tctgtatctc	tatctctttc	51060
tctctctccc	agtctctctc	tttctatctc	tctctctctc	tctctctctg	gcactcactc	51120
accacttac	ttgaagtctc	catgagcagt	gggtggcttc	acctttctgt	ttcgccactt	51180
tgcagtccga	cccgtggggc	ttgcagaccc	tcttctggcg	cacaccttca	ggagaccaac	51240
ggtgccaggg	cactcccgtg	ttcttcaggg	ttccagcccc	gagtagttgg	taaacatcag	51300
taatcgtcct	agagatccac	tgtagattcc	tcattccaggt	actgaatgag	gcccttctga	51360
gcaaatttaa	tggaatgacc	ttggtgacat	tacaagatga	cggctcatct	cctgtagtct	51420
attattgttg	ggcatttagg	ttgatactat	gtcttcacta	ttgtgagtag	tgctgcagtg	51480

aacatatatg tgcatatgtg tctttatgaa agaatgattt atagtccttt gggatatatac	51540
ccagtaatgg gattgctggg ttgagtggta tttctatttt taggtccttg aggaattgct	51600
taggtccttg aggaatcgca tttgcgtttt caaacatcat ggagaacaca ggtcttgagg	51660
atgtgcagga acatggagac caaggaatga aaaagccatc aggttggaag tgaagtcacc	51720
cagagtgatt gcaggtttat ggacagaaag gaaaatgctc tgattcctgg tgtccatgag	51780
gaaggtaggat gccaatagat ggtgctacac agacacttgt gaaaggaaga agggaaagag	51840
tctgcctcca tgggctgagt agggctgctg gggcctcagg cttcacacat agtgtcagac	51900
atgctgtgtt ggtcaggttc atctgggacc cctgtgcctc cccagggaaa ccctattggt	51960
aggtgtgggc aggcttcctt ccatgtgtcc ttaaatacac tgactgctct gtgtgtgact	52020
gtggatgagc agcgtagggc tggatcatctg gagggctggg cttgggtcca gctcttgtag	52080
tgatggattg attggttttg ggcaagtcac ttcccccat aacctttcca ttgtaccaca	52140
atcctaatag agttggatgc cagtttcttg caccctgggc ttcacaggga cacagagaaa	52200
cagagaggac cagaaaatta ggctgataat aatcatttct tctctatttg gtctggagaa	52260
gaaatgactt agggttggca caggttaata tgtcaactag aaagaacctt tgaaaatctt	52320
tagttcaaga attctaaacc taaaatctgt gggacactag gatgattttt atggactctt	52380
tgaatcacc tgaattttt taaaatattt tttctttatg ggcacatgtg tggtttttta	52440
gcagaaggat tccttagcat tgattacatt gtaaaagggc caacaatcat aaaacaatta	52500
attccccagc atcatgttat gtgattcata actgaggggt ctccgcagat ggcatattgga	52560
aacgtgcggg tccttttttt cagttgtcag catgactggg gtcacgatgg atatttagta	52620
ggtggggggc tgcgttgcat gcgtggggga ttaaaactta caataaaaaa ctgtcctgtt	52680
tgaaatccca gtagcacccc tctggagaaa cactgctcag gaactgagcc cccaagatgc	52740
acagtgattc tcccgtggcc acaaagctca ttagtggcag agttggtttt tattaacaga	52800
tcaaaacagg ggatgtgcca agaacctaca ttattttgtt tctcatcatc agctgcaa	52860
gtgctgcaat ctgtgaaaac aaaagaatca gaaatctgtc ccaccttcaa cgctccagcc	52920
atttttaaaa atgagtcttt gggttcttgg ttgtttttta ccatcaaatg agaaatgagg	52980
aaaggaatat ttacatttg gagaaactaa accataaaca attgatcca ctagcccat	53040
atcacctaga tgtgttcttc agctactgtg aattggtgac gcaagcctta gactggaaat	53100

ttcccccttta	tgttttcagag	gatctgtctt	gaatgtctct	tactctacaa	agaaagaaac	53160
attaatatac	cagcatgctg	cttgctacct	agtttactgc	aatgaagtgg	caggtgcctt	53220
agactttgga	gtgaaattga	gagatattcc	gcagtattag	ctaagagtta	ggcctgtggg	53280
atcacagaga	cagggggttg	ttccctggct	ctgtctctca	attgatgtgt	cttcttaaga	53340
aaaagtactt	aatccttctg	agcctcagtt	tccttggctg	aaaagtgggg	ataataatcg	53400
tatccacaaa	gattaattga	gatcatccat	gtgaagtgtt	ccgcacagtc	tagcatatgg	53460
ggaggctcaa	taaatgtgag	ctgttattac	aagcgattat	tatgacttgt	gtctgtgatt	53520
aaagacagcc	tgaggttcag	agtgtttgtg	ccaaaattgc	tttcagagga	caagctctgg	53580
gttggttttt	cagtcttgcc	cagctgccc	gagtcactag	agtgtttggg	gctggagtca	53640
cagtgtttat	taagtgcc	ccacacactc	cagttttgat	ggaagattgg	ttttcttcct	53700
gtgtgcacat	ccccgcatca	ctacatgagt	gggtgagaag	agactcagaa	aagcaccaga	53760
cacatgtttc	ttctgcctca	gttccctagt	ctgtaaaagt	actcaagagc	atgatggcgg	53820
ccacagggaa	cagaggggag	agacacctgg	aggagcaaaa	gacagtccca	gctgtcagtc	53880
ttcagacacc	atggttttct	ggagaaggga	tccagaccag	accgccaatc	aatatagccg	53940
cagaccgcca	gtcaatatac	ctcatccttg	tgaaggcg	ttgtctgtgt	ggcagcaggg	54000
agacaggagg	ggccatattt	gacacaagcc	tgctggccaa	gtctaaagga	gtcagcctgc	54060
catctgacca	cactttctgc	agccaagtcc	tctgggggcc	aaatcagtat	ggtttgagtt	54120
tatccccgac	ttgctgctat	ggttgatgca	cacaaaaaag	gtttggccat	tgccaagcca	54180
gctgtggtct	tggtttggaa	gaggcaacca	aagtaagaat	tgaagggaag	gcatgctttg	54240
ggatttcct	tagccttcaa	ccctctagag	gaagccaact	tctttattct	gccattttga	54300
gactcatgct	tgtctgggtg	gataagataa	atactttcat	cattatcatc	attattatta	54360
attataatta	ttgaaaatca	attactcagc	agcccctacc	ttgtaatgg	atcacttaca	54420
gttgatatagc	atgttgacga	caacgtatac	actagaccat	aaactcagct	cgaacataag	54480
cataaattta	aaacaaaaac	attgacttgg	gcagctatta	tctgagggtca	aagttattct	54540
ttaattccag	ctgtaccttc	agctgcctat	gtcgtgaag	caacttcccc	ttgtgtgttt	54600
tcacacttag	agattggagt	aatccagaca	cagaaaatga	tacccaaaatt	gaaatgtttc	54660
tgaacgttaa	atgattcaat	aatttgaaaa	tagtttcaaa	tttttcaaaa	ttcaaaaatt	54720

tgaaaatagt	tcgactagag	cccatcctca	cctcaggtag	ttacgtggtg	tgtgtgtgtg	54780
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtagg	cacacaaacc	catgcatgca	cacatgtatg	54840
cacagtgggtg	gatgagtgtg	agttaccaa	acaaatacca	ctaaatgcag	gataacacca	54900
tgatggctag	tagctggatt	gactcagatt	ttctctgcag	atccttccta	gtccaccctc	54960
cttccatgac	atgggagttt	gtgatatgca	aaggataaag	taaacctgag	tttcctgggg	55020
cttgtgtcgg	cataccaagc	tcctttaaga	tctctctttg	ccaagatggg	gctaagccag	55080
aggcttcctt	ggcctgggga	ataactgtcc	taccttcctc	cctggagcgg	ctggcgcaca	55140
ctgttttgac	gtgtgccatg	ccttccatct	tttacggcct	cggggccaga	ggagggtatg	55200
ctttgaacaa	tgtgaaattc	ctctcctaac	agctgtgcaa	aggaaactca	cagctttcta	55260
tctccatctc	ggtccactta	gctttctttt	ggggtatagg	tttcttgtct	gttcaggggt	55320
ccctgttcct	cactttctgg	ccttcagaag	gaccagtttc	agtacttcct	ttcttaggga	55380
ggcaagtcag	tcttctgagg	ttctctgttt	tattttattt	tatcctagag	tatgccttca	55440
ctgttgaccg	tgtcttgcct	catttctgct	aagaagccag	catcttctga	aaattgagtg	55500
tcttgtcctg	tttcagatct	tgggccgagg	attgaccatg	tatgttagtc	agtttcagag	55560
ccctgagttt	taagtgctaa	gttttaagga	gaggaacaac	agtagaatta	gatatcttaa	55620
agaacaggat	aggaatccac	ttccctgccc	cttctctcac	ctggtttgac	actacataga	55680
tacagacctg	tcctgagttc	agggattgtc	taatgaggcc	tcaccaaata	ccaacagaga	55740
gactagctgt	ttctcctcta	gctcactgac	tgcttctagt	tttccgtgac	tttagagcat	55800
gtgcgtcagc	aggttttggg	gaccatctct	ggcctgcttt	ctgatcttgg	ctctgtccct	55860
caatttatgt	cttttttgtg	ctccataaga	cagggagttt	catggttgcc	actacccgag	55920
agggtgtgcta	aagtgaatga	ggtggaggac	acagagctag	caggggaagaa	ggtgccgtgg	55980
aaaccacac	agctttattc	atgtgagctc	atctctcatt	tactccttgt	agcaatgcca	56040
agtcaccccc	tccttcatc	cctttccaag	catttcctc	ccttgttggg	tttgatcca	56100
ttgtaacccc	agtgagttaa	cctacatgtt	ttagtttgaa	tctaatacatg	ctatttccag	56160
gttgccaacc	cccacgggtg	cttcagccta	attggctctgt	cctcactggg	gtctgcagca	56220
ccttcccttt	gagcgtcatc	tgtgcatttt	aattaatctg	ccactttttc	ctccctcctg	56280
cttccggatc	attaataaag	agattaacca	aaaccagacc	aatccccagc	ggaggctcca	56340

ccaggatctc tctctcccag ctctcagctt tgtcatgggt cagaacatca tcatgtggcc	56400
ccattgacta attctgaggc ctcaagacct atgttcagag caaagccagt gtgaattaat	56460
ttttcaatta agatttgggg aacaagcata gccagtcttt tatatatggc cttagttaa	56520
tatttccacc acattccttt cccttttgta attctacttt aaatctatca ctttcccacc	56580
tctctccctt tccagtccct gaacaagaca ttagggttat ctggcaggat gtctcctgag	56640
cctgtagcat tagtcagaga tcatgtcata gccgaagaaa ttctccaaag aatcaggccc	56700
ttctggacac agctggccag cgtcctcttt gaccccagcc agagaatgca ttgacctca	56760
gagaatatgg ctacaccaga aataatgaaa tctcttgaat ttgcatatta aaaatttggc	56820
aaattaaaat ttgctttcac ctctgctctc atgactcctc tctgaggaaa gctaaaacag	56880
agcctttctg ctttctttgt aagtcacca gatttccagc gctgggtctca aaccaggat	56940
gtgagatggg gcactccgtg ttctgtccac tacaacacag ggctctgca ggcacagggtg	57000
acctgcagat gggctgcact ggagagcatc aggcccagga gggactaggg tgagggaatt	57060
ggattgtgtt tacttttgtg ggggatgctt tcatagggaa tacaggaggagg aggacacagg	57120
attgatggcc tcaaactctt gcaaactctg cttagagctg aggcttcttg gggaaggaca	57180
cagcctgctg acagcagccg tcctggctcc aggcttggag cttgggcttc ccctggaagt	57240
aaactctcta gaatgtttga gagcaaggac agacggagag gcagccgagt gtcattcata	57300
ataacatgag ctctagagcc agaaggtgca aatcctgact ctgttgcttg ttagctgggt	57360
gtggaacccc tctttgcctt ggtttccca ttcaaatgg gaaatataat agtacgtact	57420
tcaggatggg tgttactagg atttaagatt tatgtaagcc agtacctggc actttgcatt	57480
tttattcttg gtgatctcac ccctaagcag tctaaaaatt tacttcacac atgtgcagca	57540
agaggcgtct ttaggtacac ggtcaggagg tgctcctgtg tagggaagct ggtcctggag	57600
aggatgaattg gaacctgaag gctgcagccc agtctctgca cgtgacaggg cagacgcagg	57660
ttgctggggg ctcagattct gtgccactgt ctacctcagc ttccttttcc atgtcagctg	57720
cttcttgaat acccccaaac cctctgtgga tgctgtaaac tgagccaaaa aggaggaatc	57780
agaaattgct ctcttttgag tattttgagg gatggtgaaa cttgtagga atgttgctcc	57840
agggactgag aaaacacat atgtgggaaa gcgtttctaa gtgaggaaaa tcgcacatgc	57900
acgtcgggtc ggatcttga aaactctgag caggggaaat gggcactggg gagagtgtag	57960

aaagggggtt	cactaaggtc	tttaacaggt	ctttttctaa	ttatagccaa	gaagagcctg	58020
gggaccatgg	ccgacttcag	agccagggcc	tgtattttct	ttgggaaaag	aaggcagagt	58080
tgattggctt	ccaaaaccag	gcttggcaaa	agcccatgca	ggctctccag	ccagtaagca	58140
gtgcctgggg	taggtctcga	tcgtagcagg	cgctcattca	gtaaatagac	gtcgaaagtc	58200
cgcagcgtgt	ccagccctat	gctgggagtc	aaggatgtgt	ggatggaaaag	gacatgggtcc	58260
ctgtcgtcaa	agagaaggca	caccaacaat	aatagcacia	tatgacaatt	gcttaaataga	58320
agctctctgc	agatgacagg	gacccctaga	gaaagagcgt	agacattccc	ctaaatagca	58380
gaactcgctg	ctgtgccttg	cctggatgga	gctggcactg	ccttcctggt	tccattcccc	58440
actccaggga	actggccac	tgggggactc	ccaccccgga	catggagtga	gaaatgaagg	58500
tattggtggc	caaaaagtga	ccagagatca	gaatgcatct	attcttgccc	ccagcaaatt	58560
cagaaacacc	acctctccct	tgcccgtggt	tgctttcgtc	atttgtctgt	tctctgggct	58620
tcgagttcat	aatattccta	tcttctccca	caaccagtgt	aactggctctg	agatccaagc	58680
agtcgggcac	catggcatct	gggggaaagc	cttgtgggga	ccgtgttggt	aaagattgct	58740
gtgagccaga	tgtagaggag	ggagctctcc	agactctggg	tccctcgccc	gtgtgctgca	58800
agggcaggtg	tgcaccatct	cccgtggcca	caggccattc	agcttcatgt	tctctaakat	58860
ttctagtgtg	ctcactctgt	acctgatgct	atgctaggtg	cttccatgtg	gatttcattg	58920
tttaccctc	acgataactt	tcattggtgca	caagccatgt	gatttttggt	gcgcctacga	58980
accaaagtcc	cccgtcactc	tctgccattg	tttagcagtc	tgaccttaga	cagctcactt	59040
caccacaact	gcaggcctcc	gttcccttat	atgtataaat	gaggagggtg	tctagggtcc	59100
tttccagctg	caaacctctg	tgttttgggg	agttctgtta	gcaacattct	tggacttgct	59160
ctctacgaga	aaagctagaa	gtcgattatt	aaaatggagt	ccacaggaag	cctgcagggc	59220
tttaagacag	ccttctaaag	agtttagaaa	cactcaccag	cctgcagatg	acttactccc	59280
actacttatt	acaaagtatc	atggcaagta	agaaagaaat	ggtcagccac	tgggagagca	59340
ggctgagggt	gcaaggaagt	ggcaggtgct	atcgttggga	tcttggggga	gcattgtaga	59400
aaggtcagct	ttggaaccag	gcactgaaaa	acgggtgaaa	tttcaacagg	agaggcagtg	59460
cacatgggtt	gtgatttgag	cagcagtgag	tagtgaatgg	ggttggagca	agaggcctgt	59520
gagcccagtc	atcacatgga	aaaggaggt	gagggccaga	gtgcagagaa	cccctctcac	59580

tttcctgtgg gtctctcctc ttcttttttt tttttttttt ttttgaattg gagtcttgct	59640
ctgtcgccca ggctggagtg cagtggcgtg atgttggctc actgtaagct ccgcctcccg	59700
ggttcacacc attctcctgc ctcagcctcc cgagtagctg ggactacagg ctgccaccgt	59760
gccctgctaa tttttgtatt tttttttaga cggggtttca ccttgtagc caggatggtc	59820
tcgatctcct gacctcgtga tccaccacc tcggcctccc aaagtgcctg gattacaggc	59880
ataagccacc gtaccgggc tttctcctct tctttaaact gcatttaggt tgtcctcatc	59940
tataagatga aggaataaga ttagatggcc tctttatatg gcttatctac ttctaataaa	60000
ctttggttcc atattccagc caccacccat aaagactctt actttttcct cctcaatcca	60060
tcagcatcag ccagcattct cctctcagt tcatcctggc cattacctca ctaatctgta	60120
ccttgactct tattctacct gggacctccc ctactaacc atacttgatt tatttcctta	60180
agagcggata tcctgtgtat tcttagctta cagtggctga gcttttggga ggtttaacaa	60240
attgtttttc aatagactct cagtcctaaa tgattttatt gcagctgtat tattcttttt	60300
gaggggagta gagaatgtaa tcctagtgtg atgccatatt ttttgtattg ttcaagttgc	60360
attttgttgg gcctgtaata ttgaagaaaa tgtcacttgt atgcagaata ggaggagat	60420
tccttatgct gtagaggaa ctttttcta ggatagtaca attccttaga gtcattctta	60480
agggtgagca gcaaagctag aatccctctc ttggctttga ccttgaagga gtcagttttc	60540
ccaaagacta gttcccatgg aaagaagatg gtctccttat cacagcagca agaggaggca	60600
gtatgagcag agtgcctcac tgttgttgac ctctagggac aagtgagccg gcagtatttc	60660
agacaagctg ggaaagaggc tgtctgtgag tcctgggagc gagtgagcac tggcaggctc	60720
agacataggt gctggtagc aggactgctt ttctgtttct tgtgtcggct ttgtttattt	60780
cctctctatt ttccccctgg acttagtaaa gtctttccga aaataccaaa ggtgaaccag	60840
gggaagagtt tttattttcc atgtttggac agaactttaa agaggaaatg atgtaccccc	60900
ctgggagcca gtgaggtggc agcgatggtg attaaggagt gaatatctca aggaggtgga	60960
cgaattcggg gatcactagc tcagctgccc ctctccacct ggagcatctc cttccagtgc	61020
taccctcaga acatctgggc tttgctctag tgagggagag actagcaatg aaggtgtctt	61080
gagatcagca ctgtaattcc accaggacgc caccggagtc cgggtgttaag cttctactat	61140
ggcaacagaa tgagagcgtg gatgggttga aatgccattt caacaaggaa atagtagaat	61200

tagaagaagc	ataagcacta	aaaggaacat	tttgtagcag	aaatgttaaa	aatctaaagg	61260
aggcaagtga	atcaacaatg	actcttctat	tctgggcaag	tccagcccat	ttatgtaagg	61320
tggttattct	gcatctctgt	cttctgcaag	tagtgctgtg	gcagagctgc	gttttgtgga	61380
gagcgtcccc	ggggatggag	cagatcagtt	ggtgatgcgt	atgtatcaga	aagctcggca	61440
gagcaccctg	gaacgtaggc	cctctcgcgg	agtgggtagt	ggccctacat	gttcatttcc	61500
aagggcagga	gaatagaccg	ttccagctgc	ggcctggcca	gggatgaccc	cacatctgac	61560
actgcaatat	gggggcaact	gaaccagtcc	tcagcctcag	tgtgttccag	gggctgcagc	61620
tggggagcag	tcgaactctt	tcttgagaca	attacaaggc	caccgctgct	gctgctgaag	61680
ggaagttact	ccatgtttac	aattctcagg	tttgaagttt	tcatgctttg	ccaaggtaga	61740
gtgaaccatg	cgcttttgca	ggctcaaggg	atgtttaaag	aagcggtagg	acatcgtcca	61800
cccacaagca	gagaccgcag	gataaagcag	acatccaatg	taaatacaac	ccgtgcaaaa	61860
agcagagtcg	gcagacctgg	agtgcatctg	cagtatctcc	cgggggtggg	ggaaagaaat	61920
cacctcttca	gaatgtccag	aggggagttg	ccttgcttac	ctggggggcg	gtaccctctc	61980
tcgtgccctc	acagggctac	tcagcctcag	gtagctgggtg	ccagaataac	acagactcag	62040
ctgccagagc	ctgctcttaa	cacctgtgtt	tccttttcag	atcttacagg	tgaacaaggt	62100
gatgtccatc	ttgttttatg	tgatatttct	cgcttatctc	cgtggcatcc	aaggtaaaca	62160
catggatcaa	aggagtttgc	cagaagactc	gctcaattcc	ctcattatta	agctgatcca	62220
ggcagatatt	ttgaaaaaca	agctctccaa	gcagatgggtg	gacgttaagg	aaaattacca	62280
gagcaccctg	cccaaagctg	aggctccccg	agagccggag	cggggagggc	ccgccaagtc	62340
agcattccag	ccggtgattg	caatggacac	cgaactgctg	cgacaacaga	gacgctacaa	62400
ctcaccgcgg	gtcctgctga	gcgacagcac	ccccttgag	cccccgccct	tgtatctcat	62460
ggaggattac	gtgggcagcc	ccgtggtggc	gaacagaaca	tcacggcgga	aacggtagc	62520
ggagcataag	agtcaccgag	gggagtactc	ggtatgtgac	agtgagagtc	tgtgggtgac	62580
cgacaagtca	tcggccatcg	acattcgggg	acaccaggtc	acggtgctgg	gggagatcaa	62640
aacgggcaac	tctcccgtca	aacaatattt	ttatgaaacg	cgatgtaagg	aagccaggcc	62700
ggtcaaaaac	ggttgcaggg	gtattgatga	taaacactgg	aactctcagt	gcaaaacatc	62760
ccaaacctac	gtccgagcac	tgacttcaga	gaacaataaa	ctcgtgggct	ggcgggtggat	62820

acggatagac acgtcctgtg tgtgtgcctt gtcgagaaaa atcggaagaa catgaattgg 62880
catctctccc catatataaa ttattacttt aaattatatg atatgcatgt agcatataaa 62940
tgtttatatt gtttttatat attataagtt gacctttatt tattaaactt cagcaaccct 63000
acagtatata agcttttttc tcaataaaat cagtgtgctt gccttccctc aggcctctcc 63060
catctgttaa aacttgtttt gtgatccggc tctcaggagt cactctgtaa aatctgtgta 63120
caccagtatt ttgcattcag tattgtcaag gccatgactg ttgttttagt aaacttgta 63180
aatca 63186

<210> 31
<211> 255
<212> PRT
<213> Homo sapiens

<400> 31

Met Phe His Gln Val Arg Arg Val Met Thr Ile Leu Phe Leu Thr Met
1 5 10 15

Val Ile Ser Tyr Phe Gly Cys Met Lys Ala Ala Pro Met Lys Glu Ala
20 25 30

Asn Ile Arg Gly Gln Gly Gly Leu Ala Tyr Pro Gly Val Arg Thr His
35 40 45

Gly Thr Leu Glu Ser Val Asn Gly Pro Lys Ala Gly Ser Arg Gly Leu
50 55 60

Thr Ser Leu Ala Asp Thr Phe Glu His Val Ile Glu Glu Leu Leu Asp
65 70 75 80

Glu Asp Gln Lys Val Arg Pro Asn Glu Glu Asn Asn Lys Asp Ala Asp
85 90 95

Leu Tyr Thr Ser Arg Val Met Leu Ser Ser Gln Val Pro Leu Glu Pro
100 105 110

Pro Leu Leu Phe Leu Leu Glu Glu Tyr Lys Asn Tyr Leu Asp Ala Ala
115 120 125

Asn Met Ser Met Arg Val Arg Arg His Ser Asp Pro Ala Arg Arg Gly
130 135 140

Glu Leu Ser Val Cys Asp Ser Ile Ser Glu Trp Val Thr Ala Ala Asp
145 150 155 160

Lys Lys Thr Ala Val Asp Met Ser Gly Gly Thr Val Thr Val Leu Glu
165 170 175

Lys Val Pro Val Ser Lys Gly Gln Leu Lys Gln Tyr Phe Tyr Glu Thr
180 185 190

Lys Cys Asn Pro Met Gly Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp
195 200 205

Lys Arg His Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
210 215 220

Ala Leu Thr Met Asp Ser Lys Lys Arg Ile Gly Trp Arg Phe Ile Arg
225 230 235 240

Ile Asp Thr Ser Cys Val Cys Thr Leu Thr Ile Lys Arg Gly Arg
245 250 255

<210> 32
<211> 249
<212> PRT
<213> Mus musculus

<400> 32

Met Thr Ile Leu Phe Leu Thr Met Val Ile Ser Tyr Phe Gly Cys Met
1 5 10 15

Lys Ala Ala Pro Met Lys Glu Val Asn Val His Gly Gln Gly Asn Leu
20 25 30

Ala Tyr Pro Gly Val Arg Thr His Gly Thr Leu Glu Ser Val Asn Gly
35 40 45

Pro Arg Ala Gly Ser Arg Gly Leu Thr Thr Thr Ser Leu Ala Asp Thr
50 55 60

Phe Glu His Val Ile Glu Glu Leu Leu Asp Glu Asp Gln Lys Val Arg
65 70 75 80

Pro Asn Glu Glu Asn His Lys Asp Ala Asp Leu Tyr Thr Ser Arg Val
85 90 95

Met Leu Ser Ser Gln Val Pro Leu Glu Pro Pro Leu Leu Phe Leu Leu
100 105 110

Glu Glu Tyr Lys Asn Tyr Leu Asp Ala Ala Asn Met Ser Met Arg Val
115 120 125

Arg Arg His Ser Asp Pro Ala Arg Arg Gly Glu Leu Ser Val Cys Asp
130 135 140

Ser Ile Ser Glu Trp Val Thr Ala Ala Asp Lys Lys Thr Ala Val Asp
145 150 155 160

Met Ser Gly Gly Thr Val Thr Val Leu Glu Lys Val Pro Val Ser Lys
165 170 175

Gly Gln Leu Lys Gln Tyr Phe Tyr Glu Thr Lys Cys Asn Pro Met Gly
180 185 190

Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp Lys Arg His Trp Asn Ser
195 200 205

Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg Ala Leu Thr Met Asp Ser
210 215 220

Lys Lys Arg Ile Gly Trp Arg Phe Ile Arg Ile Asp Thr Ser Cys Val
225 230 235 240

Cys Thr Leu Thr Ile Lys Arg Gly Arg
245

<210> 33

<211> 249

<212> PRT

<213> Rattus norvegicus

<400> 33

Met Thr Ile Leu Phe Leu Thr Met Val Ile Ser Tyr Phe Gly Cys Met
1 5 10 15

Lys Ala Ala Pro Met Lys Glu Ala Asn Val His Gly Gln Gly Asn Leu
20 25 30

Ala Tyr Pro Ala Val Arg Thr His Gly Thr Leu Glu Ser Val Asn Gly
35 40 45

Pro Arg Ala Gly Ser Arg Gly Leu Thr Thr Thr Ser Leu Ala Asp Thr
50 55 60

Phe Glu His Val Ile Glu Glu Leu Leu Asp Glu Asp Gln Lys Val Arg
65 70 75 80

Pro Asn Glu Glu Asn His Lys Asp Ala Asp Leu Tyr Thr Ser Arg Val
85 90 95

Met Leu Ser Ser Gln Val Pro Leu Glu Pro Pro Leu Leu Phe Leu Leu
100 105 110

Glu Glu Tyr Lys Asn Tyr Leu Asp Ala Ala Asn Met Ser Met Arg Val
115 120 125

Arg Arg His Ser Asp Pro Ala Arg Arg Gly Glu Leu Ser Val Cys Asp
130 135 140

Ser Ile Ser Glu Trp Val Thr Ala Ala Asp Lys Lys Thr Ala Val Asp
145 150 155 160

Met Ser Gly Gly Thr Val Thr Val Leu Glu Lys Val Pro Val Ser Lys
165 170 175

Gly Gln Leu Lys Gln Tyr Phe Tyr Glu Thr Lys Cys Asn Pro Met Gly
180 185 190

Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp Lys Arg His Trp Asn Ser
195 200 205

Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg Ala Leu Thr Met Asp Ser
 210 215 220

Lys Lys Arg Ile Gly Trp Arg Phe Ile Arg Ile Asp Thr Ser Cys Val
 225 230 235 240

Cys Thr Leu Thr Ile Lys Arg Gly Arg
 245

<210> 34
 <211> 744
 <212> DNA
 <213> Homo sapiens

<400> 34
 atgaccatcc ttttccttac tatggttatt tcatactttg gttgcatgaa ggctgcccc 60
 atgaaagaag caaacatccg aggacaaggt ggcttggcct acccaggtgt gcggacccat 120
 gggactctgg agagcgtgaa tgggcccgaag gcaggttcaa gaggcttgac atcattggct 180
 gacactttcg aacacgtgat agaagagctg ttggatgagg accagaaaagt tcggcccaat 240
 gaagaaaaca ataaggacgc agacttgtag acgtccaggg tgatgctcag tagtcaagtg 300
 cctttggagc ctcctcttct ctttctgctg gaggaatata aaaattacct agatgctgca 360
 aacatgtcca tgagggtccg gcgccactct gaccctgccc gccgagggga gctgagcgtg 420
 tgtgacagta ttagtgagtg ggtaacggcg gcagacaaaa agactgcagt ggacatgtcg 480
 ggcgggacgg tcacagtcct tgaaaaggtc cctgtatcaa aaggccaact gaagcaatac 540
 ttctacgaga ccaagtgcaa tcccatgggt tacacaaaag aaggctgcag gggcatagac 600
 aaaaggcatt ggaactccca gtgccgaact acccagtcgt acgtgcgggc cttaccatg 660
 gatagcaaaa agagaattgg ctggcgattc ataaggatag acacttcttg tgtatgtaca 720
 ttgaccatta aaagggaag atag 744

<210> 35
 <211> 67164
 <212> DNA
 <213> Homo sapiens

<400> 35
 attattaaag cggtagtctg ccggcgctga taagcaacaa gttccccagc ggtcttcccc 60
 ccctagcctg acaaggcgaa ggttttctta cctggcgaca gggaaatctc ccgagccgaa 120

ttcagcttcg ccggagcccc aggtgtgacc tgcgtagtgg gcaagggagc ggtgtgcagg	180
ctgagttttt ttttttacag gggtaccctg aaactcctca ctttctctgg gaactttcag	240
tgccaggacc cagtaacggg cggttagaag gcagccctag gaaacacctg ctacatagca	300
gggcagttag gcaatcattg gtaacctcgc tcattcatta gaatcacgta agaactcaaa	360
aggaaacgtg tctctcggag tgagggcggt tgcgtaaatc tataggtttt tcgacatcga	420
tgccagttagc tttgtcttct gtagtcgcca aggtgggtga gagtttaagc ttgcggatat	480
tgcaaagggt tattagattc ataagtcaca ccaagtgggtg ggcgatccac tgagcaaagc	540
cgaacttctc acatgatgac ttcaaacaag acacattacc ttccagcatc tgttggggag	600
acgagatttt aagacacttg agtctccagg acagcaaagg cacaatgggtg agtagcaata	660
aaacctgcat tataattgaa aaatcttgac atgttgctta acaacgggca tatcacggct	720
cttcctagca cttcacacgc caaagaacag cagctactca ggccagggga atcgggtttt	780
tacacagtgc aactttaatt ggaatcattt gagatttgac acagctatgt ggaactgcgt	840
ggaacaaact tggagctggg tgggggggtg tgtgttatat tggttggtca aggctgatgc	900
ttgtctctca gcagtcttg atctattct tttccttaat gtgtatgggt tatgatcata	960
ttctatgatt tatatgtggg catgtaattg acatttgcaa ggggggttaat ttccatctaa	1020
aaacaataat gctgttagag gttgggggtta ggggggtggag tgggggtaag ggtggggtaa	1080
agactgggag tttagggtga gatgggggggt ggggttgggg ggagagaaat aagtcagaag	1140
tgcatatcac cggtaatggg taatcctctc gtagaagaaa aggttctcat caacatgtga	1200
tcaactatta acaggatggc tttggcaaag ccatccgcac gtgacaaacc gtaaggaagt	1260
ggaagaaacc gtctagagca atatcaagta tcacttaatt agagattttt aagccttttc	1320
ctcctgctgt gccgggtgtg taatccgggc gataggagtc cattcagcac cttggacaga	1380
gccaacggat ttgtccgagg tggcggtacc cccaggtagt cttcttggcc ccgctgtaaa	1440
gccaaccctg tgtcgccctt aaaaagcgtc ttttctgagg ttcggctcac actgagatcg	1500
gggctggaga gagagtcaga ttttggagcg gagcgtttgg aaagcgagcc ccagtttgggt	1560
cccctcattg agctcgtga agttggcttc ctagcggtgt aggttggaat agactcttgg	1620
caagctccgg gttggtatac tgggttaact ttgggaaatg caagtgttta tctccaggat	1680
ctagccaccg ggggtggtga agccgcaaag aaggtaagca ccagggcggg gaccccttgc	1740

atccccaatt cttgagctat tttgatactg tcttccggag aggacgcgtg gtggagggga	1800
ggaggttagag ggagagcatg agaggggggtt gtttcttggt atttgcccag tttgaattgc	1860
cctaggtgag aaccctgggg caaagggaga aagaaaaaaa agaaactcag tcttcctgcg	1920
gatataatga gtttagttaa cttggacctg caaatgtctg attcaaagt aagatttatc	1980
tctctttttc tcctcttcac ctccctcttt tccgttctct ttgctgggtg gtgtgtgtgt	2040
gtgtacagta gattcattac taattatgaa gcttttgcaa aacattcgaa ttcctaaaat	2100
ttgactttgt agcatttaga atcaggcggt ggaggtgggt tgcgggtgggg agaggaggtg	2160
gaggttggga agaggggaagg aggtaaagct aaacctccaa cacaaaaaaa tgaatcaagg	2220
taatttcagc tcttctagt agaggattc attctctctg tatccctccc tccctctctt	2280
tccccctccc tccctccttc ccgccccct tcttccaccc cgccccctcc tccagcctcc	2340
atccctccct cattctatct cttcctctcc gtcgccctcg ctctctcgctg gatgcttctt	2400
tctgggtttt cttttttttt tcccttctgt cctccctccc cgcgagtttc gggcgctggc	2460
ttagaggggt cccgctttct caagggaagg ggagctgccg agaccgcgct ccgctcccca	2520
gccgggccgg atgcctcact gagcccaggt ccgagtcagt cggggtaact cagggaagg	2580
ggagcctccg cctgggagta gaaggtcctt tccggaccga agagccagag agcgggccgg	2640
gcgagggggc ctgggcggct ggaggcgggt gagaagaaca ctttttagctc cgtgcggcgg	2700
ctggacagag ccaccaatca gctggacgcg cagaccgccc tgccagggcg aggttgctgc	2760
cggaggcgcc ggtggagggc ggccggctag tcgctgagcc gccgcccca cccgggtggg	2820
caggggactg gcggtgggtg gaggtgaggg gcttggcggg tgagatagaa gcggcgcgga	2880
gccgccaga cctgtgttct acctctcccg ccccgccctg ccccccggg ggacagcgaa	2940
ctgccggaac gcgcggctgc gttatcctct tgccactctt caggagctc agggacttag	3000
gcgcccctgg gcgggggcca ccaggctctc cacactccta taaccctcac cccaccccc	3060
ttctcaggcc ttttgttccg gccacagagc caagcccggg ggcagttttc gcccagggt	3120
ttggtcgttc tggactcctc ccccgatcag ctcttcttaa ttaaaaggca agagcttggg	3180
gtaggggtga ggggaacgta ggaaaaatct gtttccgaaa ctcaagacca ctgttttaac	3240
gaacgaaaga aagaatccca actctgcgca ggtggattca taggcgaagc gaggatattg	3300
tggaaattca gaaggaaaag ataaaaaaca ggcgctagga tcagatgacg gtgataggct	3360

gctcggcaca	caaagggagc	gtagggcagg	gtttacggag	caagcctgca	gcgaatgggg	3420
cacagattgt	tccgagatcc	agtcgttttc	tcagtcagat	ctacgcgaag	ggagggggagg	3480
ggagggggcgg	gcagggggagc	gtggcgggag	gggctgagct	tgggggcggg	gggatttctg	3540
atcagtctga	tgcaattcca	agcgtgctgc	aaaggaactc	caaggcgccc	gcatcaccat	3600
cgccacccac	ccttcccaga	tggtgctgtt	ttaaatacgg	atctgcaggg	ctgaacgcag	3660
aactgggaga	tttattgcaa	aatcccggga	ggggcggggg	ggggtggtgt	gcggaacggg	3720
gaatggagga	gcagaattta	aaggtgcaac	gcttgctttt	tccaatcagg	cggcaaccgg	3780
ccggaattat	tatttttttc	tttctgtctg	cttgtctctg	gattctaatt	caccaagaaa	3840
gaggtgtaaa	tattgtgaca	ttttgaggca	gcttgatgga	tgggaaagaa	atcatctgtc	3900
actctaaatt	gcagagttcc	ctctccccgc	gccatccctt	gctagcgaat	actcgctgct	3960
gcctaataca	gttgctaggg	cttcaaata	atgcatcggt	aagggaatat	tatcctttta	4020
gttgacttgc	caatttagtt	gacagttgaa	tcgagaaaat	tgtagatttc	gtgtctctgg	4080
gaggaaaaat	gcttaacagt	ctaagtcttg	taaccttgag	gtctttaaca	acttaaataa	4140
acctcaaaag	tgtcacgtca	tcctctacac	acacacacac	acacacacac	acacacactc	4200
aacttgtaag	atgacatgg	ttcacctaaa	ctgttggtgga	aatgaatagc	actttaaaaa	4260
tggtgcacct	gatattcact	gtttatgtgt	atttaciaag	agctcttcag	catgaaggca	4320
agacatttca	attgtcctgt	ttggaatcag	tcagaagact	agaaggtgat	ggagagaaga	4380
agggaagaaa	gaggaaagag	agagaatttt	aacctagatg	ctattaaatt	aacagtaacc	4440
tagcctactt	ttatacccct	tggtcttgca	tattaatatt	tctgtatgtg	agatttttagc	4500
ttggtctagc	tccccaatg	gagtatacca	gtattgattc	agatgagaat	atgagcatcc	4560
tgccagtagc	tttttcagtg	tcattgatag	taagacctac	aacacagcaa	tttttgaggg	4620
atagaagaga	atatatataa	gggctttgca	aactgggaag	caggcactcc	ataaatggga	4680
ggtatcatta	tgacattctc	tttgcacata	tcattttcat	ttcactgaac	cagagtacta	4740
gttattttta	aacataatct	aatgtatatg	ctcaaggtag	taagtgggga	ttttaaaagc	4800
aagtgattag	ttggcttata	aaatattatt	tttcaattgt	ctattaatgt	acattggaaa	4860
gaaggctttt	aaagatctaa	aatcaacata	aataagcttc	ccctttcatt	tgccagactc	4920
tttccctatc	agattttctga	tctaaattct	taataagaag	agaagctggt	gaatttagtt	4980

tctttccttt	tcctggcctg	tcctctaggg	gaagctttag	taagaaacaa	cattccaaaa	5040
tcaggcagtg	agcgagagag	aaggcaaggg	actggatgac	cacaaaatag	ataatcagcc	5100
aagaaacaga	aatgagggaa	aaccagcatt	aaagcatgac	ttacaaaggg	tttttatttt	5160
gtaattctgt	aattttggga	ccaggctcaa	acttgctcaa	gtaacattca	ctcgatcata	5220
ttgcttacia	tctgtcagtt	aaaatgatgt	ctaactgagc	atatttttta	ttaaataatac	5280
ttctctcaaa	ggccagtaaa	gctactcttt	ggttttaatt	agacaaacta	gtctaaccac	5340
ttaaataact	ctaatgaata	tgaactgata	tcatcagatt	taaaagctct	gctgaaaact	5400
aaatttattc	tgaaaagcac	tgacttgcca	gaaaaatatc	tatttttgca	gctttctttt	5460
cactctatgg	ataatttaat	gagttgctta	ttttaatttt	acaactgcta	cctcagaagt	5520
atctcaaatt	atctttcttt	ggctgggtgc	tttctctgct	gatctgctac	tgctgtgtgt	5580
gtgtgtgtgt	gtgtgtgtgt	gtctgtgtgt	gtgtgtgtat	gggcgtgtgt	gtctctgtgt	5640
gtgttttcta	gtgggaattt	aacaagcagt	gagtctctta	aattttacatg	ccataatcta	5700
tgtcaagaac	attgcgtact	acttagcaat	aaaaataaac	attagcatct	agtgaaagct	5760
taccatcatt	gagtgcctatg	gaaatagagg	tcttaaagaa	agattaaatt	tttcaacaaa	5820
aaaatttttc	cccttttttg	cttaaagggtg	attataattt	caaaaatatg	acatctttcc	5880
tcttttactt	tggaatgtag	agctgctgct	ttaacaagtg	tcttttgaga	aagatacacg	5940
tgtttcataa	agattaatac	ccttaaaaca	ctatgggtgca	gagagggaag	gatgaattct	6000
ttaaccctgc	ctctaattctc	atttggcaat	ttttggagta	ttcattctga	ctttttaaaa	6060
attcaggtgg	attttttttc	tgctttcttt	ccaacattat	aaaacaatcc	tataagagat	6120
ttttctgcta	tagtgcagac	tttatttgta	tttcctagta	ataacacttt	agattcatat	6180
agtactttga	cagctctcta	taggtttcat	ttgatttcct	tatcagtcac	gtaaggtagg	6240
aatcaccaat	caccttttac	agatgaggaa	agtaagggtgc	agaattatct	aacactacac	6300
tgccagtaag	atgtaaagac	taactcagag	tctttcttca	aattttaagg	aaatttgtgt	6360
ttgttcctct	ctagaccatg	ctgccttaaa	ctccactagg	gcatcagagg	gagctgtagg	6420
cattattttc	tcctattttg	atttatttta	ttaaattaaa	aacatttttt	acaaatagtt	6480
ttcaaaattt	caggcctaata	ggaaagtttt	aactagtctt	cctaagacag	tatttccttc	6540
tcccacagtt	agacactcaa	agaaagcagg	actcttcctc	tagttgacac	accatctaag	6600

tcatagttgc	taattcccca	aaaaacaaat	acaaagaata	agaagaacca	aagccaagat	6660
gtacattacc	tttacttgtg	aatcatagaa	tctgggttct	gggaagggtc	tcagaggtca	6720
cttagcccag	ctcatatctg	atacatgaat	tcatgctggc	ttgttttggt	tcaatcctct	6780
ggaatttcat	cccagtcttg	gacactgggt	acttgtttat	cttggtagat	gttctgatac	6840
ttatgtggat	catgagcttg	gctgatcatt	tcccattttg	cccaaaaggt	atactttcta	6900
tagagactct	agcattcata	attttatttt	gtaagtatta	attggcaaca	cataatttgc	6960
cattatgtgc	caggaactgt	attacaatct	ggagattcaa	gaacaagcaa	gaccaaattg	7020
tccctatctt	ggtggatctt	gtggcctaga	gacaaagact	ggtctcttta	tttgctccaa	7080
tccttagaag	gggacagacc	tctggactca	taagattcct	tttttcaaga	gaacacagt	7140
aataagatta	aaagcctgag	tttaggcttt	acttccaacc	gcttggtcagt	acccaaaaaa	7200
gtcaatcagt	cactcttgat	ctctgcttcc	tcacacataa	aacaaaagag	atgttctcag	7260
aataaactct	aaattccact	tccgttccaa	gtttgagtga	tataaaaagat	atagtctata	7320
actatcttct	gcagtgtagg	ggaaattaag	gcctaagtta	ccctaaccct	tggtgattta	7380
caattctggg	tgggtaccgc	aaatttttta	cttgtttaaga	agtatatcat	gaaaaaaatc	7440
aaagtaacat	atttgaaccc	agaatagaag	gaatctaggt	gttgaacctg	ttcatttatg	7500
gaatgatggc	taggaaagtt	ttaattttaga	gaatgatctc	aaatttctgc	ggattatttt	7560
taaaagcagg	tggtgcatg	gaagacctgc	tatcaaattt	actgcttatt	ttctttgtgc	7620
aagcagaagt	tataattttt	caggtcattt	ccttttaaaa	tcaaaaatat	tacatcctga	7680
aattgcctgg	gtctcatgaa	taatgcatta	tatacacatg	ataatagata	attagatgga	7740
caagccaaaa	gaaacatgaa	aggaagcagg	tagcccaagg	attgcagaag	gtgtgtgggc	7800
attttgacat	ccaggaaatg	ctatagatct	gtccttaact	aactcagcct	ggtggagata	7860
attaagaaaa	aaaatgtggg	tgtagaaaga	ctgcaagcca	ttccctggga	ttggctagat	7920
tgctgcagta	gttcaaaaac	aattggcaca	gccaccacac	ctagacatga	ttgccctttg	7980
atgaggcagc	tattgacttt	tataaagatg	tcatatttta	aataacttct	gatgcactag	8040
gcataacaga	catcattctt	gaattctatt	ttagacaaat	ggcaaaatgt	attacaaagt	8100
attaatttaa	aaataaaaaa	tcttttaaagt	ctagtgtcta	aaaaccagca	gtttagtaac	8160
atgcaacctc	tggatttaag	aattcagcct	gaagctggga	gaaagctgta	gcttgtatag	8220

gacatthttga tccactctgg gcattttcca gaccactaca ggaagtaaaa tgtactthttgt	8280
caaagthtttt aacctthttgag tgaatgttaa atccactcca aaatcttctgc aacctggggaa	8340
aggtgatcca acaatthttcc taaatagcgg cagaaaatgc tctgagatct ttgttcccag	8400
agtgaatgtt ataatgttat gctatctaga aatthtcttg tagcaccatg ctcatcagta	8460
ccaaaaggag ttagaattga ttcttcccgc ttcaaggaaa tatatcaacc acctctgtc	8520
tctaagtaac aaggthtactg tggggaaaaa atacacaaat taggtgattg cagaaagggtg	8580
tcacaaacat ccaaagcctt tgggataggg cattgcagtg tgagtgaata gagaaaagaa	8640
agagaatgtg ggaaaaaatt gagaaataaa aagggaagtc acagtggagt tctaattata	8700
caggggctct tgaattgact gttctctacc ttccatgctc attgttggtt tggctacttt	8760
agtaggaaac aatgattttct tctgctttca ctttctctct ccgctaagga cttcttactt	8820
gccaataact tccataatca atgtttaaga attgctctga tgcccagtgt ggtggctaac	8880
gcctgtaatc ccaacacttt gggaggccga ggtgagtggga tcacttgagg tcaggagttc	8940
cagaccagcc tggccaacat ggcaaaaacc tgtctctact aaaaatacaa aaaattagtg	9000
gttcatgcct gtaatctcag ctacttgggga ggctgaggta ggagactaac ttgaaccatg	9060
gagacagagg ctgtagtaag ccgagatcat gccattgcac tccagcctgg caaaaaaaaa	9120
aaaaaaaaaa agaaagaaag aaagaaagaa ttgctcactg taatgacttt catgccatgg	9180
actcaactct cttggcagtc tggtaaagct tatgtaaacc cttctcataa aaatgtctaa	9240
atggggccag gcgcgacggc tcacacctgt aatctcagca ctttgggagg ccaaggcggg	9300
tggatcactt gaggtcagga gttccagacc agcctggcca acatggcaaa accctgtctt	9360
tacaaaaaaa aaaaaattag ctgggcgtgg tggcatgtac ctttaatccc agctacttgg	9420
aaggctgagg cacgagaatc acttgaatct gggagggtgga ggttgacagt agccgagatt	9480
agccactgca ctccagcctg gatgacacag tgagactttg tctcaaaaaa aaaaaaaaaat	9540
tctaagtga tgaataaat gtataagatt acaaaggaag ccagtggcat tgatgtacag	9600
ttataaaaac atttaaaata atatattgtg tgatatagta atatatatgc tttttaatac	9660
attaaataag atctaacagc aaggtaaata ttataattht gaaataatga taagtatcaa	9720
tgtatthttga aatatctata aaactgacgt gatatgaagg tgtctgtgat gtatactggt	9780
gagaaagcat gcaagtacta ctgtgtaaca thtcccacac atatthtaaca acagaacact	9840

tgagaagcac	ttattaacac	agcatagatt	cagaaatatt	aatttagtaa	atgtcaacat	9900
tagccattgt	tgtttccttc	ctggcaaaag	gaaatcagca	ttgggagaaa	acttttaaaa	9960
ttcacatttg	ccattagaca	agctgtcaag	tggggaaagg	accaaatact	gagaaggcca	10020
gggtatggta	agcatgtttc	tattgactga	gcttgctatt	actctaacgt	ttatctttag	10080
catcaccagc	acaaccccat	taccctagca	atccatcact	ccattgaaaa	agataaaaag	10140
ttcagattct	ggtcattaac	tcagcattgc	ttaagatacc	tgttctgacc	tcactaacc	10200
aagagattac	tgaaactctt	cctgtttgtc	attactacac	catgggaaat	tataatgatg	10260
tgggatgaca	tttactctgc	attcatccag	tgctgttatt	tgttttgtat	ttggcatata	10320
ttacttaact	cttaaagtaa	ctctcagaga	tagattaaga	aaactagagc	tcagagaatt	10380
taagtaactt	gcccuaacta	acacagaaaa	tctgaagtgg	agaagctaaa	cttcaaacc	10440
aaagttttct	ggttccaaag	tccattatga	agttgtgcct	ccccatctta	tagctaccac	10500
ccagatttaa	tctgggtctc	ccattatcag	atggttttaca	tacacatttt	cttacaagat	10560
cttgaccaca	actctttgag	atggccatga	gtctcacaat	tcattttccag	gagtgtctact	10620
ttagaatcat	tttgatcttt	gctaaccgat	gagagatttt	caaataagcta	attgtcacct	10680
accctttttg	aagcccagtt	ttcataatca	taaaatggaa	acagtattac	aatgttttgt	10740
taggatcata	tacattaata	ataaaatcta	actttgttga	gctcactatg	gtgagcattc	10800
tgcatttcct	tagttcattg	aatcctcaca	acaatctttc	taggctaaga	cgattatfff	10860
tcttttaaaag	ataaggaaac	tgaggcatca	gtaattaatt	aactatctta	aattagcaga	10920
gccaataagt	ggcaaagctg	ggttcaaacc	taggtctgtc	taatgtcaaa	gccctttttt	10980
taatcactaa	tctgcaaatc	actattcaat	cttagctfff	attattataa	ttatcatcac	11040
acttaaaaca	ctatcaagat	acagaatgat	ccagacataa	gtatatagtc	actgaagaga	11100
ttagaatctg	aaacttttca	cctgcatggt	cttccttcca	ctttagtffa	ttaaccaat	11160
ggatgatgtc	tgactccttt	cttaacttgt	ttagggcagt	tccaagttag	ttgacttctg	11220
agagttattg	agtaagaaat	gttataaatt	gtttggatta	ggatttagta	tgtttagaag	11280
ctatttcata	agtttgcctt	tgcgaactgt	tactggctat	aatactgcag	atgctgtgat	11340
gaggaacacc	ctctccaaag	acacacagtg	gatgacaaac	ctccaaagct	aacatgttgt	11400
ttacagatat	ggagaagaag	aggatggaca	agcacagtct	aaaacgtaat	tacaaggctt	11460

atagtcctg	ttggggacta	gaatgtttat	tggctttcct	tgtgcaattc	aatgctcttc	11520
ctccaaagga	tccactccaa	acttggaaact	ttcctgaaaa	tagcatctca	tttgggagca	11580
tgccaggaat	tggtgtctgg	gtcctttgtg	tctttgcacc	aactcagaac	tctggatact	11640
agctctagaa	actaagctgg	gatataattct	gggtaaggga	gtagcatatc	tacttgggca	11700
tcttcctgat	acatttatatt	catccatctt	cctcctagag	agcacctcct	agaaagatgt	11760
ggttttaaat	gagggattgg	atgcatactg	gtatgtctta	gcacacaagt	cagtgggtctt	11820
tgagagctg	ccaaaggcat	ataagtaatc	aaagatgcgg	aagtctatga	agagacttca	11880
tcccacctcc	actctgattt	attcagggaa	ggaccccatg	aacacataat	ggatttgata	11940
cgtcccagag	ctctgaaagc	agcctagcaa	aaaaggataa	tcttgaagga	cattttgatg	12000
tatgaaaaag	tccacctaag	gctttgtcag	agataactaa	gtaatatgat	ggctggtaga	12060
ctgtaagtcc	ttaccttggc	tcaggaactg	tatatcattt	ggtaaaactaa	acttgtcgtt	12120
caaatttaga	tagaaaaagt	accttacaaa	tgatctagtt	cactgattcc	cttcatgcat	12180
tgaaatcacc	taaatcatct	cttctttctg	agataaggtc	tgaatgtgtt	gccagcttta	12240
gcaaactcag	tttgtagccc	actgacctca	tttgattgat	tgggcaactg	aggtgcacag	12300
tggtagatct	ctcaatttat	tcaataaaca	attatatggc	ccttacgata	tctatctgaa	12360
caatcttggg	ctagtgaagt	tgcttgccca	ggttacatgg	ccagaaaactg	acagttttaa	12420
attaggacca	aagttctttt	gactactatc	tgggccttaa	aataatatca	tatgacaaag	12480
atatttcttc	tgtttcctaa	tagtcacatc	aaaaggaaac	aatggacagt	ttgtgcaaga	12540
ttttagttac	tttaatgttc	aaaataaaat	taaaaacaga	ttattactaa	aacataagca	12600
taacaacact	ttaatagcat	tctaatacaga	tattattaat	ttcaaaatgg	taggacaaaa	12660
ctaattatac	tttatacttc	ttaaatatcc	tatagttact	ttatgactat	tgagacacta	12720
gctaaaactt	gaaacttcaa	gttttcattg	attcctatat	tattacttat	ttcagagtta	12780
cttcatttgg	ttcttttatc	tgagattgga	caacagcttt	atttgatttt	cagcgacaaa	12840
attcttttca	ctcctgatcc	tccaccccaa	gaaaacaaca	gctactaata	tattttccct	12900
aaagtgatca	agaaataaaa	gaggaattct	agccaggcgc	ggtggctcat	gcctgtaatc	12960
ccagctcttt	gggaggctga	ggcgggtgga	tcacctgagg	tcaggagttc	gagaccagcc	13020
tggccaacat	ggtgaaactc	cgtctctact	aaaacttaaa	aaatgagcca	agtgtggtgg	13080

cgcatgcctg	taatcccagc	tacttgggag	gctgaggcag	gagaattgct	tgaacccagg	13140
aggcagaggt	tgcagtgagc	caagattgcg	ccattgcact	ccagtctggg	tgacagagtg	13200
agactctgta	tcaaaaaaaaa	aaaaaaaaaa	aaagaggaat	tctaaaatta	attatatcta	13260
ttaatatccc	tactcttaaa	acgttagaaa	atgtttgctc	atttaaaatt	tttattttta	13320
aaaccacctt	atattccaac	taaatactct	ttggagcaat	ttctttgttc	ctcatataat	13380
atccatacat	ataattctgc	ttttgtgatt	aacttttatt	actactcttc	taaaattgtg	13440
ctcttataaa	catcagttaa	ttaagagtaa	atctgatggt	ttataaattc	tttctagaaa	13500
cagagagcaa	aatcatataa	ataacaatat	gaatttccaa	aagtacaata	ataaaaaaaaa	13560
attagaaaaa	aattaatcta	ggaaatagtc	aagaatatgt	caaacttgta	catacttttg	13620
agataaattg	gcatcatgta	gattagcatg	attcttcttt	atggaattca	acttattttt	13680
actcactttg	ctctaattag	tttttgtgtg	cggacaagat	ggaaggtaat	ggaaatttgg	13740
cttgcaaagt	agttctaaca	tgatctacat	ccacaatctg	gttataatgc	tataagaata	13800
ttatgtggga	atagtagttc	aaatcagtat	ttagtatgaa	cataaaggga	caaacaatgc	13860
aaagctaact	taagttgttt	acacttgga	cttattttaa	ttaaaaaggc	cagtggatgg	13920
tcatatgttt	ggctcattct	tctcaaggcc	ttcaggaaaa	catgcctatg	aaataaaaga	13980
tcctcaatat	taaacatttt	actgcatttg	ggggacacat	gaaatctggt	aataaaggaa	14040
gtgttggtct	tcatttttct	aattcagcat	ggaaactatc	ttgaggaaaa	ctgactatgg	14100
tcttagtttg	tgtctcagaa	atatatttag	tctgaatcat	ggcgtcgaca	tctgacttcc	14160
aaaattggat	atctagccgt	atagtacctc	acctcccaca	cacaccaccc	cccattccca	14220
ggcatgact	actgtccaag	cagcaaaaaa	agaagtaatt	tcccagagta	catacatggc	14280
agtgacaacc	aaccaaaaca	aaaacaatta	taggggctgg	aatttaaatt	aatggctgta	14340
ctctcaccaa	ttcattcccc	attccacccc	atctctctgt	cttcaacttt	tatgaaacat	14400
tatatttgtc	ctattcttct	gtatcagcat	cagcctttcc	tatatccaac	tagacttata	14460
acttcttggg	gcctctcact	ggctgactaa	ggtttcagaa	gtacctactt	acagcaaaca	14520
cttgcagcag	tctctttttg	gttacaaagt	ccctggacaa	tttctcaagg	cgatattatg	14580
aagaggaagt	aaacattctc	ctctgctacc	ccatttcttt	ttagagtgct	aactttattc	14640
tatatctggt	ttaatgtctt	cttaggccaa	ttggactgat	tttacagaca	ccatagaata	14700

tctcctgagt	aatgggaaca	atatttctgc	tgatcccatg	atttggcttc	attgggttgt	14760		
taggccataa	tggagacata	cttgatgaat	ttatgaagac	ttgattctag	gtatcatgta	14820		
ggtagcata	attctctttt	actgaattca	acttagtttt	attcacttta	ttctaactgg	14880		
attttgtgtg	cagaccaa	gaaaagaa	ggttcaattt	aggtgaaagg	taaagcttca	14940		
aaagtagtgt	agtatttcat	agaccttacc	tttgagagaa	attatatcag	tatataataa	15000		
gcacctgaga	atatgaaagc	acaaatccaa	tttaa	aatgtg	aaaggtctac	aacttgggat	15060	
tttaa	atgga	gtacagaaaa	gccactgttt	cttaaacaat	tttg	ttgagg	gggaaaacag	15120
tgaaagctaa	atgttctatt	caagagttgt	ttcttttgaa	aataatgctt	catttaaaag	15180		
ctaaggacag	aagacgtagc	tttg	ttatga	aggctcatct	ttttattaaa	caaccactac	15240	
tttg	tctcca	agttgcaaag	ggaagatttg	tcaatctgat	tgaatcttcc	ctttagtttt	15300	
tccaacagc	tgtgtccaga	taattcatga	ctcctgtgtt	tcctgagccc	tggataattt	15360		
cacacacatg	tctggtttgg	ggctccacat	tttcagaaaa	atatagaaat	cttggaggag	15420		
gtccagggta	gaccaaggga	aatgattaat	gggttgaaag	ttggagttta	tgaagaaagg	15480		
ttgtgagatc	tgatcttttg	ctacgagaaa	agtctgagtg	gtgacttaat	aacataagga	15540		
ggtagtaag	cagctgttct	ccatcttcac	taaggttgaa	tgaaatgaaa	taagatataa	15600		
attgcaacag	gaacaaaaat	gcattacaag	tgaggacttc	caagcaccag	cattgctgga	15660		
ttctagatag	ctcccaaaa	gaaggatgtg	tagtctactt	ccctgggtctg	caatgacagg	15720		
cctataaata	gtgagaaaga	tgagataatc	tcttaagatc	ccttctggac	ctatctttta	15780		
taggtctatc	tatcatat	ttt	gcctcaa	acaaaaatta	tctgatttcc	15840		
tccctctcac	cctatccact	ccttctcttt	tgtctacctt	ttgtaaaaca	ctgctaaccg	15900		
aaataactgg	ggactgatta	accgtggtgg	gccctcccc	gcctctaagt	gccactccag	15960		
cttggggagc	aagtttcttg	tccatcacta	ccacccctg	gccactaggg	gcatgtttac	16020		
catcatcttt	ctacacacca	aacctacggc	aaggga	aaacaaaac	aaaacttcct	16080		
agacttaaca	aatttgcaag	tgtcaccatg	gatta	aaaata	caactcttat	gtcctagaat	16140	
atgagcatgt	aaagggctaa	aatgtat	ttt	atgcatctgc	ctgtatcagc	ccatagaata	16200	
gcctcctgac	agatagtaga	tactcagcaa	tctttcatca	actgaatgac	tgtaactatg	16260		
aagtgaaagg	caactaaagt	tgagaaagtc	aggagtttcg	gatgtttcca	aatgattctg	16320		

tatgccagac	taatctaaag	cctaaccat	tcttcacaac	catgcactat	taaggatttc	16380
attctcacca	tgctgtgct	atctggaggt	agaaagaggg	ccagttgcac	atcctgctca	16440
agtccttggt	caaaaagacc	actaaagagt	gctttgtaga	ttcatgtatc	agaatcacat	16500
gaaagtaggc	caaattctta	gtgtgtgttt	ttaaataaag	actttaggaa	gttcacttat	16560
ttttttctaa	attatttttg	catattcttc	tttttcattt	ttttcatgaa	gaatttaaaa	16620
tttggctgta	gaaaatctct	cactccaaac	atcacacagc	ctaaataggt	gagtctcaaa	16680
aataagctaa	tgttcatctt	tcattctgatt	caatgtcctg	aaaccctttg	gtttaaattt	16740
gttaattctt	ctcatggctt	ttctcctagc	aaaaccaact	aataccacag	ctatttatta	16800
ctgtcagctc	taacttatgc	ccacaatctc	acatcccttt	tgaccacgct	tatagaacta	16860
ttacaacaag	taaaccaaat	ttattcttca	ttattaattt	ttaaatgttc	tcagcacaaa	16920
tctggtaact	tggagggcta	caagttgata	tttctcatat	gtttgggggt	ttagtctcaa	16980
cagtttctta	atggtttcta	tgccgttttt	cttgatccaa	ctaaatatta	ttcccagatg	17040
ggatcagctt	ttgaccctct	tgttctactc	tcctagtctt	ggcccttcta	aaagtttctt	17100
gctgtggttc	ctttcttttg	tctgccacta	atggctatgc	ctggttacat	aactcctgta	17160
acaggtgttg	accaatttga	acacattttg	gtatgggtatt	gagctattct	tatggttcat	17220
aaaaagctta	gtgagaacgt	aacatctcat	gaatagggaa	attacttctc	ccttaagggt	17280
tttctcagga	caggcctcat	acaagaattt	caaggattgc	gagtgacata	gtttaacatt	17340
ggaccaggcc	tttcaaatta	tccaggatga	gtttgaaaac	acctgtgcca	ctctgctcaa	17400
cagcagagtt	ttctgtttac	taagtatttt	ccctatgcta	attacggaaa	gtttcaacag	17460
tttttttagg	ccaacttatt	tgatgctaga	ctagacaact	tatttttttt	ttcttgcaag	17520
gaatactgaa	ggtaggagta	actaggaagc	ttaaataaac	ataaatataa	aatgcttata	17580
gtgatagaat	tgacctcagc	caattaaaat	tattaataga	aaaaacatgt	caatgtcaag	17640
cctactacct	ctgttctcac	ttgagtaatg	aggattagtt	tatatattcc	gacaagaata	17700
gatgggaatt	caaatttctt	cctgaccttt	gttccccctg	gaacattggg	ttaggatcat	17760
attagaacat	aaccaaaaag	aaataaagat	tcagacgaat	tcacaattaa	tttttaagcc	17820
ccacaaaagt	gaaataggta	gcattatttt	ttcaagctgt	gaaactttcc	ctcattttag	17880
taatagagaa	aatgttcaga	ttataaaactt	ggaaactttg	ctcctaacat	atcaattatg	17940

ccagaggcca	atttttaaga	agaagagaaa	tgcatgctct	atattctcag	catcatcctt	18000
gcccacaata	gggaaataat	tttgtaaaat	gtttgatttt	agacctcaa	aattatctct	18060
atatgctacc	tgaattaagc	aaataaaaaa	taatatttag	aattccatgc	aaggcactgg	18120
tacaattttg	tttatcttgg	cttcattggt	tttgaatgta	agatgtactt	ttaaggcaaa	18180
taagtacatg	ttttaagctg	gtcgcataca	gtattggcaa	tgctataatc	acaaatcaga	18240
aagtttggaa	atgcttacia	gtgttaagag	gtgtgattca	tcatggttat	ctgaattggc	18300
atctgatctt	cttttctttc	taaatatccc	tgacatttct	gactcctctg	tcttttcttc	18360
agtaaaactg	caccacacac	tggaagcga	agatacacac	atatttttat	ataatgtcaa	18420
gggagagtag	gaataagaag	attggccata	gaccaccca	atcagagtct	gggaaatgag	18480
aacacttttt	ccttcagcag	aaatgctgac	gtgccaatgt	gaatttagca	gaaaaaagat	18540
ttgccataac	ttctaagtga	gcagccttca	gaatgctagc	ttagattcct	ggcattaact	18600
tgccaggtat	tttttcagga	aggaaataaa	ttacaattga	gcttaaaaaac	ctgagggtag	18660
aactcatttt	caagcaaagt	tgaagcatca	gtttgaagtt	aacaaagtta	aagtttggag	18720
tagggttcct	ccagtccttt	ataatgtagt	acaagtattt	tttttaaagt	tataacacta	18780
gccttttaaa	ttgtattgtg	ctactaaaag	aaattgtgcc	tgcattcatc	ttacaacctg	18840
ggaaccaacg	cagagggtct	gtggggtagc	ggtatccagc	ttcatgccct	ctgtccttta	18900
ttgctttctg	gttagcctgc	gtatttcaca	tacattaaat	attccacaat	aaactctgcc	18960
atctgtgctg	tagggtagtt	tgtattggtc	atgtgctctg	tcaagttgac	agagggtcaa	19020
agctaaatgt	gtgacactcg	aagaatatgc	atatatttga	ataatttgac	tatttagtcc	19080
aacaatttgc	aaaggcgctc	tgaatgatca	cacattctga	taacacttcc	aaggaacaga	19140
tagcttcact	taggggggtg	gggagatgga	agcagggtta	tttctagcag	gaattcttga	19200
gttactgaa	gtcttgtccc	tggtacttca	ctgtgtgaac	gtgggtaaat	tatttcctgg	19260
cagaggatcg	gattcttctt	ttataaaacg	ggtaaataat	ttctgtcact	agtctttaga	19320
agttctaaaa	tagctaagt	tagtgaattc	attttgctaa	ctgtaaaccc	ttaggtaaat	19380
tgaactgagt	atgtaataat	attatatatt	cagttcaaca	gcacattctt	ggtaaccaca	19440
agagggtcca	ggaaaggaaa	ctgtttataa	atctttccct	ttagcaaaat	taatgttgga	19500
gtctttaggg	aaattcttac	agcaatagtc	ttcgcaatta	ttaggtcaaa	cccctttgag	19560

attacagaaa aacgcacaca cacagaaagc tgcctgcaga atttgggtgt gggcttggtg	19620
ggagattcct ctgataccca gtgttggtacc cccaagagag tgtttctcaa agtgtgactt	19680
cagattgtct gcattcgaat tgcttggtgt atttattaaa attataactc ctgggccctg	19740
ccccaccct actaaatcac aatttcagga ggagggacct tcattttaac actcaccag	19800
gtgattttta tgctccgagg aggtccaggg actccaagtt aagtacggta ctgctgtctt	19860
attctttatt ctaaatttta aggtctgcac aaattgggtg aactaatgag aagaaaattc	19920
agctttaaaag cagaaacaca ggtagacggt tgacagagtt catcaaatgg ataattgaaa	19980
atgtcctctg gaccctagcc atataagttc tcttcaaggg tcttggctac aggcaaatga	20040
gaacccgaaa ggctatattgc tcttttgctg cgggcagtggt tgggggtgga gggcggggga	20100
ggattaactg agccagttct gccccaccc tcgaatcacc tacccccact ctggttaaag	20160
cagaagactt tttatttatc ttggctgccc tggttcggtta ttaaaagggt tagcttatac	20220
gtgtgtttgc tggggctgga agtgaaaaca tctgcaaaag catgcaatgc cctggaacgg	20280
aactcttcta ataaaagatg tatcatttta aatgcgctga attttgattc tggtaatcgc	20340
tgactagag tgtctatttc gaggcagcgg aggtatcata tgacagcgca cgtcaaggca	20400
ccgtggagcc ctctcgtgga ctcccacca ctttccatt caccgcggag agggctgctc	20460
tcgctgccgc tcccccggc gaactagcat gaaatctccc tgcctctgcc gagatcaa	20520
ggagcttctc gctgatgggg tgcgagtatt acctccgcca tgcaatttcc actatcaata	20580
atttaacttc tttgctgcag aacagaagga gtacataccg ggcaccaaag actcgcgcc	20640
cctccccct ttaattaagc gaagggaacg tgaaaaaata atagagtgtg ggagttttg	20700
ggccgaagtc tttcccgag cagctgcctt gatgggtact ttgacaagta gtgactgaaa	20760
aggtggggtt gttttctttc tttctctttc cgtttttctg tttggtcggc tagaaagcgt	20820
gtggctttag cgaggtctgt cattgcctgg gcttcctggc tggaacaagt aacttggtgt	20880
aacgttatct gggggcgttc atcaataaaa aatgctgtta ttatcttgat tgaattccta	20940
ttaggcaaac tctagagagg tcagtgcgcg aactctgttt aagccggcgt gttaaggca	21000
gcagagtaaa ccaatagccc catgctctg tgcgatttca ttgtgtgctc gcgttcgcaa	21060
gctccgtagt gcaggaaggt gcgggaaggt gtgtctgtgg cccgggaaac gcacgccctc	21120
tcccagagaa cttgggtgct gggatgggga ggaaggggag agttgaaagc taggggagcg	21180

agacctcggg gcgtgcgatt ctcactcgct ccctcccgcc ccagcgccca cagccgggggt	21240
ttctgcagag ggcgcgggac gcgggggttcc ccggggctga ggctggggct ggaacacccc	21300
tcgaagccgc gggcgtcctg tccaaggcgc ccagggaggg cgcaggactc gcagggcgat	21360
gtcgcggggc cctaggggag gaggtgagga caggccccgg gggagcgggg agttccgggc	21420
gcccctcggg tccccgcgcg aggaaaagac gcggcggttcc ctttaagcgg ccgcctcgaa	21480
cgggtatcgg tagcgcgggc gagcggggag cggggggcgg ggggcggggg ggggggggcg	21540
gcgccgtttg accaatcgaa gctcaaccga agagctaaat aatgtctgac ccgggcgcaa	21600
ggcgcagcct ggagctccgg gtccccgacg ctgccgccgc cgcgccggg cgcacccgcc	21660
cgctcgctgt cccgcgcacc ccgtagcgcc tcgggctccc gggccggaca gaggagccag	21720
cccggtgcg cctccacct cctgctcggg gggctttaat gagacacca ccgctgctgt	21780
ggggccggcg gggagcagca ccgcgacggg gaccggggct gggcgctgga gccagaatcg	21840
gaaccacgat gtgactccgc cgccggggac ccgtgagggt tgtgtggacc ccgaggtagg	21900
caagcgctgg gaatggggct tgggtgcagga gctgcccgtc cgcgggagag agttgactgg	21960
gggatcccc accccaaagt tgtgggacga ggccagtctc cttctttcct cccctccggg	22020
agaagggacg atttggagtt actcttgggg agttttctcc cccatccac aaccagaag	22080
gtcagccggc accaccaggg aaaaagggac ccgggggaagt cacgaagtag aggagggaag	22140
gcctggagga gaccagagc tgctgatgg gagcaaagac ggcgacccgg ggatccctcg	22200
cagccctccc ccagcccagg agtagtcgag agagacttag ggggccagag ctgtcgaggg	22260
tcctgactga ggggaggggt ctggggctag gctaggaatc cttccagggg gtgggtggtc	22320
ccgcgccga cttgcggggg gagtgggagg gaagcttgcg ccttcagccc gcatcccttc	22380
cccggagctg cacacggcta cctgctcccc aggaattgag actgaagtgg acttacaagt	22440
ccgaagccaa tgtagcttgg aaaacttggg aggcggaatt cctaccgctg ggaactgaaa	22500
gggtctgcga cactctcggg caggccgaac ccacatctct acccatcctg cgcccccttt	22560
ctgaagcgcc ctccagggaa gttaagagtt ttgactttcg gggagtgggt gggatgtacg	22620
tgggggattc ttgactcggg ttagtctctg gggatgcaga gccgggaaga ggaatgggtg	22680
agtgagttac tcctggaaag aaatagctga ggattggggg ctctgtgcct gacgggcaag	22740
aagaagggga gattacagac taggggcatc cctaaggaag aagcctcggg gctgcgaggg	22800

tgaactggag gatgcagtgt ttgtgtgttg ggggtagagc ggggatgagg gaccggggtg	22860
gaggggaggc gaggaggagg aggggacca gagaacgaag ctagggaagg tagagggtgc	22920
cctctgccgg ccatgctgcc aagagcagct actgggggcg ggaggctggg ggtggggaag	22980
tggtaaagga aggttttgcg ggatccctta gagagctggg aggagggact tgttgaatgg	23040
tgctgctgac tccagctcgg tggggcgtgc gactcgtcgt cggtaggattt tgactcctcg	23100
ttcttgtttg gcttctatgc aagttttcct cgcgctgggg gagctttgat aagcctcgat	23160
tggcgggtgtg ttagggcttc ttggatctta ttttagggtc ctctagttat cctgcactta	23220
ctccttaatg tcagtagcaa ccaaagaaca ttttccgaca agcacgcagg aatgttcttg	23280
gccagaagca aagaaaggca tatttctgag tgtttattaa tcctcctagt aatcttttaa	23340
agcaaagtaa tatgtaattg ggaacgttga ttttctaact gcatataaaa ggcgacatga	23400
tattaaatga gaccctccc tactgactca atatcctgca aaatctctct ctcccctta	23460
ttattatgga aaaatctatt ttatatgag ttgtttgtaa ggtcaaaagc cattttggtc	23520
ttacaatttg atatgtcttt acattttaac ttattgaggc ataattacag atttaatttg	23580
tatgaacgtg tgtgccttca atgcttatct catgcaacat aatttttagg ttggagattt	23640
ctgatgttat ggcagttagc gtttcaaggc attacacata ataggtaaca tagcatgttg	23700
aaattacacc acaaagtttt gaccctggga acagcacctt ttaaaaaaca tcactaaact	23760
cctgttcctg ttttctgatt ttgcaaagtc cttgcttaag actttttttt tttttttttt	23820
tttttttttt ttgggaaatt tacctctggg ttagcaggag aggtaaaaaa aaggaaagag	23880
acacttggtg aaatgtaacc ataaccttta ctggaattta aaacatgttg gtcaccatta	23940
ctggaattcc agggccataa agtcgttgtc ttttttttct tctacttcat ttgtaaaaat	24000
gtgataaatg ttggtaaata tagaccagta gtaagtatta tgacactaaa agcattatgt	24060
atgtggaact attttaagtt attacagaac attttctatt tataaatgat ataagcagaa	24120
agaaatgatt tccagataaa caaggcttac gtacatgttt tgaagcatta gaacattgca	24180
gacactctta gacatcacat tttttaagc aaaataacag taatttttca catacctttg	24240
gagcctttca tagcccatc agagctgagt tagtagctgg aagtttcctt tattttaagg	24300
tgatatttta aaaccattta acatgtatag taggtcaaca ttggtgcatc cagaaaaatga	24360
agcatttagg aaatctgttt cagtgtcttt tcaatgtgtg taacttttac ttgcaaacca	24420

atggaaccaa gaaagtcatc atttgcctaa aatgcagtca tcacctcaaa tgattcattt	24480
atactatgtg agttaattgc cttcatctca ttaatggcca aggagggaag ggaggtcctg	24540
gggtatttct tgttcatttt gactcaccag gagggaaaat cctgtaaaaa aagaaatgca	24600
aatttctaaa atcctggctc aaagtccgtg ggtttcctgt ttaaaagggg cgccatgaaa	24660
atgtaagcta ttcccttttt cctggaatct ttaagagtcc cagcttttca atagtcaaaa	24720
tgtagatgat tgatatcatt tcttatatga atagcactgg tttgtagttc agcacgcaca	24780
gtgagctggg cacgcccacc tgatagtata gcagagaact tgtttacatt ctttttacat	24840
tcactttcta aaacctgggg tgctctctct ctctctctct ctctctctct ctctgtgtgt	24900
gtgtgtgtgt gtgtgcgtgc acgtgcgcgt gtgtgtagag ggggagagag agagagagag	24960
aactgtgaac tgtgaaatat aacacagcca gcagctttgg gtctcaatcg tagacttact	25020
cttaaggaaa ttacagaat ggaaaggtca tgttcaagta gtttattaac attttgagat	25080
gtaggaaatt aatcccggag tacagaagaa caatttcaga cttcctgaat aaaaacagac	25140
agcatagaga gtggatgata gctaaactct gaatatcttt tgagaagaaa ggcactccca	25200
tttcaggtgc ccataatatg gatttgattt tagtgattaa aacattaatt ttcaacttgc	25260
atctccctgt gtggaagagt tcaatttgtg tgaggggtct cgcctatcca acaaaagtga	25320
atatgtccct ttataggggt aattgctaac ttgtctcaac ttgttttcaa acaattgtta	25380
tagagcactc agtttccact aattgcaaaa ttgttgctta attgaaggac tctcagccat	25440
ctagtgcagc cattcagcca ctggcaggct ctgtgatctc aaactgtgaa ttgcatttta	25500
aagaggaatc gaggagagaa ttctgtggaa ttctaggttt taagtgtctg ctgttgttca	25560
atggaagagg aaatcatttg aacaagaatc gcatcaagtt gtgttgtgat aaattttctt	25620
tattaggatg aataacatgc acagatgagc ttcaaaagtg aatgagcaaa cttactggtt	25680
acactctgca tccatttact ctgttttagta tggagtaatg ttaggcaata aatgatgctg	25740
gcaaataaaa tccgtatgtt atttgcatgt ggtattttaa cctaggaaac atagagtggc	25800
tttgggtattt gtaggcttag tcatgtgtgt cctaaacgtc ctcttaaact tctacttaag	25860
gcatagaatt atttaatcct aaataatttt atacttaagt gcctcactgg atttccagaa	25920
tatttacact gtaaagattt agaaaggtca tgaacccaat tattgactat atggaatcat	25980
tattgatggc agatgcaaaa tggagctcac taatgtactg acattgaaaa ccttttgcag	26040

gggagaggag	ggggagtggg	aaatgtgtgt	gttcctttaag	tggaacagga	aggtattctc	26100
ttttctgtag	aaaaatttga	gtatctgggc	agataagtgt	ggaagctttc	atttaaatta	26160
agtatttaag	ttcaagtaga	agctctaggg	cacttatcct	cttgatgaga	caaatcttat	26220
caaataact	agatgctaag	aagtggctca	ttgccctgat	gtctcattta	tagattgatg	26280
tttgaggatg	ggttgcatta	agtgagttag	ggggctgagt	gtgggacagg	agaacgattg	26340
gaaggaagca	aagtaaattt	acaagcttta	gtgacagcca	taataaagta	aaagtttatt	26400
tccagagagc	ctagagagta	aggaacgtta	tatagttttc	cccaaagggt	cacttgaaag	26460
aacttttcat	tggttgtcat	ggtagtaatg	tcctgatttt	gaaatctccc	agaacctagt	26520
agctcttaaa	catgctttca	tcttggttcc	tttggctctga	cggaaacttt	atgacgaccc	26580
tctgtgtttt	tgacatgcct	ctgcattttt	ggagagagga	ggtcaggcaa	gggaggattt	26640
cttaaaacta	agacagtata	gtaaggaaac	ataaaattat	atgataaaaa	atcactgaac	26700
ttcaaattga	cttactgaaa	taaaacctag	aaggcaacct	gtcgtttaat	tacaactagc	26760
ttgtataaaa	ttaaaattta	taaaatggga	attcaaagaa	aataaacggg	cagttccaag	26820
taatttaagc	aactcaccaa	aaattgaagt	aatagtgcca	cctagagaac	aaaatcacca	26880
gctttactag	ccaaatggct	tattttcata	tgaaccattt	ttccaacgct	acagttacta	26940
ggatttcctt	gttaccatat	tcagatcttg	tgagtgtgta	tgggggtggg	ggttgcattg	27000
ggaattacag	atgaaatttt	aaaacaagca	gatccacaat	ttgatatatg	cactaaatcc	27060
ttttaacgtt	gtaatgtagc	caaatgtaga	atagcatgcc	aggaatcaac	ggctagcatc	27120
ctttttaaca	tttattattt	tcatggatat	gtaccaaacc	gaaccattga	gtataaagg	27180
tctgatttta	tttatttgct	acaggcaatt	cattatactt	tctgagatac	aataacacca	27240
aataatttga	gtagagagac	ctttaagaat	gttttcgatt	tatgatctac	ctttaacttt	27300
aatgtactca	gaagatgtga	gaataaaata	aagtcaaata	taagcaagat	tttaaacaca	27360
cacacaaaaa	acaaacaaac	aagaaaaagg	aagaaaatta	taaggattgc	cttaacctta	27420
gaatagatga	aggtatacat	ctgagccagc	accaaaaaaa	aaaaaaaaaa	aagttatgga	27480
accaggaacc	aataattaca	aattgactta	aaattcttgg	atgacaaaaa	tctatattta	27540
gttcattttt	gcatgcgccc	acaacagcat	ccaaaacagt	tctggggagg	cactttgata	27600
aatgttgctg	aatgcactaa	tagattgatt	aatggctgct	tcagattatc	actagtgatg	27660

tagacagaaa	cttcatgaaa	atggtttgtc	ttgctggaag	aaaggcagaa	attggaggaa	27720
aaggtttaaat	aatatTTTTc	cccagtacct	attataaaag	tcatttagtt	ggcttagttc	27780
tataatttct	tatgtgtaat	ttgattcact	tatgaaattg	tgaatatatg	aaatgttaaa	27840
gttgatttag	acagcaacta	taagcttggtg	gattttcttt	taaatgtcct	caaattttta	27900
aatgccagtg	gagatgccag	cgactgtgct	tcaggaggta	gaatatagta	tatcttaaat	27960
ttgtgccaat	ttctggtaag	cagagaaaaa	attgcatgat	aaccaaagaa	agtcataattg	28020
tttgtgcttt	gtgttattca	tggaagcaat	caggtgcaga	aaactttctt	tttcagaaaa	28080
aaaaaattac	taaaataaag	gtgcgtgtgt	gtgtatgcac	atatatctaa	agggagagag	28140
ggagaaggaa	acttactaaa	taaaattttt	gccacatggg	atttagtcta	atcagtcttg	28200
gttttggagt	tgctatcatc	agtagttcca	ttttgtgatt	ctttctttct	gccttcatgt	28260
gcctttgaaa	actgaaacta	tgcccaaatt	aaaacaagtt	tttctgtcct	ttcacatggt	28320
cacttatttc	ttgaatgtgt	ttttaaacac	agacaaactt	cttttacatc	atgtagaatc	28380
tgaaggtcga	gaaatttgca	gtcattttgc	tggagagaga	tgcttggcgg	agtcccaggc	28440
cacattccta	ggccaaactc	tcgaagggtat	tcctcttatg	caacattggg	aaaatacatc	28500
cagcaccgac	atgttggctg	ataatgtgtc	tgaaggcaca	gacgatatgc	ttatcatatg	28560
aaacataaag	ccagcagata	ttgcagacat	tctgttgaat	gatagaatct	ggatcattta	28620
catttactta	aatgtaaaat	actatgatta	agtacaaaaa	aatcaattta	ggagagaata	28680
gagagtgcg	ggcacggttt	taggggatga	cttatcagca	gattgtagaa	aggaagcttg	28740
aatgttttaa	attaactgca	agttcagtat	aagccagtgg	tgtgacaaga	ggctgttatc	28800
atagctactg	aaattttggg	ctgcactgct	agaaatataa	tactgaaatg	gagaagctaa	28860
taattcttca	ctttttaaat	agactgtatc	tagaatatta	tcatcagttc	aaggaaatga	28920
aataagttgt	tttaggtaca	tcatcgataa	attagtgtac	attcaaatca	ctgtgaccag	28980
gatgcatagg	gaatttgaaa	gcattgcatg	tgagcaatgg	ttgaggggac	ttggaatgca	29040
tgacttaggg	acaagaaaac	ttaggctgga	atggcaagtg	gtttttgaat	gttgggttga	29100
gaagaattct	aaaactgtga	aggattagta	aaaataacat	tcagattgct	aatgcctact	29160
gtggctggga	gattagagtg	tcaacatgtg	tgatgtattt	ttgacatcct	tattttgagg	29220
atgggcttca	aagatttgac	gaactgtcat	aagtgttaatt	tgtgttgctt	cagacagcag	29280

ttctagaacc aatgatgtaa atttagatac tctacatggg agttagaaaa ctttccatta	29340
atttaattta gcaaataattg aatgctcact gcatacagag cactttatta gaggaatata	29400
taataaagaa aaaagaggctc tgggtgtgggtg gctcatgcct gtaatcccag cactctggga	29460
ggctgagggtg ggaggatcac ttgagcccag gagatcatta cgagtctgga caacatagca	29520
agaccccatc tttacaaaag acaaaaaaaaa ttatccaagc ctggtgacag gcacttgtag	29580
tcccagctac ttgggagggtc gaggcagtag gatcgtttga gcccaggagg ttggggctgc	29640
agtgagccgt gattgtccca ctgctcttca gcctgggtga cagagtgaga ccccgttgga	29700
gggaaggaag ggaaggaagg aagaaaggta ggaaggctga aatgaaaccc ttttagaaat	29760
gacactaaaa tgggagggttg gagtaaggta ttttctgaag tgcctttgta cttgtttttt	29820
tctaatacat tggccataag tctgctcctt atttatagtc cataaacaat cctaatagaga	29880
acagttatat atttctgcct ttgaatcatc tacttgaagt gtttagcatc atgaattgag	29940
tatcagaaat ccctcccatt tctttgcaaa gcgctgtatt ttacttttcc ttatttgtat	30000
acagattctc aaaattggct atttttcctt tgggttagac agaacagaat gtctggaaaa	30060
aaaagtctt atcaaattca ggtgccccaa ttgcttaaga aattaacttt tgagggtata	30120
tttttttagg gttcagtagc taaactaaga aaacttctca ccgttcacct tcacttttgg	30180
aaaccacaaa atcttcagat attacagttt tccaaagagt ttctcttttt aaatataaac	30240
taaaaggaat tgactcctcc cccaactccc tcaggcctca gcatggatag agttactttt	30300
tttctttaat aatttattta taacttattt tgctcttctg tagaacagct ggagattaag	30360
caacatggcc atgacataaa atgcaagtta gaccataaga tgagcagccc actccaagta	30420
tgaatgagta cttattcttt gtgatctctc atactgcttt taggtattaa tagtgtcagt	30480
cagcaaagca aacagtttaa tatttacatc tccttttagga tatcatatag tttatagttt	30540
gtatgtgttc ttgcgtgtat gttttcttct gtttcaaatt cttttttctt aaagtaagaa	30600
tgttatatgt agcaaattgg tctttcatta attcatttgt taattcaccg tgcattaatt	30660
gagtgccag tgagtgtcag gcactgggct acttggaatt cttttccctg tattgatcca	30720
tatatcttca ggtgctcctt gatatggctc tccattgact ctctcatata aggctttcat	30780
tacctattca catactccct cccaagagg actggtccaa aagtaaaatc ttgagcaagt	30840
tctctgagtt atttcagaac ttcttgcccc caaactgtat ttttaattat tgactggtag	30900

cattttggaa taacttacct ctctttttta aagattgaag ttcttttatcc tctctgattt	30960
tcaggtgtca gtttgtgtac aaattgagac aataaaaatg tttgttagac attttcttaa	31020
agcattttgc tatgtgagaa tctttcatga agaactcttt ttaacaatga ctctatagca	31080
gaagccacag tagagggaga actactgaat caaagatggg gtttgagtct atgattttat	31140
gatggatttt tttttttttt acatacaagg atatgcatgg gtcttttagt attcaggaat	31200
ctgttcttca cttgacagta ttataaaatt gtgtgtttcc ccctaaaaaa acttaaattg	31260
tgagaatgct tccatttact agaagttggg taatgattat gccaaataag gaaaataaga	31320
cagaaaaatc agtttagtga atactatttt gccttttaat ttagtaattt agtaacagta	31380
tctctttggg gtttactaga aaccactttt taatccaata ggtctctttc attgtgaagt	31440
caggaggtga ttttgcttaa atgtgtagta taggaatcta tatgtgggtg tcaaggatca	31500
tgtaaatatg ctgatataat cggagcacag tttggcatca ttaactcaga aatatttaaa	31560
ctcttgctat acaacatgga agcaaacttg tgcatagttt gtgtgtgtgt gtgtgtgtga	31620
atctcaaaaa aagaaaaaaaa tcacaggatc aggaagtcgg aatagggtccc acttttcttc	31680
tagtaccaa cctacagcca tgttcctagc cttctctttt actcccaagc aagacagaca	31740
ggcaaatgac catcctgctg cccatttctg tgtatattca cttgcattga gagttgtatt	31800
cacctgcttg ttgagagtat tcacaaatgg tacctgataa agtagatact tctttaaaca	31860
tgtgaatttt tttgcattgt ataatgttta gaaataatca tgtataaatg gttgaatatt	31920
aatacaggat tgccttatca agtattttat taatcattaa aatgtgggtg cattaataca	31980
atttatttta agtgcttttc ctaaaatacc agattatttt tctgattttc acatccctga	32040
caatgacttt cttaaacttg gtagccagga acagaaaacc taacactgca tgttctcact	32100
cataagtggg agctgaacag agagaacaca tggactcagg gaggggaacc acacacactg	32160
gggcctggag cagggggcag gggaaggag agagtgcgtc aggacaaaca gacaaatagc	32220
taatgcatgc gagccttaat acctaggtga tgggttgata ggtgcagcaa accaccatga	32280
cacatattta cctatgtaac aaacctgcac attctgcaca tgtatcccgg aacttaaagt	32340
aaaataaaaa ataataataa aataaaataa acttagtagc atctattgtt ccagagcctg	32400
taattgctct tcaggcagtc tcacataaaa acctaggaga accttcactg tcaactgttc	32460
atgaggtgtt aggaaaactt gctctactgc agtgccccag taggcattgg tactgagacc	32520

aaaattcagc	tggtttgttg	ttactacgat	tcctacgtga	tttcacttgt	catgtagaca	32580
agattgcaca	cttcaataat	aatcttgtcc	aaatgtgtgg	tattccatac	atttttaaaa	32640
tgcatlcaca	tatctcattc	catttgatcc	tacaaataac	tctataaaaa	agattggcag	32700
acattatttc	tatataacag	aggaggaaac	tggagcttag	agaagctaaa	tagcaatcca	32760
aaatgcacag	ctgtagaacc	agagcaagga	tgatagccca	gtgacttcac	ctaacctagt	32820
ccccttacca	ccactccagc	tgtctataac	caaaacctgc	agtattcaag	taagaaacca	32880
tatcttgccc	ttgatgcatt	aatgtgagac	ctggagcagg	aacaggctga	tattgtcacc	32940
ctggcctact	gtccaccttt	gtctccagca	gagactggta	cccttctgtg	tgccaaggaa	33000
taaagtggta	atgggaagat	taaaaatgtt	ttttccaagg	agttttttaa	tttaattttt	33060
ttaaaaaaga	aaaaactctt	agagggaaaa	atgaatatat	gacttttgat	gtattgttcc	33120
ttagtaactt	agttataatt	ttacttaaac	ctgagactct	tgctaagtga	atgattagaa	33180
atattaggtg	gctggccaga	tggcaaatag	gaacagctgc	agtctgcagc	tcccagagag	33240
atcaatgcag	aaggtgggtg	atttctgcat	ttccaactga	ggtaccgggc	tcatctcatt	33300
gggactgggt	agacagtggg	tgcagcccat	ggagggtgag	ccaaagcagg	gtggagcatt	33360
gcctcactca	ggaagcgcaa	ggggtcaggg	gaactccctc	ccctagccaa	gggaagccct	33420
gagggactgt	gccatgaggg	acagtgctat	ctggcccaga	tactacacat	ttcctacagt	33480
ctttgcagcc	ggcagaccag	gagattccct	tgggtgccta	caccaccagg	gccctggggt	33540
tcaagtacaa	aactgggtgg	ccatttgggc	agacaccag	ttagctgcag	gagttttttc	33600
tcatacccca	gtggcacctg	aaattccagt	gagacagaac	cattcactcc	cccggaaagg	33660
ggctgaaggc	caggcagcca	agtgatctag	ctcagcagat	cccaccccca	tggagcacgg	33720
caaggtaaga	tctgctgggt	tgaaattctc	actgccagca	cagctgcctg	aagtcaacct	33780
gggatgctcc	agcttggtcg	ggggaggggc	atccgccatt	actgaggctt	gagtaggctg	33840
ttttcctctc	acaatgtaaa	caaagccact	gggaagtttg	aactgggtgg	agcctaccac	33900
agctcagcaa	agcccctgta	gccagattgc	ctctctagat	tctccctctc	tgggcagggc	33960
atctgggaaa	gaaaggcagc	agcccagtc	aggggcttat	agataaaact	cccatctcat	34020
gggacagagc	acctgggaga	gggggtggct	gtgggcccag	cttcagcaga	cttaaattgtt	34080
ctttgcctgt	tggctgtgaa	gagagcagtg	gatctcccag	cacagcactt	gagctctgct	34140

aagggacaga	ctgccttctt	aagcaggtcc	ctgaccctcg	tgattcctga	gtgggagaca	34200
cctcccagca	ggggctgaca	gacacttcat	acaggagagc	tctggctggc	atctgggtggg	34260
tgcccctctg	ggacaaacct	tccagaggaa	ggaacaggca	gcagtctttg	ctgttctgca	34320
gcttctgctg	gtgataccca	ggcaaacagg	gtctggagtg	gacctccacc	aaattccagc	34380
agacctgcag	cagagggggc	tgactgttag	aaggaaaact	aacaaacagg	aatagcatca	34440
acatcaacaa	aaaggatgtc	cacacgaaaa	ccccgtacaa	aggtcgccaa	catcaaagat	34500
caaacataga	taaatccaca	aggatgagga	aaatccagca	caaaaaggct	gaaaattcca	34560
aaaaccagaa	tgcttcttct	cctccaaggg	agcacaactc	tttgccagca	agggaacaaa	34620
actggatgga	gaatgagttt	gatgaattga	cagaagtatg	cttcagaaaag	tgggtaaaaa	34680
cagactcctc	caagctaaag	gagcatgctc	taaccaatg	caaagaagct	aagaaccttg	34740
aaaaaagggt	ggaggaattg	ctaactagaa	taactagttt	agagaagaac	ataaatgacc	34800
taatggagct	gaaaaacaca	gcacgagaac	tctgtgaagc	atacacaagc	ttcaataact	34860
gaatcgataa	agcagaggaa	aggatatcag	agattgaaga	tcaattttaat	gaaataaagc	34920
atgaagacaa	gattagagaa	aaaaagaatg	aaaaggaagg	aacaaagcct	ccaagaaatg	34980
tgggactatg	tgaaaagacc	aaacctacat	ttcattgggtg	tacctgaaag	tggcggggac	35040
aatggaacca	agtttgaaaa	cactccttag	gatattatcc	aggagaactt	ccccaaacta	35100
gcaagacaag	ccaacattca	aattcaggaa	atacagagaa	caccacaaag	atacccttca	35160
agaagagcaa	accaagaca	tgtaattgtc	agattcacca	aggttgaaat	gcaggaaaaa	35220
aagttaaggg	cagcgagaga	gaaaggtcgg	gttaccctaaa	aagggaagcc	catcagacta	35280
acagtggatc	tctcagcaga	aaccctacaa	gcctacaagc	cagaagagag	tggggggccaa	35340
tattcaacat	tcttaaagaa	aagaattttc	aaccagaat	ttcatatcca	gccaactaag	35400
cttcagaagt	gaagtagaaa	taaaatcctt	tacagacgag	caaagtctga	gagattttgt	35460
caccaccagg	catgccttac	aagagctcct	gaaggaagta	ctaaataagg	aaaggaaaaa	35520
ccggtaccag	ccactgcaga	aacataccaa	attgtaaaga	ccattgaaac	tatgaagaaa	35580
ctgcatcaac	taatgggcaa	aataaccagc	taacatcata	atgacaggat	caaattcaca	35640
cataacaata	ttaaccttaa	atataaatgg	gctaaatgcc	ccaattaaaa	gaccacagac	35700
tggcaaattg	gataaagagt	caagacccat	cagtgtgctg	tgttctggag	acccatctca	35760

catgcaaaga	cacacatagg	ctgaaaataa	agggatggag	gaagatctac	caagcaaattg	35820
gaaagcaaaa	aaaaagcagg	ggttgcaatc	ctagtctctg	ataaaacaga	ctttaaacca	35880
acaaagatca	aaagagacaa	agaaggccat	tacataatga	taaagggatc	aattcaacaa	35940
gaagagctaa	ctatcctaaa	catatatgca	ccaatacag	gagcaccag	attcataaag	36000
caagttctta	gagaccaca	aagagaccaa	gactcccaca	caataatagt	gtgagacttt	36060
aacaccccaa	tgtcaatatt	aggtcaacga	gacagaaaat	taacaagcat	attcaggatt	36120
tgaactcagc	tctggacca	gtggaactaa	tagacatcta	cagaactctc	caccccatat	36180
caacagaata	tacattcttc	tcagcaccac	atcacactta	ttctaaaatt	gaccacataa	36240
ttggaagtaa	aacactcctc	agcaaattgca	aaagaatgga	aatcataaca	aacagtctct	36300
cagaccaaag	tgcaattaaa	ttagaactca	ggattaagaa	actaactcaa	aaccatacaa	36360
ctacagtgga	aactgaacaa	cctgctcctg	aatgactact	gagtaaataa	caaaaagaag	36420
gcagaaataa	atacattatt	tgagaccaat	gagaataaag	atacaacata	ccagaatctc	36480
tgggacacag	ctaaaacagt	gtttagggga	aattcatagc	aataaatgcc	cacaggagaa	36540
agcaggaaag	agctaaaatc	aacactctaa	catcacaatt	aaaggaacta	gagaagcaag	36600
agcaaacaca	ttcaaaagct	agcagaagac	aagaaataac	taagatcaga	gcagaactga	36660
aggagattag	agacacaaaa	aacccttcaa	aaaaatcagt	gaatccagaa	gctggttttt	36720
tgaaaagatt	aacaaaatag	atagaatgct	agccagattg	ataaagaaga	aaagagagaa	36780
gaatcaaata	gacgcaataa	aagatgataa	agaggatatc	accactgatc	ccacaaaaat	36840
acaatctacc	atcagagaac	actataaaca	cctctatgca	aataaactag	aaaatctaga	36900
agaaatgaat	aaattcctgg	acacatacac	cctcctaaga	ctaaaggaag	aagtcaaatt	36960
cctgaataga	ccaataataa	gttctgaaat	cgaggcagta	attaacagcc	taccaaccaa	37020
aaaaagccca	ggaccagacg	gattcacagc	tgaattctac	cagaagtaca	aagaagagct	37080
ggtaccattc	cttctgaaac	tattccaatc	aatagaaaag	gagggaatcc	tcctaactc	37140
atthttatgag	tccggcatca	tcctgataca	aaaacctggc	agagacacag	caaaaaata	37200
aaattgtagg	ccaatatccc	tgatgaacat	tgatgcaaaa	atcttcaata	aaaaactggt	37260
aaactgaatc	cagcagcaca	tcaaaaagct	tatctaccat	gataatttgg	cttcatccct	37320
gggatgcaag	gctggttcaa	catatgcaaa	tcaataaaga	taatccatca	cataaagaga	37380

accaatgaca	aaaaccacat	gattattttca	atagatgcag	aaaaggcctt	tcataaaatt	37440
caacagccct	tcatgctaaa	aactctcaat	aaactaggta	ttgatggaac	atatctcaaa	37500
ataataagag	ctattttatga	gaaaccacaca	gccaatatca	tactgaatgg	gcaaaagctg	37560
gaagcattca	tttgaaaacc	ggcacaaaaac	aaggatgccc	tctgtcacca	ctcctattca	37620
acatagtatt	ggacgttcta	gccagggcaa	tcaggcaata	gaaagaaata	aagcatattc	37680
aaataggaag	agaggaagtc	aaattgtctc	tgtttgcaga	tgacatgatt	gtatatattag	37740
aaaaccccat	catctcagcc	caaaatctcc	ttaagctgat	aagcaacttc	agcaaagtct	37800
caggatacaa	aatcaatgtg	caaaaatcac	gagcattcct	atacaccaat	aatgacaaac	37860
agccaagtca	tgagtgaact	cccattcaca	attgctacaa	agagaataaa	atgcctagga	37920
atacaactta	caagggatgt	gaaggacctc	tttaaagaga	actacaaacc	actgctcaat	37980
gaaataagag	aggacacaaa	caaatggaag	aacattccat	tctcatggat	aggaagaatc	38040
aatatcgtga	aaatggccat	actgccccaa	gtaatttata	gatccaatgc	tatccccatc	38100
aagctacat	tgactttctt	catagaatta	gaaaaaacta	ctttaaattt	catatggaac	38160
caaaaaacag	cccgtatagc	caagacaatc	ctaagcaaaa	tgaacaagct	ggaggcatca	38220
tgctacctga	cttcaaacta	tactacaagg	ctacagtaac	caaaacatca	tggtactggg	38280
acataaacag	atagatagac	caatggaaca	gaacagaggc	ctcagaaata	acgccacaca	38340
tctacaacca	tctgatcttt	gacaaacatg	acaaaaacaa	gcaatgcaga	aaggattccc	38400
tatttaataa	atgggtgtcgg	gaaaactggc	tagccatttg	cagaaaactg	aaactggacc	38460
ccttccttac	acgttatata	aaaattaact	caagatggat	taaagactta	aacataaaac	38520
ataaaacat	aaaaacccta	gaagaaaacc	taggcaatac	cattcaggac	ataggcatgg	38580
caaagacttc	atgactaaaa	taccaaagc	aatggcaaca	aaagccaaaa	ttgacaaatg	38640
ggatctaatt	aaactaaaga	gcttctgcac	agcaaaagaa	actaacatca	gagtgaacag	38700
gcaaccgaca	gaatgggtga	aatTTTTTgc	aacgtatcca	tctgacaaaa	ggctaatatc	38760
cagaatctac	aaggaaccta	aacaagttta	caagaaaaaa	aacaacccca	tcaaaaagtg	38820
ggcgaagggt	atgaacagat	gcttctcaaa	agaagaaatt	tatgctgcca	acaaacatac	38880
gaagaaaagc	tcatcatcac	tggtcattag	agaaatgcaa	atcaaaacca	cagtgaagata	38940
ccatcttatg	ccagttagaa	tggcgatcat	taaaaagtca	ggaaacaaca	gatgcaggag	39000

aggatgtaga gaaataggaa cactttttaca ctgttggtgg gagtgtaa	tagttcaacc	39060	
attgtggaag acagtgtggt gattcctcaa ggatctagaa ccagaaat	at cttttgaccc	39120	
agccatccca ttactgggta tatactcaaa ggattataaa tcatgctact	ataaagacac	39180	
atgcacatgt atgtttattg tggcactatt cacaatagca aagacttgga	accaatccga	39240	
atgccccatca atgatagact ggataaagaa aatgtgacac acatacacca	tggaatacta	39300	
tgcagccata aaaaaggatg agttcatgtc ctttgcaggg acatggatga	agctggaaac	39360	
catcattctt ggcaaggtaa cacaggaaca gaaaaccaa caccacatgt	tctcactcat	39420	
aagtgggagt tgaacagtga gaacacatgg acactgggag gagaacatca	cacactgggg	39480	
cctgtcaggg tgtaggaggc taggggaggg atagcattag gagaaatacc	taatgtagat	39540	
gacaagttga tgagtgcagc aaaccacat ggcatctgta tacctaggta	acaaacctgc	39600	
acgttctgca catgtacccc agaacttaaa agtattatta ttattattat	aataataata	39660	
ataaaagaaa tacaataaaa tagaatgcag catacagcag tgattctcaa	acacattcag	39720	
catcagaatt acccttgaat ctttaaaata tatatacata tgagatctta	gtctccaaga	39780	
tttgtaagtt tgggtattggg tccctgggcc tatgttgggt ttagaaactt	ctacagatgg	39840	
tttggatgta tgggacagtt taagaatcgc tgaactaaaa tcaaataaac	tgaatatcct	39900	
gtgatttaga gagacttatc gtttatttca ctatccaagt acttgcatta	gagcgtggct	39960	
agaagggatt tgcagccttg taaataatca gaaattcaga cattttgaga	tgagagaact	40020	
gctgaagatt ttattctgac ttgaaataaa ttttctaatt agaaacttcc	aggtgagagc	40080	
aaaggcctgg aacaatattc ctgagccaga ggaggatcga gtttgactcc	aggcctaaca	40140	
cttactaggt ctatgacctt gggtcagtaa tttaaattct ctgtatctca	acctctcaac	40200	
agggtattgg tagggattaa atgtgttagt gtctgtgaag tgcttagagc	agtgcttggc	40260	
atagtaa	atg cttaatgaat ttcagccact gtttttattt ttagtacttt	ccagctcccc	40320
caaaaagata ctttttttag acttgtatta agacaataaa aagtttaatc	agcatgcttc	40380	
atacctaaat atgcttcact ttatagcaaa gtttacaaga ctaaaactgt	tttgttgtaa	40440	
ttctctgagt ctcattgtgtt tattaatgat ttttctgct gtttattcat	ctgaattcta	40500	
ctcattcttc aagacctagc tggaatcctg tttctagaaa gactcttgcc	cataataata	40560	
aacctgccct atctgagttc ctaggtggtc tgtacctcat aatttggtaa	ttaattgtat	40620	

atgcacttat	ataacaaaac	attatttgtgt	gtctttgctg	tatcagattc	taggctggaa	40680
gtttagata	tgatgttttt	gtctagaaaa	atgttctaga	atgtcctact	caggacagtc	40740
tgttgacttt	aaagacacat	ttcctaaca	gacacttcac	gaggcagccc	cagcctgtac	40800
ctgtgttcct	ggacctgatg	atcaagtttg	atttaagcct	caccacttac	tagctctgtg	40860
attttgggca	agttacttga	atttctctgtg	tgtagataga	acaatgttga	gggaaatccc	40920
tttcccccat	ccttgtgttt	ccacaaggga	acttgcttcc	taataagtaa	cactttcagg	40980
ggaatattct	aggcccttct	cttatcccca	ttacttgttc	tttctgtgaa	aagaggagag	41040
gttaatctga	tggatgaaat	ccttaatctt	tcattcttctg	gactgtagag	cctgtgaacc	41100
aaagcaatgg	accacttgca	ctgaaattga	ggctgaccct	gtattttgat	tcttatttgg	41160
caacttattt	ctattctgtt	cccaattcaa	aatcccaagg	ggagaaggaa	gataattgat	41220
taccagaagt	atgtaatggg	ggtaggaagt	tgaataaatg	gtaacttttt	aaaagttgca	41280
tgagatatag	tccttatccc	agagaagcta	agtttgcttt	tctttcctct	catgtatttt	41340
agtattattt	ctacaattag	attgtaaacc	ctttaaaagc	aagaatattt	ctacattttc	41400
ttactcctga	tagcacacag	tagactgctg	ggcacatata	tagtaggtgg	ctctgtaggt	41460
acttgctaaa	tgattcaaca	tgtttttccc	tcattggaaaa	gaaagatttc	agtattgttc	41520
ttatcagcta	ggaaggcact	ctgaatagga	aatcagttct	aggcaggtat	ccataaatgg	41580
gttatgattt	ccaacttact	tgccccagag	gctcgctaata	gttgaactct	tcattgggtac	41640
tttgtcttgc	ttcatgagct	atacatgcta	aggggttagc	agatcatata	atcttttgat	41700
ctacaaaata	tgatctttat	tgaacaaaaa	cttgggcca	aggcctttct	cctttgccac	41760
cttcctccct	cttttcattc	tcttttttgg	gaatgccctt	tgtgcatgtt	agttacagca	41820
tgtaccacat	tgactgttat	tgttggtttt	tgggtctaac	ccacccttaa	cactgcagtc	41880
cccaagggca	gaaattcagt	ctcattcatt	ttgatgtcct	cagtgcctgt	gctcagagaa	41940
tatctattat	ttgaaaaaat	agtgcaaaag	taaatttttag	gagactacat	cacactcatc	42000
taaactgcaa	gtttgacaag	ttgacatcca	aaagaaaggc	tctcctaaat	aacctcgcca	42060
cagaaatttg	ggtgaccttt	gtagctctgg	agaaagcaga	ggcaaaaatg	aaacctaaaa	42120
attatttgtg	ggttttttaa	aaatgttttc	tcattggagta	aaggtctaca	gctgagttct	42180
tttcatatga	gggaatgaca	gaaacacagc	tggttctgac	tttcagcttc	aactgagcga	42240

ccagagctct gctggtgaaa caggaacttg tattgtgccc ctgacgtgca ccttgaaggt	42300
gtcagctcat tgtccctttg ttcacataaa tagtttttta agaattgttt ttgatcttgt	42360
gagcctctaa ctaaattgatt aacctatgcaa agttggccat ttggggtaat actgaagcac	42420
ttctcttgag ggctattgac aggtgggaat gtgcccacct ccttgggtct ctggttttca	42480
tgtcatactt gcaaatcagt gacagtttaa acttggggca atcacttagc aagtctattg	42540
agttaccaag ttaattattc ccactttgca tgaagcaacc ttgaaaatga ttttcctaaa	42600
gcaaagtaca tccaaactca gtaccttctt aataaccttt gctgaatgaa taaatgacta	42660
attcataaaa aatgtaacat atctttaatt cttacttacg ggcagtttaa gcctcttgtg	42720
taagaggagg cctcggcttg agataacata ggatagtaag cctcctagag aaatttctat	42780
atggaaacat ggtctgctat gaagctagaa gtgagaggac attatatattg accattatat	42840
ttggcttcag agcttctcaa catggggccc aaagtcaagg tcccttgttt cattaagagg	42900
aggtccagga gtgcatgaca cccatcagac tactgagacc cagctggaac taggcacctt	42960
gcacaggggc cttgcctaata caaaatagtt cttatTTTTT ctgagttcca agtaactagt	43020
ttcctaacc agtgtctgga tagtagtgcc aagtgggagt accttcaatg aacttcctca	43080
tgaggttatt tctagcctat tggaatgttt cgTTTTtagga gggtgaggaa gggaagtctt	43140
gaatttttgt gcttagttta atgttTgtgat acagctttga ccatccgttt aatgggagat	43200
ctgTTTTcca gatgactata catgtggaaa ggagaagttt tttgagtgtt ttttttaacc	43260
cTTTTtaaag aatggTTTT catttagtct ctacatttg gggtaaaagg tcctctaggg	43320
agacttttca aaagtatttg aagtttgcac ctgatttcag aggtgagttg gaggcctatc	43380
tgtgtatgac agacacatgt ctccaacaac tatatgttca caaggactaa gagccatcct	43440
tttgggtcca tcattcaaca ttgatctcac attcgtgttc gtatcagtat ctttacagt	43500
cgctcccagt tacatctccc taatttccct tagtaggctt cacagaattt gcagtgtatg	43560
caatggcaga tgaccacatg tggagtcatt taaccacatc ttccactgca agtcagccc	43620
ctcttgatgt ctgtttatgt ttagattcca tcttttgga gatttcattc ctctgcacta	43680
tctcagtatc tcagatgctt ttgagactgg gtccttttcc cctcctatgt ttggccatgg	43740
ccaccccctc agggtttgtgt tgtgtttcac agctgctgtt tgtagggttg acctttacaa	43800
tgtacaaagc tctttcccat atgttgacaa tccctgggtgt gatgctgtga gttaggcagg	43860

gtgtgtatac	gtgtcctcat	catattacag	tggttaaggca	acagggtttt	tgaatttgat	43920
cacccatgaa	tttgtctaata	ttgtttggttaa	aaaatggtca	tgtatcagcc	gtttcacagg	43980
gtcagcttaa	tagaaaagtg	gagttaggca	ggaccagaat	tcaggacttc	agccccggt	44040
cccagggact	attctctata	cccaattgtc	ccaccttgaa	tcagtttctt	ctagggaat	44100
atctccaaaa	ctgagatggc	acccacagga	cttcttaatt	gtagtcatta	ccaggaaaaa	44160
caagcaaagg	aactggtgta	aatctctgtt	tttggtgatt	ggtggagatt	tggagattgt	44220
cttgtgtcaa	aagtaaagcc	actagattaa	atgttttgtt	aataaattgg	ttatttttaa	44280
tttaattatt	tgacagttaa	tttacattat	tcaaaaatca	aaataaaaatt	taaaagaagt	44340
ttacactgaa	aagtcttgcc	ccacttatac	cctgctcacc	tcagtatccc	ccaatacata	44400
ccatctataa	ggtgatcatt	tgtattagtt	tcttgtgaat	ccttgatagt	gtgttttata	44460
tagatacagg	taaatatgag	tatgtactat	tatttcccc	ccaccccacc	ctgttttttt	44520
tttgagacgg	agtctcgctc	tgtcgcccag	gctggagtg	agtggcacga	tctcggctca	44580
ccgcaagctc	caccattttc	ccccattttt	aaaacaaaag	gtagtagcca	tatatacact	44640
attttacacc	ttgttttatc	acttactaat	atataccaga	gagctttcca	tcattttgta	44700
catatgcacc	tatatctgtc	aattattccc	agaagtggaa	ttgctgggtc	agcaggaaaa	44760
atcatgtata	attttgatag	gtattgccta	attgtcctgc	acagggcttg	aattgtttgt	44820
actcccacct	ttagtgtatg	agaagacctg	tttctccata	gcctcatcaa	acagagtgtg	44880
tgagattaga	tgagaaatag	gaggtgagca	gtcttttacc	ccatccgtag	tttgtagtgg	44940
gaacactgca	cagttgcaag	agctggtgca	ggtatcagat	tagttccagt	ggaaacgctg	45000
cctcaccatg	gccatgggct	tgcgccagct	ctagtgcac	acacggaatg	gaccacggt	45060
gccacttgca	gaatttcctg	tagcagaaaag	ttgaacatgc	attcattatt	catctaacta	45120
gccatgctgg	atctaaagag	cacaacagtg	tttttttagaa	ccaaaaagaa	aattgtttca	45180
ctacaacaca	ctgtgtataa	ggcttttcaat	gctcttttct	cagctattaa	cattattttc	45240
aggactgagt	tcaagagatg	tatcccaa	cacagggatg	tcttgctaag	cttggaactt	45300
tcatactcaa	gggatgcttt	tttgaggaat	gatttttacac	ttactcaaca	tttgtaatta	45360
aataattagt	actttataag	ataaatttaa	actgtccaag	tacaataata	acattgaact	45420
atgatgcatt	attgctagac	tttttcctta	aagttgcca	gtggtttcct	gcattaggca	45480

aataggggat	catataaaaa	tgccatgatt	tacggcctag	ataacatctc	caccatttga	45540
gcagcatata	ttccagggtca	tccccacata	actccttacc	attctcatta	gaaaggttga	45600
ttcttagtct	tatTTTTtctc	tgaggacagc	aaaaaaaaaa	atcccccttca	gttccactgc	45660
atagaaaagt	gtggtaaagt	gagccgggca	cagtggttat	ttaatttaaa	tggacaatat	45720
tttttataga	atTTTtgacag	ggccactgta	taggggaaag	tcactcctct	tcccccttat	45780
agaagagttg	cacctggaca	gttgcatgta	tgactgtatc	cagtctacac	aagagggtcat	45840
tcctgggcat	aagaatggac	tgccaaaatc	tagctgaaac	accattgaca	aatagacatt	45900
ttcttttgtt	aataatacct	gtgaaggctt	tcataacaga	catttccagt	tttgttctca	45960
ggctccttgc	agctgctcct	ctaaaagtgt	gctctcttcc	aagagctgac	aatggccaga	46020
agcaagggtgt	tctgtctttt	gtgccatcat	catctaactt	gccacacaca	tttgggatgt	46080
cagcctaggt	ataggttttg	tatccactca	gtatggcttg	tgggtctggt	tgcccttgtt	46140
attcatgctg	agggcctctg	ggcatcagtt	tgggtgtgaga	gaaccattc	catgaccctc	46200
cttcctttgg	ctgttttgac	tcgatggctc	ttgttggcac	agtctgtgag	tgtctgatgc	46260
tctatccatg	ccggaccatc	tgttctgctg	tctctgtggt	ctgaagtcgt	tttctgaact	46320
attccttgat	aataaatttg	agatgatctt	gttctacctt	tcttttcaag	tcacatctta	46380
gccccttagc	cacattcccg	aagaacatga	caaatggatg	ggtcacaagt	cacgtagcat	46440
agggtgtcag	accacgaggc	tttgaaggga	ttctgttggg	tgctaataaag	aaagattttg	46500
tgtcaccacg	atTTTTTTta	aaggcatgtt	gacacttagg	ccttaattga	aagcgttctt	46560
actcaagtag	agttgacaga	ggagtatttg	gtagtcgcgg	ttgctggtct	gaagagcatg	46620
tggttctgtt	tcaatgccc	atgagatctt	ctcacgggaa	aatgttctga	catctcaaac	46680
aatgacctt	catgcatagt	tttgacaaaa	taccctatta	agtatgcata	tatggttggt	46740
acctgtggt	aataattcaa	tactggaaac	agagtagcaa	caaagaaaca	ttagggttat	46800
atttaacctc	tgtggaatta	gtgtgtaaac	aaactgctta	tcagaaatgc	tcatatgggg	46860
ctttgtttaa	ataaataaga	aactggcata	tagggctctgc	aggataatttc	tgccaagtag	46920
acctccctca	cattataaga	caccacatct	atgtctgacc	ccatatggaa	agaggcatag	46980
caagccagca	ctggttcata	ttccctctcc	accacataat	gggtatgtga	tcttagggaa	47040
tccaccgaaa	ctctctgggc	ctcagtttcc	tcagctataa	atgggtggata	atcaaattat	47100

ttacctcacc	attaataaat	gttagctatt	atTTTTtatt	aagttttaata	caaagagaaa	47160
cattttactt	atTTTTccag	ctatccagag	catcttccaa	aatcctatca	ccaacaaata	47220
ctgtattgta	tttattatag	caactatgta	aaaatggagt	ccctgtccta	tgcttagatg	47280
aaatatgttg	gtatttgagt	ttgcatgtct	tctataggaa	tcagtgttta	gtgaaaacgg	47340
gtggagataa	acagatgttt	tcacagtcct	gttgttcaca	gtaccgcaa	attgaatgtt	47400
tccatatagg	tgcatcttaa	tggcttaa	gatgcagata	tttcttgcc	agccatatgg	47460
atcttttgtc	atctaagatg	ttaatat	ccttatatt	tatagtagtt	ctggagtaca	47520
gccagtttct	tgaatagggt	ccacatggct	cattatgcac	agggcctgga	aactgcctta	47580
ctcgtgctgt	tgaaatgaac	cgtgacactt	cagaagagct	gggagctggg	gtagagcagt	47640
ggctaggaga	acatattcaa	ttatat	tcctgcatta	agctacaagt	aatgagcact	47700
ttcctgtgct	ttacagttaa	gtaattaaaa	gaaattatag	agtgggatgc	aaaaataacc	47760
cgaaggacaa	ctggatgtgt	ggagccacca	gttttctcca	tgagtgcaca	aggttaatcc	47820
ttgttactac	tcagaatgct	gagtttctac	agaaagggtt	gcagggtccac	acatgttttg	47880
gcgtctaccc	acacgcttct	gtatggcatg	actgtgcatc	ccagaagaag	ggctgtgctg	47940
tgtacctcca	cgtttcagt	gaatttaaca	aactgatccc	tgaaaatgg	ttcataaagg	48000
tgagtaacag	agagctaata	gccttctctt	gctaatttta	tctttcccc	aagatttctt	48060
gataatagtt	tgaaaaggag	tgttattctt	tggtctctag	aggcaactta	cctttccagt	48120
ttcttccatc	acctgttttc	atctctcttg	tttttttaaa	tttaatgctg	tatgtatttc	48180
agaggatagg	atctaata	gtgcggtccc	ttcatcagg	gagaattatt	catctcattt	48240
tcatttttagc	ccttctgaat	taatgacatt	gaagcccggc	agtttggtcc	taagatgggt	48300
ttaattatgt	acagatactc	tttctataat	ggaaattgct	cagataacta	attaaccaca	48360
agaatacact	gtctatggaa	aatttcagga	gcaccgtctg	tggaaaaact	gggaaggga	48420
tgctgtcacc	acagctctgg	ggtctattaa	aagtgtgggt	atgcagcact	ggtgtctagt	48480
ggggtgttgg	ctctcaactg	ccagaattcc	catagcattt	catggcagaa	agtcaagggtg	48540
tccagcaata	ctctgaaagt	gacctgttga	ttaaagtcgt	caattctgaa	gaaagagact	48600
gaaataagac	aaatgggtct	taactttttt	tctctttctc	tctcttgtaa	aaatgtgtga	48660
ttgttctggc	atgttcccaa	tccccacata	atgccaacat	cttttcttaa	agggggattc	48720

cctttatcct	tggatctgag	aattattgca	tgttctccct	ttagggacaa	tgaatgcagt	48780
tgcacacccc	ttgctttttt	tttttttttt	gtacacagca	tgcttattct	tggatgcagg	48840
gacttgaaag	acaaagcccc	acctggcttt	cacaacatct	cctattagta	ggtgtgcctt	48900
gtgtgtaatt	tgaaggaggc	ggtcccttag	ctgtgtttac	actgtacttt	taaatgtggg	48960
gctgaaggta	gaatcaacca	tacttaagat	gccacctggg	aaaatagggt	tctgtgtcat	49020
ctcagcccca	cccatttgca	aatgacttaa	cagcagcact	attagggttc	ctagtgtgag	49080
tcatttgcac	ttggactggg	gaacttgggt	acttcttggg	gtttggaaac	aaacaacctt	49140
tgcagtcttt	cgtaaaaagc	ctgaacagt	gaccagtctc	cagttctact	tgcaaagctg	49200
cccccatcaa	atccctcata	atgttcaact	taaaaaatgt	tacacttttc	tctggaaatc	49260
taaccttttt	tcctttttta	aaagccattt	taagtacttc	agtcttgaat	caaagtatcc	49320
caaatatagg	acaccaacct	agaaattggg	ttacctcctg	ggaactttat	cgaagaagag	49380
agattttggg	tggagagggg	gttttgatgt	ttgatactta	tatttactat	tttaatat	49440
cattgttggt	gttgctgctg	ctgctgtatt	atgttgcgag	tttcgtttgt	ttaaatttca	49500
tggtatttgg	taggagagag	ctggatctgt	tggtttcagg	acaagtctag	aaataagaaa	49560
tctgccttga	gtgagtgagt	tggttccctc	tggttgctatt	tcaccattaa	ggacgaaagg	49620
aactcacaag	gaccagagac	atctggctga	aagcaatact	agtgtgactg	gacatctact	49680
acctgccata	gttggtcata	tcgtttccag	tatgattctg	attgagttag	tgatattagg	49740
ctatgttcag	ggatcaggga	ggctaattat	gcttatattg	ccttgtagca	ttttggtaag	49800
aattaatgat	tgtgtagatg	tccagattta	ggtcagcaat	attctaaaag	ttctcattga	49860
actaatcatg	tttataagta	gcctgtactt	tctatcataa	taacaatagt	ggaaaagcta	49920
gttgacataa	aaggagccca	gattttactt	aagtaaaaac	acaaaagcaa	agatattttc	49980
ccacataaat	tacaaaagca	aagatatatt	cccacataaa	tgtccccata	aaacaagttg	50040
aaccaaagag	gaaagatgac	aggtaaccgt	atgacacgct	aagaaagtat	cataatactt	50100
aagttaactt	caacctttta	tttccttata	ctaagcagcc	tcttttctct	ttatcattta	50160
gtcctgtgct	tctcaacttt	gataagtaaa	aaagttattg	cactaaataa	atcttattga	50220
aatgcaggat	ctgattgagt	gggtggggta	ggtggaatga	gggtggggaa	gttgagattc	50280
tgcatttctt	agaagtttct	actttatgtt	aaaatggcta	atccatctca	acattgagaa	50340

gtaaggtttc acttaatttc agcctgtgta agtttatccc atatgtacat ttcctaaaac	50400
tctaattctca ggccccagga atttctcctt tagttaaaat attttttagga ataaatttga	50460
attgcattaa tacacaattt ataaatttaa cacaaaaaat tattttgaagt ttgagacttt	50520
aggttgcatg aaatcaattt catacttgaa aattttctat aaattcaaaa gtctgtgtat	50580
ttaaatacaa tttaaatacc tgtgttacag tgacatttgt ttttctgtct ctctctccac	50640
catttccaga gtcacatcc ctgtacagaa aaatTTTTcc cacatgattt caccataaat	50700
tcattaaata tgatgcttac ttgataattt ctccaggttc tttttttttt taattatact	50760
ttaagttcta gggtagatgt gcacaacctg caggtttggt acatatgtat acatgtgcca	50820
tgttgggtgtg ctgcacccat taactcgtca ttacattag gtatatctcc taatgctatc	50880
cctccccct acccctactc catgacaggt ccagtggtgt gatgttcccc accctgtgtc	50940
caagtgttct cattgttcag ttcccaccta tgagtgagaa catgcggtgt ttggttttct	51000
gtccttgca tagtttgctc agaatgatgt ctttgctcac tgatggacat ttggttggct	51060
ccaagtattt gctattgtaa atagtgccgc aataaacata cgtgtgcatg tgtctttata	51120
gtagcatgat ttataatcct ctgggtatat acccagtaat gggatggctg gctcaaagg	51180
tatttctagt tctagatcct agaggaatcg ccacactgtc ttccacaatg tttgaactag	51240
ttacagtcc catcaacagt gtaaaagtgt tcctatttct ctacatcctt tccagcacct	51300
gttgtttccg gactttaatg atcgccattc taactgggtgt gagatgggtat ctcatgtgg	51360
ttttgatttg catttctctg atggccagtg atgatgagca ttttttcatg tgtcttttgg	51420
ctacataaat gtcttctttt gagaagtgtc tgttcatatc cttcacccac tttttgatgg	51480
ggtcatttga ttttttcttg taaatttggt taagttcttt tagattctgg atattagccc	51540
tttgtcagat gggtagattg taaaaatttt ctcccattcc gtaggtttcc tattcactct	51600
gatggtagtt tcttttgctg tgcagaagct ctttagttta attagatccc atttgtcaat	51660
tttggctttt gttgccattg cttttgggtgt tttagtcatg aagtccttgt ccatgcctat	51720
gtcctgaatg gtattgccta ggTTTTcttc tagggTTTT atggTTTTag gtctaactgt	51780
taagtcttta attcatcttg aattaatttt tgtataaggt gtaaagaagg gatccagttt	51840
cagctttcta catatggcta gccagttttc ccagcaccat ttattaaata gagaatcctt	51900
tccccatttc ttgtttttgt caggtttgtc aaagatcaga tggttgtaga tgtgtggtat	51960

tgtttctgag ggctctgttc tgttccattg gtctatatct ctgttttggg accagtagca	52020
tgctgttttg gttactgtag ccttgtaata tagtttgaag tcaggtagcg tgatgcctcc	52080
agctttgttc ttttggttta ggattgtctt ggcgatgcgg gctctttttt ggttccatat	52140
gaactttaaa gtagtttttt tccaattctg tggagaaagt cattggtagc ttgatgggga	52200
tggcattgaa tctataaatt accttgggta gtatggccat tttcatgata ttgattcttc	52260
ctacccatga gaatggaatg ttcttccatt tgtttgcgtc ctcttttatt tccttgagca	52320
gtggtttgta gttctccttg aagaggctct ccacatccct tgtaagttgg attcctaagt	52380
attttattct ctttgaaaca attgtgaatg ggagttcact catgatttgg ctctctgttt	52440
gtctgttatt ggtgtatagg aatgcttgtg atttttgcac attgattttg tatcctgaga	52500
ctttgctgaa gttgcttatc agcttaagga gattttgggc tgagatgatg gggttttcta	52560
aatatacaat catgtcatct gcaaacagag acaatttgac ttcctctctt cctatttgaa	52620
tatcctttat ttctttctat tgcctgattg ccctggctag aacgtccaat actatgttga	52680
ataggagtgg tgacagagga catccttggt ttgtgccagt tttcaaaggg aatgcttcca	52740
gcttttgccc attcagtatg acattggctg tgggtttgtc gtgaatagct cttattattt	52800
tgagatatgt cccatcaata cctagtttat ttagagtttt tagcacaaag gctgttgaat	52860
tttgtcaaag gccttttctg catctattga gataatcatg gtttttgtct ttgattctgt	52920
ttatatgatg gattatattt attgatttgc atatgttgaa ccagccttgc atcccaggga	52980
tgaagccaac ttgatcatgg tggataagct ttttgatggt ctgctggatt cggtttgcca	53040
gtattttact gaggattttt ccatcgatct tcatcaggga tattggcctg aaattctctt	53100
ttttgttgt gtctctgtca ggctgtggta tcaggatgat gctggcctca taaaatgagt	53160
tagggaggat tccctctttt tctattgatt agaatagttt cagaatggta ccagctcctc	53220
cttatacctc tggtagaatt cagctgtgaa tccatctggg cctgatggat ttttttggtt	53280
ggtaggctat taattattgc ctcaatttca gagcctgtta ttggtctatt aagagattca	53340
acttcttcct ggtttagtcc tgggaggggtg tgtgtgtcca ggaatttata aatttctttt	53400
aggttttcta gtttatttgc atagaagtgt ttatagtgtt ctctgatggg agtttgtatt	53460
tctgtgggat tgggtggtgat atccccctta tcacctttta ttgcatctat ttgattcttt	53520
tctcttttct tctttattag tcttgctagt gatctatcaa ttttgttgat ctttttaaaa	53580

aaccagctcc tgggttcatt gatTTTTtga aggagTTTT ctgtctctat ctccttcagt	53640
tctactctga tcttagttat ttcttgtctt ctgctagctt ttgaatgtgt ttgctcttgc	53700
ttctctaaat tgtgatgtta ggggtgtcaat tttagatctt tcctgctttc tcttgtgggc	53760
atttagtgct ataaatttcc ctctacacac tgctttaaat gtgtcccaga gattctggta	53820
tgttgtgtct ttgttctcat tggtttcaaa gaacatcttt atttctgcct tcacttcgtt	53880
aagtaccag tagtcactca ggagcagggt gctcagtttc catgtagttg agtggttctg	53940
agtgagtttc ttaatcctga gttctagttt gaaagcactg tagtctgaga ggagtttgt	54000
tataatttct gttcttttac atttgctgag gagtgcttta cttccaacta tgtagtcaat	54060
ttttggaata agtgtgatgt ggtgccgaga agaattgata ttctgttgat ttggagtgga	54120
gagttctgta gatgtctatt aggtccgctt ggtgcagagc tgagttcaat ttctggatat	54180
ctttgttaat ttctgtctt gttgatctgt ctaatatga ccgtgggggtg ataaagtctc	54240
ccattattat tgtgtgggag tctaagtctc tttgtaggtc tctaaggact tgctttgtga	54300
atctggtgct cctgtattag gtgcatatat ttttaggata gttagctctt cttgttgaat	54360
tgatccctt atcattatgt aatggccttc tttgtctctt ttgatctttg ttggtttaaa	54420
gtctgtttta tcagagacta ggattgcaac tcctgctttt ttttgctttc catttccttg	54480
gtagatcttc ctccatccct ttattttgag cctatgtgcg tctctgcaca tgagatgggt	54540
ctgctgaata cagcacactg atgggtcttg actctttatc caatttgcca gtccatgtct	54600
tttaactgga gcatttagcc catttacatt taaggttaat attgttatgt gtgaatttga	54660
tcctgtcatt atgatgttag ctggttattt tgctcgtag ttgatgcagt ttcttcctag	54720
cctcaatgat ctttacaatt tggcatgttt ttgcagtggc tggtagtggg tgttcctttc	54780
catgtttagt gcttccttca ggagctcttg taaggcaggc ctggtggtga caaatctct	54840
cagcatttgc ttgtctgtaa aggattttat ttctccttca cttatgaagc ttagtttggc	54900
tggatatgaa attctgggtt gaaaattctt ttctttaaga atgttgaata ttggcccca	54960
ctctcttctg gctttagtag tttctgccga aagatgctgt tagtctgatg gacttcctt	55020
tgtgggtaac ctgcccttc tctctcgtg cacttaatgt ttttctctc atttcaactt	55080
tggatgaatct gacaattatg tgtctttgag ttactcttct tgaggagtat ctttgcgga	55140
ttctctgtat ttctgaatt tgaatgctgg cctgcctcac tagattgggg aagttctcct	55200

ggataatatc	ctgcagagcg	ttttccaact	tggttccatt	ctcccatca	ctttcaggt	55260
caccaatcag	atgtagat	gggtcttttca	catagtccca	tatttcttgg	aggctttggt	55320
catttctttt	tactcttttt	tctctaaact	tctcttcttg	cttcatttca	ttcatttgat	55380
cttcaatccc	tttctttccac	ttgattgaat	cagctactga	agcttggtga	tgtgtcacat	55440
agttctcgtg	ccatgggtttt	cagctccatc	aggctattta	aggctcttctc	tatgctgttt	55500
tttctagtta	gccattcgtc	taatgttttt	tcaagggttt	tagcttcttt	gctaaaagggt	55560
tcaaacatcc	tccttttagct	cggaggaggt	tgttattact	gatcatctga	agccttcttc	55620
tctcaacttg	tgaaagtc	tctctgtcca	gctttgttcc	attgctggcg	aggagctgca	55680
ttcctttgga	ggagaagacg	tgctctgatt	tttagaattt	tcagcttctc	tgctctgggt	55740
tctcccatc	ttattgggtt	tatctacctt	tggtctttga	tgatgggtgac	gtacagatgg	55800
ggttttgggtg	tggtatgttct	ttctctttgt	tagttttcct	tctaacagtc	aggaccctca	55860
gctgcaggtc	tggtggaggt	tgctggaggt	ccactccaga	ccctgtttgc	ttgggtatca	55920
ccagcagagg	ctgcagaaca	gcaaattattg	cagaacggca	aatgttgctc	cctgattggt	55980
cctctggaag	cttcgtctca	gaggggcacc	tggccgtatg	agggtgtcagt	cggcccctac	56040
tgggaggtgc	ctcccagtta	ggctactcag	gggtcaggaa	cccacttgaa	gaggcagact	56100
gtccattctc	agatatcata	ttccatgctg	ggaggacccc	tactcttttc	aaagctgtca	56160
gacagggaca	tttaagtctg	cagaagtttc	tgctgtcttt	tgttcagctg	tgccctgccc	56220
ctagagggtg	agtctacaga	ggcaggcagg	cctccttgag	ctgcggtggg	ctccacccat	56280
ttcgagcttc	ctggctgctt	tgtttaccta	ctcaagtctc	agcaatgggtg	gacaccctc	56340
ccccagcctc	gctgctgctt	tgcatgttca	tctcagactg	ctgtgctagc	agtgagccag	56400
gctccgtggg	catgggaccc	tccgagccag	gcctgggaca	taatctcctg	gtgtgccgtt	56460
tgctaagacc	attggaaaag	cacagtatta	gggtggggag	tgtcctgatt	ttccaggtac	56520
cgtcagtc	ggcttccctt	ggctaggaaa	gggaattccc	caacccttg	tgcttcttg	56580
gtgaggtgat	gccccaccct	gctttggctc	atgctccgtg	ggttgtaccc	actgtctgac	56640
aagccccagt	gagatgaacc	cgggtaccta	gttggaatg	cagaaatcac	ccgtcttctg	56700
catcactcac	gctgggggct	gtagactgga	gctgttcata	tttgccatc	ttggaacctc	56760
cctttccaag	ttctttatta	cagagtgggt	cactgaaact	tcatggaaca	aattggaaat	56820

tatcttctta	attaatgtca	ctgtctacca	tgtatgggaa	tttggtaa	attatatggt	56880
ttcaataaca	tagtagatag	aacattgtca	aatctaaact	tcagtgaatt	gtaacagatc	56940
ccacctgaaa	ttctaaagaa	aacagaattc	taattgaaga	ggttaaactt	ttacagggaa	57000
tgtcaactgc	catttgggtc	ctgtaaacia	aaaactgttt	tttaaaaaag	taaactttaa	57060
aagtattttc	agatgacctc	atttgctatc	caagtggctt	gagtatgctt	gatgctaaga	57120
cttctttgtt	acagactgga	gatgtgtgct	actggggcag	tgttgctctg	tgacaaggag	57180
gcagaggatg	agggcaaggt	tcgatgtgac	tgtgaattct	gggtggctct	ggctatcggg	57240
agccttcatt	gattacagca	aaacagttgc	tttcctaggg	caatagtgtc	tctgtcacc	57300
aggctggagt	tcagtggcat	gatcaatcgc	tcactgtagc	ctcaacttct	tagactcaag	57360
taatcctccc	acctcagcct	ccaagtagc	tgaaactaca	ggtgtgcacc	accacaccta	57420
atTTTTTTaa	TTTTtaagtt	TTTgtagaga	catggctctca	ctgtgttgcc	caggttgatc	57480
tcgaattcct	gggctctagt	gacccctccg	cctcggcctc	ccaaagtgtt	gggattacaa	57540
atgtgagcca	ctgcacctgg	ccctttgcaa	ccttcttgac	aatgcattcc	tttattccct	57600
aactggaagt	aacttctttc	tctttataaa	attgtatctg	taccttttct	gggtcatttc	57660
tacctttata	ttctagttac	gtatgtccta	cctccctcct	agggaggag	gtaagtaaga	57720
ctggaaagta	gacttcatgt	gtgatgaatg	aatgaacaaa	aggaagtcta	acatatggat	57780
atagtcaact	ggatgcaaat	taaaaatttt	taaatattga	tttgcaagat	ttcattaagg	57840
tcaactctta	atagtttgta	tcatatatgt	taggaaccaa	atattaataa	cttcttcagc	57900
attaccatta	tctttatagg	actgtctaaa	atgagcagcc	atatctttaa	actgtgtttt	57960
ctctgattac	acgctcacag	gtaaaacca	aaggggctgg	gaacaaacaa	gacttttttt	58020
TTTTtctgta	tgctgaatt	atctgtactg	ttgcttggtt	tcccaccttt	ggccatagaa	58080
acttagttct	aacatgctac	aatttttgca	gttctttctc	ttagaaaaag	accacattgt	58140
ctgaaatttc	atccatttaa	gtaatcaagc	cttaaagttg	aaggatcttg	gtcatgatta	58200
atctagacct	acaaagtagt	atcttaatgg	cactcctttt	agaaagttag	gttcaggagc	58260
acacatagct	gcagtgtcca	cattttgtaa	gctccttcgt	tgtcacagcc	actctcttct	58320
ctgtggctga	tattctaaaa	ctggcaacac	atcctgatgg	taaaagcttg	gttcaggaga	58380
caggtgacct	actagcttta	tggcatttga	caggttacct	aacctctctg	acgcataatt	58440

gcctcatcta	tataatgggg	ataataatac	ccatcctgtc	tccttgtaaa	aatcaaatta	58500
gatgacgcct	gtgaatgttc	tatagtctct	tagacaaatg	taagttatga	ctacagcaag	58560
agtaaaagag	catggttgta	tggacattct	ttcagtgaag	tgtctaagac	ttgtgagtca	58620
cacttaaagc	taaacttgat	atctacttca	ttgattttct	tttttagttct	atgtactata	58680
ttgaatttcc	tgacagtggg	gctatgaaag	ccttcctagc	attttataga	tgtggttgaa	58740
ttaatggctg	taagccttaa	agcagaatta	gacagcatca	atgaatttat	taagtataaa	58800
taaatatata	atctgcttag	caatattaca	cagcctcttt	atccttatgtg	tgataaagag	58860
tcatccgaag	gttgaaaatg	aagaattgtc	ctggaagctc	ttacttaatc	ttttattatt	58920
tcctaataca	gtatataaaa	ttactcattg	aaagccttagc	agaataagaa	acaagaagtt	58980
aaaaggctga	aaactacaaa	ttttgctatt	attattgtta	ttacttccca	agtctcttat	59040
tgatctgtta	gaaatagagc	tacacaggaa	attgtaggac	agttagtatg	tggtagtgtt	59100
atctgctttt	taattattca	agtaaggttt	tattccatta	gaggaactca	agaagttggt	59160
catggctgat	aattgctatc	tgtcaaattc	ccttagagcag	ggatccgcaa	caccagggcc	59220
atggattgtt	accagtcctt	ggcctgttag	gaaccaggct	gcacagtagg	aagtgagcgg	59280
cgggtgagca	aacattgcca	tctgagctct	gtctcctttc	agatcagcag	cagcattaga	59340
ttctcataga	agcatgagcc	ctgttgtaga	ctgcacatgc	aagggatcta	ggttggtgtgc	59400
tccttatgag	aatccaattc	cttatgataa	tttaactgat	gatctaaggt	ggaaccattt	59460
catcccaaaa	ccatccaccc	tgctactccc	agaccgtgga	aaaattgtct	tccacaaaac	59520
tggttcctgc	tgccaaaaag	gttgggacca	ctgccttaga	gtttataatt	tgggggttagc	59580
acagcctata	tttacctgag	aatttcaatg	ggttcactga	tctttccaaa	tgaaaaggct	59640
tcttacgaaa	attatatcca	aactgtcttt	tctcttagtt	taataaacct	atcagtaagt	59700
ttttactgag	tactgctatt	acatttttct	ctgttaagca	ttatgggggc	tcagacatga	59760
tccattccct	caaagaactt	acctttcagc	tgaagactga	ctagaatgag	caaatacggt	59820
taacaattaa	caagttagta	ggccagctcg	gccaacatgg	tgaaaccctg	tctctactaa	59880
aaatacaaaa	attagccggg	catggtggtg	ggcgcctgta	atcccagcta	cctcctgctg	59940
aggcaggaga	attgcttgaa	cccaggaggt	ggagattgca	gtgagccgag	attgcaccat	60000
tgcactccag	cctgggagac	agagcaagac	tctgtcttgg	gaaaaaaaca	aaacaaaaca	60060

agtgagaagg	gaatcaagta	ctacgtaaga	tgtaatgtgg	aatttttaggg	aaggaaggca	60120
gtgtatgctg	gagtaattag	agaaaggggc	atgcatgatt	tgatgcttga	actggatcat	60180
gaaggataag	caagattttg	gcagcaagtg	agagggagag	aggagtttgt	cagaggaatg	60240
gaacaagtca	ggaggcagta	atgtgtacgg	cactccaagg	actctgtcct	gatcggagca	60300
gaagtgatga	agcattgagt	agtttgagaa	agttagctaa	gaagggtgag	ggcagatgtt	60360
ggagagagtt	gagtatcaga	gagaagacat	tagatttgag	cagatagaac	aaaaatgcca	60420
ttgccagttt	ttgtgtaaga	gaatagtatt	agggtgggta	cccaggaagt	ggtcatcaga	60480
gttgaatgga	gcagaaagag	tgtctgatac	agcatcggga	cgttggtggtc	acaaaatgag	60540
gtggtgaggg	ctgagccaag	atggtggcag	tgggatggat	aaaaagggat	ggccaggaca	60600
aatattttta	aggaaaaatt	aacaggacat	ctttactgac	tggattggaa	ggctatgcag	60660
gaaatatatt	gtcaaacttg	attccaggat	ttctatccta	tgcctggggt	gccccaaata	60720
tcaggaacc	attgttagaa	aaggtaggag	ataccactgt	tccaacaaaa	agtattgagt	60780
ttggtgttgc	accacttaa	cttcaaggcc	ttacaagtga	gtagacagtt	agttagaatt	60840
gcagaagtgc	cactcagaga	gcagggcttg	caatgtgggg	ttggactttg	tcaccattgt	60900
gttaattcct	aattctatgc	agatgctcag	cttgaggaat	acccatgttt	gggcttcaga	60960
atgaaagcca	agtaatat	actcagatgc	caattttccc	tctgaaatat	ttgctcatgg	61020
aactgagaga	acaatatata	aagcattaat	tatttttctc	ataaagttat	taataaaaag	61080
ataagatcag	tgaaaggcag	agtaaaactag	aagccaagta	tagaaaatgg	tatcattcaa	61140
agactcatta	ctgtagtgg	gaaaacaaaa	caattttcca	acagcttaag	atgcctcagt	61200
attttggacc	atttttaagt	agttagtgtg	ggcacttagt	aaatatgtat	taaactatag	61260
ttcattaatt	cttttttttt	tttttttgag	atggagtttc	actctgggtca	cccaggctgc	61320
atttttgctt	tcttagtgat	atataaaatg	tcgagtttca	caatgatggg	atcttagatt	61380
tgattaaata	tggtattaaa	aaatagctga	tcacagaaag	tctctaccag	tgtgatgtag	61440
atggctaaag	tattccacat	ttgcaaactt	ttattgacct	aaataagagg	tgccccttgg	61500
gttgttttta	tttggactgg	gaatattagg	agaaagcttt	ttcattcagt	gtgtaagtac	61560
aatctaccag	aaatagaaac	ccccatggac	gatctatttc	tttgatggta	caggactcag	61620
aacattcaca	aagatttagt	tgttagcggg	atagacatct	gtatttttatt	caaaccaatt	61680

ttcccttcct aatctgagaa cattgtgcaa tctaagcagt tctaagcatg tttgctattc	61740
gtgcaaagt agagtaaadc taaaagaaat ttttttgtgt gtttagggat ggtaataaag	61800
tctcttagtg gttgaaaatg ttattttcta caaaagtgga gaacatttgc ttttcaatac	61860
cagagttttc agccatttct gcattctgac ctattgactg gaggtagggt gcctttgaat	61920
tcagtaaaac ttcattgggca gaaacacagt tccttttcct acttatttgg atatcatgat	61980
ggccattgca tgtatgtgtc tttttgtaag tccatgcctc agaactgaga agtaggaata	62040
aaattagggt cagggctggg gatgctactc tttgctgctg agaaacacaa tgcttcaggt	62100
aagtgattct gaagtccttc accacctgac ggtaaccttg ggttgggtcca taggtatgtt	62160
ttcatTTTTgc ttgttcatcc attttaattg gcttcctaga gcatgcttgt agatgtagag	62220
ccaaatttag agtagagcaa ccctctggca aacaggaaga gattaatttt gtggtatgct	62280
tttaagggaac ttcccaggaa acttcaaaag cagaaaaaga agcactagct gcctattcca	62340
aaatgtgtaa aacaccactc agctttttta aagtaggata aactcagagc gcgcgcacac	62400
gcgcgcgcgc acacacacac acacacacag agagaacatc tctagtaaaa agaaaagttg	62460
agcttttcta gctagatgtg tgtattagcc agaaaaagcc aaggagtga gggttttaga	62520
gaactggagg agataaagt gagtctgcat atgggaggca tttgaaatgg acttaaagt	62580
ctttttaatg ctgacttttt cagttttctc cttaccagac acattgtttt catgacatta	62640
gccccaggca tagacacatc attaaaatga acatgtcaaa aaatgatttc tgtttagaaa	62700
taagcaaaac attttcagtt gtgaccaccc aggtgtagaa taaagaacag tggaattggg	62760
agccctgagt tctaacataa actttcttca tgacataagg caagtcttct atggcctttg	62820
gtttccttac ctgtaaaaca ggatggctca atgaaattat ctttcttctt tgctataata	62880
gagtatctct gtgggaagag gaaaaaaaaa gtcaatttaa aggctcctta tagttcccca	62940
actgctgttt tattgtgcta ttcattgccta gacatcacat agctagaaag gcccatcaga	63000
cccctcaggc cactgctgtt cctgtcacac attcctgcaa aggaccatgt tgctaacttg	63060
aaaaaaatta ctattaatta cacttgcagt tgttgcttag taacatttat gattttgtgt	63120
ttctcgtgac agcatgagca gagatcatta aaaattaaac ttacaaagct gctaaagtgg	63180
gaagaaggag aacttgaagc cacaattttt gcacttgctt agaagccatc taatctcagg	63240
tttatatgct agatcttggg ggaaacactg catgtctctg gtttatatta aaccacatac	63300

agcacactac	tgacactgat	ttgtgtcttg	tgacagctgga	gtttatcacc	aagacataaa	63360
aaaaccttga	ccctgcagaa	tggcctggaa	ttacaatcag	atgggccaca	tggcatcccc	63420
gtgaaagaaa	gccctaacca	gttttctgtc	ttgtttctgc	tttctcccta	cagttccacc	63480
aggtgagaag	agtgatgacc	atccttttcc	ttactatggg	tatttcatac	tttggttgca	63540
tgaaggctgc	ccccatgaaa	gaagcaaaca	tccgaggaca	aggtggcttg	gcctaccag	63600
gtgtgcggac	ccatgggact	ctggagagcg	tgaatgggcc	caaggcaggt	tcaagaggct	63660
tgacatcatt	ggctgacact	ttcgaacacg	tgatagaaga	gctgttggat	gaggaccaga	63720
aagttcggcc	caatgaagaa	aacaataagg	acgcagactt	gtacacgtcc	agggtgatgc	63780
tcagtagtca	agtgcccttg	gagcctcctc	ttctctttct	gctggaggaa	tacaaaaatt	63840
acctagatgc	tgcaaacatg	tccatgaggg	tccggcgcca	ctctgaccct	gcccgccgag	63900
gggagctgag	cgtgtgtgac	agtattagtg	agtgggtaac	ggcggcagac	aaaaagactg	63960
cagtggacat	gtcgggcggg	acggtcacag	tccttgaaaa	ggcccctgta	tcaaaaggcc	64020
aactgaagca	atacttctac	gagaccaagt	gcaatcccat	gggttacaca	aaagaaggct	64080
gcaggggcat	agacaaaagg	cattggaact	cccagtgccg	aactaccag	tcgtacgtgc	64140
gggcccttac	catggatagc	aaaaagagaa	ttggctggcg	attcataagg	atagacactt	64200
cttgtgtatg	tacattgacc	attaaaagg	gaagatagtg	gatttatgtt	gtatagatta	64260
gatttatattg	agacaaaaat	tatctatttg	tatatataca	taacagggta	aattattcag	64320
ttaagaaaaa	aataatttta	tgaactgcat	gtataaatga	agttttataca	gtacagtggg	64380
tctacaatct	atttattgga	catgtccatg	accagaagg	aaacagtc	ttgcgcacaa	64440
cttaaaaagt	ctgcattaca	ttccttgata	atgttgtggg	ttgttgccgt	tgccaagaac	64500
tgaaaacata	aaaagttaaa	aaaaataata	aattgcatgc	tgctttaatt	gtgaattgat	64560
aataaactgt	cctctttcag	aaaacagaaa	aaaacacaca	cacacacaac	aaaaatttga	64620
acaaaaacat	tccgtttaca	ttttagacag	taagtatctt	cgcttctgtt	agtactatat	64680
ctgttttact	gcttttaact	tctgatagcg	ttggaattaa	aacaatgtca	aggtgctgtt	64740
gtcattgctt	tactggctta	ggggatgggg	gatggggggg	atatttttgt	ttgttttgtg	64800
tttttttttc	gtttgtttgt	tttgtttttt	agttcccaca	gggagtagag	atggggaaaag	64860
aattcctaca	atatatatct	tggctgataa	aagatacatt	tgtatgttgt	gaagatgttt	64920

gcaatatcga tcagatgact agaaagtga taaaaattaa ggcaactgaa caaaaaaatg	64980
ctcacactcc acatcccgtg atgcacctcc caggccccgc tcattctttg ggcgttggtc	65040
agagtaagct gcttttgacg gaaggacctg tgtttgctca gaacacattc tttccccccc	65100
tccccctctg gtctcctctt tgttttgttt taaggaagaa aaatcagttg cgcgttctga	65160
aatattttac cactgctgtg aacaagtga cacttggtgt cacatcatga cactcgtata	65220
agcatggaga acagtgattt ttttttagaa cagaaaacaa caaaaaataa ccccaaatg	65280
aagattattt tttatgagga gtgaacattt gggtaaatac tggctaagct taaaaaaac	65340
tcatggtgag gcttaacaat gtcttgtaag caaaaggtag agccctgtat caaccagaa	65400
acacctagat cagaacagga atccacattg ccagtgacat gagactgaac agccaaatgg	65460
aggctatgtg gagttggcat tgcatttacc ggcagtgcgg gaggaatttc tgagtggcca	65520
tccaaggtc taggtggagg tggggcatgg tatttgagac attccaaaac gaaggcctct	65580
gaaggaccct tcagaggtgg ctctggaatg acatgtgtca agctgcttgg acctcgtgct	65640
ttaagtcct acattatcta actgtgctca agaggttctc gactggagga ccacactcaa	65700
gccgacttat gccaccatc ccacctctgg ataattttgc ataaaattgg attagcctgg	65760
agcaggttgg gagccaatg tggcatttgt gatcatgaga ttgatgcaat gagatagaag	65820
atgtttgcta cctgaacact tattgctttg aaactagact tgaggaaacc agggtttatc	65880
ttttgagaac ttttggtgaa ggaaaaggga acaggaaaag aaaccccaaa ctcaggccga	65940
atgatcaagg ggaccatag gaaatcttgt ccagagacaa gacttcggga aggtgtctgg	66000
acattcagaa caccaagact tgaaggtgcc ttgctcaatg gaagaggcca ggacagagct	66060
gacaaaattt tgctccccag tgaaggccac agcaaccttc tgcccatcct gtctgttcat	66120
ggagagggtc cctgcctcac ctctgccatt ttgggttagg agaagtcaag ttgggagcct	66180
gaaatagtgg ttcttggaag aatggatccc cagtgaatac tagagctcta agccattca	66240
gcccatttca cacctgaaaa tgttagtgat caccacttgg accagcatcc ttaagtatca	66300
gaaagcccca agcaattgct gcatcttagt aggggtgaggg ataagcaaaa gaggatgttc	66360
accataaccc aggaatgaag ataccatcag caaagaattt caatttgttc agtctttcat	66420
ttagagctag tctttcacag taccatctga atacctcttt gaaagaagga agactttacg	66480
tagtgtagat ttgttttgtg ttgtttgaaa atattatctt tgtaattatt ttaatatgt	66540

```

aaggaatgct tggaatatct gctatatgtc aactttatgc agcttccttt tgagggacaa 66600
atttaaaaca aacaaccccc catcacaac ttaaaggatt gcaagggcca gatctgttaa 66660
gtggtttcat aggagacaca tccagcaatt gtgtggtcag tggctctttt acccaataag 66720
atacatcaca gtcacatgct tgatggttta tgttgacctt agatttatatt tgtaaaatc 66780
tctctctgtt gtgttcgttc ttgttctgtt ttgttttggt ttttaaagtc ttgctgtggt 66840
ctctttgtgg cagaagtgtt tcatgcatgg cagcaggcct gttgcttttt tatggcgatt 66900
cccatgaaa atgtaagtaa atgtctgtgg ccttgttctc tctatggtaa agatattatt 66960
caccatgtaa aacaaaaaac aatatttatt gtattttagt atatttatat aattatgtta 67020
ttgaaaaaaa ttggcattaa aacttaaccg catcagaacc tattgtaaat acaagttcta 67080
tttaagtgtt ctaattaaca tataatatat gttttaaata tagaattttt aatgttttta 67140
aatatatattt caaagtacat aaaa 67164

```

```

<210> 36
<211> 280
<212> PRT
<213> Homo sapiens

```

```

<400> 36

```

```

Met Pro Ala Asp Ile Met Glu Lys Asn Ser Ser Ser Pro Val Ala Ala
1           5           10           15

```

```

Thr Pro Ala Ser Val Asn Thr Thr Pro Asp Lys Pro Lys Thr Ala Ser
          20           25           30

```

```

Glu His Arg Lys Ser Ser Lys Pro Ile Met Glu Lys Arg Arg Arg Ala
          35           40           45

```

```

Arg Ile Asn Glu Ser Leu Ser Gln Leu Lys Thr Leu Ile Leu Asp Ala
          50           55           60

```

```

Leu Lys Lys Asp Ser Ser Arg His Ser Lys Leu Glu Lys Ala Asp Ile
65           70           75           80

```

```

Leu Glu Met Thr Val Lys His Leu Arg Asn Leu Gln Arg Ala Gln Met
          85           90           95

```


Thr Ala Ala Leu Ser Thr Asp Pro Ser Val Leu Gly Lys Tyr Arg Ala
100 105 110

Gly Phe Ser Glu Cys Met Asn Glu Val Thr Arg Phe Leu Ser Thr Cys
115 120 125

Glu Gly Val Asn Thr Glu Val Arg Thr Arg Leu Leu Gly His Leu Ala
130 135 140

Asn Cys Met Thr Gln Ile Asn Ala Met Thr Tyr Pro Gly Gln Pro His
145 150 155 160

Pro Ala Leu Gln Ala Pro Pro Pro Pro Pro Gly Pro Gly Gly Pro
165 170 175

Gln His Ala Pro Phe Ala Pro Pro Pro Pro Leu Val Pro Ile Pro Gly
180 185 190

Gly Ala Ala Pro Pro Pro Gly Gly Ala Pro Cys Lys Leu Gly Ser Gln
195 200 205

Ala Gly Glu Ala Ala Lys Val Phe Gly Gly Phe Gln Val Val Pro Ala
210 215 220

Pro Asp Gly Gln Phe Ala Phe Leu Ile Pro Asn Gly Ala Phe Ala His
225 230 235 240

Ser Gly Pro Val Ile Pro Val Tyr Thr Ser Asn Ser Gly Thr Ser Val
245 250 255

Gly Pro Asn Ala Val Ser Pro Ser Ser Gly Pro Ser Leu Thr Ala Asp
260 265 270

Ser Met Trp Arg Pro Trp Arg Asn
275 280

<210> 37
<211> 843
<212> DNA
<213> Homo sapiens

<400> 37
atgccagctg atataatgga gaaaaattcc tcgtccccgg tggctgctac cccagccagt 60
gtcaacacga caccggataa accaaagaca gcatctgagc acagaaagtc atcaaagcct 120
attatggaga aaagacgaag agcaagaata aatgaaagtc tgagccagct gaaaacactg 180
atdddggatg ctctgaagaa agatagctcg cggcattcca agctggagaa ggaggacatt 240
ctggaaatga cagtgaagca cctccggaac ctgcagcggg cgagatgac ggctgcgctg 300
agcacagacc caagtgtgct ggggaagtac cgagccggct tcagcgagtg catgaacgag 360
gtgacccgct tcctgtccac gtgcgagggc gttataaccg aggtgcgcac tcggctgctc 420
ggccacctgg ccaactgcat gaccagatc aatgccatga cctaccccgg gcagccgcac 480
cccgcccttg aggcgccgc accgccccca ccgggacccg gcggccccca gcacgcgccc 540
ttcgcgccgc cgccgccact cgtgcccatc ccggggggcg cggcgcccc tcccggcggc 600
gccccctgca agctgggcag ccaggctgga gaggcggcta aggtgtttgg aggcttccag 660
gtggtaccgg ctcccgatgg ccagtttgct ttctcattc ccaacggggc cttcggcac 720
agcggccctg tcatccccgt ctacaccagc aacagcggca cctccgtggg cccaacgca 780
gtgtcacctt ccagcggccc ctgcttacg gcggactcca tgtggaggcc gtggcggaac 840
tga 843

<210> 38
<211> 1475
<212> DNA
<213> Homo sapiens

<400> 38
gggatcacac aggatccgga gctggtgctg ataacagcgg aatccccctg ctacctctct 60
ccttggtcct ggaacagcgc tactgatcac caagtagcca caaaatataa taaacctca 120
gcacttgctc agtagttttg tgaaagtctc aagtaaaaga gacacaaaca aaaaattctt 180
tttcgtgaag aactccaaaa ataaaattct ctagagataa aaaaaaaaaa aaaaggaaaa 240
tgccagctga tataatggag aaaaattcct cgtccccggg ggctgctacc ccagccagt 300
tcaacacgac accggataaa ccaaagacag catctgagca cagaaagtca tcaaagccta 360
ttatggagaa aagacgaaga gcaagaataa atgaaagtct gagccagctg aaaacactga 420
ttttggatgc tctgaagaaa gatagctcgc ggcattccaa gctggagaag gcggacattc 480

tggaatgac agtgaagcac ctccggaacc tgcagcgggc gcagatgacg gctgcgctga	540
gcacagaccc aagtgtgctg gggaagtacc gagccggctt cagcgagtgc atgaacgagg	600
tgacccgctt cctgtccacg tgcgagggcg ttaataccga ggtgcgcact cggctgctcg	660
gccacctggc caactgcatg acccagatca atgccatgac ctaccccggg cagccgcacc	720
ccgccttgca ggcgccgcca ccgccccac cgggacccgg cggccccag cacgcgccgt	780
tcgcgccgcc gccgccactc gtgcccatcc ccgggggcgc ggcgccccct cccggcggcg	840
ccccctgcaa gctgggcagc caggctggag aggcggctaa ggtgtttgga ggcttccagg	900
tggtaccggc tcccgatggc cagtttgctt tcctcattcc caacggggcc ttcgcgcaca	960
gcggccctgt catccccgtc tacaccagca acagcggcac ctccgtgggc cccaacgcag	1020
tgtcaccttc cagcggcccc tcgcttacgg cggactccat gtggaggccg tggcggaact	1080
gagggggctc aggccacccc tcctcctaaa ctccccaacc cacctctctt ccctccggac	1140
tctaaacagg aacttgaata ctgggagaga agaggacttt tttgattaag tggttacttt	1200
gtgttttttt aatttctaag aagttacttt ttgtagagag agctgtatta agtgactgac	1260
catgcactat atttgtatat attttatatg ttcatatagg attgcgcctt tgtattataa	1320
aagctcagat gacatttcgt tttttacacg agatttcttt tttatgtgat gccaaagatg	1380
tttgaaaatg ctcttaaaat atcttccttt ggggaagttt atttgagaaa atataataaa	1440
agaaaaaagt aaaggctttt aaaaaaaaaa aaaaa	1475

<210> 39
 <211> 166
 <212> PRT
 <213> Homo sapiens

<400> 39

Met	Ala	Pro	Ser	Thr	Val	Ala	Val	Glu	Leu	Leu	Ser	Pro	Lys	Glu	Lys
1				5				10						15	

Asn	Arg	Leu	Arg	Lys	Pro	Val	Val	Glu	Lys	Met	Arg	Arg	Asp	Arg	Ile
		20						25					30		

Asn	Ser	Ser	Ile	Glu	Gln	Leu	Lys	Leu	Leu	Leu	Glu	Gln	Glu	Phe	Ala
		35					40					45			

Arg His Gln Pro Asn Ser Lys Leu Glu Lys Ala Asp Ile Leu Glu Met
 50 55 60

Ala Val Ser Tyr Leu Lys His Ser Lys Ala Phe Val Ala Ala Ala Gly
 65 70 75 80

Pro Lys Ser Leu His Gln Asp Tyr Ser Glu Gly Tyr Ser Trp Cys Leu
 85 90 95

Gln Glu Ala Val Gln Phe Leu Thr Leu His Ala Ala Ser Asp Thr Gln
 100 105 110

Met Lys Leu Leu Tyr His Phe Gln Arg Pro Pro Ala Ala Pro Ala Ala
 115 120 125

Pro Ala Lys Glu Pro Lys Ala Pro Gly Ala Ala Pro Pro Pro Ala Leu
 130 135 140

Ser Ala Lys Ala Thr Ala Ala Ala Ala Ala His Gln Pro Ala Cys
 145 150 155 160

Gly Leu Trp Arg Pro Trp
 165

<210> 40
 <211> 501
 <212> DNA
 <213> Homo sapiens

<400> 40
 atggcccca gcactgtggc cgtggagctg ctcagcccca aagagaaaaa ccgactgcgg 60
 aagccggtgg tggagaagat gcgccgacgc cgcatcaaca gcagcatcga gcagctgaag 120
 ctgctgctgg agcaggagtt cgcgcggcac cagcccaact ccaagctgga gaaggccgac 180
 atcctggaga tggctgtcag ctacctgaag cacagcaaag ccttcgtcgc cgccgccggc 240
 cccaagagcc tgcaccagga ctacagcgaa ggctactcgt ggtgcctgca ggaggccgtg 300
 cagttcctga cgctccacgc cgccagcgac acgcagatga agctgctgta ccacttccag 360
 cggcccccgg ccgcgcccgc cgcgcccgcc aaggagccca aggcgcccgg cgccgcgccc 420
 ccgcccgcgc tctccgcaa ggccaccgcc gccgcgccc ccgcgcacca gcccgcctgc 480

ggcctctggc ggccctggtg a

501

<210> 41

<211> 1319

<212> DNA

<213> Homo sapiens

<400> 41

cgcgcttggc cttgcccgcg cccgctcgcc tcgtctcgcc cggcctcccc gcgtcgcctc	60
gtcgcctgtt ccgcgccagg catggccccc agcactgtgg ccgtggagct gtcagcccc	120
aaagagaaaa accgactgcg gaagccggtg gtggagaaga tgcgccgcga ccgcatcaac	180
agcagcatcg agcagctgaa gctgctgctg gagcaggagt tcgcgcggca ccagcccaac	240
tccaagctgg agaaggccga catcctggag atggctgtca gctacctgaa gcacagcaaa	300
gccttcgtcg ccgccgccgg ccccaagagc ctgcaccagg actacagcga aggctactcg	360
tggtgcctgc aggaggccgt gcagttcctg acgctccacg ccgccagcga cacgcagatg	420
aagctgctgt accacttcca gcggcccccg gccgcgcccc ccgcgccccg caaggagccc	480
aaggcgccgg gcgccgcgcc cccgccccgc ctctccgcca aggccaccgc cgccgccgcc	540
gccgcgcacc agcccgccctg cggcctcttg cggccctggt gacccggcgg gacctgcggg	600
cgcgcgggccc gacgaccaga gggcgagcct gctcctctcg cctgtaggga agcgccttcc	660
cgccgtcgtc cgccccgggc ttggacgcgc ctttctccgg aaggctcttg cccaagctg	720
gccggccccg aggagcccca ttctcagaga atgtgtgtgc agagtccctg ccgttttagg	780
acaatcaggg cccatcttct gccaaagtgtc tgaccccatg gggttgttct gtgtttgcat	840
ttaagcaagt gacttctggg aagtccccgg ccgcccgggg ttctatgata tttgtagtgc	900
cggggctcgc aactgctgc cccagcctg tagaggactt tcttcagggc ccgtagctgc	960
tgggcgtacc cctggcaggc gggctgtgcc gcgggcacat ttgccttttg tgaaggccga	1020
actcgagctg tatcctcata ggaaacagt atcaccccg acgggcgtcc aggaccctga	1080
gggcatggc caaaaggctc ctgagtgtgc ctggtggtct ggctggggct cacggtgggc	1140
tgtctgggga ggggtgggtgc ctccactatg atccttaaag gattcctctg tgtgggtgga	1200
tgcgtgtggg cacgactttg tactcagaaa ttgaactctc agtcacgtgg aagccacggg	1260
actgctccga agccgccata ataaaatctg attgttcagc ccccaaaaaa aaaaaaaaa	1319

<210> 42
<211> 198
<212> PRT
<213> Homo sapiens

<400> 42

Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
1 5 10 15

Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
20 25 30

Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
35 40 45

Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
50 55 60

Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
65 70 75 80

Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
85 90 95

Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
100 105 110

Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
115 120 125

His Leu Val Asp Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu
130 135 140

Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
145 150 155 160

Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
165 170 175

Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
180 185 190

Leu Arg Arg Arg Gln Thr
195

<210> 43
<211> 597
<212> DNA
<213> Homo sapiens

<400> 43
atgtcaaacg tgcgagtgtc taacgggagc cctagcctgg agcggatgga cgccaggcag 60
gcggagcacc ccaagccctc ggccctgcagg aacctcttcg gcccgggtgga ccacgaagag 120
ttaacccggg acttggagaa gcactgcaga gacatggaag aggcgagcca gcgcaagtgg 180
aatttcgatt ttcagaatca caaaccccta gagggcaagt acgagtggca agaggtggag 240
aagggcagct tgcccagatt ctactacaga ccccgcggc ccccaaaagg tgcctgcaag 300
gtgccggcgc aggagagcca ggatgtcagc gggagccgcc cggcggcgcc tttaattggg 360
gctccggcta actctgagga cacgcatttg gtggacccaa agactgatcc gtcggacagc 420
cagacggggt tagcggagca atgctgcagga ataaggaagc gacctgcaac cgacgattct 480
tctactcaaa acaaaagagc caacagaaca gaagaaaatg tttcagacgg ttccccaaat 540
gccggttctg tggagcagac gcccaagaag cctggcctca gaagacgtca aacgtaa 597

<210> 44
<211> 5114
<212> DNA
<213> Homo sapiens

<400> 44
ttaaggccgc gctcgccagc ctcggcgggg cggtctccgc cgccgcaacc aatggatctc 60
ctcctctgtt taaatagact cgccgtgtca atcattttct tcttcgtcag cctcccttcc 120
accgccatat tgggccacta aaaaaagggg gctcgtcttt tcggggtgtt tttctcccc 180
tcccctgtcc ccgcttgctc acggctctgc gactccgacg ccggcaaggt ttggagagcg 240
gctgggttcg cgggacccgc gggcttgac ccgcccagac tcggacgggc ttgcccaccc 300
tctccgcttg cctgggtcccc tctcctctcc gccctccgc tcgccagtcc atttgatcag 360
cggagactcg gcggccgggc cggggcttcc ccgcagcccc tgcgcgctcc tagagctcgg 420
gccgtggctc gtcggggtct gtgtcttttg gctccgaggg cagtcgctgg gcttccgaga 480

ggggttcggg	ctgcgtaggg	gcgctttgtt	ttgttcgggt	ttgttttttt	gagagtgcga	540
gagaggcggg	cgtgcagacc	cgggagaaaag	atgtcaaacg	tgcgagtgtc	taacgggagc	600
cctagcctgg	agcggatgga	cgccaggcag	gcggagcacc	ccaagccctc	ggcctgcagg	660
aacctcttcg	gcccgggtgga	ccacgaagag	ttaaccgggg	acttgagaa	gcactgcaga	720
gacatggaag	aggcgagcca	gcgcaagtgg	aatttcgatt	ttcagaatca	caaacccta	780
gagggcaagt	acgagtggca	agagggtggag	aagggcagct	tgcccagatt	ctactacaga	840
ccccgcggc	ccccaaaagg	tgctgcaag	gtgccggcgc	aggagagcca	ggatgtcagc	900
gggagccgcc	cggcgggcgcc	tttaattggg	gctccggcta	actctgagga	cacgcatttg	960
gtggacccaa	agactgatcc	gtcggacagc	cagacggggg	tagcggagca	atgcgcagga	1020
ataaggaagc	gacctgcaac	cgacggtaat	gaccctttcc	caaccataga	atgtgtttgg	1080
ggccccgctt	tgcttgctgg	agggtgttaa	ccttagcttg	cttttcggcg	tattctgatt	1140
tagctttggg	agagctaact	ttattggtct	taggtgttca	gtgctacctg	gcccactgct	1200
tgtctgtttg	tgacttttaa	gtcagaaact	ggagatggta	agatccgata	atttccctaa	1260
cttaatacat	cgcggtccct	ctcactagca	actcctaggt	atgtgacaaa	gttgggatgt	1320
ttatcaacgg	tccgcctcct	ggctagggaa	agagctctgg	ggcggagaat	gcactttctg	1380
ttttttgaaa	acaacctcat	tttgtgccct	taaaagccac	tggggatgac	ggatccagga	1440
ttgtgggtgg	aggtagtggg	tttttcatcc	cctgactatg	gggccaactt	ctgccagcca	1500
ttgttttttc	taataaagat	tgtgtgttct	ttttaaaaat	ttcccctgcg	cttagattct	1560
tctactcaaa	acaaaagagc	caacagaaca	gaagaaaatg	tttcagacgg	ttccccaat	1620
gccggttctg	tggagcagac	gccaagaag	cctggcctca	gaagacgtca	aacgtaaaca	1680
gctcgggtggg	ttgatcacta	aaggagcacg	cactggaacc	cggggccttc	agacctcacg	1740
atactgatc	ttactggttg	ctggcaaatt	aaaagcttat	ggggttttgt	ttgtttata	1800
cttcgtgagg	tcaaaaaagt	agcaatgggg	aaggctgggg	atacggtaat	tcctcagagt	1860
ttctatgccc	agagatactt	tctcttcaaa	ctgttgacca	gagcagctac	ttgtaacca	1920
ggcccatcgc	ggtaggaagg	tcgtttccct	gtgagtccca	ctaaaacgtg	ttgggagcaa	1980
taggttcttt	gcccattccga	acaagaacta	gggtactccc	tcagtccgaa	ttaatgagaa	2040
ttaatttcct	agaggttcag	cttgagtcgg	taacagattt	tgagccatac	atggaaaaat	2100

ggcaaataca tgattaagtt tcaatthttga gggggaatgt ttggtagaaa ttgctcatct	2160
ttggttatgc aagggttag agatgtgaat aggatggtat gttgtgttct ttgacattht	2220
aataaactgt cactthccct gttgtctcct aagthtggag agagaaggaa ccagtatthg	2280
caaaaaccaa atggaaagat aaaaaagtha ctaaagthtc tacagaatth ctggtaacac	2340
tgaagttgca aagcagaagt taaatthaact cttgtcagta agcaatccag gaacacgtca	2400
gccagtgtat gctaattgtg ccgtaacagg gtgatttgga tathtgtagg ggaaatgggt	2460
agtaaatatc aagactgggt accgtaggtc agcccagcac aaagggaagt gagatthtthc	2520
catgcacaag aatctgatca ctgtaaatag ctaatthgaa taattcagtc cccagataac	2580
caacatgggt tggttattca taataaacta catatthtaa tagthtatta gctthcttha	2640
gaccaagact gtgacctctt taththctaa agcacacacg tagthtagca tatgaggcga	2700
taaaatattg atgtthaactt thtaaatccc cagthataaa aatthtaaaa taacagggt	2760
taaggtgaga thcaggthtg ttgtgtctth aaattgtata tgtgacttca catatcttht	2820
tcagcgctta taaaaacgg cactatagaa cthccattht acagcaccat atgaagtggg	2880
aaaattaggt gaaaatthtth ctgaagcaac cthaacatgc gcagcccttg ttggthtgtg	2940
acttgtggcc tagctcatca gatgagccac gagaatcaga cthggattht gatctggccc	3000
thtthctgaca tgcaatgagg catthtgtagg atthtagtaat attgctagth caaagaatac	3060
tagaaatatt agtaagaacc tathcaaaag tathcatgag taththctgc atatgaatca	3120
ggaattagaa taththgaaa atgatgttha taaaatthtth ctctggaagg cththataat	3180
ththattccc aatcatthtth caaatthaga aagththaatc tgtcacagga gaaaaaaat	3240
taaaatthtth caaaaattha gaaaatthtth acccgtaagt attacagtht cthaattatc	3300
ctaththtth cccacttgcc thtgacctag atththtaga gtagggthtth tcagcctctg	3360
cactaatgac atthtgggcc gaataattct gthtgtaggag gctgtcctgt gtgththtaga	3420
thgthtgga thtatccctgg cctctcacac tagatgccag cagtatctc thccccagth	3480
gtgacaacct aaaatgtctc cagacattgc caaatgtthc ctgtggggcg ggggcaacat	3540
tgcctactgt taagaactac tgctctagac caaagaacac agcacagagg aaaggaaaaa	3600
aaaatccagth taagagaatg thaggtggag atgactatag tcatcaaaact thththccca	3660
tcaagthtth ccaagctaac atagtgacaa aataatthct gtactctact ggtaacgtha	3720

atctagtgct cttcctttaa ttcttccggt ttgttttctt ttgcagaatt aagaatatgt	3780
ttccttggtt atcagataca tcaactgcttg atgaagcaag gaagatatac atgaaaattt	3840
taaaaataca tatcgctgac ttcatggaat ggacatcctg tataagcact gaaaaacaac	3900
aacacaataa cactaaaatt ttaggcactc tttaatgatac tgcctctaaa agcgttggt	3960
gtagcattat gcaattaggt ttttccttat ttgcttcatt gtactacctg tgtatatagt	4020
ttttacctt tatgtagcac ataaactttg gggaaggag ggcagggtgg ggctgaggaa	4080
ctgacgtgga gcggggtatg aagagcttgc ttgatttac agcaagtaga taaatatttg	4140
acttgcata agagaagcaa ttttggggaa gggtttgaat tgttttctt aaagatgtaa	4200
tgtccctttc agagacagct gatacttcat ttaaaaaaat cacaaaaatt tgaacactgg	4260
ctaaagataa ttgctattta tttttacaag aagtttattc tcatttggga gatctggtga	4320
tctccaagc tatctaaagt ttgttagata gctgcatgtg gcttttttaa aaaagcaaca	4380
gaaacctatc ctcaactgcc tccccagtct ctcttaaagt tggaatttac cagttaatta	4440
ctcagcagaa tggatgacac tccaggtagt ttggggcaaa aatccgaggt gcttgggagt	4500
tttgaatgtt aagaattgac catctgctt tattaaattt gttgacaaaa ttttctcatt	4560
ttcttttcac ttcgggctgt gtaaacaacag tcaaaataat tctaaatccc tcgatatttt	4620
taaagatctg taagtaactt cacattaaaa aatgaaatat tttttaattt aaagcttact	4680
ctgtccattt atccacagga aagtgttatt tttcaaggaa gggtcatgta gagaaaagca	4740
cactttagg ataagtgaat tggatactac atctttaaac agtatttcat tgcctgtgta	4800
tggaaaaacc atttgaagtg tacctgtgta cataactctg taaaaaact gaaaaattat	4860
actaacttat ttatgttaaa agattttttt taatctagac aatatacaag ccaaagtggc	4920
atgttttgtg cttttgtaaa tgctgtgttg ggtagaatag gttttcccct cttttgttaa	4980
ataatatggc tatgcttaaa aggttgcata ctgagccaag tataattttt tgtaatgtgt	5040
gaaaaagatg ccaattattg ttacacatta agtaatcaat aaagaaaact tccatagcta	5100
ttcattgagt caaa	5114

<210> 45
 <211> 317
 <212> PRT
 <213> Homo sapiens

<400> 45

Met Tyr Asn Met Met Glu Thr Glu Leu Lys Pro Pro Gly Pro Gln Gln
1 5 10 15

Thr Ser Gly Gly Gly Gly Gly Asn Ser Thr Ala Ala Ala Ala Gly Gly
20 25 30

Asn Gln Lys Asn Ser Pro Asp Arg Val Lys Arg Pro Met Asn Ala Phe
35 40 45

Met Val Trp Ser Arg Gly Gln Arg Arg Lys Met Ala Gln Glu Asn Pro
50 55 60

Lys Met His Asn Ser Glu Ile Ser Lys Arg Leu Gly Ala Glu Trp Lys
65 70 75 80

Leu Leu Ser Glu Thr Glu Lys Arg Pro Phe Ile Asp Glu Ala Lys Arg
85 90 95

Leu Arg Ala Leu His Met Lys Glu His Pro Asp Tyr Lys Tyr Arg Pro
100 105 110

Arg Arg Lys Thr Lys Thr Leu Met Lys Lys Asp Lys Tyr Thr Leu Pro
115 120 125

Gly Gly Leu Leu Ala Pro Gly Gly Asn Ser Met Ala Ser Gly Val Gly
130 135 140

Val Gly Ala Gly Leu Gly Ala Gly Val Asn Gln Arg Met Asp Ser Tyr
145 150 155 160

Ala His Met Asn Gly Trp Ser Asn Gly Ser Tyr Ser Met Met Gln Asp
165 170 175

Gln Leu Gly Tyr Pro Gln His Pro Gly Leu Asn Ala His Gly Ala Ala
180 185 190

Gln Met Gln Pro Met His Arg Tyr Asp Val Ser Ala Leu Gln Tyr Asn
195 200 205

Ser Met Thr Ser Ser Gln Thr Tyr Met Asn Gly Ser Pro Thr Tyr Ser
210 215 220

Met Ser Tyr Ser Gln Gln Gly Thr Pro Gly Met Ala Leu Gly Ser Met
225 230 235 240

Gly Ser Val Val Lys Ser Glu Ala Ser Ser Ser Pro Pro Val Val Thr
245 250 255

Ser Ser Ser His Ser Arg Ala Pro Cys Gln Ala Gly Asp Leu Arg Asp
260 265 270

Met Ile Ser Met Tyr Leu Pro Gly Ala Glu Val Pro Glu Pro Ala Ala
275 280 285

Pro Ser Arg Leu His Met Ser Gln His Tyr Gln Ser Gly Pro Val Pro
290 295 300

Gly Thr Ala Ile Asn Gly Thr Leu Pro Leu Ser His Met
305 310 315

<210> 46
<211> 954
<212> DNA
<213> Homo sapiens

<400> 46
atgtacaaca tgatggagac ggagctgaag ccgccgggcc cgcagcaaac ttcggggggc 60
ggcggcggca actccaccgc ggcggcggcc ggcggcaacc agaaaaacag cccggaccgc 120
gtcaagcggc ccatgaatgc cttcatggtg tgggtcccgcg ggcagcggcg caagatggcc 180
caggagaacc ccaagatgca caactcggag atcagcaagc gcctgggcgc cgagtggaaa 240
cttttgtcgg agacggagaa gcggccgttc atcgacgagg ctaagcggct gcgagcgctg 300
cacatgaagg agcacccgga ttataaatac cggccccggc ggaaaaccaa gacgctcatg 360
aagaaggata agtacacgct gcccggcggg ctgctggccc ccggcggcaa tagcatggcg 420
agcgggggtcg ggggtgggcgc cggcctgggc gcgggcgtga accagcgcac ggacagttac 480
gcgcacatga acggctggag caacggcagc tacagcatga tgcaggacca gctgggctac 540
ccgcagcacc cgggcctcaa tgcgcacggc gcagcgcaga tgcagcccat gcaccgctac 600

gacgtgagcg ccctgcagta caactccatg accagctcgc agacctacat gaacggctcg	660
cccacctaca gcatgtccta ctcgcagcag ggcacccctg gcatggctct tggctccatg	720
ggttcggtgg tcaagtccga ggccagctcc agccccctg tggttacctc ttcctcccac	780
tccagggcgc cctgccaggc cggggacctc cgggacatga tcagcatgta tctccccggc	840
gccgaggtgc cggaaccgc cgcaccagc agacttcaca tgtcccagca ctaccagagc	900
ggcccgggtgc ccggcacggc cattaacggc aactgcccc tctcacacat gtga	954

<210> 47
 <211> 2513
 <212> DNA
 <213> Homo sapiens

<400> 47	
ggatggttgt ctattaactt gttcaaaaaa gtatcaggag ttgtcaaggc agagaagaga	60
gtgtttgcaa aagggggaaa gtagtttgct gcctctttaa gactaggact gagagaaaga	120
agaggagaga gaaagaaagg gagagaagtt tgagccccag gcttaagcct ttccaaaaaa	180
taataataac aatcatcggc ggccggcagga tcggccagag gaggagggaa gcgctttttt	240
tgatcctgat tccagtttgc ctctctcttt ttttccccca aattattctt cgcctgattt	300
tcctcgcgga gccctgcgt cccgacacc cgcgccct cccctcctcc tctcccccg	360
cccgccggcc ccccaaagtc ccggccgggc cgagggtcgg cggccgccgg cgggccgggc	420
ccgcgcacag cgcccgcag tacaacatga tggagacgga gctgaagccg ccgggcccgc	480
agcaaacttc ggggggcggc ggccggcaact ccaccgcggc ggccggccggc ggcaaccaga	540
aaaacagccc ggaccgcgtc aagcggccca tgaatgcctt catggtgtgg tcccgcgggc	600
agcggcgcaa gatggcccag gagaaccca agatgcacaa ctcggagatc agcaagcgcc	660
tgggcgccga gtggaaactt ttgtcggaga cggagaagcg gccgttcac gacgaggcta	720
agcggctgcg agcgtgcac atgaaggagc acccggatta taaataccgg ccccgccgga	780
aaaccaagac gtcataag aaggataagt acacgctgcc cggcgggctg ctggcccccg	840
gcggcaatag catggcgagc ggggtcgggg tgggcgccgg cctgggcgcg ggcgtgaacc	900
agcgcatagga cagttacgc cacatgaac gctggagcaa cggcagctac agcatgatgc	960
aggaccagct gggctacccg cagcaccgg gcctcaatgc gcacggcgca gcgcagatgc	1020
agcccatgca ccgctacgac gtgagcgccc tgcagtacaa ctccatgacc agctcgcaga	1080

cctacatgaa cggctcgccc acctacagca tgtcctactc gcagcagggc acccctggca	1140
tggctcttgg ctccatgggt tcggtggtca agtccgaggc cagctccagc ccccctgtgg	1200
ttacctcttc ctcccactcc agggcgccct gccaggccgg ggacctccgg gacatgatca	1260
gcatgtatct ccccggcgcc gaggtgccgg aaccgcgcgc ccccgagcaga cttcacatgt	1320
cccagcacta ccagagcggc cgggtgcccg gcacggccat taacggcaca ctgcccctct	1380
cacacatgtg agggccggac agcgaactgg aggggggaga aattttcaaa gaaaaacgag	1440
ggaaatggga ggggtgcaaa agaggagagt aagaaacagc atggagaaaa cccggtacgc	1500
tcaaaaagaa aaaggaaaaa aaaaaatccc atcacccaca gcaaatgaca gctgcaaaag	1560
agaacaccaa tcccatccac actcacgcaa aaaccgcgat gccgacaaga aaacttttat	1620
gagagagatc ctggacttct ttttggggga ctatttttgt acagagaaaa cctggggagg	1680
gtggggaggg cgggggaatg gaccttgtat agatctggag gaaagaaagc tacgaaaaac	1740
tttttaaaag ttctagtggc acggtaggag ctttgcagga agtttgcaaa agtctttacc	1800
aataatattt agagctagtc tccaagcgac gaaaaaatg ttttaatat tgcaagcaac	1860
ttttgtacag tatttatcga gataaacatg gcaatcaaaa tgtccattgt ttataagctg	1920
agaatttgcc aatatttttc aaggagaggc ttcttgctga attttgattc tgcagctgaa	1980
atttaggaca gttgcaaacg tgaaaagaag aaaattattc aaatttggac attttaattg	2040
tttaaaaatt gtacaaaagg aaaaaattag aataagtact ggccaaccat ctctgtggtc	2100
ttgtttaaaa agggcaaaag ttttagactg tactaaattt tataacttac tgttaaaagc	2160
aaaaatggcc atgcaggttg acaccgttgg taatttataa tagcttttgt tcgatcccaa	2220
ctttccattt tgttcagata aaaaaaacca tgaaattact gtgtttgaaa tattttctta	2280
tggtttgtaa tattttctgta aatttattgt gatatttttaa ggttttcccc cttttatttt	2340
ccgtagttgt attttaaaag attcggctct gtattatttg aatcagtctg ccgagaatcc	2400
atgtatatat ttgaactaat atcatcctta taacaggtag attttcaact taagttttta	2460
ctccattatg cacagtttga gataaataaa tttttgaaat atggacactg aaa	2513

<210> 48
 <211> 159
 <212> PRT
 <213> E.coli

<400> 48

Met Ile Ser Leu Ile Ala Ala Leu Ala Val Asp Tyr Val Ile Gly Met
1 5 10 15

Glu Asn Ala Met Pro Trp Asn Leu Pro Ala Asp Leu Ala Trp Phe Lys
20 25 30

Arg Asn Thr Leu Asn Lys Pro Val Ile Met Gly Arg His Thr Trp Glu
35 40 45

Ser Ile Gly Arg Pro Leu Pro Gly Arg Lys Asn Ile Ile Leu Ser Ser
50 55 60

Gln Pro Ser Thr Asp Asp Arg Val Thr Trp Val Lys Ser Val Asp Glu
65 70 75 80

Ala Ile Ala Ala Cys Gly Asp Val Pro Glu Ile Met Val Ile Gly Gly
85 90 95

Gly Arg Val Ile Glu Gln Phe Leu Pro Lys Ala Gln Lys Leu Tyr Leu
100 105 110

Thr His Ile Asp Ala Glu Val Glu Gly Asp Thr His Phe Pro Asp Tyr
115 120 125

Glu Pro Asp Asp Trp Glu Ser Val Phe Ser Glu Phe His Asp Ala Asp
130 135 140

Ala Gln Asn Ser His Ser Tyr Cys Phe Glu Ile Leu Glu Arg Arg
145 150 155

<210> 49

<211> 112

<212> PRT

<213> A.thaliana

<400> 49

Met Gly Val Glu Lys Gln Val Ile Arg Pro Gly Asn Gly Pro Lys Pro
1 5 10 15

Ala Pro Gly Gln Thr Val Thr Val His Cys Thr Gly Phe Gly Lys Asp
20 25 30

Gly Asp Leu Ser Gln Lys Phe Trp Ser Thr Lys Asp Glu Gly Gln Lys
35 40 45

Pro Phe Ser Phe Gln Ile Gly Lys Gly Ala Val Ile Lys Gly Trp Asp
50 55 60

Glu Gly Val Ile Gly Met Gln Ile Gly Glu Val Ala Arg Leu Arg Cys
65 70 75 80

Ser Ser Asp Tyr Ala Tyr Gly Ala Gly Gly Phe Pro Ala Trp Gly Ile
85 90 95

Gln Pro Asn Ser Val Leu Asp Phe Glu Ile Glu Val Leu Ser Val Gln
100 105 110

<210> 50
<211> 1805
<212> DNA
<213> Artificial Sequence

<220>
<223> pITR-CMV-mScarlet

<400> 50
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagttag cgagcgagcg cgagagagg gaggggccaa ctccatcact 120
aggggttcct ctatgccca tatatggagt tccgcgttac ataacttacg gtaaatggcc 180
cgcctggctg accgcccac gacccccgcc cattgacgtc aataatgacg tatgttccca 240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg 300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg 360
acggtaaatg gcccgctgg cattatgccc agtacatgac cttatgggac tttcctactt 420
ggcagtacat ctacgtatta gtcacgcta ttaccatggt gatgcgggtt tggcagtaca 480
tcaatgggag tggatagcgg tttgactcac ggggatttcc aagtctccac ccattgacg 540
tcaatgggag tttgttttgg caccaaatc aacgggactt tccaaatgt cgtacaact 600
ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat ataagcagag 660

ctcgttttagt gaaccgtcag atcgcttgga gacgcatgcc taagaagaag cggaaagtcg	720
gctccggcgt gagcaagggc gaggcagtga tcaaggagtt catgcggttc aaggtgcaca	780
tggagggctc catgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct	840
acgagggcac ccagaccgcc aagctgaagg tgaccaaggg tggccccctg cctttctcct	900
gggacatcct gtcccctcag ttcatgtacg gctccagggc cttcatcaag caccccgccg	960
acatccccga ctactataag cagtccttcc ccgagggcctt caagtgggag cgcgtgatga	1020
acttcgagga cggcggcgcc gtgaccgtga cccaggacac ctccctggag gacggcaccc	1080
tgatctacaa ggtgaagctc cgcggcacca acttccttcc tgacggcccc gtaatgcaga	1140
agaagacaat gggctgggaa gcgtccaccg agcggttgta ccccgaggac ggcgtgctga	1200
agggcgacat taagatggcc ctgcgctga aggacggcgg ccgctacctg gcggacttca	1260
agaccaccta caaggccaag aagcccgtgc agatgcccgg cgcctacaac gtcgaccgca	1320
agttggacat cacctcccac aacgaggact acaccgtggt ggaacagtac gaacgctccg	1380
agggccgcca ctccaccggc ggcatggacg agctgtacaa gtaagctgat cagcctcgac	1440
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct	1500
ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct	1560
gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg	1620
ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggaggaac ccctagtgat	1680
ggagttggcc actccctctc tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt	1740
cggccgacgc ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc gcagctgcct	1800
gcagg	1805

<210> 51
 <211> 130
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5' ITR cDNA

<400> 51	
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgccccg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120

aggggttcct 130

<210> 52
<211> 328
<212> DNA
<213> Artificial Sequence

<220>
<223> CMV enhancer

<400> 52
ctagatccca tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg 60
accgccaac gacccccgcc cattgacgtc aataatgacg tatgttcca tagtaacgcc 120
aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc 180
agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaagt 240
gccccgctgg cattatgccc agtacatgac cttatgggac tttcctactt ggagtagcat 300
ctacgtatta gtcacgcta ttaccatg 328

<210> 53
<211> 237
<212> DNA
<213> Artificial Sequence

<220>
<223> CMV promoter

<400> 53
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 60
ccaagtctcc acccattga cgtcaatggg agtttggtttt ggcacacaaa tcaacgggac 120
tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg 180
tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgctg gagacgc 237

<210> 54
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> SV40-NLS

<400> 54
atgcctaaga agaagcggaa agtcggctcc ggc 33

<210> 55
<211> 696
<212> DNA
<213> Artificial Sequence

<220>
<223> mScarlet

<400> 55
gtgagcaagg gcgaggcagt gatcaaggag ttcatgcggt tcaagggtgca catggaggggc 60
tccatgaacg gccacgagtt cgagatcgag ggcgagggcg agggccgccc ctacgagggc 120
accagaccg ccaagctgaa ggtgaccaag ggtggccccc tgcccttctc ctgggacatc 180
ctgtcccctc agttcatgta cggctccagg gccttcatca agcaccgccg cgacatcccc 240
gactactata agcagtcctt ccccgagggc ttcaagtggg agcgcgtgat gaacttcgag 300
gacggcggcg ccgtgaccgt gacccaggac acctccctgg aggacggcac cctgatctac 360
aaggtgaagc tccgcggcac caacttcctt cctgacggcc ccgtaatgca gaagaagaca 420
atgggctggg aagcgtccac cgagcggttg taccgagagg acggcgtgct gaagggcgac 480
attaagatgg ccctgcgcct gaaggacggc ggccgctacc tggcggactt caagaccacc 540
tacaaggcca agaagcccgt gcagatgccc ggcgcctaca acgtcgaccg caagttggac 600
atcacctccc acaacgagga ctacaccgtg gtggaacagt acgaacgctc cgaggggccg 660
cactccaccg gcggcatgga cgagctgtac aagtaa 696

<210> 56
<211> 240
<212> DNA
<213> Bovidae

<400> 56
gctgatcagc ctgcactgtg ctttctagtt gccagccatc tgttggttgc ccctcccccg 60
tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa 120
ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtggggtg gggcaggaca 180
gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg 240

<210> 57
<211> 141
<212> DNA

<213> Artificial Sequence

<220>

<223> 3' ITR

<400> 57

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgag ctgcctgcag g 141

<210> 58

<211> 2285

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-mScarlet-DD

<400> 58

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgctcg ggcgaccttt 60

ggtcgcccgg cctcagttag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120

aggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaatggcc 180

cgcctggctg accgcccac gacccccgcc cattgacgtc aataatgacg tatgttccca 240

tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg 300

cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg 360

acggtaaatg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt 420

ggcagtacat ctacgtatta gtcacgcta ttaccatggt gatgcggttt tggcagtaca 480

tcaatgggcg tggatagcgg ttgactcac ggggatttcc aagtctccac ccattgacg 540

tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact 600

ccgccccatt gacgcaaag ggcggtaggc gtgtacgggt ggagggtctat ataagcagag 660

ctcgtttagt gaaccgtcag atcgctgga gacgcatgcc taagaagaag cggaaagtcg 720

gctccggcgt gagcaagggc gaggcagtga tcaaggagtt catgcggttc aaggtgcaca 780

tggagggctc catgaacggc cacgagttcg agatcgaggg cgagggcgag ggccgcccct 840

acgagggcac ccagaccgcc aagctgaagg tgaccaaggg tggccccctg cctttctcct 900

gggacatcct gtcccctcag ttcatgtacg gctccagggc cttcatcaag caccgcccg 960

acatccccga ctactataag cagtccttcc ccgagggcctt caagtgggag cgcgtgatga	1020
acttcgagga cggcggcgcc gtgaccgtga cccaggacac ctccctggag gacggcaccc	1080
tgatctacaa ggtgaagctc cgcggcacca acttccctcc tgacggcccc gtaatgcaga	1140
agaagacaat gggctgggaa gcgtccaccg agcggttgta ccccgaggac ggcgtgctga	1200
agggcgacat taagatggcc ctgcgctga aggacggcgg ccgctacctg gcggacttca	1260
agaccaccta caaggccaag aagcccgtgc agatgcccgg cgcctacaac gtcgaccgca	1320
agttggacat cacctccac aacgaggact acaccgtggt ggaacagtac gaacgctccg	1380
agggccgcca ctccaccggc ggcatggacg agctgtacaa ggggtaccatc agtctgattg	1440
cggcgttagc ggtagattac gttatcggca tggaaaacgc catgccgtgg aacctgcctg	1500
ccgatctcgc ctggtttaaa cgcaacacct taaataaacc cgtgattatg ggccgccata	1560
cctgggaatc aatcggtcgt ccgttgccag gacgcaaaaa tattatcctc agcagtcaac	1620
cgagtacgga cgatcgcgta acgtgggtga agtcgggtgga tgaagccatc gcggcgtgtg	1680
gtgacgtacc agaaatcatg gtgattggcg gcggtcgcgt tattgaacag ttcttgccaa	1740
aagcgcaaaa actgtatctg acgcatatcg acgcagaagt ggaaggcgac acccatttcc	1800
cggattacga gccggatgac tgggaatcgg tattcagcga attccacgat gctgatgcgc	1860
agaactctca cagctattgc tttgagattc tggagcggcg ataagctgat cagcctcgac	1920
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct	1980
ggaagggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat cgcatgtgt	2040
gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg	2100
ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggaggaac ccctagtgat	2160
ggagttggcc actccctctc tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt	2220
cgcccgacgc ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc gcagctgcct	2280
gcagg	2285

<210> 59
 <211> 483
 <212> DNA
 <213> E. coli

<400> 59	
ggtaccatca gtctgattgc ggcgttagcg gtagattacg ttatcggcat ggaaaacgcc	60

atgccgtgga acctgcctgc cgatctcgcc tggtttaaac gcaacacctt aaataaaccc	120
gtgattatgg gccgccatac ctgggaatca atcggtcgtc cgttgccagg acgcaaaaat	180
attatcctca gcagtcaacc gagtacggac gatcgcgtaa cgtgggtgaa gtcggtggat	240
gaagccatcg cggcgtgtgg tgacgtacca gaaatcatgg tgattggcgg cggtcgcgtt	300
attgaacagt tcttgccaaa agcgcaaaaa ctgtatctga cgcatatcga cgcagaagt	360
gaaggcgaca cccatttccc ggattacgag ccggatgact gggaatcggg attcagcgaa	420
ttccacgatg ctgatgcga gaactctcac agctattgct ttgagattct ggagcggcga	480
taa	483

<210> 60

<211> 3440

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-hPou4F3-T2A-mScarlet-DD

<400> 60

cctgcaggca gctgcgcgtc cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
agggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaattggc	180
cgcctggctg accgccaac gacccccgcc cattgacgtc aataatgacg tatgttccca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaatg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggag tggatagcgg ttgactcac ggggatttcc aagtctccac ccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaatg ggcggtaggc gtgtacgggt ggagggtctat ataagcagag	660
ctcgtttagt gaaccgtcag atcgcttggg gacgcatgat ggccatgaac tccaagcagc	720
ctttcggcat gcacccggtg ctgcaagaac ccaaattctc cagtctgcac tctggctccg	780
aggccatgcg ccgagtctgt ctcccagccc cgcagctgca gggtaatat tttggaagct	840

ttgatgagag cctgctggca cgcgccgaag ctctggcggc ggtggatatc gtctcccacg	900
gcaagaacca tccgttcaag cccgacgcca cctaccatac catgagcagc gtgccctgca	960
cgtccacttc gtccaccgtg cccatctccc acccagctgc gctcacctca caccctcacc	1020
acgccgtgca ccagggcctc gaaggcgacc tgctggagca catctcgccc acgctgagtg	1080
tgagcggcct gggcgctccg gaacactcgg tgatgcccgc acagatccat ccacaccacc	1140
tgggcgccat gggccacctg caccaggcca tgggcatgag tcacccgcac accgtggccc	1200
ctcatagcgc catgcctgca tgcctcagcg acgtggagtc agacccgcgc gagctggaag	1260
ccttcgccga gcgcttcaag cagcggcgca tcaagctggg ggtgaccag gcggacgtgg	1320
gcgcggctct ggctaattct aagatccccg gcgtgggctc gctgagccaa agcaccatct	1380
gcaggttcga gtctctcact ctctcgaca acaacatgat cgctctcaag ccggtgctcc	1440
aggcctggtt ggaggaggcc gaggccgcct accgagagaa gaacagcaag ccagagctct	1500
tcaacggcag cgaacggaag cgcaaacgca cgtccatcgc ggcgccggag aagcgttcac	1560
tcgaggccta tttcgctatc cagccacgtc cttcatctga gaagatcgcg gccatcgctg	1620
agaaactgga ccttaaaaag aacgtggtga gagtctgggt ctgcaaccag agacagaaac	1680
agaaacgaat gaagtattcg gctgtccacg ttaacgatta caaggatgac gacgataagg	1740
actataagga cgatgatgac aaggactaca aagatgatga cgataaagga tccggcgagg	1800
gcagaggaag tctgctaaca tgcggtgacg tcgaggagaa tcctggccca atgcctaaga	1860
agaagcgga agtcggctcc ggcgtgagca agggcgaggc agtgatcaag gagttcatgc	1920
ggttcaaggt gcacatggag ggctccatga acggccacga gttcgagatc gagggcgagg	1980
gcgagggccg cccctacgag ggcacccaga ccgccaagct gaaggtgacc aagggtggcc	2040
ccctgccctt ctcttgggac atcctgtccc ctacgttcat gtacggctcc agggccttca	2100
tcaagcacc cgccgacatc cccgactact ataagcagtc cttccccgag ggcttcaagt	2160
gggagcgcgt gatgaacttc gaggacggcg gcgccgtgac cgtgaccag gacacctccc	2220
tggaggacgg caccctgatc tacaaggtga agctccgcgg caccaacttc cctcctgacg	2280
gccccgtaat gcagaagaag acaatgggct gggaagcgtc caccgagcgg ttgtaccccg	2340
aggacggcgt gctgaagggc gacattaaga tggccctgcg cctgaaggac ggcgggcgct	2400
acctggcgga cttcaagacc acctacaagg ccaagaagcc cgtgcagatg cccggcgcct	2460

acaacgtcga ccgcaagttg gacatcacct cccacaacga ggactacacc gtggtggaac	2520
agtacgaacg ctccgagggc cgccactcca ccggcggcat ggacgagctg tacaagggta	2580
ccatcagtct gattgcggcg ttagcggtag attacgttat cggcatggaa aacgccatgc	2640
cgtggaacct gcctgccgat ctgcctggt ttaaacgcaa caccttaaataa aaacccgtga	2700
ttatgggccc ccatacctgg gaatcaatcg gtcgtccgtt gccaggacgc aaaaatatta	2760
tcctcagcag tcaaccgagt acggacgata gcgtaacgtg ggtgaagtcg gtggatgaag	2820
ccatcgcggc gtgtggtgac gtaccagaaa tcatggtgat tggcggcggc cgcgttattg	2880
aacagttctt gccaaaagcg caaaaactgt atctgacgca tatcgacgca gaagtggaag	2940
gcgacacca tttcccggat tacgagccgg atgactggga atcggtattc agcgaattcc	3000
acgatgctga tgcgcagaac tctcacagct attgctttga gattctggag cggcgataag	3060
ctgatcagcc tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctccccctg	3120
gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaaa atgaggaaat	3180
tgcacgcgat tgtctgagta ggtgtcattc tattctgggg ggtgggggtg ggaggacag	3240
caagggggag gattgggaag acaatagcag gcatgctggg gatgcgggtg gctctatgga	3300
ggaacccta gtgatggagt tggccactcc ctctctgcgc gctcgtcgc tctactgaggc	3360
cgggcgacca aaggtcgccc gacgccggg ctttgcccgg gcggcctcag tgagcgagcg	3420
agcgcgcagc tgcctgcagg	3440

<210> 61
 <211> 1020
 <212> DNA
 <213> Homo sapiens

<400> 61	
atgatggcca tgaactccaa gcagcctttc ggcatgcacc cggctgctgca agaaccctaaa	60
ttctccagtc tgcactctgg ctccgaggcc atgcgccgag tctgtctccc agccccgcag	120
ctgcagggtg atatattttg aagctttgat gagagcctgc tggcacgcgc cgaagctctg	180
gcggcgggtg atatcgtctc ccacggcaag aaccatccgt tcaagcccga cgccacctac	240
cataccatga gcagcgtgcc ctgcacgtcc acttcgtcca ccgtgcccac ctcccacca	300
gctgcgctca cctcacaccc tcaccacgcc gtgcaccagg gcctcgaagg cgacctgctg	360

gagcacatct cgccccacgct gagtgtgagc ggcctgggcg ctccggaaca ctcggtgatg	420
cccgcacaga tccatccaca ccacctgggc gccatgggcc acctgcacca ggccatgggc	480
atgagtcacc cgcacaccgt ggcccctcat agcgccatgc ctgcatgcct cagcgacgtg	540
gagtcagacc cgcgcgagct ggaagccttc gccgagcgct tcaagcagcg gcgcatcaag	600
ctgggggtga cccaggcgga cgtgggcgcg gctctggcta atctcaagat ccccggcgtg	660
ggctcgctga gccaaagcac catctgcagg ttcgagtctc tcaactctctc gcacaacaac	720
atgatcgctc tcaagccggt gctccaggcc tggttggagg aggccgaggc cgcctaccga	780
gagaagaaca gcaagccaga gctcttcaac ggcagcgaac ggaagcgcaa acgcacgtcc	840
atcgcggcgc cggagaagcg ttcaactcgag gcctatttcg ctatccagcc acgtccttca	900
tctgagaaga tcgcggccat cgctgagaaa ctggacctta aaaagaacgt ggtgagagtc	960
tggttctgca accagagaca gaaacagaaa cgaatgaagt attcggctgt ccacgttaac	1020

<210> 62
 <211> 81
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 3x FLAG

<400> 62	
gattacaagg atgacgacga taaggactat aaggacgatg atgacaagga ctacaaagat	60
gatgacgata aaggatccgg c	81

<210> 63
 <211> 54
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> T2A

<400> 63	
gagggcagag gaagtctgct aacatgcggt gacgtcgagg agaatcctgg ccca	54

<210> 64
 <211> 3692
 <212> DNA
 <213> Artificial Sequence

<220>

<223> pITR-CMV-hGF11-T2A-mScarlet-DD

<400> 64

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgctc ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaattggcc	180
cgcctggctg accgcccac gacccccgcc cattgacgtc aataatgacg tatgttccca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaatg gcccgctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg ttgactcac ggggatttcc aagtctccac cccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaag ggcggtaggc gtgtacgggt ggaggtctat ataagcagag	660
ctcgtttagt gaaccgtcag atcgctgga gacgcatgcc gcgctcattt ctcgtcaaaa	720
gcaagaaggc tcacagctac caccagccgc gctccccagg accagactat tccctccgtt	780
tagagaatgt accggcgcct agccgagcag acagcacttc aaatgcaggc ggggcgaagg	840
cggagccccg ggaccgtttg tccccgaat cgcagctgac cgaagcccca gacagagcct	900
ccgcatcccc agacagctgc gaaggcagcg tctgcgaacg gagctcggag tttgaggact	960
tctggaggcc cccgtcacc tccgcgtctc cagcctcgga gaagtcaatg tgcccatcgc	1020
tggacgaagc ccagcccttc cccctgcctt tcaaaccgta ctcatggagc ggcttggcgg	1080
gttctgacct gcggcacctg gtgcagagct accgaccgtg tggggccctg gagcgtggcg	1140
ctggcctggg cctcttctgc gaaccgccc cggagcctgg ccacccggcc gcgctgtacg	1200
gcccgaagcg ggctgccggc ggcgcggggg ccggggcgcc agggagctgc agcgcagggg	1260
ccggtgccac cgctggccct ggcctagggc tctacggcga cttcgggtct gcggcagccg	1320
ggctgtatga gaggcccacg gcagcggcgg gcttgctgta ccccgagcgt ggccacgggc	1380
tgcacgcaga caaggcgct ggcgtcaagg tggagtcgga gctgctgtgc accgcctgc	1440
tgctgggcgg cggctctac aagtgcacat agtgcagcaa ggtgttctcc acgccgcacg	1500

ggctcgaggt gcacgtgcgc aggtcccaca gcggcaccag accctttgcc tgcgagatgt	1560
gcggcaagac cttcgggcac gcggtgagcc tggagcagca caaagccgtg cactcgcagg	1620
aacggagctt tgactgtaag atctgtggga agagcttcaa gaggtcatcc aactgtcca	1680
cacacctgct tatccactca gacactcggc cctaccctg tcagtactgt ggcaagaggt	1740
tccaccagaa gtcagacatg aagaaacaca ctttcatcca cactggtgag aagcctcaca	1800
agtgccaggt gtgcggcaag gcattcagcc agagctcaa cctcatcacc cacagccgca	1860
aacacacagg cttcaagccc ttcggctgcg acctctgtgg gaagggtttc cagaggaagg	1920
tggacctccg aaggcaccgg gagacgcagc atgggctcaa agttaacgat tacaaggatg	1980
acgacgataa ggactataag gacgatgatg acaaggacta caaagatgat gacgataaag	2040
gatccggcga gggcagagga agtctgctaa catgcggtga cgtcaggag aatcctggcc	2100
caatgcctaa gaagaagcgg aaagtcggct ccggcgtgag caagggcgag gcagtgatca	2160
aggagtcat gcggttcaag gtgcacatgg agggctccat gaacggccac gagttcgaga	2220
tcgagggcga gggcgagggc cgccccctacg agggcaccca gaccgccaag ctgaaggatga	2280
ccaaggggtg cccctgccc ttctcctggg acatcctgtc ccctcagttc atgtacggct	2340
ccagggcctt catcaagcac cccgccgaca tccccgacta ctataagcag tccttccccg	2400
agggcttcaa gtgggagcgc gtgatgaact tcgaggacgg cggcgccgtg accgtgaccc	2460
aggacacctc cctggaggac ggcaccctga tctacaaggt gaagctccgc ggcaccaact	2520
tccctcctga cggccccgta atgcagaaga agacaatggg ctgggaagcg tccaccgagc	2580
ggttgtagcc cgaggacggc gtgctgaagg gcgacattaa gatggccctg cgcctgaagg	2640
acggcggccg ctacctggcg gacttcaaga ccacctaca ggccaagaag cccgtgcaga	2700
tgcccggcgc ctacaacgtc gaccgcaagt tggacatcac ctcccacaac gaggactaca	2760
ccgtgggtgga acagtacgaa cgctccgagg gccgccactc caccggcggc atggacgagc	2820
tgtacaaggg taccatcagt ctgattgcgg cgttagcggg agattacgtt atcggcatgg	2880
aaaacgccat gccgtggaac ctgcctgccg atctcgcctg gtttaaagc aacaccttaa	2940
ataaaccctg gattatgggc cgccatacct gggaatcaat cggctgtccg ttgccaggac	3000
gcaaaaatat tatcctcagc agtcaaccga gtacggacga tcgcgtaacg tgggtgaagt	3060
cgggtgatga agccatcgcg gcgtgtggtg acgtaccaga aatcatggtg attggcggcg	3120

gtcgcgttat tgaacagttc ttgccaaaag cgcaaaaact gtatctgacg catatcgacg	3180
cagaagtgga aggcgacacc catttcccgg attacgagcc ggatgactgg gaatcggtat	3240
tcagcgaatt ccacgatgct gatgcgcaga actctcacag ctattgcttt gagattctgg	3300
agcggcgata agctgatcag cctcgactgt gccttctagt tgccagccat ctgttgtttg	3360
cccctcccc gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata	3420
aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtgggg	3480
ggggcaggac agcaaggggg aggattggga agacaatagc aggcattgctg gggatgcgg	3540
gggctctatg gaggaacccc tagtgatgga gttggccact ccctctctgc gcgctcgtc	3600
gctcactgag gccgggagc caaaggctgc cgcagcccg ggctttgccc gggcggcctc	3660
agtgagcgag cgagcgcgca gctgcctgca gg	3692

<210> 65
 <211> 1272
 <212> DNA
 <213> Homo sapiens

<400> 65	
atgccgcgct catttctcgt caaaagcaag aaggctcaca gctaccacca gccgcgctcc	60
ccaggaccag actattccct ccgttttagag aatgtaccgg cgcctagccg agcagacagc	120
acttcaaatt caggcggggc gaaggcggag ccccgaggacc gtttgtcccc cgaatcgag	180
ctgaccgaag cccagacag agcctccgca tcccagaca gctgcgaagg cagcgtctgc	240
gaacggagct cggagtttga ggacttctgg aggccccgt caccctccgc gtctccagcc	300
tcggagaagt caatgtgccc atcgctggac gaagcccagc ctttccccct gcctttcaaa	360
ccgtactcat ggagcggcct ggcggttct gacctgcggc acctggtgca gagctaccga	420
ccgtgtgggg ccctggagcg tggcgctggc ctgggcctct tctgcgaacc cgccccggag	480
cctggccacc cggccgcgct gtacggcccg aagcgggctg ccggcggcgc gggggccggg	540
gcgccaggga gctgcagcgc aggggcccgt gccaccgctg gccctggcct agggctctac	600
ggcgacttcg ggtctgcggc agccgggctg tatgagaggc ccacggcagc ggcgggcttg	660
ctgtaccccg agcgtggcca cgggctgcac gcagacaagg gcgctggcgt caaggtggag	720
tcggagctgc tgtgcacccg cctgctgctg ggcggcggct cctacaagt catcaagtgc	780
agcaaggtgt tctccacgcc gcacgggctc gaggtgcacg tgcgcaggct ccacagcggc	840

accagaccct ttgcctgcga gatgtgcggc aagaccttcg ggacgcggt gagcctggag	900
cagcaciaag ccgtgcactc gcaggaacgg agctttgact gtaagatctg tgggaagagc	960
ttcaagaggt catccacact gtccacacac ctgcttatcc actcagacac tcggccctac	1020
ccctgtcagt actgtggcaa gaggttccac cagaagtcag acatgaagaa acacactttc	1080
atccacactg gtgagaagcc tcacaagtgc cagggtgtgcg gcaaggcatt cagccagagc	1140
tccaacctca tcacccacag ccgcaaacac acaggcttca agcccttcgg ctgcgacctc	1200
tgtgggaagg gtttccagag gaaggtggac ctccgaaggc accgggagac gcagcatggg	1260
ctcaaagtta ac	1272

<210> 66

<211> 3488

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-hATOH1-T2A-mScarlet-DD

<400> 66

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
agggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaattggcc	180
cgcttggttg accgcccac gacccccgcc cattgacgtc aataatgacg tatgttccca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaatg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac ccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaag ggcggttaggc gtgtacgggt ggagggtctat ataagcagag	660
ctcgtttagt gaaccgtcag atcgcttga gacgatgtc ccgcctgctg catgcagaag	720
agtgggctga agtgaaggag ttgggagacc accatcgcca gcccagccg catcatctcc	780
cgcaaccgcc gccgccgcc cagccacctg caactttgca ggcgagagag catcccgtct	840

acccgcctga gctgtccctc ctggacagca ccgaccacg cgcctggctg gctcccactt	900
tgcagggcat ctgcacggca cgcgccgccc agtatttgct acattccccg gagctgggtg	960
cctcagaggc cgctgcgccc cgggacgagg tggacggccg gggggagctg gtaaggagga	1020
gcagcggcgg tgccagcagc agcaagagcc ccgggccggt gaaagtgcgg gaacagctgt	1080
gcaagctgaa aggcgggggtg gtggtagacg agctgggctg cagccgcaa cggggccctt	1140
ccagcaaaca ggtgaatggg gtgcagaagc agagacggct agcagccaac gccagggagc	1200
ggcgcaggat gcatgggctg aaccacgcct tcgaccagct gcgcaatgtt atcccgtcgt	1260
tcaacaacga caagaagctg tccaaatatg agaccctgca gatggcccaa atctacatca	1320
acgccttgtc cgagctgcta caaacgcca gcggagggga acagccaccg ccgcctccag	1380
cctcctgcaa aagcgaccac caccaccttc gcaccgcggc ctcttatgaa gggggcgcgg	1440
gcaacgcgac cgcagctggg gctcagcagg cttccggagg gagccagcgg ccgacccgc	1500
ccgggagttg ccggactcgc ttctcagccc cagcttctgc gggagggtac tcggtgcagc	1560
tggacgctct gcacttctcg actttcgagg acagcgccct gacagcgatg atggcgcaaa	1620
agaatttgtc tccttctctc cccgggagca tcttgagcc agtgcaggag gaaaacagca	1680
aaacttcgcc tcggtccac agaagcgacg gggaattttc ccccatc cattacagt	1740
actcggatga ggcaagtgtt aacgattaca aggatgacga cgataaggac tataaggacg	1800
atgatgaaa ggactacaaa gatgatgacg ataaaggatc cggcgagggc agaggaagtc	1860
tgctaacatg cggtgacgtc gaggagaatc ctggcccaat gcctaagaag aagcggaaag	1920
tcggctccgg cgtgagcaag ggcgaggcag tgatcaagga gttcatgcgg ttcaaggtgc	1980
acatggaggg ctccatgaac ggccacgagt tcgagatcga gggcgagggc gagggccgcc	2040
cctacgaggg caccagacc gccaaagctga aggtgaccaa ggggtggccc ctgcccttct	2100
cctgggacat cctgtcccct cagttcatgt acggctccag ggccttcac aagcaccgc	2160
ccgacatccc cgactactat aagcagtcct tccccgaggg cttcaagtgg gagcgcgtga	2220
tgaacttcga ggacggcggc gccgtgaccg tgaccagga cacctccctg gaggacggca	2280
ccctgatcta caaggtgaag ctccgcggca ccaacttccc tcctgacggc cccgtaatgc	2340
agaagaagac aatgggctgg gaagcgtcca ccgagcgggt gtaccccgag gacggcgtgc	2400
tgaagggcga cattaagatg gccctgcgcc tgaaggacgg cggccgctac ctggcggact	2460

tcaagaccac ctacaaggcc aagaagcccg tgcagatgcc cggcgcctac aacgtcgacc	2520
gcaagttgga catcacctcc cacaacgagg actacaccgt ggtggaacag tacgaacgct	2580
ccgagggccg ccactccacc ggcggcatgg acgagctgta caagggtagc atcagtctga	2640
ttgcggcggt agcggtagat tacgttatcg gcatggaaaa cgccatgccg tggaacctgc	2700
ctgccgatct cgcctggttt aaacgaaca ccttaaataa acccgtgatt atgggcccgc	2760
atacctggga atcaatcggg cgtccgttgc caggacgcaa aaatattatc ctcagcagtc	2820
aaccgagtac ggacgatcgc gtaacgtggg tgaagtcggg ggatgaagcc atcgcggcgt	2880
gtggtgacgt accagaaatc atggtgattg gcggcggtcg cgttattgaa cagttcttgc	2940
caaaagcgca aaaactgtat ctgacgcata tcgacgcaga agtggaaggc gacacccatt	3000
tcccggatta cgagccggat gactgggaat cggtattcag cgaattccac gatgctgatg	3060
cgcagaactc tcacagctat tgctttgaga ttctggagcg gcgataagct gatcagcctc	3120
gactgtgcct tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac	3180
cctggaagggt gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg	3240
tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga	3300
ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggagg aaccctagt	3360
gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaa	3420
ggtcgcccga cgccggggt ttgcccgggc ggcctcagtg agcgagcgag cgcgcagctg	3480
cctgcagg	3488

<210> 67
 <211> 1068
 <212> DNA
 <213> Homo sapiens

<400> 67	
atgtcccgcc tgctgcatgc agaagagtgg gctgaagtga aggagttggg agaccacat	60
cgccagcccc agccgcatca tctcccga cgcgccgc cgccgagcc acctgcaact	120
ttgcaggcga gagagcatcc cgtctacccg cctgagctgt ccctcctgga cagcaccgac	180
ccacgcgcct ggctggctcc cactttgcag ggcatctgca cggcacgcgc cgcccagtat	240
ttgctacatt ccccggagct ggggtgcctca gaggccgctg cgccccggga cgaggtggac	300

ggccgggggg agctggtaag gaggagcagc ggcggtgcc a gcagcagcaa gagccccggg	360
ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg gggtgggtgt agacgagctg	420
ggctgcagcc gccaacgggc cccttccagc aaacaggtga atgggggtgca gaagcagaga	480
cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac	540
cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagacc	600
ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gccagcgga	660
ggggaacagc caccgccgcc tccagcctcc tgcaaaagcg accaccacca ctttcgcacc	720
gcggcctcct atgaaggggg cgcgggcaac gcgaccgcag ctgggggtca gcaggcttcc	780
ggagggagcc agcggccgac cccgcccggg agttgccgga ctgccttctc agccccagct	840
tctgcgggag ggtactcggg gcagctggac gctctgcact tctcgacttt cgaggacagc	900
gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg	960
cagccagtgc aggaggaaaa cagcaaaact tcgcctcggg cccacagaag cgacggggaa	1020
ttttcccccc attcccatta cagtgactcg gatgaggcaa gtgttaac	1068

<210> 68

<211> 3690

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-Luc2-T2A-mScarlet-U6-Hes1-S3

<400> 68

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagttag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaatggcc	180
cgcctggctg accgccaac gacccccgcc cattgacgtc aataatgacg tatgttccca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaatg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac cccattgacg	540

tcaatgggag	tttgttttgg	caccaaaatc	aacgggactt	tccaaaatgt	cgtaacaact	600
ccgccccatt	gacgcaaatg	ggcggtaggc	gtgtacgggtg	ggaggtctat	ataagcagag	660
ctcgttttagt	gaaccgtcag	atcgcctgga	gacgcatgga	agatgccaaa	aacattaaga	720
agggccccagc	gccatttctac	ccactcgaag	acgggaccgc	cggcgagcag	ctgcacaaaag	780
ccatgaagcg	ctacgccctg	gtgcccggca	ccatcgcctt	taccgacgca	catatcgagg	840
tggacattac	ctacgccgag	tacttcgaga	tgagcgttcg	gctggcagaa	gctatgaagc	900
gctatgggct	gaatacaaac	catcggatcg	tggtgtgcag	cgagaatagc	ttgcagttct	960
tcatgcccgt	gttgggtgcc	ctgttcatcg	gtgtggctgt	ggccccagct	aacgacatct	1020
acaacgagcg	cgagctgctg	aacagcatgg	gcatcagcca	gcccaccgtc	gtattcgtga	1080
gcaagaaagg	gctgcaaaag	atcctcaacg	tgcaaaagaa	gtaccgatc	atacaaaaga	1140
tcatcatcat	ggatagcaag	accgactacc	agggcttcca	aagcatgtac	accttcgtga	1200
cttcccattt	gccacccggc	ttcaacgagt	acgacttcgt	gcccagagagc	ttcgaccggg	1260
aaaaaacat	cgccctgatc	atgaacagta	gtggcagtac	cggattgccc	aagggcgtag	1320
ccctaccgca	ccgcaccgct	tgtgtccgat	tcagtcatgc	ccgcgacccc	atcttcggca	1380
accagatcat	ccccgacacc	gctatcctca	gcgtgggtgcc	atttcaccac	ggcttcggca	1440
tgttcaccac	gctgggctac	ttgatctgcg	gctttcgggt	cgtgctcatg	taccgcttcg	1500
aggaggagct	attcttgcg	agcttgcaag	actataagat	tcaatctgcc	ctgctgggtgc	1560
ccacactatt	tagcttcttc	gctaagagca	ctctcatcga	caagtacgac	ctaagcaact	1620
tgcacgagat	cgccagcggc	ggggcgccgc	tcagcaagga	ggtaggtgag	gccgtggcca	1680
aacgcttcca	cctaccaggc	atccgccagg	gctacggcct	gacagaaaca	accagcgcca	1740
ttctgatcac	ccccgaaggg	gacgacaagc	ctggcgcagt	aggcaagggtg	gtgcccttct	1800
tcgaggctaa	ggtggtggac	ttggacacag	gtaagacact	gggtgtgaac	cagcgcgggcg	1860
agctgtgcgt	ccgtggcccc	atgatcatga	gcggctacgt	taacaacccc	gaggctacaa	1920
acgctctcat	cgacaaggac	ggctggctgc	acagcggcga	catcgcctac	tgggacgagg	1980
acgagcactt	cttcatcgtg	gaccggctga	agagcctgat	caaatacaag	ggctaccagg	2040
tagccccagc	cgaactggag	agcatcctgc	tgcaacaccc	caacatcttc	gacgccgggg	2100
tcgccggcct	gcccgacgac	gatgccggcg	agctgcccgc	cgcagtcgtc	gtgctggaac	2160

acggtaaaac catgaccgag aaggagatcg tggactatgt ggccagccag gttacaaccg	2220
ccaagaagct gcgcggtggt gttgtgttcg tggacgaggt gcctaaagga ctgaccggca	2280
agttggacgc ccgcaagatc cgcgagattc tcattaaggc caagaagggc ggcaagatcg	2340
ccgtgggctc cggagagggc agaggaagtc tgctaacatg cggtgacgtc gaggagaatc	2400
ctggcccaat ggtgagcaag ggcgaggcag tgatcaagga gttcatgcgg ttcaagggtc	2460
acatggaggg ctccatgaac ggccacgagt tcgagatcga gggcgagggc gaggggccgc	2520
cctacgaggg caccagacc gccaaagctga aggtgaccaa ggggtggcccc ctgcccttct	2580
cctgggacat cctgtccctt cagttcatgt acggctccag ggccttcac aagcaccctg	2640
ccgacatccc cgactactat aagcagtcct tccccgaggg cttcaagtgg gagcgcgtga	2700
tgaacttcga ggacggcggc gccgtgaccg tgaccagga cacctccctg gaggacggca	2760
ccctgatcta caaggtgaag ctccgcggca ccaacttccc tcctgacggc cccgtaatgc	2820
agaagaagac aatgggctgg gaagcgtcca ccgagcgggt gtaccccgag gacggcgtgc	2880
tgaagggcga cattaagatg gccctgcgcc tgaaggacgg cggccgctac ctggcggact	2940
tcaagaccac ctacaaggcc aagaagcccg tgcagatgcc cggcgcctac aacgtcgacc	3000
gcaagttgga catcacctcc cacaacgagg actacaccgt ggtggaacag tacgaacgt	3060
ccgagggccg ccactccacc ggcggcatgg acgagctgta caagtaagct gatcagctc	3120
gataagatac attgatgagt ttggacaaac cacaactaga atgcagtga aaaaatgctt	3180
tatttgtgaa atttgtgatg ctattgcttt atttgttaacc attataagct gcaataaaca	3240
agttaaggtc gggcaggaag agggcctatt tcccatgatt cttcatatt tgcatatacg	3300
atacaaggct gttagagaga taattagaat taatttgact gtaaacacaa agatattagt	3360
acaaaatacg tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg	3420
ttttaaaatg gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt	3480
atatatcttg tggaaaggac gaaacaccga aagtcacaa agcctatcga aataggcttt	3540
gatgactttc ttttttagga acccctagtg atggagttgg ccactccctc tctgcgcgct	3600
cgctcgctca ctgaggccgg gcgaccaaag gtcgcccgc gcccgggctt tgcccgggcg	3660
gcctcagtga gcgagcgagc gcgcagctgc	3690

<211> 1659
<212> DNA
<213> Artificial Sequence

<220>
<223> luciferase sequence

<400> 69
atggaagatg ccaaaaacat taagaagggc ccagcgccat tctaccact cgaagacggg 60
accgccggcg agcagctgca caaagccatg aagcgctacg ccctggtgcc cggcaccatc 120
gcctttaccg acgcacatat cgaggtggac attacctacg ccgagtactt cgagatgagc 180
gttcggctgg cagaagctat gaagcgctat gggctgaata caaaccatcg gatcgtggtg 240
tgcagcgaga atagcttgca gttcttcatg cccgtgttgg gtgccctgtt catcgggtgtg 300
gctgtggccc cagctaacga catctacaac gagcgcgagc tgctgaacag catggggcatc 360
agccagccca ccgtcgtatt cgtgagcaag aaagggtgc aaaagatcct caacgtgcaa 420
aagaagctac cgatcataca aaagatcatc atcatggata gcaagaccga ctaccagggc 480
ttccaaagca tgtacacctt cgtgacttcc catttgccac ccggcttcaa cgagtacgac 540
ttcgtgcccc agagcttcga ccgggacaaa accatcgccc tgatcatgaa cagtagtggc 600
agtaccgatg tgcccaaggg cgtagcccta ccgcaccgca ccgcttgtgt ccgattcagt 660
catgcccgcg accccatctt cggcaaccag atcatccccg acaccgctat cctcagcgtg 720
gtgccatttc accacggctt cggcatgttc accacgctgg gctacttgat ctgcggcttt 780
cgggtcgtgc tcatgtaccg cttcaggag gagctattct tgcgcagctt gcaagactat 840
aagattcaat ctgccctgct ggtgcccaca ctatttagct tcttcgctaa gagcactctc 900
atcgacaagt acgacctaag caacttgac gagatcgcca gcggcggggc gccgctcagc 960
aaggaggtag gtgaggccgt ggccaaacgc ttccacctac caggcatccg ccagggtac 1020
ggcctgacag aaacaaccag cgccattctg atcacccccg aaggggacga caagcctggc 1080
gcagtaggca aggtggtgcc cttcttcgag gctaagggtg tggacttgga cacaggtgag 1140
aactgggtg tgaaccagcg cggcgagctg tgcgtccgtg gcccacatgat catgagcggc 1200
tacgttaaca accccgaggc tacaacgct ctcatcgaca aggacggctg gctgcacagc 1260
ggcgacatcg cctactggga cgaggacgag cacttcttca tcgtggaccg gctgaagagc 1320
ctgatcaaat acaagggtc ccaggtagcc ccagccgaac tggagagcat cctgctgcaa 1380

cacccaaca tcttcgacgc cggggtcgcc ggcctgcccg acgacgatgc cggcgagctg	1440
cccgccgcag tcgtcgtgct ggaacacggt aaaacatga ccgagaagga gatcgtggac	1500
tatgtggcca gccaggttac aaccgccaag aagctgcgcg gtggtgttgt gttcgtggac	1560
gaggtgccta aaggactgac cggcaagttg gacgcccgca agatccgcga gattctcatt	1620
aaggccaaga agggcggcaa gatcgccgtg ggctccgga	1659

<210> 70
 <211> 137
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> SV40pA

<400> 70	
gctgatcagc ctcgataaga tacattgatg agtttggaca aaccacaact agaatgcagt	60
gaaaaaaaaatg ctttatttgt gaaatttgtg atgctattgc tttatttgtta accattataa	120
gctgcaataa acaagtt	137

<210> 71
 <211> 264
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> U6

<400> 71	
aaggtcgggc aggaagaggg cctatttccc atgattcctt catatttgca tatacgatac	60
aaggctgtta gagagataat tagaattaat ttgactgtaa acacaaagat attagtacaa	120
aatacgtgac gtagaaagta ataatttctt gggtagtttg cagttttaaa attatgtttt	180
aaaatggact atcatatgct taccgtaact tgaaagtatt tcgatttctt ggctttatat	240
atcttgtgga aaggacgaaa cacc	264

<210> 72
 <211> 50
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> siRNA DNA sequence

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (25)..(44)
<223> n is a, c, g, or t

<400> 72
nnnnnnnnnnn nnnnnnnnnn cgaannnnnn nnnnnnnnnn nnnntttttt 50

<210> 73
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> S3 shRNA

<400> 73
gaaagtcatc aaagcctat 19

<210> 74
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> S5 shRNA

<400> 74
actgcatgac ccagatcaa 19

<210> 75
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Kop shRNA

<400> 75
actgcatgac ccagatcaa 19

<210> 76
<211> 1444

<212> DNA
<213> Artificial Sequence

<220>
<223> pITR-CMV-mScarlet-bGHpA

<400> 76
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagttag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120
aggggttcct gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc 180
acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcacaaaaa 240
tcaacgggac tttccaaaat gtcgtaacaa ctccgcccc a ttgacgcaaa tgggcggtag 300
gcgtgtacgg tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgctg 360
gagacgcgtg agcaaggcg aggagtgat caaggagttc atgcggttca aggtgcacat 420
ggagggtctc atgaacggcc acgagttcga gatcgagggc gagggcgagg gccgccccta 480
cgagggcacc cagaccgcca agctgaaggt gaccaagggt ggccccctgc ctttctcctg 540
ggacatcctg tcccctcagt tcatgtacgg ctccagggcc ttcacaaagc accccgccga 600
catccccgac tactataagc agtccttccc cgagggtctc aagtgggagc gcgtgatgaa 660
cttcgaggac ggcggcgccg tgaccgtgac ccaggacacc tccctggagg acggcacctt 720
gatctacaag gtgaagctcc gcggcaccaa cttccctcct gacggccccg taatgcagaa 780
gaagacaatg ggctgggaag cgtccaccga gcggttgtag cccgaggacg gcgtgctgaa 840
gggcgacatt aagatggccc tgcgcctgaa ggacggcggc cgctacctgg cggacttcaa 900
gaccacctac aaggccaaga agcccgtgca gatgcccggc gcctacaacg tcgaccgcaa 960
gttggacatc acctcccaca acgaggacta caccgtgggtg gaacagtacg aacgctccga 1020
gggccgccac tccaccggcg gcatggacga gctgtacaag taagctgatc agcctcgact 1080
gtgccttcta gttgccagcc atctgttggt tggccctccc ccgtgccttc cttgacctg 1140
gaaggtgcca ctcccactgt cttttcctaa taaaatgagg aaattgcatc gcattgtctg 1200
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg 1260
gaagacaata gcaggcatgc tggggatgag gtgggctcta tggaggaacc ctagtgatg 1320
gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggag accaaaggctc 1380
gcccgcagcc cgggctttgc ccgggcggcc tcagttagcg agcgagcgcg cagctgcctg 1440

cagg

1444

<210> 77

<211> 1918

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-mScarlet-DD-bGHPA

<400> 77

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgctc ggcgaccttt	60
ggtcgccccg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
agggggttcct gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc	180
acgggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt ggccacaaaa	240
tcaacgggac tttccaaaat gtcgtaacaa ctccgcccc a ttgacgcaaa tgggcggtag	300
gcgtgtacgg tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgcttg	360
gagacgcgtg agcaagggcg aggcagtgat caaggagttc atgcggttca aggtgcacat	420
ggaggggctcc atgaacggcc acgagttcga gatcgagggc gagggcgagg gccgccccta	480
cgagggcacc cagaccgcca agctgaaggt gaccaagggt ggccccctgc ctttctcttg	540
ggacatcctg tcccctcagt tcatgtacgg ctccagggcc ttcacaaagc accccgccga	600
catccccgac tactataagc agtcctttcc cgagggttcc aagtgaggagc gcgtgatgaa	660
cttcgaggac ggcggcgccg tgaccgtgac ccaggacacc tccctggagg acggcacctt	720
gatctacaag gtgaagctcc gcggcaccaa cttccctcct gacggccccg taatgcagaa	780
gaagacaatg ggctgggaag cgtccaccga gcggttgtag cccgaggacg gcgtgctgaa	840
gggcgacatt aagatggccc tgcgcctgaa ggacggcggc cgctacctgg cggacttcaa	900
gaccacctac aaggccaaga agcccgtgca gatgcccggc gcctacaacg tcgaccgcaa	960
gttgacatc acctcccaca acgaggacta caccgtgggtg gaacagtacg aacgctccga	1020
gggcccgcac tccaccggcg gcatggacga gctgtacaag taaatcagtc tgattgcggc	1080
gtagcggtg gattacgtta tcggcatgga aaacgccatg ccgtggaacc tgcctgccga	1140
tctgcctgg tttaaacgca acaccttaaa taaaccctg attatgggcc gccatacctg	1200
ggaatcaatc ggtcgtccgt tgccaggacg caaaaatatt atcctcagca gtcaaccgag	1260

tacggacgat cgcgtaacgt ggggtgaagtc ggtggatgaa gccatcgcgg cgtgtggtga	1320
cgtaccagaa atcatggtga ttggcggcgg tcgcgttatt gaacagttct tgccaaaagc	1380
gcaaaaactg tatctgacgc atatcgacgc agaagtggaa ggcgacaccc atttcccga	1440
ttacgagccg gatgactggg aatcgggtatt cagcgaattc cacgatgctg atgcgcagaa	1500
ctctcacagc tattgctttg agattctgga gcggcgagct gatcagcctc gactgtgcct	1560
tctagttagc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt	1620
gccactcca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg	1680
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac	1740
aatagcaggc atgctgggga tgcgggtgggc tctatggagg aaccctagt gatggagttg	1800
gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaa ggtcgcccga	1860
cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag cgcgagctg cctgcagg	1918

<210> 78

<211> 3690

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-Luc2-T2A-GFP-U6-Hes1-S5

<400> 78

cctgcaggca gctgcgcgt cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgagagagg gaggggccaa ctccatcact	120
aggggttcct ctagatcca tatatggagt tccgcgttac ataacttac gtaaattggc	180
cgcttggtg accgccaac gacccccgcc cattgacgtc aataatgacg tatgttcca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaatg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac ccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaag ggcggtaggc gtgtacgggt ggaggtctat ataagcagag	660

ctcgtttagt gaaccgtcag atcgccctgga gacgcatgga agatgccaaa aacattaaga	720
agggcccagc gccattctac ccactcgaag acgggaccgc cggcgagcag ctgcacaaaag	780
ccatgaagcg ctacgccctg gtgcccggca ccatcgcctt taccgacgca catatcgagg	840
tggacattac ctacgccgag tacttcgaga tgagcgttcg gctggcagaa gctatgaagc	900
gctatgggct gaatacaaac catcggatcg tgggtgtgcag cgagaatagc ttgcagttct	960
tcatgcccgt gttgggtgcc ctgttcacg gtgtggctgt ggccccagct aacgacatct	1020
acaacgagcg cgagctgctg aacagcatgg gcatcagcca gcccaccgtc gtattcgtga	1080
gcaagaaagg gctgcaaaaag atcctcaacg tgcaaaaaga gctaccgatc atacaaaaga	1140
tcatcatcat ggatagcaag accgactacc agggcttcca aagcatgtac accttcgtga	1200
cttcccattt gccacccggc ttcaacgagt acgacttcgt gcccgagagc ttcgaccggg	1260
acaaaaccat cgccctgatc atgaacagta gtggcagtac cggattgccc aagggcgtag	1320
ccctaccgca ccgcaccgct tgtgtccgat tcagtcatgc ccgcgacccc atcttcggca	1380
accagatcat ccccgacacc gctatcctca gcgtgggtgcc atttcaccac ggcttcggca	1440
tgttcaccac gctgggctac ttgatctgcg gctttcgggt cgtgctcatg taccgcttcg	1500
aggaggagct attcttgcg agcttgcaag actataagat tcaatctgcc ctgctgggtgc	1560
ccacactatt tagcttcttc gctaagagca ctctcatcga caagtacgac ctaagcaact	1620
tgcacgagat cgccagcggc ggggcgccgc tcagcaagga ggtaggtgag gccgtggcca	1680
aacgcttcca cctaccaggc atccgccagg gctacggcct gacagaaaca accagcgcca	1740
ttctgatcac ccccgaggc gacgacaagc ctggcgcagt aggcaagggtg gtgcccttct	1800
tcgaggctaa ggtggtggac ttggacacag gtaagacact ggggtgtgaac cagcgcggcg	1860
agctgtgcgt ccgtggcccc atgatcatga gcggctacgt taacaacccc gaggctacaa	1920
acgctctcat cgacaaggac ggctggctgc acagcggcga catcgcctac tgggacgagg	1980
acgagcactt cttcatcgtg gaccggctga agagcctgat caaatacaag ggctaccagg	2040
tagccccagc cgaactggag agcatcctgc tgcaacaccc caacatcttc gacgccgggg	2100
tcgccggcct gcccgacgac gatgccggcg agctgcccgc cgcagtcgtc gtgctggaac	2160
acggtaaaac catgaccgag aaggagatcg tggactatgt ggccagccag gttacaaccg	2220
ccaagaagct gcgcggtggt gttgtgttcg tggacgaggt gcctaaagga ctgaccggca	2280

agttggacgc cgcgaagatc cgcgagattc tcattaaggc caagaagggc ggcaagatcg	2340
ccgtgggctc cggagagggc agagggaagtc tgctaacatg cggtgacgtc gaggagaatc	2400
ctggcccaat ggtgagcaag ggcgaggcag tgatcaagga gttcatgcgg ttcaagggtgc	2460
acatggaggg ctccatgaac ggccacgagt tcgagatcga gggcgagggc gagggccgcc	2520
cctacgaggg caccagacc gccaaagctga aggtgaccaa ggggtggccc ctgcccttct	2580
cctgggacat cctgtcccct cagttcatgt acggctccag ggccttcac aagcaccgcc	2640
ccgacatccc cgactactat aagcagtcct tccccgaggg cttcaagtgg gagcgcgtga	2700
tgaacttcga ggacggcggc gccgtgaccg tgaccagga cacctccctg gaggacggca	2760
ccctgatcta caaggtgaag ctccgaggca ccaacttccc tcctgacggc cccgtaatgc	2820
agaagaagac aatgggctgg gaagcgtcca ccgagcgggt gtaccccgag gacggcgtgc	2880
tgaagggcga cattaagatg gccctgcgcc tgaaggacgg cggccgctac ctggcggact	2940
tcaagaccac ctacaaggcc aagaagccc tgcagatgcc cggcgcctac aacgtcgacc	3000
gcaagttgga catcacctcc cacaacgagg actacaccgt ggtggaacag tacgaacgt	3060
ccgagggccg ccactccacc ggcggcatgg acgagctgta caagtaagct gatcagctc	3120
gataagatac attgatgagt ttggacaaac cacaactaga atgcagtga aaaaatgctt	3180
tatttgtgaa atttgtgatg ctattgcttt atttgtacc attataagct gcaataaaca	3240
agttaaggtc gggcaggaag agggcctatt tcccatgatt cttcatatt tgcatatacg	3300
atacaaggct gttagagaga taattagaat taatttgact gtaaacacaa agatattagt	3360
acaaaatacg tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg	3420
ttttaaatg gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt	3480
atatatcttg tggaaaggac gaaacaccac tgcattgaccc agatcaacga attgatctgg	3540
gtcatgcagt ttttttagga acccctagtg atggagttgg ccactccctc tctgcgcgct	3600
cgctcgctca ctgaggccgg gcgaccaaag gtcgcccgc gcccgggctt tgccggggcg	3660
gcctcagtga gcgagcgagc gcgcagctgc	3690

<210> 79

<211> 3691

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-CMV-Luc2-T2A-GFP-U6-Hes1-KOP

<400> 79

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgctc ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgagagagg gagtggccaa ctccatcact	120
aggggttcct ctagatccca tatatggagt tccgcgttac ataacttacg gtaaattggcc	180
cgcctggctg accgcccac gacccccgcc cattgacgtc aataatgacg tatgttccca	240
tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg	300
cccacttggc agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg	360
acggtaaattg gccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt	420
ggcagtacat ctacgtatta gtcacgcta ttaccatggt gatgcgggtt tggcagtaca	480
tcaatgggcg tggatagcgg ttgactcac ggggatttcc aagtctccac cccattgacg	540
tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact	600
ccgccccatt gacgcaaatt ggcggttaggc gtgtacggtg ggaggtctat ataagcagag	660
ctcgtttagt gaaccgtcag atcgctgga gacgcatgga agatgccaaa aacattaaga	720
agggcccagc gccattctac ccactcgaag acgggaccgc cggcgagcag ctgcacaaag	780
ccatgaagcg ctacgccctg gtgcccggca ccatcgcctt taccgacgca catatcgagg	840
tggacattac ctacgccgag tacttcgaga tgagcggttc gctggcagaa gctatgaagc	900
gctatgggct gaatacaaac catcggatcg tgggtgtgcag cgagaatagc ttgcagttct	960
tcatgcccgt gttgggtgcc ctgttcacg gtgtggctgt ggcccagct aacgacatct	1020
acaacgagcg cgagctgctg aacagcatgg gcatcagcca gccaccgtc gtattcgtga	1080
gcaagaaagg gctgcaaaag atcctcaacg tgcaaaagaa gctaccgatc atacaaaaga	1140
tcatcatcat ggatagcaag accgactacc agggcttcca aagcatgtac accttcgtga	1200
cttcccattt gccaccggc ttcaacgagt acgacttcgt gcccagagc ttcgaccggg	1260
acaaaaccat cgccctgac atgaacagta gtggcagtac cggattgccc aagggcgtag	1320
ccctaccgca ccgcaccgt tgtgtccgat tcagtcatgc ccgcgacccc atcttcggca	1380
accagatcat ccccgacacc gctatcctca gcgtgggtgcc atttcaccac ggcttcggca	1440
tgttcaccac gctgggctac ttgatctgcg gctttcgggt cgtgctcatg taccgcttcg	1500

aggaggagct attcttgcgc agcttgcaag actataagat tcaatctgcc ctgctggtgc	1560
ccacactatt tagcttcttc gctaagagca ctctcatcga caagtacgac ctaagcaact	1620
tgcacgagat cgccagcggc ggggcgccgc tcagcaagga ggtaggtgag gccgtggcca	1680
aacgcttcca cctaccaggc atccgccagg gctacggcct gacagaaaca accagcgcca	1740
ttctgatcac ccccgaaggg gacgacaagc ctggcgcagt aggcaagggtg gtgcccttct	1800
tcgaggctaa ggtggtggac ttggacacag gtaagacact gggtgtgaac cagcgcggcg	1860
agctgtgcgt ccgtggcccc atgatcatga gcggctacgt taacaacccc gaggctacaa	1920
acgctctcat cgacaaggac ggctggctgc acagcggcga catcgcctac tgggacgagg	1980
acgagcactt cttcatcgtg gaccggctga agagcctgat caaatacaag ggctaccagg	2040
tagccccagc cgaactggag agcatcctgc tgcaacaccc caacatcttc gacgccgggg	2100
tcgccggcct gcccgacgac gatgccggcg agctgcccgc cgcagtcgtc gtgctggaac	2160
acggtaaaac catgaccgag aaggagatcg tggactatgt ggccagccag gttacaaccg	2220
ccaagaagct gcgcgggtgt gttgtgttcg tggacgaggt gcctaaagga ctgaccggca	2280
agttggacgc ccgcaagatc cgcgagattc tcattaaggc caagaagggc ggcaagatcg	2340
ccgtgggctc cggagagggc agaggaagtc tgctaacatg cggtgacgtc gaggagaatc	2400
ctggcccaat ggtgagcaag ggcgaggcag tgatcaagga gttcatgcgg ttcaagggtc	2460
acatggaggg ctccatgaac ggccacgagt tcgagatcga gggcgagggc gagggccgcc	2520
cctacgaggg caccagacc gccaaagctga aggtgaccaa ggggtggcccc ctgcccttct	2580
cctgggacat cctgtcccct cagttcatgt acggctccag ggcccttcac aagcaccgcc	2640
ccgacatccc cgactactat aagcagtcct tccccgaggg cttcaagtgg gagcgcgtga	2700
tgaacttcga ggacggcggc gccgtgaccg tgaccagga cacctccctg gaggacggca	2760
ccctgatcta caaggtgaag ctccgcggca ccaacttccc tcctgacggc cccgtaatgc	2820
agaagaagac aatgggctgg gaagcgtcca ccgagcgggt gtaccccgag gacggcgtgc	2880
tgaagggcga cattaagatg gccctgcgcc tgaaggacgg cggccgctac ctggcggact	2940
tcaagaccac ctacaaggcc aagaagcccg tgcagatgcc cggcgcctac aacgtcgacc	3000
gcaagttgga catcacctcc cacaacgagg actacaccgt ggtggaacag tacgaacgct	3060
ccgagggccg ccactccacc ggcggcatgg acgagctgta caagtaagct gatcagcctc	3120

gataagatac attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt	3180
tatttgtgaa atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca	3240
agttaaggct gggcaggaag agggcctatt tcccatgatt ccttcatatt tgcatatacg	3300
atacaaggct gttagagaga taattagaat taatttgact gtaaacacaa agatattagt	3360
acaaaatacg tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg	3420
ttttaaaatg gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt	3480
atatatcttg tggaaggac gaaacaccgc agctgatata atggagaacg aattctccat	3540
tatatcagct gttttttagg aacccttagt gatggagttg gccactccct ctctgcgcgc	3600
tcgctcgctc actgaggccg ggcgaccaa ggctgcccga cgcccgggct ttgcccgggc	3660
ggcctcagtg agcgagcgag cgcgagctg c	3691

<210> 80
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Hes1-KOP

<400> 80	
gcagctgata taatggagaa	20

<210> 81
 <211> 4221
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pITR-CMV-mScarlet

<400> 81	
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
aggggttcct gcggccgcac gcgtctagat cccatatatg gaggttccgcg ttacataact	180
tacggtaaata ggccgcctg gctgaccgcc caacgacccc cgccattga cgtcaataat	240
gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggagta	300
tttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa gtacgcccc	360

tattgacgtc aatgacggta aatggcccg	ctggcattat gccagtaga tgaccttatg	420
ggactttcct acttggcagt acatctacgt	attagtcatc gctattacca tggatgacg	480
gttttggcag tacatcaatg ggcgtggata	gcggtttgac tcacggggat ttccaagtct	540
ccacccatt gacgtcaatg ggagtttgtt	ttggcaccaa aatcaacggg actttccaaa	600
atgtcgtaac aactccgccc cattgacgca	aatgggcggg aggcgtgtac ggtgggaggt	660
ctatataagc agagctcgtt tagtgaaccg	tcagatcgcc tggagacgca ccggtgccac	720
catgcctaag aagaagcgga aagtcggctc	cggcgtgagc aaggcgagg cagtgatcaa	780
ggagttcatg cggttcaagg tgcacatgga	gggctccatg aacggccacg agttcgagat	840
cgagggcgag ggcgagggcc gcccctacga	gggcacccag accgccaagc tgaaggtgac	900
caagggtggc cccctgccct tctcctggga	catcctgtcc cctcagttca tgtacggctc	960
cagggccttc atcaagcacc ccgccgacat	ccccgactac tataagcagt cttccccga	1020
gggcttcaag tgggagcgcg tgatgaactt	cgaggacggc ggcgccgtga ccgtgacca	1080
ggacacctcc ctggaggacg gcaccctgat	ctacaagggtg aagctccgcg gcaccaactt	1140
ccctcctgac ggccccgtaa tgcagaagaa	gacaatgggc tgggaagcgt ccaccgagcg	1200
gttgtagccc gaggacggcg tgctgaaggg	cgacattaag atggccctgc gcctgaagga	1260
cggcggccgc tacctggcgg acttcaagac	cacctacaag gccaagaagc ccgtgcagat	1320
gcccggcgcc tacaacgtcg accgcaagtt	ggacatcacc tcccacaacg aggactacac	1380
cgtggtggaa cagtacgaac gctccgaggg	ccgccactcc accggcggca tggacgagct	1440
gtacaagtaa taagagctcg ctgatcagcc	tcgactgtgc cttctagttg ccagccatct	1500
gttgtttgcc cttccccgt gccttccttg	accctggaag gtgccactcc cactgtcctt	1560
tcctaataaa atgaggaaat tgcacgcac	tgtctgagta ggtgtcattc tattctgggg	1620
ggtggggtgg ggcaggacag caagggggag	gattgggaag acaatagcag gcatgctggg	1680
gatgcggtgg gctctatgga agcttgaatt	cagctgacgt gcctcgacc gctaggaacc	1740
cctagtgatg gagttggcca ctccctctct	gcgcgctcgc tcgctcactg aggccggg	1800
accaaaggtc gcccgacgcc cgggctttgc	ccgggcggcc tcagtgagcg agcgagcgcg	1860
cagctgcctg caggggcgcc tgatgcggta	ttttctcctt acgcatctgt gcggtatttc	1920
acaccgcata cgtcaaagca accatagtag	gcgccctgta gcggcgcat aagcgggcg	1980

ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgcctcct	2040
ttcgcttttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat	2100
cgggggctcc ctttaggggtt ccgatttagt gcttttacggc acctcgaccc caaaaaactt	2160
gatttggttg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg	2220
acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac	2280
cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta	2340
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca	2400
attttatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga	2460
cacccgccaa cacccgctga cgcgccctga cgggcttgct tgctcccggc atccgcttac	2520
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg	2580
aaacgcgcga gacgaaagg cctcgtgata cgcctatttt tatagggtta tgatcatgaac	2640
aataaaaactg tctgcttaca taaacagtaa tacaaggggt gttatgagcc atattcaacg	2700
ggaaacgtcg aggccgcgat taaattccaa catggatgct gatttatatg ggtataaatg	2760
ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat cgcttgatat ggaagcccga	2820
tgcgccagag ttgtttctga aacatggcaa aggtagcggt gccaatgatg ttacagatga	2880
gatggtcaga ctaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat	2940
ccgtactcct gatgatgcat ggttactcac cactgcgatc cccggaaaaa cagcattcca	3000
ggtattagaa gaatatcctg attcaggtga aaatattggt gatgcgctgg cagtgttcct	3060
gcgccgggtg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg	3120
tctcgctcag gcgcaatcac gaatgaataa cggtttggt gatgcgagtg attttgatga	3180
cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa atgcataaac ttttgccatt	3240
ctcaccggat tcagtcgtca ctcatggtga tttctcactt gataacctta tttttgacga	3300
ggggaaatta ataggttgta ttgatgttg acgagtcgga atcgagacc gataccagga	3360
tcttgccatc ctatggaact gcctcgggtga gttttctcct tcattacaga aacggctttt	3420
tcaaaaatat ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga	3480
tgagtttttc taatctcatg accaaaatcc cttaacgtga gttttcggtc cactgagcgt	3540
cagaccccgt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct	3600

gctgcttgca aacaaaaaaaa ccaccgctac cagcgggtggt ttgtttgccg gatcaagagc	3660
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc	3720
ttctagtgtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc	3780
tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg	3840
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt	3900
cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg	3960
agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggat cgggtaagcg	4020
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt	4080
atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag	4140
gggggaggag cctatggaaa aacgccagca acgcggcctt tttacgggtt ctggcctttt	4200
gctggccttt tgctcacatg t	4221

<210> 82
 <211> 4701
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pITR-CMV-mScarlet-DD

<400> 82	
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagttag cgagcgagcg cgcagagagg gtagtgccaa ctccatcact	120
aggggttcct gcggccgcac gcgtctagat cccatatatg gagttccgcg ttacataact	180
tacggtaa at ggcccgcctg gctgaccgcc caacgacccc cgcccattga cgtcaataat	240
gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggagta	300
tttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa gtacgcccc	360
tattgacgtc aatgacggta aatggcccgc ctggcattat gcccagtaca tgaccttatg	420
ggactttcct acttggcagt acatctacgt attagtcac gctattacca tgggtgatgcg	480
gttttggcag tacatcaatg ggcgtaggata gcggtttgac tcacggggat ttccaagtct	540
ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg actttccaaa	600
atgtcgtaac aactccgccc cattgacgca aatgggcggg aggcgtgtac ggtgggaggt	660

ctatataagc agagctcgtt tagtgaaccg tcagatcgcc tggagacgca ccggtgccac	720
catgcctaag aagaagcgga aagtcggctc cggcgtgagc aagggcgagg cagtgatcaa	780
ggagttcatg cggttcaagg tgcacatgga gggctccatg aacggccacg agttcgagat	840
cgagggcgag ggcgagggcc gcccctacga gggcacccag accgccaagc tgaaggtgac	900
caagggtggc cccctgccct tctcctggga catcctgtcc cctcagttca tgtacggctc	960
cagggccttc atcaagcacc ccgccgacat ccccgactac tataagcagt cttccccga	1020
gggcttcaag tgggagcgcg tgatgaactt cgaggacggc ggcgccgtga ccgtgacca	1080
ggacacctcc ctggaggacg gcaccctgat ctacaagggtg aagctccgcg gcaccaactt	1140
ccctcctgac ggccccgtaa tgcagaagaa gacaatgggc tgggaagcgt ccaccgagcg	1200
gttgtagccc gaggacggcg tgctgaaggg cgacattaag atggccctgc gcctgaagga	1260
cggcggccgc tacctggcgg acttcaagac cacctacaag gccagaagc ccgtgcagat	1320
gcccggcgcc tacaacgtcg accgcaagtt ggacatcacc tcccacaacg aggactacac	1380
cgtggtggaa cagtacgaac gctccgaggg ccgccactcc accggcggca tggacgagct	1440
gtacaagggt accatcagtc tgattgcggc gttagcggtg gattacgtta tcggcatgga	1500
aaacgccatg ccgtggaacc tgccctgccga tctcgccctgg tttaaacgca acaccttaaa	1560
taaacccgtg attatgggcc gccatacctg ggaatcaatc ggctgtccgt tgccaggacg	1620
caaaaatatt atcctcagca gtcaaccgag tacggacgat cgcgtaacgt ggggtgaagtc	1680
gggtggatgaa gccatcgcgg cgtgtggtga cgtaccagaa atcatggtga ttggcggcgg	1740
tcgcgttatt gaacagttct tgccaaaagc gcaaaaactg tatctgacgc atatcgacgc	1800
agaagtggaa ggcgacaccc atttcccgga ttacgagccg gatgactggg aatcggtatt	1860
cagcgaattc cacgatgctg atgcgcagaa ctctcacagc tattgctttg agattctgga	1920
gcggcgataa taagagctcg ctgatcagcc tcgactgtgc cttctagtgt ccagccatct	1980
gttgtttgcc cttccccgt gccttccttg accctggaag gtgccactcc cactgtcctt	2040
tcctaataaa atgaggaaat tgcacgcgat tgtctgagta ggtgtcattc tattctgggg	2100
ggtgggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg	2160
gatgcggtgg gctctatgga agcttgaatt cagctgacgt gcctcggacc gctaggaacc	2220
cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	2280

accaaaggtc gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg	2340
cagctgcctg cagggggcgcc tgatgcggta ttttctcctt acgcatctgt gcggtatttc	2400
acaccgcata cgtcaaagca accatagtac gcgccctgta gcggcgcatt aagcgcggcg	2460
ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgtcct	2520
ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat	2580
cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt	2640
gatttgggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg	2700
acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac	2760
cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta	2820
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca	2880
attttatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga	2940
cacccgcaa caccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac	3000
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcacg	3060
aaacgcgcga gacgaaagg cctcgtgata cgcctatttt tataggttaa tgtcatgaac	3120
aataaaaactg tctgcttaca taaacagtaa tacaaggggt gttatgagcc atattcaacg	3180
ggaaacgtcg aggccgcgat taaattccaa catggatgct gatttatatg ggtataaatg	3240
ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat cgcttgatg ggaagcccga	3300
tgcgccagag ttgtttctga aacatggcaa aggtagcggt gccaatgatg ttacagatga	3360
gatggtcaga ctaaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat	3420
ccgtactcct gatgatgcat gggttactcac cactgcgatc cccggaaaaa cagcattcca	3480
ggtattagaa gaatatcctg attcaggtga aaatattggt gatgcgctgg cagtgttcct	3540
gcgccgggtg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg	3600
tctcgctcag gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg attttgatga	3660
cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa atgcataaac ttttgccatt	3720
ctcaccggat tcagtcgtca ctcatggtga tttctcactt gataacctta tttttgacga	3780
ggggaaatta ataggttgta ttgatgttg acgagtcgga atcgagacc gataccagga	3840
tcttgccatc ctatggaact gcctcgggtga gttttctcct tcattacaga aacggctttt	3900

tcaaaaatat ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga	3960
tgagtttttc taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt	4020
cagaccccggt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct	4080
gctgcttgca aacaaaaaaaa ccaccgctac cagcgggtggg ttgtttgccg gatcaagagc	4140
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc	4200
ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc	4260
tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg	4320
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt	4380
cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg	4440
agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacagggtat ccggtaagcg	4500
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt	4560
atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag	4620
gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt	4680
gctggccttt tgctcacatg t	4701

<210> 83

<211> 4218

<212> DNA

<213> Artificial Sequence

<220>

<223> pITR-U6-shHES1-S5-CMV-3xFLAG-hATOH1-DD-T2A-hPOU4F3-U6-shHES1-S3

<400> 83

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttcct gcggccgcaa aaaaactgca tgaccagat caattcgttg atctgggtca	180
tgcagtcggt gtttcgtcct ttccacaaga tatataaagc caagaaatcg aaatactttc	240
aagttacggg aagcatatga tagtccattt taaaacataa ttttaaaact gcaaactacc	300
caagaaatta ttactttcta cgtcacgtat tttgtactaa tatctttgtg tttacagtca	360
aattaattct aattatctct ctaacagcct tgtatcgtat atgcaaatat gaaggaatca	420
tgggaaatag gccctcttcc tgcccgaacca cgcgtctaga tcccatatat ggagttccgc	480

gttacataac ttacggtaaa tggcccgccct ggctgaccgc ccaacgaccc ccgcccattg	540
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa	600
tgggtggagt atttacggta aactgcccac ttggcagtac atcaagtgt tcatatgcca	660
agtacgcccc ctattgacgt caatgacggt aaatggcccc cctggcatta tgcccagtac	720
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc	780
atggtgatgc ggttttggca gtacatcaat gggcgtggat agcggtttga ctcacgggga	840
tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcacca aaatcaacgg	900
gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggcgg taggcgtgta	960
cggtgggagg tctatataag cagagctacc ggtgccacca tggattacaa ggatgacgac	1020
gataaggact ataaggacga tgatgacaag gactacaaag atgatgacga taaagttaac	1080
tcccgctgc tgcatgcaga agagtgggct gaagtgaagg agttgggaga ccaccatcgc	1140
cagccccagc cgcacatctt cccgcaaccg ccgccgccgc cgcagccacc tgcaactttg	1200
caggcgagag agcatcccggt ctacccgcct gagctgtccc tcctggacag caccgacca	1260
cgcgcctggc tggctccac tttgcagggc atctgcacgg cacgcgccgc ccagtatttg	1320
ctacattccc cggagctggg tgcctcagag gccgctgcgc cccgggacga ggtggacggc	1380
cggggggagc tggtaaggag gagcagcggc ggtgccagca gcagcaagag ccccgggccg	1440
gtgaaagtgc gggaacagct gtgcaagctg aaaggcgggg tggtaggtaga cgagctgggc	1500
tgcagccgcc aacgggcccc ttccagcaaa cagggtgaatg gggtagcaga gcagagacgg	1560
ctagcagcca acgccaggga gcggcgcagg atgcatgggc tgaaccacgc cttcgaccag	1620
ctgcgcaatg ttatcccgtc gttcaacaac gacaagaagc tgtccaaata tgagaccctg	1680
cagatggccc aaatctacat caacgccttg tccgagctgc tacaacgcc cagcggaggg	1740
gaacagccac cgccgcctcc agcctcctgc aaaagcgacc accaccacct tcgcaccgcg	1800
gcctcctatg aagggggcgc gggcaacgcg accgcagctg gggctcagca ggcttccgga	1860
gggagccagc ggccgacccc gcccgggagt tgccggactc gcttctcagc cccagcttct	1920
gcgggaggggt actcgggtgca gctggacgct ctgcacttct cgactttcga ggacagcgcc	1980
ctgacagcga tgatggcgca aaagaatttg tctccttctc tccccgggag catcttgacg	2040
ccagtgcagg aggaaaacag caaaacttcg cctcgggtccc acagaagcga cggggaattt	2100

tccccccatt cccattacag tgactcggat gaggcaagtg ttaacatcag tctgattgcg	2160
gcgtttagcgg tagattacgt tatcggcatg gaaaacgcca tgccgtggaa cctgcctgcc	2220
gatctcgcct ggtttaaacg caacacctta aataaaccg tgattatggg ccgccatacc	2280
tgggaatcaa tcggtcgtcc gttgccagga cgcaaaaata ttatcctcag cagtcaaccg	2340
agtacggacg atcgcgtaac gtgggtgaag tcggtggatg aagccatcgc ggcgtgtggt	2400
gacgtaccag aaatcatggt gattggcggc ggtcgcgtta ttgaacagtt cttgccaaaa	2460
gcgcaaaaac tgtatctgac gcatatcgac gcagaagtgg aaggcgacac ccatttccc	2520
gattacgagc cggatgactg ggaatcggtta ttcagcgaat tccacgatgc tgatgcgcag	2580
aactctcaca gctattgctt tgagattctg gagcggcgag gatccggcga gggcagagga	2640
agtctgctaa catgcggtga cgtcaggag aatcctggcc caatgatggc catgaactcc	2700
aagcagcctt tcggcatgca cccggtgctg caagaacca aattctccag tctgcactct	2760
ggctccgagg ccatgcgccg agtctgtctc ccagccccgc agctgcaggg taatatattt	2820
ggaagctttg atgagagcct gctggcacgc gccgaagctc tggcggcggg ggatatcgtc	2880
tcccacggca agaaccatcc gttcaagccc gacgccacct accataccat gagcagcgtg	2940
ccctgcacgt ccacttcgtc caccgtgccc atctcccacc cagctgcgct cacctcacac	3000
cctcaccacg ccgtgcacca gggcctcgaa ggcgacctgc tggagcacat ctgcccacg	3060
ctgagtgtga gcggcctggg cgctccggaa cactcgggtga tgcccgcaca gatccatcca	3120
caccacctgg gcgccatggg ccacctgcac caggccatgg gcatgagtca cccgcacacc	3180
gtggcccctc atagcgccat gcctgcatgc ctcagcgacg tggagtcaga cccgcgcgag	3240
ctggaagcct tcgccgagcg cttcaagcag cggcgcacat agctgggggt gaccagggcg	3300
gacgtggggc cggtcttggc taatctcaag atccccggcg tgggctcgct gagccaaagc	3360
accatctgca ggttcgagtc tctcactctc tcgcacaaca acatgatcgc tctcaagccg	3420
gtgctccagg cctggttgga ggaggccgag gccgcctacc gagagaagaa cagcaagcca	3480
gagctcttca acggcagcga acggaagcgc aaacgcacgt ccatcgcggc gccggagaag	3540
cgttcactcg aggccatatt cgctatccag ccacgtcctt catctgagaa gatcgcggcc	3600
atcgctgaga aactggacct taaaaagaac gtggtgagag tctggttctg caaccagaga	3660
cagaaacaga aacgaatgaa gtattcggct gtccactaaa taataaaaata tctttatttt	3720

cattacatct gtgtgttggt tttttgtgtg ttaattaaaa aaaagaaagt catcaaagcc	3780
tatttcgata ggctttgatg actttcggtg tttcgtcctt tccacaagat atataaagcc	3840
aagaaatcga aatactttca agttacggta agcatatgat agtccatttt aaaacataat	3900
tttaaaactg caaactaccc aagaaattat tactttctac gtcacgtatt ttgtactaat	3960
atctttgtgt ttacagtcaa attaattcta attatctctc taacagcctt gtatcgtata	4020
tgcaaatatg aaggaatcat gggaaatagg ccctcttcct gcccgacccg gaccgctagg	4080
aaccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg	4140
ggcgaccaa ggtcgccga cgcccgggct ttgcccgggc ggccctcagt agcgagcgag	4200
cgcgagctg cctgcagg	4218

<210> 84
 <211> 263
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> U6

<400> 84	
cggtgtttcg tcctttccac aagatatata aagccaagaa atcgaaatac tttcaagtta	60
cggtaaagcat atgatagtc attttaaaac ataattttta aactgcaaac tacccaagaa	120
attattactt tctacgtcac gtattttgta ctaatatctt tgtgtttaca gtcaaattaa	180
ttctaattat ctctctaaca gccttgtatc gtatatgcaa atatgaagga atcatgggaa	240
ataggccctc ttctgccc acc	263

<210> 85
 <211> 263
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> shHES1-1

<400> 85	
cggtgtttcg tcctttccac aagatatata aagccaagaa atcgaaatac tttcaagtta	60
cggtaaagcat atgatagtc attttaaaac ataattttta aactgcaaac tacccaagaa	120
attattactt tctacgtcac gtattttgta ctaatatctt tgtgtttaca gtcaaattaa	180

ttctaattat ctctctaaca gccttgtatc gtatatgcaa atatgaagga atcatgggaa	240
ataggccctc ttcctgccc acc	263

<210> 86
 <211> 75
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 3x FLAG

<400> 86	
atggattaca aggatgacga cgataaggac tataaggacg atgatgacaa ggactacaaa	60
gatgatgacg ataaa	75

<210> 87
 <211> 1065
 <212> DNA
 <213> Homo sapiens

<400> 87	
gttaactccc gcctgctgca tgcagaagag tgggctgaag tgaaggagtt gggagaccac	60
catcgccagc cccagccgca tcattctccg caaccgccgc cgccgccgca gccacctgca	120
actttgcagg cgagagagca tcccgtctac ccgcctgagc tgtccctcct ggacagcacc	180
gaccacgcg cctggctggc tcccactttg cagggcatct gcacggcacg cgccgccag	240
tatttgctac attccccgga gctgggtgcc tcagaggccg ctgcgccccg ggacgaggtg	300
gacggccggg gggagctggt aaggaggagc agcggcgggt ccagcagcag caagagcccc	360
gggccggtga aagtgcggga acagctgtgc aagctgaaag gcgggggtgt ggtagacgag	420
ctgggctgca gccgccaacg ggcccccttc agcaaacagg tgaatggggt gcagaagcag	480
agacggctag cagccaacgc caggagcgcg cgaggatgc atgggctgaa ccacgccttc	540
gaccagctgc gcaatgttat cccgtcgttc aacaacgaca agaagctgtc caaatatgag	600
accctgcaga tggcccaa atctacatcaac gccttgtccg agctgtaca aacgcccagc	660
ggaggggaac agccaccgcc gcctccagcc tcctgcaaaa gcgaccacca ccaccttcgc	720
accgcggcct cctatgaagg gggcgcgggc aacgcgaccg cagctggggc tcagcaggct	780
tccggaggga gccagcggcc gaccccgccc gggagttgcc ggactcgctt ctacgcccc	840

gcttctgcgg gaggtactc ggtgcagctg gacgctctgc acttctcgac tttcgaggac	900
agcgccctga cagcgatgat ggcgcaaaag aatttgtctc cttctctccc cgggagcatc	960
ttgcagccag tgcaggagga aaacagcaaa acttcgcctc ggtcccacag aagcgacggg	1020
gaattttccc cccattccca ttacagtgac tcggatgagg caagt	1065

<210> 88
 <211> 474
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Destabilizing Domain

<400> 88	
atcagtctga ttgcggcggtt agcggtagat tacgttatcg gcatggaaaa cgccatgccg	60
tggaacctgc ctgccgatct cgcctggttt aaacgcaaca ccttaaataa acccgtgatt	120
atgggccgcc atacctggga atcaatcggg cgtccgttgc caggacgcaa aaatattatc	180
ctcagcagtc aaccgagtac ggacgatcgc gtaacgtggg tgaagtcggg ggatgaagcc	240
atcgcggcgt gtggtgacgt accagaaatc atggtgattg gcggcggtcg cgttattgaa	300
cagttcttgc caaaagcgca aaaactgtat ctgacgcata tcgacgcaga agtggaaggc	360
gacacccatt tcccggatta cgagccggat gactgggaat cggatttcag cgaattccac	420
gatgctgatg cgcagaactc tcacagctat tgctttgaga ttctggagcg gcga	474

<210> 89
 <211> 56
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> T2A

<400> 89	
gcgagggcag aggaagtctg ctaacatgcg gtgacgtcga ggagaatcct ggccca	56

<210> 90
 <211> 51
 <212> DNA
 <213> Bovidae

<400> 90	
ataataaaat atctttatct tcattacatc tgtgtgttgg ttttttgtgt g	51

<210> 91
<211> 262
<212> DNA
<213> Artificial Sequence

<220>
<223> U6

<400> 91
ggtgtttcgt cttttccaca agatatataa agccaagaaa tcgaaatact ttcaagttac 60
ggtaagcata tgatagtcca ttttaaaaca taatttttaa actgcaaact acccaagaaa 120
ttattacttt ctacgtcacg tttttgtac taatatcttt gtgtttacag tcaaattaat 180
tctaattatc tctctaacag ctttgtatcg tatatgcaaa tatgaaggaa tcatgggaaa 240
taggccctct tcctgcccga cc 262

<210> 92
<211> 262
<212> DNA
<213> Artificial Sequence

<220>
<223> shHES1-2

<400> 92
ggtgtttcgt cttttccaca agatatataa agccaagaaa tcgaaatact ttcaagttac 60
ggtaagcata tgatagtcca ttttaaaaca taatttttaa actgcaaact acccaagaaa 120
ttattacttt ctacgtcacg tttttgtac taatatcttt gtgtttacag tcaaattaat 180
tctaattatc tctctaacag ctttgtatcg tatatgcaaa tatgaaggaa tcatgggaaa 240
taggccctct tcctgcccga cc 262

<210> 93
<211> 4930
<212> DNA
<213> Artificial Sequence

<220>
<223> pITR-U6-shHES1-S5,
hATOHessps-3xFLAG-hATOH1-T2A-hPOU4F3-US-shHES1-S3

<400> 93
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt 60

ggtcgccccg cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
agggggttcct gcggccgcaa aaaaactgca tgaccagat caattcgttg atctgggtca	180
tgagtcggt gtttcgtcct ttccacaaga tatataaagc caagaaatcg aaatactttc	240
aagttacggt aagcatatga tagtccattt taaaacataa ttttaaaact gcaaactacc	300
caagaaatta ttactttcta cgtcacgtat ttgtactaa tatctttgtg ttacagtca	360
aattaattct aattatctct ctaacagcct tgtatcgtat atgcaaatat gaaggaatca	420
tgggaaatag gccctcttcc tgccgacca cgcgtctatg gaggttgcat aacaaacgtt	480
tggcagctcg ctctcttaca ctccattaac aagctgtaac atatagctgc aggttgctat	540
aatctcatta atattttgga aacttgaata ttgagtattt ctgagtgtc attccccata	600
tgccagccac ttctgccatg ctgactgggt ctttctctc cattattagc aattagcttc	660
ttaccttcca aagtcagatc caaggtatcc aagatactag caaaggaatc aactatgtgt	720
gcaagttaag catgcttaat atcacccaaa caaacaaga ggcagcattt cttaaagtaa	780
tgaagataga taaatcgggt tagtcctttg cgacactgct ggtgctttct agagttttat	840
atattttaag cagcttgctt tatattctgt ctttgccctc caccaccca gcacttttat	900
ttgtggaggg ttttggtcg ccacactttg ggaaacttat ttgatttcac ggagagctga	960
aggaagatca tttttggcaa cagacaagtt taaacacgat ttctatggga cattgctaac	1020
tggggcccct aaggagaaag gggaaactga gcggagaatg ggttaaattc ttggaagcag	1080
gggagaggca ggggaggaga gaagtcggag gaggataaag aaaaggacag gaaccaagaa	1140
gcgtgggggt ggtttgccgt aatgtgagtg tttcttaatt agagaacggt tgacaataga	1200
gggtctggca gaggctcctg gccgcggtgc ggagcgtctg gagcggagca cgcgctgtca	1260
gctggtgagc gcactctcct ttcaggcagc tccccgggga gctgtgcggc cacatttaac	1320
accatcatca cccctccccg gcctcctcaa cctcggcctc ctctcgtcg acagccttcc	1380
ttggcccca ccagcagagc tcacagtagc gagcgtctct cgccgtctcc cgactcggc	1440
cggggcctct ctctcccc agctgcgcag cgggagccgc cactgcccac tgcacctccc	1500
agcaaccagc ccagcacgca aagaagctgc gcaaagttaa agccaagcaa tgccaagggg	1560
aggggaagct ggaggcgggc tttgagtggc ttctgggcgc ctggcgggtc cagaatcgcc	1620
cagagccgcc cgcggtcgtg cacatctgac ccgagtcagc ttgggcacca gccgagagcc	1680

ggctccgcac cgctcccga cccagccgc cggggtggtg acacacaccg gagtccaatt	1740
acagccctgc aattaacata tgaatctgac gaatttaaaa gaaggaaaaa aaaaaaaaaa	1800
cctgagcagg cttgggagtc ctctgcacac aagaactttt ctcggggtgt aaaaactctt	1860
tgattggctg ctgcacgcg cctgcccgcg ccctccattg gctgagaaga cacgcgaccg	1920
gcgcgaggag ggggttggga gaggagcggg gggagactga gtggcgcgtg ccgcttttta	1980
aaggggcgca gcgccttcag caaccggaga agcatagttg cacgcgacct ggtgtgtgat	2040
ctccgagtgg gtgggggagg gtcgaggagg gaaaaaaaaa taagacgttg cagaagagac	2100
ccggaaaggg cttttttttt ggttgagctg gtgtcccagt gctgcctccg atcctgagcc	2160
tccgagcctt tgcagtgcaa ccggtgccac catggattac aaggatgacg acgataagga	2220
ctataaggac gatgatgaca aggactacaa agatgatgac gataaagtta acatgtcccg	2280
cctgctgcat gcagaagagt gggctgaagt gaaggagttg ggagaccacc atcgccagcc	2340
ccagccgcat catctcccgc aaccgccgcg gccgccgcag ccacctgcaa ctttgaggc	2400
gagagagcat cccgtctacc cgctgagct gtccctcctg gacagcaccg acccacgcgc	2460
ctggctggct cccactttgc agggcatctg cacggcacgc gccgcccagt atttgctaca	2520
ttccccggag ctgggtgcct cagaggccgc tgcgccccgg gacgaggtgg acggccgggg	2580
ggagctggta aggaggagca gcggcgggtgc cagcagcagc aagagccccg ggccggtgaa	2640
agtgcgggaa cagctgtgca agctgaaagg cggggtggtg gtagacgagc tgggctgcag	2700
ccgccaacgg gcccttcca gcaaacaggt gaatggggtg cagaagcaga gacggctagc	2760
agccaacgcc agggagcggc gcaggatgca tgggtgaac cacgccttcg accagctgcg	2820
caatgttatc ccgtcgttca acaacgacaa gaagctgtcc aaatatgaga ccctgcagat	2880
ggcccaaadc tacatcaacg ctttgtccga gctgctacaa acgcccagcg gaggggaaca	2940
gccaccgccg cctccagcct cctgcaaaag cgaccaccac caccttcgca ccgcggcctc	3000
ctatgaaggg ggcgcgggca acgcgaccgc agctggggct cagcaggctt ccggaggag	3060
ccagcggccg accccgcccg ggagttgccg gactcgcttc tcagccccag cttctgcggg	3120
agggtactcg gtgcagctgg acgctctgca cttctcgact ttcgaggaca gcgccctgac	3180
agcgatgatg gcgcaaaaga atttgtctcc ttctctcccc gggagcatct tgcagccagt	3240
gcaggaggaa aacagcaaaa cttgcctcg gtcccacaga agcgacgggg aattttcccc	3300

ccattcccat tacagtgact cggatgaggc aagtgttaac gagggcagag gaagtctgct	3360
aacatgcggt gacgtcgagg agaatcctgg cccaatgatg gccatgaact ccaagcagcc	3420
tttcggcatg cacccggtgc tgcaagaacc caaattctcc agtctgcact ctggctccga	3480
ggccatgcg cagagtctgtc tcccagcccc gcagctgcag ggtaatatat ttggaagctt	3540
tgatgagagc ctgctggcac gcgccgaagc tctggcggcg gtggatatcg tctcccacgg	3600
caagaacat ccgttcaagc ccgacgccac ctaccatacc atgagcagcg tgccctgcac	3660
gtccacttcg tccaccgtgc ccatctcca cccagctgcg ctcacctcac accctcacca	3720
cgccgtgcac cagggcctcg aaggcgacct gctggagcac atctcgcca cgctgagtgt	3780
gagcggcctg ggcgctccgg aacactcggt gatgcccga cagatccatc cacaccacct	3840
gggcgccatg ggccacctgc accaggccat gggcatgagt caccgcaca ccgtggcccc	3900
tcatagcgcc atgcctgcat gcctcagcga cgtggagtca gaccgcgcg agctggaagc	3960
cttcgccgag cgcttcaagc agcggcgcat caagctgggg gtgaccagg cggaagtggg	4020
cgcggctctg gctaattctca agatccccgg cgtgggctcg ctgagccaaa gcaccatctg	4080
caggttcgag tctctcactc tctcgacaaa caacatgatc gctctcaagc cgggtgctcca	4140
ggcctggttg gaggaggccg aggccgccta ccgagagaag aacagcaagc cagagctctt	4200
caacggcagc gaacggaagc gcaaacgcac gtccatcgcg gcgccggaga agcgttcact	4260
cgaggcctat ttcgctatcc agccacgtcc ttcacttgag aagatcgcgg ccatcgctga	4320
gaaactggac cttaaaaaga acgtggtgag agtctggttc tgcaaccaga gacagaaaca	4380
gaaacgaatg aagtattcgg ctgtccacta aataataaaa tatctttatt ttcattacat	4440
ctgtgtgttg gttttttgtg tgttaattaa aaaaaagaaa gtcataaag cctatttcga	4500
taggctttga tgactttcgg tgtttcgtcc tttccacaag atatataaag ccaagaaatc	4560
gaaatacttt caagttacgg taagcatatg atagtccatt ttaaaacata attttaaac	4620
tgcaaaactac ccaagaaatt attactttct acgtcacgta ttttgtacta atatctttgt	4680
gtttacagtc aaattaattc taattatctc tctaacagcc ttgtatcgta tatgcaaata	4740
tgaaggaatc atgggaaata ggccctcttc ctgcccgaac cggaccgcta ggaacccta	4800
gtgatggagt tggccactcc ctctctgcgc gctcgctcgc tctactgaggc cgggcgacca	4860
aaggtcgccc gacggcggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcagc	4920

tgcctgcagg

4930

<210> 94
<211> 1725
<212> DNA
<213> Homo sapiens

<400> 94
ctatggagtt tgcataacaa acgtttggca gctcgctctc ttacactcca ttaacaagct 60
gtaacatata gctgcagggtt gctataatct cattaatatt ttggaaactt gaatattgag 120
tatttctgag tgctcattcc ccatatgccca gccacttctg ccatgctgac tggttccttt 180
ctctccatta ttagcaatta gcttcttacc ttccaaagtc agatccaagg tatccaagat 240
actagcaaag gaatcaacta tgtgtgcaag ttaagcatgc ttaatatcac ccaaacaac 300
aaagaggcag catttcttaa agtaatgaag atagataaat cgggttagtc ctttgcgaca 360
ctgctgggtgc tttctagagt tttatatatt ttaagcagct tgctttatat tctgtctttg 420
cctcccaccc caccagcact tttatttgtg gagggttttg gctcgccaca ctttgggaaa 480
cttatttgat ttcacggaga gctgaaggaa gatcattttt ggcaacagac aagtttaaac 540
acgatttcta tgggacattg ctaactgggg cccctaagga gaaaggggaa actgagcgga 600
gaatgggtta aatccttgga agcaggggag aggcagggga ggagagaagt cggaggagta 660
taaagaaaag gacaggaacc aagaagcgtg ggggtggttt gccgtaatgt gagtgtttct 720
taattagaga acggttgaca atagagggtc tggcagaggc tcctggccgc ggtgcggagc 780
gtctggagcg gagcacgcgc tgtcagctgg tgagcgcact ctcctttcag gcagctcccc 840
ggggagctgt gcggccacat ttaacaccat catcaccct ccccgccctc ctcaacctcg 900
gcctcctcct cgtcgacagc cttccttggc cccaccagc agagctcaca gtagcgagcg 960
tctctcgccg tctccgcac tcggccgggg cctctctcct cccccagctg cgcagcggga 1020
gccgccactg cccactgcac ctcccagcaa ccagcccagc acgcaaagaa gctgcgcaaa 1080
gttaaagcca agcaatgcca aggggagggg aagctggagg cgggctttga gtggcttctg 1140
ggcgcctggc ggggtccagaa tcgcccagag ccgcccgcgg tcgtgcacat ctgacctgag 1200
tcagcttggg caccagccga gagccggctc cgcaccgctc ccgcacccca gccgccgggg 1260
tggtgacaca caccggagtc gaattacagc cctgcaatta acatatgaat ctgacgaatt 1320
taaaagaagg aaaaaaaaaa aaaaacctga gcaggcttgg gagtcctctg cacacaagaa 1380

cttttctcgg ggtgtaaaaa ctctttgatt ggctgctcgc acgcgcctgc ccgcgccctc	1440
cattggctga gaagacacgc gaccggcgcg aggaggggggt tgggagagga gcgggggggag	1500
actgagtggc gcgtgccgct ttttaaaggg gcgcagcgcc ttcagcaacc ggagaagcat	1560
agttgcacgc gacctggtgt gtgatctccg agtgggtggg ggaggggtcga ggagggaaaa	1620
aaaaataaga cgttgcagaa gagacccgga aagggccttt tttttggttg agctggtgtc	1680
ccagtgtcgc ctccgatcct gagcctccga gcctttgcag tgcaa	1725

<210> 95
 <211> 1017
 <212> DNA
 <213> Homo sapiens

<400> 95	
atgatggcca tgaactccaa gcagcctttc ggcatgcacc cgggtgtgca agaaccctaaa	60
ttctccagtc tgcactctgg ctccgaggcc atgcgccgag tctgtctccc agccccgcag	120
ctgcagggtg atatatttgg aagctttgat gagagcctgc tggcacgcgc cgaagctctg	180
gcggcggtgg atatcgtctc ccacggcaag aaccatccgt tcaagcccga cgccacctac	240
cataccatga gcagcgtgcc ctgcacgtcc acttcgtcca ccgtgcccac ctccccacca	300
gctgcgctca cctcacaccc tcaccacgcc gtgcaccagg gcctcgaagg cgacctgctg	360
gagcacatct cgcccacgct gagtgtgagc ggcctgggag ctccggaaca ctcggtgatg	420
cccgcacaga tccatccaca ccacctgggc gccatgggac acctgcacca ggccatgggc	480
atgagtcacc cgcacaccgt ggccccctcat agcgccatgc ctgcatgcct cagcgacgtg	540
gagtcagacc cgcgcgagct ggaagccttc gccgagcgct tcaagcagcg gcgcatcaag	600
ctgggggtga cccaggcgga cgtgggcgcg gctctggcta atctcaagat cccggcggtg	660
ggctcgctga gccaaagcac catctgcagg ttcgagtctc tcaactctct gcacaacaac	720
atgatcgctc tcaagccggt gctccaggcc tggttggagg aggccgaggc cgcctaccga	780
gagaagaaca gcaagccaga gctcttcaac ggcagcgaac ggaagcgcaa acgcacgtcc	840
atcgcggcgc cggagaagcg ttcaactgag gcctatctcg ctatccagcc acgtccttca	900
tctgagaaga tcgcggccat cgctgagaaa ctggacctta aaaagaacgt ggtgagagtc	960
tggttctgca accagagaca gaaacagaaa cgaatgaagt attcggctgt ccactaa	1017

<210> 96
<211> 508
<212> DNA
<213> Artificial Sequence

<220>
<223> CMV enhancer-promoter

<400> 96
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tattttacgt aaactgcca cttggcagta catcaagtgt atcatatgcc 180
aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagct 508