US 20060160764A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2006/0160764 A1**

Friedman et al. (43) **Pub. Date: Jul. 20, 2006**

(54) **LIMKS AS MODIFIERS OF P53 PATHWAY AND METHODS OF USE**

(75) Inventors: **Lori Friedman**, San Francisco, CA (US); **Gregory D. Plowman**, San Carlos, CA (US); **Marcia Belvin**, Albany, CA (US); **Helen Francis-Lang**, San Francisco, CA (US); **Danxi Li**, San Francisco, CA (US); **Roel P. Funke**, South San Francisco, CA (US)

Correspondence Address:
MCDONNELL BOEHNEN HULBERT & BERGHOFF LLP
300 S. WACKER DRIVE
32ND FLOOR
CHICAGO, IL 60606 (US)

(73) Assignee: **Exelixis, Inc.**, South San Francisco, CA

(21) Appl. No.: **11/378,923**

(22) Filed: **Mar. 17, 2006**

Related U.S. Application Data

(62) Division of application No. 10/161,453, filed on Jun. 3, 2002.

(60) Provisional application No. 60/296,076, filed on Jun. 5, 2001. Provisional application No. 60/328,605, filed on Oct. 10, 2001. Provisional application No. 60/357,253, filed on Feb. 15, 2002.

Publication Classification

(51) **Int. Cl.**

A61K	***48/00***	(2006.01)
C12Q	***1/68***	(2006.01)
G01N	***33/574***	(2006.01)
C12Q	***1/48***	(2006.01)
A61K	***39/395***	(2006.01)

(52) **U.S. Cl.** **514/44**; 424/155.1; 435/6; 435/7.23; 435/15

(57) **ABSTRACT**

Human LIMK genes are identified as modulators of the p53 pathway, and thus are therapeutic targets for disorders associated with defective p53 function. Methods for identifying modulators of p53, comprising screening for agents that modulate the activity of LIMK are provided.

LIMKS AS MODIFIERS OF P53 PATHWAY AND METHODS OF USE

REFERENCE TO RELATED APPLICATIONS

[0001] This application is a divisional application of U.S. patent application Ser. No. 10/161,453, filed Jun. 3, 2002, which claims priority to U.S. provisional patent application No. 60/296,076 filed Jun. 5, 2001, No. 60/328,605 filed Oct. 10, 2001, and No. 60/357,253 filed Feb. 15, 2002. The contents of the prior applications are hereby incorporated by reference in their entirety.

BACKGROUND OF THE INVENTION

[0002] The p53 gene is mutated in over 50 different types of human cancers, including familial and spontaneous cancers, and is believed to be the most commonly mutated gene in human cancer (Zambetti and Levine, FASEB (1993) 7:855-865; Hollstein, et al., Nucleic Acids Res. (1994) 22:3551-3555). Greater than 90% of mutations in the p53 gene are missense mutations that alter a single amino acid that inactivates p53 function. Aberrant forms of human p53 are associated with poor prognosis, more aggressive tumors, metastasis, and short survival rates (Mitsudomi et al., Clin Cancer Res 2000 October; 6(10):4055-63; Koshland, Science (1993) 262:1953).

[0003] The human p53 protein normally functions as a central integrator of signals including DNA damage, hypoxia, nucleotide deprivation, and oncogene activation (Prives, Cell (1998) 95:5-8). In response to these signals, p53 protein levels are greatly increased with the result that the accumulated p53 activates cell cycle arrest or apoptosis depending on the nature and strength of these signals. Indeed, multiple lines of experimental evidence have pointed to a key role for p53 as a tumor suppressor (Levine, Cell (1997) 88:323-331). For example, homozygous p53 "knockout" mice are developmentally normal but exhibit nearly 100% incidence of neoplasia in the first year of life (Donehower et al., Nature (1992) 356:215-221).

[0004] The biochemical mechanisms and pathways through which p53 functions in normal and cancerous cells are not fully understood, but one clearly important aspect of p53 function is its activity as a gene-specific transcriptional activator. Among the genes with known p53-response elements are several with well-characterized roles in either regulation of the cell cycle or apoptosis, including GADD45, p21/Waf1/Cip1, cyclin G, Bax, IGF-BP3, and MDM2 (Levine, Cell (1997) 88:323-331).

[0005] Eukaryotic LIM proteins contain LIM domains, highly conserved cysteine-rich structures containing two zinc fingers (Mizuno, K. et al. (1994) Oncogene 9: 1605-1612, 1994). Zinc fingers usually function by binding to DNA or RNA. However, it is believed that the LIM motif mediates protein-protein interactions. LIM kinase (LIMK) belongs to a small subfamily of kinases with a unique combination of two N-terminal LIM motifs and a C-terminal protein kinase domain. These LIM motifs and kinase domains are linked by a proline-serine-rich region containing several putative casein kinase and map kinase recognition sites (Mizuno, K. et al. (1994) supra). LIM kinases and their pathway proteins are believed to contribute to Rho-induced reorganization of the actin cytoskeleton (Maekawa et al. (1999) Science 285: 895-898).

[0006] LIM kinase 1 (LIMK1) contains a serine/threonine kinase with two LIM domains. It mediates signaling from Rho to the actin cytoskeleton via phosphorylation of cofilin, is involved in T cell chemotaxis, is activated during mitosis, and may regulate cell cycle progression. LIMK1 lacks a signal peptide and putative transmembrane domains, and therefore it may be a component of an intracellular signaling pathway and is likely involved in brain development (Tassabehji, M., et al. (1996) Nature Genet. 13: 272-273). LIMK1 demonstrates 95% homology with the murine LIMK1 gene and therefore the function of this gene should also be highly conserved (Proschel, C., et al. (1995) Oncogene 11: 1271-1281). Limk1 is expressed in the central nervous system during embryogenesis (Proschel, C., et al. (1995) supra). LIMK1 has also been identified as a potential activator of serum response factor, which regulates transcription of many serum-inducible and muscle-specific genes (Sotiropoulos, A., et al. (1999) Cell 98: 159-169). LIMK1 has been identified as a candidate for the unexplained neurologic features of William's syndrome (Tassabehji, M. et al. (1996) supra).

[0007] LIM kinase 2 (LIMK2), a second member of the LIMK family, has a domain structure similar to LIMK 1 with a serine/threonine protein kinase with two LIM motifs (Okano, I. et al. (1995) Science 285: 895-898). It also phosphorylates cofilin and regulates Rho family-dependent actin cytoskeletal rearrangement. LIMK2 has two alternative transcripts, LIMK2a and LIMK2b. Both forms of LIMK2 are expressed in a wide variety of tissues, with placenta showing the highest expression levels (Osada, H. et al. (1996) Biochem. Biophys. Res. Commun. 229:582-589). LMK2a contains two LIM domains, a PDZ domain, and a kinase domain. LIMK2b has only 1.5 LIM domains, which is thought to be involved in protein-protein interactions; the PDZ domain was originally recognized in a family of proteins related to the *Drosophila* lethal discs-large tumor suppressor gene product Dlga which functions in in protein-protein interactions targeting the protein to the submembranous compartment. Several studies have shown that LIMK2a (63 kD protein) resides in the cytoplasm and nucleus while the LIMK2b (58 kD protein) is found mainly in the cytoplasm (Okano, I. et al. (1995) supra).

[0008] Testis-specific protein kinase (TESK) is a probable member of the LIMK subfamily of serine/threonine protein kinases, and has strong similarity to rat Rn.7006, which is expressed specifically in testicular germ cells and may have a role in spermatogenesis. The protein kinase of the human TESK1 is structurally similar to both LIMK1 and LIMK2 (Toshima, J. et al. (1995) J. Biol. Chem. 270: 31331-31337). Testis-specific kinase 2 protein kinase (TESK 2) is a member of the LIMK/TESK subfamily of serine/threonine protein kinases, and may play a role in spermatogenesis. TESK2 exhibits dual specific protein kinase activity on both serine/threonine and tyrosine. It is predominantly expressed in the testis and prostate (Rosok, O. et al. (1999) Genomics 61: 44-54).

[0009] The ability to manipulate the genomes of model organisms such as *Drosophila* provides a powerful means to analyze biochemical processes that, due to significant evolutionary conservation, has direct relevance to more complex vertebrate organisms. Due to a high level of gene and pathway conservation, the strong similarity of cellular processes, and the functional conservation of genes between

these model organisms and mammals, identification of the involvement of novel genes in particular pathways and their functions in such model organisms can directly contribute to the understanding of the correlative pathways and methods of modulating them in mammals (see, for example, Mechler B M et al., 1985 EMBO J. 4:1551-1557; Gateff E. 1982 Adv. Cancer Res. 37: 33-74; Watson K L., et al., 1994 J Cell Sci. 18: 19-33; Miklos G L, and Rubin G M. 1996 Cell 86:521-529; Wassarman D A, et al., 1995 Curr Opin Gen Dev 5: 44-50; and Booth D R. 1999 Cancer Metastasis Rev. 18: 261-284). For example, a genetic screen can be carried out in an invertebrate model organism having underexpression (e.g. knockout) or overexpression of a gene (referred to as a "genetic entry point") that yields a visible phenotype. Additional genes are mutated in a random or targeted manner. When a gene mutation changes the original phenotype caused by the mutation in the genetic entry point, the gene is identified as a "modifier" involved in the same or overlapping pathway as the genetic entry point. When the genetic entry point is an ortholog of a human gene implicated in a disease pathway, such as p53, modifier genes can be identified that may be attractive candidate targets for novel therapeutics.

[0010] All references cited herein, including sequence information in referenced Genbank identifier numbers and website references, are incorporated herein in their entireties.

SUMMARY OF THE INVENTION

[0011] We have discovered genes that modify the p53 pathway in *Drosophila*, and identified their human orthologs, hereinafter referred to as LIMK. The invention provides methods for utilizing these p53 modifier genes and polypeptides to identify candidate therapeutic agents that can be used in the treatment of disorders associated with defective p53 function. Preferred LIMK-modulating agents specifically bind to LIMK polypeptides and restore p53 function. Other preferred LIMK-modulating agents are nucleic acid modulators such as antisense oligomers and RNAi that repress LIMK gene expression or product activity by, for example, binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA).

[0012] LIMK-specific modulating agents may be evaluated by any convenient in vitro or in vivo assay for molecular interaction with a LIMK polypeptide or nucleic acid. In one embodiment, candidate p53 modulating agents are tested with an assay system comprising a LIMK polypeptide or nucleic acid. Candidate agents that produce a change in the activity of the assay system relative to controls are identified as candidate p53 modulating agents. The assay system may be cell-based or cell-free. LIMK-modulating agents include LIMK related proteins (e.g. dominant negative mutants, and biotherapeutics); LIMK-specific antibodies; LIMK-specific antisense oligomers and other nucleic acid modulators; and chemical agents that specifically bind LIMK or compete with LIMK binding target. In one specific embodiment, a small molecule modulator is identified using a kinase assay. In specific embodiments, the screening assay system is selected from a binding assay, an apoptosis assay, a cell proliferation assay, an angiogenesis assay, and a hypoxic induction assay.

[0013] In another embodiment, candidate p53 pathway modulating agents are further tested using a second assay system that detects changes in the p53 pathway, such as angiogenic, apoptotic, or cell proliferation changes produced by the originally identified candidate agent or an agent derived from the original agent. The second assay system may use cultured cells or non-human animals. In specific embodiments, the secondary assay system uses non-human animals, including animals predetermined to have a disease or disorder implicating the p53 pathway, such as an angiogenic, apoptotic, or cell proliferation disorder (e.g. cancer).

[0014] The invention further provides methods for modulating the p53 pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a LIMK polypeptide or nucleic acid. The agent may be a small molecule modulator, a nucleic acid modulator, or an antibody and may be administered to a mammalian animal predetermined to have a pathology associated the p53 pathway.

DETAILED DESCRIPTION OF THE INVENTION

[0015] Genetic screens were designed to identify modifiers of the p53 pathway in *Drosophila* in which p53 was overexpressed in the wing (Ollmann M, et al., Cell 2000 101: 91-101). The CG6027/cdi/tsk gene was identified as a modifier of the p53 pathway. Accordingly, vertebrate orthologs of these modifiers, and preferably the human orthologs, LIMK genes (i.e., nucleic acids and polypeptides) are attractive drug targets for the treatment of pathologies associated with a defective p53 signaling pathway, such as cancer.

[0016] In vitro and in vivo methods of assessing LIMK function are provided herein. Modulation of the LIMK or their respective binding partners is useful for understanding the association of the p53 pathway and its members in normal and disease conditions and for developing diagnostics and therapeutic modalities for p53 related pathologies. LIMK-modulating agents that act by inhibiting or enhancing LIMK expression, directly or indirectly, for example, by affecting a LIMK function such as enzymatic (e.g., catalytic) or binding activity, can be identified using methods provided herein. LIMK modulating agents are useful in diagnosis, therapy and pharmaceutical development.

Nucleic Acids and Polypeptides of the Invention

[0017] Sequences related to LIMK nucleic acids and polypeptides that can be used in the invention are disclosed in Genbank (referenced by Genbank identifier (GI) number) as GI#s 11421438 (SEQ ID NO:1), 8051614 (SEQ ID NO:2), 8051616 (SEQ ID NO:3), 565279 (SEQ ID NO:4), 10439425 (SEQ ID NO:5), 2078469 (SEQ ID NO:6), 8051617 (SEQ ID NO:7), 8051619 (SEQ ID NO:8), 15341773 (SEQ ID NO:9), 13928783 (SEQ ID NO:10), 5454109 (SEQ ID NO:11), 14739649 (SEQ ID NO:12), 15080032 (SEQ ID NO:13), 1805594 (SEQ ID NO:14), 6005895 (SEQ ID NO:15), 18549250 (SEQ ID NO:16), 15207800 (SEQ ID NO:17), and 14042344 (SEQ ID NO:18) for nucleic acid, and GI#s 11421439 (SEQ ID NO:19), 4505001 (SEQ ID NO:20), 2078472 (SEQ ID NO:21), 5031869 (SEQ ID NO:22), 2499660 (SEQ ID NO:23), 5454110 (SEQ ID NO:24) and 6005896 (SEQ ID NO:25) for polypeptides.

[0018] LIMKs are kinase proteins with kinase domains. The term "LIMK polypeptide" refers to a full-length LIMK protein or a functionally active fragment or derivative thereof. A "functionally active" LIMK fragment or derivative exhibits one or more functional activities associated with a full-length, wild-type LIMK protein, such as antigenic or immunogenic activity, enzymatic activity, ability to bind natural cellular substrates, etc. The functional activity of LIMK proteins, derivatives and fragments can be assayed by various methods known to one skilled in the art (Current Protocols in Protein Science (1998) Coligan et al., eds., John Wiley & Sons, Inc., Somerset, New Jersey) and as further discussed below. For purposes herein, functionally active fragments also include those fragments that comprise one or more structural domains of a LIMK, such as a kinase domain or a binding domain. Protein domains can be identified using the PFAM program (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2; http://pfam.wustl.edu). For example, the LIM domains of LIMK from GI# 4505001 (SEQ ID NO:20) are located at approximately amino acid residues 80 to 136 and 139 to 198 (PFAM 00412). The LIM domains of LIMK from GI# 5031869 (SEQ ID NO:22) are located at approximately amino acid residues 12 to 69, 72 to 129 (PFAM 00412). Further, the kinase domains of LIMK from GI#s 4505001, 5031869, 5454110 and 6005896 (SEQ ID NOs: 20, 22, 24, and 25, respectively) are located at approximately amino acid residues 394 to 652 for GI# 4505001, 310-518 for GI#5031869, 57 to 314 for GI#5454110, and 62 to 117 and 120 to 289 for GI#6005896, respectively (PFAM00069). Methods for obtaining LIMK polypeptides are also further described below. In some embodiments, preferred fragments are functionally active, domain-containing fragments comprising at least 25 contiguous amino acids, preferably at least 50, more preferably 75, and most preferably at least 100 contiguous amino acids of any one of SEQ ID NOs:19, 20, 21, 22, 23, 24, or 25 (a LIMK). In further preferred embodiments, the fragment comprises the entire kinase (functionally active) domain.

[0019] The term "LIMK nucleic acid" refers to a DNA or RNA molecule that encodes a LIMK polypeptide. Preferably, the LIMK polypeptide or nucleic acid or fragment thereof is from a human, but can also be an ortholog, or derivative thereof with at least 70% sequence identity, preferably at least 80%, more preferably 85%, still more preferably 90%, and most preferably at least 95% sequence identity with LIMK. Normally, orthologs in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. Orthologs are generally identified by sequence homology analysis, such as BLAST analysis, usually using protein bait sequences. Sequences are assigned as a potential ortholog if the best hit sequence from the forward BLAST result retrieves the original query sequence in the reverse BLAST (Huynen M A and Bork P, Proc Natl Acad Sci (1998) 95:5849-5856; Huynen M A et al., Genome Research (2000) 10:1204-1210). Programs for multiple sequence alignment, such as CLUSTAL (Thompson J D et al, 1994, Nucleic Acids Res 22:4673-4680) may be used to highlight conserved regions and/or residues of orthologous proteins and to generate phylogenetic trees. In a phylogenetic tree representing multiple homologous sequences from diverse species (e.g., retrieved through BLAST analysis), orthologous sequences from two species generally appear closest on the tree with respect to all other sequences from these two species. Structural threading or other analysis of protein folding (e.g., using software by ProCeryon, Biosciences, Salzburg, Austria) may also identify potential orthologs. In evolution, when a gene duplication event follows speciation, a single gene in one species, such as *Drosophila*, may correspond to multiple genes (paralogs) in another, such as human. As used herein, the term "orthologs" encompasses paralogs. As used herein, "percent (%) sequence identity" with respect to a subject sequence, or a specified portion of a subject sequence, is defined as the percentage of nucleotides or amino acids in the candidate derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul et al., J. Mol. Biol. (1997) 215:403-410; http://blast.wustl.edu/blast/README-.html) with all the search parameters set to default values. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A % identity value is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation.

[0020] A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the protein is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

[0021] Alternatively, an alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman (Smith and Waterman, 1981, Advances in Applied Mathematics 2:482-489; database: European Bioinformatics Institute http://www.ebi.ac.uk/MPsrch/; Smith and Waterman, 1981, J. of Molec.Biol., 147:195-197; Nicholas et al., 1998, "A Tutorial on Searching Sequence Databases and Sequence Scoring Methods" (www.psc.edu) and references cited therein; W. R. Pearson, 1991, Genomics 11:635-650). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff (Dayhoff: Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA), and normalized by Gribskov (Gribskov 1986 Nucl. Acids Res. 14(6):6745-6763). The Smith-Waterman algorithm may be employed where default parameters are used for scoring (for example, gap open penalty of 12, gap extension penalty of two). From the data generated, the "Match" value reflects "sequence identity."

[0022] Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that hybridize to the nucleic acid sequence of any of SEQ ID NOs:1 through 18. The stringency of hybridization can be controlled by temperature, ionic strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing. Conditions routinely used are set out in readily available procedure texts (e.g., Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook et al., Molecular Cloning, Cold Spring Harbor (1989)). In some embodiments, a nucleic acid molecule of the invention is capable of hybridizing to a nucleic acid molecule containing the nucleotide sequence of any one of SEQ ID NOs:1 through 18 under stringent hybridization conditions that comprise: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C. in a solution comprising 6× single strength citrate (SSC) (1×SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5× Denhardt's solution, 0.05% sodium pyrophosphate and 100 μg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C. in a solution containing 6×SSC, 1× Denhardt's solution, 100 μg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C. for 1 h in a solution containing 0.2×SSC and 0.1% SDS (sodium dodecyl sulfate).

[0023] In other embodiments, moderately stringent hybridization conditions are used that comprise: pretreatment of filters containing nucleic acid for 6 h at 40° C. in a solution containing 35% formamide, 5×SSC, 50 mM Tris-HCl (pH7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA; hybridization for 18-20 h at 40° C. in a solution containing 35% formamide, 5×SSC, 50 mM Tris-HCl (pH7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 μg/ml salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C. in a solution containing 2×SSC and 0.1% SDS.

[0024] Alternatively, low stringency conditions can be used that comprise: incubation for 8 hours to overnight at 37° C. in a solution comprising 20% formamide, 5×SSC, 50 mM sodium phosphate (pH 7.6), 5× Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1×SSC at about 37° C. for 1 hour.

Isolation, Production, Expression, and Mis-expression of LIMK Nucleic Acids and Polypeptides

[0025] LIMK nucleic acids and polypeptides, useful for identifying and testing agents that modulate LIMK function and for other applications related to the involvement of LIMK in the p53 pathway. LIMK nucleic acids and derivatives and orthologs thereof may be obtained using any available method. For instance, techniques for isolating cDNA or genomic DNA sequences of interest by screening DNA libraries or by using polymerase chain reaction (PCR) are well known in the art. In general, the particular use for the protein will dictate the particulars of expression, production, and purification methods. For instance, production of proteins for use in screening for modulating agents may require methods that preserve specific biological activities of these proteins, whereas production of proteins for antibody generation may require structural integrity of particular epitopes. Expression of proteins to be purified for screening or antibody production may require the addition of specific tags (e.g., generation of fusion proteins). Overexpression of a LIMK protein for assays used to assess LIMK function, such as involvement in cell cycle regulation or hypoxic response, may require expression in eukaryotic cell lines capable of these cellular activities. Techniques for the expression, production, and purification of proteins are well known in the art; any suitable means therefore may be used (e.g., Higgins S J and Hames B D (eds.) Protein Expression: A Practical Approach, Oxford University Press Inc., New York 1999; Stanbury P F et al., Principles of Fermentation Technology, 2nd edition, Elsevier Science, New York, 1995; Doonan S (ed.) Protein Purification Protocols, Humana Press, New Jersey, 1996; Coligan J E et al, Current Protocols in Protein Science (eds.), 1999, John Wiley & Sons, New York). In particular embodiments, recombinant LIMK is expressed in a cell line known to have defective p53 function (e.g. SAOS-2 osteoblasts, H1299 lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, available from American Type Culture Collection (ATCC), Manassas, Va.). The recombinant cells are used in cell-based screening assay systems of the invention, as described further below.

[0026] The nucleotide sequence encoding a LIMK polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native LIMK gene and/or its flanking regions or can be heterologous. A variety of host-vector expression systems may be utilized, such as mammalian cell systems infected with virus (e.g. vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g. baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, plasmid, or cosmid DNA. A host cell strain that modulates the expression of, modifies, and/or specifically processes the gene product may be used.

[0027] To detect expression of the LIMK gene product, the expression vector can comprise a promoter operably linked to a LIMK gene nucleic acid, one or more origins of replication, and, one or more selectable markers (e.g. thymidine kinase activity, resistance to antibiotics, etc.). Alternatively, recombinant expression vectors can be identified by assaying for the expression of the LIMK gene product based on the physical or functional properties of the LIMK protein in in vitro assay systems (e.g. immunoassays).

[0028] The LIMK protein, fragment, or derivative may be optionally expressed as a fusion, or chimeric protein product (i.e. it is joined via a peptide bond to a heterologous protein sequence of a different protein), for example to facilitate purification or detection. A chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other using standard methods and expressing the chimeric product. A chimeric product may also be made by protein synthetic techniques, e.g. by use of a peptide synthesizer (Hunkapiller et al., Nature (1984) 310:105-111).

[0029] Once a recombinant cell that expresses the LIMK gene sequence is identified, the gene product can be isolated and purified using standard methods (e.g. ion exchange, affinity, and gel exclusion chromatography; centrifugation; differential solubility; electrophoresis, cite purification ref-

erence). Alternatively, native LIMK proteins can be purified from natural sources, by standard methods (e.g. immunoaffinity purification). Once a protein is obtained, it may be quantified and its activity measured by appropriate methods, such as immunoassay, bioassay, or other measurements of physical properties, such as crystallography.

[0030] The methods of this invention may also use cells that have been engineered for altered expression (mis-expression) of LIMK or other genes associated with the p53 pathway. As used herein, mis-expression encompasses ectopic expression, over-expression, under-expression, and non-expression (e.g. by gene knock-out or blocking expression that would otherwise normally occur).

Genetically Modified Animals

[0031] Animal models that have been genetically modified to alter LIMK expression may be used in in vivo assays to test for activity of a candidate p53 modulating agent, or to further assess the role of LIMK in a p53 pathway process such as apoptosis or cell proliferation. Preferably, the altered LIMK expression results in a detectable phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal LIMK expression. The genetically modified animal may additionally have altered p53 expression (e.g. p53 knockout). Preferred genetically modified animals are mammals such as primates, rodents (preferably mice), cows, horses, goats, sheep, pigs, dogs and cats. Preferred non-mammalian species include zebrafish, *C. elegans*, and *Drosophila*. Preferred genetically modified animals are transgenic animals having a heterologous nucleic acid sequence present as an extrachromosomal element in a portion of its cells, i.e. mosaic animals (see, for example, techniques described by Jakobovits, 1994, Curr. Biol. 4:761-763.) or stably integrated into its germ line DNA (i.e., in the genomic sequence of most or all of its cells). Heterologous nucleic acid is introduced into the germ line of such transgenic animals by genetic manipulation of, for example, embryos or embryonic stem cells of the host animal.

[0032] Methods of making transgenic animals are well-known in the art (for transgenic mice see Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985), U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Pat. No. 4,873,191 by Wagner et al., and Hogan, B., Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1986); for particle bombardment see U.S. Pat. No., 4,945,050, by Sandford et al.; for transgenic *Drosophila* see Rubin and Spradling, Science (1982) 218:348-53 and U.S. Pat. No. 4,670,388; for transgenic insects see Berghammer A. J. et al., A Universal Marker for Transgenic Insects (1999) Nature 402:370-371; for transgenic Zebrafish see Lin S., Transgenic Zebrafish, Methods Mol. Biol. (2000);136:375-3830); for microinjection procedures for fish, amphibian eggs and birds see Houdebine and Chourrout, Experientia (1991) 47:897-905; for transgenic rats see Hammer et al., Cell (1990) 63:1099-1112; and for culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection see, e.g., Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E. J. Robertson, ed., IRL Press (1987)). Clones of the nonhuman transgenic animals can be produced according to available methods (see Wilmut, I. et al. (1997) Nature 385:810-813; and PCT International Publication Nos. WO 97/07668 and WO 97/07669).

[0033] In one embodiment, the transgenic animal is a "knock-out" animal having a heterozygous or homozygous alteration in the sequence of an endogenous LIMK gene that results in a decrease of LIMK function, preferably such that LIMK expression is undetectable or insignificant. Knock-out animals are typically generated by homologous recombination with a vector comprising a transgene having at least a portion of the gene to be knocked out. Typically a deletion, addition or substitution has been introduced into the transgene to functionally disrupt it. The transgene can be a human gene (e.g., from a human genomic clone) but more preferably is an ortholog of the human gene derived from the transgenic host species. For example, a mouse LIMK gene is used to construct a homologous recombination vector suitable for altering an endogenous LIMK gene in the mouse genome. Detailed methodologies for homologous recombination in mice are available (see Capecchi, Science (1989) 244:1288-1292; Joyner et al., Nature (1989) 338:153-156). Procedures for the production of non-rodent transgenic mammals and other animals are also available (Houdebine and Chourrout, supra; Pursel et al., Science (1989) 244:1281-1288; Simms et al., Bio/Technology (1988) 6:179-183). In a preferred embodiment, knock-out animals, such as mice harboring a knockout of a specific gene, may be used to produce antibodies against the human counterpart of the gene that has been knocked out (Claesson M H et al., (1994) Scan J Immunol 40:257-264; Declerck P J et al., (1995) J Biol. Chem. 270:8397-400).

[0034] In another embodiment, the transgenic animal is a "knock-in" animal having an alteration in its genome that results in altered expression (e.g., increased (including ectopic) or decreased expression) of the LIMK gene, e.g., by introduction of additional copies of LIMK, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the LIMK gene. Such regulatory sequences include inducible, tissue-specific, and constitutive promoters and enhancer elements. The knock-in can be homozygous or heterozygous.

[0035] Transgenic nonhuman animals can also be produced that contain selected systems allowing for regulated expression of the transgene. One example of such a system that may be produced is the cre/loxP recombinase system of bacteriophage P 1 (Lakso et al., PNAS (1992) 89:6232-6236; U.S. Pat. No. 4,959,317). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) Science 251:1351-1355; U.S. Pat. No. 5,654,182). In a preferred embodiment, both Cre-LoxP and Flp-Frt are used in the same system to regulate expression of the transgene, and for sequential deletion of vector sequences in the same cell (Sun X et al (2000) Nat Genet 25:83-6).

[0036] The genetically modified animals can be used in genetic studies to further elucidate the p53 pathway, as

animal models of disease and disorders implicating defective p53 function, and for in vivo testing of candidate therapeutic agents, such as those identified in screens described below. The candidate therapeutic agents are administered to a genetically modified animal having altered LIMK function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered LIMK expression that receive candidate therapeutic agent.

[0037] In addition to the above-described genetically modified animals having altered LIMK function, animal models having defective p53 function (and otherwise normal LIMK function), can be used in the methods of the present invention. For example, a p53 knockout mouse can be used to assess, in vivo, the activity of a candidate p53 modulating agent identified in one of the in vitro assays described below. p53 knockout mice are described in the literature (Jacks et al., Nature 2001;410:1111-1116, 1043-1044; Donehower et al., supra). Preferably, the candidate p53 modulating agent when administered to a model system with cells defective in p53 function, produces a detectable phenotypic change in the model system indicating that the p53 function is restored, i.e., the cells exhibit normal cell cycle progression.

Modulating Agents

[0038] The invention provides methods to identify agents that interact with and/or modulate the function of LIMK and/or the p53 pathway. Such agents are useful in a variety of diagnostic and therapeutic applications associated with the p53 pathway, as well as in further analysis of the LIMK protein and its contribution to the p53 pathway. Accordingly, the invention also provides methods for modulating the p53 pathway comprising the step of specifically modulating LIMK activity by administering a LIMK-interacting or -modulating agent.

[0039] In a preferred embodiment, LIMK-modulating agents inhibit or enhance LIMK activity or otherwise affect normal LIMK function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In a further preferred embodiment, the candidate p53 pathway-modulating agent specifically modulates the function of the LIMK. The phrases "specific modulating agent", "specifically modulates", etc., are used herein to refer to modulating agents that directly bind to the LIMK polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the LIMK. The term also encompasses modulating agents that alter the interaction of the LIMK with a binding partner or substrate (e.g. by binding to a binding partner of a LIMK, or to a protein/binding partner complex, and inhibiting function).

[0040] Preferred LIMK-modulating agents include small molecule compounds; LIMK-interacting proteins, including antibodies and other biotherapeutics; and nucleic acid modulators such as antisense and RNA inhibitors. The modulating agents may be formulated in pharmaceutical compositions, for example, as compositions that may comprise other active ingredients, as in combination therapy, and/or suitable carriers or excipients. Techniques for formulation and administration of the compounds may be found in "Remington's Pharmaceutical Sciences" Mack Publishing Co., Easton, Pa., 19th edition.

[0041] Small Molecule Modulators

[0042] Small molecules, are often preferred to modulate function of proteins with enzymatic function, and/or containing protein interaction domains. Chemical agents, referred to in the art as "small molecule" compounds are typically organic, non-peptide molecules, having a molecular weight less than 10,000, preferably less than 5,000, more preferably less than 1,000, and most preferably less than 500. This class of modulators includes chemically synthesized molecules, for instance, compounds from combinatorial chemical libraries. Synthetic compounds may be rationally designed or identified based on known or inferred properties of the LIMK protein or may be identified by screening compound libraries. Alternative appropriate modulators of this class are natural products, particularly secondary metabolites from organisms such as plants or fungi, which can also be identified by screening compound libraries for LIMK-modulating activity. Methods for generating and obtaining compounds are well known in the art (Schreiber S L, Science (2000) 151: 1964-1969; Radmann J and Gunther J, Science (2000) 151:1947-1948).

[0043] Small molecule modulators identified from screening assays, as described below, can be used as lead compounds from which candidate clinical compounds may be designed, optimized, and synthesized. Such clinical compounds may have utility in treating pathologies associated with the p53 pathway. The activity of candidate small molecule modulating agents may be improved several-fold through iterative secondary functional validation, as further described below, structure determination, and candidate modulator modification and testing. Additionally, candidate clinical compounds are generated with specific regard to clinical and pharmacological properties. For example, the reagents may be derivatized and re-screened using in vitro and in vivo assays to optimize activity and minimize toxicity for pharmaceutical development.

[0044] Protein Modulators

[0045] Specific LIMK-interacting proteins are useful in a variety of diagnostic and therapeutic applications related to the p53 pathway and related disorders, as well as in validation assays for other LIMK-modulating agents. In a preferred embodiment, LIMK-interacting proteins affect normal LIMK function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In another embodiment, LIMK-interacting proteins are useful in detecting and providing information about the function of LIMK proteins, as is relevant to p53 related disorders, such as cancer (e.g., for diagnostic means).

[0046] An LIMK-interacting protein may be endogenous, i.e. one that naturally interacts genetically or biochemically with a LIMK, such as a member of the LIMK pathway that modulates LIMK expression, localization, and/or activity. LIMK-modulators include dominant negative forms of LIMK-interacting proteins and of LIMK proteins themselves. Yeast two-hybrid and variant screens offer preferred methods for identifying endogenous LIMK-interacting proteins (Finley, R. L. et al. (1996) in DNA Cloning-Expression Systems: A Practical Approach, eds. Glover D. & Hames B. D (Oxford University Press, Oxford, England), pp. 169-203; Fashema S F et al., Gene (2000) 250:1-14; Drees B L Curr Opin Chem Biol (1999) 3:64-70; Vidal M and Legrain P Nucleic Acids Res (1999) 27:919-29; and U.S. Pat. No.

5,928,868). Mass spectrometry is an alternative preferred method for the elucidation of protein complexes (reviewed in, e.g., Pandley A and Mann M, Nature (2000) 405:837-846; Yates J R 3rd, Trends Genet (2000) 16:5-8).

[0047] An LIMK-interacting protein may be an exogenous protein, such as a LIMK-specific antibody or a T-cell antigen receptor (see, e.g., Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory; Harlow and Lane (1999) Using antibodies: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press). LIMK antibodies are further discussed below.

[0048] In preferred embodiments, a LIMK-interacting protein specifically binds a LIMK protein. In alternative preferred embodiments, a LIMK-modulating agent binds a LIMK substrate, binding partner, or cofactor.

[0049] Antibodies

[0050] In another embodiment, the protein modulator is a LIMK specific antibody agonist or antagonist. The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify LIMK modulators. The antibodies can also be used in dissecting the portions of the LIMK pathway responsible for various cellular responses and in the general processing and maturation of the LIMK.

[0051] Antibodies that specifically bind LIMK polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of LIMK polypeptide, and more preferably, to human LIMK. Antibodies may be polyclonal, monoclonal (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab').sub.2 fragments, fragments produced by a FAb expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Epitopes of LIMK which are particularly antigenic can be selected, for example, by routine screening of LIMK polypeptides for antigenicity or by applying a theoretical method for selecting antigenic regions of a protein (Hopp and Wood (1981), Proc. Natl. Acad. Sci. U.S.A. 78:3824-28; Hopp and Wood, (1983) Mol. Immunol. 20:483-89; Sutcliffe et al., (1983) Science 219:660-66) to the amino acid sequence shown in any of SEQ ID NOs:19, 20, 21, 22, 23, 24, or 25. Monoclonal antibodies with affinities of 10^8 M^{-1} preferably 10^9 M^{-1} to 10^{10} M^{-1}, or stronger can be made by standard procedures as described (Harlow and Lane, supra; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed) Academic Press, New York; and U.S. Pat. Nos. 4,381, 292; 4,451,570; and 4,618,577). Antibodies may be generated against crude cell extracts of LIMK or substantially purified fragments thereof. If LIMK fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of a LIMK protein. In a particular embodiment, LIMK-specific antigens and/or immunogens are coupled to carrier proteins that stimulate the immune response. For example, the subject polypeptides are covalently coupled to the keyhole limpet hemocyanin (KLH) carrier, and the conjugate is emulsified in Freund's complete adjuvant, which enhances the immune response. An appropriate immune system such as a laboratory rabbit or mouse is immunized according to conventional protocols.

[0052] The presence of LIMK-specific antibodies is assayed by an appropriate assay such as a solid phase enzyme-linked immunosorbant assay (ELISA) using immobilized corresponding LIMK polypeptides. Other assays, such as radioimmunoassays or fluorescent assays might also be used.

[0053] Chimeric antibodies specific to LIMK polypeptides can be made that contain different portions from different animal species. For instance, a human immunoglobulin constant region may be linked to a variable region of a murine mAb, such that the antibody derives its biological activity from the human antibody, and its binding specificity from the murine fragment. Chimeric antibodies are produced by splicing together genes that encode the appropriate regions from each species (Morrison et al., Proc. Natl. Acad. Sci. (1984) 81:6851-6855; Neuberger et al., Nature (1984) 312:604-608; Takeda et al., Nature (1985) 31:452-454). Humanized antibodies, which are a form of chimeric antibodies, can be generated by grafting complementary-determining regions (CDRs) (Carlos, T. M., J. M. Harlan. 1994. Blood 84:2068-2101) of mouse antibodies into a background of human framework regions and constant regions by recombinant DNA technology (Riechmann L M, et al., 1988 Nature 323: 323-327). Humanized antibodies contain ~10% murine sequences and 90% human sequences, and thus further reduce or eliminate immunogenicity, while retaining the antibody specificities (Co MS, and Queen C. 1991 Nature 351: 501-501; Morrison S L. 1992 Ann. Rev. Immun. 10:239-265). Humanized antibodies and methods of their production are well-known in the art (U.S. Pat. Nos. 5,530,101, 5,585,089, 5,693,762, and 6,180,370).

[0054] LIMK-specific single chain antibodies which are recombinant, single chain polypeptides formed by linking the heavy and light chain fragments of the Fv regions via an amino acid bridge, can be produced by methods known in the art (U.S. Pat. No. 4,946,778; Bird, Science (1988) 242:423-426; Huston et al., Proc. Natl. Acad. Sci. USA (1988) 85:5879-5883; and Ward et al., Nature (1989) 334:544-546).

[0055] Other suitable techniques for antibody production involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors (Huse et al., Science (1989) 246:1275-1281). As used herein, T-cell antigen receptors are included within the scope of antibody modulators (Harlow and Lane, 1988, supra).

[0056] The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, antibodies will be labeled by joining, either covalently or non-covalently, a substance that provides for a detectable signal, or that is toxic to cells that express the targeted protein (Menard S, et al., Int J. Biol Markers (1989) 4:131-134). A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, fluorescent emitting lanthanide metals, chemiluminescent moieties, bioluminescent moieties, magnetic particles, and the like (U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241). Also, recombinant immunoglobulins may be produced (U.S. Pat. No. 4,816,567). Antibodies to cytoplasmic polypeptides may be delivered and reach their targets by conjugation with membrane-penetrating toxin proteins (U.S. Pat. No. 6,086,900).

[0057] When used therapeutically in a patient, the antibodies of the subject invention are typically administered parenterally, when possible at the target site, or intravenously. The therapeutically effective dose and dosage regimen is determined by clinical studies. Typically, the amount of antibody administered is in the range of about 0.1 mg/kg—to about 10 mg/kg of patient weight. For parenteral administration, the antibodies are formulated in a unit dosage injectable form (e.g., solution, suspension, emulsion) in association with a pharmaceutically acceptable vehicle. Such vehicles are inherently nontoxic and non-therapeutic. Examples are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Nonaqueous vehicles such as fixed oils, ethyl oleate, or liposome carriers may also be used. The vehicle may contain minor amounts of additives, such as buffers and preservatives, which enhance isotonicity and chemical stability or otherwise enhance therapeutic potential. The antibodies' concentrations in such vehicles are typically in the range of about 1 mg/ml to about 10 mg/ml. Immunotherapeutic methods are further described in the literature (U.S. Pat. No. 5,859,206; WO0073469).

[0058] Nucleic Acid Modulators

[0059] Other preferred LIMK-modulating agents comprise nucleic acid molecules, such as antisense oligomers or double stranded RNA (dsRNA), which generally inhibit LIMK activity. Preferred nucleic acid modulators interfere with the function of the LIMK nucleic acid such as DNA replication, transcription, translocation of the LIMK RNA to the site of protein translation, translation of protein from the LIMK RNA, splicing of the LIMK RNA to yield one or more mRNA species, or catalytic activity which may be engaged in or facilitated by the LIMK RNA.

[0060] In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to a LIMK mRNA to bind to and prevent translation, preferably by binding to the 5' untranslated region. LIMK-specific antisense oligonucleotides, preferably range from at least 6 to about 200 nucleotides. In some embodiments the oligonucleotide is preferably at least 10, 15, or 20 nucleotides in length. In other embodiments, the oligonucleotide is preferably less than 50, 40, or 30 nucleotides in length. The oligonucleotide can be DNA or RNA or a chimeric mixture or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appending groups such as peptides, agents that facilitate transport across the cell membrane, hybridization-triggered cleavage agents, and intercalating agents.

[0061] In another embodiment, the antisense oligomer is a phosphothioate morpholino oligomer (PMO). PMOs are assembled from four different morpholino subunits, each of which contain one of four genetic bases (A, C, G, or T) linked to a six-membered morpholine ring. Polymers of these subunits are joined by non-ionic phosphodiamidate intersubunit linkages. Details of how to make and use PMOs and other antisense oligomers are well known in the art (e.g. see WO99/18193; Probst J C, Antisense Oligodeoxynucleotide and Ribozyme Design, Methods. (2000) 22(3):271-281; Summerton J, and Weller D. 1997 Antisense Nucleic Acid Drug Dev.:7:187-95; U.S. Pat. No. 5,235,033; and U.S. Pat. No. 5,378,841).

[0062] Alternative preferred LIMK nucleic acid modulators are double-stranded RNA species mediating RNA interference (RNAi). RNAi is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene. Methods relating to the use of RNAi to silence genes in *C. elegans, Drosophila*, plants, and humans are known in the art (Fire A, et al., 1998 Nature 391:806-811; Fire, A. Trends Genet. 15, 358-363 (1999); Sharp, P. A. RNA interference 2001. Genes Dev. 15, 485-490 (2001); Hammond, S. M., et al., Nature Rev. Genet. 2, 110-1119 (2001); Tuschl, T. Chem. Biochem. 2, 239-245 (2001); Hamilton, A. et al., Science 286, 950-952 (1999); Hammond, S. M., et al., Nature 404, 293-296 (2000); Zamore, P. D., et al., Cell 101, 25-33 (2000); Bernstein, E., et al., Nature 409, 363-366 (2001); Elbashir, S. M., et al., Genes Dev. 15, 188-200 (2001); WO0129058; WO9932619; Elbashir S M, et al., 2001 Nature 411:494-498).

[0063] Nucleic acid modulators are commonly used as research reagents, diagnostics, and therapeutics. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used to elucidate the function of particular genes (see, for example, U.S. Pat. No. 6,165,790). Nucleic acid modulators are also used, for example, to distinguish between functions of various members of a biological pathway. For example, antisense oligomers have been employed as therapeutic moieties in the treatment of disease states in animals and man and have been demonstrated in numerous clinical trials to be safe and effective (Milligan J F, et al, Current Concepts in Antisense Drug Design, J Med. Chem. (1993) 36:1923-1937; Tonkinson J L et al., Antisense Oligodeoxynucleotides as Clinical Therapeutic Agents, Cancer Invest. (1996) 14:54-65). Accordingly, in one aspect of the invention, a LIMK-specific nucleic acid modulator is used in an assay to further elucidate the role of the LIMK in the p53 pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, a LIMK-specific antisense oligomer is used as a therapeutic agent for treatment of p53-related disease states.

Assay Systems

[0064] The invention provides assay systems and screening methods for identifying specific modulators of LIMK activity. As used herein, an "assay system" encompasses all the components required for performing and analyzing results of an assay that detects and/or measures a particular event. In general, primary assays are used to identify or confirm a modulator's specific biochemical or molecular effect with respect to the LIMK nucleic acid or protein. In general, secondary assays further assess the activity of a LIMK modulating agent identified by a primary assay and may confirm that the modulating agent affects LIMK in a manner relevant to the p53 pathway. In some cases, LIMK modulators will be directly tested in a secondary assay.

[0065] In a preferred embodiment, the screening method comprises contacting a suitable assay system comprising a LIMK polypeptide with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. kinase activity), which is based on the particular molecular event the screening method detects. A statistically significant difference between

the agent-biased activity and the reference activity indicates that the candidate agent modulates LIMK activity, and hence the p53 pathway.

[0066] Primary Assays

[0067] The type of modulator tested generally determines the type of primary assay.

[0068] Primary Assays for Small Molecule Modulators

[0069] For small molecule modulators, screening assays are used to identify candidate modulators. Screening assays may be cell-based or may use a cell-free system that recreates or retains the relevant biochemical reaction of the target protein (reviewed in Sittampalam G S et al., Curr Opin Chem Biol (1997) 1:384-91 and accompanying references). As used herein the term "cell-based" refers to assays using live cells, dead cells, or a particular cellular fraction, such as a membrane, endoplasmic reticulum, or mitochondrial fraction. The term "cell free" encompasses assays using substantially purified protein (either endogenous or recombinantly produced), partially purified or crude cellular extracts. Screening assays may detect a variety of molecular events, including protein-DNA interactions, protein-protein interactions (e.g., receptor-ligand binding), transcriptional activity (e.g., using a reporter gene), enzymatic activity (e.g., via a property of the substrate), activity of second messengers, immunogenicty and changes in cellular morphology or other cellular characteristics. Appropriate screening assays may use a wide range of detection methods including fluorescent, radioactive, colorimetric, spectrophotometric, and amperometric methods, to provide a read-out for the particular molecular event detected.

[0070] Cell-based screening assays usually require systems for recombinant expression of LIMK and any auxiliary proteins demanded by the particular assay. Appropriate methods for generating recombinant proteins produce sufficient quantities of proteins that retain their relevant biological activities and are of sufficient purity to optimize activity and assure assay reproducibility. Yeast two-hybrid and variant screens, and mass spectrometry provide preferred methods for determining protein-protein interactions and elucidation of protein complexes. In certain applications, when LIMK-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the LIMK protein may be assayed by various known methods such as substrate processing (e.g. ability of the candidate LIMK-specific binding agents to function as negative effectors in LIMK-expressing cells), binding equilibrium constants (usually at least about $10^7 M^{-1}$, preferably at least about $10^8 M^{-1}$, more preferably at least about $10^9 M^{-1}$), and immunogenicity (e.g. ability to elicit LIMK specific antibody in a heterologous host such as a mouse, rat, goat or rabbit). For enzymes and receptors, binding may be assayed by, respectively, substrate and ligand processing.

[0071] The screening assay may measure a candidate agent's ability to specifically bind to or modulate activity of a LIMK polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The LIMK polypeptide can be full length or a fragment thereof that retains functional LIMK activity. The LIMK polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The LIMK polypeptide is preferably human LIMK, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of LIMK interaction with a binding target, such as an endogenous or exogenous protein or other substrate that has LIMK-specific binding activity, and can be used to assess normal LIMK gene function.

[0072] Suitable assay formats that may be adapted to screen for LIMK modulators are known in the art. Preferred screening assays are high throughput or ultra high throughput and thus provide automated, cost-effective means of screening compound libraries for lead compounds (Fernandes P B, Curr Opin Chem Biol (1998) 2:597-603; Sundberg S A, Curr Opin Biotechnol 2000, 11:47-53). In one preferred embodiment, screening assays uses fluorescence technologies, including fluorescence polarization, time-resolved fluorescence, and fluorescence resonance energy transfer. These systems offer means to monitor protein-protein or DNA-protein interactions in which the intensity of the signal emitted from dye-labeled molecules depends upon their interactions with partner molecules (e.g., Selvin P R, Nat Struct Biol (2000) 7:730-4; Fernandes P B, supra; Hertzberg R P and Pope A J, Curr Opin Chem Biol (2000) 4:445-451).

[0073] A variety of suitable assay systems may be used to identify candidate LIMK and p53 pathway modulators (e.g. U.S. Pat. No. 6,165,992 (kinase assays); U.S. Pat. Nos. 5,550,019 and 6,133,437 (apoptosis assays); and U.S. Pat. No. 6,020,135 (p53 modulation), among others). Specific preferred assays are described in more detail below.

[0074] Kinase assays. In some preferred embodiments the screening assay detects the ability of the test agent to modulate the kinase activity of a LIMK polypeptide. In further embodiments, a cell-free kinase assay system is used to identify a candidate p53 modulating agent, and a secondary, cell-based assay, such as an apoptosis or hypoxic induction assay (described below), may be used to further characterize the candidate p53 modulating agent. Many different assays for kinases have been reported in the literature and are well known to those skilled in the art (e.g. U.S. Pat. No. 6,165,992; Zhu et al., Nature Genetics (2000) 26:283-289; and WO0073469). Radioassays, which monitor the transfer of a gamma phosphate are frequently used. For instance, a scintillation assay for p56 (lck) kinase activity monitors the transfer of the gamma phosphate from gamma—^{33}P ATP to a biotinylated peptide substrate; the substrate is captured on a streptavidin coated bead that transmits the signal (Beveridge M et al., J Biomol Screen (2000) 5:205-212). This assay uses the scintillation proximity assay (SPA), in which only radio-ligand bound to receptors tethered to the surface of an SPA bead are detected by the scintillant immobilized within it, allowing binding to be measured without separation of bound from free ligand.

[0075] Another example of antibody based assays for protein kinase activity is TRF (time-resolved fluorometry). This method utilizes europium chelate-labeled anti-phosphotyrosine antibodies to detect phosphate transfer to a polymeric substrate coated onto microtiter plate wells. The amount of phosphorylation is then detected using time-resolved, dissociation-enhanced fluorescence (Braunwalder A F, et al., Anal Biochem 1996 Jul. 1;238(2):159-64).

[0076] Apoptosis assays. Assays for apoptosis may be performed by terminal deoxynucleotidyl transferase-medi-

ated digoxigenin-11-dUTP nick end labeling (TUNEL) assay. The TUNEL assay is used to measure nuclear DNA fragmentation characteristic of apoptosis (Lazebnik et al., 1994, Nature 371, 346), by following the incorporation of fluorescein-dUTP (Yonehara et al., 1989, J. Exp. Med. 169, 1747). Apoptosis may further be assayed by acridine orange staining of tissue culture cells (Lucas, R., et al., 1998, Blood 15:4730-41). An apoptosis assay system may comprise a cell that expresses a LIMK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the apoptosis assay system and changes in induction of apoptosis relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, an apoptosis assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using a cell-free assay system. An apoptosis assay may also be used to test whether LIMK function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express LIMK relative to wild type cells. Differences in apoptotic response compared to wild type cells suggests that the LIMK plays a direct role in the apoptotic response. Apoptosis assays are described further in U.S. Pat. No. 6,133,437.

[0077] Cell proliferation and cell cycle assays. Cell proliferation may be assayed via bromodeoxyuridine (BRDU) incorporation. This assay identifies a cell population undergoing DNA synthesis by incorporation of BRDU into newly-synthesized DNA. Newly-synthesized DNA may then be detected using an anti-BRDU antibody (Hoshino et al., 1986, Int. J. Cancer 38, 369; Campana et al., 1988, J. Immunol. Meth. 107, 79), or by other means.

[0078] Cell Proliferation may also be examined using [^{3}H]-thymidine incorporation (Chen, J., 1996, Oncogene 13:1395-403; Jeoung, J., 1995, J. Biol. Chem. 270:18367-73). This assay allows for quantitative characterization of S-phase DNA syntheses. In this assay, cells synthesizing DNA will incorporate [^{3}H]-thymidine into newly synthesized DNA. Incorporation can then be measured by standard techniques such as by counting of radioisotope in a scintillation counter (e.g., Beckman L S 3800 Liquid Scintillation Counter).

[0079] Cell proliferation may also be assayed by colony formation in soft agar (Sambrook et al., Molecular Cloning, Cold Spring Harbor (1989)). For example, cells transformed with LIMK are seeded in soft agar plates, and colonies are measured and counted after two weeks incubation.

[0080] Involvement of a gene in the cell cycle may be assayed by flow cytometry (Gray J W et al. (1986) Int J Radiat Biol Relat Stud Phys Chem Med 49:237-55). Cells transfected with a LIMK may be stained with propidium iodide and evaluated in a flow cytometer (available from Becton Dickinson).

[0081] Accordingly, a cell proliferation or cell cycle assay system may comprise a cell that expresses a LIMK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the assay system and changes in cell proliferation or cell cycle relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the cell proliferation or cell cycle assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system such as a cell-free kinase assay system. A cell proliferation assay may also be used to test whether LIMK function plays a direct role in cell proliferation or cell cycle. For example, a cell proliferation or cell cycle assay may be performed on cells that over- or under-express LIMK relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the LIMK plays a direct role in cell proliferation or cell cycle.

[0082] Angiogenesis. Angiogenesis may be assayed using various human endothelial cell systems, such as umbilical vein, coronary artery, or dermal cells. Suitable assays include Alamar Blue based assays (available from Biosource International) to measure proliferation; migration assays using fluorescent molecules, such as the use of Becton Dickinson Falcon HTS FluoroBlock cell culture inserts to measure migration of cells through membranes in presence or absence of angiogenesis enhancer or suppressors; and tubule formation assays based on the formation of tubular structures by endothelial cells on Matrigel® (Becton Dickinson). Accordingly, an angiogenesis assay system may comprise a cell that expresses a LIMK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the angiogenesis assay system and changes in angiogenesis relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the angiogenesis assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system. An angiogenesis assay may also be used to test whether LIMK function plays a direct role in cell proliferation. For example, an angiogenesis assay may be performed on cells that over- or under-express LIMK relative to wild type cells. Differences in angiogenesis compared to wild type cells suggests that the LIMK plays a direct role in angiogenesis.

[0083] Hypoxic induction. The alpha subunit of the transcription factor, hypoxia inducible factor-1 (HIF-1), is upregulated in tumor cells following exposure to hypoxia in vitro. Under hypoxic conditions, HIF-1 stimulates the expression of genes known to be important in tumour cell survival, such as those encoding glyolytic enzymes and VEGF. Induction of such genes by hypoxic conditions may be assayed by growing cells transfected with LIMK in hypoxic conditions (such as with 0.1% O2, 5% CO2, and balance N2, generated in a Napco 7001 incubator (Precision Scientific)) and normoxic conditions, followed by assessment of gene activity or expression by Taqman®. For example, a hypoxic induction assay system may comprise a cell that expresses a LIMK, and that optionally has a mutated p53 (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the hypoxic induction assay system and changes in hypoxic response relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the hypoxic induction assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system. A hypoxic induction assay may also be used to test whether LIMK function plays a direct role in the hypoxic response. For example, a hypoxic induction assay may be performed

on cells that over- or under-express LIMK relative to wild type cells. Differences in hypoxic response compared to wild type cells suggests that the LIMK plays a direct role in hypoxic induction.

[0084] Cell adhesion. Cell adhesion assays measure adhesion of cells to purified adhesion proteins, or adhesion of cells to each other, in presence or absence of candidate modulating agents. Cell-protein adhesion assays measure the ability of agents to modulate the adhesion of cells to purified proteins. For example, recombinant proteins are produced, diluted to 2.5 g/mL in PBS, and used to coat the wells of a microtiter plate. The wells used for negative control are not coated. Coated wells are then washed, blocked with 1% BSA, and washed again. Compounds are diluted to 2×final test concentration and added to the blocked, coated wells. Cells are then added to the wells, and the unbound cells are washed off. Retained cells are labeled directly on the plate by adding a membrane-permeable fluorescent dye, such as calcein-AM, and the signal is quantified in a fluorescent microplate reader.

[0085] Cell-cell adhesion assays measure the ability of agents to modulate binding of cell adhesion proteins with their native ligands. These assays use cells that naturally or recombinantly express the adhesion protein of choice. In an exemplary assay, cells expressing the cell adhesion protein are plated in wells of a multiwell plate. Cells expressing the ligand are labeled with a membrane-permeable fluorescent dye, such as BCECF, and allowed to adhere to the monolayers in the presence of candidate agents. Unbound cells are washed off, and bound cells are detected using a fluorescence plate reader.

[0086] High-throughput cell adhesion assays have also been described. In one such assay, small molecule ligands and peptides are bound to the surface of microscope slides using a microarray spotter, intact cells are then contacted with the slides, and unbound cells are washed off. In this assay, not only the binding specificity of the peptides and modulators against cell lines are determined, but also the functional cell signaling of attached cells using immunofluorescence techniques in situ on the microchip is measured (Falsey J R et al., Bioconjug Chem. 2001 May-June;12(3):346-53).

[0087] Primary Assays for Antibody Modulators

[0088] For antibody modulators, appropriate primary assays test is a binding assay that tests the antibody's affinity to and specificity for the LIMK protein. Methods for testing antibody affinity and specificity are well known in the art (Harlow and Lane, 1988, 1999, supra). The enzyme-linked immunosorbant assay (ELISA) is a preferred method for detecting LIMK-specific antibodies; others include FACS assays, radioimmunoassays, and fluorescent assays.

[0089] Primary Assays for Nucleic Acid Modulators

[0090] For nucleic acid modulators, primary assays may test the ability of the nucleic acid modulator to inhibit or enhance LIMK gene expression, preferably mRNA expression. In general, expression analysis comprises comparing LIMK expression in like populations of cells (e.g., two pools of cells that endogenously or recombinantly express LIMK) in the presence and absence of the nucleic acid modulator. Methods for analyzing mRNA and protein expression are well known in the art. For instance, Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR (e.g., using the TaqMan®, PE Applied Biosystems), or microarray analysis may be used to confirm that LIMK mRNA expression is reduced in cells treated with the nucleic acid modulator (e.g., Current Protocols in Molecular Biology (1994) Ausubel F M et al., eds., John Wiley & Sons, Inc., chapter 4; Freeman W M et al., Biotechniques (1999) 26:112-125; Kallioniemi O P, Ann Med 2001, 33:142-147; Blohm D H and Guiseppi-Elie, A Curr Opin Biotechnol 2001, 12:41-47). Protein expression may also be monitored. Proteins are most commonly detected with specific antibodies or antisera directed against either the LIMK protein or specific peptides. A variety of means including Western blotting, ELISA, or in situ detection, are available (Harlow E and Lane D, 1988 and 1999, supra).

[0091] Secondary Assays

[0092] Secondary assays may be used to further assess the activity of LIMK-modulating agent identified by any of the above methods to confirm that the modulating agent affects LIMK in a manner relevant to the p53 pathway. As used herein, LIMK-modulating agents encompass candidate clinical compounds or other agents derived from previously identified modulating agent. Secondary assays can also be used to test the activity of a modulating agent on a particular genetic or biochemical pathway or to test the specificity of the modulating agent's interaction with LIMK.

[0093] Secondary assays generally compare like populations of cells or animals (e.g., two pools of cells or animals that endogenously or recombinantly express LIMK) in the presence and absence of the candidate modulator. In general, such assays test whether treatment of cells or animals with a candidate LIMK-modulating agent results in changes in the p53 pathway in comparison to untreated (or mock- or placebo-treated) cells or animals. Certain assays use "sensitized genetic backgrounds", which, as used herein, describe cells or animals engineered for altered expression of genes in the p53 or interacting pathways.

[0094] Cell-Based Assays

[0095] Cell based assays may use a variety of mammalian cell lines known to have defective p53 function (e.g. SAOS-2 osteoblasts, H1299 lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, available from American Type Culture Collection (ATCC), Manassas, Va.). Cell based assays may detect endogenous p53 pathway activity or may rely on recombinant expression of p53 pathway components. Any of the aforementioned assays may be used in this cell-based format. Candidate modulators are typically added to the cell media but may also be injected into cells or delivered by any other efficacious means.

[0096] Animal Assays

[0097] A variety of non-human animal models of normal or defective p53 pathway may be used to test candidate LIMK modulators. Models for defective p53 pathway typically use genetically modified animals that have been engineered to mis-express (e.g., over-express or lack expression in) genes involved in the p53 pathway. Assays generally require systemic delivery of the candidate modulators, such as by oral administration, injection, etc.

[0098] In a preferred embodiment, p53 pathway activity is assessed by monitoring neovascularization and angiogen-

esis. Animal models with defective and normal p53 are used to test the candidate modulator's affect on LIMK in Matrigel® assays. Matrigel® is an extract of basement membrane proteins, and is composed primarily of laminin, collagen IV, and heparin sulfate proteoglycan. It is provided as a sterile liquid at 4° C., but rapidly forms a solid gel at 37° C. Liquid Matrigel® is mixed with various angiogenic agents, such as bFGF and VEGF, or with human tumor cells which overexpress the LIMK. The mixture is then injected subcutaneously(SC) into female athymic nude mice (Taconic, Germantown, N.Y.) to support an intense vascular response. Mice with Matrigel® pellets may be dosed via oral (PO), intraperitoneal (IP), or intravenous (IV) routes with the candidate modulator. Mice are euthanized 5-12 days post-injection, and the Matrigel® pellet is harvested for hemoglobin analysis (Sigma plasma hemoglobin kit). Hemoglobin content of the gel is found to correlate the degree of neovascularization in the gel.

[0099] In another preferred embodiment, the effect of the candidate modulator on LIMK is assessed via tumorigenicity assays. In one example, xenograft human tumors are implanted SC into female athymic mice, 6-7 week old, as single cell suspensions either from a pre-existing tumor or from in vitro culture. The tumors which express the LIMK endogenously are injected in the flank, 1×10^5 to 1×10^7 cells per mouse in a volume of 100 μL using a 27 gauge needle. Mice are then ear tagged and tumors are measured twice weekly. Candidate modulator treatment is initiated on the day the mean tumor weight reaches 100 mg. Candidate modulator is delivered IV, SC, IP, or PO by bolus administration. Depending upon the pharmacokinetics of each unique candidate modulator, dosing can be performed multiple times per day. The tumor weight is assessed by measuring perpendicular diameters with a caliper and calculated by multiplying the measurements of diameters in two dimensions. At the end of the experiment, the excised tumors maybe utilized for biomarker identification or further analyses. For immunohistochemistry staining, xenograft tumors are fixed in 4% paraformaldehyde, 0.1M phosphate, pH 7.2, for 6 hours at 4° C., immersed in 30% sucrose in PBS, and rapidly frozen in isopentane cooled with liquid nitrogen.

Diagnostic and Therapeutic Uses

[0100] Specific LIMK-modulating agents are useful in a variety of diagnostic and therapeutic applications where disease or disease prognosis is related to defects in the p53 pathway, such as angiogenic, apoptotic, or cell proliferation disorders. Accordingly, the invention also provides methods for modulating the p53 pathway in a cell, preferably a cell pre-determined to have defective p53 function, comprising the step of administering an agent to the cell that specifically modulates LIMK activity. Preferably, the modulating agent produces a detectable phenotypic change in the cell indicating that the p53 function is restored, i.e., for example, the cell undergoes normal proliferation or progression through the cell cycle.

[0101] The discovery that LIMK is implicated in p53 pathway provides for a variety of methods that can be employed for the diagnostic and prognostic evaluation of diseases and disorders involving defects in the p53 pathway and for the identification of subjects having a predisposition to such diseases and disorders.

[0102] Various expression analysis methods can be used to diagnose whether LIMK expression occurs in a particular sample, including Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR, and microarray analysis. (e.g., Current Protocols in Molecular Biology (1994) Ausubel F M et al., eds., John Wiley & Sons, Inc., chapter 4; Freeman W M et al., Biotechniques (1999) 26:112-125; Kallioniemi O P, Ann Med 2001, 33:142-147; Blohm and Guiseppi-Elie, Curr Opin Biotechnol 2001, 12:41-47). Tissues having a disease or disorder implicating defective p53 signaling that express a LIMK, are identified as amenable to treatment with a LIMK modulating agent. In a preferred application, the p53 defective tissue overexpresses a LIMK relative to normal tissue. For example, a Northern blot analysis of mRNA from tumor and normal cell lines, or from tumor and matching normal tissue samples from the same patient, using full or partial LIMK cDNA sequences as probes, can determine whether particular tumors express or overexpress LIMK. Alternatively, the TaqMane is used for quantitative RT-PCR analysis of LIMK expression in cell lines, normal tissues and tumor samples (PE Applied Biosystems).

[0103] Various other diagnostic methods may be performed, for example, utilizing reagents such as the LIMK oligonucleotides, and antibodies directed against a LIMK, as described above for: (1) the detection of the presence of LIMK gene mutations, or the detection of either over- or under-expression of LIMK mRNA relative to the non-disorder state; (2) the detection of either an over- or an under-abundance of LIMK gene product relative to the non-disorder state; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by LIMK.

[0104] Thus, in a specific embodiment, the invention is drawn to a method for diagnosing a disease in a patient, the method comprising: a) obtaining a biological sample from the patient; b) contacting the sample with a probe for LIMK expression; c) comparing results from step (b) with a control; and d) determining whether step (c) indicates a likelihood of disease. Preferably, the disease is cancer, most preferably a cancer as shown in TABLE 1. The probe may be either DNA or protein, including an antibody.

EXAMPLES

[0105] The following experimental section and examples are offered by way of illustration and not by way of limitation.

[0106] I. *Drosophila* p53 Screen

[0107] The *Drosophila* p53 gene was overexpressed specifically in the wing using the vestigial margin quadrant enhancer. Increasing quantities of *Drosophila* p53 (titrated using different strength transgenic inserts in 1 or 2 copies) caused deterioration of normal wing morphology from mild to strong, with phenotypes including disruption of pattern and polarity of wing hairs, shortening and thickening of wing veins, progressive crumpling of the wing and appearance of dark "death" inclusions in wing blade. In a screen designed to identify enhancers and suppressors of *Drosophila* p53, homozygous females carrying two copies of p53 were crossed to 5663 males carrying random insertions of a piggyBac transposon (Fraser M et al., Virology (1985) 145:356-361). Progeny containing insertions were com-

pared to non-insertion-bearing sibling progeny for enhancement or suppression of the p53 phenotypes. Sequence information surrounding the piggyBac insertion site was used to identify the modifier genes. Modifiers of the wing phenotype were identified as members of the p53 pathway. CG6027/cdi/tsk was an enhancer of the wing phenotype. Human orthologs of the modifiers are referred to herein as LIMK.

[0108] BLAST analysis (Altschul et al., supra) was employed to identify Targets from *Drosophila* modifiers. For example, representative sequences from LIMK GI#5454110 (SEQ ID NO:24) and GI#6005896 (SEQ ID NO:25) share 59% and 54% amino acid identity, respectively, with the *Drosophila*. CG6027/cdi/tsk.

[0109] Various domains, signals, and functional subunits in proteins were analyzed using the PSORT (Nakai K., and Horton P., Trends Biochem Sci, 1999, 24:34-6; Kenta Nakai, Protein sorting signals and prediction of subcellular localization, Adv. Protein Chem. 54, 277-344 (2000)), PFAM (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2; http://pfam.wustl.edu), SMART (Ponting C P, et al., SMART: identification and annotation of domains from signaling and extracellular protein sequences. Nucleic Acids Res. 1999 Jan. 1;27(1):229-32), TM-HMM (Erik L. L. Sonnhammer, Gunnar von Heijne, and Anders Krogh: A hidden Markov model for predicting transmembrane helices in protein sequences. In Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182 Ed J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen Menlo Park, Calif.: AAAI Press, 1998), and clust (Remm M, and Sonnhammer E. Classification of transmembrane protein families in the *Caenorhabditis elegans* genome and identification of human orthologs. Genome Res. 2000 November;10(11):1679-89) programs. For example, the LIM domains of LIMK from GI# 4505001 (SEQ ID NO:20) are located at approximately amino acid residues 80 to 136 and 139 to 198 (PFAM 00412). The LIM domains of LIMK from GI# 5031869 (SEQ ID NO:22) are located at approximately amino acid residues 12 to 69 and 72 to 129 (PFAM 00412). Further, the kinase domains of LIMK from GI#s 4505001, 5031869, 5454110 and 6005896 (SEQ ID NOs: 20, 22, 24, and 25, respectively) are located at approximately amino acid residues 394 to 652 for GI# 4505001, 310-518 for GI#5031869, 57 to 314 for GI#5454110, and 62 to 117 and 120 to 289 for GI#6005896, respectively (PFAM00069). Still further, the PDZ domains of LIMK from GI#4505001 (SEQ ID NO: 20) and GI#5031869 (SEQ ID NO: 22) are located at approximately amino acids 165-257 and 131-217, respectively (PFAM 00595).

[0110] II. High-Throughput In Vitro Fluorescence Polarization Assay

[0111] Fluorescently-labeled LIMK peptide/substrate are added to each well of a 96-well microtiter plate, along with a test agent in a test buffer (10 mM HEPES, 10 mM NaCl, 6 mM magnesium chloride, pH 7.6). Changes in fluorescence polarization, determined by using a Fluorolite FPM-2 Fluorescence Polarization Microtiter System (Dynatech Laboratories, Inc), relative to control values indicates the test compound is a candidate modifier of LIMK activity.

[0112] III. High-Throughput In Vitro Binding Assay.

[0113] ^{33}P-labeled LIMK peptide is added in an assay buffer (100 mM KCl, 20 mM HEPES pH 7.6, 1 mM $MgCl_2$, 1% glycerol, 0.5% NP-40, 50 mM beta-mercaptoethanol, 1 mg/ml BSA, cocktail of protease inhibitors) along with a test agent to the wells of a Neutralite-avidin coated assay plate and incubated at 25° C. for 1 hour. Biotinylated substrate is then added to each well and incubated for 1 hour. Reactions are stopped by washing with PBS, and counted in a scintillation counter. Test agents that cause a difference in activity relative to control without test agent are identified as candidate p53 modulating agents.

[0114] IV. Immunoprecipitations and Immunoblotting

[0115] For coprecipitation of transfected proteins, 3×10^6 appropriate recombinant cells containing the LIMK proteins are plated on 10-cm dishes and transfected on the following day with expression constructs. The total amount of DNA is kept constant in each transfection by adding empty vector. After 24 h, cells are collected, washed once with phosphate-buffered saline and lysed for 20 min on ice in 1 ml of lysis buffer containing 50 mM Hepes, pH 7.9, 250 mM NaCl, 20 mM-glycerophosphate, 1 mM sodium orthovanadate, 5 mM p-nitrophenyl phosphate, 2 mM dithiothreitol, protease inhibitors (complete, Roche Molecular Biochemicals), and 1% Nonidet P-40. Cellular debris is removed by centrifugation twice at 15,000×g for 15 min. The cell lysate is incubated with 25 μl of M2 beads (Sigma) for 2 h at 4° C. with gentle rocking.

[0116] After extensive washing with lysis buffer, proteins bound to the beads are solubilized by boiling in SDS sample buffer, fractionated by SDS-polyacrylamide gel electrophoresis, transferred to polyvinylidene difluoride membrane and blotted with the indicated antibodies. The reactive bands are visualized with horseradish peroxidase coupled to the appropriate secondary antibodies and the enhanced chemiluminescence (ECL) Western blotting detection system (Amersham Pharmacia Biotech).

[0117] V. Kinase Assay

[0118] A purified or partially purified LIMK is diluted in a suitable reaction buffer, e.g., 50 mM Hepes, pH 7.5, containing magnesium chloride or manganese chloride (1-20 mM) and a peptide or polypeptide substrate, such as myelin basic protein or casein (1-10 μg/ml). The final concentration of the kinase is 1-20 nM. The enzyme reaction is conducted in microtiter plates to facilitate optimization of reaction conditions by increasing assay throughput. A 96-well microtiter plate is employed using a final volume 30-100 μl. The reaction is initiated by the addition of ^{33}P-gamma-ATP (0.5 μCi/ml) and incubated for 0.5 to 3 hours at room temperature. Negative controls are provided by the addition of EDTA, which chelates the divalent cation ($Mg2^+$ or Mn^{2+}) required for enzymatic activity. Following the incubation, the enzyme reaction is quenched using EDTA. Samples of the reaction are transferred to a 96-well glass fiber filter plate (MultiScreen, Millipore). The filters are subsequently washed with phosphate-buffered saline, dilute phosphoric acid (0.5%) or other suitable medium to remove excess radiolabeled ATP. Scintillation cocktail is added to the filter plate and the incorporated radioactivity is quantitated by scintillation counting (Wallac/Perkin Elmer). Activity is defined by the amount of radioactivity detected following subtraction of the negative control reaction value (EDTA quench).

[0119] VI. Expression Analysis

[0120] All cell lines used in the following experiments are NCI (National Cancer Institute) lines, and are available from ATCC (American Type Culture Collection, Manassas, Va. 20110-2209). Normal and tumor tissues were obtained from Impath, UC Davis, Clontech, Stratagene, and Ambion.

[0121] TaqMan analysis was used to assess expression levels of the disclosed genes in various samples.

[0122] RNA was extracted from each tissue sample using Qiagen (Valencia, Calif.) RNeasy kits, following manufacturer's protocols, to a final concentration of 50 ng/μl. Single stranded cDNA was then synthesized by reverse transcribing the RNA samples using random hexamers and 500 ng of total RNA per reaction, following protocol 4304965 of Applied Biosystems (Foster City, Calif., http://www.appliedbiosystems.com/).

[0123] Primers for expression analysis using TaqMan assay (Applied Biosystems, Foster City, Calif.) were prepared according to the TaqMan protocols, and the following criteria: a) primer pairs were designed to span introns to eliminate genomic contamination, and b) each primer pair produced only one product.

[0124] Taqman reactions were carried out following manufacturer's protocols, in 25 μl total volume for 96-well plates and 10 μl total volume for 384-well plates, using 300 nM primer and 250 nM probe, and approximately 25 ng of cDNA. The standard curve for result analysis was prepared using a universal pool of human cDNA samples, which is a mixture of cDNAs from a wide variety of tissues so that the chance that a target will be present in appreciable amounts is good. The raw data were normalized using 18S rRNA (universally expressed in all tissues and cells).

[0125] For each expression analysis, tumor tissue samples were compared with matched normal tissues from the same patient. A gene was considered overexpressed in a tumor when the level of expression of the gene was 2 fold or higher in the tumor compared with its matched normal sample.

[0126] Results are shown in Table 1. Data presented in bold indicate that greater than 50% of tested tumor samples of the tissue type indicated in row 1 exhibited over expression of the gene listed in column 1, relative to normal samples. Underlined data indicates that between 25% to 49% of tested tumor samples exhibited over expression. A modulator identified by an assay described herein can be further validated for therapeutic effect by administration to a tumor in which the gene is overexpressed. A decrease in tumor growth confirms therapeutic utility of the modulator. Prior to treating a patient with the modulator, the likelihood that the patient will respond to treatment can be diagnosed by obtaining a tumor sample from the patient, and assaying for expression of the gene targeted by the modulator. The expression data for the gene(s) can also be used as a diagnostic marker for disease progression. The assay can be performed by expression analysis as described above, by antibody directed to the gene target, or by any other available detection method.

TABLE 1

	breast			colon			lung			ovary	
GI#1805594 (SEQ ID NO: 14)	0	11	.	3	30	.	0	13	.	1	7
GI#6005895 (SEQ ID NO: 15)	3	11	.	1	30	.	2	13	.	2	7

[0127]

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 25

<210> SEQ ID NO 1
<211> LENGTH: 3332
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

gcgccgagcc ggtttccccg ccggtgtccg agaggcgccc ccggcccggc ccggcccggc 60

ccgcgccctc cgccccccgcc tccccgggcc ggcggcggtg ggcgagctcg cgggcccggc 120

cgcccccagc cccagccccg ccgggccccg ccccccgtcg agtgcatgag gttgacgcta 180

ctttgttgca cctggaggga agaacgtatg ggagaggaag gaagcgagtt gcccgtgtgt 240

gcaagctgcg gccagaggat ctatgatggc cagtacctcc aggccctgaa cgcggactgg 300

cacgcagact gcttcaggtg ttgtgactgc agtgcctccc tgtcgcacca gtactatgag 360

aaggatgggc agctcttctg caagaaggac tactgggccc gctatggcga gtcctgccat 420

gggtgctctg agcaaatcac caagggactg gttatggtgg ctggggagct gaagtaccac 480

cccgagtgtt tcatctgcct cacgtgtggg acctttatcg gtgacgggga cacctacacg 540

ctggtggagc actccaagct gtactgcggg cactgctact accagactgt ggtgaccccc 600

-continued

gtcatcgagc agatcctgcc tgactcccct ggctcccacc tgccccacac cgtcaccctg 660

gtgtccatcc cagcctcatc tcatggcaag cgtggacttt cagtctccat tgacccccgg 720

cacggcccac cgggctgtgg caccgagcac tcacacaccg tccgcgtcca gggagtggat 780

ccgggctgca tgagcccaga tgtgaagaat tccatccacg tcggagaccg gatcttggaa 840

atcaatggca cgcccatccg aaatgtgccc ctggacgaga ttgacctgct gattcaggaa 900

accagccgcc tgctccagct gaccctcgag catgaccctc acgatacact gggccacggg 960

ctggggcctg agaccagccc cctgagctct ccggcttata ctcccagcgg ggaggcgggc 1020

agctctgccc ggcagaaacc tgtcttgagg agctgcagca tcgacaggtc tccgggcgct 1080

ggctcactgg gctccccggc ctcccagcgc aaggacctgg gtcgctctga gtccctccgc 1140

gtagtctgcc ggccacaccg catcttccgg ccgtcggacc tcatccacgg ggaggtgctg 1200

ggcaagggct gcttcggcca ggctatcaag gtgacacacc gtgagacagg tgaggtgatg 1260

gtgatgaagg agctgatccg gttcgacgag gagacccaga ggacgttcct caaggaggtg 1320

aaggtcatgc gatgcctgga acaccccaac gtgctcaagt tcatcggggt gctctacaag 1380

gacaagaggc tcaacttcat cactgagtac atcaagggcg gcacgctccg gggcatcatc 1440

aagagcatgg acagccagta cccatggagc cagagagtga gctttgccaa ggacatcgca 1500

tcagggatgg cctacctcca ctccatgaac atcatccacc gagacctcaa ctcccacaac 1560

tgcctggtcc gcgagaacaa gaatgtggtg gtggctgact tcgggctggc gcgtctcatg 1620

gtggacgaga agactcagcc tgagggcctg cggagcctca agaagccaga ccgcaagaag 1680

cgctacaccg tggtgggcaa cccctactgg atggcacctg agatgatcaa cggccgcagc 1740

tatgatgaga aggtggatgt gttctccttt gggatcgtcc tgtgcgagat catcgggcgg 1800

gtgaacgcag accctgacta cctgccccgc accatggact ttggcctcaa cgtgcgagga 1860

ttcctggacc gctactgccc cccaaactgc cccccgagct tcttccccat caccgtgcgc 1920

tgttgcgatc tggaccccga gaagaggcca tcctttgtga agctggaaca ctggctggag 1980

accctccgca tgcacctggc cggccacctg ccactgggcc cacagctgga gcagctggac 2040

agaggtttct gggagaccta ccggcgcggc gagagcggac tgcctgccca ccctgaggtc 2100

cccgactgag ccagggccac tcagctgccc ctgtccccac ctctggagaa tccacccccа 2160

ccagattcct ccgcgggagg tggccctcag ctgggacagt ggggacccag gcttctcctc 2220

agagccaggc cctgacttgc cttctcccac cccgtggacc gcttcccctg ccttctctct 2280

gccgtggccc agagccggcc cagctgcaca cacacaccat gctctcgccc tgctgtaacc 2340

tctgtcttgg cagggctgtc ccctcttgct tctccttgca tgagctggag ggcctgtgtg 2400

agttacgccc ctttccacac gccgctgccc cagcaaccct gttcacgctc cacctgtctg 2460

gtccatagct ccctggaggc tgggccagga ggcagcctcc gaaccatgcc ccatataacg 2520

cttgggtgcg tgggagggcg cacatcaggg cagaggccaa gttccaggtg tctgtgttcc 2580

caggaaccaa atggggagtc tggggcccgt tttcccccca gggggtgtct aggtagcaac 2640

aggtatcgag gactctccaa acccccaaag cagagagagg gctgatccca tggggcggag 2700

gtccccagtg gctgagcaaa cagccccttc tctcgctttg ggtctttttt ttgtttcttt 2760

cttaaagcca ctttagtgag aagcaggtac caagcctcag ggtgaagggg gtcccttgag 2820

ggagcgtgga gctgcggtgc cctggccggc gatggggagg agccggctcc ggcagtgaga 2880

-continued

ggataggcac agtggaccgg gcaggtgtcc accagcagct cagcccctgc agtcatctca 2940

gagccccttc ccgggcctct cccccaaggc tccctgcccc tcctcatgcc cctctgtcct 3000

ctgcgttttt tctgtgtaat ctatttttta agaagagttt gtattatttt ttcatacggc 3060

tgcagcagca gctgccaggg gcttgggatt ttatttttgt ggcgggcggg ggtgggaggg 3120

ccattttgtc actttgcctc agttgagcat ctaggaagta ttaaaactgt gaagctttct 3180

cagtgcactt tgaacctgga aaacaatccc aacaggcccg tgggaccatg acttagggag 3240

gtgggaccca cccacccca tccaggaacc gtgacgtcca aggaaccaaa cccagacgca 3300

gaacaataaa ataaattccg tactccccac cc 3332

<210> SEQ ID NO 2
<211> LENGTH: 3271
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

gcgccgagcc ggtttccccg ccggtgtccg agaggcgccc ccggcccggc ccggcccggc 60

ccgcgccctc cgccccgcc tccccgggcc ggcggcggtg ggcgagctcg cgggcccggc 120

cgcccccagc cccagccccg ccgggccccg ccccccgtcg agtgcatgag gttgacgcta 180

ctttgttgca cctggaggga agaacgtatg ggagaggaag gaagcgagtt gcccgtgtgt 240

gcaagctgcg gccagaggat ctatgatggc cagtacctcc aggccctgaa cgcggactgg 300

cacgcagact gcttcaggtg ttgtgactgc agtgcctccc tgtcgcacca gtactatgag 360

aaggatgggc agctcttctg caagaaggac tactgggccc gctatggcga gtcctgccat 420

gggtgctctg agcaaatcac caagggactg gttatggtgg ctggggagct gaagtaccac 480

cccgagtgtt tcatctgcct cacgtgtggg acctttatcg gtgacgggga cacctacacg 540

ctggtggagc actccaagct gtactgcggg cactgctact accagactgt ggtgaccccc 600

gtcatcgagc agatcctgcc tgactccct ggctcccacc tgccccacac cgtcaccctg 660

gtgtccatcc cagcctcatc tcatggcaag cgtggacttt cagtctccat tgaccccccg 720

cacggcccac cgggctgtgg caccgagcac tcacacaccg tccgcgtcca gggagtggat 780

ccgggctgca tgagcccaga tgtgaagaat tccatccacg tcggagaccg gatcttggaa 840

atcaatggca cgcccatccg aaatgtgccc ctggacgaga ttgacctgct gattcaggaa 900

accagccgcc tgctccagct gaccctcgag catgaccctc acgatacact gggccacggg 960

ctggggcctg agaccagccc cctgagctct ccggcttata ctcccagcgg ggaggcgggc 1020

agctctgccc ggcagaaacc tgtcttcgca aggacctggg tcgctctgag tccctccgcg 1080

tagtctgccg gccacaccgc atcttccggc cgtcggacct catccacggg gaggtgctgg 1140

gcaagggctg cttcggccag gctatcaagg tgacacaccg tgagacaggt gaggtgatgg 1200

tgatgaagga gctgatccgg ttcgacgagg agacccagag gacgttcctc aaggaggtga 1260

aggtcatgcg atgcctggaa caccccaacg tgctcaagtt catcggggtg ctctacaagg 1320

acaagaggct caacttcatc actgagtaca tcaagggcgg cacgctccgg ggcatcatca 1380

agagcatgga cagccagtac ccatggagcc agagagtgag ctttgccaag gacatcgcat 1440

cagggatggc ctacctccac tccatgaaca tcatccaccg agacctcaac tcccacaact 1500

gcctggtccg cgagaacaag aatgtggtgg tggctgactt cgggctggcg cgtctcatgg 1560

-continued

tggacgagaa gactcagcct gagggcctgc ggagcctcaa gaagccagac cgcaagaagc 1620

gctacaccgt ggtgggcaac ccctactgga tggcacctga gatgatcaac ggccgcagct 1680

atgatgagaa ggtggatgtg ttctcctttg ggatcgtcct gtgcgagatc atcgggcggg 1740

tgaacgcaga ccctgactac ctgccccgca ccatggactt tggcctcaac gtgcgaggat 1800

tcctggaccg ctactgcccc ccaaactgcc ccccgagctt cttccccatc accgtgcgct 1860

gttgcgatct ggaccccgag aagaggccat cctttgtgaa gctggaacac tggctggaga 1920

ccctccgcat gcacctggcc ggccacctgc cactgggccc acagctggag cagctggaca 1980

gaggtttctg ggagacctac cggcgcggcg agagcggact gcctgcccac cctgaggtcc 2040

ccgactgagc cagggccact cagctgcccc tgtccccacc tctggagaat ccacccccac 2100

cagattcctc cgcgggaggt ggccctcagc tgggacagtg gggacccagg cttctcctca 2160

gagccaggcc ctgacttgcc ttctcccacc ccgtggaccg cttcccctgc cttctctctg 2220

ccgtggccca gagccggccc agctgcacac acacaccatg ctctcgccct gctgtaacct 2280

ctgtcttggc agggctgtcc cctcttgctt ctccttgcat gagctggagg gcctgtgtga 2340

gttacgcccc tttccacacg ccgctgcccc agcaaccctg ttcacgctcc acctgtctgg 2400

tccatagctc cctggaggct gggccaggag gcagcctccg aaccatgccc catataacgc 2460

ttgggtgcgt gggagggcgc acatcagggc agaggccaag ttccaggtgt ctgtgttccc 2520

aggaaccaaa tggggagtct ggggcccgtt ttccccccag ggggtgtcta ggtagcaaca 2580

ggtatcgagg actctccaaa cccccaaagc agagagaggg ctgatcccat ggggcggagg 2640

tccccagtgg ctgagcaaac agccccttct ctcgctttgg gtctttttt tgtttctttc 2700

ttaaagccac tttagtgaga agcaggtacc aagcctcagg gtgaaggggg tcccttgagg 2760

gagcgtggag ctgcggtgcc ctggccggcg atggggagga gccggctccg gcagtgagag 2820

gataggcaca gtggaccggg caggtgtcca ccagcagctc agcccctgca gtcatctcag 2880

agccccttcc cgggcctctc ccccaaggct ccctgccccc cctcatgccc ctctgtcctc 2940

tgcgtttttt ctgtgtaatc tatttttaa gaagagtttg tattattttt tcatacggct 3000

gcagcagcag ctgccagggg cttgggattt tatttttgtg gcgggcgggg gtgggagggc 3060

cattttgtca ctttgcctca gttgagcatc taggaagtat taaaactgtg aagctttctc 3120

agtgcacttt gaacctggaa aacaatccca acaggcccgt gggaccatga cttagggagg 3180

tgggacccac ccacccccat ccaggaaccg tgacgtccaa ggaaccaaac ccagacgcag 3240

aacaataaaa taaattccgt actccccacc c 3271

<210> SEQ ID NO 3
<211> LENGTH: 3332
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

gcgccgagcc ggtttccccg ccggtgtccg agaggcgccc ccggcccggc ccggcccggc 60

ccgcgccctc cgccccccgcc tccccgggcc ggcggcggtg ggcgagctcg cgggcccggc 120

cgcccccagc cccagccccg ccgggccccg ccccccgtcg agtgcatgag gttgacgcta 180

ctttgttgca cctggaggga agaacgtatg ggagaggaag gaagcgagtt gcccgtgtgt 240

gcaagctgcg gccagaggat ctatgatggc cagtacctcc aggccctgaa cgcggactgg 300

-continued

```
cacgcagact gcttcaggtg ttgtgactgc agtgcctccc tgtcgcacca gtactatgag    360

aaggatgggc agctcttctg caagaaggac tactgggccc gctatggcga gtcctgccat    420

gggtgctctg agcaaatcac caagggactg gttatggtgg ctggggagct gaagtaccac    480

cccgagtgtt tcatctgcct cacgtgtggg acctttatcg gtgacgggga cacctacacg    540

ctggtggagc actccaagct gtactgcggg cactgctact accagactgt ggtgaccccc    600

gtcatcgagc agatcctgcc tgactccct ggctcccacc tgccccacac cgtcaccctg    660

gtgtccatcc cagcctcatc tcatggcaag cgtggacttt cagtctccat tgaccccccg    720

cacggcccac cgggctgtgg caccgagcac tcacacaccg tccgcgtcca gggagtggat    780

ccgggctgca tgagcccaga tgtgaagaat tccatccacg tcggagaccg gatcttggaa    840

atcaatggca cgcccatccg aaatgtgccc ctggacgaga ttgacctgct gattcaggaa    900

accagccgcc tgctccagct gaccctcgag catgaccctc acgatacact gggccacggg    960

ctggggcctg agaccagccc cctgagctct ccggcttata ctcccagcgg ggaggcgggc   1020

agctctgccc ggcagaaacc tgtcttgagg agctgcagca tcgacaggtc tccgggcgct   1080

ggctcactgg gctccccggc ctcccagcgc aaggacctgg gtcgctctga gtccctccgc   1140

gtagtctgcc ggccacaccg catcttccgg ccgtcggacc tcatccacgg ggaggtgctg   1200

ggcaagggct gcttcggcca ggctatcaag gtgacacacc gtgagacagg tgaggtgatg   1260

gtgatgaagg agctgatccg gttcgacgag gagacccaga ggacgttcct caaggaggtg   1320

aaggtcatgc gatgcctgga acaccccaac gtgctcaagt tcatcggggt gctctacaag   1380

gacaagaggc tcaacttcat cactgagtac atcaagggcg gcacgctccg gggcatcatc   1440

aagagcatgg acagccagta cccatggagc cagagagtga gctttgccaa ggacatcgca   1500

tcagggatgg cctacctcca ctccatgaac atcatccacc gagacctcaa ctcccacaac   1560

tgcctggtcc gcgagaacaa gaatgtggtg gtggctgact tcgggctggc gcgtctcatg   1620

gtggacgaga agactcagcc tgagggcctg cggagcctca agaagccaga ccgcaagaag   1680

cgctacaccg tggtgggcaa cccctactgg atggcacctg agatgatcaa cggccgcagc   1740

tatgatgaga aggtggatgt gttctccttt gggatcgtcc tgtgcgagat catcgggcgg   1800

gtgaacgcag accctgacta cctgccccgc accatggact ttggcctcaa cgtgcgagga   1860

ttcctggacc gctactgccc cccaaactgc cccccgagct tcttccccat caccgtgcgc   1920

tgttgcgatc tggaccccga gaagaggcca tcctttgtga agctggaaca ctggctggag   1980

accctccgca tgcacctggc cggccacctg ccactgggcc cacagctgga gcagctggac   2040

agaggtttct gggagaccta ccggcgcggc gagagcggac tgcctgccca ccctgaggtc   2100

cccgactgag ccagggccac tcagctgccc ctgtccccac ctctggagaa tccaccccca   2160

ccagattcct ccgcgggagg tggccctcag ctgggacagt ggggacccag gcttctcctc   2220

agagccaggc cctgacttgc cttctcccac cccgtggacc gcttcccctg ccttctctct   2280

gccgtggccc agagccggcc cagctgcaca cacacaccat gctctcgccc tgctgtaacc   2340

tctgtcttgg cagggctgtc ccctcttgct tctccttgca tgagctggag ggcctgtgtg   2400

agttacgccc ctttccacac gccgctgccc cagcaaccct gttcacgctc cacctgtctg   2460

gtccatagct ccctggaggc tgggccagga ggcagcctcc gaaccatgcc ccatataacg   2520

cttgggtgcg tgggagggcg cacatcaggg cagaggccaa gttccaggtg tctgtgttcc   2580
```

-continued

caggaaccaa atggggagtc tggggcccgt tttcccccca gggggtgtct aggtagcaac 2640

aggtatcgag gactctccaa acccccaaag cagagagagg gctgatccca tggggcggag 2700

gtccccagtg gctgagcaaa cagccccttc tctcgctttg ggtctttttt ttgtttcttt 2760

cttaaagcca ctttagtgag aagcaggtac caagcctcag ggtgaagggg gtcccttgag 2820

ggagcgtgga gctgcggtgc cctggccggc gatggggagg agccggctcc ggcagtgaga 2880

ggataggcac agtggaccgg gcaggtgtcc accagcagct cagcccctgc agtcatctca 2940

gagccccttc ccgggcctct cccccaaggc tccctgcccc tcctcatgcc cctctgtcct 3000

ctgcgttttt tctgtgtaat ctatttttta agaagagttt gtattatttt ttcatacggc 3060

tgcagcagca gctgccaggg gcttgggatt ttatttttgt ggcgggcggg ggtgggaggg 3120

ccattttgtc actttgcctc agttgagcat ctaggaagta ttaaaactgt gaagctttct 3180

cagtgcactt tgaacctgga aaacaatccc aacaggcccg tgggaccatg acttagggag 3240

gtgggaccca cccaccccca tccaggaacc gtgacgtcca aggaaccaaa cccagacgca 3300

gaacaataaa ataaattccg tactccccac cc 3332

<210> SEQ ID NO 4
<211> LENGTH: 3262
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

gcgccgagcc ggtttccccg ccggtgtccg agaggcgccc ccggcccggc cgcccccagc 60

cccagccccg ccgggccccg ccccccgtcg agtgcatgag gttgacgcta ctttgttgca 120

cctggaggga agaacgtatg ggagaggaag gaagcgagtt gcccgtgtgt gcaagctgcg 180

gccagaggat ctatgatggc cagtacctcc aggccctgaa cgcggactgg cacgcagact 240

gcttcaggtg ttgtgactgc agtgcctccc tgtcgcacca gtactatgag aaggatgggc 300

agctcttctg caagaaggac tactgggccc gctatggcga gtcctgccat gggtgctctg 360

agcaaatcac caagggactg gttatggtgg ctggggagct gaagtaccac cccgagtgtt 420

tcatctgcct cacgtgtggg acctttatcg gtgacgggga cacctacacg ctggtggagc 480

actccaagct gtactgcggg cactgctact accagactgt ggtgaccccc gtcatcgagc 540

agatcctgcc tgactcccct ggctcccacc tgccccacac cgtcaccctg gtgtccatcc 600

cagcctcatc tcatggcaag cgtggacttt cagtctccat tgacccccgg cacggcccac 660

cgggctgtgg caccgagcac tcacacaccg tccgcgtcca gggagtggat ccgggctgca 720

tgagcccaga tgtgaagaat tccatccacg tcggagaccg gatcttggaa atcaatggca 780

cgcccatccg aaatgtgccc ctggacgaga ttgacctgct gattcaggaa accagccgcc 840

tgctccagct gaccctcgag catgaccctc acgatacact gggccacggg ctggggcctg 900

agaccagccc cctgagctct ccggcttata ctcccagcgg ggaggcgggc agctctgccc 960

ggcagaaacc tgtcttgagg agctgcagca tcgacaggtc tccgggcgct ggctcactgg 1020

gctccccggc ctcccagcgc aaggacctgg gtcgctctga gtccctccgc gtagtctgcc 1080

ggccacaccg catcttccgg ccgtcggacc tcatccacgg ggaggtgctg ggcaagggct 1140

gcttcggcca ggctatcaag gtgacacacc gtgagacagg tgaggtgatg gtgatgaagg 1200

agctgatccg gttcgacgag gagacccaga ggacgttcct caaggaggtg aaggtcatgc 1260

-continued

gatgcctgga acaccccaac gtgctcaagt tcatcggggt gctctacaag gacaagaggc 1320

tcaacttcat cactgagtac atcaagggcg gcacgctccg gggcatcatc aagagcatgg 1380

acagccagta cccatggagc cagagagtga gctttgccaa ggacatcgca tcagggatgg 1440

cctacctcca ctccatgaac atcatccacc gagacctcaa ctcccacaac tgcctggtcc 1500

gcgagaacaa gaatgtggtg gtggctgact tcgggctggc gcgtctcatg gtggacgaga 1560

agactcagcc tgagggcctg cggagcctca agaagccaga ccgcaagaag cgctacaccg 1620

tggtgggcaa cccctactgg atggcacctg agatgatcaa cggccgcagc tatgatgaga 1680

aggtggatgt gttctccttt gggatcgtcc tgtgcgagat catcgggcgg gtgaacgcag 1740

accctgacta cctgccccgc accatggact ttggcctcaa cgtgcgagga ttcctggacc 1800

gctactgccc cccaaactgc cccccgagct tcttccccat caccgtgcgc tgttgcgatc 1860

tggaccccga gaagaggcca tcctttgtga agctggaaca ctggctggag accctccgca 1920

tgcacctggc cggccacctg ccactgggcc cacagctgga gcagctggac agaggtttct 1980

gggagaccta ccggcgcggc gagagcggac tgcctgccca ccctgaggtc cccgactgag 2040

ccagggccac tcagctgccc ctgtccccac ctctggagaa tccaccccca ccagattcct 2100

ccgcgggagg tggccctcag ctgggacagt ggggacccag gcttctcctc agagccaggc 2160

cctgacttgc cttctcccac cccgtggacc gcttcccctg ccttctctct gccgtggccc 2220

agagccggcc cagctgcaca cacacaccat gctctcgccc tgctgtaacc tctgtcttgg 2280

cagggctgtc ccctcttgct tctccttgca tgagctggag ggcctgtgtg agttacgccc 2340

ctttccacac gccgctgccc cagcaaccct gttcacgctc cacctgtctg gtccatagct 2400

ccctggaggc tgggccagga ggcagcctcc gaaccatgcc ccatataacg cttgggtgcg 2460

tgggagggcg cacatcaggg cagaggccaa gttccaggtg tctgtgttcc caggaaccaa 2520

atggggagtc tggggcccgt tttcccccca gggggtgtct aggtagcaac aggtatcgag 2580

gactctccaa acccccaaag cagagagagg gctgatccca tggggcggag gtccccagtg 2640

gctgagcaaa cagccccttc tctcgctttg ggtctttttt ttgtttcttt cttaaagcca 2700

ctttagtgag aagcaggtac caagcctcag ggtgaagggg gtcccttgag ggagcgtgga 2760

gctgcggtgc cctggccggc gatggggagg agccggctcc ggcagtgaga ggataggcac 2820

agtggaccgg gcaggtgtcc accagcagct cagcccctgc agtcatctca gagccccttc 2880

ccgggcctct cccccaaggc tccctgcccc tcctcatgcc cctctgtcct ctgcgttttt 2940

tctgtgtaat ctatttttta agaagagttt gtattatttt ttcatacggc tgcagcagca 3000

gctgccaggg gcttgggatt ttatttttgt ggcgggcggg ggtgggaggg ccattttgtc 3060

actttgcctc agttgagcat ctaggaagta ttaaaactgt gaagctttct cagtgcactt 3120

tgaacctgga aaacaatccc aacaggcccg tgggaccatg acttagggag gtgggaccca 3180

cccaccccca tccaggaacc gtgacgtcca aggaaccaaa cccagacgca gaacaataaa 3240

ataaattccg tactccccac cc 3262

<210> SEQ ID NO 5
<211> LENGTH: 2456
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

-continued

```
cagctgaccc tcgagcatga ccctcacgat acactgggcc acgggctggg gcctgagacc     60

agccccctga gctctccggc ttatactccc agcggggagg cgggcagctc tgcccggcag    120

aaacctgtct tgaggagctg cagcatcgac aggtctccgg gcgctggctc actgggctcc    180

ccggcctccc agcgcaagga cctgggtcgc tctgagtccc tccgcgtagt ctgccggcca    240

caccgcatct tccggccgtc ggacctcatc cacggggagg tgctgggcaa gggctgcttc    300

ggccaggcta tcaaggtgac acaccgtgag acaggtgagg tgatggtgat gaaggagctg    360

atccggttcg acgaggagac ccagaggacg ttcctcaagg aggtgaaggt catgcgatgc    420

ctggaacacc ccaacgtgct caagttcatc ggggtgctct acaaggacaa gaggctcaac    480

ttcatcactg agtacatcaa gggcggcacg ctccggggca tcatcaagag catggacagc    540

cagtacccat ggagccagag agtgagcttt gccaaggaca tcgcatcagg gatggcctac    600

ctccactcca tgaacatcat ccaccgagac ctcaactccc acaactgcct ggtccgcgag    660

aacaagaatg tggtggtggc tgacttcggg ctggcgcgtc tcatggtgga cgagaagact    720

cagcctgagg gcctgcggag cctcaagaag ccagaccgca agaagcgcta caccgtggtg    780

ggcaacccct actggatggc acctgagatg atcaacggcc gcagctatga tgagaaggtg    840

gatgtgttct cctttgggat cgtcctgtgc gagatcatcg ggcgggtgaa cgcagaccct    900

gactacctgc cccgcaccat ggactttggc ctcaacgtgc gaggattcct ggaccgctac    960

tgccccccaa actgcccccc gagcttcttc cccatcaccg tgcgctgttg cgatctggac   1020

cccgagaaga ggccatcctt tgtgaagctg gaacactggc tggagaccct ccgcatgcac   1080

ctggccggcc acctgccact gggcccacag ctggagcagc tggacagagg tttctgggag   1140

acctaccggc gcggcgagag cggactgcct gcccaccctg aggtccccga ctgagccagg   1200

gccactcagc tgcccctgtc cccacctctg gagaatccac ccccaccaga ttcctccgcg   1260

ggaggtggcc ctcagctggg acagtgggga cccaggcttc tcctcagagc caggccctga   1320

cttgccttct cccaccccgt ggaccgcttc ccctgccttc tctctgccgt ggcccagagc   1380

cggcccagct gcacacacac accatgctct cgccctgctg taacctctgt cttggcaggg   1440

ctgtcccctc ttgcttctcc ttgcatgagc tggagggcct gtgtgagtta cgcccctttc   1500

cacacgccgc tgccccagca accctgttca cgctccacct gtctggtcca tagctccctg   1560

gaggctgggc caggaggcag cctccgaacc atgccccata taacgcttgg gtgcgtggga   1620

gggcgcacat cagggcagag gccaagttcc aggtgtctgt gttcccagga accaaatggg   1680

gagtctgggg cccgttttcc ccccaggggg tgtctaggta gcaacaggta tcgaggactc   1740

tccaaacccc caaagcagag agagggctga tcccatgggg cggaggtccc cagtggctga   1800

gcaaacagcc ccttctctcg ctttgggtct ttttttttgtt tctttcttaa agccacttta   1860

gtgagaagca ggtaccaagc ctcagggtga agggggtccc ttgagggagc gtggagctgc   1920

ggtgccctgg ccggcgatgg ggaggagccg gctccggcag tgagaggata ggcacagtgg   1980

accgggcagg tgtccaccag cagctcagcc cctgcagtca tctcagagcc ccttcccggg   2040

cctctccccc aaggctccct gcccctcctc atgcccctct gtcctctgcg ttttttctgt   2100

gtaatctatt ttttaagaag agtttgtatt attctttcat acggctgcag cagcagctgc   2160

caggggcttg ggattttatt tttgtggcgg gcgggggtgg gagggccatt ttgtcacttt   2220

gcctcagttg agcatctagg aagtattaaa actgtgaagc tttctcagtg cactttgaac   2280
```

-continued

ctggaaaaca atcccaacag gcccgtggga ccatgactta gggaggtggg acccacccac 2340

ccccatccag gaaccgtgac gtccaaggaa ccaaacccag acgcagaaca ataaaataaa 2400

ttccgtactc cccaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaca aaaaaa 2456

<210> SEQ ID NO 6
<211> LENGTH: 128978
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

gatcctactt cgcactcaca aatatggcca taataaaaaa ggaaaaaaat caacaacatc 60

cttgttggca aggatgtgga aaaattagaa tcctcataca ttgctggtgg gatgtaaact 120

ggtacagcca ccagtacctg actacagcca taccaccacg aacatgtcca atcttgtcta 180

aactttgatg tggaaaaact gatacggaaa gtagtttgat cagccgggca cagtggctca 240

cgtctgcaat ctcagaactt tggggagctg agggggggtgg atcacctgag gttgggagat 300

gaagaccagc ctggccaaca tggcaaaacc tgtctctact aaaaatacaa atttttgtat 360

ttttagtaga gccgggtgtg gtggctcaca cctgtaatcg cagctgctcg tgaggctgag 420

gcaggagagt catttgaacc caggaggcgg agattgcagt gaaccaagat cacatcactg 480

cactccagcc tgggtgacag agcgagactc catctcaaaa aaaaaaaaaa aaagaaaaag 540

aaaagaaaat agtttgatgg ttcctcgaaa agttaaacac agaataacca gcaatgccac 600

tcctaggtat atactcaaaa gaaatgaaaa caggtatcaa acaaaaacct gtacatgaat 660

attcatagta gcattattca taatagccaa aaggtggaaa caactcaaat gtccaccaac 720

tgatgactag ataaacaaaa tatggtacat ccatacacag gaatattatt cagccataaa 780

aagaaatgct gatacatgct acaacttaga tgaaccttca aaacatcatg cttagtagaa 840

gaagccagac acaaaaggtt gcatatagta tgattccatt tacataaaat gtctggaata 900

ggcatatcca tagaaataca aagagatcag tatttacttg gaacacttgg gacaataaca 960

aataaaataa taaaaatttt aaaaaagagt ccaggcgcag tggctcacgc ctgtaatccc 1020

agcactttgg gaggccaagg tgggtggatc ataaggtcag gagatcgaga ccgtggtgaa 1080

accccgtctc tactaaaaat acaaaaaatt agccgggcac gctggcgggc gcctgtagtc 1140

ccagctactc gggaggctga ggcaggagaa tggcgtgaac ctgggaggcg gagcttgcag 1200

tgagccgaga tcatgccact gcactccagc ctggctgaca gagcgagact ccgtctcaaa 1260

ataataataa caataataat aataataata ataataatta ttattattat tatttaaaaa 1320

agaaggagat cagaggttgc cagagggtga ggagaaagga aaatgaagag tggctgccaa 1380

caggtactag gtttctccca tctgagtact aaccagactc aaccctgctt agcttctgag 1440

atcagacaag attgcacatg ttcagggtgg tatggctgta gacaggtacc agcgtttctt 1500

tttggggtga tgaagtgttc tggaattaga tagcggtgat ggttacacaa ccttgtgaat 1560

atacagacat gagccacata atgatgtttc agtcaacaga ccgcatatat gacagtggtc 1620

ctataacaca gtggtcccca acctttttgg caccagggac tggtttcatg gaagacaatt 1680

tttccatgga ctgggggtga tgggggggatg gtttcaggat gaaactgttc cacctcagat 1740

catcaggcat tagttagatt cttggaagga gtgcgcaacc tagatcccta gcatgtgcag 1800

ttcacaatag ggttcacact cctattagaa tctaatgtca ccactgatct gacaagaggc 1860

-continued

agagctcaga tggtaatgtt tgctcacctg ctgctcacct cctgctgtgc tactgggttc 1920

ctaactggcc acagccctgg ggtttgggga tccccgctat aagattataa taccatattt 1980

ttactgtacc ttttctatgt ttagatacac taatacttat cattgtgtta aagttgccta 2040

tagtatttag tacagttaga tgctgtacag gtttgtagcc tgagatcagt aggttacacc 2100

atatagccta gatatgtagt aggctgtccc gtctaggttt gtgtaattat actctatgat 2160

gtttgcatga cgcatttctc agaacctatc cttgttgtta agtgacgcac aggtactaac 2220

aactactgaa ctgtatgctg taaaatgatg aatttcatga tatgtgaatt gtatctcaat 2280

agaaaagagc atttataaag tttttgttgc taaatttatt tactttcggt aaattttctt 2340

ttaaaagatg tctatgtttt actgtcttcg tggtgctggg acataactga catctgaagc 2400

cttgaatgcc tctctttgca aactcatttg ttactactca gttaccatct cacctccaat 2460

acccaactgc ataccaaacc tctagccatt actagaacac atactctgag ctctgccata 2520

ctttagcgcc tgcacacatg ctgtttcttc tgctcagaaa cttttgtgcc ttctgtccat 2580

cccatccaag cctgtttttc aatgcccaac acaaatgtca ctggtaccct gaaacttcaa 2640

ctcctccttg ccccatctca gaaaatgtaa ctttctgtat tcttctctct gtatattcct 2700

ctaaagtggt atttcttata ctataccata atgacctggt cccatatttg tcctcttctt 2760

ccctcttcat cccagactgg gaccacaagg aggtagaact ttttctgttt ttttttttt 2820

ttgagacaga gtcttgctct gtcgccaggc tggagtgcaa tggcacgatc tcggctcact 2880

gcaacctctg cctcccgggt tcaagtgatt ctcctacctc agcctcccca gtagctggga 2940

tttcaggcgc acatcaccat gcccagctaa tttttgtatt tttagtagag acgggatttc 3000

accatgttga ccgggatggt ctcgatctct tgtcctcgtg atccgcccgt ctcagcctcc 3060

caaagtgttg ggattacggc catgagccac tgcgcccggc cggaggtaca actttcttag 3120

tcaatcttga gcaccagtgt agtgcttaga atcgggtagt gagagaagct aacatttact 3180

gaatgcctaa gaagtccagg gcacataaca gtcttcacaa ccatcctgtg aagtacgtat 3240

taccctatcc attttacaga taaagaaact tgagtttcaa gaaaggtata cctaaggtca 3300

ttcaacttgt taaatgatgg aatggaggag ttgagaccag attggacttc tgcctctttt 3360

ttttttttt ttttttttga gacaggctgg agtgcagtgg tgtgatctca gctcactgca 3420

agctctgcct cccaggttca cgccattctc ctgcctcagc ctcccaagta gctgggacta 3480

caggcgcctg ccaccacgcc cggctaattt tttctacttt tagtagagac ggggtttcac 3540

cgtgttagcc aggatggtct caatctcctg acctcgtgat ctgcccacct cggcctccca 3600

aagtgctggg attacaggcg tgagcgaccg cgcccagccg aacttctgcc tcttaaatcc 3660

agggttctcc ctgtcagtac agtgaggtgg taactagcaa aagctatgag atatgactgc 3720

ctgggtacat atcccagctc tttcacttat ctttgtggct ttacgcaaat tacttaacct 3780

ctttatgatt gtttcttcat ttgtaaaagg aagataataa cagtgcctat atatagggtt 3840

tttatgaaga ataaatgaga tagtatatat aaagcactta gaacagtatc tggcacatac 3900

taggtgctca ataaatgtta gcgatgacat ttattactgt cccacattga gctggtgttt 3960

gttaaacact ggatgactgg ataaaatgta catctttcat gcagagattg ctcgattatc 4020

catttcactg attcaggatc agagtctctt aagaatccag agaggggatg ttgaagccct 4080

aagaagtaaa aacactaaaa tgataacatt ttgacctttt gaaaactctt tgaagtggca 4140

-continued

atggccacaa aattgagcgg aaaccaaaca tattcagccc catttccaga taaattggat 4200

ttgctaaacc actcagaata acaagagcaa gaattagatg ttcagctaag ctctagcttc 4260

tgatcctaga tggtggcaat aacaacgata gtattaatga taatagcaga aactcttaat 4320

atgtgtgagg cattcttcta aatgctttac atatattaat ttaatgttca tcggagcaga 4380

ttctcctact gtatccaaga gatcagaaat tcagagcaaa aagcagattg ctgcctgata 4440

gcagatccac tctaggaatg gggagatttt ctcataaatc taatgactgg tgaaagtgac 4500

aagaggttag tccactcagt gttggaaagg acaagatcct agaaagtgga aatgctccac 4560

tattgcttta gggggaacag cctggctaca actgggcaga ctttctatta ctaacatcag 4620

gataataggg aagctcctgc caatgctgca tagaactaag aacacttagc agtgatggct 4680

gacttccaac tgagccatta aggtaagtgt agaagaaagc tccaggtaga aatatctttg 4740

ggaactcaga gaacacactg agtgtttgaa gctggtactt aatctattat tgaactcaga 4800

gtttgccatt ctggggccaa tttgagtcca cttatgacac ctcccacagc actgctcatt 4860

tctgttaact gggcttcaca gagggtgaga agggcagcaa tcagatcata taattgggcc 4920

tcattgtaat aaaatgttga tttctgttgt ccttggactt ccttccagcc acaatcactc 4980

tatgatcttg tctgagattc tttttccttt ctctctccat tttgctgact gtaatggaat 5040

aatagcagaa gctctcaatc tgtatgagga atatatatga agaatgtgta acattcactg 5100

tgtttcctaa tagatgaatt atctaaagga cacctcttaa gaagagatga gatggagcca 5160

actgtcagaa aaaaaaatgg gtgaagaaag tttcaaggag aagactccac agagcagaat 5220

gggaaagatg gccaggagaa aagtgagata ataaaaggaa aagcagtccc tgacacagaa 5280

ggtactcaac agctatttgc tgactgaaag aataaacact aataagagaa tgaaatacca 5340

ggattataaa gagggaacaa agatattttc ccatctactt tgttgttacc ccacccagga 5400

tatatttagg ttggtgcaaa agtaattgcg gtttttgaat cgttaataga ttaacatcta 5460

atttatgtat tgtttacata catacaaata caaactacaa ccatatatga tacatttaaa 5520

gtaatacgct aggcacagtg gtgcatatct atagtcccag ctgaggtggg aagactgctt 5580

gagcatagga gttcaagtcc ccttgggcaa catagcaaga ccctgtctct aaacaaagaa 5640

aacaaaaaaa gtgaagtaat taatgccaaa cattgcaatc agggtttggt gcacttgaaa 5700

tacacaaatg agtgttttta catgctgaag acacagcaac caacaactgg ggctaaagtg 5760

ctgcattcct aaagttaacc cacagagcct atcatgccat gccaaagttc actgctccgg 5820

ctccaaaccg gggactttcc tttcctattt cctctcccta acagcttaga gctactggac 5880

tccatccttg aactcctctg tttactaaaa atgttcttta taagctatga aatctggatt 5940

tgaggtctgg tttcttactt caactttgca actcatttcc caccccgagt tcacacaatc 6000

tcaaaatgtt catgaatttt atggctagac ctaaagcttc tttgtgcctt cctgccagcc 6060

tgcctttgtt ttggtacctt ctagcctatc ctcactcact tgagttttta ctctaatttt 6120

ctgtccagtc ttcccaccac tagtccccta tcaaattgct tgctggacat ttctatatac 6180

atagcttgct tgtccttacg tcaaattcaa aactaccaag agcaaatcca atgtgtgttc 6240

ccacatcagc tcccctttct cacctccctg gtatggttac tgaccaatta gaatgaatgt 6300

ccaaccactc agaatgccac aggcaatagg ccaaatcagt gtactgggtc caatgtcagc 6360

cctgtccata atccttctac ccagctaaac ttggaggaat tgtctgttct cacagtctcc 6420

-continued

cttcttcatc tcttattcct tccataatcc atacattttt agccctcctt ccttatcata 6480

aaaattcctc ctgacaaggt caccaatggc cttcttgttg ctaaaaccaa ttaattcttt 6540

ttagctttta tttcacttga tcatccttgg taacatttaa caatgctgac cattccctcc 6600

tccctgaaac actctcctta gcttctctga cattatactc ttctggttgc aatccagtca 6660

cctctggtgg tgccttctca gcctccccca aacctgcttt tcctcaggtt tcttatctca 6720

gtgaaagtca ccaccatccc ctaagctcct aaaccagaaa ccgaattcct tattcttctt 6780

atttgcttct aaccttcctt cagtcatgga atcctgctta ttcttagttt gcttctcttc 6840

actgcctaca ttttttttttt ttttttttga gatggagtct cgctctgtcg cccaggctgg 6900

agtgcagtgg catgatctca gctcacggca acctccgcct ccctggttca agtgattcgt 6960

ctgtctcagc ctcccgagta gctgggacta caggcacgtg ccaccaggcc cggctaattt 7020

ttgtattttt agtagagaca gggtttcacc atattggcca ggctggtctc aaactcctga 7080

cctcgtgatc cgcccgcctc ggcctcccaa agtgctggga ttacagacgt gagccactgt 7140

gaccagcctt cactgcctac atttttaaag cacttaaata tgtgctaggc aggtgtgttc 7200

taagtgattt atcctcacaa taagcccata aggtaaattt tattatctcc aattttcaga 7260

ttaggtaaca gaggcaaaga gtggttaaat aacttgcata aagtcacata actactaagt 7320

ggcagagcag gattcaaacc caggtggcct gactcaattc tggatccagg tcactatcct 7380

atctcactga aatctctgca acagcctcct tggtttccag ccttccatca tggatggctt 7440

ccctgtagcc acagtgatct ttgttaaaac agaagtctga ctcttgctac tcttttgcat 7500

gaaaatcttt actggctctt tattgccttc agaataaagt ctacactcaa tatgacattc 7560

agggacctcc ataacctagc cagaaagtgt atacttcttt ggcttcatgt ttttggtttt 7620

tggccccatt ttaaccacga ccctaaaagc tccactagca caatgatagc ctatacttcc 7680

aggtttctgc actgtttgat gtgccaagaa cactctttcc caccttcaca ttcgacgact 7740

tctgataagc cttttctgat ttcctaagac cagggtatat gcctctcctg tgagctccca 7800

cagctcccag ttctaattct cactgtagca tttatttcac tgtgttatat gattctacta 7860

cccatctact aatattaata tctgctttat cacaagattg tgagcttcct gagaatggga 7920

actgtgtagg attcactgtt gtatcctcag tactcactgc tgcctacctg gtgcacggta 7980

ggcaatattc atttaatgaa taaaacttta ttgtctatat cacactcttt aacacttact 8040

tatataatgc cctatattga ttttattttc atatgtgtta gcctgatctt tataaccaga 8100

atgtcagttt cttgaaagat agcaattata cccagtactt cttttataaa tttatttact 8160

cttaggctgg gtgcagtggc tcacgcctat aatcgccaaa ttttgggagg ctgaggcagg 8220

aggatcactt gagtcgagga gtttgaggtc agcctgggta tagtgagacc tcatctctat 8280

aaaaaataag caagattagc tgggcatggt ggcatgcacc tgtggtccca gctattcagg 8340

agagctgagg tgggacggct tgagctccag agcctgaggc tacagtaagc catgatcatg 8400

ccactccagc ctcggtgaca aagcgagatc ctgtctcata aaagaaaaaa aaattatctt 8460

tagagatggg gtctcactct gtcccccagg atggagtgca gtgtcatgat catagctcac 8520

tgcagccttg aactcttggg ctcaagagat cctcccacct cagcctccca agtagcgggg 8580

actacaggtc tgtgctgcca tgcccagctg attcttttat ttgcagagat gggggtctca 8640

ctatgttgcc caggctggtc ttgaacttct gacctcaagc aatactcctg cctcagcctc 8700

-continued

ccaaaatgtt ggaattatag gcaaaaccca ccatgcctgg cctataccta gtacttacaa 8760

gttaggaggt actcacaaaa attcctgata acagctaaca tttattcagg gatcaatgaa 8820

tgtcaacaat ccaatgattt tggtaatttt tctctctctc tctgcctccc tccatatgag 8880

gaaactgagc catagtttag tagcttgcct aagatgacat agctaggaag caacagagct 8940

gggatttgaa tcctcatgtt ggtgcctgca ttcttaacca ctatgataac aggctaagaa 9000

agcttaaaat tatagatgct aagtaactta cctaagatct ttttaaacat ttaaaaagtt 9060

tttttttcc tttgttagta ggcaaggaac taagatcata gtgcagtgct aagactagaa 9120

cccaagtatt ttcaaaggct acgctcggct gggcgtggtg gctcatgcct ataataccag 9180

cactttggga ggctgaggtg ggtggatcac gaggtcagga gtttgagatc agcctggcca 9240

acatggtgaa accccatctc tactaaaaat acaaaaaata gctgggcgtg gtggcgggtg 9300

cctgtaatcc cagttattcg ggaggctgag gcaagagaat ctcgcttgaa cctgggaggc 9360

ggaggttgca gtgagccaag atcgcaccac tgcactccag tctgggcgac agagcgagac 9420

tctgtctcaa aaacaaaaac aaaaacaaaa aaaaacccca aagcctacgc tcataaatta 9480

tttggtatac tgcttctcct actttaaaag tcagtactaa gctcatactg ggtgttaata 9540

cattttaaaa ctgcaaatcc agtttactga aattggagaa tatcagatcc ccttccaaga 9600

gaggcataat tgtagcttac tgggtgaatt caccttttat attgacagtc tagccttatg 9660

aaccaattaa gtttgtagtt tctgcattct tcctgtgaca ctgaagactg ttttcatcaa 9720

tgtgaaaacc atgttatcta agctaaatca aaatccactc tcattcagaa tttattcttc 9780

cttttatttc tcaacaccag tgacatgttc ccaagttctc tgcctttgtg gctccttagt 9840

ttccagtgtg gaatgtctta tttctgaaag caacctgtga ttgacaagcc ttcttggcta 9900

ttctgtagtt taccctaaaa aaaatctgtt ctctctgtcg gctcattaga ctgcttttcc 9960

ttacatggta ctctgctttt ccccttttatt tcaagcattt attaaaacta aactatgtac 10020

caagcataca caaaatggat ataacctaac tctttgaagg ctacacttca aagaaacagc 10080

cttggtgttt acatattttg actagaatat tatttcccta acaagctgca taaaaccatt 10140

cctgctcagt ttgtaattat acaccaaaat attgagcaca atttcagtag ctacttctac 10200

agcacatcat ataaagtgaa ttgctggagc caactcatta ggtcttattc caaagatttt 10260

taccaacact ttggccctct tttgctgaca aaaaaggatt cttaagcctt ttaacagctt 10320

tataacttcc ttattagagt tcactcacta acatatcctg ttaacaaaat gaagtgccca 10380

tccttcttgg gaaacactct ggagaatatt tcctgatttc atcgcacagc tttgtctctg 10440

gattagtaat catctaatac gaaactgagg atatgagtat tatgtttcac aatgttgtct 10500

tcaaggatat aaacaactca aaatttaagt atttgagtag cactgtatca cattattatt 10560

gccgtgatat tttcagaaga aatctgaagt ctgcgctaac tacaggccat aaaatatggc 10620

catcttgccc caataaatgg agcaaagttc aaatctggat catcagaaaa agacaagagg 10680

cacatttcta aaccaaatct gattgccttg aggaaggaag gtactgaacc tgttagagta 10740

catctctgaa ggactgagaa tcaagtcctt cattccctcg ccccctaaaa aaatccaatc 10800

aatcatgctg ataaatgtct tagttacagg ggagaaaaaa tagacatgac taagctgaaa 10860

aaaatcttct ggctttgtat cttttctagg cagtgctgtt cctgataaga cacatgacac 10920

aaaatatggt tttgctgagt atcaaggtaa gttagggaaa atagtatttt ttaaatagga 10980

-continued

aaattagatg acaacagcaa caggtaaaga agccttgggc ttctgaggta gaaatactag 11040

catccctaca taggcatggt agataatctg aaactagtac aagtgctggc ctgttatagc 11100

ctgtctctac cgtaatcaga gggaaaagta ttacaatcca gcaaaagtag cccttcaaaa 11160

tctaaggagg agcatttcat cagatatcaa taaatgaacc ctcgctgcat acaaaggcaa 11220

tacattacgt gctgagtgac caacttcaaa gaaccttctg ctgttctgaa atttaaaact 11280

actctcattt cctcaagcct gtattttatg taatgaaaca aaatgcacat catatgtctc 11340

ctgctgcact gtaagagtaa acataggacc cctcccacca aaaattgtgc aaagagccct 11400

tccccttcag ctttttgcta tagaaatatc ctactacttc cctgaaaata gtatatgctt 11460

tctttggcaa gtggaatatc accctgttac tctctccctt actctgcttc actttttttc 11520

gcataacgct taagattacc taatatttta catgtttatc atgtctctcc caattagaat 11580

ataagcttca tgagggcaga gatttcaaat attttgttca cagatgtatt cccagtatct 11640

acaactatat ttagcatgta gcagattctc aaaaatattt gtaaaattaa taaaatttaa 11700

aaatccctat tatattggaa ttgttaaatt aaaaaaaaat tttggccagg cgcagtggct 11760

cacacctata atcccagcac cttaggaggc caagacaggc ggattacgag gtcaggagat 11820

cgagaccatc ctggctaaca cagtgaaacc ccgtctctac taaacaaaat acaaaaaatt 11880

agccgggcgt ggtggcgggc gcctgtagtc ccagctactc gggaggctga ggcagaatgg 11940

cgtgaaccca ggaggcggag cttgcagcga gccgagatgg cgccactgca ttccagcctg 12000

ggcaacagag caagactctg tctcaaaaaa aataaatttt ttttttaaga cagggtctct 12060

ctttgttgcc caagctggag tgcagtggca tgatcacagg tcactgcagc ctccaactcc 12120

tgaactcaaa tgatcctccc acctcagtag gtgggactac agatgagtgc caccacacct 12180

acctattttt tttttttgt agagatgagg gtcttattat gttgcccagg ctggtctcaa 12240

actcctgggc tcatcctccc accttggcct cccaagtgct gggatttaca agtgtgagcc 12300

actatacctg gccaaatttt aaaatgctac cgtatctgtg tatgaatatg tgtgaataaa 12360

tttacagact ctagggtgta attcacaggg tagaccaagt caagagtcag aggatcagaa 12420

gcttgggtgg ggagatagga gagtgattgc aaaaacaaag caatctatct tattaaatgg 12480

acgaaaggga tcaaacatta tcagtaaact attttaaaga gtgaattaaa tgtatctctc 12540

tatcctttct atgagtgtag tgaattcaga ttcacctcag attagttgaa ttgacactag 12600

ttttcttccc tgataaagct tttaaaattt atttagtaaa gtggccaggc atggtggctc 12660

acacctgtaa tcccagaact ttgggaggcc aaggtgggtg gatcacctga ggtcaggagt 12720

ttgagaccag cctggccaac atggtgaaac cccatctcta ctaaaaaaat acaaaaaaaa 12780

atttagccag gcgtggtggc gcacaccttt aatcccagct acttgggagg ctgaggcagg 12840

agaattgctt gaacctggga ggtggaggtt gcagtaagcc gagatggtgc cactgcactc 12900

cagcctggga gacagagact ctgtcccccc cacccaaaaa aaaattagtt gttttcctcc 12960

ttaaattaca gcctgcaaga ttgaaacaga gtttccccat ttctattttg ggatctgact 13020

gagtcacaag gtagtcttga cgaatcactc tgctctatct ctcagtttac tgaggctgcc 13080

tgggaacttg agctgtcagc caaacaggaa gaagggccct tctgcctact ccaaaaggcc 13140

tcagtcaaat ggatatggtg agatttaatg atggatataa attacagcaa atggatttcc 13200

tatagggtga gggaaaaaaa aaaaagcaaa cattttactt aaattcctct atacaattta 13260

-continued

gggctggcag aaagggtaga ctattacaat gtagctgagc aaatttttcc atgtttgcat 13320

aacatgtgtt ttcaactaaa gatactcctt ctatagtcct actatataca acaccccaaa 13380

aaacctcaaa tgtaagaaaa agcaatgtga cagatttgac cctggccaag ttacactgtt 13440

ttttttttt ttttttttt ttttttttt taagtgtcag tgttcataaa ggccctttt 13500

ctttttcaag gatgggtata aagtgttact cggccgaacg cggtggctca cacctgtaat 13560

tccaacactt tgggaggccg aggcaggtgg attacgaggt caggagttca agaccagtct 13620

ggccaacaca gtgaaacccc cgtctctatt aaaaatacaa aaaattagct gggttgtggt 13680

ggtgtgcccc tgtcacaagg tagtcttgac gaatcactct gctctctctc tcagtttact 13740

gaggctgcct gggagctgtc agccaaacag gaagaagggc ccttctgcct aactccaaaa 13800

ggcctaagtc aaatggatat ggtgagattt aatgatggat ataaattaca gcaaatggat 13860

ttcctataga gtgagggaaa aaaaaaatcc agctactccg gagactgagg cagaattgcg 13920

agaacccggc aggtggttgc agtgagccga gattgtgcca ttgcactcca gcccgggcga 13980

cagtgtgaaa ctctgtctca aaaagaaaag aaaagaaaaa aaaaaaaact attactccac 14040

tccccaccag aaactcttat tttttccccc ctatgtggtc atgaattcct cctgggtgca 14100

tagctcccta agttgttctg tgttttagtg cttataattt cccacatccc atgatgcccc 14160

ttgatagaat tatttttttct gtactcacac aaatggaaac agaagagatg aaaaaccctg 14220

gggatgcggc ctcgcaggtc tcaaactgat cggctgctaa gcacaacctg catatctctt 14280

ggctgagttc ctttcttggc tcagttagat tttgcatgac ctaggaggct taggacccag 14340

ggggcgcctt tcagctgaaa aacagctcgc gctgcagcaa gctagctggg aagctcccag 14400

ttctaaagag aggctgttta ccagaacagc ataacaaggg caggtctgac tgcaaggctg 14460

ggactgggag gcagagccgc cgccaagggg gcctcggtta aacactggtc gttcaatcac 14520

ctgcaagacg aaggaggcaa ggatgctgtt ggcctgggta caagcattcc tcgtcagcaa 14580

catgctccta gcagaagcct atggatctgg aggtgagatt acaatctttc ctttccgcag 14640

caacttcatg aaacagaagg ggtaaagccc cctatttccc aaacagggcc aggatgacag 14700

gaaacgaagc gtttagccgg tttctgttgt cccgtgaaag cttttccacc cctcctttcc 14760

ttgtcagggg cttgcagctc catctttgcc attaagaaca aaggcaagaa gccactttct 14820

ctctgccttt ccttgctcat gagtggatgg agcttggggg ttcatctccc ctactcccca 14880

agtcctgttc tttctaagcc ctttgcatgg tgtggtcctc cttccccagc ctccttctgt 14940

ggttctcctg cagaggggag tgggagcttg gagtctcatc cagtcagcga tgaggccggt 15000

ttcctccgtt tgaggctgtc ttgaatcttt tgcctgacgg agtcaaggct cctggagaat 15060

tctcctggga gggagtattt gccttttaaa cgggcagaag gtggccgtat ccgtgaagga 15120

ggaaacaccg gagtttatgt aacgtggcca ttacaacaca gggacaaatg tcacgtctgg 15180

gcgccccccg tcccaaaact cccagctcct caaagggacc ctaagccagg cctcctcctc 15240

tccatctcct cctccaaagt ggtcacaagt tcaagagtaa caacgctttg cacttctgca 15300

gcgctttcca ggtttcaaaa ataaataaaa cctaacagtc tctgagcgtt gactgggggа 15360

caggcacagg ctgggggctt cacaccctga ggtgtgcgtg gagttggggg gggggggtgc 15420

gtggggcgtg ggggtgttgc cacgagagga gagggaggcc cagcagggcc cggggtcccc 15480

taaccaggct gaggttcagc ccgggtggtc ccggctcccc tgtccgccac gctgggcgtc 15540

-continued

```
ttccgttcgc tgcccgctat tctgtgtccg ggggcttctt ttcctcacag gctgtttctg  15600

ggacaacggc cacctgtacc gggaggacca gacctccccc gcgccgggcc tccgctgcct  15660

caactggctg gacgcgcaga gcgggctggc ctcggccccc gtgtcgggta agtgtcctcc  15720

gggggacggc cccggggctg gcgggcggcg gccgcgttcc cggtgccagc cccggcacac  15780

atttccttct ctcctctctt aggggccggc aatcacagtt actgccgaaa cccggacgag  15840

gacccgcgcg ggccctggtg ctacgtcagt ggcgaggccg gcgtccctga gaaacggcct  15900

tgcgaggacc tgcgctgtcc aggtacctgc ccgggacccc gggaatcccc gctcctggcg  15960

ctgacaggaa gcggctgaga cactcgtggc gcacgtgcgg ccgcgcgatg ccggccggcc  16020

ccgctccgcc tccgcgccgt ccgattggcc agggctgtct ccaggccgcg gggctcaaag  16080

cctattggcc gccgccgcca ctcgtatgag ctcatccgcc gcgagcgctt gtaaacaact  16140

gcgcggggcg aggctggggt gaggctaggt ctctccgccg cgcgtcctga ggagtggctt  16200

ttggcccctc agcgtttatt gaccgtttgc ggagcgtctg cagtgtgcta gacccacact  16260

gctcacaacc ccagctcccc aacctcccct gacacacacc gcgggaagaa gtcgtttgag  16320

gcatctgagc ttcagagaag ttaagcgatt taccaaaggc aatggcaggg gacagatttg  16380

aattcacgtt tcttagattc cagagcgcag gctttttgct cccaaagcag caagagagtt  16440

ttctaggttt aaatgggata atcaaaataa agcatttatt ttttattttt tatttttga  16500

gacggagtct cgttctgtca cccaggctgg agtgcagtgg cgcaatttcg gctcactgca  16560

acttccgcct tccgcgttac agtgattctc ctgcctcagc cgcttgagta gctgggataa  16620

caggcgtgca ccaccacgcc cagctaattt ttgtattttt agtagagacg gggtttcacc  16680

atgttggccg ggctgctctc taactcctga cctcaggtga tccgaccgtc tggcctccca  16740

aagtgctggg attacaggcg tgagccaccg ctcccggcac aaagtaaaac attttaccta  16800

aggactggca cagagcgtac gcttagatgt gagctgctca aaggaagcat gggttgtttc  16860

ctgaagtaga agtgggggga atcccctgca gaagaacgca accgagaggg gccttgggaa  16920

atagcaagga tttggacaga taaagacgtg ggaggtagga gagaagccct tccctacggc  16980

cttcctctct catttgcctt gtccctgaga tggttcagcc ctctcaggcc tgatcttacc  17040

tctggggcag aatttgactt tgggttgtgt gtgggggggaa attcttaggt gacccacctt  17100

catctaccct aacactcagg ccaggttcag ctgtggaatt cactcagggc agtccctgat  17160

gagagctgtt tcccacctca tctccagaga ccacctccca ggccctgcca gccttcacga  17220

cagaaatcca ggaagcgtct gaagggccag gtgcagatga ggtgcaggtg ttcgctcctg  17280

ccaacgccct gcccgctcgg agtgaggcgg cagctgtgca gccagtgatt gggatcagcc  17340

agcgggtgcg gatgaactcc aaggagaaaa aggacctggg aactctgggt atgacggtcc  17400

cccacccctg ccctcgttgg gattcatcaa gagatgtcat ttgctgattg tctagggtgt  17460

ggctaatggg accttgtgtc ctatccttgg caggctacgt gctgggcatt accatgatgg  17520

tgatcatcat tgccatcgga gctggcatca tcttgggcta ctcctacaag aggtcagtag  17580

cttctcttct gggccctctt aggaggaggg gaggaaggta cacaaagtca aactttgtgg  17640

ctttcttacc caaaaggaag caagatgttg aaaatcagct caagatgtag agacttggtc  17700

ctagtcttta tctgctgacc ttggacaaac cctttccttt ggtgccttta cacctcactt  17760

aaaagccttg tgagggatgt tcttgttccg cattttacaa atgaagaagc ctctgtcaga  17820
```

-continued

gacttgctca aagtcacata gcctaggagt ggcagaacct ggattcagac ccagctcttt 17880

tgatagcaaa ttccctgctg tctgatccca catgcctggc attgaagtga gtggcactgc 17940

agaaatcagg cttgacactg attgcaggga ccttaagacc aaattctttc tgtgagttat 18000

aaaattctgc tactagaaaa ggtttccagc cactgttgca gttgaggctt aaatttcctg 18060

agctataagc agtgggatct cctctgctag ggaagaaaaa agtgagttgg agaacagaaa 18120

gtctgtttgc cactcttcac cacgtgtgag aaagccaaag aggctacttg tgaaactcaa 18180

ggacatgacc tgtggctggg acagcaggag aagaatgctt attttgaaag gtcccacgag 18240

atgtccgtca gtgactggtg catgatggta agccagactg gcattcatct ttacagcact 18300

ggctgattgt ggaaaattat ttgccattct tgggcaaaat cgggggattt gacttctagc 18360

cctgggtgtc tgtaatattc tcctctgtgg tctgagaaaa cttcctgtga ggaaagtgac 18420

acagtagcct gcatttttcc ttcctttcct ccctccctcc ctcccttcct tccttcccct 18480

ttctttcctt cctttctttg attttttttt ttcagggtct cgctctattg cctaggcttg 18540

agtgcagtgg tgcgatcacg gctcactgca gcctcaacct cccagggcgc aagtaattac 18600

ccccacccac ccccactgag tagctgggac tacaggcatg tcccaccatg cctggctgat 18660

ttttaatttt tttatagaga cagggtctca ctgtgttgcc caggctagtc tcaaactcct 18720

gggctcaagc agtcctccta cctcagcctc ccaaagtgta taggcatgaa ccaccatgcc 18780

tggcatgcag cctgcatttt ccagtgatga gggccatgtc tgagcccaca gcctcagtgg 18840

atcacccagc agggcatctt taaaggcctt gtagtcaacc tcctcattac agatgaaaaa 18900

ggggaagttt gggtataata ataacagcag tagcaaacat gtaccgtatg ccagacactg 18960

cttgaagcac atgctttaac tcataatcac ccttatgagg taaatgctgt catctccaat 19020

ttacagatga ggaaactgag gcaagagggg attatatacg ttacctgagg ttacacacct 19080

agcatttggt ggagagaagt cctttaacta gctatctctg tgaccctggg caaattgctt 19140

caatttccta ctccacatca tggtgtggag ctatggcttt tgactactta gctcagggtg 19200

tggcacatag taggtggtca gtaaatgatg gctgtttttc ttttctataa aatgtgtggc 19260

tgggcgtggt ggctcacacc tgtaatccta gcactttggg aggccgaggc gggcagatca 19320

tgaggtcagg agttcgagac cagcctggcc aacatagtga aatcctgtct ctactaaaaa 19380

tacaaaaaat tagccaggcg tggtggcggg cacttgtaat cccggctact tgggaggctg 19440

aggcaggaga atcgcttgaa cctgggaggc tgaggttgca gtgagccgag attgtgccac 19500

tgcactccag cctgggcgac agagccatac tccgtctcaa aaaaaaaaaa aaaaaaatgt 19560

aggtaatgga ccaggcgtag tggctcacac ctgtaatcct agcactttgg gaggcagagg 19620

tgggtggatc gcttgagccc aagagttcga gaccagcatg ggcaacatgg caaaaccctg 19680

tctctacaaa aaaatacaaa aattagccag gcatggtggt gggcgcctgt agtcctagtt 19740

actaaggagg ctgagatggg aggatcacct gagcctggga agtcgctatg gtgagccatg 19800

attgtgccac tgcactccag cctgggtgac agagtgagat gctgtctcaa aaaaaaaaaa 19860

aaaaaaaaaa aagtgggtaa tggactagta ctctctgatt tcttatgatg caggaagcag 19920

aggtccagag agggaaagta ttgacttgag aggacatcct gacgcccagt ctgaggttgt 19980

ttcccacaat gggatgcttg acatgcaggt ccagggacac ctgcttgttc atttcagagt 20040

taatgacata gtggccctgg ccttacaatt cataagcaga cctaagccac attggctgag 20100

-continued

tttagtcctt gaactatctc ctttttcctc agccacccaa cagcatttgg ctgtcctgca 20160

atcctgtgac agtttccagg ccttccaggt gcttgaactg actaaacatt cttaattctg 20220

agctcagttc tggtgtccag tggaagtctt ttttcttcta gatttgttcc tctaggctgt 20280

cacgagactg ccctgaggcc taaaatttat tttcctgtaa actcagggca aggcacccag 20340

tcatcagact ccccgtcact aacgaccaga tcaccccttt cttcagcaga actgctctga 20400

gaacagactg tacaacagga aaccgcatgt gggcttaaaa acaagcttgc agccttcatc 20460

tcttcctaaa gtcagagagc agagtgtcct gaaaagagct gtaggttgag tgtgggttca 20520

ggtgccctct gagccttggt atcttgtctg ttaaatgggg atatatcctc cctccacatc 20580

tacccatcta agattataag tattaaatgt aaacagaaac aatagcacac caatcaggta 20640

ataatcacaa tagcagccaa catttatttc atgctttctt aataagtaag gccctactgt 20700

atgctcttcc ctgattatct ttttttctta attttttttt ttttgagaca gagtctcgct 20760

gtgtcggctc actgcaaccg ccgcctcctg ggttcaagtg attctcctgc ctcagcctcc 20820

cgagtagctg ggattagagg cacccgccac catgtcctgc taatttttgt atttttagta 20880

gagacggggt ttcaccatgt tggccaggct ggtcttgaac ttctaacctc aggcaatcca 20940

cccgcctcgg cctcccaaag tgttgggatt ataggcatga gccactgcac ccagcctcct 21000

gattaccttg actaagcctc acaaccacac taggagcact gttagatttt acagtgagga 21060

gctgagacct agcagagtta agtacaatgg ccaggctaca cagctagtaa tgtatgtctg 21120

aggagtatta gtatctccag caagtccagt gaggccaagt agcctcatgt aaatttgctt 21180

tgaattgcaa aggaaagtag gaaaaataga aatttaaaaa aaggctgggt ttaaagtcag 21240

aagaaacaag atctagttct ggaactttca cctattctgg atgtgacact ggacaagtta 21300

tttaccttct ctggacctca atttcttcat ctgtactgta agggggttag atatggtgat 21360

actcccaact ctatttccta aaaagacaat attgccggtg ctaagactgg gttcaaatcc 21420

cagctctgcc atttgcttga acaaatagtg taacttctct gcatctgcct tcctatctat 21480

gaaacagatc aaaatagata ccacataggt tgtgaggatg aaattggaca gtatagttta 21540

ggcatacaat tggtgttcta tagcctggga gaggaagttg gaggcagcat acagtagggc 21600

cttacttatt aagaaagcag gaaataaatg ttggctgctc ttgtgattat tacctgtagc 21660

tgccttggga acttttcctg agcagcagtg tggcacagca ctgcctctga ctgggaaaat 21720

agcagcggct cagagtaata attgaagacg acctgtgtag tctgaatgag gtctgtggtg 21780

cttccatact ctgggagagg aagttggagg taggcagcag cttcctctac acctgctctc 21840

tctggccagc tcagctccca gtcccaccct ggccttggag caagcagtca gggtaggtac 21900

tgtaagcctt gagctctgcc aaaggccatg aggcttgcct aatgggatgg agcttttccc 21960

tggctgggat gggtgggcag gcaccaggct tcctcttagg gagggagggc tgtaaatggg 22020

ggtgggtggg agctcagcct ggccgggtct ttgtcctggt agtctaggcc acgaggttca 22080

aatggccagt ctgtgactga ggagctaagt gtactgtctg cactagcttc agagggaggt 22140

cttattttct gataaaaagg ggagggaaaa actgtaagtc aacaccactc cactccctta 22200

ggagaggaac taagtcagtg aacaagcctt ttgttctttc ttggctgcca cacaacaccc 22260

aggctaacca ccccctaccc ccaagcagtg agaaggggct aggctgcctg atggtcagtg 22320

tagcaggcct agtggcctct caaaggtcac ccaagggagc tgggagcagc cactgcatct 22380

-continued

ctaggactca caggcaccat taacagcagg cacatgggat tgagtgtgct tccaggcttc 22440

caaatggatg agactaatcc agactagtgc cagcatctgc tttctttacc cttgactgct 22500

cttccagccc tgctctgcct cattaataca aattgatttt tttcccctaa agacttgtat 22560

atgatgcagc ctggccatta gcagcagctc agctggccac aggtaaagga ggttgctgag 22620

gaaagcaagg cgcggaagtt aaactcaaaa gcctcagttt cttcatctat aaaacaggga 22680

ttaggctggg tgtggtggct cacgcctgta atcccagcac tttgggatgc tgaggcgggt 22740

ggatcacctg agatcgggag ttcgagacca gcctgaccac catggagaaa ccccatctct 22800

actaaaaata caaaatactc tgggcgtggt ggcgcatgcc tgtaatcctg gctacttggg 22860

aggctgaggc aggagaatcg cttgaacctg agaggcagag gttgcagtga gccgagattg 22920

tgccattgta ctccagcctg ggcaaaaagc acgaaactcc atctccaaaa aaaaaagcct 22980

gtagctgaca tttagtagat gcttattaaa agtaataagc atctctggga ggctgtctca 23040

cagcaggatg atcaggtcag ttttctggac aactttatgg gttaggtccg aagttcagat 23100

agttctctgg gtctctgggg ctgaaagaac aaaactgtgt cttgctttaa gagtctctat 23160

ctggccaggt gcagtggctc acgcctgtaa tcccagtact ttgggggggtg ggaggatcgc 23220

ttgagcccag tagtttgaga ccagcctagg caacatgttg agaccctgtc tcaaccataa 23280

aaaaaaaaaa aaagtctctg tcctttactt tattttcatt gagccccggt gaagctgtat 23340

gaagtccata ctgctgtcac cctattttac agatggagaa cctaagacca gaggtgcagg 23400

agcttcttgt tgagtttaac gccgtcttgc agaatggata tcctagaatc ctggacctga 23460

gtttgggtcc tggctcagcc actgatttgt tgtgacaatg gcaagctgct gctcttgctg 23520

agcctcaggg atgctatcta atgagcaaag ctatgcagga ggcatagcct gaatgtttgt 23580

ttgtttttt cctgccaggg ggaaggattt gaaagaacag catgatcaga aagtatgtga 23640

gagggagatg cagcgaatca ctctgccctt gtctgccttc accaacccca cctgtgagat 23700

tgtggatgag aagactgtcg tggtccacac cagccagact ccagttgacc ctcaggaggg 23760

caccaccccc cttatgggcc aggccgggac tcctggggcc tgagcccccc cagtgggcag 23820

gagcccatgc agacactggt gcaggacagc ccaccctcct acagctagga ggaactacca 23880

ctttgtgttc tggttaaaac cctaccactc ccccgctttt ttggcgaatc ctagtaagag 23940

tgacagaagc aggtggccct gtgggctgag ggtaaggctg ggtagggtcc taacagtgct 24000

ccttgtccat cccttggagc agattttgtc tgtggatgga gacagtggca gctcccacag 24060

tgatgctgct gctaagggct tccaaacatt gcctgcaccc ctggaactga accagggata 24120

gacggggagc tcccccaggc tcctctgtgc tttactaaga tggcctcagt ctccactgtg 24180

ggcttgagtg gcatacactg ttattcatgg ttaaggtaaa gcaggtcaag ggatggcatt 24240

gaaaaaatat atttagtttt taaaatattt gggatggaac tccctactga cctctgagaa 24300

ctggaaacga gtttgtacag aagtcagaac tttgggttgg gaatgagatc taggttgtgg 24360

ctgctggtat gcttcagctt gctggcaatg atgtgccttg acaaccgtgg gccaggcctg 24420

ggcccaggga ctcttcctgt ttcataagga aaggaagaat tgcactgagc attccactta 24480

ggaagaggat agagaaggat ctgctccgcc tttggccaca ggagcagagg cagacctggg 24540

atgccccagt ttctcttcag ggatggatag tgacctgtct tcattttgca caggtaagag 24600

agtagttagc taacctatgg gaattatact gtggggcctt gtgagctgct tctaagaggc 24660

-continued

taacctggaa actaagctca gaggcaaggt aataaagcac ttcagggctt gctccccaag 24720

tgggcctgat ttagcaggtg gtcctgcggg cgtccaggtc agcaccttcc tgtagggcac 24780

tggggctagg gtcacagccc ctaactcata aagcaatcaa agaaccatta gaaagggctc 24840

attaagcctt ttggacacag gaccccagag aggaaaaagt gacttgccca aggtcgtaag 24900

caagctactg gcatggcaag agcccagctt cctgacggag cgcaacattt ctccactgca 24960

ctgtgctagc agctcagcag ggcctctaac ctgtgatgtc acactcaaga ggccttggca 25020

gctcctagcc atagagcttc ctttccagaa cccttccact gcccaatgtg gagacgggtt 25080

agtggggctt tctatggagc catctgcttt ggggacctag acctcaggtg gtctcttggt 25140

gttagtgatg ctggagaaga gaatattact ggtttctact tttctataaa ggcatttctc 25200

tatatacatg ttttatatac ctcattctga cacctgcata tagtgtggga aattgctctg 25260

catttgactt aattaaaaaa aaaaaaagac tccacattgc caagtttttg aggggtaaca 25320

ggaaccctcc gtgtaagttg agaagctttg ggttaccttg tcaacaccta tgtggcaagc 25380

tctgggcctt ttcttcattc atcctcgccc tgttactgga aagtgacaac tgcagcctgt 25440

gtccagattt tttttttga gatggagttt ggctcttgtt gcacaggctg gagtgcaatg 25500

gcatgatctc gcctcactgc aacctctgcc ttgtgggttc aagcgattct cctgcctcag 25560

cctccagagt agctgggatt acaggtatgt gccaccacac ctggctaatt tttgtatttt 25620

tagtagagac ggggtttcac catgttggtc aggctgttct caaactcttg atctcaggtg 25680

atccacctgc tgaggcctcc caaagtgctg agattacagg ctagagccac cgtgcccggc 25740

ttgtgtgtcc agattatcat catcaggaat ctttcaggtt gaagcaggaa agacaggtgt 25800

gcaggcgtcc caacaaggtg gttggtcatt gcaaatcctc tccccttcag caattgtttc 25860

caaggtggag gtagttaaaa tgaatatcta gacagccctt gttaaaggct ggagtgcggc 25920

aggggttggg tggtgggtga cgggggtggt gctgcagcag ccaaaggggg aataactgca 25980

gaggaggaag ggtcactggg ctgcctgcag cagccctggg ctgaggaagt tccctgcaag 26040

gtgactacta tgactccaga agaaaagata ccaacaggag tgtctcctgg acttttagca 26100

agggatttca gccacaggtt atcagcctca ggaacaccca gatgacctct gcccatgccc 26160

agccattgct ctgtggttct gccagcaatg cgtactccag atcatcagtt tttttcggga 26220

accatattct gtccttaagt tggattcttc aagctttgag agtgcctggg agaaaacacc 26280

tttgggttta cagatttata ttgtgcccaa ggccccacta accagatagc attgtccttc 26340

ccaatggtca ctcttctacg acatccccca catgcagaag caagaagaaa ctggtcttct 26400

cctctcaagc acttgtaaag aaacacaagg ggtggggtga cttgcattca tttcttccca 26460

ttgcaaaggt ggcacccagc actgaggaaa gcaagccagg atacccaggc tatggaagta 26520

ttcatataga gaggatgaaa gggtagggct cctaaggaat gactcagctg gcatgtgaga 26580

gccaggccag gaaaaggtca tagtcagtca cacctgctgc aaggaagggg ctgcctcagg 26640

tgggatgtac atattctcta gagcagatgg gtgtctgctt gagggtttta gacttgggtg 26700

catgaaaact ctctgctatg gggagtttgt gtgggctagg ccagaaggtg gctagcagac 26760

cctctacatt tatgtccctg agccaatcat gttttcactc tccatattgc cagttatgtt 26820

aactgcttaa tttatttagt ctgatcaatg gcccagttca ggcccaccac tcctacactt 26880

gctggggaca gccaggaaac aaccagccac acaatgctga actgtgctgc tgctgggagt 26940

-continued

```
ccagggctca gccctaaagc aagcttgcaa acttcacaca taagtacagt ctatatagca  27000
agtaaactct gaccagagat gacatctggt cccacaactc atcaggtcta tgtacaatat  27060
ttcacatacc acccaataga taagataata ttaacagcaa ccactctcct ttatcaattc  27120
cccctgctcc aatacaacca ccacacattg cattaatacc ccaaacccat tcccaattta  27180
ttaaatatgg tgcaagctca tagacactta gaagaggcaa atctagttgt gatgaagagt  27240
tcctagagct ctgggagcca agatggaggt tttccagtac ctgcacatgt ggctcaggag  27300
gatgctgccc aggagctaat gagttgggag agcaaacatg ggaggtagaa gtcagatggc  27360
ccagctcagg gagctatctc tctcagcatc tcagctttga gactctgcca ccacctcttc  27420
ccagcccaag ctgctgccta aaccaggcat gttgaagggt gagcagtggt tgccatgaag  27480
ccaagaccaa gagattgctg agactcccac tcccctccct cagactctag gcctgtgaca  27540
agccacactg tcctccagaa cccatcgagc tttaggcaaa atgttttagg catctgacta  27600
aggagcccac ccgagtatga gtaacagaag ccaagatctg agctttctag agggcagggc  27660
ctctttctag tcccccagcc tcttcctttg cttgtgactg tttgtttcag acagaaagga  27720
tgttgtccgt aagttctcag ccagtctcca gtattaagtt tgaaaacatg aggtggaggc  27780
ttcctctcac attgatgtgg tccctgggct ggctagagcc agtgagcagg taaacagaag  27840
ccagcccttt ctctaactcc tggcctgttc caccacatat taagggactc ttcaaaacct  27900
actccctcaa ccttgctcca ggaaggacag gatctggagt agaaaggggg agcatcactg  27960
acttttcac acccctttgag ccatcagtct ttctttcgtc cacacacttg aagtcctcct  28020
gcacccagct ggactagtga cttcagagtt cacaggcagg ttcaccctgg gatcctgccc  28080
agcaggtcca ggtcacgctg attagtgtaa gcaagagcct cccacaatcc actgccaggg  28140
agttcctggg tggattcgag taacagattc ctcctggccc aaaggccagg ggtttttttc  28200
ttcttcacag ctttcaggca agtattggat ttacagacag taactaggcc ccagaacctg  28260
tggagataca tttccctggt gctgcgccca cttgcctcct ggggaggtaa taggaatggc  28320
ttgttctgct tctaaacatt ccgccaatcc acaggaagcc cggacggccc tgctcacagc  28380
aggaatgggg cacagagggg caatgctggc tgtagaacac ccccctgcag ggggctgggc  28440
cagggctagg gaggtgagtc ccgggtcagg ccgtactgca tgctcacagt gtggtccaac  28500
tcctccagct ctgcaggcag cgggatgccc agctccccca ggtacaggga gagggcctca  28560
aaggagtcct ccaatttcga gaatgctggt ctagaaggaa aagaggtaaa ttagaggcag  28620
gtggaacatg agcagggcca tggggtgggt ggaagggatc taggctcagc tacttcctgg  28680
ttgtataacg tcaggcaaac tccttcacct agtagtgcct ctgtgagaaa tgggcataat  28740
gttaccggta cctcatatat ttgctataaa gattaggtaa cactataaca aggctggact  28800
agctggagtt aatacttaaa aacaatagga gaaaaacttc cctcccatat cttagagaaa  28860
atgcagttat caaaggtgga atcggaaaca ccaggctcct agtgccacgg aaatggcttg  28920
gctgccccgg aagcctaaga cagctcaggc ttactcctcc tcctccatgt ttatgcttaa  28980
atgtcatgtc acctcagtga ggcctgctct gacgaccata tttaaaattg caactggtac  29040
cctgaacaat cgtctttact tgtcttgact ctttttccat agcatctgcc acattctaat  29100
attcaatata gtttacttat tgtatctact gtatgataca gtagatacca tgctagaatc  29160
tatgttctat gagggtaagc acctttattt tgttcgctgc tgtaacccaa gtgtttagaa  29220
```

-continued

ctgtacctag cacacagtag gcattcaata ctgtatatat atatttttta ataaattaaa 29280

gatatttaga atgacaagag gtattagaaa acaaatactg tgtagcttca caatttctcc 29340

tcaagctttt acactggtgt tactgaagga ctgcggtcct tggcctataa aggaacaagg 29400

tatagattct cccaccagaa aatttcaact ctaatgtgga ccgctggaaa tcatttttt 29460

ttttttttt gagctgcacc aaaggcgtac cctacctcgc tttattaaag ggcccgtgcc 29520

gcagagtcaa tataaaaaca caaagtccca tcagtttaat aacaataaaa aaatccaaaa 29580

gtggaaaact gaggggcag gggaagagac ccctgggcca ggggcacgag gagccctgct 29640

catggaacca ggcctggccg cagggtcccc cggtattgct gttgctacga ggtcgggggg 29700

tagcgattgt cctatgggag ccaccgttcg cctgggtcgg ggaccctcac ttcttctggg 29760

gtgtgctcag cttctgcatg gcccggatct tgtccagcag gccagagatg aaggcctctg 29820

tgggtttgta acagtcaacc agcagctcct tgaccctgga agcccaggca tcgtcactct 29880

ccatgtccag gagctcatcc acgtcaatct ctagttctga gatctcctct tcctggcagt 29940

cgtagaggcg cgtgagctgc tccaggatcc actcctctag gttgaggtgc ttccgtagct 30000

ccttggggtc atacttgatg gtgaccttcc cttggcgcct cactgggccc tcatcatccg 30060

cgcagcccgg gccctctcct gcggccccgg gggggctct gaaagtagat gcgtggtcct 30120

tggccgccac tgcctggcct gggggccggg gccggggccg ccaacgccgc gccccccgcg 30180

gtgccgctgt ccgccacggc ggcctccggg gccacgtgcg cgcgtggggc ccctccttgt 30240

gccgccgcct ccgtgacgcc cgccggctgg ctcaggttag ctcgccggct ccgcggcgag 30300

ggcggcggcg ggggcgcccg ggggcagctg cagcatgcgg agcccccggg cctggcctgg 30360

tcgcgccccg ccccctcccc gccgatccgc ccgccctttg tcccctgcgg ccgccgccgg 30420

ggctgccgcc gccgaagcgc tttctctatc cggaaaccat ttttaaataa agattatgtc 30480

ccaactggtt atcctgggag ggtggagttg ggggtgagga gaatgttgag ggatatattg 30540

gttaagggag attacagaag acgtatttcc tgtgctataa actttggtaa tgtttgaact 30600

tgtttttaaa agaaacatga tttttgatgg caggaaaaga accccatttt tgagacagag 30660

tttcactctt gttgccgagg ctggagtgca atggcgtgat atcggctcac tgcaacctcc 30720

gcccctgggt tcaagcaatt ctcctgcctc agcctcccaa gtagctgggg ttacaggcac 30780

ctgccaccat gcctggctaa gttttgtatt tttagtagag atggggtttc accatgttgg 30840

ccaggctggt ctcgaactcc agacctcagg tgatccacca gcctcagcct cccaaagtgc 30900

tgggattaca ggtgtgagcc atcgcgccca gccagattct cgtttttttaa ccacacactt 30960

gaaaagacct agggtaactt tggtaataat gaaattatta cctgagaaga ataaaggtaa 31020

ttttgaaaat acttcctcaa aggagaaatt agctaaatgc ataaaactgc tcactgtagt 31080

taaataccta ttgtaaagag aaactggaaa aacctgaaca tctacattac aatatgttca 31140

gacagtctct caatagaatt tcatgttgtc gacaaaaatg ataaacagga agactggata 31200

gtaaacacaa gggaaatgca gggcaaaata cagaactatt tactataaaa aaaagtatat 31260

gtgtagatat ggacagggga atatcgaaag atgaatacac ttactgtatt caaatggttt 31320

gggtgaattt ccttctttta aactttcttt taatggctgt gtttctgtgt cctatccaga 31380

ttacttgctg ctatgtcttt acctctctct tctccaacct gtacatatgt gttgtcagat 31440

ggctcagggg ttgacagcca cattcttgac ctcaaggctt ctctcacccc tgtgcttgca 31500

-continued

gagggctcag gggtggcctt agacagaggc aaacgtccca ggaccctgtg agctgggaga 31560

aaaaggcagg ataccaacct gctctcaggc tccagtctgc agcagatggc ggccagcggg 31620

aagaaggccg ggggacaatc tgtgggaaca aacttctccc agaaaagctt cacgttgagg 31680

ccaaagtcca gtgttcgggg aaggcagtca ggatctgcat acacctgccc aatgatctgt 31740

agggtaagaa gacagctgtt tacaaaggac tacagaggcc acagatactt ggcaagggaa 31800

agggatgctg ggatgtggac aggaagaatg gcagaagagg aggtggtaaa ataggttaga 31860

gagctctccc caatccccaa tcatctctcc aaatcttact aatccagacc tgtatccact 31920

gagagctggg ttactgaaga ggactaagaa ggggcttcac tgcagaagcc aagtgttcca 31980

aatcagcatc tgctcctagc ttccactcat cctaccctgc cttgtctgtg tctccttgtg 32040

gctgtgtgtg tgtgccagag ttaaggctct atgaggtgct ggatctgatt catgtctgaa 32100

accctctggc accagcccag tgctctgaac agaggtgttc aattaatatt aaagaaaaca 32160

atttttaaa acaaatcttc cttcatggcc aggcacggta gctcatgcct gtaatctcag 32220

cactttggga ggctgaagcg ggtggatcac ctgaggtcag gcatttgaga ccagcctggc 32280

caatatggcg aaaccctgtc tctactgaaa atacaaaaat tagccaggcg tggtggcgca 32340

cacctgtagt cccagctact tgggaggctg aggcaggagg attgcttgaa cctgggaggc 32400

ggaggctgca gtgagccaag attgcacctc tgcactccag cctgggtgac agagtgaaac 32460

tccatctcaa aaacaaaaac aaaaacaaaa cattccttca ttatcctaac catcccttgt 32520

ctctcctctc ctctacccct cacaactttg taggaattac agaaggcttt gaatgtgtgt 32580

gtgtttgagt cgctgagata tgctgtcacc tgactaaact cctaaaggta ggaatctttc 32640

caagattcct cactctgtcc acgtggtacc tggcacaggg ctgacacttc ctagagtggc 32700

tacttctact tgaagggaac tgagaagggc tacatttggc ccaagagcta gctagacagc 32760

cacatgaaca gtctcttgcc cttccttctg gaattctgag tctcatctgc cccttatcga 32820

ccatgtgcaa gccactgccc tcttccaagc ctcagtttcc tcaactaaaa aaaaaaaaaa 32880

aaaaacctgg ccgggcgagg tagctcacgc ctattaattc cagcactttt agaggctgag 32940

gcaggcggat cacctgaggt ccggagttca agaccagcct ggccaacatg gtgaaaccct 33000

gtctctacta aaaatacaaa aattagccgg gcatggtggt gtgtgccggt agtcccagct 33060

acttgggagg ctgaggcagg agaatcactt gaacccagaa ggcagaggtt gcagtgagcc 33120

cagattgtgc cactgcactc cagcctgggc aacgaagtga ggctccatct aaaacaaaaa 33180

aaccctgcca gtccttctct taagaccaat gaaacacgaa cgtggcaaag atgtgcagat 33240

gcgttataaa ccataaaacc ctacacccat gttgatttta gtcaagctag tcatccctag 33300

gaggagatgc cccagttccg agggaaggca gagctgctag gcctgctgcc tcgggcatgg 33360

ccttggtgcc agagctcacc tcacagagaa cgatcccaaa ggagaagata tccaccgtct 33420

catcatagct ctttccttgg ggaacacagg agagcacact gttaagttta catcccttac 33480

ctgctctgca ctgggtctag ggtcaatcta agccacagct aagaactagg gctagaacag 33540

gttctgtagg gacaccccac aggctgtcct gcgcagacac cacctgaaaa tagtccttcc 33600

tcagctgtag gtgaactgcc ttggagagcc cttataacat aacactgcct aggatgaaac 33660

tgcacgttaa agtggaaaag aaccagacta gggttcatct agcctctgcc attaggcagc 33720

tgtatgctcc taggcaattc actgaacttc tctgggcctc agactccaaa atggattgta 33780

-continued

ttcagaccat ttctgaggcc tcacaaccct gatttcaaga atcccaacca gcatctctgg 33840

ttggtttgag cattatccac cagagggcac cacacttcag ttctcaccat ggagagagaa 33900

ggctaaagct gggacaatag cctgagctct tagcgatccc tgaggcacaa atgccacagt 33960

cagaactcac cttatccccc atgatgggac actgatcagt tcccagcccc ctgagctcag 34020

gagctaaatc gtgaggcccc acctggcaag ggaaaaatat cctcagctct gggtatccgg 34080

tcctacgctt cacaggcctc ccctggaatt cccagggctc tgatgcaagg gcagcatctg 34140

tcctccctct gcgggtgtcc cctccagggc ttcaggactc accgttcagc atctcagggg 34200

ccatccagta ggggtttccc accaccgtgt agcgcttctt gcggtcgttc ttgcgcaagg 34260

tgcgtttctt ggtggtggcc ttctccatgg gggccctttt cctctcttcc actatgagcc 34320

gtgacagccc aaagtctgcc accaccacag tcttgtcctg gtaggataaa agacaggtca 34380

ggacttcatg gggatgttct caaacctggc cctggggagg aagcataggc aagagaagcc 34440

actgagaagg ggagaaaagg tgtatagcaa gggagaatca gagcagaaag ctctatcaaa 34500

tagacacact caggccaggc acggtggctt ccacctgtaa tcccagaact ttgggaggcc 34560

aaggcgggtg aatcacctga ggtcaggagt tcaagaccag cctggccaac atggcaaaac 34620

cccgtctcta ctaaaaatac aaaaaaaaaa aaaaaaaaat tagctgggcg tggtggcaca 34680

cgcctgtagt cctagctact tgggaggctg aagcagaaga atcactagaa cccgggaggt 34740

agaggctgca gtaagccaag attgcaccac tgcactccag cctgggaaac agaatgagac 34800

tccatctaaa aaacaaaaca aaacaaaaca aaaaaacaga cacacttggc atttgtgttt 34860

ctgagggtct ggtatctgtg cacagcagac tcccagccca cagtgggttc atggcagagc 34920

cactggagcc ctggaggtga ggctcaggga cttggagagt ggaactcaga ggtgggacag 34980

ggggttgctt ttgtttttcc ctgtcctgtc ttattcctaa aatgatttca ggcagtggct 35040

agaaccagga taagacagaa aaaggcaagg tggaggcagt agatggggag acttgggggа 35100

gacaagctct ttgaatagct acagggcagt catgtgggaa aggaatgaag attgcgtggt 35160

gtgggtctag agggcagagc cagggccagc cggtagacac ccaggaagcc agatagcaag 35220

tcagtaagac aaacaacatt gtaaccacca cagctgaggg gcagtgggag gaatgctggt 35280

ctcaaactgg tccctgtaaa tttactagct cggtggcctc aggctgaggc ttcaggtgcc 35340

accactgtaa aatggggtca gcaatgaact gttttacaag actactatgg ggcaaaagca 35400

ggtaaaatga cagggttttc tacatatagt caacaaactt ctgaagtccc ttctaatcaa 35460

attatttcaa aggttgctaa tgagagaagg tgaaattatg tctgtgtgtg tctcggagct 35520

cccagaggac agagaagggc attggaagaa gcaggaaggg aaagatagga gctcctaggc 35580

aacacctctg ccctctagcc agaggggaca agtcagcctc ctttgtgaca aggaagccag 35640

gaaggatagg accctggagg ccaggcccag agcagtggga cataccaact tgatgaggca 35700

gttgtgcgag ttcagatccc ggtggatgat gcacatagag tgcaaatagg cctggaacag 35760

aagcatgagc tgaggctaca gggccagcaa atcaactaga ccctgaatct ggaattgggc 35820

gccatagatg ctctcctacc aggccatgcc cacccagctt cagaacccaa cctggacaac 35880

agggccaggg cctcctatgg cctctccctt gtcccaatac caacaggaag cctggacttc 35940

ctctgcttcc ctgctgccat ctcctgccca tcctttggtt caccaatcac tagcacagtg 36000

gctggccaaa gccctcacag ctctctccta atgtccccaa cttcagaaaa ccctggccgt 36060

-continued

tgcagcaccc cctccacctc tacccatttt acttagatcc aggctgctcc ctcctcccca 36120

ctctcccaat ctctggcttg agccactttg gatccttcct gcctgcaaca gagaagtgac 36180

ttcttgctcc agctagctct gtctttcaga cattcacctc ctctaagact ccaccttccc 36240

cactgtctaa actttaagct ctccccctc ctaccaaaat gtttgtcaac aaatttgttt 36300

gttatgtctc ctctccaggc ctccaccact gactcttacc actctcagga gcaacttcct 36360

ttctccttcc cttgcaaatg tagaaggagc cactccttgc tctccagccc cctgcaatac 36420

tactttcatc ccttaggctc ctggagctgc tcacacaccc agggagctct tcctgaatcg 36480

ctatagctca tgacctgcct gcctgcctgc ctgagggagt gctcctcaaa gactgtccac 36540

aactttccta tgctacctga tcctgagact cttgtgtgtc tgttttttc ttccagcctt 36600

ttcagtgtca gccctgctca aagccatctc ctcttcacaa gctgagacac ccccaccagc 36660

tacatggatg cctcttctgt cccacagctc cagctgcccc ttctctgcag attaggtgac 36720

tagcagcctc tccctaagcc agtctcatgg agccagtttc ctgcccagaa gatgctgata 36780

ggttctgcta gcacagaaca gcagggatca atgctagctc ttcccaaacc agttgctcaa 36840

caatcatggc tatcccccat cccccaaatg tctggtaagt gacacctcaa gccactcgta 36900

actcctcctc ctcctcctcc tgttccctca ttgccaacca gtctctccat tctgactcca 36960

cccggcctcc caggctcctg tccccacagc cactacctta attcagaaat ttgaaatgtt 37020

tcattagccc cctatgcagg gctccattcc ctggaccctc cgcattccac cccaatgtac 37080

taaccatatc agtctctctt ttttttttt tttgagatga agtctcactc tgtcacccag 37140

gctggaatgc agtggtgtga tcttggctca ctgcaacttc tgcctcccag gttcaagcta 37200

ttctcctgcc tcggcctcct gagtagctgg gactacaggc gtgcaccacc acgcccagct 37260

aatttttgca ttttttttctt tagtagaaac agggtttcac cagttggcca ggctggtctc 37320

aaactcctga cctcaagtga tccactcacc tcggcctccc aaaaaagtgc tgggattaca 37380

ggcatgagcc actgtgccag cccagactca ttttggaaaa accacagctc ttaagctttc 37440

ttgcttaaaa atcttttcct gccgggcatg gtggctcaca actgtaatcc caacactttg 37500

ggagggcgag gcaggaggaa tgcttgagcc caggagtttg agatcagcct gggcaacaaa 37560

gtgagaacct gtctctacca aaaaaaaaaa aaaaaaaaaa gaaaaaaaaa ttcaaaaaat 37620

taggtgtggt ggtacacacc tgtagtccag ctacttggga ggctaaggca ggaggattgc 37680

ttgagcccag gaggtcgaag ctgtagtgag ccatgatcgt gccactgcac tccagtctgg 37740

gcaacagagc aatatccttt ctcaaaacaa caacaacaaa aaaacctcct tcctgattca 37800

agccccagaa acaggaaaaa aaaacaaaaa aacaaacaaa caaacaaaca aaaaaaccaa 37860

aaagccagtc aaccaaaaaa ctcttttaca tttgtctcct atagtccaga ctccatagct 37920

taaaacacag cgaccttgat gtacccccaa actgttccag actactacta ttttcatgga 37980

cccccatgtg aagggcagtg gctaaagtct caggcttggc gaccagactg attagcccag 38040

gatcccagct ctgtcccttt ccagtccagc ttatctggaa accaaggttg tcatgaggaa 38100

taaacatgca catgcatgtt aatccctcag cccagtgcct agcatataga gggcactcag 38160

tacttgtcag ctgttattat ttctatgctt ggaatcattt gtttccaggt gaaatagatg 38220

attttctcct ctttcctccc aggtctttgc tcctgtcact tcaatcctgg tgcctctctg 38280

ccaataagaa ctttatccaa ttctgatcct tcagggctcc gctagaactc atttcccctt 38340

-continued

cattgatttc ctaggcagtg aggtgctcta gttggtctaa gggccctcag cacctgaata 38400

cctgtgatgg caagcccagg ccatctcagg ctacagtttg cctctccagt ttgccttgtc 38460

cgtgccctgt ttgcctttag agatgtgtaa ttctttgaaa acagggctgt gaactgagaa 38520

cagttggctg gggtcagttt cagtgacaca gacaaactca atgaatagct gctgacccag 38580

cccacacaag ttgcatagga aggggtctct accttccagt gaagcccaca attctcacac 38640

ctctccctac tctcgccctg ctggcaggtt tgttggtggg actcaccatt ccggaggcga 38700

ttcctttggc aaacctgacc ttctgctgcc aggggaacgg atcctgcaga atggggaatg 38760

attttcagag gctgttgagc cctgtcccac tcgtggtcca ggccccatgg gtcaatgtgg 38820

tctcccagtg aagcctccct gtgtatgctc agtcccaccc agcatggtcc tttggaggca 38880

aagcagggcc cctctgccct ttaagctcct agggtagtga tagcataggt gtctgacaac 38940

ccaccaaggc tcctggagac tatggggtgg tgtgctcacc atactgcgca gaaagtcctt 39000

cagtgtgccc ccctcaatgt actctgtcag gaggttcagc ttcttatcct tgtacagcac 39060

accaatgaac ttgagcacat tggggtggtc caggctgcgc atcactttca cctgggggtg 39120

ggggaagcca aggaaagctg gtggttaact tctcactgtc ttccttacct cttccctcca 39180

aggaatcagc ttgacaagtg ttaccacctg tatctcctac ctcctatctt gtgatatgac 39240

agggcactaa gatcaggatg agggaagacc accaccatat agatggagat ggcttcatag 39300

gccttaccct tggaactgca ttgggagaac tggggtggga gtggccaagt ggcttggagg 39360

gcaataccca gaccacaggt gtccccaact taagttactc catgaagtaa cttgggaggg 39420

gtagtgaaaa ctgtaaatgc ccaaggagtg tgagtgctca gggagctccc ctccaggctc 39480

ccttagcagg gcgttcaggt gaggattata acattgagtt gtactaaatt ggagtgtttt 39540

catgtccatc atcctcataa cctgagagct cttctgaacc ctgtatcttt tgttactgca 39600

tcccttgtac cctgtttgtg ttgcttttag taagtgttca ttgaacaaat aattacgctg 39660

taatttagag ggagtgggat attcctctag ggctcttttc tgcttagtct ttaagacagt 39720

aattccaaac tgtattttct cttgaagcca ttatttgtaa ggtcttctct cttccaatgg 39780

tgacaccaac ctcccgggcc ccctccatct tcttacctca gtcagaaaag ttttctgggt 39840

ctcctcatca catcgaatta actctttcat gaccatcact ttgcccgtgg ctttgtgtgt 39900

cacctacagg gagggtgact gttaatggtg ctgagaccag aggaagttct ggagaactgt 39960

ttgagaaaat ctatgcagct ggccatccca tcccttggag ggcacactga gccgggtcag 40020

aatgacagag ttcatgtgtg gaagatgcag cgtggagcaa catgggactc acacggaagc 40080

aaaccctctt tcctggagag ctgggcatgc acagaggaca ggcaagaacc cagaaaaggg 40140

aagtggggaa gaaggcagga ggggatgctg ccaagccaag caggggctct ggtcatttgg 40200

catagaagtg agtctcctga gtgaacagaa gtctgggggc aggaaggtca gtgccaaaga 40260

catgcaccag ggagctgggt ggatcggggg tgagctctcc agccctgctg cccccaagct 40320

gagacacacg tcttacctga ccccccctcc cagttctcag cagaagagga ggaggaggag 40380

agaaagagag aggttagtcc ctgagactgt tcaccaggga ggcaggagcc attcccctgt 40440

gtgctccttg tagcatgccc ggtactgagg tctgcaggca cctaacaaac acccactgac 40500

tggaggcagg agggccacag gcacagctca acccaacctg gatctccagg tccacaagct 40560

cactgtaaca gcaacattga ccttaggtgg cctgggtggg ggtgaaggat acctctcaca 40620

-continued

ccaacaccca aagccaggaa ggaggcctgg catcctgcat ggtgtggggt ggggccaagt 40680

gatgagaaat ccccgaaaga cagtgacaga gggactgggg gcagaagagc aaagcaattg 40740

ttgcctgcgc tcaccttgat agcctgccca aagaagccct tccccaggac ctccccatgg 40800

attaggtcac agggccggaa gatctgctgt gaatagctgc tggaacaacg aagggattct 40860

gagcggctga tgtcacggct gaacagcagg ggctcctttg gggagctggg gccaggggac 40920

ttggagatac tgttactgcg cctgaagatg aggaaagacg ggagtaggga tgtgtagtcc 40980

tcagagaccg gccaagaaac cctggaggga ggaaacttag gggcacctgt tcattccaat 41040

gtctatagtt tagatgttat tattggggaa caggggctct tacatcagat gtatcatttt 41100

ggggttgctg acatcagtca tttatttaag catagtaaaa acacattcaa aatgggaact 41160

ctgcatacat cacacataga acaattaact agattctact agctgataat gctcagtatg 41220

tgacatatat gacaacactg tagatggtaa ttcaccatct cccatatctg cttgatttaa 41280

gtactaactg gtttagagtg atagttatac tagtatccaa aaaatgagca aatgggattt 41340

ttttttttg agacagagtc tcgctctgtc aaccaggctg gagtgcagtg gtgtgatctt 41400

ggctcactgc aacctctgtc tccctagttc acctgagtct catgcctccg cctcctgagt 41460

aactgggact acaggcatgt gccaccaggc ccggctagtt attgttatta ttattatttt 41520

tgagacagag tctcgctctg tcgcccaggc tgaagtgcaa tggcacgatc tcggctcact 41580

gcgacctccg ccttccaggt tcaagcgatt ctcctgcctc agcctcccaa gtagctggga 41640

gtataggtgc acgccaccac gcccggctaa ttttttttt ttttttttt ttttgatgg 41700

agtctcactc tgtcgcccag gctggagtgc aacggcggga tctccactca ctgcaacctc 41760

tgcctcccag gttcaagaga ttctcctgcc tcagcctccc aagtagctgg gattacaggc 41820

gtgtgccacc acaaccgaat aatttttgta tttttttagt agagatgggg tttcaccatg 41880

ttggctaggt tggtctcgaa ctcctgatct cgcgatccat ctgtctcggc ctccctaagt 41940

gctgggatta caggcgtgag ccaccacacc cggccctaat ttttgtattt ttagtagaga 42000

tggggtttca ccatattggc caggctagtc tcaaactcct gaccttgtga ttcgcccacc 42060

tcggcctccc aaagtgctgg gattataggc gtgagccact gcgcccagcc taattattgt 42120

atttttagta gagatagggt ttcaccatgt tggccaggct ggtctcaaac tcctgacctc 42180

aagtgatttg cccgcctcgg cctcccaatg tgctgggatt acaggcgtga gccaccgtgc 42240

ctggccagca aatgggattt tggatgatat tttgcacatc tttgtacctg tccatgtttc 42300

ctgaatttta agatgataaa cacgtatgat ttttacagca ataaagtcac tatgtttaca 42360

ttttcaaaag cggtcgataa tagtcatggt caatgactgt tattttcact tactgaagaa 42420

ggaacaagaa ggttgaagtc cacagcactg caggtgaaca gccctgcgaa ctttctctct 42480

ggggtatgct cctgctttac ttgggctacc cagagaggac tgacgctgga tggctcttgg 42540

cttgtcccga cctacaggtt tacttaccct tgtccttttg ctctgcctcc tgctgaccat 42600

gcccaccaga gcacacaagg ctagctgccc ccctcctctg catggtggct ctacagatac 42660

ctggccactc ccaaatcttc ccttttccag gccatggatt tctagctctt ccctgccctc 42720

tgttgccctg cctcccacta atttcatcat actgaccaag gccaagaggc aatgggagtg 42780

agggcagcac tttaatgaga gggcactgtt tccaacacat gccagctgag agggctctaa 42840

cttactcagt ttgtttggtg aggagagtgg ataaaaaaaa acaaactttt ttttttggag 42900

-continued

acaaggtctc actctgttgc ccaagctgta gtacagtggt gtgatcaagg ctcactgcag 42960

tcttgaactc ctgggctcag gcaatcctcc tcctcagcct cttgagtagc tgggattaca 43020

ggcatgtgcc accatgccca gttaagtttt ttattctttg tagagacagg aatgtggccc 43080

aggctgctcc tgaactcctg gcctcaagca atcctccggc cttggcctcc caaagtgctg 43140

ggattacagg tgtgagccac agtgccctgc ccatctaatt aaaaacaatt attattatta 43200

ttttttttt gtagagacag tgtctcacta tgttgctcag gctggtctca aactcctgtg 43260

ttcaagccat cctccgacct aggcctccca gagtgttggg attacaggca tgagccacgg 43320

tgcctggcca agaacctgat tcttagtgat actaagaatg tgtcacgttt actgaacact 43380

gtccctatgg cagcactata caactttatg tccttttcca catttaatcc ttacataatt 43440

ggtgagacag ccacctttat cttaccatct tatgtcacag atgcagaaac aggctcagag 43500

agatgaagtg atttgtgaag ggaggagctc agcctgaagc cccagtaggg tatcagtatt 43560

tctgcctggg agccagtgct cttgacttct tttctgtaag tctccctact tggggcttcc 43620

actcaaggca acttctcctt ccccatggac ctggcagaat ctggggtttg ggggttgttg 43680

gtccctcaag tctaaaaaaa acgttcattc atcttttgct ttattgtttt caaagtcttt 43740

ctacacaatt tattcaatcc tatctgtgcc cctactccag gctctgagct ccatgctttc 43800

tccataggcc atgttgccct tcagggttta gcatcactgc ctcctgtcca gcttaggagt 43860

tcagtggaca gagactgact ctcaggtttc taaggaaacc tggcagacag ctgggtccgc 43920

acaggctctg gctgtaaact gtgatcccca caggggcagc tgggtcagca caatcctcat 43980

agtctgtccc tggccagctg gttcctttcc taactctgca ggcttctgaa agccagctca 44040

gcttcatccg agagacagac ataggacaga acagagccag ggtgggaggt ggcaccttag 44100

ggaacgtctc ctcagtgtcc cctccagatt ctccttggtg tccagggtgc tgagggcgtg 44160

ggggtgtccg gcattctgca tgtgaggagc gagccgggcc tccagccgca gctggtccag 44220

gcgttgggag acggggtcat gttcaatcaa cagctgaagt gtctggctcg tctggctaat 44280

tgcatcctcc acctgtggcc aagagcagta taaagctccc acaggctgcc ccatcccagg 44340

ccacccatgg ccaacccaga tgactcaggc acagtgtggc cagtgctcaa cttgtctaat 44400

ggtggccttt agctgccagt ggcctcacag tgccctctgc ctggcatcct gtcagctctc 44460

ccgtccatca accatccctg ctaccctggg aacctgttac gctgaaatag gtcaccagtg 44520

gtcacaggct caatgctaag acaaagagat gggctggggg taggaaaagg taaaggctac 44580

agcctgattt ctcagaggat ctgttccatg tcccaccctc acaagacaga ttagacacac 44640

actctacctc ctccactcga agtgtgcgga cgggggtccc attgatctcc aggatgcggt 44700

ccccagggtg gatggcgttt cgattgttgg gactgatgtg catccggttg accctgggcc 44760

agacaagagt tgaggcaggg taaagatgtc ctaaaaacct gcttgaccaa ttctgcttgt 44820

gcagggaagg tttaccaggt cccaggccac cagtggcccc taaggggccc aggataccca 44880

cacagaaagg caacaccttt tccaggatgc cccatctcag tgaagggaat tgccactcac 44940

cctgtggcct aagccagaaa cttgaacatc atcttcccaa tctggagttc cctcatctct 45000

cctactccca tctctaagaa ctctgccttc tttccatcat tgtagttatt cacatctctt 45060

acctggatca cggtaatagc ttctattttt tttttaaact ttgtttttaa tttttatttg 45120

agatggcatc ttactctgtc atccaggctg gagtgcagtg gcatgatctg ggctcactgc 45180

-continued

aacctccacc tcccaggctc aagcaattct catgcctcag cctcctgagt aactgggatt 45240

acaggcgcgc accaccatac ccagctaatt tttttgtact ttttgtagag atggggtttt 45300

gccatgttgg ccaggctaat cccaaactct ttttttttt tgagatggag tctcgctctg 45360

tcgctcaggc tggagtgcag tggcgtgatc tctgctcaca gcaagctccg cctcctaggt 45420

tcatgccatt ctcctgcctc agcctcccga gtagctggga ctacaggcgt ccgccaccat 45480

gcccggctaa tttttgtact tttagtagag acggggtttc attgtgttag ccaggatggt 45540

cttgatctcc tgaccttgtg atccgcccac ctcggcctcc caaagtgctg ggattacagg 45600

cgtgagccac tgcgcctggc ctaatcccaa actcttgacc tcaagcaatc tgcccacctt 45660

ggcctcccaa agtgctggca ttacaggcgt gagccactat gtctacgtct ggacttttct 45720

tttaattaag aaaaacagag acggggtctt gctatgttgg ccaggttggt cttcaactcc 45780

tggcctcaag caatcctccc acctcagcct cccaaagtgc tgagattaca ggcatgaacc 45840

accatgccca gcctacaata gcttcctaac tagtctctat acctttaggc tttcccctct 45900

ccaatccaac ctccatacca gtgatccttc taaaatgtta acttggtcac aacaccttgc 45960

tgcttaaact ctcaatggat ctgctggatt ttcataatcc aggccctcta tacttaacag 46020

ccccttttct ggccacactc aacctacatc agacatgatt aagactccta cttttggaga 46080

aatgcttctc cctctgccct gaacacaatt ccctctcctc ttcatctaat aaggttgtag 46140

aacagacctt agctcctcct gacttcccaa actctagtat ctttcacctg ctgtgcaagc 46200

actagcctag tgtccaacac atggtagata cacaatcaat atttacatgg taagcggcca 46260

gaggtacgac attactggta ttagatagac acagactctt tctccatctg gtctctggcc 46320

tgaatctagg atagggtctt aacttctttt gggccatgga ccacattttc tccattcatt 46380

ctctcgactg acaaaataat aatcaatcaa ctcctactat atgccaggta ctcttccaag 46440

ccctgggcat gcagcactga ataagacagc aagtttcttc tggaatttac attctagtgg 46500

aggggataga agataaacaa ctaatacaca taatcaagcc agatgatttc aaagagcaaa 46560

catgacctga aacctgacac ctgaaaggtg agtgattctt gccttaagta tcgggtagga 46620

agtagcctgg gcagagagaa cagtaagtgc caaggcccag gtacggttac agcatagtgg 46680

ggtccaggag aaagatatga gatgagaata aagggagttc tgcatgagac tttggaggtc 46740

taagccctct gaggcccatt tacagactct gctcaagaac ccctgctgaa gggttctcaa 46800

aatacttact ctttcacttg cacagtggtg gcgtagttgg agcaggcact ctccacggac 46860

acggagaagc cccgcctgcc ttcagtggtg gccggcatgg agatgagcgt gacagagtag 46920

ggcagctgct cctgaacaga ctctgtggag agtctctcaa acatgggtgc cagcaccacc 46980

tcattgtggc acttcccact ggggagaggg cacaggggg aataagatgc tagagtccca 47040

cacagaagcc ccccagcaac cactgcacca tgtggcttta agcaacagta cccatagtct 47100

cactaccacc ctctttgtca tgattctagc tggctgctca gaatgaggag cctgggttct 47160

cagggtgggc tgggactggg ggtgctgagc tagatctgca gcaagtggta tatttggaaa 47220

aacactggcc tgggttgggg tgagatggtc tgggttctct tcactctacc tttgactcac 47280

aatgagaacc ttgcccttgt actattctat tccatgggcc ccaggctgac ttcttttatc 47340

aaccaaggtg gttggacagg gtgagctcta aggggtcttt gactttaatg ttctgggatc 47400

gtccttgaag agtgcctccc actgggagga ttgcagggca gagtgaagac attttcaagg 47460

-continued

```
ttggtaagga aaaagggtag agacatgtgt aaatcattct gcttacaagt taaagagctg  47520

gttaagagca ctgctgtggt aagaatccaa tgctctgact accctaacac tccttctgcc  47580

caactcaccc tgccactgtc cctccccgcc agccactctt atatgggaga ggatagacaa  47640

aggaccacta tcttaccagt agagggtggc atgctgcacc agtgcatatg catccccatc  47700

ctcaatgatc accttgcagc tcatacaggc aaagcactct gggtggtact tgaactcccc  47760

agccacctgt gagcaggtgg ggagaagggt gtacaaaggg aggaaggaag tgagttatca  47820

ctgttgctgg agtgaggaca ggggtgggtg aatgacctcc aagcaaggcc cccgtcaaaa  47880

tccactcttt cactctctgc cctagagctt ccaaaggaag tgtcagattg gtacctcctg  47940

tcatttctga gactgcacaa aactgttttg gtctcagtta tgagaagaaa tccccagaag  48000

taatcatccc cactaaaaaa gatcagttct tccaatacag ggttggcaag tctgtcacta  48060

tgcttaaagt cgtttgtcca gagcaggtat aggaggtagc ccctctctct gccagcctat  48120

tccagagtct ccttttctga ttccctaaag acaacctaag tatatcctcc tctcagcctg  48180

taacactctc cttggaggtc cttatccaac tatattgtgt ctcctctctg cctgcctctc  48240

ctcccactcc caacgttcaa ctattaagtc tttgagggaa ggatcttacc tttcacttag  48300

taggtgctca attgatacct gctagttgaa tgaatgaata aataaaggga gatgccctgg  48360

ctgatggaaa gactaaaaga ttaacaggcc acagggaact ggaagcactg tcaatggact  48420

ctgaagacca agaggggcag atatgaaggg attcactcac cataaaaggc cctgtcatca  48480

gcagggagca cccatgacag aactccccaa acttccccca gtagtccttg gggcagtaga  48540

gcttcccatc cttctcatag taccagttgg tgagggaatc ctggcattct gaacacctga  48600

aaggcaagac aagaactgaa attctcctgg atccctggaa ccaggccaca taagctgtcc  48660

ctggttccta ggtatgacaa caaaatggac agtcctcttt gatccaaagc caagatgggg  48720

caaagttaca gccttgtgct atctccccaa gggaagctgg ttcagactga acaaggcccc  48780

agaaaccagc tctgtgttgc ctgcctcagt ggacagacta gtgtttcata acctctttgg  48840

gtcttggggc cctctgagaa tttgataaaa gccaagaagc ctcccctaga aaaatgaata  48900

tataaagaag atgtgtatat aatttaggtg ggggacaggg tgtatatata aaggattctc  48960

tgaaacctat ttatggcacc agttaaggat ctctggtctt gacagatact tcaatgaaca  49020

aaagcagttt aatcatctcc actcttcaaa actgacccat caccaacagc gttaccaatg  49080

taatgttctt gccaaaaata cgttaagtat aaaaaaacta atcaggcaaa tctagaatgt  49140

gggttattct ataagaataa ccggtcttga ctcttctaaa aagtcatggg gagaagaaaa  49200

ggcacagatt ttaaaaaccc tagagacaca gcaaccaaat acaatgcctc aaacttggat  49260

tcaaatttag attaaaaaag gaaacagcca tgaacatttt ttggaacaac tgaggaaata  49320

tggatttgga ctataatatt atattgtaaa attgccgtga attttcttag atgtaatgat  49380

atctgactta tataggagaa ttcttagcag aaacatattg aactgttcaa gggtaaaatc  49440

aagatgcctt taagttactc tgaaatagtt cagcaaaaat aaataaataa aagtatacag  49500

atagagggag agagatatat acaaagtggc ataacattaa caaatcatga atctatgtga  49560

aaggtatgta ggtgttcttt gaactacatt ttcaactttt ctatatattt gaatttttt  49620

tcaaaataaa atattgggga agaaacccaa cattctatta gcactgtgta agatgtacat  49680

gttagatcca tagctcgctg ggcggggtgg ctgatgcctc taatcccagc actttgggag  49740
```

-continued

gctgaggtgg gaggatcact tgagcccagg agttcaagac cagcctgggc aacatagtaa 49800

gaccccgtct ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaatcta tagctccctt 49860

ctcccctaat gacctggcta tgaagctcca cctctcaaaa gtctggtgtg ttctcaggac 49920

ttacctgcac agacaacctc ttcctccctt ctgagcttcc caacacacac agctaagagc 49980

tccctggcct gctgagccca tgccatctct ttccttctag ctccttcttt tccctattct 50040

tgttgtcaag gccttgcacc ttccacagtg gccacattct tgcttccaag aggtttgaac 50100

attgcttggt ttgaatcttg atcctgccac ttactagtta catcatcttg ggcaagatat 50160

tttacttctc tggcctagtt tcctcatcta taaaatgaac tgctgtgaaa attaatggaa 50220

ataatatgta ttaaaaatat agagagcagc tgggcacagt ggcttatgct tataatccca 50280

gcactttggg aggccgaggc gggtggatca cttgaagtca ggagttcgag accagcctgg 50340

ccaacatggt gaaacctggt ctactaaaaa tacaaaaaat tagctgggcg tggtggcatg 50400

cacctataat cccagctact cgggaggctg aggcaggaga atcactggaa cccaggaggt 50460

ggaggttgca gtgagtcgag atcacgccac tgcactccag cctgggtgac agagtcagac 50520

tctgtctcaa aaaacattat agctagctag ctagctagct agagatagct tatataaaag 50580

tcctaatata ggcagggcgc ggtggctcac acctgtaatc ccagcacttt gggaggctga 50640

ggcgggcaga tcacaaggtc aggagttcaa gaccagcctg gccaacatgg tgaaaccctg 50700

tctctactaa aaatacaaaa attagccggg cgtggtggcg tgcactggta gtcccagcta 50760

ctcaggaggc tgaggcagga gaatcgcttg aacctgggaa gcggaggttg tggtgaccca 50820

aaattgcacc actgcactcc agcctggaca acagagcaag attccgtctc aaaaaaaaaa 50880

aaaaaaaaaa aaaaaaaaaa aaaaaaaaga cccaatatag tagaaactca gtatgtgtta 50940

acttctatta aagatgatga caatgtaatg gcagctcact ttatcagcca cttactgtca 51000

ctagctggac aacttctctt agtagctttc tctcattgaa aataactaat agcatcctat 51060

tactacaaag agccaatgtg aagaggaaat aagctaatgt atgagagaaa gccttttgcc 51120

aggcaagtgc attattataa cttgctaagc actgcagata taaaagataa atacacaggg 51180

caccctactt caccтttcat gagcttatga agagttgagg aagagaaagt ctaaataaaa 51240

ctcaacaatg caaggtgaga ggccacagtg agttctacta cttctgagca ctgcgggagg 51300

tagagaagag gggcaggaat gagttgaagc agctgagcta gggaagtcta gagaagcaag 51360

tccctgagaa tccctttatc tttctggagt gcccaattcc tcctgggtgc tcaacaaaca 51420

ctgatggcta tttttaccct atagcttttc ctacttgtct tatggaaggg gaacttcagg 51480

gtcttaaggg ctgggtatgg catgaagagt atggctgtga agtgtctcaa aactgcacag 51540

tgaggttgga ggcagagggg aagactgttc ttggtatgtg gctcacactc ctctcaatat 51600

aagatacatg tgggatgagc tctaaatgag gaggagagag atatgggttc aagttcttgg 51660

gtgctaaatt gccttggaca agccccтttc cctctctggg ttgtttcccc acctgtgaaa 51720

caagaggaca aacaataccc tgggtcccct taactctgtg catatctgtg tttttcccta 51780

gatgatctga ggttctaatc ttagcaccac catttaccac ctaagtgatc ttgggctagc 51840

cccttcactt ctcatttgtg aaatggcaat gtcatcacct actttcaagg tagcagtgag 51900

gatttttct tttttctttt tttttttgag atggaatttt gctcttatcg cccaggctgg 51960

agtgcagtgg cacaatctcg gctcactgca acctctgctt cccgagttca agtgattttc 52020

-continued

```
ccacctcagc ctcctgagta gctgggatta ctggcgcatg ccaccacacc cggctaattt  52080
ttctattttt agtagagacg gggtttcccc atgttggtga ggctggtctc gaactcccga  52140
cctcaagtga tctgcccgcc ttggcctccc gaagtgctgc gattacaggt gtgagccact  52200
atgcctggcc tgcagtgagg attaaatgtg gttctgaaga actatggggt gcccaacaga  52260
gggctggcag taggagcagg ggcttccttc ttgcctcctg ccttctgact tagaagttcc  52320
tgggcaacca tctttcttct tcactagact gttagctccc tgaggacaaa aacagcatcc  52380
cacttactct tggtagactc agcaccaagc aaaggacctg gaagagcaaa gccttaacaa  52440
atatttccaa actgaaaatg gtaggagtgg caagaaccag tagcccaggt aggggagcac  52500
aagggttcca tggctcctgg tcaccaggtc cttcttttct tctgtgtgca gacttatttg  52560
cacccaaatg caggcctggg ctctgttacg cccttgcaat ggtatgtgcg ccgcagggca  52620
tgggagcacc agaacacacc ctcctttcca ttgccaatat cctgcttttg tttgctgcaa  52680
gtagctctgg ccaggcccag gagagggagc agggagacaa acaagtacag gcatgcagga  52740
gacagccagg ccagcctgct gggaagaaca cctggctgga gagcaacgtc accagtccac  52800
atatctgtgt caaggaatgc ctttaaaggg cacagccagc tatgtggcca aagtcactgc  52860
cacacatgtc tctgaagaac tggaaggtga acagctgaca ggagggctgg gtatggatcc  52920
ctaagctgtg ttttgccaca gaacacaact gggttctcag accacttaag aagttgccat  52980
ggtagacttc agagttctag cttagctatc aattaacagc tactatttat tgaatgccag  53040
ccataatcca ggtcctatat gaggtacctt atttatatta tctctaatcc tgacaaccct  53100
acggtagcag gattatccct gttttacaga agactgaaaa gcagagttaa agtcaaatca  53160
ttctgcaaag ctgcccaata cataattgat ggagctggaa tgtgaagcta tgtgtgcttg  53220
gctgcaaagc cacatccact gtactatgtc accggagtga gacagcacat ctcaacagct  53280
gagcctcagg attctcacct gtgaaaataa ggaagctgga gccctgacag ctcaagaccc  53340
tttgagtgtt aacacttcgt gatgcttctc ctttagccca agtactgaat gttaaatagc  53400
acaattctcc tccttcatat acttcacagt gttcaattcc ccagcagtct ctgttctttt  53460
aaacaaagaa cccagaacac actttggaac tgtaaggtgc tactcgaaca tgtcccacct  53520
cagatgctta gaggccatgg acatcatacg ttcttcccaa agccaaggct gtttttctac  53580
ctttcaattt ttggaagttt caagtgaaag tcaaaatctg ccctggctgt aggtagaaga  53640
gtacccagct gggaggtggt aaagctgagg tctgaaccag tcatcccact gtcctgagga  53700
gccagagtga attgagctga aatcctcagg ttaggggaga tcacctagag gcctacatac  53760
gctgaggcta tgcctaagga caacctccaa aactcatctt ccaaaacatt actgaggggt  53820
tctgagggct tgggcagata caaagtgcat ggaaccctga tggtcagttc agtgggtcaa  53880
cccagatcca cacagagacc ttttgagtca gcagagaaca tatgagtttt ttaaaaaaga  53940
taatcagttg agattttctc tttcaagacc tggatccata ttccttctct gccatttacc  54000
atctgtgtca cccaaggcca gttacttcat gactctgagc ctcagtttcc tcacttgtac  54060
attgctaata atatcttatg gggttatgag aattaaatta tgtacagatg tcttgcatga  54120
tgcctggctc acagaaggta gccagtaagt gttcatttct ttccctcccc aatttgagac  54180
caccaagggc aaatgcagac agaccccttt ttctcctatg cctagcaata tggacctgct  54240
tctcagccca gaatgaaaca gatttctgta agtgaaaaag aaaggtccta aaatgacctt  54300
```

-continued

```
tatttatcta tccaaataga aatcaacttt ttttttttt ttttaagaga cagagtctta  54360

ctctgtcacc caggctggag tgcagtggca caatcattgc tcactgtagt ctgacctccc  54420

aggctcaaga gatcctcctg cctcagcctc ctgagtagct gggactacag gtgcacacca  54480

ccatgcctgg ctaatttttt atttttgta gagacagggt ctccctatgt tgcccaggct  54540

ggtcttaaac tcctggactc aggtgatcct ccagccttgg cctcccaaag tgctgggatt  54600

acaggaagga gccaccaagc ccagctgaaa tcaattgatt ctaatagatt tctaataatt  54660

aggaaatagt tactgtgcat ctaggatgtg taagaacaag caggcaaggt gctgtgctga  54720

gagaagccaa gtgagataag atagaggccc agccctcact tacatccaaa ttcttcagac  54780

aaaaatgttt acaggtaaat tctcaacctt atggggtaaa agaaccagtc attattgcag  54840

tggatttaca catgttaaag cataagccct cccatgtcct gaggaaggcc tggctggatg  54900

gtcatccctt aacctcccag tgaacagcca gttgtagaga agcttctttc agcccctagc  54960

ctctgaattg agtgggctga aggagagaat gggaagctcc ctccaccaca gcctgagggc  55020

acctggagag caccacaact ggggatgcca acttttcccc agggcccaga agggcagata  55080

tgacaggggc ctagagagac tctgttccat aagagctcag tcaaaaatcc ctttattact  55140

cctccttcct gcagatccta tagactgaaa cagtctgcca attaagctgt aattattatg  55200

gaatatgtta ttgttgtttt aaaaaattaa ccaaacacat aatgaagcct acgaacagca  55260

agaaataagt gttgctagag tgcactgatt cagttccaac caaaaaccag gaagactgta  55320

gtttcaagta ttttgtgtag ttgtactgca gttaaagaca gggcctccat tatttggcaa  55380

ataatggaaa agagcttgtg tgcctctgag gagccacata agttgcagag ctgggaacca  55440

ttcacctgat ccaagtcctt catttttcag ataaggaaaa gagtcccaga ggacctaagg  55500

cacctgccaa aggttactgt gaagctgggc agagttatgc ataatttttt ccctttaaac  55560

atggattgct ttggaacaca aaggttcatt tgttttcttc aagagataat gtgtaaaccc  55620

aagctcatta aatcagtatt ggacattaag tgaaattcta taaaaagaaa aatgctttaa  55680

agatgtccga caaaagccac tccctaattt acctcttttg taaattccaa tttgaggtat  55740

tttcttactt ggaagcacag tggtaattgc agctcacttc tcacgctagt gtggtaggaa  55800

agacagtgca ttctctgctg agtatgtatg gcctgtcagg caggggtgta tgggacacag  55860

aagttgtatt cttcattccc ttgctccctc ttcctccctc acatgtcacg cctgatcatg  55920

tccccatgga atctggttgc gtcactttcc ttgcaaccac agccccagaa ctcccataga  55980

ctttgtactc ccccagaggc ctagggtcaa gaacttaatt tgtttcacag aatccagcaa  56040

ggaccctgtt caggagccca agaagtctag cttctacttt ctgccaagga tgaaactgag  56100

gcccaaagag gagcccttga ggatatccta taataaaaga gggaggggct aatggcagaa  56160

acagctttct gaccaggtag tgatgtggta acggatacct gggagagcag agggtctggg  56220

tggtccctcc ctgggggggaa agagccctgt gtggcagcta tctcttcaga agtgacttca  56280

ggaatgtggc ctcttatttg gggaagggaa gaatgtctac gaaagtcagt cagtaccccc  56340

cagactctag ctgggcaaac ggagcttggg ggtaggggca agaagatcag gttctccaga  56400

aagcctggga cagaacaaga caaacggcag gtggacttgg gtggcattag tggtctggcc  56460

cttaaaggtt caggttgaga gttcactatc tgctctagct tctcctcttc ctgagaagtt  56520

gtgtcctccc tggcttccct gcctctccaa gggtaaaagg ggtatagttc ctggctccta  56580
```

-continued

aaggaacaac ccttctccta agctgccttc ccaggattaa gagaaattaa tgaggacctg 56640

aaatgtggct acttcaacac ttggcatctt aattccacat gtaaaagctt tctcctaaac 56700

tcaagatcaa tttccattgg ctctttagaa aacaatctca gccacacctt caacaagtca 56760

ggcaaactat ccctttttt actctctgct aggagaggaa aacaaacctg agaatttcct 56820

ctccactagg agaagaaaac aaacctgcag ccccctgaag gatgagagaa agtcaggaag 56880

acagaaaaag ctgcccagtg aggctccata tgttaactct gttgaaactg gttgcctcat 56940

ctgtaaaatg gggatatata atctctttag gctgttgaga gattaaagga taaagtgtct 57000

agcagggcgc ttggtagagt cagcattgag gaccagttcc cttccccctt ccccctgtcc 57060

agacacttcc cagcaatcta ccagggtccg gtcattccac aaaaccacat ggtgcactgc 57120

ttttctgccc agggcctggc cacctgacta gtatcccggc agcaagatct gggcactgga 57180

cctagcttag agaagccaga agggaaaagt cgaccaggac cactggtgag ggggaggcag 57240

ccttttgga acaacaggct gctcccaaag aatgaaaaca aaccccaccc tagcctggaa 57300

tggctccatt cagaagggac ccagagggcc agaaaccccc tggggagttt ctatgggagg 57360

tggagtgggg caatgagcaa agaaagccag gttagaggta gagtttccac atctccctct 57420

gagaagcccc acagagagag gctcgtggag gtacttaggc ctagggactg aggaggtgtg 57480

ggggctggcc cagcagctca ggaacactct aaagcccatt gctccagtct tagaagcact 57540

cgcccaggtc tttccccagc aactacattc acctgggagc tcagaagccc ttcaaggcca 57600

gaggcaaatg tgcctgccat taagagaaca aacgtttata aggtgcctac tgtgtaccag 57660

gctacccaca cccaggactt gccagtttcc acagtttggc aaaacctgcc catcctgctt 57720

gttgggacaa attcctgggc aggcagcatg gcccgcaggt aagacatgtg agagctggaa 57780

acagcactct atttgggcta ttccccccgg aagcagtaac atctacaatg ggcaggtcag 57840

tggcatcttt tggaggctcc ttcaggtttg tctgggatag aaggggagct ggagctccct 57900

aagaagttcc ttctgggtat atccctcccc tcaccacttc caggacataa tgggggtacc 57960

tcacaaggcc gattctcctt gggtataacc tccgctggct tctgcccggc agaaccgatg 58020

gtaaaggact aagtgagaag gagatttagg gagaccaccc tcctctcagg accaggggcc 58080

aggagagatg gagaggggaa ctcggagacc aactcaccga aacaggtctc tggaggtgaa 58140

gtaagccggg actgacaagt aactccccat cttcttccca caccagatag acaggctcag 58200

ctgcaaatcc gtctgtgact gcctcaggcc ggctgtagat ggccaccatg acaggcccgg 58260

ggctcccggc acctctggct gcaggcaggc tgccctgagc tcacctctcc ctccggagca 58320

gctgcgggag cctgaggact ccaggccact tggcccccag cagattccca caggggctca 58380

tggctgggat ggctcctccc tagttctcag aagctcaggc ggcccctctg gactccgaga 58440

aaacgacctc tctgtggagg ctacacacac actggcattc ttgagcaggg gccggtgact 58500

agctaagaaa gaagcgctgc tgtcaagtta gccttcttct tcagtcccac ccctgaatga 58560

gggagggagg gaggaagagg aggggctccc gcagaaccga ggtgtgtggt tggagcagct 58620

gttaggaaac agaccttgca ggcaaggggt ggagctgggc tggatcgagg ctgcaggcag 58680

gcgctgcaga caataactgg gaccaattga gcgtccatcc tgagctcagg gaagggcaag 58740

gcaggataga gtgtgaacat atctggcccc agatgtcctc aagcagagga gctgcctaca 58800

ggacagtaat tctgcttgga ggtggaaaaa aaaaaaaagc agaggatgcc aggtccagga 58860

-continued

agagccacac cagacggggg gacctctggt aggagagaca ggtccagaga gccatcctag 58920

ctctgagctg ccatctgttc tgtcctggca ctagggctcc cattcattcc ctccctctgc 58980

ctgtcctggg gacagcctca agagctgacc aatgggaggt gccacttcca ctccagggca 59040

gggctgtgtt gcaaataagg aaacaaggcc cagagtatta acctacacct gatgtgtctc 59100

catgtggcaa acactatatg cactctttta tttctctgag acaggaatta ttcccactct 59160

acagacaaga aaagtgaggt acagagagat taggtaactc actcaaggtc aaaaagggag 59220

tcagtggcag ggctggaatt taaatgtagg ctcatgtgac tctagagctg gagctctccc 59280

actaccaggc ttcattctgg ctctttaagg agaaaaatct tttccaaagc tctctgcaaa 59340

taggatggag cccaggttac tctcattcac tcactattca ctgagtgttc acatgcagaa 59400

caccatccat tggcctcagt ggacagtgga gcttcttcta cttgtcctgt gataaccatg 59460

cttttcaaac tgaaaaatgg ggcatgaggg agaaaaaggc ctggggtgcc aaaagggatt 59520

gataagccta ggtatctgga atgggagccc acctgaacag cccccagttg tgacgatgac 59580

caaaccagag cctaccatgt agaaagccct tgtctccctg cctgcaggat tgtaaaggaa 59640

aaagcaaggc tgctggagga aaccagaaca ctcagatctc actactggat ttggggcagt 59700

gacatcagct ctcccttacc aggcacacac cgctcaccat tgtaaagagc cagaagtcac 59760

cgtgcagctt ctatctctta ctgtgaacca tggcaggaaa gtctgataaa gatggtagat 59820

ggggttacaa gggagccttt tcccattcca ccaactggac acagggacaa caatgcatgt 59880

gataagagca catgggtttg aatactgtta agaagctacc atttattgag cagttaatat 59940

cttctaggcc ctgtgttaag tgcttagcat ttattatcca tcatgtttaa attgccagta 60000

aaaggaggaa acgtcatcag cctcatttaa cagattagga aataggttct gaggggttat 60060

gtcactattc tcaagacaaa tagctgataa gtggtgaggg tgtttttttt ttttgttttt 60120

tttttttga gtcagggtct ctctctgtca ctcaggctgg agtgcggtga tgccatcatg 60180

gctcactaca atcttgacct cgaccttctg gactcaagtg atcctctcac ctcagcctcc 60240

agagtagctg ggactacaga tgcgtgccat cacacccagc tgtgaggctg agtttcaaac 60300

ctggtacata tggctctaga gcacaggtca aggcactttt tctgtaaagg gccagagagc 60360

aaatatctta ggctctgcag gccatacaga ctgtcgtggc cactcaattg tgtcattata 60420

gggcaaaagc agccatatat aaaacataaa ctaagggaca tggctatgtt ccaacagata 60480

tttacttatg gaaagatttt tttctttttt tttttttttct tttgagattc agtctcacac 60540

tgttgccctg gctaaagtgc agtggtataa acatggttca ttgtagtctc aaccacctaa 60600

gctcaagtga tcctcctacc taagcctcca gtgcatcctg ggaccacacg tgctggccac 60660

catacccagc gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgcgtgcgt gcgtgtgtcg 60720

gagtctcgct ctgttgccca ggctgcagtg tagtgtcgcg atcttggctc actgcaacct 60780

ccgcctccca ggttcaagca attctcttgt ctcagcctcc caagtagctg ggactacagg 60840

cgcatgccac catgctcggc taatttttgt gtttttagta cagacagggt ttcaccatat 60900

tggtcaggct ggtcttgaac tcctgacctc aggtgatcca cttgcctcgg cctcccaaag 60960

tgctgagatt ataggtgtga gccactgcgc ccggccctgg cttatttttt tttatttttt 61020

gtagagacag ggtctcactt tgttgcccag gctgggacac tgatatttga gtttcatgta 61080

atttccacat gtcacaaagc attcttttgc tttttttccc ctaaccatta aaaatcacag 61140

-continued

gtcatatagg ccaggcatag tggctcatgc ctataatccc agcattcgag aggccaagga 61200

gggaggactg cttgaagcca gaagtttgag accagcctgg gtaacaaacc aagaccctat 61260

ctctactaaa aaaatcagtc aatcaggcca gtcatggtgg ctcacgcctg taatcccagc 61320

actttggaag gctgaggcgg gcagatcacc tgaggttgga agttcgagac tagcctgacc 61380

aacatggaga aaccctgtct ctactaaaaa tacaaaaatt agccaggcgt ggtggcacat 61440

gcctgtaatc ccagctactc gagaggctga gacaggagaa tcgcttgaac ccaggaggca 61500

gaggttgcgg tgagctgagt tcgtgccatt gcattccagc ctgggcaaca agagtgaaac 61560

tctgtctcaa aaagaaaaaa aaaatcaatc aatcacaggc cgtattaaaa taggcaatgg 61620

gccatagctt gccggcctct gttctagaac cacagagctc aaccactctg tttactacct 61680

ctaatagtga agaacagctc ttggtgtttg tgcagagctc agccatatct atctatctat 61740

accacgaaga caccacactc tagccccaga ggcagccaga agaaggaagc actgaacagg 61800

tgctcaccct gtattaaaca ttgcttgcag cactacttgt atgaatgaat tttaaccccc 61860

actattccac agtgctgtga tttttctctt tacagataaa gcaactgagg tacagggagg 61920

ttaaatcact tgcctgaggt cactcagctg taaatgtgag aagtgaagcc aggatttgaa 61980

cccagactag acattataga caacgttcta ccatgctata ccaaagctca gcctgggccc 62040

agagccagac tgtctgggtt tgaatgcgac aatgcactgc atgtaagccc ttggcacaaa 62100

gtctgggaca cattaaaggt tgataaatgt tagccattaa tatcattaag actattgtta 62160

caactactgt gggctagcta tgtggccttt agcaagtcac tctctctgtc tgggccttgg 62220

ttaggtctct aaatagctcc cagcatcaac atttttcacc cctatgaggt tgcagagggc 62280

tctctatttt ccttgggtct caagttgagt ttaaactctt gagcaggagg tcaggatcct 62340

agccgagggg acacacaagg gccacagagg aggtgaggtg ggctggaagt ggggaaataa 62400

catacgggat actgaaaggg ctgcccagct ggagagctgg gagcaatctt tctcccagcc 62460

acacagacaa ctatttctca ccaaaaatct attatgttcc cctgtaattc aatcaggcaa 62520

gaaggggcc ccaaattaag tctgggacta gcaaattatc ttttaattgt accctctgat 62580

ttccatgtga tgaacacagt actaatggaa gaatatgcaa gccacaaaga gactgaggcc 62640

agggagccta tgagcataac tcctcttgac atttcacaca gctctgctgt tccagcggcg 62700

acgggctcag tgagtcggtg tctaggaatg tgagcaggag gctgcaggcg ctctcctctc 62760

cccctactcc tcttacaaag gacttggtgg gtgggtgggg aataagtgag aggacatctc 62820

aatgccatga aaactgcctc cccctggaag aggaacagcc ttctccccat gtcctgagtt 62880

gaagcagaag cccctcccag gtggaatcca tttgacctcc aaaaattcag acttcaggag 62940

gaacctttga aggattgatc accagatgtg ataagctgct ggctgctcag acccaagtgt 63000

gtggttcctc agagccaact tcaaagggcc caagccaccc atcactggga gatgaaatgt 63060

cacactggag gatatatttc accctcccac ccttggaaac tggtcctgga tagaaaactg 63120

ctcttctagg tcctcagagc agctggcaga acattcatca gttatggctg ccagaggcaa 63180

ctgcaggaat cctttgactg gcaattatag tagggataaa atggcaggag aggtgttaat 63240

ctccagcaca tttcataagt tggaaacctg gcacggagct gaacaaactg actccttaga 63300

accagctgag gcacaggtct tcaaagcgtc tgtgacaggc tggagaacaa tagccttcat 63360

aaccggcagc cccaagtgaa tgttcaggca aggtgtggta agtgaatgaa tcaccccaac 63420

-continued

```
cttttgtcgg aagaaagctc tggtcaaaca attcaaaacc caagaaactg tttgtgtgaa  63480
gccctctatt ggctgggagg actccaatca aggaagagaa tttttgagga ctccgaagcc  63540
tccaaatact gaaaggtaag cattaccacc cccatctact aaccatcttt cgtgccccac  63600
tcctcccctc ttccaaatca ctttctacat tatagcagca gtcatccttc taaaacaagg  63660
gtcccgggca ggcatggtgg ctcatgcctg taatcccagc actttgggag gctgaggcag  63720
gcagatcact tgaggtcaag tgttcgtgac caacctggcc aacatggaga aactccatct  63780
ctgctaaaaa tacaaaatta gctggatgtg gtggtgtgtg cccgtaatcc cagctactgc  63840
ggaggctgag gcaagagaat cgcttgaacc cgggaggcgg aggttgcggt gagccaagat  63900
catgccattg cacaccagcc taggtgacag agccagactc cgtctcaata aataaataaa  63960
taaaatgtcc cagcaattcc cagcttcaca gggttatgag cattaaatga gaacatatgt  64020
gaacaagctc acttcagtat ctggctcatg gaactcagtc agtgttttac agagtgagag  64080
tttagtatct gctactcctg taatgtgatg tcatgttcac aaacaccagg taaaagacaa  64140
tatttggcat ctatcttagg tatcatggaa agagaaaaga atttggcatc tggatagaac  64200
tttagagatt tagtctaatc ctttgtttta cagttcatga acttgaaact aggaatggag  64260
aattggaagc ccaagacata tgaggatacc aagtcacatg aaatgtgttc agattttcag  64320
actgggagaa ttatttccca atatattgaa aaaagaaaat atgttggcac agtgatggaa  64380
ccattgttag taactttgga gaaactggag aatgggtgca atatctccaa agtgtaatgt  64440
tccatatcta ctatatacca ggcgctgggg atgtattcat gaaataaaaa tggagatagc  64500
taactggccc agttttcaaa aaggaggaaa aacatactgg aaaaaggaag ccttatgatc  64560
cttttgcttt gattccttca cataaattct aggcagatga tggaataagt ggacagtaag  64620
cacttagaaa aggagctgat catcaggagt cagcagtagt gaaatagtag tgactagtta  64680
gacttgactt catgttttca ctgggttact catcatgaga ccttatcttg gatttccaca  64740
ggcatttgat aaactctttc ctgatatcca gacagtaaga atgtgggcag gaaatctaga  64800
aaaatcagat acaagaactc aacataccca gcaatttgag gctctaaact gtagaggcag  64860
gattcctgaa taccttgcta gagctcacat tgaggaaagg gcatgggatt tgagctcaga  64920
aaacttttct tttaatttca gtgctgccct tcccagcagc ataaacttgg ataaattatt  64980
taacatctct gggcttccct tgtcctgctt gtgctatatt cctcaagagg ttgttgcagg  65040
aactgaacta gatgatgagt acaaaagccc tctgtaaact gaagatgcat tataaacatt  65100
agttattatc actgttctag tgctatagtt gctcatcatg ggggtcctga atcataggtt  65160
tctttactag gcatctgtgg acacctggta tgggccaggc ctatgctggg cactgacagt  65220
acctagtgcc tgccttcacg tagtttgatc tggtgataga cagatatggc cacagggtcc  65280
ttcctgggtg tgatgtgctt ggagacacat tgaaggaacc cttcccagag gagaatgaat  65340
gctgtgtcag agcagggcct ggaaaaagga ggagaagaca gcacctgcag gagatgtagg  65400
aggctgtaaa aggagttaag agagcaaagg tttgggggac tcccagggag ccttatgtgg  65460
tgagtcatca atggagtaca gcaggaggtg accagagccg ttcctgaggg aactaaggtt  65520
cacagagaag gttgctgttt aaggtcacag gtctaggaac tgacaaagct ggagtttgac  65580
ccaagtcttt ggactccagg tccaacatac taaaactgcc tcaaataact attttattct  65640
tgccatttga tttacagtaa tatttctgaa agtgtaacta aaggaatagc agcattggac  65700
```

-continued

cctctggtca gcttgttgaa aatcaagatt cctgagtccc acttcagata tacagaatca 65760

gaatcgctga gactggggct aagaatctgc attttaaaca agccccagtg tttctgatcc 65820

acactaacaa ttgaaaacca atgcttttct ttttcctaag tgcttttcca cttgatttca 65880

ttttattctc acaataaccc tcagaggcag gcaggataga tcactatcat taccttgtta 65940

caaacgagaa aatgaatgta cagatagctt gttcatcttc ttatcctgaa cgtagcttaa 66000

taaataaatg tttggtgatt gacattttgt taagccttgg ctaagtagag actttcacct 66060

tcctggtgcc actcttaccc atttgtaacc ccacagtagc tgcctctctc tttccttggt 66120

gacgcggcag tggtgctcac gttatggtat aaaaaggtga ttctaggcca ggcgcggtga 66180

ctcacacctg taatcccagc actttgggag gctgaagcag gcagatcaca aagtcaggag 66240

atcgagacca tcctggccaa cacggtgaaa ccccgtctct actaaaacac ataaaattag 66300

ccaggagtgg gggtgtgcgc ctgtagtaac agctcctcgg gaggctgagg caggggaatc 66360

acttgaaccc ggaaggtgga gattgcagtg agctgagatt gtgccactgc actccagcct 66420

ggcgacacag cgagactcca actcaaaaaa aaaaaaaaag tgattctaga agatagcgtc 66480

attccatgaa agtgctatct ctggaactaa aacaaatttc aaacaaacaa acgtgaggga 66540

tgacattagc caacttcaga aacggtgctt ctgcaagagt cagtgctact gatggcagtg 66600

gagcaccctc tgatctcagc aaaatctcta tagggcatca tcaggaagtc atgaccagaa 66660

tacagctgac acacaaaaga aaacaacctt aactcaagtc tggcatgaga tgctgagggt 66720

ttttttgta atcgtaagca gtatggagaa agaatttgtt agcacacttc cagcactgcc 66780

catgggagaa gggactctag cttttggggg aaaaatgaaa aacacggcaa aataagcaca 66840

gcaggttgct acagttgcag aaggcatgcc attccttgag gccctggttt tgatggttta 66900

agacttaggg atgacattta cagactgagg tcttagtcta cataagtttg agcttatctt 66960

atcacctcaa aagataagct cttgcctaaa agaatattgt agagcgaaga gatggtaatg 67020

gtgcatactg gatggacttt attcaagtgt ccaaagagtc gttcaacaaa tacatcattc 67080

acccagcaaa cactctgagc acccactaca ttctaggcac tttgctaagc tctgggatac 67140

aaagacaaac tagatacggc tccacacttg gagaatttca cactctaaga gtgatgctag 67200

atgtgtaaac aaatactgaa atggatataa ctgaggtatg aagagaataa cgaagagaat 67260

aacaaagaga gaagaagtct ctgattggtg gggagaggta tgggtggatt tggcgggggg 67320

ctcctccaca gaagatggga tggtattggc agactctaat aggctgattg ccaggcagta 67380

acaggcattc tggcctgaag gaacaacatg tgcaaatgca ctgaggcata aagagtatgg 67440

tatgctcata taacgctgag caggtctgta agctctagtg catggcacat atgggcagca 67500

attagacagg aggccagaat ggtggaagag agtaggggca gagcactaaa gaccaagtct 67560

gctaggctaa agagttgcct acaggcagtg aagagctatc agaggtattt aagcaagcag 67620

tgatttttaa cagaacgcaa agctgaaatg ccaaagagaa tgctgatctt tagttgccag 67680

tttggctctc tgaatgggta taaatgccta aagtctacag ttggccgtgt atatgcccat 67740

gtgaaatttg ctggccaact agtgttcact aacagcagag tatgcctttg tcctggggcc 67800

acatctctac tgccttgtca ttttaacagc tgatatttgt ctatatgggc tgtaagcaca 67860

taagcatgga attgcaaccc cttatagata atcatccact gttgctccat gatactaatt 67920

agaaaattga atgtctaaca ttgagtaaca aatgctgtgt tgacagggac atccttctcc 67980

-continued

aggctgcact gggaaaaatg cttcggatat ctcaagaaat ctctaggcaa ttcaaggaaa 68040

aaagccctcg tctggcaggg aacacctgaa aaatcaccct ggttaacaga aaatctgcgt 68100

gtgatcttag gcaagccatt aaatctctct gggcttcaac ttcttcctct gtaaaaggaa 68160

ggagttaggt gagcaaaaaa ttccataccc attacaggtc aatgcttaaa gaagcaacga 68220

ggaaggaggt gtgatcaaag gaaggttgac tggtggacca cccagagaac tcttcaaggc 68280

tgatgttttc ttggtgaatg accctaatct aggcctagct tttggccttt caggtttttg 68340

accttcctaa tgtccaaaga ggtgtagaca ttctgtgatg gttctgacct aaccatttat 68400

caaaccctct gtacacttat gaattgacta ttaaaatcag aacactcata tcccctatac 68460

aggtactact tccagaccct tcctgttatt aagtatccca ggaagctgaa ttaggcatca 68520

gaaggttgct ctgtgagcat ggacttaaaa taaagggaaa tgagtagcca ctgaaagaaa 68580

tttcatcctg gcagtggtga gtcagccaaa ctagcccagg ttgccacaaa ggaccctatt 68640

ctggtgacca aatctagatt cagcatcttt atctgaatac agtactagtt ttgcttcact 68700

aacttttcac tttctgtttc attgctggcc acttatttat tatattggtt ttggtttttt 68760

ttttttttt ttgagacagt gtctccctct gtcacccagg ctgaagtgca gtggcacaaa 68820

ctcggctcac tacaacctcc acctcctggg ttcaagcaat tctcctgcct cggcctccca 68880

agtagctggg attacaggga cctgccatca tgcctggcta atttttgtac ttttttttt 68940

ttttagcaga gacagggttt caccatgttg gccaggctag tctcgaactc ctgacctcaa 69000

atgatctgcc cacctcggcc tcccaaagtg ctgggattac aggtgtgagc cactgcgctc 69060

ggccttatat tggtttttaa actaagtttt aattgaaaaa aataactact tctttagata 69120

tgtgcgacaa tcaccaattc acacatatta ttcagccctg gcagacattt gcctctctag 69180

taactaataa caaagaacaa acattcatga gtgctctgtg aatagtgttc ccagtgctac 69240

acacactttt attcacagta atactgtgag ggtaggttgt aaggaggtag gtagtaagtt 69300

tgtatctcca tttcttcttt ttttttttt ttgagacgga atctcactct actggagggc 69360

agtggcgcga tctcagcaca ctgcaacctc catctcccag gttcaagcta ttctcctgcc 69420

ttagtctcct aagtagctgg gattacaggt gcctgccatc gtgcctggct aatttttgta 69480

attttagtag agacaaggtt tcaccatgtt ggccaggctg gtctggaact cctgacctca 69540

ggtgatccgc ccgcctcgac ctcccaaagt gctgggatta caggcgcgag ccaccgcgcc 69600

caacctccat ttttttttt ttttttgaga cagagtcttg ctctgtcgcc caggctggac 69660

tgcagtggcg cgatctcagc tcactgcaag ctccgcctcc cgggttcacg ccattctcct 69720

gcctcagcct cctgagtagc agggactaca ggtgcctgcc accatgcctg gctaattttg 69780

tgtattttta gtagagacgc ggtttcactg tgttagccag gatggtctcg atctcctgac 69840

cttgtgatcc gccccctcg gcctcccaaa gtgctgggat tacatgcgtg agccaccgtg 69900

cctgctggcc tccatatttt tttattttat tttattttat tttattttat tttatgttta 69960

tttatttatt gagacagagt ctcgctctgt ctcccaggct ggagtgcagt gacatgatgt 70020

tggcccactg caacctccgc ctcctgggtt caagcgattc tcgtgcctca gcctccaagt 70080

gagtagctgg gactacagga gcatgccacc atgcctggct aatttttgta ttttcagttg 70140

agatggggtt tcaccatgtt ggccaggcta gtctcgatct cctgacctca agtgatccac 70200

ctgccttggc ctcccaaagt gctgggatta ctggcgtgag caaacgcgcc tggctcgttt 70260

-continued

ctttttttttt ttttttttga gatagggtct tgctctgtca tccaggctgg catgtagtgg 70320

tgcagtcatg gttcactgca gcctcaacct cctgggctca agcaatcctc tcgcctcagc 70380

ctcctgagta gctgggacca caggcatgag ccaccatgcc tggctaaatt tttgattttt 70440

tgtggggatg aggtcttgct ctgttaccca gactgatctt gaaactcctg agttcaagca 70500

atcctcccgc ctaggcctcc caaagtgctg atattataag tgtaagccct gcacccagcc 70560

tgtatctcca gtttgttttt ttttttttt tttagactga gtctagctct tgtcacccag 70620

gctggagtgc agtggtgcaa tcctggctca ctacaacctc tgccttccag gttcaagtgt 70680

ttctcccgcc tcaggctcct gggtagctgg gattacaggc acccgccatc atgcctggct 70740

aatttttgta tttttgtaga gacagggttt caccatgttg gccaggctgg tcttgaactt 70800

ctcacctcag gtgatctgcc cacctcggcc tcccaaagtg ctgggattat aggcatgagc 70860

cactacgcct ggcctctcca ttttctagat gagaaaaatg aagcacacaa agggcaagta 70920

acttgcccaa agtacacaga taataaaagg cagaaccagg atttgaaact aagcaatctg 70980

gttctagagt gtgctcttaa ctacaacaca atactgcttt cattacttgt ggatctaatt 71040

agttcaactc acagaggaca aacttcattt ccaacttctt gggggaactg ccaacctatc 71100

caccccactc cccatcataa atcacctata agtaaaagat actatgcaag tattccattt 71160

tttccactct atattgaata tatatcttta gttaatgcca tctttcaaat atacaattta 71220

agaaaacaaa atgaaacaaa acaaacttac gattggaatt gaaatttttc agatttcagg 71280

gcaattcata atcagtttgg tggctagagt agctgaactc attcttcctc tcttgccctc 71340

ctctccggag gtacctaggg cttgttcctg cctggattgt agctacagag aagtgcagct 71400

caagacctgc tgctggaaag aaggactcct ggggtcagct tctctggctt ctactcacta 71460

tgggatagct cctctgaggg caaaagacaa agctttaact tatcaccctg acacttgcag 71520

gcaattttct aaaacaaccc aatgtccgga actgattccg tcttgccatt tacagtaact 71580

gcagccacgg tggctgctat aggaaggatg tggggaaggg gaaggggaag ggcagggaaa 71640

agcctgactt acatcatatc tcttggccac ctctagcaat tcattcaaca aatatccctc 71700

ctatgccaac aatatgctca gtgctgacac aatgaggggc aaaaacccca cagtccatga 71760

ccaaattgtt ctattagagg tcttaagata tcattaagga gatgatcaca tgaataaatg 71820

tgagattaaa actacttttt taaaaaaagc ctcaaaggaa aggtctatga aaacataaaa 71880

cagggaccta acctattttt cagggtgagg aattacttct atatttctct gagccatcca 71940

ggtaaagatc taaacgataa gcaggagtta agtgaacaga gtaaggggggc attctaggca 72000

gagggagcag aatgaacaaa gcccttgtca gaaggagggc ctgtccagaa gagaggtgag 72060

gctgaagtgc agggagcagg aggttgtgga gatgtggaga gaggccacat tacggagggc 72120

tctactgtag gccaaattaa ggaatttttt ttttttttt ttttttgag atggagtctt 72180

actctgtcac caggctggag tgcagtggcg tgatctgggc tcactgcaac ctctgcctcc 72240

caggttcaag tgattctcct gcctcagcct cccaagtagc tgggactaca ggcacacgcc 72300

accacgacca gctaattttt gtatttttag tagagatggg gtttcaccat gttggccagg 72360

atggtctcaa tctcttgact tcatgatctg cctgcctcgg cctcccaaag tgctgagatg 72420

ataggtgtga gccaccgcgc ctggccaagt taaggatttt ggtctttatc caagaacact 72480

gggaagttgt ggaagtatac aaggtaggag agacaacaca gagtgacaca gtcaagtttt 72540

-continued

```
tttttgtttt ttttttttt aagatcactc cttccagcca ggtgcagtgg ctcacacctg  72600
taatcccagc actttgggag gccgaggcag gtggatcact tgaggtcagg agttcgagac  72660
tagcttggcc aacatggcaa aaccccatct ctactaaaaa tataaaaatt agctgggtat  72720
gggggccggt gccttaatcc cagctacttg ggaggctgag gcaggagaat tgcttaaacc  72780
cgggaggcag aagttgcagt gagccgaaat tgtgccactg cactccagcc tggggaacag  72840
agtgagactc tgtctaaaac aaaacaaaac aaacaacaaa aatcactcct tccaactgct  72900
ctatagagaa cagattaggg ggtgcagagt ggctgtggga gatgaggcta ttgcagacat  72960
ccaggtgaga tgatggtagc ttggagcagg gtggtggtgg gagaaaagga aagaagtaaa  73020
cctaccaaaa tataatttag agataaaact gataggcttg gatatgcagg atgagagaga  73080
acaggtatca aggattccta ggtttctggc ttaggcactg ggtggatggt ggtgcctctc  73140
aatgacatac atgtttcagt ggaagaggag ccagttttag ggggaagagt gaaagtccag  73200
caaagtctag agaagagagt ctgaagtgcc tggaagacat tcaaatggac aagccaggtt  73260
gatagtacaa catatgggtc tagaactcag tggaaaggtt taggctggag atactgattt  73320
tagggtcatc agtgtgtgga ggacaatgga aggaagtcat gagtgtgtgt aagactgcct  73380
aaggagaaaa cagcacgcct gaaggaatgg taacatccca gagctggaca aagaaggata  73440
agcaaacagg tattaccaga gaggtaagag aaaaccagga gagtgtggaa tgtcatgaaa  73500
gccaagggaa gagaacattt caagactgac agatactaat aataaatgct aacatctgtt  73560
aagtcccagt catggccaga cattctgctc aggactttga atacatgatt gaatttaatg  73620
ttaacaactt tttgaggtgg gagccatttc ccttttcata gacaaggaaa ctgtggcaaa  73680
gaatggttaa gcagctgact catggtcagg tggctggtaa gtgaagctgg catttaagct  73740
caagcagcct gactccgaac tttatccact tttctctaat acaaccctta gtgcagtttg  73800
ggaggtcata taagatgaag cctgaagagt ggtcggtaga cagactgatg aagctgttat  73860
tcggaatatt actaaaggaa tgttgtttca ggggagaaac agaatgaaag ccaggctgga  73920
gcccatggtt aggaatgctc catgacattc tctatttcca gctgtcttag ggaaccagtt  73980
aagagaccag gccaggtgcc gtggctcata cctgtaatcc caacactttg tgaggcccag  74040
gtgggaggat ctcgagccca ggagtttgag accagcctgg gcaacagagc aagaccctgt  74100
ctctacaaaa aattaaaaaa aaattagctg ggcatggtag tgtgtgcctg tagtcccagc  74160
tactcaggag gattgtttaa gcctgggaga ttgaggctat agtaagctat catcatgcca  74220
ctgcactcca gcctgggcaa tagagcaaga cctatctcaa aaataaataa gctaattaat  74280
gaattaatta agaggccagg ctagtatgcc aactccgtat cccccacagt gcctaatttg  74340
atgactttta tatatagcaa atgacgcttg atatatttgg atagttaata taagatgcac  74400
aatgttgcca tgtgctgctg ctaatccttc acattaaacc attctaatgt tttcagagtc  74460
cttgatcatc acaactactc catgagatgg atttctttta tgacatttta cagataagaa  74520
actgaggctc acaggggcta aggaacttgc ccacagttac acagttagcc tagctcaact  74580
aagactcaaa tcagattttc tgtctccaag tctactgttt ttttttttt tttctcctgg  74640
gagtcctagt ggaaagcaga ctctcttggg agccagggag acctgagttc aaatatcatc  74700
tgtgtgaccc tggacaagtt actccaaacc tttctgggcc ttcctgtttt agtataaaat  74760
gggggttgca ctggccaggc acagtggctt acttccgtaa tcccaagcac tttgggaggc  74820
```

-continued

tggggcagga ggatcgctgg agttcaagac cagcctgggc aacatagtat taataagacc 74880

ttgtttctac ttaaaataaa aaaaaatggc aaggtgtggt ggcacgcctg tagtccaagc 74940

tgcttgcaag gctgaggtgg gaggattgct tgagctcagg aggttgacac tgcagtgaac 75000

tatgattgta ccagtgcact gtagcctgga caacagagca agaccctgtc tcaaaaaaag 75060

gtgggggttg cagggttatt atgaagattg aatcagataa tctaagtgag gttcttaata 75120

caacgtttaa ttaatatcag ttatctatct ctggcctcat ctctctccat ttttctacct 75180

ctgcctccta tagaagctgt gaaacatctg tgggtcctta gttctctctt ctcaaagtaa 75240

agagcaggat taaggcacat tcaagctcat atgctgctgc tacatcctag gagtaccact 75300

ctgtggtcag gctggagtct cttgtagggg cagaggaaga accagatgtc tgcccttcct 75360

tcctctgtca ggcaggcccc agccccctgt ccaatagcca gccttaatga tctacgggct 75420

ggctgtttcc taaaagataa actggacaaa tgagtaaaca ccttctggag aaggccacag 75480

ccagaatcaa aggtttaggt tagtaagtgt ggcagcaaca tcaacagcat aacagcagat 75540

ggcagggttg ctccactgac tctacatcag gagcctgaga cccagaagaa agttaacagc 75600

acttttgttg acattataga gtccatgaag cctttaaact aacacccttc aatgttatct 75660

gacactacca gtgtgtggca ggcagggtga gggctcttta ccccatatta tataagagga 75720

aactgagcct cagtgaggtt aagaaatatg cttgcattca agaatgccaa gagactcttg 75780

caggcaagga gtaagaatgc ctgagttcaa atctcagcta tgctacctat tagctctgtg 75840

actttaaaca agtaacctcc atttccctat tctcgaagta ggtaagacta gtactacatt 75900

ttattaagtc attgtaagga ttaaatgtgc tatgcatgtc gtatgctcag gagggtgtct 75960

ggcacagagt actcaataag tggtacctat tactgtaatt atgattctac caggcttgct 76020

tagatgggga tgaggggaga gaaaatgaat gtggctttca gcatcacaga aagtaccctg 76080

ctgtcagcca acatattaca acacagtctc tggttgctac gtgtgttatt tcttcactgg 76140

ccttcctgga gtcactaaat ttagcggaaa tggtctgttg cctgatgcag tgaccacagt 76200

cttctatctg cattagccct ttcccagagt ataggtatta gtaaccgctt ttcaaaagtg 76260

aaagctcaaa tagtcttatg tttgcaatcc tgagaaagga agtcctgaaa gttcccagga 76320

gccttaagag cggtggattt tccttctgat aatgctgcta ggaaccaggt caagcaatag 76380

agacccacgc tggggccatc aagactaata ggggtagctt ttcttccctc ctcctaagcc 76440

aaacaccttt gatcagggtg aacggacaat catcaaaaaa gttgtcaagg tcctctggct 76500

tgctaacatg ccaagaggcc aggtcaccat tgatagctcc acactttaac ctcagctaca 76560

aatgacatgc tctctggcta ctctcctggc ctctcctaag gagagagagc caggtccaga 76620

gaacatcaga aaggattgcc ccagatacag aattaatggc actgctcaag gaaaggaaag 76680

agaggtaact cagagtttta aaaaaggcct actgggctgg tccaagtaca atagtgttta 76740

caagtaattg atcactacca gttaccgatt tctttgttcc ttttccactc ccacaatttc 76800

acttcactag cctaaaaaaa aaaaaaaaaa gaaagaaaga aagaaagaaa aaaaagaaaa 76860

gaaaaaaagg aaaaaaaaga ggcctgttgg gaaggcttgg gaggactagt cagccatatg 76920

tccaggagat ccaaaagatt ctggccagct gcctaatcgc accagcacaa ttactgacag 76980

agctactttc agttttgtca ctggcctcca ttgcacagtc tgaggttcca gctttaagta 77040

gcttcctgtt gtcctgggca cctgagggtc tgtcggaagg ctaacagagg caggattatg 77100

-continued

gatgtggctg cagacagaag aatggggtaa tgtaagggag agatgagggg aggagtttag 77160

gatggacccc agcaatttct ccattcaata tccagcccca ccctgtgact aggaaaacaa 77220

aggtactggg ggaaggggag aagaaaagtc atttccttta ttgtaactgt agcccagctg 77280

ggtaagagca aggctggaag ggggccttga agaatgaatt tctctgagta gcacagaggg 77340

agagtcctcc tcccagcagg cctgcactac tcaggtactc taaattcatc agagaaacac 77400

tggcaagctg gcgtgggaga gatcttacaa gaccattcag gttgatgccc catgttcctc 77460

aaaacccagg ttatctctgg ggtcagacaa atacaatcct tcctagccct gaagattttc 77520

ctccaactgg ggcaaaagtc aaattagatc caaaagacag tcactctccc cagggcaaat 77580

taccaggctc tatctgaggg cgccgtgaca tccagtttga atttaccagg ggaaagctca 77640

gaagtcagga gatgtgctcc ttcagcacag ctggattaga gtctctgtcc tcaaggtaaa 77700

gctgcaacta ggagacagaa cctagatact ctacttcctc cacgggaatg atatttggct 77760

tgattgtgtc agggtagttc aaagttgccc ttaatccctg gggctccaca tcagatcacc 77820

aatgaatgtg taataagaga aatgtataca tataggtgta tatagacaca cactcacaac 77880

ataatataaa gtgacagtta ctgattatta agggtttttt ttttcatttg ttttgttttg 77940

agacaggggt ttcactctgt cccccagctg gatggagtac agtggcatga tcacagctca 78000

ctacagcctt gatatccctg ggctcaggta atcatcgctc ctcagcttcc caagtagctg 78060

ggactacagg tgtgtgccat cacgcctggc taatattttt gtatttttg tagaggtgga 78120

gttttgccat gttggccagg ctggtctcaa actcctgggc tcatgggctc aagcgatcct 78180

tctgcctcgg cctcccagag tgctaggatt acaggcataa gccaccacgt ccggccctat 78240

taagagtttt ctatatgtca gatgctatgc tagccacttt atataggtta tcacatttaa 78300

tcctccctaa caaccatgtg gcatgagtaa tattcttatc ctcattttac agattggtgc 78360

ctgaagctca gggaagtcac ttaaggctag ttaataaatt gcaaagccag gatttaaaac 78420

catagtctca ttccatagcc tgtactcttg gtggtggttg ttttttgcg atggagtttc 78480

cctcttgttg cccaggctag actgtaatgg cacaatctcg gctcactgca acttctgctt 78540

cccggattca agcgattctc ctgcttcagc ttcccaagta gctgggatta caggcatgtg 78600

ccaccatgcc aagctaattt ttttctgtat ttagtagaga tggggtttca ccatgttggt 78660

caggctggtc tcgaattcct gacctcaggt gatccacctg ccttggcctc ccaaagtgct 78720

gggattacag gcgtgagcca ccacgcctgg cctgcctgta ctcttatcta ccatgctaac 78780

gtgtctactt cctctttatg tatttgtgta catatgttat ctttctgtct ctatctcaca 78840

cacacacaca cacacacaca cacacgcacg cacagagcta acaagctagc atctgactag 78900

acctacaaat gacccaagag ttcagtgcta gggccagctc ttgtcatcca tggcctgagc 78960

aagtcttttt ggtttcaggc ctcatctgtg atatgaatgg gttataccat atagtgggtg 79020

gtttcttcta gttctgaagt tttttagctc taagagaatc cagaaattcc aaagattaac 79080

ttgctctacc ctgtcctgct tattttggat agaaaatatt catacacata ggaaaacagc 79140

aagaatatga acatatgggt tacatgccaa ggcgagaccc aatctgtgaa ctcattgggt 79200

tttgacgtta cttattagtc ccagctataa aatgaaagga ccagccaaag gctttaacta 79260

taaggtccct tacagctctg acattctctg ctcctgtcat actaaatgga acagaaaaga 79320

cagaatcaac ttccctggga gtaaaagggg caagaaccaa ggcaataaga tccacagaaa 79380

-continued

agacttgaga gctctgagtt atagctgaaa ggaaaggata ataatgatgg ttgctttcca 79440

actcttccag aggccagacg tctccaggcc aatcccctat atggttccca gggaagaggc 79500

aagtgtttaa aggtcaactc aattacttgg tataagcaca gggatgtttt acaacaagag 79560

ctatgtgcag gagctggaat ctacagcaca aggtgcctgc ctagtctatc tcattctcag 79620

gatataggca cagccctacc acttcagacc ttgggggcca ggcatgccca tccagcctct 79680

ggagtggact gctcttcttt ccaccaacaa atttcaacta ttttacgttg tgaagaatat 79740

aaagacaagt tcccaccctt gagaattcca atccgatggg aggaaggaat aaccattcac 79800

tgggtgtcta ccatgtactg gatacactgc atatattatt tcattcaatc tcaacaagat 79860

tattaccatc cattttccag atgaagaaac tgagcctcag acagcaggaa caaatctgat 79920

ttgaagtctg tataattttt aagaaactga gctgcctaca aaaataatgt agctataggc 79980

agggcgcggt ggctcatgcc tgtaatccca acattttggg aggccgaggc aggcggatca 80040

cctgaggtca ggagtttgag accagcctgg ccaacaaggt gaaaccccgc ctctactaaa 80100

aatacaaaaa ttagctgggc gtggtggcag gtgcctggag tcccacctac taggggggct 80160

gaagcaggag aattgcttga actcgggggg cagaggttgc agtgagccaa gatcacatca 80220

ctgtacacca gcctgggcga cagagcgaga ctctgtctca ataataataa caataataat 80280

aataatgtag ctataataaa gctagtaaca aaaagtagga acctagggag gcaacctgtg 80340

gtgaggaaga gctggctctg ccaaccacac taaaaacaac tcttggctgg gcacagtggt 80400

tcacgcctgt gattccagca ctttgggagg tcgaggtggg tgcatcattt gaggtcagga 80460

gttcaagacc agcctgacca acatgttgaa acccccgtct ccactaaaat acaaaaatta 80520

gctgcgcatg gtggcaggtg cctgtcatct cagctactca ggaagctaag gcagaaaaat 80580

cacttgaacc cagcaggcag agattgcagt gagccgagat tgcaccactg cactccagcc 80640

tgggcaacag caagaatccc tctcataaaa taaaataaaa taaaataaaa taaaataaaa 80700

taaaataaga atacaaacaa ctcttgaagt acctgcataa tctgtagtgc agccacagga 80760

aagtcactct tcctgggcct cagtttcctc acccataaag tgagggatat gggacaaccc 80820

tggaaagggc tcctctcatc cccccaccct gcctccagct ggcccattta tggctcttcc 80880

aatttaacct gggttcaggg gggccacact tgaggtttgg tcatgagctc ttccttggta 80940

cattcctttc cccaccctaa gattgaaatt ctcaactcag aagcttgttt ctgtgttttc 81000

catcagaaca tgaaatagtg cttggcccat agtacactgg taaagatggg aggataggcc 81060

cacctaccgg aagcaagagc cgtgccaggt ttcgttgaca gtcctgtacc atatctggct 81120

tggagcaatg tggtccccac agcctggaca cctccagaca tcttcacctg cacacaaaac 81180

caagaggttg agaggcccaa agtagcagga ggcaaaggga caggattcct accatgtaga 81240

ataagactga gttaacagca tctcatttga aagggttagg ccaatgctct gctagatgca 81300

cagtttctgc ctccacaaga tctctggtaa gcacttgtcc aacctctgct catctctcca 81360

gggacaagga actcatggcc tcaaaagtag tctcattcca ttttctaaaa atggttctga 81420

ttatcctttt ttctactgaa acaaaacttg ccttttcta gaactttcag cctttgctta 81480

tagctctgag aagaagaaac gaacacaatc tccttctttt cagtgtcagc ccttcctatg 81540

actattccca gccaacagcc agaaacttct ctctgggtta aacatgttac tgccttttcc 81600

tattcctcaa aagccctgtt ttttcagatt tttccatcac ctaatccaag agctcattta 81660

-continued

ctcaaccaac attgaataaa ttggtcctct agatcagcag tcctcaacct ttttggcacc 81720

aggggccaat ttttccacgg accctaggct gggggttggg gggcgggggg atggttttgg 81780

tatcaaactg ttccaactca gatcatcagc cattagtgag attctcataa ggagcatgca 81840

acccagatcc cttgcatgtg cagttcacaa tagggttcac actcctatga gaatctaatg 81900

ccaccactga tctgacagga ggcagagctc aggcagcaat gctcactggc ccgccgctca 81960

cctcctgctg tgcagcctgg ttcctaacag gctacagacc cgtaccagtt ggcggtccag 82020

gggctgggga cccctgttct agatacaaag ctctttgtat gttggaggct gttagattac 82080

ataatctctg ctccttaaga agcaaatcca ctggaaaaga actgctcttt catgttgttg 82140

atttttcact tggatgttat ttcagagata gaaccaagta ttattttagt tcagaaaaag 82200

aaaagtaatg ggtataaaag gggaagaaca aatatttaat gaatgcctcc tatatgccaa 82260

gaactctgct aggtgctttc atagacctaa tctcatttaa ttctcaaaat aatcctgtga 82320

cacaggtatt attatattca cttaatggat aggaaaatta aaacttgagg caggggcag 82380

ggttgactag ctaagtggtg gatctagcat tcatactcag gtctggctca ttctaaagct 82440

tctgctcttt ccattctacc atgtcaactc ttgttcactc caatgagggg aaccaggaaa 82500

gtctttgtag aggaagctga cacttgagct gagccttgca agatagggaa gacttaccta 82560

tcactggctg agaaaggagg gtgaagggtt gtgtgagctc agagaccagc atagacaaaa 82620

gcacagagtc ggtccatgtt cctccaaaaa gagaggcctc aaacaaaata caatactcct 82680

ggtatggctc agatgtcaac ttcttgaaat aaagaggaca gctgagatca agctgtggta 82740

cctgactcag ccctagaaag caggacaagg ttgttacaag actcccttgt gcgctctgtg 82800

tatgcagctt tgttctctct gggtaactac agactgccac aggcacagtg acagtcaccg 82860

aaccctgttt cacaggcaca ggggccatga atagacaaag cacagctcct ggcacttcaa 82920

gattgctcaa catcattctg gagcagccaa tgtacacagt tgtccccctc gttccccaga 82980

gaagagttaa acaagctggc aagaggcctt aaaacaccac ctcaacccac ccattccact 83040

cttctcatgc ctcagtaata ggatatatat atgcccatca ctcattcttc ctgggtctaa 83100

aatcctcaga ctccttgtca agaggggcct ttttggtttt agggtgacag tggtcacaga 83160

gtatgccatc tctctcacaa actccaaggc agaaagtctg agagactcac tcaaaccaat 83220

cccaggttgg gttacggaca gtggacaaga caatgaatta gtgagagatt ttaatgaaca 83280

caaacatcca aatgaatctt ccaagttttg gaaatatttt tggaactctt acatgtaaga 83340

gctttcaggg tcagtttatg agctgccatg ccccaccccc agcacatctc attattttag 83400

agccacacct tgtttcttgt ttttactttt taacccaaaa ttaagtctgt aacctgacca 83460

cccacttact caacaattca atctcttggc tctgaatgat ttctatttcc aaaattcaaa 83520

caggatcagg actttccacc atgaggatat tcaagataac ttgttattgc cataccctcc 83580

agagcttgcc ctgctttctc cagtcttgag aaacccaaag acattcaatc aacacatact 83640

ttagtacctg ctttgggtta agcatgaaag acccgaaatc aagaccctct actctctcct 83700

ctcatatcca cctggtttca ctttccaaat gtctctctag cccttccctg tccttcaccc 83760

cagccctact tcagattttt ggcttctttt ccctatctgc tgtactcagc tggctcatcc 83820

acactgcaag ctatctcagc tctgggttag actccccagt gtccacggaa caaaggctgg 83880

gctactcggt atgacacccc tgttctggat cccattccat cttacttctt agaagaagtg 83940

-continued

cacttttagt tatctttaca gcacccctct ccctcactgg ctcctcctta tcttttaaat 84000

aagttcaagt cactgccatc aaaaaaacct ctggtactct cgcttcagca gcacatacat 84060

taaaattgga ctgagacaca gaaggttagc atggcccctg cgcaaggatg acatgcaaac 84120

tcgttaagtt ttaaaatcta atttagtttt aaaatctaat tttaaaaaaa accaaaccac 84180

tggatccttt tctcctctag ctaccatcct tgtttctccc cttcacagcc acatgtatga 84240

aaaagttgtc ttgggaggcc aaggtgggtg gatcacttga ggtcaggagt ttgagaccag 84300

cctggccaac atggtaaaac cccacctcta ctaaaactac aaaaattggc tgggcatggt 84360

ggtgcacgcc tgtagtccca gctactcagg aagctgaggc aggagaatcg cttgaaccca 84420

ggaggtggag gttgcagtaa tctcgccact gcactcacgc ctgggcaaca gagcgagact 84480

ctgtctcaaa aaaaaaaaaa aaaaaaaaaa aaaagaggct gggcacggtg gctcatgcct 84540

gtaatcccag tactttggga ggctgaggcg ggtggatcac ctgaggtcgg gaatttgaga 84600

ccagccatga ccaacatgga gaaaccacat ctctaccaaa aatacaaaat tagccaggca 84660

tggtggtgca tgcctgtaat cccagctact caggaggctg aggcaggaga atggcttgaa 84720

cccgggaggc agaggttgca gtgagctgag atctcgccat tgcactccag cctgggcaag 84780

agtgaaactc caactcaaaa aaaaagaaaa aagaaaaaaa gttgtctaca catcgctgcc 84840

ctcatttcct cacttcccac tttctgcata atgcactgca atctggcttc cacccccacc 84900

atttctctga aactgtctca ccaaggtcaa cagtaacagt aactgcttat tgctaaattc 84960

aatggacccg cttcagtcat tctttcagca aatatttatt gcatgccaac tatgcatctg 85020

acactgtgtt ggggaaagag gagcaaacaa gagagagatt aaacacaaag caaatattaa 85080

acaataatat aattatttat tgagggccgg gcgcagtggc tcacgcctgt aatcccagca 85140

ctttgggagg ctgaggcagg tggatcacaa ggtcaggagt tcgagaccag cctggccaat 85200

atggtgaaac cttgtctcta ctacaaatac aaaaattagc caggcatggt ggtgggcgcc 85260

tgtagtccca gctactcagg aggctgaggc aagagaattg cttgaacctg gtaggcggag 85320

gttgcagtga gccaagatca ttccactgca ttccagcttg ggcgacagag cgagactctg 85380

tctcaaaaaa taataattat tattattatt tatttataat tggggcaagt ggtacgaaga 85440

tatcaagttc ataactagga aaagggcggg cggccaaagt attccttgag gaagtaattt 85500

ccaagctgag ctctaaagga tgaatacaaa ttaacggggc aaaaaagggg ggaaggaagg 85560

ctaacactgg agaggggga catgtgcaaa ggccctgggg caggaacaaa aacttctcac 85620

acatcattaa agttggcttt tgctacctcc aaagggctag caaaccatgt ctggcccctg 85680

gtctgctaca atccttcaat aaacagttac tgaaagagca aatggtcctt acaatgtggt 85740

cccaaaggac ccttctctct caccattcct tacaaaagaa tgctctagtt acattttgac 85800

ctattttcta ttcctgccac ttttattcat gtacttgcat cctgacattc acacctccag 85860

ttattctcta tgtataagct tctctatctt ttcccttaat caaaaaaaaa attttttttt 85920

ttgaaacaga gttttgcttt tgttgcccag gctggagtgc aatggcacaa tctcggctca 85980

ccacaacctc cgcctcccag gttcaagtga ttctcctgcc tcagcctccc gagtatagct 86040

gggattgtgg gcatgcgcca cgacgcccgg ctaattttgt atttttagta gagacggagt 86100

ttctccatgt tggtcaggct ggtctcaaac tcccaacctc aggtggtctg cccacctcgg 86160

cctcccaaag tgctgggatt acaggcgtga gccactgctc ccggcccctt aatcaaattt 86220

-continued

ctatcctttc tttaaagttc cctaaaatca ggcagtgttt gtttcatttt gtctatacct 86280

attttacata atttgttatt taattacagt tcactgcata tgtttctctc tccttcagtc 86340

cctctctctc cactctcctc ccccacaccc cattcagaga actaaggtgt aagaactgta 86400

tctttattat gttcttatct ccagggccta accttgggcc tggtgcaaaa cagacattcc 86460

tttgtgtggt gaacccaact gaaaatttgg atccaactgc cagctataaa agcaatttcc 86520

aaagataagt tccaaaggtc agcacagatg attccactgg aataactgtt ctggctccca 86580

gattacctat gagtagtctc atgtggataa acctggcctc agtttacctg ttatgcattc 86640

acactgcttt tctgttagga agtacctgag caaagatatt cacccaaagc acacagcctc 86700

ctacgcccag accctcttca agagaacact cttgacttct ttagctctct tacagtctta 86760

aaaattcatt gatctattct cttaccacct ttattctagg aaatgtaaaa gaaatgtaat 86820

actgtaatgc aaagtcttag ggatttctgt gaagcctagt ggaggatcaa tccaaagtcc 86880

agtcagcctg ggaagttctt tcttaagaaa atatctcatt gtaccaggct gttcattctc 86940

tccctctata gtgtttacct ctcatatttg cagacttatc aacagtggct actgattgtt 87000

aagagtccag aaaagagccg ggcgtggtgg ctcacgcctg taatcccagc actttgggag 87060

gccgaggcgg gtggatcacg aggtcaagag atcaaggcca tcctggctaa cacggtgaaa 87120

ccctatctct actaaaaaat acaaaaaatt agccgggcat ggtgatgggc gcctgtagtc 87180

ccagctactc gggaggctga agcaggagaa tggcatgaac ccgggaggtg gagcttgcag 87240

tgagccgaga tcgtgccact gcgctccagc ttgggcgaca gagccagact ctgtctcaaa 87300

aaaataaaat aaaataaaaa aaagaagagt ccagaaaagg agcttgattt aagtaaatgt 87360

tcacagtacc ttcaaacaat aaaggaccct tattaggtgt ttaataaaca gaaaactgtg 87420

ttttcagtta ttattggaat ttgttgcttt taataagtgc aactgtgcaa cgtaaggtga 87480

gcactgcaac aaatgtctca ttataggggt tctcaccttt aagacctttt caatcttggt 87540

ataaaataca ttcactttac attcacttta tgaatgcttc tctctcccaa cttaattagc 87600

aatgtggccc tcaaacaaat gtgaatccac tggggcaggc gccaaaaaaa aaagaaggaa 87660

aatagacttg cctaaccacc acaaagctac tgaaagatgg tccaggaggg attcaaaccc 87720

atgctcagag actaccatgc cagtcatatt tttaataccc tatgcagcct cagacaagtt 87780

gagacttgct caagaacaca cagctattca gtggaagggt tggcctggta tcccagctca 87840

gagccaatgc agtccccaca atgttcacct taaatttggt ctaagagctt ggtttctgca 87900

aatgagtcct ttttaggctt caaaatggtg agaaatcctc aagcactgag agctcatgag 87960

actctaatgc ctagcttaaa tcaacatcct atcctttcat ttaaaaaaat ctttgaagct 88020

tctcacgttc caccagacaa cgagaaccag gaacaaaact cagccaattc agtgcagcat 88080

aaaacaagta tggtaacccc tgggcccctt gggaactggc gggtcaggtg tataaatccc 88140

acttgttctg ggacaaagca gaaagcagaa gggcagggca gcagctgctg gggaaaagcc 88200

aagagaaaaa gatcacattt taatgaagtg actatcaaaa agactaacag taaggattaa 88260

catgcctcta gatactctcc ctgtaaccca aagggttttc ttcctcacta gcttaccaaa 88320

ttttaccaaa aggagtgtct caggaaaaaa aaagattaga tgatttttaa aagagctttc 88380

gagcttctct aaaccttgga ttgttctttt gtaaaaggag ggatcaagtg atttcttcca 88440

gctctgagag actgacacca aaccaagggc ctctcctttt tcaggttaag gcacagacac 88500

-continued

aggtatctca gaaagtttcc ttcttgctgt tattctcttc ttcccttctt tggtttattt 88560

ggaaaaattt gtgaggggta aatgtatgtg tgtttgaaac ctgaatctgt cactgtttgc 88620

ctctgtgtcc ttgggcaaat tatttgacct ctctgagctt gtctcctaat ctttaaaaag 88680

aggataatac cacctgcttt acaggattac caagtggatt aatagaaata atgtaataga 88740

tgtgaaaata tctcccatgg tgccaggcac ttagttggac cgatttgtga ctaacttctc 88800

ttccttcctg cgtgtgtatc taggaagttg gctggggaaa aggttaagct gaatacagaa 88860

acttctaaaa tctgaaggga aggttttcaa aacaaactca gcaaactttt ttctatgtgc 88920

ccatgtacaa ggcacagagt aggccaggaa ttactggggc ccttggtgtg tactcatggg 88980

gtgtgaagga catggatata cacagaaata acatgtcaga gagggacaaa ccaggtgcta 89040

caggcaaggc aaactcagta cattaacaaa atacactgca gtgacaaatg agacagatgg 89100

ttcttatgtc tttcaggcca cccagactgc aacttccaac cagaagttca ctcaactgtg 89160

aatgagtagg cattttctta tctatggagg gcctataaaa atatctttat cagagaactg 89220

gttaaccaat tgatatattc tcatacagca gttctctgcc agttgtcaat gatagaggaa 89280

aaaaactact ttatgccaga caatgtccta agcactacgc aagtatgagc taactgaata 89340

caaagatccc atgagatcag gactagtctt aaagaactat tattgttctc attttacgga 89400

agaggaaact gaaatactca gctaataaat gtttgagctg ggattagacc acaggtctag 89460

tctccaaagc ctgtcttttt tttttttct gagacggagt ctcgctctgt cacccaggct 89520

ggagtgcggt ggtgcaacct cggctcactg caacctctgc ctcccaggtt caagcaattc 89580

tctgcctcag cctcccgagt agctgggatt acaggcacct gcccccatgc ccggctaatt 89640

tttgtatttt agtagagatg gggtttcacc atcttggcca gactggtttg aactcctgac 89700

cttgtgatcc acccacctca gcctcccaaa gtgctgggat tacaggcgtg agccactgca 89760

cctggctcaa agcctgtctt taactacatt ttagactttc tttcctggac aaaacaaatt 89820

acaggcagaa agacatgata aaaaagaatc ccagaaagaa tgttcaagga acactttaat 89880

gaaaccaggc attggtttca tgttctagtt actattcaat gtaaattaat aggtaaaata 89940

gtctaattct ttctactata taaaaataga taaaggatga gaaagaaggt agacttgtaa 90000

gagactgcta cagtttcagg tcctctcatt gaaatgagat gtgtgtgatg gtgggcccta 90060

tcagacccgc tcaaggttac actaacaaga aggcagccct gttctgggaa tcagctcagc 90120

ctccagaaga tgcagaaaat ccaccatgaa tgtggaaagc ctaatacagc cccttgatca 90180

tgacggattt ttatttagtt ttgaaatctt gtggtcattt tggaccctac cctgctccat 90240

cacatatcca ggaccatcca agtcaccctg gcctgcctat gtcaggaggc atcttttatt 90300

ggttggtgcc ttcttcattg ctaccactaa gacctcatct cagggcccac tatacccacc 90360

tccagattgg tggccctgtg cccacatcat gagtcttcct ttctctaaac ctacggtgcc 90420

tactgtagac tagaaacttc caaaatcctc agcctggcat taaaccctct agtccggttg 90480

cacccacatt ctcagctgta ccttccactt cctatgaatc ttccacctcc caatccttgc 90540

ctctaccaca cgtatttgct acccttttatg aagacctgaa acttgcctac ttttgtctcc 90600

tactcccaaa gtttcttccc tactaagata tcatcttccc caccttcttt aagcctattt 90660

aagatttctc aagtcttacc cacttcaaga agtattccct aaccaaaact gatggccacc 90720

ctggataccc tgctattctt taaacatgcc aaacacattc cctgacaact cacttcttct 90780

-continued

atcccttac cctacttcac ttctcttcaa tggatttatc cctgaatatg atgttcatac 90840

atttacttgc tgtctatctc tccccaccag aacatacact acatgaggga agggatttgg 90900

tctatcttgt tccttgctgt atccctaggg gccaaaagag tttagcacat agcaggaact 90960

caatatatat ttgtggaata aatgtttgaa gagaattctc cctcatcttt gaaattctat 91020

aatgttatca tacgcaagcc ttataacagc ttgtcttatc aatttttttc ttttttttaa 91080

gagatgggga tcttgctatg tagcccaggc tggagtgcgg tggcttttca caggcgcaat 91140

cccactactg atcagcatgg cagttttgtt ttttgcctta attaccttaa ctagacagta 91200

agtttctcag aggcacaaac tatatataac tcatttccac aaatacttat taagggctta 91260

ccatgtgcca gaccctgtac gtcttaagca gcctggaata ctggaaagaa cactgcagaa 91320

gtcttacctt ggttctggct ctaccactag ttgagtgacc ctggacaact ccgtttttt 91380

tccctgagcc tcagtttctt cataggtaaa acaaggagac tatacattat ctcaaaactc 91440

tctcctggct ctaacacctg ttaaaattta taagaggtaa aacttggtcc tgcccagtgc 91500

cacagacatc acaggcactc agtaaatatt tgatttgtta tcgattctcc ggttacacga 91560

ggaatctgag gtgtctcaga acatggtttg aatggaataa caagtgaaag tttaagtccc 91620

acttccttct cagactccat agtcaacttc attctgcttt ttgtgctaca gagaggagga 91680

agcactaaac ttgaagttgg aaaatatgaa ctctgaccac cactgctcat ttagctatgg 91740

aaccttgtca tttctgggcc tgtctcctca tctataaaaa agccatactt ggctgggcat 91800

ggtggctcac gcctataatc ctagcatttt gggagtccaa ggcgggtgga ctgcctgagc 91860

tcaggagttt gagaccagcc tgggcaacat ggtaagaccc cgtctctaat aaaaatacaa 91920

aaaattagct gggcgtggta gcacttgcct gtagtcccag ctactcagga agctgaggca 91980

gaagattcac ttgaacctgg gaggtggagg gtacagtgag tctggatcgc gccactgtac 92040

tccagcatgg gcaacagagc aagactcttc ctccccaccc cctacaaaaa aatcaaaaaa 92100

ccatattgac ctcccagctc caaggctcaa aggcactaca tgccttccta ccctagggcc 92160

tttttatgct gtttcctctg cctggaatgc tctgtgagct ttttttgcca ggctaacttt 92220

tacttatctc caggggtcag cataaacttg catttttaag aaaagtattc ccttcaaaaa 92280

aaaaaaaagt agaggatggg catggtggtt catgcctgta atcccagcac tttgggaggt 92340

tgagacagga ggatggcttg aaaccaccgt gggcaacata gtgagaccct gtctctacca 92400

aaacagaaaa aaaagtagag aaaaccctca cccctggtgc actcttacag cacttatctc 92460

aattgtttac cttaactatt atagcattta ttgcaattgt tacaattatt tattgtccat 92520

ctctaggatt attaagcacc acaaagccaa gaactatgtc accactgtac tcagcatttg 92580

gctaagtata tgacacagaa gaatccaaca aatattcttt gaatgaaaga aagaatgcat 92640

aaactaatgc atctagaggc tttatgcctc tagactattg tactgtgaaa cctttcacaa 92700

ttttctaaga aacagggttc aagacaggca acctatatta cacagctgtt agaaaaagtt 92760

actggaaatc ggccgggtgc ggtggctcac gcctgtaatc ccagcacttt gggaggccga 92820

agcgggcgga tcacaaggtc aggagttcaa gaccagcctg gccaatatgg tgaaacccca 92880

tctctactaa atataccaaa aaaattagcc gggcgtggtg gcgggcccct gttgtcccag 92940

ctactcggga ggctgaggta ggagaatcgc ttgaacccgg gaggcagagg ttgcaatgag 93000

ccaagattgt gccactgcac cccagcctgg gcgacagagc aagactccat ctaagaaaaa 93060

-continued

aaaaaaaaag ttactggaaa tcattcgagt tgaaggactc tgtaagctta aaaaaaatct 93120

tttttaagag acaaggtttt gtttttaaga gatgaagtct cgctatattg cccaagctgg 93180

tctcaaactc ctggactcaa gcgatcctcc tgccctgccc tcccaaagtg ctgggattac 93240

aggcatgagc caccacgccc agcctaaaaa aatttgtttt aggcatttcc acccctagct 93300

ctctccagaa gtctgttatc accagtgaag agaaaatccc atgctccctt gctgaggaag 93360

tgacccactg gtagcccagt ccttgcattc aagatgaagg aagtgactgt ccacagacgg 93420

aaagccagga agcagggggc agatagccag ccccctcca cctcaatacc ctcttgccgc 93480

ctctccaaag cttctggacg ctgaccctgg tggacccaga ggtgggccac ccacacaggt 93540

ggctgggaaa cataggagac agaaggttcc tactgctggg ggaaaagatg gggaaggcag 93600

gaagctgcag ggtggtgtgt gaggtggcgg ggaggtgact gcaggtgcaa ctaacccaac 93660

tctgaagcta cccttcaccc ttagacaact cttccttcca ggaggtgttt ggccgaatcc 93720

ctctagtctt tagtctagag aggccaaggg agcttccaca ggaggactca gaaccccagg 93780

tggcctactt tcgccacctg tagcagagat ctttgggccc caaaaagaat ggtccgatct 93840

ctccttaccc ttagtcctgt cctaaacctc aaggggctcc atactccctt attttctgca 93900

aagagctccc agccccgggg gactcccaga caccaagatg cccgcgagaa gccttgggtc 93960

ttggcaggga agtcacatcc tccatttccc tcagaaaacc cagaacccgg ggtaactctc 94020

ctaattgtcc tgaaccctag aggatcctgt actcctcaaa ttccacagag cttccactat 94080

aagagaaccc catgcttttc aacttcctta aaaagggtta atttgggggt ctgcggccgc 94140

caggtccctc caaaaagctg gattcctggc agacccagaa tctcttcctc cccgggacct 94200

caactcgcgg gaagcctccc aacccgcctg cccctcggag ccctctgtct ccaggatgcc 94260

ctgcctcccc attcccggag tcccaccctc caaatgcctg gggagcgctg tgcccacccc 94320

gcctcccagc ttttctgagg agctcggacc cacgaggacc caaccagccc gaaacccccg 94380

agagacagcc ggtttccctc aggcgccatg gaaccggact tcggagataa cagcgggaca 94440

gggcggagca gcccttcctt acccgccagc gcggacatgg tcccgggagc gcggaaatgg 94500

ggaggaggag gcgggggaca cagcagctcc cctcagttcc ctacaactcc cctcagctcc 94560

tgccgccgcc gcctcaggcg cgggaagacc acagccccag gcgcgaacgg gatgcaccgc 94620

ctacggcgcc ctgggaggac caaaccacac ctggcacagg ggactgcccg ccgacgcccc 94680

gccccgccgc tctcaggcct cgctcccgag acgcggcccc gccccctggc tctggctcgg 94740

gctctgcagc ggccctggcg ccccctagtg cgagcggcgt cccctcgagg tgctcccgcc 94800

ggtaaggttg attctgcggg ggccggaagt ctcaacacgg atgacagtcc gcacccatct 94860

agtcgccctc atttacagat agggaaagtt aaggatcaga ggctaagttc gactacattg 94920

aaatggcgtc gctcattagg aggagagccg ggatgggagc ccaggcctgc taactgccag 94980

ttcagagctc tttccagtag tgtttggaat caggaatatt acaatggagg agccaggtct 95040

atcccacaaa cagctttgag gatggacttt tcttagaagc ttctagctca gagaaaccac 95100

ttctctgcgc cactatgagt aagctgacag taaataatct ctaagtcaaa gacagttgga 95160

ggcgaccaac cagcattcca ggagcagatc ctcagccagg ccaattaggt ttagcgcagg 95220

acgctggagc ttgtgcgccg cattggaaga cgagcactcc cctgcaggaa catgggccta 95280

ccattcccag gtcttccaac ctttcaaaac atgccaaaaa tccagatatt taggtgaaat 95340

-continued

```
ttgccacttt taaaaaatgt tgatagcatc aaaaagttct taaaatgcca tgtagaccaa  95400

atgaaacaca cctggccact agatttgccc tgtgggctgc cagtttacaa tcttctatga  95460

ggaggaggga ttacaggact aagagttagg acagctgact gggttatggc agctgctagc  95520

tggctgagta gctttgagaa ccacttccgc tcattaggtt tggaaggaga ttggggtcct  95580

ctgtgtgatt aagatattaa tacatcatca gccaggccag ctggtccctg atgacccacc  95640

ctcccatcag attctctcac gctgggtctt tcacaattgt catttcctta catcagattt  95700

agatcacagg aaacatgtcc aaatgaccag aaagaggatg gagcccaggg aaatggattc  95760

tgtcccctcc acccccaccc ctgaccaaag acaaagttca acccagctgc ctctggtccc  95820

agcagttggg taagcaaagg caaaggcaaa gacaaaggcg agcacaattc tcttccttcc  95880

tgctctaatt caaaccggat gccggtcaga tttttctctt ctgggtgtgg caggtggctc  95940

acgatgtcct gagtgtgcac aaagtgctgg ggaggagggc tctgtcccat ttccgttaat  96000

tcctgggaac tcttttcagt actcagtttt gtagtttctg aggaaatcat tatagtaaac  96060

acaagattgt gtattgaatt aaagtgacag gaatccagag actctcctga caagctgcct  96120

catggctagc aaacagggga cagggacttt ccccatcctt cctacagaaa gatgtaacac  96180

tgtattttcc caggcaacaa agcactggca tcctgggaac caaaacctcc tgggtgtctt  96240

ccctctccct ccgttccaca ccaaggcacc caggtctatc tcttctccct cctcattctc  96300

tctcacattc cacttcctcc tacctccagc ctcatggtct tcaccagcag ggagctgggg  96360

aatagccaag tcccccaagt tcctcccagg gaagttcagg cgactctgag aaacacaagt  96420

tgaaaatgta atggaaggga atagctttct atagaaatgc tatagccaag agatatattt  96480

ttaatatgga cccaaagaac attccagaaa tcacttgtga actcaagggg gacttctgga  96540

gggtagagaa agcctgggca gggagtagac caactgagtt tcaatagcac ccaggttacc  96600

tgtccacagt ttgtttcact ggccatagct tggtcagtcc ctagggccag atcctgccag  96660

ccgcctcaca gtcaactctc tgagcctagg atcctgaaac ctggaccttg aaaaacaact  96720

caaggttctg ccctcttgaa ggtactggtc cactggtacc caaccagact tgcccagtgt  96780

cacttccact tgcttcaaac tgagccagtc tctgccctcc ctgaatcatc ctacccaaaa  96840

ctgactgtgt taatcttcca aaagtatgtt tctggacctg tcatctctct gttcaaaacc  96900

cttcaacagc accccactgc ctacaaaata aaagccaaac cagtcctcat gagctgacta  96960

catctcaccc tgcggacttg tctccctatt cccttcactc caggactctg ctttttcctg  97020

agcacatcct gccttgcgcc acctgccatt cctacaacaa ccagctctcc aggttctctg  97080

taattttgtg ccccactctg cccagagtgc agtcttgttt ctcaactcta aatgatatag  97140

tggcatctgt aacatcacaa agtaggacta ccagacatta agtgcctcca ggtaagacgc  97200

accaagtata tactatcacc tatgacgtag tcttacaaaa caaaaataat aagcttgaat  97260

ctggtcaagc ttctagatct aacaaccaac ttacaggaaa tacatggaga agaggagctt  97320

aataacattt ccatgatgca aatagcaaaa atgatgctgt gggaaatgca acaggacaaa  97380

taacctagtt tctttaacag ataaactgca gagtagaaaa aatttcatag acaggccagg  97440

caaggtggct caccctgta atcctagcac tttgggaggc tgaggcgggt ggatcatgag  97500

gtcaggagct tgagaccatc ctggctgaca tggtgaaacc ccgtctgtac taaaaataca  97560

aaaaaaatta gccgggcgtg gtggcgggca cctgtagtct cagctacttg ggaggctgag  97620
```

-continued

gcaggagaat tgtgtgaacc caggaggtgg agcttgcagt tagccgagat cccgccactg 97680

cacttcagcc tgggcaacag agcgagactc tgtcacacaa acacacacac aaaaaaacaa 97740

aaaattagcc aggcatgatg gcgggcgcct gtagtcccag ctaccaagga ggctgaggca 97800

tgagaatcac ttgaacccgg gagacggggg ctgtagtgag ctgggatcgc atcactgcac 97860

tccagtctgg gcggcaagtg aaactctgtc tcaaaaaaaa aaaaaaaaaa aattcagaga 97920

caaataatag caatgtatga aactgatttg gatctcaatt gaaactacac aagttatgag 97980

acaactaggg aaatctgaac aatgtgattg gatatgatga ttttaagaaa ttttttaaag 98040

gtattttgct ttttttttt ttttgagatg gagtcttgct ctgtcgcccc gactggagtg 98100

cactggtgcg atctcggctc actgcaacct ctgcctcccg ggttcaagcg attctcctgc 98160

ctcagtctcc cgagttgctg agactacagg cgcctaccgc tacgcctggc taatttttg 98220

tattttagta gagacggggt ttcactatgt tgcccaggtt ggtctcaaac tcctgagctc 98280

tggcaatccg cccgcctcac cctcccaaag tgctaggatc acaggtgtgt gccatcgtgc 98340

ccggcatatt ttgcttgttt ttaaaatact tgagatgcat atcgaaatac ttaataaatg 98400

acatgactgg ggtttgtttc cgaatagttc agacagtggg agaggaacag cgtgcgggag 98460

gtatagatgc aacaagactg atctgagttg atagttactg gagttgggtt atggctacat 98520

ggaggcttta ttgtcctgtt ctactactgt ttttttttt ttttccccat cgaaaacaaa 98580

aacctaaggc cgggcatggt ggcttaaacc tgtaattcca acactttggg aggctgaggt 98640

gggagaacct actgagaccg ggagttccag gctgcagtaa gctgtgatca caccactgca 98700

ctctagccag ggcacccgag taagaccctg tttctaaaag ggtctgttgc cctggctgga 98760

atgcactagc atgatcacaa ctcactgcag ccttgacctc ccaggctcag gtgatcttcc 98820

cacctcagcc tccctcccgg gtagctggga ccacaagcat gcaccactgt gcctggctaa 98880

tatttttat ttttgtagag atggagtctc cctgtgttgc ccaggctggt ctcaaactct 98940

tggcctcaag cgatcctccc acctcagcct cacaaagtgc tggggtttgt aggtgtaagc 99000

caccatgcct ggcctaaaat taaaatcttt caaaaaaaaa aaaaaaacct ggtcttttct 99060

cattgggctt atgagctcct gaagttccct cacctttgca ttccccactg tatcttggta 99120

ctattcaatt tgataactgt tagatcaaaa tctaagagga agaaaaaaag ctcttttcag 99180

aggaattact tgtagctctg gcatgttggg aaggctgagg ctgggccaga tttgtaacct 99240

cagcaacatc ttccatctca cactacaccc tgctactaag gaaccttagg gagggacaca 99300

ccttctgctc gagagagagg acttggctga cagcccacct gccctgttac taaacagatg 99360

gagagacgct gacagtgagt aacctgattg ttctaaaggc tctgataagt gacaagctaa 99420

gctctattgt ggcagcttgt cacttacgaa gctcaggaat gccatgtact tgtcaggact 99480

ggggctgcct gtcagacagg ttttatattg ggcaaggctg tgtatgaagc caggcctcca 99540

gcctacattt atcatctcta attatcaata gtcgtttaaa gactttcttt ccttttttt 99600

aagagacagt ctcgctttgt cacccaggct ggagttcagt ggcatgatca tagctcaatg 99660

cagccttgaa actcagggct caagcaatcc tcttgcctca gcctccagag taactaggac 99720

tacaggcata agtcatgccc agcttctttg aagactttcg ttctcagtgt gacattacgg 99780

aggatgcaca ggagcctcca ggagggcata ccagtgtgac tactccccac ccactcgccc 99840

tccacaacac agacaagaag tccaagctca tagtaatgta aaaccatttg tttaattcta 99900

-continued

aatcaaatca ctttcacaac agtgaaaatt agtgactggt taaggtgtgc cactgtacat 99960

atcatcattt tctgactggg gtcaggacct ggtcctagtc cacaagggtg gcaggaggag 100020

ggtggaggct aagaacacag aaaacacaca aaagaaagga aagctgcctt ggcagaagga 100080

tgaggtggtg agcttgccga gggatggtgg gaaggggct ccctgttggg gccgagccag 100140

gagtcccaag tcagctctcc tgccttactt agctcctggc agagggtgag tggggaccta 100200

cgaggttcaa aatcaaatgg catttggcca gcctggcttt actaacaggt tcccagagtg 100260

cctctgttgg ctgagctctc ctgggctcac tccatttcat tgaagagtcc aaatgattca 100320

ttttcctacc cacaactttt cattattctt ctggaaaccc atttctgttg agtccatctg 100380

acttaagtcc tctctccctc cactagttgg ggccactgca ctgaggggg tcccaccaat 100440

tctctctaga gaagagacac tccagaggcc cctgcaactt tgcggatttc cagaaggtga 100500

taaaaagagc actcttgagt gggtgcccag gaatgtttaa aatctatcag gcacactata 100560

aagctggtgg tttcttccta ccaagtggat tcggcatatg aaccacctac tcaatacttt 100620

atattttgtc tgtttaaaca ctgaactctg gtgttgacag gtacaaagga gaagagatgg 100680

ggactgtgaa gaggggaggg cttccctcat cttcctcaag atctttgttt ccataaacta 100740

tgcagtcata attgagaaaa agcaatagat ggggcttcct accatttgtt ggttattgct 100800

ggggttagcc aggagcagtg tggatggcaa agtaggagag aggcccagag gaaagcccat 100860

ctccctccag ctttggggtc tccagaaaga ggctggattt ctgggatgaa gcctagaagg 100920

cagagcaaga actgttccac caggtgaaca gtcctacctg cttggtacca tagtccctca 100980

ataagattca gaggaagaag cttatgaaac tgaaaatcaa atcaaggtat tgggaagaat 101040

aatttcccct cgattccaca ggagggaaga ccacacaata tcattgtgct ggggctcccc 101100

aaggccctgc cacctggctt tacaaatcat caggggttgc ctgcttggca gtcacatgct 101160

tccctggttt tagcacacat acaaggagtt ttcagggaac tctatcaagc cataccaaaa 101220

tcagggtcac atgtgggttt ccccttttcct tgcctcttca taaaagacaa cttggcttct 101280

gaggatggtg gtcttttgca tgcagttggg ctgacctgac aaagccccca gtttcctgtg 101340

gcaggttctg ggagaggatg cattcaagct tctgcagcct aggggacagg gctgcttgtt 101400

cagttattac tgcctcggag ctccaaatcc caccaaagtc ctgactccag gtctttccta 101460

atgcacagta gtcagtctca gcttcggcag tattctcggc tgtatgttct ctggcagaga 101520

gaggcagatg aacatagttt tagggagaaa gctgatggga aacctgtgag ttaagccaca 101580

tgtctcacca ggaataattt atgccaggaa accaggaagt cattcaagtt gttctctgag 101640

gccaaagaca ctgagcacag cccagagcca ataaaagatc tttgagtctc tggtgaattc 101700

acgaagtgac cccagcttta gctactgcaa ttatgatttt tatgggacag caatttcttg 101760

catctctaca gaggaagaag aggggagtg ggaggggaag gaaagagaac agagcggcac 101820

tgggatttga aaggggaacc tctctatctg aggagccccc actggcttca gaagcaactt 101880

accaaggggt atttaaagac atgaaaattt ccagaaatac catttggtgc atccctttgt 101940

ttctgtaata ttaaactcag gtgaaattat actctgacag tttctctctt tctgcctctt 102000

ccctctgcag agtcaggacc tgcagaactg gctgaaacaa gatttcatgg tgtcacccat 102060

gagagatgac tcaatgccaa ggcctgaagt tatagagtgt ttacagcggt ggcgatattc 102120

aggggtcatc gccaactggt ctcgagttcc aaagctctga tgaagaaaca agactccttg 102180

-continued

atgtgttact gatcccactg attccaggag tcaagattag ccaggaagcc aaacaccagg 102240

agttggggtg gcacgtcacc agtccagagc cctgccacgg atgtaggcag gagcccagca 102300

ttaggcaatc aggagccaga acatgatcac cagggccaca aataggaaga ggcgtgacag 102360

gaactgctcg tccacatact ggggtgtccc agggacagct ggagagacag aaaggagaaa 102420

acaggctatt agaacactgc agctatgcac actgcccact caggccagaa gttgacagca 102480

tcaatcttta tctctttaca acatcagaat cctacacaca aaaagatttc ttttttttt 102540

ttttgaaatg tagtctcgct ctgttgccca ggctagagtg cagtggcacg atctcagctc 102600

actgcaacct ccgcttccca ggttcaagcg attctcctgc ctcagcctcc taagtagctg 102660

ggattacagg tgtgcgccac caagcctggc taatttttgt atttttagta gagacggggt 102720

ttcaccatgt tggtcaggct ggtctcgaac tcctgacctt gtgatccgcc tgcctcagcc 102780

tccaaaagtg ctgggattac aggcatgagc caccatgcct ggccttttt ttttttttc 102840

tttttgaggt ggagttttgc tcttgttgcc caggctggag tgcaatggca cgatctcggc 102900

tcactgcaac ctccacctcc taggttcaag cgattctgct gcctcagcct cctgagtagc 102960

tgggattaca ggcacgcacc accacacccg gctaattttt tttttgtatt tttagtagga 103020

acagggtttc accatgttag ccaggctggt ctcaaactcc tgacctcaga tgatccgccc 103080

acctcagcct cccaaagcgt tggcattaca ggcatgagcc atcgcgccca gccaaaaaaa 103140

atttctaaac cccaaataat atcaggcaaa atgaaactat gtggaacaat aacaataagg 103200

acaagacacg gtccaattta tcataaaagg tttggcctgc acatgcaaat aaaccaggtg 103260

attaacagga aactcttgca ggaaataatg aaagcaagcc tactatttgt taagtatctt 103320

ctatctgcta agtgttttgt acatgttatt atttggtgga tattattatt tccattttga 103380

gtaggaagca gctgaaacat aggtttgagg tgacctgttc aagattgaaa ggctcagaaa 103440

ttagtggcag agctctgcag cctctccaca cactgcctct caagggcccc acagggattg 103500

ccttcctttt ccagaagaga actaggatcc agacaggcaa aatgacttgt ttttcaaggt 103560

agaggcagaa ctggggtgag aacctaagtc tccactcctc tctgctttag gaaaagctac 103620

catttagagc tccctgtcct gactttgggg gaaggaggat aaaagaatct gtgtccagga 103680

tgggcaaatt gtaccaagca tgacaaggga aaactgtgac tgtggcaggg ggataaagta 103740

atccagagtc tagcctggct gtggaactct gcttagctgg ggaatccaag gttagccttc 103800

gaccaggtta gaaggatgca agctatcagg caggttcttg ctggcttcac aggaatttgc 103860

cctttcttca gggcagagta caccctatct tttggggtat acttcctctg aaaaacaaat 103920

gtgaacagag gataaacctt caaggcagat aaagtaagaa agggcatccc cagggccggg 103980

cgcggtggct cacgcctgta atcccagcac attgggaggc ccaggcgggc ggatcacgag 104040

gtcaggagat cgagaccatc ctggctaaca cagtgaaacc ctgtctctac aaaaaaatac 104100

aaaaaaatta gccaggcgtg gtggcaggcg cctgtagtcc cagctactcg ggaggctgag 104160

gcaggagaat ggcgtaaacc tgggaggtgg aacttgcagt gagccgagat tgtgccactg 104220

cactccagcc tgggggacag agcgagactc catctcaaaa aaaaaaaaaa aaaaaaaaag 104280

gcattcccaa gatatcttgg tgctgagatt cctcagctag atgctctgat ggattcacga 104340

agacagacag gcaggcaggc aggcagtcaa gaaagactta tagatcacct gctgtgtacc 104400

aaggaccagg gtggggatga gaataactgg gatgacccag ttaccatctt tagggcataa 104460

-continued

gctcttggcg tgaaccacac aatggtcact tctgtaccat atgctgtgtc cagtgcacat 104520

gaagctccaa caagtgcagg cagagctcaa ccgatcatgg catctctgca ggctccacag 104580

taaggatctt ggcttgggtg tcagactgtg ataagagttc aaggcaccca cctcatagag 104640

gtggcttgca cctatagagc aagggctcta taccgaatgc ggatggcact ttcaagcagt 104700

aggagggaga actgcagcag cagttggtag acctaggcat ggttcccaca gggctgaggc 104760

tgcttctgac cagcctcatg ttgtgtcctg aaggaagcta gacagagatg ccagagaatg 104820

gagatgcctt cactccttat gggtcatctc acatttccag gcccccaaat ggtcaaagtg 104880

ggaagtaaat gtgcagaact aggcttttcc tcagctgagt atcacttaca gttttccatt 104940

ttatagacat ggggtatgtc ttggtaacat tactgagcat tttctgatct gttaaacgca 105000

gctaatcaca cacctgttcc atagggtggt tgtgatgatc ccctgagctt ggaatggaaa 105060

gtgctatata tatgcatgcc taatagttgg taagtgctca gcattgttac catgtctgta 105120

agtgatctaa aaacagctta agatattttg aaaaggctga ttatccatgt agtgacaatt 105180

atctgtaact tatatactgc tgaattccaa ggagctttca agcaatatta ctaagatatc 105240

ttcctaattt ctcaaggaaa tagggtctta cctggaggag gccgcccatc atttatatta 105300

aatgctgtgg caaatatccc aaagggaaat gccccaattc caaaagacat ctggaagcca 105360

ccatctccaa atccaaatcc ttgaaatccc tgtgtggaag aaaggcaatt aagggcagga 105420

gaatcaggca ggtttctcct gtgcctgtca aaactcaaca tacctcacgt tgtcttgcat 105480

gtggcacctg cctcttctcc ctaccttgtg ccacatgttt cgaaggtggg ggacagcttt 105540

tcctcccata tataccctga agtagaggaa ctctggcaag ccctctgccc ttttagcctt 105600

agcttccagc taagaatcta tgaatgggcc aaaaaggact ggctctgacc cagtgtgaaa 105660

cactctgctg acataagtca gcactcacag ggcttgagga tattttccat agtcccacat 105720

tagctagtaa caccttctac atgctttata ggagcaatat tgacagccac ccctgtgccc 105780

aagggaacat caaacaagct acacattatc tttaatagca aaaatgttga aactctaatt 105840

agccaacaat ggaagtcatg atcaaaatct tgatacatgc attcaaagaa ccattaccct 105900

gacttaagcg gtaacactga acaactcttg tgaccaagtt atcagacaaa attgtttgaa 105960

cactgtggtc acaactacat gacaatatgt aggtatatgg acaagttacg gattaaaatt 106020

tcttttaata ctgttacata gaatttttac attttaaaac actataatgt tttctggtta 106080

taaaagtaat atatacttgc tataagaaac tcacatatta cagaaataat tgacagagtc 106140

agagatcttc tattcacctt cccacttcgc agccactcca ctttacctct gaaataacca 106200

ctatcagcag tttggtatat caccctcagt tttctttcca tgtgcatgtt ctacactgca 106260

tcatctttca catgaaaaga ttcagcagga acaatactgt gaataaataa ctccctttt 106320

taaagcgtga gctctgtgct aggcaccatg ttaagcacta ggacatataa atcatgttta 106380

atttttacat tggcttcatc tgcatttcac agatgaccaa actgaggctt ggagaggtga 106440

aataacttga ccaaggagta cagtttagat ctaaaactag gtcaatttga tctgaaaaca 106500

catgcttttc tttttctttt ctttttttt gagacagggt ctcgctctgt caccaagcct 106560

ggagtgcaat ggtgcaatca tggctcactg tagactcaac ctcccaggct caagcaatgc 106620

tcccatctca gcctcttgag tagctgggac cacaggtgta cccctccatg ctttgctaat 106680

ttttaaattt tttgtacaga tagggtctca ctttgttgct caggctggtc tcaaactcct 106740

-continued

gggctcaagt gatcctccca cctcagcctc ccaaagtgct gggattagag gccagagcca 106800

ctgtgcccgg ccttcatgct cttaattact gtgcgcatct ttcaaatatt tttttctctt 106860

cccctgcacc cctgatgaga ggtttagaac aaatgatttt aaaaataata aaaagcagac 106920

agactattgg gtcaaatgct cttcacaaag agctcaatgg ccttcattta tgagtttttt 106980

gcaagtcaca aaaatagctg ctctccatca cacactgatc agcctggcat ctagttccaa 107040

acttcttgcc agatggcatc acctataaag cctaacagct gacttaatga gagggcaaac 107100

ctaaagggag ggaaaacaaa gagtttccat ctcctcaata caaaatgcct cctactcagt 107160

tcatggttga atctaatggc tcctgggact gaggctgcct tacttgggaa ggtgcaattg 107220

tttcatcagg aaccacagat tcattcagtg caatcctgag gactaaccgt tctgcttccc 107280

ttctagcaac atctttatga tctcaccagt ccatacattt ggtgggatta tgttcctggg 107340

ccagagtatg caaaaaaaca tcgctagtag ctgcaggtgc ttataaatgc tagaggtatg 107400

cactttggtc aacaaggcca gggcctctgg cttctggggg cctctgattc ctgcattagt 107460

tgctaatgct ctgtaggaaa ccaaaaccta ggaatgttaa agtggtgtca cctggtggtc 107520

aagagcagag acttcaatac cagccagatc tgggtttcag cactgtacgt gtggccctgg 107580

acagatcact ttacctatct gggcctcagt ttcctcctcc gttaaatgag ttgatgtgag 107640

gcttaaatga gatcatgcat ataaagctct taacatatag ttatttccat atagtaacca 107700

ttcaacaagt gttcactgct agtattatta ttaagggtaa tccggagaag agaaccttaa 107760

tttctttctt tctttttttt ttttttgag acagagtctt gctctgtcgc caaggctgga 107820

gtgcagtggt gtgatctcag ctcactgcaa cctctgcctc ccgagttcaa gtgattctcc 107880

tgtttcagcc tcccgagtag ctgggactac aggcacatgc caccacaccc ggctaatttt 107940

tatgttttta agaaagcaga gacagggttt caccatattg gtcaggctgg tctcgaattc 108000

ctgacctcag gtgatccacc cgcctcagcc tcccaaagtg ctgagattac aggtgtgagc 108060

caccacgccc agccgagaac cttaatttct ttcaagatcc ctgactcaat aatcctactc 108120

ctggaaatcc taataaaatg ccattattca aactcaaaaa aaaaaaaaaa aaaaaaaaaa 108180

aagaaagtgg cagtgcgcgg tggctcacgc ctgtaatccc agcactttgg gaggccaagg 108240

agggcggatc atctgagctt gggagttcga gaccagcctg accaacatgg agaaacctcg 108300

tctctactaa aaatacaaaa attagctggg catggtggcc catgcctata atcccagcta 108360

ctctggaggc tgaggcagaa gaattgttta aacccgggag gcagaggctg tggtgagccg 108420

agatcgcact actacactcc agttgggcaa caagagtgaa acttcatctc aaaaaaaaaa 108480

aaaaaaaaaa aagtgtgcaa agatggttat tgtagcatta tttaaacagc aaaatgggga 108540

aacaactaaa aggtccttcg atgaggaaaa cccaaattac aatgaaccaa aactcaatgg 108600

gttcccagca atagcatgga aaaataaaaa ataataaaaa actcaatagg ataaatggca 108660

gcctaaataa taaagaggac acttcagagc atcatggaaa acctgccaaa attaaatttt 108720

aaagcaaaca gaatggtgat ggctacataa aaatacatat gcacagccta cagatgaaga 108780

gaacatgagg atatgagaat atacttgttc tgggctcttg gcaatggtgg gtaaatattt 108840

tgcttttaaa atttctttaa cattttaaca gaatttttt aagagtaaca tttgatgtaa 108900

atatgcaaaa aatgtaaatg tttttgttcc tagccactag accactagag acaaatgtgc 108960

aatttttttt tagccagtaa tcttacttct aagaatttat cctaagccag gtacagtggt 109020

-continued

tcacgcctgt aatcctaaca ctttgggagg ctgaggtggg aggattgctt gagcccagga 109080

gttagagacc agcctgggca acatagtgag accttgtttc tacccaaaca aaacaaaaca 109140

aaaaaaagaa agacaaaaaa aaactcaaaa aactagctgg gcatggtgcc gtgcacctgt 109200

ggtcccagct acatgggagg ctgaggtaga ctgagcccag gaggtcgagg ctacagtgaa 109260

ctgtgatcac accactgcac tccagcctgg gctacagagt gagaccctgt cttatcctaa 109320

taaaataata atgaacatgc tcacaaattt atctgtaagt tcatcaggat ttttttaac 109380

ttggaaaaag aaagccgcta atgtctaaca aaggattgaa ttatgttaaa tccacacaac 109440

taaatgctaa acagctagtt agtaaaatat tgctggcaat gtaaaatctc aagataatac 109500

taaaagactc ccactcacac aaagaaggta ctaagtatct gtagaacaaa tgaaatttta 109560

agagagggtt acaaaacagg agccacatat tgtaaaaaac aaaaaacaaa aactgtaaaa 109620

agtatacaca catctgtaaa atacacgcac agatttgtaa aatttgtgtg tatatataat 109680

tatagaaaaa tggaaaagaa aatagaacca aaatgtttcc agtggttacc tctggattgg 109740

tgtggttata agtgattttt ttccccctct tttcctttct gtattttaca agttttcctt 109800

tacttttata atcaaggaga aaaattgtaa ggctaagaag gaaacttcta aagccatcag 109860

ctcatttgta gaagcttctc ctagaatgtt cctcacccct ctattctccg gctctggcct 109920

ctgtccttga ggacgaggag gggtcttctc tctgaaacag aaatggtgta aaataagtct 109980

attacaaatt gctttgagtt gctccctgtg aggacctgga gctctttacc aacagggtca 110040

gggccatagc ctctttcttg tccaggacat gaacttatgc cagattaatt ctccaaatac 110100

gaactcttca ttctacctgc aactgttctt cccacgtgta tagagaacca gggagaaaat 110160

caactcgtta tgagatattc accagattgc actacagcct actcctctta cagagatcct 110220

gactcagatc ggttaggaat catggccaga tgccttccct gggcctgtta cagataacag 110280

aaatcccaaa tccatactgg aacagagggt gggttgtcag cagtaagggc tgaggttcat 110340

ggtttctctg accttcccct attaacttca gctctatctc ctactattct tgggcaaata 110400

cccaatgcta ttatttttga gtctcacccc ctcaaatcac tcaacccatg aatactcaat 110460

tgagggctta gttatacaat aagcattatg caggatgctg taatactgat gacagaccca 110520

acccttgaag tgttcatggt ttcaaggggg ataaagccag tcatatgaca ctgtggtggg 110580

aactgggtga gagcaactaa tgtgaagaag acagaccact tcaaggaaga gggaacttta 110640

gggctgctca taagagatga gtagaagtta gaggagcagg tgggggtaga ggggtagctt 110700

gagcaggact ggggagggaa gaaggcagaa gagagagcaa atacacaggc tgataaatgt 110760

gaaagagcaa gtccctttaa gagaaagtct aatatgtcgg gtacacagca tatgctaggg 110820

tagctggata aaggtaagag attcttctgg agagcttggg gatgccagac catgaagggc 110880

tctcaatact tgtgtttgtt gaatgctgga cccatgcaca tctccatttc tagaaatccc 110940

aaactccaca aagtcaccca gatcctgttt cccctcccgc catgctttat acatcactga 111000

tgcagcaacc ctgctccctg gagcctctcc tgctacatat cctgcttccc aggttgcaga 111060

aatgctcttt cacagggctc tgcacacttg tgattgtgat gtctgttcac tgctgggtcc 111120

tagcaagtgg caacttggat cccctctctt ccctcccctc ctagctgact ctcatctgga 111180

tgagactgct tcgaaaccaa gagatcatct agtactaagt aaagggtggg gttaggggat 111240

gctgaatggg agtgaattca catccaaggg gagaagtggt tgggagaggc atcctgagtc 111300

-continued

ctcacctggg gtcctgttgc ccagtgctgc cccttccata gaggggggatg accttgtctc 111360

ggctgatgcc agctttgcaa acaggacaca cctgtctgtt aggtctggtc tccaaccact 111420

gcaaacagga gagaaaaatg catgtggatg gcccaagacc caccagagtg atcagccaga 111480

gccaccagaa caacacaggc cagaggcctc aaactgggaa acaccagcat gagcagggaa 111540

catctcaagg aacctgactc ccattccttc tcctaccatt cagcaaatac ctaccagtgg 111600

taaaatcaaa attgccttta aatatcatca taattggccg ggtgcggtgc tcatgcctgt 111660

aatcccagcg ctttgggagg cggaggtggg tggatcacct gaggtcagga gttcaagaga 111720

agcctggcca acaaggtgaa accccgtctc tactaaaaat acaaaaatta gccagacatg 111780

gtggcgcatg cctgtagttc cagctactcg ggaggctgag ggaggagaat cgcttgaacc 111840

cgggaggcgg aggttgcagt gagccgagat tgcaccactg cattccagcc tgggcaacag 111900

agcaagactc tgtctcaaaa aaaagaaaaa aaatcatcac aattaggctt cttcatgctt 111960

tgtttctctc tagggcctcc tagcaacaca gggggaggga aactgcaagt ctggggctgt 112020

caagtctgac atgcccagcc tcaatttagg ctgctggtgt aagtacagca aactttaaat 112080

ccccaccttc ccagtaggag caactgtatt ttcctgtcag tcatattctt ctttctgtga 112140

tttctcattt ttttttcccc agacagagtc ttgctctgtc acccaggctg gagtgtagtg 112200

gcatgatctc aggtcactgc aacccctgcc tcctgggttc aagcaattct cctgtctcag 112260

ccttccaagt agctgggatt acaggcgccc accactgcgc ccagctaatt tttgtatttt 112320

tagtagagac agggtttcac catgttggcc aggctggtct caaactcctg acctcgtgat 112380

cttcctccct tggtctgcca aagtgctggg atcacaggca tgagccaccg cgcccggcca 112440

atttctcatt tttatagcat tttcccaaca actgtgtatc ccaggcatca gaatcaggag 112500

ataagctgat cattttctgg tgacttacca catgctaaag gataggtttt gctcattatc 112560

taaatttgct taatgctttt tgtgtacttt cattgatttt ttgcattcac ctatttcttt 112620

caatataaaa aaaatcccgt tttgggtttc tcaggccctt tcactggtta cattaaaatt 112680

ttttgtttag gccaggcgcg gtggctcacg cttgtaatcc cagcactttg ggaagctgag 112740

gcgggtggat cacctgaggt cagcagttca agaccagcgt ggggaacatg gtgaaactcc 112800

gttgtctcta ctaaaaagac aaaaaaacta gccgggtgtg gtggtgtgca cctgtaatcc 112860

cagctaccca ggaggctgag gcacaagaat cgcttgaacc cgggaggcgg aggttccagt 112920

gagccgagat ctcaccattg cactctagcc tgggcgacag aatgagactg tctcaaaaaa 112980

aaaaaaaaat taattttctt aaaaaaacaa gtgataaaaa atgtttttgg tttgtttttg 113040

tttgtttttt gcttgttttt gatttttttt tcggtttttt ggttttttttt ctcgaaacag 113100

agtctcacta tgtcacccag gctggagtgc agtggtgcaa tctcagctca ctgcaacctc 113160

tgcctcccag attcaagcga ttctcgtgcc tcagcctccc tagtagctcg gactataggc 113220

gtgcaccacc acgcccagct aatttttata ttttttgtag agacggggtt ttgccatatt 113280

ggccaggctg ttctcaaact ccttgcctca agtgatccac ctgccttggc ctccctacgt 113340

gctggaatta taggtgtgag ccaccacacc cggccaaaat gtgtttttat agacattcta 113400

taaacatgaa tctaggactt cagggctgag gagatggcag cactgaatga ttttatttac 113460

tactcttact gttcatagct ccaaaactta atgtgttatg tgttcttcat aaactcatag 113520

tcaatatgta ctaccaggtc ccgctgggga cagaatgtgc agccctgttc ccaccttctt 113580

-continued

tctcccatcc agattctgaa gctttagcag tctggaactt ctgtcctccc ccacattccc 113640

aggccaccta cgagtgtggt actctactgc tgacgaccag ggagactttc agcaggctgt 113700

gattattcac agccacttgt gaatgccagg cccagctctg tctcaccaag taggtgctaa 113760

agctaggaaa atagactttt ctctcctcac aaaagctctt tgttgacacc aaaaaaattt 113820

tccccttcct aatattttca taacctttaa gggctggggg actgttgata tgataggagg 113880

cactgtttta atagaagcag aatctcaaaa ataaaaagga cctaactaac atactatcaa 113940

tccaatccaa tttataaatg agtaactgag atccaaaaag gccctctgtg gcagaattag 114000

gaccatcccg atttcatcaa gtgtctgtca gtggacaggt atgtaaagaa ggctgtgggg 114060

gtaagtaagc aaatgagccc acaagcaggg cagattgcag cttctgaaag cagtctcctg 114120

agactagggc tctcttgtgc aaccagagct gatgaaagac aaagtaaaat gaaatgcatc 114180

ttacctgatg taaacacggc caactgccag agtccccagg gcaaccagtc atcaaggaga 114240

aggagaaaag gggaaaggaa gaaaaatgtt ttcattagca gaaacacttg atgggtggag 114300

tggtagtagc agaatgagaa aacaggagta aactatggag cactgcatgc ccattgaggt 114360

tctgtaggga ttactggtat cggaaattaa aaaaaaactg gggtaaaaat cacagaaaat 114420

aaataacgca atggagataa gaccccagcc tctgcccctg agatctcagt atatgtgaaa 114480

tccccagggc tcagaccaca cagacccagc ttctctacaa atcaccctgt gagaggttac 114540

aacatctgaa ggaatccaat aaaagctaag gacaccctgt gcaaaaaatt gatgggggca 114600

ttcacataca catgtactac acactatttt acactcaatt tcagtgggtt catagtctcc 114660

tcgggctcac aagggatgtg tggactctca ggtagtccag atccagttta acaaaatctt 114720

caggctgtcc tcatccttga gaactctcct agcaggaagc agcttaaaag ttcaggtcta 114780

ctctaaaagc atctttaagt aatagaaagg atggggatgc tacttgacag aaagggttac 114840

cttcagctta gagagctgct gtccaatctc tctccatttt caaatgagga aaccaaagct 114900

cagaaacata aaatgaactg catcaggatt aaaaaaaaaa gtctcctgac ttctagttca 114960

ttgccctttc tataatctaa caccacatac aaaaaaaaaa tgatatctac agaattgaaa 115020

gaaaagtatt caactccact ttttttttt ttttttttg agacgaagtt ttgttcttgt 115080

tgcccaagct ggagtgcagc tcactcagct cactgcaacc tccacctccc aggttcaagc 115140

gattctcctg tctcagcctc ctgagtagct gggattatag gtgtgtgcca ccacgcccag 115200

ctaatttttt gtatttttag cagaaacggg gtttcaccgt gttagtcagg ctggtctcaa 115260

actcctgacc tcaggtgatc cgcccacctt ggcctcccaa agtgctggga ttaaaggtgt 115320

gagccaccgt gtctggccca cattttaat ttctttctct cgtttattta gttaacacat 115380

atctactgtg tacctactat gtaggccagg acagtggaaa atgcaggcat ggcttctgta 115440

cctcacatac caagtatctt tagctgagta cactttgaac tcaaaggata gttccttttc 115500

caattcctga accacagttt agcttcacaa accaaataca caatcttcct ggccataaca 115560

tctgactcca acattaatct tgcaatgtca ttctaaactg agtggagaat aaccccaaat 115620

catggcagat ctccactggc aagttccctt aaagactcta ggggagaagt ggaggtgaat 115680

gaatgaccct caggggaaag tcaggcttta taacaacgac cattcctccc atttattaag 115740

caacgatagt catgtatcac tgaacgacag ggataattct gagaaatgtg tcattaggca 115800

atttcatcat tgtgagcatc acagagtgta cttacacaaa cctagctgaa atagcttact 115860

-continued

acacacctag gctacatagt atagcctatt gctcctaagc tacaaacctg aatagcatgt 115920

gagtatcata ggtaactgta acagagtggt aaatatttgt gtatttaaac atacctaaca 115980

tagagccaag gtgtggtggc tcatgcctgt aatcccagca ctttgggagc tgaggtgggc 116040

agattatttg aggtcaggag ttcgagacca gcctggccaa catggtgaaa ccccatctct 116100

attaaaaata caaaaactag ccatgtgtgg tggcacatgc ctgtaatccc agctacatag 116160

gaggctgaag cacgagaatc acttgaacct gggaggtgga ggttgcagtg agctgagatc 116220

gtgctactca atctggagtg ctactgcact ccagcatggg tgacgcagtg agactccatc 116280

tcaaaaaaaa ccaaaaacca ggctgggcgt ggtggctcac gcctgtaatc ccagcacttt 116340

ggaaggctga ggcggttgga tcacctgaga tcaggagttc aagaccagcc tggccaacat 116400

ggtaaaaccc cgtctctact aaaaatacaa aaattagtca ggtgtggtgg cacgtatgtg 116460

taatcccagc ttctcgggag gctgaagcag gagaatcgtt tgaacctggg aggcagaggt 116520

tgcagtgagc caagatcacg ccactgcact ccagcctggg caacagagtg agactcctca 116580

aaaaaacaac aaacaaacaa acaaaaaatc caacatacct aacacagaaa aggtaacgca 116640

ctgctctata acattttaaa atcactagat tataggaact tttcaggtcc atcatcatct 116700

tatggaaccg ctgttggata tgtaatccat tgttgaccaa aatgtcatta tgtactgcat 116760

gactgtactg tgtgccacgc tgtggctgaa taccttttca tttcttaatg tcactggacg 116820

gtcacattaa cctatgaggc agtaccatca tctccatctt acagatgagg aaaccaaagt 116880

acagggaagc tagcagattt gctcagggat gcaccatcca gactggaagc atctgtttca 116940

ttttagagtc tgtgttcttt ccaatggcat tctgccttta aaaaaaaaat tgatagcagg 117000

catggtggct cacaccagta atcccagcac tatgggaggc cgaggcgggc agataacgag 117060

gtcagaagat cgagaccatc ctggccaaca tggtgaaaac ctctctctac taaaaataca 117120

aaaattagct gggtgcacct gtaatcccag ctacccggga ggctgaggca ggagaattgc 117180

ttgaacttgg gaggcagaga ttgcagtgag ccgagattgt gccactgcac tccagcctgg 117240

tgacaaagca agacttcgtc tctaaaaaaa aaaaaaaaaa tttgacacag ggccaggcat 117300

ggtggctcac gcctgtaatc ctagcacttt gggaggtcgg ggtgggcaga tcacttgagg 117360

tgaggagttt gaaaccagct tggccagcat ggcgaagccc catctctact aaaaatacaa 117420

aaagaaaatt agccgggcgt ggtggtgcat tcctgcagtc ccagctactt gggaggctga 117480

gacaggagaa ccacttgagc ccaggaggca gaggttgcaa gtgagccaag actgcaccac 117540

tgtactccag cctgggcaag agagcaaggc tgtctccaca aaaaaaaaaa aaaaaaaaaa 117600

aaaaaagata tacaacagct gtacatttaa atatatgtat agttaatata tacataaaat 117660

acataattta taatcttaaa atttataatt ataaaattta ataaatacat acacaaacac 117720

aaacacacac acacacacac acacacacac attttttttt tgagacagag tcttgctcca 117780

tcacccaggc tggagtgcag tggcgcgatc tcagctcgct gcaagctctg cctcccgggt 117840

tcacgccatt ctcctgcctc agcctcccca gcagctggga ctacaggtgc actccaccac 117900

acccggctaa tttttttttt ttttttttt tagtagaaac gggtttcacc gtgttagcca 117960

ggatggtctc gatctcctga ccttgtgatc cgcccacctt ggcctcccaa agtgctggga 118020

ttacaggcgt gagccaccgc gcccagccca gaatatatat atatttttaa gagacagggt 118080

ctctctctgt cgcctagact ggagtgcagt gatacaatca tagctcactg tagcctcgaa 118140

-continued

ttcctgggct caagccttcc tcccacctca gtctcccgag tagctaggac tacaggtacc 118200

gccaccatgc ctgactagtt ttaaacatag tggggcctca ctatgttgcc cagactcatc 118260

ttgcactcct gggctcaagt gatcctccca ccatggcctc tcaaagtgtt tggattacag 118320

gtgtgagcca ctgagcccag ccataatagc tgtacacatt ttttatgtta tatgtgataa 118380

tttgatacat tcacataatc aaatcagggg tatcctgtct ttgaattctt aattttgtca 118440

atcatttcaa gtcctggttt tacgataccc tgaagtatcc tgaatttcta ggagaatgtg 118500

gatcacaaca gcttttacac aagatggtaa ttagggcaag agcctcaaga acaaaaatct 118560

ctgctccatt ctatttattc accaataaat tgatgtttta cgctgttgcc aacttaaaat 118620

cccttaaata ttgtttaaac ggccttttct tcacacgatt tctctgctca aaaatattca 118680

ttacctctat gtccctttag tcgaatctgg gtttgataat ctggctcaac actactgtaa 118740

ccattatttt tcctaacact cccagaccct catccagagt gctgtgagag gtatacagta 118800

gcctcaaggg tctcagatga gtccaacagt gcatgtgcct gtccctagaa actccctgcc 118860

ttggagtgac acacacaccc ttcgccttta cctagcattt tttgatcgac ctttcaaggt 118920

tcagctcaaa atctcaagac aaaaaaaaac acacacacac acacaaaaaa ccaaaatccg 118980

accttcacct agcacctagg ttctcagagc cagaaagggc tcctgtaacc gtcaaggctg 119040

gaacagacac cgaccttgac acttcatttt gatgagtgta tctatctata tataaacata 119100

aaatatagta acaattatag tcatagctaa cacatagtat ttactgcaag ctatgtgctc 119160

ttctaagcct tttacatatt ttgagtcatt taatgcttgc aaacgaccct gtaaaatgag 119220

gcagaatgtg tgtaagtacc ttgtccaagg tcaaacagga aggtggcaga gtcaagcttc 119280

aaaaggagga agcccagttc cacagtcctt gcttataacc attatgctat actgcctcta 119340

ggctcagcac agcaagtagc aggtagtaag tgctcactaa atggtaagta ttattacttt 119400

agagatgggg agaacgagct tagagtggag acaggatttg cccaaggtca tccagctcat 119460

tggtagcaga gctgggatgg gaactcaggt ctcttgtccc ttgagcctga gtgtattccc 119520

aaagtgagca ggaagcattc tgctccaaac cacccagggc cacaggcttt ccaaacttgg 119580

caatgtgctc tctggggtgg aagtaggggt actgaccaga agaggtggcc acacaggctg 119640

atgacggcat ccttggctgt gtccaagcag atgttgcact cgaaagtgct gtcctgccct 119700

ccgctctcgc cagcgccatt gctgctccca ctgggccccc ctgcactgga gttctcagga 119760

gatgcagagg ccgagggccc cttgcttgcc atccttggct gtcagcgaag ctctgagtga 119820

agactttccc cagggaaatc ctaaaaggca gagaactctt gactgtccat ttctgcagcc 119880

ctctcccccc agcttccatg ccaacattcc tgtcttctct aggctccaga gctgattatg 119940

actgaatccc tgagcagact tcccaaagcg ctacagaaga caagagttgt agggcactac 120000

ccaaagggca gtgacccaga aacttctcat gaatatggtc cagaatatat cacacagata 120060

cgcaaactcc accaaataca gagcagcctc cagatggttc tccacaggaa ccaggacctt 120120

gtctcacttg tctttagatt cccccatgac aatcaaccca ggtgacaaat acatgctaca 120180

ttaaatggaa tgttcaggct gtccagcctc tatattgatg ggtgaatctg ctgaggcagg 120240

aaggttgtga ttgtcatatg ccagaacata tatatatatt tttaagagat ggagtttcac 120300

tctgtcgccc aggctggagt gcagtggcac aatcttggct cactgcaacc tctgcctccc 120360

aggttcaagt gattctcctg cctcagcctc ccgagtagct gggattacag gtgcctgcca 120420

-continued

ccacacctgg ctaatttttg tatttttagt agagatgggg tttcatcatg ttggccaggc 120480

tggtcttgaa ctcctgacct caggtgatcc acctgcctca gtttcccaaa gtgctaggat 120540

tacaggtgtg agccactgtg cccagcccat acgccagaat attaagaaat tcctgtaact 120600

atgaaaggat aggaagctaa taaggaccag aaagatcatc tggtccaagg tcttccatat 120660

tgcaatcttt ttcaatcctc cttctgccac attggaatac tattcacctt attttatcta 120720

aatatatcaa ctttttagcc tcagtacatt taaagaagag gtttatgtca ctgtaataaa 120780

tggaaaacca gtaatactta ccataagtaa aaaggttaat ataaaagtaa caaatgaaaa 120840

gtttaccctc ataccacctg acatcacctg gcacactcat actatgtttt ggacaatctt 120900

gatctggtcc aatttcttta ggaaactgca gttaccttgg ctagctgact aatgtggggc 120960

tgagactggg actcatggac accctgtttg caccgtacca ctacatacta aactggccac 121020

taaatgctgg tataaacaga ccactgttga ccaccacctc caagtcaaag aacccagcca 121080

ggagttgaac aagaaagagg ataaaaccaa caccaaggga tacccaggaa ttctggatgc 121140

acattaacaa tctgtcagga gccttatctg gccaccctca tcacctcaga aattccaagt 121200

aggaaatgtt cgctggctac cacagtcaga gaagtggagg cctggatact ggactcagcc 121260

tctagcttgc taggtggttc tgggaacctc cctttccctc tactgatcat aattttctca 121320

tctgtaaaat gaggcactgg actccctggt gtcctgggtg gccccctcca gctcttagac 121380

tgtgcctctc caggtcatct taacctactg ttacctccaa aacaagcggc cccttcacaa 121440

atttaggact taccactgga acaccagaag gtcaccctga tgtcagcttc cagagtatga 121500

gaagacaagg accaaaaaag ttacagcctt ctcagaccca gaaagcttcc tgccttcaca 121560

cggtggactc agagctctgg cctcatgagt atccctcacc acgataccag ctcagtgact 121620

tgagcagaga gaggaggcca gctcaccact ttttcactca gttacatcat tttacataat 121680

gtcccagaag ggtagctgga tgctggtggc accagtttcg tgcgctggcc tcggctccaa 121740

gaccttctct ctgaccttcc tcttcttcat attacatgtg ctctcttcat cttaggataa 121800

aacccacttt tgagggccac attactatct ttcctcatta ttccagctca acaatcaatc 121860

ctaagtaccc aacctatgtg tgtttctttt taaaaaaaaa aaaaaatctt aaaaaattaa 121920

aaattgctct ccctcccccc ccccctccct ctccccttg cacggtcctc gtctcccctt 121980

tgcacggtct ccctctgatg ccgagctgag gctggacagt actgccgcca tcttggctca 122040

ctgcagcctc cctgcctgat tctcctgcct cagcctgccg agtgcctggg attgcaggcg 122100

cgcgccgcca cacctgactg gttttcatat tttttggtgg agacggggtt ttgccgtgtt 122160

ggccgggctg gtctccagct cctgaccgcc agtggtctgc cagcctcggc ctcccgaggt 122220

gccaggattg cagacggagt ctcgctcact cagtgctcaa tgttgcccag gctggagtgc 122280

agtggcgtga tctcggatcg ctacaacctc cacctcccag ccgcctgcct tggcctccca 122340

aagtgccgag attgcagcct ctgcccggcc gccaccccat ctaggaagtg aggagcgtct 122400

ctgcctggcc gcccatcgtc tgggatgtga ggagcccctc tgcccagccg cccagtctgg 122460

gaagtgagga gcgcctctcc ccggcggtca tcccgtctag gaagtgagga gagtctctcc 122520

ctggccgccc atcctctggg atgtggggag cgcctctgcc ccgccacccc gtctgagatg 122580

tgaggagtgc ctctgccagg ccgcgacccc gtctgggaac tgaggagtgt ctctgcccca 122640

ccgccacccg tctgggaggt gaggagcgtc tctgacctgc caccctgtct gagaagtgag 122700

-continued

gagcccctct gcccggcagc cgccccgtcc gggaagtgag gagcgtctct gcccggcagc 122760

tgccccgtcc aggaggtggt gggcagcccc cgcccggcca gccgccccgt ccgggaaggg 122820

aggggcagcc cccgaccggc cagccgcccc gtccgggagg tgggggggggg cagcccccgc 122880

ccggccagct gccccgtccg ggagctcggg gcagcccccg cctggacagc tgccccgtcc 122940

gggaggtggg agcccctctg cccggccgcc accccgtctg ggaggtgtac ccaacagctc 123000

attgagaacg ggccatgatg acaatagcgg ttttgtcgaa tagaaaaggg ggaaatgtgg 123060

ggaaaagaaa gagagatcag attgttattg tgtctgtgta gaaagaagta gacatagcag 123120

actccatttt gttctgtact aggaaaaatt cttctgcctt gggatgctgt taatctataa 123180

ccttaccccc aacccctgct ctctgaaaca tgtgctgtgt ccactaaggg ttaaatggat 123240

taagggcggt gcaagatgtg ctttgttaaa cagatgcttg aaggcagcat actcgttaag 123300

agtcatcacc actccctaat ctcaactacc cagggacaca aacactgcgg aaggcggcag 123360

ggccctctgc ctaggaaaac cagagacctt tgttcacatg tttatctgct gaccttccct 123420

ccactattgt cctatgaccc tgccaaatcc ccctctccga gaaacaccca agaatgatca 123480

ataaatacta aaattaaaaa aaaaaaaatt aaaaattatg catatagatt ttggaaagac 123540

aaccataaag tcagatcagt cacatcagcc tcctgatgtg atgcaacaga agtacacagc 123600

accaccaatg aaacatgttt gcgttaaaac tcagacactg atcaagcatc taaaatgaat 123660

taccagccag gcatggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcagg 123720

tggatacctg agatcaggag ttcaagacaa gcctggccaa catggtgaaa ccctgtctct 123780

actaaaaaca caaaaattag ctgggcatgg tggcacacac ctgtaatccc agctacttga 123840

gaagctgagg catgagaatc gcttgaaccc gggaggcgga ggttgcagtg aactgagatt 123900

gtgtcatcgc actccaacct gggtgacaag agagagactc catctcaaaa ataataataa 123960

taataataat aataataaat aaaataaatt acctatttac agaaaatatg gaggataggt 124020

gcaagtatta aacgactcca tgaagatgtc atcagccaaa tccagaatgt gggactttct 124080

acaggacaaa caataaagtt ctacaaaaaa taaatagcat ggacaggatt ggggttatag 124140

attgagatac gagatattcc aaccacttgt agtatgtggg tcttattgga acaagattca 124200

aacaaaccaa ctgtacaaag acacttgtga catgattgac aaaaactgtt tagcctaggt 124260

attactcgat atgaaggaat tattgttggc tgggcatggt ggctcatgtc tataattcca 124320

gcactttggg aggccaaggt aggtggatcg cttgagctca ggaggtggag accagcctgg 124380

gcaacatggt gaaaccctat ctctaccaaa acaaaataaa acaaaacaaa aattagctgg 124440

gtgtggtggt gcatgtctat ggtcccagct actccagagg ctgaggcagg agaatcgctt 124500

aagcttgcga ggcggaagtt gcagtgagcc aagattatac cactgcattc cagcctgagt 124560

gacatagtga ctctgtctca aaaaaaataa ataaataaaa ataaaaaatt agctgggtgt 124620

ggtggtacac acctgtagcc ccagacagtc aggaggctga ggtgggagga tcacttgggc 124680

ctgggaggtg gaggctgcag tgagccgtga ttgtgctgct gcactccagc ctgggtgaca 124740

gagtggaaac cctgtctcaa aaaaaaaaaa aaaaaaagga attactgtta attttgttgg 124800

gtataataat gatattgtgt ttttttttt aaacaaaacc tgatctgcta gagacacatt 124860

tttttaaata tatatatatt tttattatac tttaagttct agggtacatg tgcacaacgt 124920

gcaggtttgt tacatatgta tacatgtgct atgttggtgt gctgcaccca ttaactcgtc 124980

-continued

atttacatta ggtatacctc ctaatgctat cccttccccc tccccccacc ccacaacagg 125040

cccaggtgtg tgatgttccc cttcctgtgt ccaagtgttc tcgttgttca tttcccacct 125100

atgagtgaga acatgaggtg tttggttttt tgtccttgcg atagtttgct gagaatgatg 125160

gtttccagct tcatccacgt ccctacaaag gacatgaact catcattttt tatggctgca 125220

tagtattcca tggtgtatat gtgccacatt ttcttaatcc agtctaccat tgttggacat 125280

ctggcttggt tccaagtctt tactattgtg agtagtcccg caataaacat acgtgtgcat 125340

gtagagacac atatttttga agcaatacaa tgtttgggat ttgctttaaa atactccagg 125400

ggaaaaaaat aaaaatggag agaaaagatg aaacaagaac aaggaagttg gttgttgaag 125460

ctgggtgatg aatggaggtt cactgtaata gtctgtgtgt gtgtttgaaa attacaataa 125520

aaggtatttt ttttttcttt tttttgagat ggtctcgctc tgttgcccag gctggagtgc 125580

agtggtgcaa tcacagctca ctgcaagctc aacctcctgg gctcaagcaa tcccctgcct 125640

cagcctccca agcagctagg actacaggtg cacaccacca tgctcagcta atttttaaat 125700

tattttgtag aaacaaggtc tatgttgccc agggtagtct caaactcctg ggctcaagca 125760

attctcctgc ctcagcctcc cagtgagcca ccatgcctag ccccaataaa aggtattttt 125820

taaaaagatt aagttgtaga agagtatgta tataaactta tttgggtaaa aactaagtat 125880

gtatatgagt atgtatgtgt gtggatccat acacagacag acaggcagac acacacacac 125940

acacacacac acacacacac acacacacac agagagagag agaaaatgcc tagaaaatac 126000

acattaaacc atcaacagtt atttctggga ttataagggc taggagacat tgaattatta 126060

tattcttacc ctgtttcaat gttttataac aaataagtat tactttcaca agaacatttt 126120

taaaaaggaa atcataaaat gtgaagtatt ttacatatgt tacaataatg aatataggta 126180

ttagaaaatg gctttaggcc aggcatggtg gctcatgcct gtaaccctag tactttggga 126240

ggccaaggta ggaggattac ttatgcccag gagtttgagg ctgcagagag ctatgaatgt 126300

accactgcac tccagcctgg gcaattagag cgagactctg attgtattaa aaaaaaaaaa 126360

aaagtcttta acaatgtgga agaacatact caactgtatt actttttttt cttctttttt 126420

tttgagacaa gtctcagtct gtcacccaag ctggaatgca gtggtgtgat ctcggctcac 126480

tgcaacctcc acctcccggg ttcaagcaat tctccctgcc tcagcttcct gagtagctgg 126540

gcttataggt gcccaccacc acgcccagct aatttttgta ctttttagta gagacagggt 126600

ttcgccatgt tggccaggct gctcttgaac tcctgacctc aggtgatcca cccacctcag 126660

ccttccaaag tgctgggatt acacgtgtga gccactgcac ccggccttta catgtattac 126720

ttttgaaatc agtcaaaaga ataataatac ataaaatttc ctacccccca ccccccccgcc 126780

atgttctggt ttccacccca tgttctggtt ttccccaatc ctgaaagttc tgagttcaag 126840

tcctggttcc atcactgact gtccatatgg ccacggtaga tctcttgatg atttgaatct 126900

cttttctcac ctgaaaagtg agagtatcat gcctcctaca tggctactga ggtggcctta 126960

tgaaataaca cagtgacctg gagggccctc caaggcagga cagggctgtt gctgttgggg 127020

aggcacacat ttcataccat gtaacctggt ccactgccac cccacttcgt ggacccaaca 127080

aagctgtgca tggcccatac tcctacacac aggcatgagg ggacaaccag gatcaaggta 127140

ggtgttggag taggaatggt taagggacag taacaggtca gctcagagtt tgcagggagt 127200

ggcagccttg tgccaggcca ggcactgaga caggcctgca tgtcccaact gtaggcaggc 127260

-continued

catatcactc tccttgtaaa tcttagtaga caccacatgg tccccaccca ggatctacag 127320

cctcatccta gcaggcactc aaatatttgc tgagtgagga tgggcacgca tgactgactt 127380

tggtcaactc ccagctggct ggtcaccaac ttttctttgt aacatattct gggccccaca 127440

gctgcatact catccctgga gccacctctt ctcaccttct cctacaccta tgcacaggaa 127500

tctcagtaca ccaaggtggc atggcagaaa tcagcagctt cccctgggga ttgtctactg 127560

agcacctaca gaaccatatg ggtctaccag gactgggaca gtggcccaaa acatagaggt 127620

tcctgatcgg ctcaggaact aggcagaaag agaatcaggt cctgagcata cctatatggc 127680

catcctgcta aaatttctca tggtccctgt agaagctata gactttcacc taaccacaga 127740

gcaaccttgt tgtggctagt tctgtgcctc tgtccctacc ctgaccattc ttttctcctg 127800

accagaatcc actctctagt tttcaaagca tttcacatct atcgctgcat tttatcctca 127860

tgccatttta cagatgagga cactgggcca gggacattag aggactgcct aaggccagga 127920

caccacgcaa tgccaaattc agacttccta acacaatatt cagcactctc ttcattctta 127980

tgtttttcta ttaacatgct gctttacaaa attaaatacc tggtttttgt cttccctcaa 128040

atttatcgtg caatgttgta acttttgtgc tgtttgaaat gagtacagag taaattaatc 128100

actcatccac ttattcattc aaccaaccag cagttttcca atgccaatta catttgaggt 128160

ctggtgctac caactgtgga ggagttacaa agaggaataa gaccaagttg aggtcttgga 128220

gtctagtaca ggagctgaag gcacatagat taattttaag aattctaaac aaggggctcc 128280

atcaaactca cctgtttttc ctttcgttgt tgtttttgcc tgtttttctt aaacaaagca 128340

tcctctcacc ccacgaaaga tctggttccc agatcccttc agagatttaa ataggatctg 128400

gacctagatc aggcctaaga atctgtatgc ttttaatgta gccctcactt tctaccctag 128460

atgattctct aattaacaga gttttggaaa cccagctcta gaagagggta cagctctatc 128520

aatagcacaa agcaatacca acaaaataca gtattttaga aggacagaaa cctcagggcc 128580

tatccagcta caatccacag ggactacaga gaggtcagat aaacttgggg ggtaagggac 128640

aaatctgctg aattcagatt aatcagcatt tagtccaact ggaacacagc tgatcatttt 128700

aagaaataag ttcacaccta gttttaaaaa acagtcctaa ctttgagatg tccagggcaa 128760

cacatgtggt gctgggacta acgcggattt gcaatactta atgatcctgt aggaacacag 128820

agcctggcgg ctgaagttga ctgacccagt atcttcaata tattaccttc ccagagctgg 128880

cctttggtct acttggtctg gcccacctca caggaattga gcaagaaaga gaacacgaga 128940

gttgaatatc cctgcaggaa ccaccaggga acttgatc 128978

<210> SEQ ID NO 7
<211> LENGTH: 3806
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

ctcaagaatg ccagtgtgtg tgtagcctcc acagagaggt cgttttctcg gagtccagag 60

gggccgcctg agcttctgag aactagggag gagccatccc agccatgagc ccctgtggga 120

atctgctggg ggccaagtgg cctggagtcc tcaggctccc gcagctgctc cggagggaga 180

ggtgagctca gggcagcctg cctgcagcca gaggtgccgg gagccccggg cctgtcatgg 240

tggccatcta cagccggcct gaggcagtca cagacggatt tgcagctgag cctgtctatc 300

-continued

tggtgtggga agaagatggg gagttacttg tcagtcccgg cttacttcac ctccagagac 360

ctgtttcggt gttcagaatg ccaggattcc ctcaccaact ggtactatga gaaggatggg 420

aagctctact gccccaagga ctactggggg aagtttgggg agttctgtca tgggtgctcc 480

ctgctgatga cagggccttt tatggtggct ggggagttca agtaccaccc agagtgcttt 540

gcctgtatga gctgcaaggt gatcattgag gatggggatg catatgcact ggtgcagcat 600

gccaccctct actgtgggaa gtgccacaat gaggtggtgc tggcacccat gtttgagaga 660

ctctccacag agtctgttca ggagcagctg ccctactctg tcacgctcat ctccatgccg 720

gccaccactg aaggcaggcg gggcttctcc gtgtccgtgg agagtgcctg ctccaactac 780

gccaccactg tgcaagtgaa agaggtcaac cggatgcaca tcagtcccaa caatcgaaac 840

gccatccacc ctggggaccg catcctggag atcaatggga cccccgtccg cacacttcga 900

gtggaggagg tggaggatgc aattagccag acgagccaga cacttcagct gttgattgaa 960

catgaccccg tctcccaacg cctggaccag ctgcggctgg aggcccggct cgctcctcac 1020

atgcagaatg ccggacaccc ccacgccctc agcaccctgg acaccaagga gaatctggag 1080

gggacactga ggagacgttc cctaaggcgc agtaacagta tctccaagtc ccctggcccc 1140

agctccccaa aggagcccct gctgttcagc cgtgacatca gccgctcaga atcccttcgt 1200

tgttccagca gctattcaca gcagatcttc cggccctgtg acctaatcca tggggaggtc 1260

ctggggaagg gcttctttgg gcaggctatc aaggtgacac acaaagccac gggcaaagtg 1320

atggtcatga aagagttaat tcgatgtgat gaggagaccc agaaaacttt tctgactgag 1380

gtgaaagtga tgcgcagcct ggaccacccc aatgtgctca agttcattgg tgtgctgtac 1440

aaggataaga agctgaacct gctgacagag tacattgagg ggggcacact gaaggacttt 1500

ctgcgcagta tggatccgtt cccctggcag cagaaggtca ggtttgccaa aggaatcgcc 1560

tccggaatgg cctatttgca ctctatgtgc atcatccacc gggatctgaa ctcgcacaac 1620

tgcctcatca agttggacaa gactgtggtg gtggcagact ttgggctgtc acggctcata 1680

gtggaagaga ggaaaagggc ccccatggag aaggccacca ccaagaaacg caccttgcgc 1740

aagaacgacc gcaagaagcg ctacacggtg gtgggaaacc cctactggat ggcccctgag 1800

atgctgaacg gaaagagcta tgatgagacg gtggatatct tctcctttgg gatcgttctc 1860

tgtgagatca ttgggcaggt gtatgcagat cctgactgcc ttccccgaac actggacttt 1920

ggcctcaacg tgaagctttt ctgggagaag tttgttccca cagattgtcc cccggccttc 1980

ttcccgctgg ccgccatctg ctgcagactg gagcctgaga gcagaccagc attctcgaaa 2040

ttggaggact cctttgaggc cctctccctg tacctggggg agctgggcat cccgctgcct 2100

gcagagctgg aggagttgga ccacactgtg agcatgcagt acggcctgac ccgggactca 2160

cctccctagc cctggcccag ccccctgcag gggggtgttc tacagccagc attgcccctc 2220

tgtgccccat tcctgctgtg agcagggccg tccgggcttc ctgtggattg gcggaatgtt 2280

tagaagcaga acaagccatt cctattacct ccccaggagg caagtgggcg cagcaccagg 2340

gaaatgtatc tccacaggtt ctggggccta gttactgtct gtaaatccaa tacttgcctg 2400

aaagctgtga agaagaaaaa aacccctggc ctttgggcca ggaggaatct gttactcgaa 2460

tccacccagg aactccctgg cagtggattg tgggaggctc ttgcttacac taatcagcgt 2520

gacctggacc tgctgggcag gatcccaggg tgaacctgcc tgtgaactct gaagtcacta 2580

-continued

gtccagctgg gtgcaggagg acttcaagtg tgtggacgaa agaaagactg atggctcaaa 2640

gggtgtgaaa aagtcagtga tgctcccccc ttctactcca gatcctgtcc ttcctggagc 2700

aaggttgagg gagtaggttt tgaagagtcc cttaatatgt ggtggaacag gccaggagtt 2760

agagaaaggg ctggcttctg tttacctgct cactggctct agccagccca gggaccacat 2820

caatgtgaga ggaagcctcc acctcatgtt ttcaaactta atactggaga ctggctgaga 2880

acttacggac aacatccttt ctgtctgaaa caaacagtca caagcacagg aagaggctgg 2940

gggactagaa agaggccctg ccctctagaa agctcagatc ttggcttctg ttactcatac 3000

tcgggtgggc tccttagtca gatgcctaaa acattttgcc taaagctcga tgggttctgg 3060

aggacagtgt ggcttgtcac aggcctagag tctgagggag gggagtggga gtctcagcaa 3120

tctcttggtc ttggcttcat ggcaaccact gctcaccctt caacatgcct ggtttaggca 3180

gcagcttggg ctgggaagag gtggtggcag agtctcaaag ctgagatgct gagagagata 3240

gctccctgag ctgggccatc tgacttctac ctcccatgtt tgctctccca actcattagc 3300

tcctgggcag catcctcctg agccacatgt gcaggtactg gaaaacctcc atcttggctc 3360

ccagagctct aggaactctt catcacaact agatttgcct cttctaagtg tctatgagct 3420

tgcaccatat ttaataaatt gggaatgggt ttggggtatt aatgcaatgt gtggtggttg 3480

tattggagca gggggaattg ataaaggaga gtggttgctg ttaatattat cttatctatt 3540

gggtggtatg tgaaatattg tacatagacc tgatgagttg tgggaccaga tgtcatctct 3600

ggtcagagtt tacttgctat atagactgta cttatgtgtg aagtttgcaa gcttgcttta 3660

gggctgagcc ctggactccc agcagcagca cagttcagca ttgtgtggct ggttgtttcc 3720

tggctgtccc cagcaagtgt aggagtggtg ggcctgaact gggccattga tcagactaaa 3780

taaattaagc agttaacata actggc 3806

<210> SEQ ID NO 8
<211> LENGTH: 3699
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

taggcggtgc atcccgttcg cgcctggggc tgtggtcttc ccgcgcctga ggcggcggcg 60

gcaggagctg aggggagttg tagggaactg aggggagctg ctgtgtcccc cgcctcctcc 120

tccccatttc cgggctcccg ggaccatgtc cgcgctggcg ggtgaagatg tctggaggtg 180

tccaggctgt ggggaccaca ttgctccaag ccagatatgg tacaggactg tcaacgaaac 240

ctggcacggc tcttgcttcc ggtgttcaga atgccaggat tccctcacca actggtacta 300

tgagaaggat gggaagctct actgccccaa ggactactgg gggaagtttg gggagttctg 360

tcatgggtgc tccctgctga tgacagggcc ttttatggtg gctggggagt tcaagtacca 420

cccagagtgc tttgcctgta tgagctgcaa ggtgatcatt gaggatgggg atgcatatgc 480

actggtgcag catgccaccc tctactgtgg gaagtgccac aatgaggtgg tgctggcacc 540

catgtttgag agactctcca cagagtctgt tcaggagcag ctgccctact ctgtcacgct 600

catctccatg ccggccacca ctgaaggcag gcggggcttc tccgtgtccg tggagagtgc 660

ctgctccaac tacgccacca ctgtgcaagt gaaagaggtc aaccggatgc acatcagtcc 720

caacaatcga aacgccatcc accctgggga ccgcatcctg gagatcaatg ggacccccgt 780

-continued

ccgcacactt cgagtggagg aggtggagga tgcaattagc cagacgagcc agacacttca 840

gctgttgatt gaacatgacc ccgtctccca acgcctggac cagctgcggc tggaggcccg 900

gctcgctcct cacatgcaga atgccggaca cccccacgcc ctcagcaccc tggacaccaa 960

ggagaatctg gaggggacac tgaggagacg ttccctaagg cgcagtaaca gtatctccaa 1020

gtcccctggc cccagctccc caaaggagcc cctgctgttc agccgtgaca tcagccgctc 1080

agaatccctt cgttgttcca gcagctattc acagcagatc ttccggccct gtgacctaat 1140

ccatggggag gtcctgggga agggcttctt tgggcaggct atcaaggtga cacacaaagc 1200

cacgggcaaa gtgatggtca tgaaagagtt aattcgatgt gatgaggaga cccagaaaac 1260

ttttctgact gaggtgaaag tgatgcgcag cctggaccac cccaatgtgc tcaagttcat 1320

tggtgtgctg tacaaggata agaagctgaa cctgctgaca gagtacattg agggggcac 1380

actgaaggac tttctgcgca gtatggatcc gttcccctgg cagcagaagg tcaggtttgc 1440

caaaggaatc gcctccggaa tggcctattt gcactctatg tgcatcatcc accgggatct 1500

gaactcgcac aactgcctca tcaagttgga caagactgtg gtggtggcag actttgggct 1560

gtcacggctc atagtggaag agaggaaaag ggcccccatg gagaaggcca ccaccaagaa 1620

acgcaccttg cgcaagaacg accgcaagaa gcgctacacg gtggtgggaa acccctactg 1680

gatggcccct gagatgctga acggaaagag ctatgatgag acggtggata tcttctcctt 1740

tgggatcgtt ctctgtgaga tcattgggca ggtgtatgca gatcctgact gccttccccg 1800

aacactggac tttggcctca acgtgaagct tttctgggag aagtttgttc ccacagattg 1860

tcccccggcc ttcttcccgc tggccgccat ctgctgcaga ctggagcctg agagcagacc 1920

agcattctcg aaattggagg actcctttga ggccctctcc ctgtacctgg gggagctggg 1980

catcccgctg cctgcagagc tggaggagtt ggaccacact gtgagcatgc agtacggcct 2040

gacccgggac tcacctccct agccctggcc cagccccctg caggggggtg ttctacagcc 2100

agcattgccc ctctgtgccc cattcctgct gtgagcaggg ccgtccgggc ttcctgtgga 2160

ttggcggaat gtttagaagc agaacaagcc attcctatta cctccccagg aggcaagtgg 2220

gcgcagcacc agggaaatgt atctccacag gttctggggc ctagttactg tctgtaaatc 2280

caatacttgc ctgaaagctg tgaagaagaa aaaaacccct ggcctttggg ccaggaggaa 2340

tctgttactc gaatccaccc aggaactccc tggcagtgga ttgtgggagg ctcttgctta 2400

cactaatcag cgtgacctgg acctgctggg caggatccca gggtgaacct gcctgtgaac 2460

tctgaagtca ctagtccagc tgggtgcagg aggacttcaa gtgtgtggac gaaagaaaga 2520

ctgatggctc aaagggtgtg aaaaagtcag tgatgctccc cctttctact ccagatcctg 2580

tccttcctgg agcaaggttg agggagtagg ttttgaagag tcccttaata tgtggtggaa 2640

caggccagga gttagagaaa gggctggctt ctgtttacct gctcactggc tctagccagc 2700

ccagggacca catcaatgtg agaggaagcc tccacctcat gttttcaaac ttaatactgg 2760

agactggctg agaacttacg gacaacatcc tttctgtctg aaacaaacag tcacaagcac 2820

aggaagaggc tgggggacta gaaagaggcc ctgccctcta gaaagctcag atcttggctt 2880

ctgttactca tactcgggtg ggctccttag tcagatgcct aaaacatttt gcctaaagct 2940

cgatgggttc tggaggacag tgtggcttgt cacaggccta gagtctgagg gaggggagtg 3000

ggagtctcag caatctcttg gtcttggctt catggcaacc actgctcacc cttcaacatg 3060

-continued

cctggtttag gcagcagctt gggctgggaa gaggtggtgg cagagtctca aagctgagat 3120

gctgagagag atagctccct gagctgggcc atctgacttc tacctcccat gtttgctctc 3180

ccaactcatt agctcctggg cagcatcctc ctgagccaca tgtgcaggta ctggaaaacc 3240

tccatcttgg ctcccagagc tctaggaact cttcatcaca actagatttg cctcttctaa 3300

gtgtctatga gcttgcacca tatttaataa attgggaatg ggtttggggt attaatgcaa 3360

tgtgtggtgg ttgtattgga gcaggggaa ttgataaagg agagtggttg ctgttaatat 3420

tatcttatct attgggtggt atgtgaaata ttgtacatag acctgatgag ttgtgggacc 3480

agatgtcatc tctggtcaga gtttacttgc tatatagact gtacttatgt gtgaagtttg 3540

caagcttgct ttagggctga gccctggact cccagcagca gcacagttca gcattgtgtg 3600

gctggttgtt tcctggctgt ccccagcaag tgtaggagtg gtgggcctga actgggccat 3660

tgatcagact aaataaatta agcagttaac ataactggc 3699

<210> SEQ ID NO 9
<211> LENGTH: 2805
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

gggaggagcc atcccagcca tgagcccctg tgggaatctg ctgggggcca agtggcctgg 60

agtcctcagg ctcccgcagc tgctccggag ggagaggtga gctcagggca gcctgcctgc 120

agccagaggt gccgggagcc ccgggcctgt catggtggcc atctacagcc ggcctgaggc 180

agtcacagac ggatttgcag ctgagcctgt ctatctggtg tgggaagaag atggggagtt 240

acttgtcagt cccggcttac ttcacctcca gagacctgtt tcggtgttca gaatgccagg 300

attccctcac caactggtac tatgagaagg atgggaagct ctactgcccc aaggactact 360

gggggaagtt tggggagttc tgtcatgggt gctccctgct gatgacaggg ccttttatgg 420

tggctgggga gttcaagtac cacccagagt gctttgcctg tatgagctgc aaggtgatca 480

ttgaggatgg ggatgcatat gcactggtgc agcatgccac cctctactgt gggaagtgcc 540

acaatgaggt ggtgctggca cccatgtttg agagactctc cacagagtct gttcaggagc 600

agctgcccta ctctgtcacg ctcatctcca tgccggccac cactgaaggc aggcggggct 660

tctccgtgtc cgtggagagt gcctgctcca actacgccac cactgtgcaa gtgaaagagg 720

tcaaccggat gcacatcagt cccaacaatc gaaacgccat ccaccctggg gaccgcatcc 780

tggagatcaa tgggaccccc gtccgcacac ttcgagtgga ggaggtggag gatgcaatta 840

gccagacgag ccagacactt cagctgttga ttgaacatga ccccgtctcc caacgcctgg 900

accagctgcg gctggaggcc cggctcgctc ctcacatgca gaatgccgga cacccccacg 960

ccctcagcac cctggacacc aaggagaatc tggaggggac actgaggaga cgttccctaa 1020

ggcgcagtaa cagtatctcc aagtcccctg gccccagctc cccaaaggag cccctgctgt 1080

tcagccgtga catcagccgc tcagaatccc ttcgttgttc cagcagctat tcacagcaga 1140

tcttccggcc ctgtgaccta atccatgggg aggtcctggg gaagggcttc tttgggcagg 1200

ctatcaaggt gacacacaaa gccacgggca aagtgatggt catgaaagag ttaattcgat 1260

gtgatgagga gacccagaaa acttttctga ctgaggtgaa agtgatgcgc agcctggacc 1320

accccaatgt gctcaagttc attggtgtgc tgtacaagga taagaagctg aacctcctga 1380

-continued

cagagtacat tgaggggggc acactgaagg actttctgcg cagtatggat ccgttcccct 1440

ggcagcagaa ggtcaggttt gccaaaggaa tcgcctccgg aatggcctat ttgcactcta 1500

tgtgcatcat ccaccgggat ctgaactcgc acaactgcct catcaagttg gacaagactg 1560

tggtggtggc agactttggg ctgtcacggc tcatagtgga agaggaggaa agggccccca 1620

tggagaaggc caccaccaag aaacgcacct tgcgcaagaa cgaccgcaag aagcgctaca 1680

cggtggtggg aaacccctac tggatggccc ctgagatgct gaacggaaag agctatgatg 1740

agacggtgga tatcttctcc tttgggatcg ttctctgtga gatcattggg caggtgtatg 1800

cagatcctga ctgccttccc cgaacactgg actttggcct caacgtgaag cttttctggg 1860

agaagtttgt tcccacagat tgtcccccgg ccttcttccc gctggccgcc atctgctgca 1920

gactggagcc tgagagcaga gccccccccg gggccgcagg agagggcccg ggctgcgcgg 1980

atgatgaggg cccagtgagg cgccaaggga aggtcaccat caagtatgac cccaaggagc 2040

tacggaagca cctcaaccta gaggagtgga tcctggagca gctcacgcgc ctctacgact 2100

gccaggaaga ggagatctca gaactagaga ttgacgtgga tgagctcctg gacatggaga 2160

gtgacgatgc ctgggcttcc agggtcaagg agctgctggt tgactgttac aaacccacag 2220

aggccttcat ctctggcctg ctggacaaga tccgggccat gcagaagctg agcacacccc 2280

agaagaagtg agggtccccg acccaggcga acggtggctc ccataggaca atcgctaccc 2340

cccgacctcg tagcaacagc aataccgggg gaccctgcgg ccaggcctgg ttccatgagc 2400

agggctcctc gtgcccctgg cccaggggtc tcttcccctg ccccctcagt tttccacttt 2460

tggatttttt tattgttatt aaactgatgg gactttgtgt ttttatattg actctgcggc 2520

acgggccctt taataaagcg aggtagggta cgcctttggt gcagctcaaa aaaaaaaaaa 2580

aaaatgattt ccagcggtcc acattagagt tgaaattttc tggtgggaga atctatacct 2640

tgttccttta taggccaagg accgcagtcc ttcagtaaca ccagtgtaaa agcttgagga 2700

gaaattgtga agctacacag tatttgtttt ctaatacctc ttgtcattct aaatatcttt 2760

aatttattaa aaaatatata tatacagtaa aaaaaaaaaa aaaaa 2805

<210> SEQ ID NO 10
<211> LENGTH: 3581
<212> TYPE: DNA
<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 10

gggctgtaga ggcctctgca cctcctcctc tggcctgagg ggggcatacc cgcgtctaca 60

gacgcggggc gcaacctgcc gaaaggggtc acaccgctgg gcgagagcta ggggaggagt 120

caccttgaga ggaggggtca caatgattag ccgagtcact ctgaaaggtc tcttttagtg 180

aaaggaacac aaagaggagg gggtgcagca ctaaggggct ctggatacac actgatgtgc 240

cagtgtcatc ccaaggaaag gtccccctga aggagagcgg gctccgaggg gcagcggcct 300

ttgaggggggc gggatcctgt cgaggggtgg gctgtgacgg agggcggggc tttcttgcag 360

gtgcaggctt cccctaagag ggttattagg acaaagaagt aggagggtgg tcaccgggcg 420

gagcgcagga atctcaactt tactcagggg aatggagatg gagtgaggag gggtcgggcc 480

aggtcaggga tgcggaggcc agaggctccc tacagccctg ggtaccacga ccggagaaaa 540

caggtgttag cgagccccga gatttagggt gtcggggtca gaagcccctt gaaatgggca 600

-continued

gtcggagtgg ggcgcagact gggcaggcag cgggcttgat ccgtggagcc cagggcgcag 660

gagaattcgc gcctgcagtc aacccttaag tcgtcctgtt attcacgccc cggcctacgc 720

gccgggacgc acctgagggg gcgtggccgg gccacagctc gggctccaag gagttaaccg 780

cagacccggc cggccacgcc ccgggcaggc cgcgcagcct gagcccggcg gcgaagcgga 840

gcagcagccg cccgcgggga gcccgccgcc ggcgcccggg agcgccgccg ccgcccggcc 900

caggcccagg cccaggcccc gcgcagagcg gagagccgcg gcgttccacc aggcccaact 960

cgccggcgcg gcgggggcgg gcccaggacc ttcggcaggt cctcccggag tccgccgggc 1020

gggagtcccg gcacccggga tcaggttggg ccggcgccca tggcgagcgc ggcctgcccg 1080

ggtcccgggg atccggccat gtgaagcagg ccggggctgg gggcccggcc atggccgggg 1140

aacggccacc gctgcggggc cccgggcccg gagagactcc ggtggagggg ccgggggcg 1200

caggcggagg cccgggccgg ggccggccct cctcctaccg ggccctccgc agcgccgtgt 1260

ccagcctggc gcgcgtggac gatttcgact gcgcagagaa gatcggggct ggtttcttct 1320

ctgaggtcta taaggttcgg caccgacagt cgggacaagt catggtgctg aagatgaaca 1380

agctccccag taaccggagc aacactctaa gggaggtgca gctgatgaac cggctccggc 1440

atcctaacat cctcaggttc atgggggtct gtgtgcacca agggcagctg cacgcgctta 1500

cagagtatat gaatgggggt accctggaac agctgctgag ctccccagaa cccctctcct 1560

ggccagtcag gctccaccta gcactggaca ttgcacaagg cctacggtac ctacacgcca 1620

aaggtgtgtt tcaccgagac ctcacatcca agaactgtct ggtccgacgg gaagacggag 1680

gcttcacagc tgttgtgggt gacttcggac tggctgaaaa gatccctgtg tatagggaag 1740

gggcaaggaa ggagccattg gctgtggtgg gctccccata ctggatggcc ccagaggtgt 1800

tgcggggaga gctgtatgat gagaaggccg acgtcttcgc ctttgggatt gtcctctgtg 1860

agctcatcgc ccgagtacct gcagaccctg actacctacc ccgtactgag gactttggtc 1920

tggatgtgcc tgcttttcga actctggtgg gaaatgactg tccgctacct ttcctgcttc 1980

tggccatcca ctgctgcagc atggaaccca gcgcccgtgc ccctttttact gaaattaccc 2040

agcacctgga acaaatcctg gagcagctgc ctgagcccac gcccctcgcc aagatgcccc 2100

ttgccaaggc tcccttgaca tacaatcagg ggtctgttcc aagaggaggt ccctctgcca 2160

cgcttcccag gtcagacccc cggctctccc gaagccggtc agacctcttc ctgccacctt 2220

caccagaatc acccccccagt tgggggggaca atctaacacg agtcaacccc ttctcactgc 2280

gggaagacct caggggtggc aagatcaaac ttctggacac accctgcaag ccggccactc 2340

cactgcccct ggtgccacca tcaccactga cctctaccca gctgcccttg gtggcctctc 2400

cagagtccct ggtccagcct gagactcccg tccgccgctg tcgttcactc ccttcatccc 2460

ctgagctccc ccgacggatg gaaactgcac tgccaggtcc tggcccgtcc cccgtgggcc 2520

cctcgactga ggaaaggatg gactgtgagg gcagcagccc tgagccagag cccccagggc 2580

cggctcccca gctgcctctg gctgtggcca ctgacaactt catcagtacg tgttcctcag 2640

cctctcagcc ctggtctgct agaccaggac cttccctcaa taataacccc cctgctgtgg 2700

tggtgaattc cccacaagga tgggcgaggg agccctggaa ccgggcacag catagccttc 2760

cccgggcagc agccctggag cggacagaac cctcaccacc cccatcagct ccccgggagc 2820

aggaggaggg gctgccctgc cctggctgct gcctcagccc cttcagcttt ggcttcctat 2880

-continued

ccatgtgccc ccgccctacc ccagctgttg cccgctaccg caacttgaac tgtgaggcgg 2940

gcagtcttct ctgccaccga gggcatcatg ccaagccacc cacacccagc ctgcagctgc 3000

ctggggcacg ctcttagcac tgaggcctgt gatctcaggc ttccaacctt ggccttcggg 3060

ataccctgta aggacagagt gcacttgctg gacaatgcca gccccagatg ggctgactgg 3120

ctcttctccc cgtatagggg agcctcagca tggtctcaat ggacagagca cttcctgtcc 3180

aacccctggg atcactggac tagacacagc attgctgaca cacaagacta acacgtgcaa 3240

aatatttaaa aatatttcaa taaaactgcc tggctggctc cgggccatcc atccactcag 3300

cggagcgccc tcccctgacc catctctagt tcgggaagtt cttaaggaag gtcaaagtct 3360

tttctggagg cttctttact cttcccaaga atagtttctg gagcagcttt caagcaagca 3420

gcaggtcgcc agaccagaat cagcccttag ctgggcgtgc ccaagagtga gggagtgctg 3480

gaggaatagc agtcttgaga attggcaggc cataggttaa ggatgtaggt cacaaggtct 3540

tgttcaagtg tctggaaaaa taaagcacca gtgagctccc c 3581

<210> SEQ ID NO 11
<211> LENGTH: 2443
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

caccgccacc gcctggtcca ggcccaagcc caggccctgc gcggagcgga gcggcgcggc 60

gtccacccgg gcccagccag ccggcgcggc gggggcgggt ccaggatctt cggcggatct 120

tcggcggatc ttcggcggat cttccattct cagggcggga gccggagtcc gggcgcccgg 180

gatcgggctg ggcccgcgcc catggcaagc gcggcctgcc cgggccctgg ggaccctgcc 240

atgtgaggca ggcccgggct gggggcccgg ccatggccgg ggaacggccc ccactgcggg 300

gccctgggcc cgggcctgga gaggtgccgg gggaggggcc cccggggccg gggggcacgg 360

gcggaggccc gggccggggc cgcccctcct cctaccgggt tctccgcagc gccgtgtcta 420

gcctggcgcg tgtggacgat tttcactgcg cggagaagat cggggccggc ttcttctctg 480

aggtctacaa ggttcggcac cgacagtcag ggcaagtcat ggtgctgaag atgaacaagc 540

tccccagtaa ccggggcaac acactacggg aagtgcagct gatgaaccgg ctcaggcacc 600

ccaacatcct aaggttcatg ggagtctgtg tgcaccaggg acagctgcac gctcttacag 660

agtatatgaa tgggggggaca ttggaacagc tgctcagctc ccctgaaccc ctatcctggc 720

cggtcaggct ccacctggcc ctggacattg cccgaggcct gcggtacctg cactccaaag 780

gtgtatttca ccgcgacctc acatccaaga actgtctagt ccgacgggaa gatcgaggct 840

tcaccgctgt cgtgggtgac ttcgggctgg ccgaaaagat tcctgtgtat agggaggggg 900

caaggaagga gccattggcc gtggtgggct ccccatactg gatggctcca gaggtgttac 960

ggggtgagct gtatgatgag aaggctgatg tctttgcctt cgggattgtc ctctgtgagc 1020

tcatcgcccg agtacctgca gaccctgact acctaccacg cactgaggac tttggcctgg 1080

atgtgcctgc tttccgaact ctggtggggg atgactgccc actgcctttc ttgctcctgg 1140

ccatccactg ctgcaacctg gaacccagca cccgtgcccc cttcaccgaa attacccagc 1200

acctggaatg gatcctggag cagctgcctg agccagcccc actcaccagg accgccctga 1260

cacacaatca gggctctgtt gcaagagggg gtccctctgc cacgcttccc aggccagatc 1320

-continued

cccggctttc ccgaagccgg tcagacctct tcctgccccc atcaccagaa tcacccccca 1380

actgggggga caatctgact cgagtcaacc ccttctcact acgggaagac ctcaggggtg 1440

gcaagatcaa gctcttagac acacccagca agccagtcct gcctcttgtg ccaccatcac 1500

cattcccatc cactcagctg cccttggtga ccactccgga gaccctggtc cagcctggga 1560

cacctgcccg ccgctgccgc tcactaccct catccccga gctcccccgc cgtatggaga 1620

cagcactgcc aggtcctggc cctcccgctg tgggcccctc ggctgaagag aagatggagt 1680

gcgagggcag cagccctgag ccggaacctc cagggccagc gccccagctg cctctggctg 1740

tggccacaga caacttcatc agcacctgtt cctcggcctc ccaaccctgg tcccctagat 1800

caggacccgt cctcaataac aatcccccag ctgtggtggt gaactcccca caaggctggg 1860

ctggggagcc ctggaaccgg gcccagcata gcctgccccg ggcggcagcc ctggagcgga 1920

cagaaccctc gccacccccc tcagctcccc gggagcccga tgaggggctg ccctgtcctg 1980

gctgctgcct cggccccttc agctttggct tcctgtccat gtgcccccgc cccacaccag 2040

ctgttgcccg ctaccgcaac ctgaactgtg aggcgggcag tctcctctgc caccgagggc 2100

accacgccaa gccacccaca cccagcctgc agctgcctgg ggcacgctct tagcagtgga 2160

gcccgtggtc tcaggcctcc aactttggcc ttcaggacac cctgtaagaa cagagcacac 2220

ttgctggaca gtgccagttc cagatgggct gaccggctct tctccccgtg taggggagcc 2280

ccagcatgga ctcaagggac agagcacttc cagtcgaccc cccggctcgc gttcccgtgg 2340

ggatcactga accagacaca gcattgctga cacatgagac taacacgtgc aattatttaa 2400

aaagatttca ataaaactgc ctggctggct ccgggccgcc cct 2443

<210> SEQ ID NO 12
<211> LENGTH: 2421
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

caccgccacc gcccggtcca ggcccaagcc caggccctgc gcggagcgga gcggcgcggc 60

gtccacccgg gcccagccag ccggcgcggc gggggcgggt ccaggatctt cggcggatct 120

tccattctca gggcgggagc cggagtccgg gcgcccggga tcgggctggg cccgcgccca 180

tggcaagcgc ggcctgcccg ggccctgggg accctgccat gtgaggcagg cccgggctgg 240

gggcccggcc atggccgggg aacggccccc actgcggggc cctgggcccg ggcctggaga 300

ggtgccgggg gaggggcccc cggggccggg gggcacgggc ggaggcccgg gccggggccg 360

cccctcctcc taccgggctc tccgcagcgc cgtgtctagc ctggcgcgtg tggacgattt 420

tcactgcgcg gagaagatcg gggccggctt cttctctgag gtctacaagg ttcggcaccg 480

acagtcaggg caagtcatgg tgctgaagat gaacaagctc cccagtaacc ggggcaacac 540

actacgggaa gtgcagctga tgaaccggct caggcacccc aacatcctaa ggttcatggg 600

agtctgtgtg caccagggac agctgcacgc tcttacagag tatatgaatg gggggacatt 660

ggaacagctg ctcagctccc ctgaacccct gtcctggccg gtcaggctcc acctggccct 720

ggacattgcc cgaggcctgc ggtacctgca ctccaaaggt gtatttcacc gcgacctcac 780

atccaagaac tgtctagtcc gacgggaaga tcgaggcttc accgctgtcg tgggtgactt 840

cgggctggcc gaaaagattc ctgtgtatag ggaggggggca aggaaggagc cattggccgt 900

-continued

ggtgggctcc ccatactgga tggctccaga ggtgttacgg ggtgagctgt atgatgagaa 960

ggctgatgtc tttgccttcg ggattgtcct ctgtgagctc atcgcccgag tacctgcaga 1020

ccctgactac ctaccacgca ctgaggactt tggcctggat gtgcctgctt tccgaactct 1080

ggtgggggat gactgcccac tgcctttctt gctcctggcc atccactgct gcaacctgga 1140

acccagcacc cgtgccccct tcaccgaaat tacccagcac ctggaatgga tcctggagca 1200

gctgcctgag ccagccccac tcaccaggac cgccctgaca cacaatcagg gctctgttgc 1260

aagaggggt ccctctgcca cgcttcccag gccagatccc cggctttccc gaagccggtc 1320

agacctcttc ctgcccccat caccagaatc acccccccaac tgggggggaca atctgactcg 1380

agtcaacccc ttctcactac gggaagacct caggggtggc aagatcaagc tcttagacac 1440

acccagcaag ccagtcctgc ctcttgtgcc accatcacca ttcccatcca ctcagctgcc 1500

cttggtgacc actccggaga ccctggtcca gcctgggaca cctgcccgcc gctgccgctc 1560

actaccctca tcccccgagc tcccccgccg tatggagaca gcactgccag gtcctggccc 1620

tcccgctgtg ggcccctcgg ctgaagagaa gatggagtgc gagggcagca gccctgagcc 1680

ggaacctcca gggccagcgc cccagctgcc tctggctgtg gccacagaca acttcatcag 1740

cacctgttcc tcggcctccc aaccctggtc ccctagatca ggacccgtcc tcaataacaa 1800

tcccccagct gtggtggtga actccccaca aggctgggct ggggagccct ggaaccgggc 1860

ccagcatagc ctgccccggg cggcagccct ggagcggaca gaaccctcgc caccccttc 1920

agctccccgg gagcccgatg aggggctgcc ctgtcctggc tgctgcctcg gcccccttcag 1980

ctttggcttc ctgtccatgt gcccccgccc cacaccagct gttgcccgct accgcaacct 2040

gaactgtgag gcgggcagtc tcctctgcca ccgagggcac cacgccaagc cacccacacc 2100

cagcctgcag ctgcctgggg cacgctctta gcagtggagc ccgtggtctc aggcctccaa 2160

ctttggcctt caggacaccc tgtaagaaca gagcacactt gctggacagt gccagttcca 2220

gatgggctga ccggctcttc tccccgtgta ggggagcccc agcatggact caagggacag 2280

agcacttcca gtcgaccccc cggctcgcgt tcccgtgggg atcactgaac cagacacagc 2340

attgctgaca catgagacta acacgtgcaa ttatttaaaa agatttcaat aaaactgcct 2400

ggctggctcc gggccgcccc t 2421

<210> SEQ ID NO 13
<211> LENGTH: 2299
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

ggcacgaggc cattctcagg gcgggagccg gagtccgggc gcccgggatc gggctgggcc 60

cgcgcccatg gcaagcgcgg cctgcccggg ccctggggac cctgccatgt gaggcaggcc 120

cgggctgggg gcccggccat ggccgggggaa cggcccccac tgcggggccc tgggcccggg 180

cctggagagg tgccgggggga ggggccccccg gggccgggggg gcacgggcgg aggcccgggc 240

cggggccgcc cctcctccta ccgggctctc cgcagcgccg tgtctagcct ggcgcgtgtg 300

gacgattttc actgcgcgga gaagatcggg gccggcttct tctctgaggt ctacaaggtt 360

cggcaccgac agtcagggca agtcatggtg ctgaagatga acaagctccc cagtaaccgg 420

ggcaacacac tacgggaagt gcagctgatg aaccggctca ggcaccccaa catcctaagg 480

-continued

ttcatgggag tctgtgtgca ccagggacag ctgcacgctc ttacagagta tatgaatggg 540

gggacattgg aacagctgct cagctcccct gaacccctat cctggccggt caggctccac 600

ctggccctgg acattgcccg aggcctgcgg tacctgcact ccaaaggtgt atttcaccgc 660

gacctcacat ccaagggagg gggcaaggaa ggagccattg gccgtggtgg gctccccata 720

ctggatggct ccagaggtgt tacggggtga gctgtatgat gagaaggctg atgtctttgc 780

cttcgggatt gtcctctgtg agctcatcgc ccgagtacct gcagaccctg actacctacc 840

acgcactgag gactttggcc tggatgtgcc tgctttccga actctggtgg gggatgactg 900

cccactgcct ttcttgctcc tggccatcca ctgctgcaac ctggaaccca gcacccgtgc 960

ccccttcacc gaaattaccc agcacctgga atggatcctg gagcagctgc ctgagccagc 1020

cccactcacc aggaccgccc tgacacacaa tcagggctct gttgcaagag ggggtccctc 1080

tgccacgctt cccaggccag atccccggct ttcccgaagc cggtcagacc tcttcctgcc 1140

cccatcacca gaatcacccc ccaactgggg ggacaatctg actcgagtca acccccttctc 1200

actacgggaa gacctcaggg gtggcaagat caagctctta gacacaccca gcaagccagt 1260

cctgcctctt gtgccaccat caccattccc atccactcag ctgcccttgg tgaccactcc 1320

ggagaccctg gtccagcctg ggacacctgc ccgccgctgc cgctcactac cctcatcccc 1380

cgagctcccc cgccgtatgg agacagcact gccaggtcct ggccctcccg ctgtgggccc 1440

ctcggctgaa gagaagatgg agtgcgaggg cagcagccct gagccggaac ctccagggcc 1500

agcgccccag ctgcctctgg ctgtggccac agacaacttc atcagcacct gttcctcggc 1560

ctcccaaccc tggtccccta gatcaggacc cgtcctcaat aacaatcccc cagctgtggt 1620

ggtgaactcc ccacaaggct gggctgggga gccctggaac cgggcccagc atagcctgcc 1680

ccgggcggca gccctggagc ggacagaacc ctcgccaccc ccttcagctc cccgggagcc 1740

cgatgagggg ctgccctgtc ctggctgctg cctcggcccc ttcagctttg gcttcctgtc 1800

catgtgcccc cgccccacac cagctgttgc ccgctaccgc aacctgaact gtgaggcggg 1860

cagtctcctc tgccaccgag ggcaccacgc caagccaccc acacccagcc tgcagctgcc 1920

tggggcacgc tcttagcagt ggagcccgtg gtctcaggcc tccaactttg gccttcagga 1980

caccctgtaa gaacagagca cacttgctgg acagtgccag ttccagatgg gctgaccggc 2040

tcttctcccc gtgtagggga gccccagcat ggactcaagg gacagagcac ttccagtcga 2100

ccccccggct cgcgttcccg tggggatcac tgaaccagac acagcattgc tgacacatga 2160

gactaacacg tgcaattatt taaaaagatt tcaataaaac tgcctggctg gctccgggcc 2220

gcccctatcc gctcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2280

aaaaaaaaaa aaaaaaaaa 2299

<210> SEQ ID NO 14
<211> LENGTH: 2443
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

caccgccacc gcctggtcca ggcccaagcc caggccctgc gcggagcgga gcggcgcggc 60

gtccacccgg gcccagccag ccggcgcggc gggggcgggt ccaggatctt cggcggatct 120

tcggcggatc ttcggcggat cttccattct cagggcggga gccggagtcc gggcgcccgg 180

-continued

gatcgggctg ggcccgcgcc catggcaagc gcggcctgcc cgggccctgg ggaccctgcc 240

atgtgaggca ggcccgggct gggggcccgg ccatggccgg ggaacggccc ccactgcggg 300

gccctgggcc cgggcctgga gaggtgccgg gggaggggcc cccggggccg gggggcacgg 360

gcggaggccc gggccggggc cgcccctcct cctaccgggt tctccgcagc gccgtgtcta 420

gcctggcgcg tgtggacgat tttcactgcg cggagaagat cggggccggc ttcttctctg 480

aggtctacaa ggttcggcac cgacagtcag ggcaagtcat ggtgctgaag atgaacaagc 540

tccccagtaa ccggggcaac acactacggg aagtgcagct gatgaaccgg ctcaggcacc 600

ccaacatcct aaggttcatg ggagtctgtg tgcaccaggg acagctgcac gctcttacag 660

agtatatgaa tgggggggaca ttggaacagc tgctcagctc ccctgaaccc ctatcctggc 720

cggtcaggct ccacctggcc ctggacattg cccgaggcct gcggtacctg cactccaaag 780

gtgtatttca ccgcgacctc acatccaaga actgtctagt ccgacgggaa gatcgaggct 840

tcaccgctgt cgtgggtgac ttcgggctgg ccgaaaagat tcctgtgtat agggaggggg 900

caaggaagga gccattggcc gtggtgggct ccccatactg gatggctcca gaggtgttac 960

ggggtgagct gtatgatgag aaggctgatg tctttgcctt cgggattgtc ctctgtgagc 1020

tcatcgcccg agtacctgca gaccctgact acctaccacg cactgaggac tttggcctgg 1080

atgtgcctgc tttccgaact ctggtggggg atgactgccc actgcctttc ttgctcctgg 1140

ccatccactg ctgcaacctg gaacccagca cccgtgcccc cttcaccgaa attacccagc 1200

acctggaatg gatcctggag cagctgcctg agccagcccc actcaccagg accgccctga 1260

cacacaatca gggctctgtt gcaagagggg gtccctctgc cacgcttccc aggccagatc 1320

cccggctttc ccgaagccgg tcagacctct tcctgccccc atcaccagaa tcacccccca 1380

actgggggga caatctgact cgagtcaacc ccttctcact acgggaagac ctcaggggtg 1440

gcaagatcaa gctcttagac acacccagca agccagtcct gcctcttgtg ccaccatcac 1500

cattcccatc cactcagctg cccttggtga ccactccgga gaccctggtc cagcctggga 1560

cacctgcccg ccgctgccgc tcactaccct catcccccga gctcccccgc cgtatggaga 1620

cagcactgcc aggtcctggc cctcccgctg tgggcccctc ggctgaagag aagatggagt 1680

gcgagggcag cagccctgag ccggaacctc cagggccagc gccccagctg cctctggctg 1740

tggccacaga caacttcatc agcacctgtt cctcggcctc ccaaccctgg tcccctagat 1800

caggacccgt cctcaataac aatcccccag ctgtggtggt gaactcccca caaggctggg 1860

ctggggagcc ctggaaccgg gcccagcata gcctgccccg ggcggcagcc ctggagcgga 1920

cagaaccctc gccacccccct tcagctcccc gggagcccga tgaggggctg ccctgtcctg 1980

gctgctgcct cggccccttc agctttggct tcctgtccat gtgcccccgc cccacaccag 2040

ctgttgcccg ctaccgcaac ctgaactgtg aggcgggcag tctcctctgc caccgagggc 2100

accacgccaa gccacccaca cccagcctgc agctgcctgg ggcacgctct tagcagtgga 2160

gcccgtggtc tcaggcctcc aactttggcc ttcaggacac cctgtaagaa cagagcacac 2220

ttgctggaca gtgccagttc cagatgggct gaccggctct tctccccgtg taggggagcc 2280

ccagcatgga ctcaagggac agagcacttc cagtcgaccc cccggctcgc gttcccgtgg 2340

ggatcactga accagacaca gcattgctga cacatgagac taacacgtgc aattatttaa 2400

aaagatttca ataaaactgc ctggctggct ccgggccgcc cct 2443

-continued

<210> SEQ ID NO 15
<211> LENGTH: 3016
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

gaattcgcgg ccgctcgacg ctcagcagag ctaccagctg ccctgttggc ttcgctggtc 60

ggatcgtcct cctggccccg ccaaacaggc gagcggcccc gactgtgggg catggcagta 120

gtctcctcgt tctccgccgc cgctagccta gctgagtcgc cggcttctgc gctaggggct 180

cccaccgcct ccgcaggcta aggagccgct gccaccaacg agctgtgagg gttactatgc 240

tccctctttg ccgccgtctc ctcctcttgc ccgcgcaggc acccctctgg ctgctcagtc 300

ctgcctcagt gtcaaaccag aagagaagta aaattcaaca aaaatttatg tgtggagttc 360

cttcttaaaa gaagaaaaaa gtgattattt agactatgga tcggagcaaa cggaattcaa 420

ttgcaggatt tcctccacgt gtggagcgtc ttgaagagtt tgaaggaggt ggtggaggag 480

aaggaaatgt gagccaggtg ggaagagttt ggccatcttc gtatcgagct cttataagtg 540

ccttttccag actgacgcgt ttggatgatt tcacctgtga aaaaataggg tctggcttct 600

tttctgaagt gttcaaggta cgacaccgag cttctggtca ggtgatggct cttaagatga 660

acacattgag cagtaaccgg gcaaacatgc tgaaagaagt acagctcatg aatagactct 720

cccatcccaa catccttagg tatatcaact ccgggaacct ggaacagttg ctagacagta 780

acctgcattt gccttggact gtgagggtaa aactggccta tgacatagca gtgggcctca 840

gctaccttca cttcaaaggc atttttcatc gggacctcac atctaagaac tgcctgataa 900

agagggatga gaatggttac tctgcagtgg tagctgactt tggcctggct gagaagatcc 960

ccgatgtcag catggggagt gagaagctgg ccgtggtggg ttccccattc tggatggcac 1020

ctgaggttct ccgagatgag ccctataatg aaaaggcaga tgtgttctct tatggtatca 1080

tcctctgcga gatcatcgcc cgcatccagg ccgatccgga ctatcttccc cgcacagaga 1140

atttcgggct ggactatgat gctttccagc acatggtggg agactgtccc ccagattttc 1200

tgcaacttac tttcaactgc tgtaacatgg atcccaaact gcgcccatct tttgtggaga 1260

ttgggaagac cctggaggaa attctgagcc gcctacagga agaagagcag gagagggata 1320

ggaagctgca gcccacagcc aggggactct tggagaaagc acctggggtg aagcgactaa 1380

gctcactgga tgacaagatc ccccacaagt caccatgccc aagacgtacc atctggctgt 1440

ctcgaagcca gtcagatatc ttttcccgta agccccccacg tacagtgagt gtcttggacc 1500

catactaccg gccacgagat ggtgctgccc gcacccccaa agtcaaccct tttagtgctc 1560

gccaggacct catgggggggc aagatcaagt tttttgacct gcccagcaag tctgtcatct 1620

ctctggtatt tgacctggat gcaccagggc ccggaactat gcccctggct gactggcagg 1680

agcccctggc cccacctatt cgccggtggc gttccttgcc tggttcgcct gagttcttgc 1740

atcaagaggc ttgtccattt gtgggccggg aagaatcgct atctgatggg cccccaccac 1800

gcctaagtag tctcaagtac agagttaaag agatcccacc attccgggca tctgccctac 1860

cagctgctca agcccatgag gctatggact gctccattct ccaggaagaa aatggttttg 1920

ggtccaggcc ccaggggacc agtccatgcc ctgcgggtgc ttctgaggag atggaggtag 1980

aagaaaggcc agcaggctca actccagcca ccttctccac ctcaggcata ggcctgcaaa 2040

-continued

cccagggaaa gcaggatggg tgagggggtt tagtccctgc ctcaccttgg ggatggacct 2100

tcagctgaaa ccatatggcc ccctaggtgc acagccttga ttcttccctg gagcctacag 2160

agcaggcagg ctaggccaag ccaggctcaa cttctgggct cccagtgccc attggctgtg 2220

tatgacggga ggcagcagtg agaggccttc ctagttaggg ccaacagctg ataccaagcc 2280

tctgaaatcc agcaaggagg tctgcctccc accagaccct ctccagtgta cttccccaga 2340

taggaccaga ggatgtctag ttctaggctg agctggcagg cagctattac cccggttctt 2400

tccccacccc aggtctgtct cttgcctttt cttggggcat ataagctact gagtggaaca 2460

tggagctgat caagaggccg taatggtcat ggctgtttcc agacctgaat attgggtgct 2520

tcttgccagt attctaagac atttgagtaa ttgctgtttg cacttactgc atggtcagac 2580

cacgtcacta catttctatg caaggggaca gcaaggcagc gtggtggtca tggctcttag 2640

ctaacctatt caaagacctt ttcctgttga ttaatctatt ttcatattta taaaggagtc 2700

ttaatgttct gccccataag actttcaacc ttgtggttgg gagtggggct ggttttgtag 2760

gccctagggc ctgcttctat gtatttatca acatgtgata cattcaattg gttaaatggt 2820

ttatacaggg actgatttgc ttcccttcct gccatggctg gagctttggg aacagtctgt 2880

ccttacagag ctgcaataag aaataaccaa agatgaagct ggtcaaatat tttcataact 2940

tgcttctgtt gatttttttt ttgtaaaact ttcccaagac attttcagac ttaaaaataa 3000

agtcagtgtt acaggt 3016

<210> SEQ ID NO 16
<211> LENGTH: 2280
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

gtatatcaac tccgggaacc tggaacagtt gctagacagt aacctgcatt tgccttggac 60

tgtgagggta aaactggcct atgacatagc agtgggcctc agctaccttc acttcaaagg 120

catttttcat cgggacctca catctaagaa ctgcctgata aagagggatg agaatggtta 180

ctctgcagtg gtagctgact ttggcctggc tgagaagatc cccgatgtca gcatggggag 240

tgagaagctg gccgtggtgg gttccccatt ctggatggca cctgaggttc tccgagatga 300

gccctataat gaaaaggcag atgtgttctc ttatggtatc atcctctgcg agatcatcgc 360

ccgcatccag gccgatccgg actatcttcc ccgcacagag aatttcgggc tggactatga 420

tgctttccag cacatggtgg gagactgtcc cccagatttt ctgcaactta ctttcaactg 480

ctgtaacatg gatcccaaac tgcgcccatc ttttgtggag attgggaaga ccctggagga 540

aattctgagc cgcctacagg aagaagagca ggagagggat aggaagctgc agcccacagc 600

caggggactc ttggagaaag cacctggggt gaagcgacta agctcactgg atgacaagat 660

cccccacaag tcaccatgcc caagacgtac catctggctg tctcgaagcc agtcagatat 720

cttttcccgt aagcccccac gtacagtgag tgtcttggac ccatactacc ggccacgaga 780

tggtgctgcc cgcaccccca aagtcaaccc ttttagtgct cgccaggacc tcatgggggg 840

caagatcaag ttttttgacc tgcccagcaa gtctgtcatc tctctggtat ttgacctgga 900

tgcaccaggg cccggaacta tgcccctggc tgactggcag gagcccctgg ccccacctat 960

tcgccggtgg cgttccttgc ctggttcgcc tgagttcttg catcaagagg cttgtccatt 1020

-continued

tgtgggccgg gaagaatcgc tatctgatgg gcccccacca cgcctaagta gtctcaagta 1080

cagagttaaa gagatcccac cattccgggc atctgcccta ccagctgctc aagcccatga 1140

ggctatggac tgctccattc tccaggaaga aaatggtttt gggtccaggc cccaggggac 1200

cagtccatgc cctgcgggtg cttctgagga gatggaggta gaagaaaggc cagcaggctc 1260

aactccagcc accttctcca cctcaggcat aggcctgcaa acccagggaa agcaggatgg 1320

gtgagggggt ttagtccctg cctcaccttg gggatggacc ttcagctgaa accatatggc 1380

cccctaggtg cacagccttg attcttccct ggagcctaca gagcaggcag gctaggccaa 1440

gccaggctca acttctgggc tcccagtgcc cattggctgt gtatgacggg aggcagcagt 1500

gagaggcctt cctagttagg gccaacagct gataccaagc ctctgaaatc cagcaaggag 1560

gtctgcctcc caccagaccc tctccagtgt acttccccag ataggaccag aggatgtcta 1620

gttctaggct gagctggcag gcagctatta ccccggttct ttccccaccc caggtctgtc 1680

tcttgccttt tcttggggca tataagctac tgagtggaac atggagctga tcaagaggcc 1740

gtaatggtca tggctgtttc ccagacctga atattggggt gcttcttgcc agtattctaa 1800

gacatttgag taattgctgt ttgcacttac tgcatggtca gaccacgtca ctacatttct 1860

atgcaagggg acagcaaggc agcgtggtgg tcatggctct tagctaacct attcaaagac 1920

cttttcctgt tgattaatct attttcatat ttataaagga gtcttaatgt tctgccccat 1980

aagactttca accttgtggt tgggagtggg gctggttttg taggccctag ggcctgcttc 2040

tatgtattta tcaacatgtg atacattcaa ttggttaaat ggtttataca gggactgatt 2100

tgcttccctt cctgccatgg ctggagcttt gggaacagtc tgtccttaca gagctgcaat 2160

aagaaataac caaagatgaa gctggtcaaa tattttcata acttgcttct gttgattttt 2220

ttttttgtaa aactttccca agacattttc agacttaaaa ataaagtcag tgttacaggt 2280

<210> SEQ ID NO 17
<211> LENGTH: 3071
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

gaattcgcgg ccgctcgacg ctcagcagag ctaccagctg ccctgttggc ttcgctggtc 60

ggatcgtcct cctggccccg ccaaacaggc gagcggcccc gactgtgggg catggcagta 120

gtctcctcgt tctccgccgc cgctagccta gctgagtcgc cggcttctgc gctaggggct 180

cccaccgcct ccgcaggcta aggagccgct gccaccaacg agctgtgagg gttactatgc 240

tccctctttg ccgccgtctc ctcctcttgc ccgcgcaggc acccctctgg ctgctcagtc 300

ctgcctcagt gtcaaaccag aagagaagta aaattcaaca aaaatttatg tgtggagttc 360

cttcttaaaa gaagaaaaaa gtgattattt agactatgga tcggagcaaa cggaattcaa 420

ttgcaggatt tcctccacgt gtggagcgtc ttgaagagtt tgaaggaggt ggtggaggag 480

aaggaaatgt gagccaggtg ggaagagttt ggccatcttc gtatcgagct cttataagtg 540

ccttttccag actgacgcgt ttggatgatt tcacctgtga aaaaatagggg tctggcttct 600

tttctgaagt gttcaaggta cgacaccgag cttctggtca ggtgatggct cttaagatga 660

acacattgag cagtaaccgg gcaaacatgc tgaaagaagt acagctcatg aatagactct 720

cccatcccaa catccttagg ttcatgggtg tatgtgttca tcaaggacaa ttgcatgcac 780

-continued

ttacagagta tatcaactcc gggaacctgg aacagttgct agacagtaac ctgcatttgc 840

cttggactgt gagggtaaaa ctggcctatg acatagcagt gggcctcagc taccttcact 900

tcaaaggcat ttttcatcgg gacctcacat ctaagaactg cctgataaag agggatgaga 960

atggttactc tgcagtggta gctgactttg gcctggctga gaagatcccc gatgtcagca 1020

tggggagtga gaagctggcc gtggtgggtt ccccattctg gatggcacct gaggttctcc 1080

gagatgagcc ctataatgaa aaggcagatg tgttctctta tggtatcatc ctctgcgaga 1140

tcatcgcccg catccaggcc gatccggact atcttccccg cacagagaat ttcgggctgg 1200

actatgatgc tttccagcac atggtgggag actgtccccc agattttctg caacttactt 1260

tcaactgctg taacatggat cccaaactgc gcccatcttt tgtggagatt gggaagaccc 1320

tggaggaaat tctgagccgc ctacaggaag aagagcagga gagggatagg aagctgcagc 1380

ccacagccag gggactcttg gagaaagcac ctggggtgaa gcgactaagc tcactggatg 1440

acaagatccc ccacaagtca ccatgcccaa gacgtaccat ctggctgtct cgaagccagt 1500

cagatatctt ttcccgtaag cccccacgta cagtgagtgt cttggaccca tactaccggc 1560

cacgagatgg tgctgcccgc acccccaaag tcaacccttt tagtgctcgc caggacctca 1620

tgggggcaa gatcaagttt tttgacctgc ccagcaagtc tgtcatctct ctggtatttg 1680

acctggatgc accagggccc ggaactatgc ccctggctga ctggcaggag cccctggccc 1740

cacctattcg ccggtggcgt tccttgcctg gttcgcctga gttcttgcat caagaggctt 1800

gtccatttgt gggccgggaa gaatcgctat ctgatgggcc cccaccacgc ctaagtagtc 1860

tcaagtacag agttaaagag atcccaccat tccgggcatc tgccctacca gctgctcaag 1920

cccatgaggc tatggactgc tccattctcc aggaagaaaa tggttttggg tccaggcccc 1980

aggggaccag tccatgccct gcgggtgctt ctgaggagat ggaggtagaa gaaaggccag 2040

caggctcaac tccagccacc ttctccacct caggcatagg cctgcaaacc cagggaaagc 2100

aggatgggtg aggggttta gtccctgcct caccttgggg atggaccttc agctgaaacc 2160

atatggcccc ctaggtgcac agccttgatt cttccctgga gcctacagag caggcaggct 2220

aggccaagcc aggctcaact tctgggctcc cagtgcccat tggctgtgta tgacgggagg 2280

cagcagtgag aggccttcct agttagggcc aacagctgat accaagcctc tgaaatccag 2340

caaggaggtc tgcctcccac cagaccctct ccagtgtact tccccagata ggaccagagg 2400

atgtctagtt ctaggctgag ctggcaggca gctattaccc cggttctttc cccaccccag 2460

gtctgtctct tgccttttct tggggcatat aagctactga gtggaacatg gagctgatca 2520

agaggccgta atggtcatgg ctgtttccag acctgaatat tgggtgcttc ttgccagtat 2580

tctaagacat ttgagtaatt gctgtttgca cttactgcat ggtcagacca cgtcactaca 2640

tttctatgca aggggacagc aaggcagcgt ggtggtcatg gctcttagct aacctattca 2700

aagacctttt cctgttgatt aatctatttt catatttata aaggagtctt aatgttctgc 2760

cccataagac tttcaacctt gtggttggga gtggggctgg ttttgtaggc cctagggcct 2820

gcttctatgt atttatcaac atgtgataca ttcaattggt taaatggttt atacagggac 2880

tgatttgctt cccttcctgc catggctgga gctttgggaa cagtctgtcc ttacagagct 2940

gcaataagaa ataaccaaag atgaagctgg tcaaatattt tcataacttg cttctgttga 3000

ttttttttt gtaaaacttt cccaagacat tttcagactt aaaaataaag tcagtgttac 3060

-continued

aggtaaaaaa a 3071

<210> SEQ ID NO 18
<211> LENGTH: 3050
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

agagagctac cagctgccct gttggcttcg ctggtcggat cgtcctcctg gccccgccaa 60

acaggcgggg ggagcggccc cgactgtggg gccatggcag tagtctcctc gttctccgcc 120

gccgctagcc tagctgagtc gccggcttct gcgctagggg ctcccaccgc ctccgcaggc 180

taaggagccg ctgccaccaa cgagctgtga gggttactat gctccctctt tgccgccgtc 240

tcctcctctt gcccgcgcag gcacccctct ggctgctcag tcctgcctca gtgtcaaacc 300

agaagagaag taaaattcaa caaaaattta tgtgtggagt tccttcttaa aagaagaaaa 360

aagtgattat ttagactatg gatcggagca aacggaattc aattgcagga tttcctccac 420

gtgtggagcg tcttgaagag tttgaaggag gtggtggagg agaaggaaat gtgagccagg 480

tgggaagagt ttggccatct tcgtatcgag ctcttataag tgccttttcc agactgacgc 540

gtttggatga tttcacctgt gaaaaaatag ggtctggctt cttttctgaa gtgttcaagg 600

tacgacaccg agcttctggt caggtgatgg ctcttaagat gaacacattg agcagtaacc 660

gggcaaacat gctgaaagaa gtacagctca tgaatagact ctcccatccc aacatcctta 720

ggttcatggg tgtatgtgtt catcaaggac aattgcatgc acttacagag tatatcaact 780

ccgggaacct ggaacagttg ctagacagta acctgcattt gccttggact gtgagggtaa 840

aactggccta tgacatagca gtgggcctca gctaccttca cttcaaaggc atttttcatc 900

gggacctcac atctaagaac tgcctgataa agagggatga gaatggttac tctgcagtgg 960

tagctgactt tggcctggct gagaagatcc ccgatgtcag catggggagt gagaagctgg 1020

ccgtggtggg ttccccattc tggatggcac ctgaggttct ccgagatgag ccctataatg 1080

aaaaggcaga tgtgttctct tatggtatca tcctctgcga gatcatcgcc cgcatccagg 1140

ccgatccgga ctatcttccc cgcacagaga atttcgggct ggactatgat gctttccagc 1200

acatggtggg agactgtccc ccagattttc tgcaacttac tttcaactgc tgtaacatgg 1260

atcccaaact gcgcccatct tttgtggaga ttgggaagac cctggaggaa attctgagcc 1320

gcctacagga agaagagcag gagagggata ggaagctgca gcccacagcc aggggactct 1380

tggagaaagc acctggggtg aagcgactaa gctcactgga tgacaagatc ccccacaagt 1440

caccatgccc aagacgtacc atctggctgt ctcgaagcca gtcagatatc ttttcccgta 1500

agcccccacg tacagtgagt gtcttggacc catactaccg gccacgagat ggtgctgccc 1560

gcacccccaa agtcaaccct tttagtgctc gccaggacct catgggggggc aagatcaagt 1620

tttttgacct gcccagcaag tctgtcatct ctctggtatt tgacctggat gcaccagggc 1680

ccggaactat gcccctggct gactggcagg agcccctggc cccacctatt cgccggtggc 1740

gttccttgcc tggttcgcct gagttcttgc atcaagaggc ttgtccattt gtgggccggg 1800

aagaatcgct atctgatggg cccccaccac gcctaagtag tctcaagtac agagttaaag 1860

agatcccacc attccgggca tctgccctac cagctgctca agcccatgag gctatggact 1920

gctccattct ccaggaagaa aatggttttg ggtccaggcc ccaggggacc agtccatgcc 1980

-continued

ctgcgggtgc ttctgaggag atggaggtag aagaaaggcc agcaggctca actccagcca 2040

ccttctccac ctcaggcata ggcctgcaaa cccagggaaa gcaggatggg tgaggggttt 2100

tagtccctgc ctcaccttgg ggatggacct tcagctgaaa ccatatggcc ccctaggtgc 2160

acagccttga ttcttccctg gagcctacag agcaggcagg ctaggccaag ccaggctcaa 2220

cttctgggct cccagtgccc attggctgtg tatgacggga ggcagcagtg agaggccttc 2280

ctagttaggg ccaacagctg ataccaagcc tctgaaatcc agcaaggagg tctgcctccc 2340

accagaccct ctccagtgta cttccccaga taggaccaga ggatgtctag ttctaggctg 2400

agctggcagg cagctattac cccggttctt tccccacccc aggtctgtct cttgcctttt 2460

cttggggcat ataagctact gagtggaaca tggagctgat caagaggccg taacggtcat 2520

ggctgtttcc cagacctgaa tattggggtg cttcttgcca gtattctaag acatttgagt 2580

aattgctgtt tgcacttact gcatggtcag accacgtcac tacatttcta tgcaagggga 2640

cagcaaggca gcgtggtggt catggctctt agctaaccta ttcaaagacc ttttcctgtt 2700

gattaatcta ttttcatatt tataaaggag tcttaatgtt ctgccccata agactttcaa 2760

ccttgtggtt gggagtgggg ctggttttgt aggccctagg gcctgcttct atgtatttat 2820

caacatgtga tacattcaat tggttaaatg gtttatacag ggactgattt gcttcccttc 2880

ctgccatggc tggagctttg ggaacagtct gtccttacag agctgcaata agaaataacc 2940

aaagatgaag ctggtcaaat attttcataa cttgcttctg ttgatttttt tttttgtaaa 3000

actttcccaa gacattttca gacttaaaaa taaagtcagt gttacaggtg 3050

<210> SEQ ID NO 19
<211> LENGTH: 702
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

Ala Pro Ser Arg Phe Pro Arg Arg Cys Pro Arg Gly Ala Pro Gly Pro
1 5 10 15

Ala Arg Pro Gly Pro Arg Pro Pro Pro Pro Pro Pro Arg Ala Gly Gly
20 25 30

Gly Gly Arg Ala Arg Gly Pro Gly Arg Pro Gln Pro Gln Pro Arg Arg
35 40 45

Ala Pro Pro Pro Val Glu Cys Met Arg Leu Thr Leu Leu Cys Cys Thr
50 55 60

Trp Arg Glu Glu Arg Met Gly Glu Glu Gly Ser Glu Leu Pro Val Cys
65 70 75 80

Ala Ser Cys Gly Gln Arg Ile Tyr Asp Gly Gln Tyr Leu Gln Ala Leu
85 90 95

Asn Ala Asp Trp His Ala Asp Cys Phe Arg Cys Cys Asp Cys Ser Ala
100 105 110

Ser Leu Ser His Gln Tyr Tyr Glu Lys Asp Gly Gln Leu Phe Cys Lys
115 120 125

Lys Asp Tyr Trp Ala Arg Tyr Gly Glu Ser Cys His Gly Cys Ser Glu
130 135 140

Gln Ile Thr Lys Gly Leu Val Met Val Ala Gly Glu Leu Lys Tyr His
145 150 155 160

Pro Glu Cys Phe Ile Cys Leu Thr Cys Gly Thr Phe Ile Gly Asp Gly
165 170 175

-continued

```
Asp Thr Tyr Thr Leu Val Glu His Ser Lys Leu Tyr Cys Gly His Cys
            180                 185                 190

Tyr Tyr Gln Thr Val Val Thr Pro Val Ile Glu Gln Ile Leu Pro Asp
        195                 200                 205

Ser Pro Gly Ser His Leu Pro His Thr Val Thr Leu Val Ser Ile Pro
    210                 215                 220

Ala Ser Ser His Gly Lys Arg Gly Leu Ser Val Ser Ile Asp Pro Pro
225                 230                 235                 240

His Gly Pro Pro Gly Cys Gly Thr Glu His Ser His Thr Val Arg Val
                245                 250                 255

Gln Gly Val Asp Pro Gly Cys Met Ser Pro Asp Val Lys Asn Ser Ile
            260                 265                 270

His Val Gly Asp Arg Ile Leu Glu Ile Asn Gly Thr Pro Ile Arg Asn
        275                 280                 285

Val Pro Leu Asp Glu Ile Asp Leu Leu Ile Gln Glu Thr Ser Arg Leu
    290                 295                 300

Leu Gln Leu Thr Leu Glu His Asp Pro His Asp Thr Leu Gly His Gly
305                 310                 315                 320

Leu Gly Pro Glu Thr Ser Pro Leu Ser Ser Pro Ala Tyr Thr Pro Ser
                325                 330                 335

Gly Glu Ala Gly Ser Ser Ala Arg Gln Lys Pro Val Leu Arg Ser Cys
            340                 345                 350

Ser Ile Asp Arg Ser Pro Gly Ala Gly Ser Leu Gly Ser Pro Ala Ser
        355                 360                 365

Gln Arg Lys Asp Leu Gly Arg Ser Glu Ser Leu Arg Val Val Cys Arg
    370                 375                 380

Pro His Arg Ile Phe Arg Pro Ser Asp Leu Ile His Gly Glu Val Leu
385                 390                 395                 400

Gly Lys Gly Cys Phe Gly Gln Ala Ile Lys Val Thr His Arg Glu Thr
                405                 410                 415

Gly Glu Val Met Val Met Lys Glu Leu Ile Arg Phe Asp Glu Glu Thr
            420                 425                 430

Gln Arg Thr Phe Leu Lys Glu Val Lys Val Met Arg Cys Leu Glu His
        435                 440                 445

Pro Asn Val Leu Lys Phe Ile Gly Val Leu Tyr Lys Asp Lys Arg Leu
    450                 455                 460

Asn Phe Ile Thr Glu Tyr Ile Lys Gly Gly Thr Leu Arg Gly Ile Ile
465                 470                 475                 480

Lys Ser Met Asp Ser Gln Tyr Pro Trp Ser Gln Arg Val Ser Phe Ala
                485                 490                 495

Lys Asp Ile Ala Ser Gly Met Ala Tyr Leu His Ser Met Asn Ile Ile
            500                 505                 510

His Arg Asp Leu Asn Ser His Asn Cys Leu Val Arg Glu Asn Lys Asn
        515                 520                 525

Val Val Val Ala Asp Phe Gly Leu Ala Arg Leu Met Val Asp Glu Lys
    530                 535                 540

Thr Gln Pro Glu Gly Leu Arg Ser Leu Lys Lys Pro Asp Arg Lys Lys
545                 550                 555                 560

Arg Tyr Thr Val Val Gly Asn Pro Tyr Trp Met Ala Pro Glu Met Ile
                565                 570                 575
```

-continued

```
Asn Gly Arg Ser Tyr Asp Glu Lys Val Asp Val Phe Ser Phe Gly Ile
            580                 585                 590

Val Leu Cys Glu Ile Ile Gly Arg Val Asn Ala Asp Pro Asp Tyr Leu
        595                 600                 605

Pro Arg Thr Met Asp Phe Gly Leu Asn Val Arg Gly Phe Leu Asp Arg
    610                 615                 620

Tyr Cys Pro Pro Asn Cys Pro Pro Ser Phe Phe Pro Ile Thr Val Arg
625                 630                 635                 640

Cys Cys Asp Leu Asp Pro Glu Lys Arg Pro Ser Phe Val Lys Leu Glu
                645                 650                 655

His Trp Leu Glu Thr Leu Arg Met His Leu Ala Gly His Leu Pro Leu
            660                 665                 670

Gly Pro Gln Leu Glu Gln Leu Asp Arg Gly Phe Trp Glu Thr Tyr Arg
        675                 680                 685

Arg Gly Glu Ser Gly Leu Pro Ala His Pro Glu Val Pro Asp
    690                 695                 700


<210> SEQ ID NO 20
<211> LENGTH: 647
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

Met Arg Leu Thr Leu Leu Cys Cys Thr Trp Arg Glu Glu Arg Met Gly
1               5                   10                  15

Glu Glu Gly Ser Glu Leu Pro Val Cys Ala Ser Cys Gly Gln Arg Ile
            20                  25                  30

Tyr Asp Gly Gln Tyr Leu Gln Ala Leu Asn Ala Asp Trp His Ala Asp
        35                  40                  45

Cys Phe Arg Cys Cys Asp Cys Ser Ala Ser Leu Ser His Gln Tyr Tyr
    50                  55                  60

Glu Lys Asp Gly Gln Leu Phe Cys Lys Lys Asp Tyr Trp Ala Arg Tyr
65                  70                  75                  80

Gly Glu Ser Cys His Gly Cys Ser Glu Gln Ile Thr Lys Gly Leu Val
                85                  90                  95

Met Val Ala Gly Glu Leu Lys Tyr His Pro Glu Cys Phe Ile Cys Leu
            100                 105                 110

Thr Cys Gly Thr Phe Ile Gly Asp Gly Asp Thr Tyr Thr Leu Val Glu
        115                 120                 125

His Ser Lys Leu Tyr Cys Gly His Cys Tyr Tyr Gln Thr Val Val Thr
    130                 135                 140

Pro Val Ile Glu Gln Ile Leu Pro Asp Ser Pro Gly Ser His Leu Pro
145                 150                 155                 160

His Thr Val Thr Leu Val Ser Ile Pro Ala Ser Ser His Gly Lys Arg
                165                 170                 175

Gly Leu Ser Val Ser Ile Asp Pro Pro His Gly Pro Pro Gly Cys Gly
            180                 185                 190

Thr Glu His Ser His Thr Val Arg Val Gln Gly Val Asp Pro Gly Cys
        195                 200                 205

Met Ser Pro Asp Val Lys Asn Ser Ile His Val Gly Asp Arg Ile Leu
    210                 215                 220

Glu Ile Asn Gly Thr Pro Ile Arg Asn Val Pro Leu Asp Glu Ile Asp
225                 230                 235                 240
```

-continued

```
Leu Leu Ile Gln Glu Thr Ser Arg Leu Leu Gln Leu Thr Leu Glu His
                245                 250                 255

Asp Pro His Asp Thr Leu Gly His Gly Leu Gly Pro Glu Thr Ser Pro
            260                 265                 270

Leu Ser Ser Pro Ala Tyr Thr Pro Ser Gly Glu Ala Gly Ser Ser Ala
        275                 280                 285

Arg Gln Lys Pro Val Leu Arg Ser Cys Ser Ile Asp Arg Ser Pro Gly
    290                 295                 300

Ala Gly Ser Leu Gly Ser Pro Ala Ser Gln Arg Lys Asp Leu Gly Arg
305                 310                 315                 320

Ser Glu Ser Leu Arg Val Val Cys Arg Pro His Arg Ile Phe Arg Pro
                325                 330                 335

Ser Asp Leu Ile His Gly Glu Val Leu Gly Lys Gly Cys Phe Gly Gln
            340                 345                 350

Ala Ile Lys Val Thr His Arg Glu Thr Gly Glu Val Met Val Met Lys
        355                 360                 365

Glu Leu Ile Arg Phe Asp Glu Glu Thr Gln Arg Thr Phe Leu Lys Glu
    370                 375                 380

Val Lys Val Met Arg Cys Leu Glu His Pro Asn Val Leu Lys Phe Ile
385                 390                 395                 400

Gly Val Leu Tyr Lys Asp Lys Arg Leu Asn Phe Ile Thr Glu Tyr Ile
                405                 410                 415

Lys Gly Gly Thr Leu Arg Gly Ile Ile Lys Ser Met Asp Ser Gln Tyr
            420                 425                 430

Pro Trp Ser Gln Arg Val Ser Phe Ala Lys Asp Ile Ala Ser Gly Met
        435                 440                 445

Ala Tyr Leu His Ser Met Asn Ile Ile His Arg Asp Leu Asn Ser His
    450                 455                 460

Asn Cys Leu Val Arg Glu Asn Lys Asn Val Val Val Ala Asp Phe Gly
465                 470                 475                 480

Leu Ala Arg Leu Met Val Asp Glu Lys Thr Gln Pro Glu Gly Leu Arg
                485                 490                 495

Ser Leu Lys Lys Pro Asp Arg Lys Lys Arg Tyr Thr Val Val Gly Asn
            500                 505                 510

Pro Tyr Trp Met Ala Pro Glu Met Ile Asn Gly Arg Ser Tyr Asp Glu
        515                 520                 525

Lys Val Asp Val Phe Ser Phe Gly Ile Val Leu Cys Glu Ile Ile Gly
    530                 535                 540

Arg Val Asn Ala Asp Pro Asp Tyr Leu Pro Arg Thr Met Asp Phe Gly
545                 550                 555                 560

Leu Asn Val Arg Gly Phe Leu Asp Arg Tyr Cys Pro Pro Asn Cys Pro
                565                 570                 575

Pro Ser Phe Phe Pro Ile Thr Val Arg Cys Cys Asp Leu Asp Pro Glu
            580                 585                 590

Lys Arg Pro Ser Phe Val Lys Leu Glu His Trp Leu Glu Thr Leu Arg
        595                 600                 605

Met His Leu Ala Gly His Leu Pro Leu Gly Pro Gln Leu Glu Gln Leu
    610                 615                 620

Asp Arg Gly Phe Trp Glu Thr Tyr Arg Arg Gly Glu Ser Gly Leu Pro
625                 630                 635                 640
```

-continued

Ala His Pro Glu Val Pro Asp
645

<210> SEQ ID NO 21
<211> LENGTH: 638
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

Met Ser Ala Leu Ala Gly Glu Asp Val Trp Arg Cys Pro Gly Cys Gly
1 5 10 15

Asp His Ile Ala Pro Ser Gln Ile Trp Tyr Arg Thr Val Asn Glu Thr
20 25 30

Trp His Gly Ser Cys Phe Arg Cys Ser Glu Cys Gln Asp Ser Leu Thr
35 40 45

Asn Trp Tyr Tyr Glu Lys Asp Gly Lys Leu Tyr Cys Pro Lys Asp Tyr
50 55 60

Trp Gly Lys Phe Gly Glu Phe Cys His Gly Cys Ser Leu Leu Met Thr
65 70 75 80

Gly Pro Phe Met Val Ala Gly Glu Phe Lys Tyr His Pro Glu Cys Phe
85 90 95

Ala Cys Met Ser Cys Lys Val Ile Ile Glu Asp Gly Asp Ala Tyr Ala
100 105 110

Leu Val Gln His Ala Thr Leu Tyr Cys Gly Lys Cys His Asn Glu Val
115 120 125

Val Leu Ala Pro Met Phe Glu Arg Leu Ser Thr Glu Ser Val Gln Glu
130 135 140

Gln Leu Pro Tyr Ser Val Thr Leu Ile Ser Met Pro Ala Thr Thr Glu
145 150 155 160

Gly Arg Arg Gly Phe Ser Val Ser Val Glu Ser Ala Cys Ser Asn Tyr
165 170 175

Ala Thr Thr Val Gln Val Lys Glu Val Asn Arg Met His Ile Ser Pro
180 185 190

Asn Asn Arg Asn Ala Ile His Pro Gly Asp Arg Ile Leu Glu Ile Asn
195 200 205

Gly Thr Pro Val Arg Thr Leu Arg Val Glu Glu Val Glu Asp Ala Ile
210 215 220

Ser Gln Thr Ser Gln Thr Leu Gln Leu Leu Ile Glu His Asp Pro Val
225 230 235 240

Ser Gln Arg Leu Asp Gln Leu Arg Leu Glu Ala Arg Leu Ala Pro His
245 250 255

Met Gln Asn Ala Gly His Pro His Ala Leu Ser Thr Leu Asp Thr Lys
260 265 270

Glu Asn Leu Glu Gly Thr Leu Arg Arg Arg Ser Leu Arg Arg Ser Asn
275 280 285

Ser Ile Ser Lys Ser Pro Gly Pro Ser Ser Pro Lys Glu Pro Leu Leu
290 295 300

Phe Ser Arg Asp Ile Ser Arg Ser Glu Ser Leu Arg Cys Ser Ser Ser
305 310 315 320

Tyr Ser Gln Gln Ile Phe Arg Pro Cys Asp Leu Ile His Gly Glu Val
325 330 335

Leu Gly Lys Gly Phe Phe Gly Gln Ala Ile Lys Val Thr His Lys Ala
340 345 350

-continued

```
Thr Gly Lys Val Met Val Met Lys Glu Leu Ile Arg Cys Asp Glu Glu
        355                 360                 365

Thr Gln Lys Thr Phe Leu Thr Glu Val Lys Val Met Arg Ser Leu Asp
    370                 375                 380

His Pro Asn Val Leu Lys Phe Ile Gly Val Leu Tyr Lys Asp Lys Lys
385                 390                 395                 400

Leu Asn Leu Leu Thr Glu Tyr Ile Glu Gly Gly Thr Leu Lys Asp Phe
                405                 410                 415

Leu Arg Ser Met Asp Pro Phe Pro Trp Gln Gln Lys Val Arg Phe Ala
            420                 425                 430

Lys Gly Ile Ala Ser Gly Met Ala Tyr Leu His Ser Met Cys Ile Ile
        435                 440                 445

His Arg Asp Leu Asn Ser His Asn Cys Leu Ile Lys Leu Asp Lys Thr
    450                 455                 460

Val Val Val Ala Asp Phe Gly Leu Ser Arg Leu Ile Val Glu Glu Arg
465                 470                 475                 480

Lys Arg Ala Pro Met Glu Lys Ala Thr Thr Lys Lys Arg Thr Leu Arg
                485                 490                 495

Lys Asn Asp Arg Lys Lys Arg Tyr Thr Val Val Gly Asn Pro Tyr Trp
            500                 505                 510

Met Ala Pro Glu Met Leu Asn Gly Lys Ser Tyr Asp Glu Thr Val Asp
        515                 520                 525

Ile Phe Ser Phe Gly Ile Val Leu Cys Glu Ile Ile Gly Gln Val Tyr
    530                 535                 540

Ala Asp Pro Asp Cys Leu Pro Arg Thr Leu Asp Phe Gly Leu Asn Val
545                 550                 555                 560

Lys Leu Phe Trp Glu Lys Phe Val Pro Thr Asp Cys Pro Pro Ala Phe
                565                 570                 575

Phe Pro Leu Ala Ala Ile Cys Cys Arg Leu Glu Pro Glu Ser Arg Pro
            580                 585                 590

Ala Phe Ser Lys Leu Glu Asp Ser Phe Glu Ala Leu Ser Leu Tyr Leu
        595                 600                 605

Gly Glu Leu Gly Ile Pro Leu Pro Ala Glu Leu Glu Glu Leu Asp His
    610                 615                 620

Thr Val Ser Met Gln Tyr Gly Leu Thr Arg Asp Ser Pro Pro
625                 630                 635


<210> SEQ ID NO 22
<211> LENGTH: 638
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

Met Ser Ala Leu Ala Gly Glu Asp Val Trp Arg Cys Pro Gly Cys Gly
1               5                   10                  15

Asp His Ile Ala Pro Ser Gln Ile Trp Tyr Arg Thr Val Asn Glu Thr
            20                  25                  30

Trp His Gly Ser Cys Phe Arg Cys Ser Glu Cys Gln Asp Ser Leu Thr
        35                  40                  45

Asn Trp Tyr Tyr Glu Lys Asp Gly Lys Leu Tyr Cys Pro Lys Asp Tyr
    50                  55                  60

Trp Gly Lys Phe Gly Glu Phe Cys His Gly Cys Ser Leu Leu Met Thr
```

-continued

```
65                  70                  75                  80

Gly Pro Phe Met Val Ala Gly Glu Phe Lys Tyr His Pro Glu Cys Phe
                85                  90                  95

Ala Cys Met Ser Cys Lys Val Ile Ile Glu Asp Gly Asp Ala Tyr Ala
            100                 105                 110

Leu Val Gln His Ala Thr Leu Tyr Cys Gly Lys Cys His Asn Glu Val
        115                 120                 125

Val Leu Ala Pro Met Phe Glu Arg Leu Ser Thr Glu Ser Val Gln Glu
    130                 135                 140

Gln Leu Pro Tyr Ser Val Thr Leu Ile Ser Met Pro Ala Thr Thr Glu
145                 150                 155                 160

Gly Arg Arg Gly Phe Ser Val Ser Val Glu Ser Ala Cys Ser Asn Tyr
                165                 170                 175

Ala Thr Thr Val Gln Val Lys Glu Val Asn Arg Met His Ile Ser Pro
            180                 185                 190

Asn Asn Arg Asn Ala Ile His Pro Gly Asp Arg Ile Leu Glu Ile Asn
        195                 200                 205

Gly Thr Pro Val Arg Thr Leu Arg Val Glu Glu Val Glu Asp Ala Ile
    210                 215                 220

Ser Gln Thr Ser Gln Thr Leu Gln Leu Leu Ile Glu His Asp Pro Val
225                 230                 235                 240

Ser Gln Arg Leu Asp Gln Leu Arg Leu Glu Ala Arg Leu Ala Pro His
                245                 250                 255

Met Gln Asn Ala Gly His Pro His Ala Leu Ser Thr Leu Asp Thr Lys
            260                 265                 270

Glu Asn Leu Glu Gly Thr Leu Arg Arg Arg Ser Leu Arg Arg Ser Asn
        275                 280                 285

Ser Ile Ser Lys Ser Pro Gly Pro Ser Ser Pro Lys Glu Pro Leu Leu
    290                 295                 300

Phe Ser Arg Asp Ile Ser Arg Ser Glu Ser Leu Arg Cys Ser Ser Ser
305                 310                 315                 320

Tyr Ser Gln Gln Ile Phe Arg Pro Cys Asp Leu Ile His Gly Glu Val
                325                 330                 335

Leu Gly Lys Gly Phe Phe Gly Gln Ala Ile Lys Val Thr His Lys Ala
            340                 345                 350

Thr Gly Lys Val Met Val Met Lys Glu Leu Ile Arg Cys Asp Glu Glu
        355                 360                 365

Thr Gln Lys Thr Phe Leu Thr Glu Val Lys Val Met Arg Ser Leu Asp
    370                 375                 380

His Pro Asn Val Leu Lys Phe Ile Gly Val Leu Tyr Lys Asp Lys Lys
385                 390                 395                 400

Leu Asn Leu Leu Thr Glu Tyr Ile Glu Gly Gly Thr Leu Lys Asp Phe
                405                 410                 415

Leu Arg Ser Met Asp Pro Phe Pro Trp Gln Gln Lys Val Arg Phe Ala
            420                 425                 430

Lys Gly Ile Ala Ser Gly Met Ala Tyr Leu His Ser Met Cys Ile Ile
        435                 440                 445

His Arg Asp Leu Asn Ser His Asn Cys Leu Ile Lys Leu Asp Lys Thr
    450                 455                 460

Val Val Val Ala Asp Phe Gly Leu Ser Arg Leu Ile Val Glu Glu Arg
465                 470                 475                 480
```

-continued

```
Lys Arg Ala Pro Met Glu Lys Ala Thr Thr Lys Lys Arg Thr Leu Arg
                485                 490                 495

Lys Asn Asp Arg Lys Lys Arg Tyr Thr Val Val Gly Asn Pro Tyr Trp
            500                 505                 510

Met Ala Pro Glu Met Leu Asn Gly Lys Ser Tyr Asp Glu Thr Val Asp
        515                 520                 525

Ile Phe Ser Phe Gly Ile Val Leu Cys Glu Ile Ile Gly Gln Val Tyr
    530                 535                 540

Ala Asp Pro Asp Cys Leu Pro Arg Thr Leu Asp Phe Gly Leu Asn Val
545                 550                 555                 560

Lys Leu Phe Trp Glu Lys Phe Val Pro Thr Asp Cys Pro Pro Ala Phe
                565                 570                 575

Phe Pro Leu Ala Ala Ile Cys Cys Arg Leu Glu Pro Glu Ser Arg Pro
            580                 585                 590

Ala Phe Ser Lys Leu Glu Asp Ser Phe Glu Ala Leu Ser Leu Tyr Leu
        595                 600                 605

Gly Glu Leu Gly Ile Pro Leu Pro Ala Glu Leu Glu Glu Leu Asp His
    610                 615                 620

Thr Val Ser Met Gln Tyr Gly Leu Thr Arg Asp Ser Pro Pro
625                 630                 635


<210> SEQ ID NO 23
<211> LENGTH: 626
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 23

Met Ala Gly Glu Arg Pro Pro Leu Arg Gly Pro Gly Pro Gly Pro Gly
1               5                   10                  15

Glu Val Pro Gly Glu Gly Pro Pro Gly Pro Gly Gly Thr Gly Gly Gly
            20                  25                  30

Pro Gly Arg Gly Arg Pro Ser Ser Tyr Arg Val Leu Arg Ser Ala Val
        35                  40                  45

Ser Ser Leu Ala Arg Val Asp Asp Phe His Cys Ala Glu Lys Ile Gly
    50                  55                  60

Ala Gly Phe Phe Ser Glu Val Tyr Lys Val Arg His Arg Gln Ser Gly
65                  70                  75                  80

Gln Val Met Val Leu Lys Met Asn Lys Leu Pro Ser Asn Arg Gly Asn
                85                  90                  95

Thr Leu Arg Glu Val Gln Leu Met Asn Arg Leu Arg His Pro Asn Ile
            100                 105                 110

Leu Arg Phe Met Gly Val Cys Val His Gln Gly Gln Leu His Ala Leu
        115                 120                 125

Thr Glu Tyr Met Asn Gly Gly Thr Leu Glu Gln Leu Leu Ser Ser Pro
    130                 135                 140

Glu Pro Leu Ser Trp Pro Val Arg Leu His Leu Ala Leu Asp Ile Ala
145                 150                 155                 160

Arg Gly Leu Arg Tyr Leu His Ser Lys Gly Val Phe His Arg Asp Leu
                165                 170                 175

Thr Ser Lys Asn Cys Leu Val Arg Arg Glu Asp Arg Gly Phe Thr Ala
            180                 185                 190

Val Val Gly Asp Phe Gly Leu Ala Glu Lys Ile Pro Val Tyr Arg Glu
```

-continued

```
        195                 200                 205

Gly Ala Arg Lys Glu Pro Leu Ala Val Val Gly Ser Pro Tyr Trp Met
    210                 215                 220

Ala Pro Glu Val Leu Arg Gly Glu Leu Tyr Asp Glu Lys Ala Asp Val
225                 230                 235                 240

Phe Ala Phe Gly Ile Val Leu Cys Glu Leu Ile Ala Arg Val Pro Ala
                245                 250                 255

Asp Pro Asp Tyr Leu Pro Arg Thr Glu Asp Phe Gly Leu Asp Val Pro
            260                 265                 270

Ala Phe Arg Thr Leu Val Gly Asp Asp Cys Pro Leu Pro Phe Leu Leu
        275                 280                 285

Leu Ala Ile His Cys Cys Asn Leu Glu Pro Ser Thr Arg Ala Pro Phe
    290                 295                 300

Thr Glu Ile Thr Gln His Leu Glu Trp Ile Leu Glu Gln Leu Pro Glu
305                 310                 315                 320

Pro Ala Pro Leu Thr Arg Thr Ala Leu Thr His Asn Gln Gly Ser Val
                325                 330                 335

Ala Arg Gly Gly Pro Ser Ala Thr Leu Pro Arg Pro Asp Pro Arg Leu
            340                 345                 350

Ser Arg Ser Arg Ser Asp Leu Phe Leu Pro Pro Ser Pro Glu Ser Pro
        355                 360                 365

Pro Asn Trp Gly Asp Asn Leu Thr Arg Val Asn Pro Phe Ser Leu Arg
    370                 375                 380

Glu Asp Leu Arg Gly Gly Lys Ile Lys Leu Leu Asp Thr Pro Ser Lys
385                 390                 395                 400

Pro Val Leu Pro Leu Val Pro Pro Ser Pro Phe Pro Ser Thr Gln Leu
                405                 410                 415

Pro Leu Val Thr Thr Pro Glu Thr Leu Val Gln Pro Gly Thr Pro Ala
            420                 425                 430

Arg Arg Cys Arg Ser Leu Pro Ser Ser Pro Glu Leu Pro Arg Arg Met
        435                 440                 445

Glu Thr Ala Leu Pro Gly Pro Gly Pro Pro Ala Val Gly Pro Ser Ala
    450                 455                 460

Glu Glu Lys Met Glu Cys Glu Gly Ser Ser Pro Glu Pro Glu Pro Pro
465                 470                 475                 480

Gly Pro Ala Pro Gln Leu Pro Leu Ala Val Ala Thr Asp Asn Phe Ile
                485                 490                 495

Ser Thr Cys Ser Ser Ala Ser Gln Pro Trp Ser Pro Arg Ser Gly Pro
            500                 505                 510

Val Leu Asn Asn Asn Pro Pro Ala Val Val Val Asn Ser Pro Gln Gly
        515                 520                 525

Trp Ala Gly Glu Pro Trp Asn Arg Ala Gln His Ser Leu Pro Arg Ala
    530                 535                 540

Ala Ala Leu Glu Arg Thr Glu Pro Ser Pro Pro Pro Ser Ala Pro Arg
545                 550                 555                 560

Glu Pro Asp Glu Gly Leu Pro Cys Pro Gly Cys Cys Leu Gly Pro Phe
                565                 570                 575

Ser Phe Gly Phe Leu Ser Met Cys Pro Arg Pro Thr Pro Ala Val Ala
            580                 585                 590

Arg Tyr Arg Asn Leu Asn Cys Glu Ala Gly Ser Leu Leu Cys His Arg
        595                 600                 605
```

-continued

```
Gly His His Ala Lys Pro Pro Thr Pro Ser Leu Gln Leu Pro Gly Ala
    610                 615                 620

Arg Ser
625


<210> SEQ ID NO 24
<211> LENGTH: 626
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 24

Met Ala Gly Glu Arg Pro Pro Leu Arg Gly Pro Gly Pro Gly Pro Gly
1               5                   10                  15

Glu Val Pro Gly Glu Gly Pro Pro Gly Pro Gly Gly Thr Gly Gly Gly
            20                  25                  30

Pro Gly Arg Gly Arg Pro Ser Ser Tyr Arg Val Leu Arg Ser Ala Val
        35                  40                  45

Ser Ser Leu Ala Arg Val Asp Asp Phe His Cys Ala Glu Lys Ile Gly
    50                  55                  60

Ala Gly Phe Phe Ser Glu Val Tyr Lys Val Arg His Arg Gln Ser Gly
65                  70                  75                  80

Gln Val Met Val Leu Lys Met Asn Lys Leu Pro Ser Asn Arg Gly Asn
                85                  90                  95

Thr Leu Arg Glu Val Gln Leu Met Asn Arg Leu Arg His Pro Asn Ile
            100                 105                 110

Leu Arg Phe Met Gly Val Cys Val His Gln Gly Gln Leu His Ala Leu
        115                 120                 125

Thr Glu Tyr Met Asn Gly Gly Thr Leu Glu Gln Leu Leu Ser Ser Pro
    130                 135                 140

Glu Pro Leu Ser Trp Pro Val Arg Leu His Leu Ala Leu Asp Ile Ala
145                 150                 155                 160

Arg Gly Leu Arg Tyr Leu His Ser Lys Gly Val Phe His Arg Asp Leu
                165                 170                 175

Thr Ser Lys Asn Cys Leu Val Arg Arg Glu Asp Arg Gly Phe Thr Ala
            180                 185                 190

Val Val Gly Asp Phe Gly Leu Ala Glu Lys Ile Pro Val Tyr Arg Glu
        195                 200                 205

Gly Ala Arg Lys Glu Pro Leu Ala Val Val Gly Ser Pro Tyr Trp Met
    210                 215                 220

Ala Pro Glu Val Leu Arg Gly Glu Leu Tyr Asp Glu Lys Ala Asp Val
225                 230                 235                 240

Phe Ala Phe Gly Ile Val Leu Cys Glu Leu Ile Ala Arg Val Pro Ala
                245                 250                 255

Asp Pro Asp Tyr Leu Pro Arg Thr Glu Asp Phe Gly Leu Asp Val Pro
            260                 265                 270

Ala Phe Arg Thr Leu Val Gly Asp Asp Cys Pro Leu Pro Phe Leu Leu
        275                 280                 285

Leu Ala Ile His Cys Cys Asn Leu Glu Pro Ser Thr Arg Ala Pro Phe
    290                 295                 300

Thr Glu Ile Thr Gln His Leu Glu Trp Ile Leu Glu Gln Leu Pro Glu
305                 310                 315                 320

Pro Ala Pro Leu Thr Arg Thr Ala Leu Thr His Asn Gln Gly Ser Val
```

-continued

```
               325                 330                 335

Ala Arg Gly Gly Pro Ser Ala Thr Leu Pro Arg Pro Asp Pro Arg Leu
            340                 345                 350

Ser Arg Ser Arg Ser Asp Leu Phe Leu Pro Pro Ser Pro Glu Ser Pro
        355                 360                 365

Pro Asn Trp Gly Asp Asn Leu Thr Arg Val Asn Pro Phe Ser Leu Arg
    370                 375                 380

Glu Asp Leu Arg Gly Gly Lys Ile Lys Leu Leu Asp Thr Pro Ser Lys
385                 390                 395                 400

Pro Val Leu Pro Leu Val Pro Pro Ser Pro Phe Pro Ser Thr Gln Leu
               405                 410                 415

Pro Leu Val Thr Thr Pro Glu Thr Leu Val Gln Pro Gly Thr Pro Ala
            420                 425                 430

Arg Arg Cys Arg Ser Leu Pro Ser Ser Pro Glu Leu Pro Arg Arg Met
        435                 440                 445

Glu Thr Ala Leu Pro Gly Pro Gly Pro Pro Ala Val Gly Pro Ser Ala
    450                 455                 460

Glu Glu Lys Met Glu Cys Glu Gly Ser Ser Pro Glu Pro Glu Pro Pro
465                 470                 475                 480

Gly Pro Ala Pro Gln Leu Pro Leu Ala Val Ala Thr Asp Asn Phe Ile
               485                 490                 495

Ser Thr Cys Ser Ser Ala Ser Gln Pro Trp Ser Pro Arg Ser Gly Pro
            500                 505                 510

Val Leu Asn Asn Asn Pro Pro Ala Val Val Val Asn Ser Pro Gln Gly
        515                 520                 525

Trp Ala Gly Glu Pro Trp Asn Arg Ala Gln His Ser Leu Pro Arg Ala
    530                 535                 540

Ala Ala Leu Glu Arg Thr Glu Pro Ser Pro Pro Pro Ser Ala Pro Arg
545                 550                 555                 560

Glu Pro Asp Glu Gly Leu Pro Cys Pro Gly Cys Cys Leu Gly Pro Phe
               565                 570                 575

Ser Phe Gly Phe Leu Ser Met Cys Pro Arg Pro Thr Pro Ala Val Ala
            580                 585                 590

Arg Tyr Arg Asn Leu Asn Cys Glu Ala Gly Ser Leu Leu Cys His Arg
        595                 600                 605

Gly His His Ala Lys Pro Pro Thr Pro Ser Leu Gln Leu Pro Gly Ala
    610                 615                 620

Arg Ser
625


<210> SEQ ID NO 25
<211> LENGTH: 555
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

Met Asp Arg Ser Lys Arg Asn Ser Ile Ala Gly Phe Pro Pro Arg Val
1               5                   10                  15

Glu Arg Leu Glu Glu Phe Glu Gly Gly Gly Gly Gly Glu Gly Asn Val
            20                  25                  30

Ser Gln Val Gly Arg Val Trp Pro Ser Ser Tyr Arg Ala Leu Ile Ser
        35                  40                  45
```

-continued

```
Ala Phe Ser Arg Leu Thr Arg Leu Asp Asp Phe Thr Cys Glu Lys Ile
    50                  55                  60

Gly Ser Gly Phe Phe Ser Glu Val Phe Lys Val Arg His Arg Ala Ser
65                  70                  75                  80

Gly Gln Val Met Ala Leu Lys Met Asn Thr Leu Ser Ser Asn Arg Ala
                85                  90                  95

Asn Met Leu Lys Glu Val Gln Leu Met Asn Arg Leu Ser His Pro Asn
            100                 105                 110

Ile Leu Arg Tyr Ile Asn Ser Gly Asn Leu Glu Gln Leu Leu Asp Ser
        115                 120                 125

Asn Leu His Leu Pro Trp Thr Val Arg Val Lys Leu Ala Tyr Asp Ile
    130                 135                 140

Ala Val Gly Leu Ser Tyr Leu His Phe Lys Gly Ile Phe His Arg Asp
145                 150                 155                 160

Leu Thr Ser Lys Asn Cys Leu Ile Lys Arg Asp Glu Asn Gly Tyr Ser
                165                 170                 175

Ala Val Val Ala Asp Phe Gly Leu Ala Glu Lys Ile Pro Asp Val Ser
            180                 185                 190

Met Gly Ser Glu Lys Leu Ala Val Val Gly Ser Pro Phe Trp Met Ala
        195                 200                 205

Pro Glu Val Leu Arg Asp Glu Pro Tyr Asn Glu Lys Ala Asp Val Phe
    210                 215                 220

Ser Tyr Gly Ile Ile Leu Cys Glu Ile Ile Ala Arg Ile Gln Ala Asp
225                 230                 235                 240

Pro Asp Tyr Leu Pro Arg Thr Glu Asn Phe Gly Leu Asp Tyr Asp Ala
                245                 250                 255

Phe Gln His Met Val Gly Asp Cys Pro Pro Asp Phe Leu Gln Leu Thr
            260                 265                 270

Phe Asn Cys Cys Asn Met Asp Pro Lys Leu Arg Pro Ser Phe Val Glu
        275                 280                 285

Ile Gly Lys Thr Leu Glu Glu Ile Leu Ser Arg Leu Gln Glu Glu Glu
    290                 295                 300

Gln Glu Arg Asp Arg Lys Leu Gln Pro Thr Ala Arg Gly Leu Leu Glu
305                 310                 315                 320

Lys Ala Pro Gly Val Lys Arg Leu Ser Ser Leu Asp Asp Lys Ile Pro
                325                 330                 335

His Lys Ser Pro Cys Pro Arg Arg Thr Ile Trp Leu Ser Arg Ser Gln
            340                 345                 350

Ser Asp Ile Phe Ser Arg Lys Pro Pro Arg Thr Val Ser Val Leu Asp
        355                 360                 365

Pro Tyr Tyr Arg Pro Arg Asp Gly Ala Ala Arg Thr Pro Lys Val Asn
    370                 375                 380

Pro Phe Ser Ala Arg Gln Asp Leu Met Gly Gly Lys Ile Lys Phe Phe
385                 390                 395                 400

Asp Leu Pro Ser Lys Ser Val Ile Ser Leu Val Phe Asp Leu Asp Ala
                405                 410                 415

Pro Gly Pro Gly Thr Met Pro Leu Ala Asp Trp Gln Glu Pro Leu Ala
            420                 425                 430

Pro Pro Ile Arg Arg Trp Arg Ser Leu Pro Gly Ser Pro Glu Phe Leu
        435                 440                 445

His Gln Glu Ala Cys Pro Phe Val Gly Arg Glu Glu Ser Leu Ser Asp
```

-continued

```
    450                 455                 460

Gly Pro Pro Pro Arg Leu Ser Ser Leu Lys Tyr Arg Val Lys Glu Ile
465                 470                 475                 480

Pro Pro Phe Arg Ala Ser Ala Leu Pro Ala Ala Gln Ala His Glu Ala
                485                 490                 495

Met Asp Cys Ser Ile Leu Gln Glu Glu Asn Gly Phe Gly Ser Arg Pro
            500                 505                 510

Gln Gly Thr Ser Pro Cys Pro Ala Gly Ala Ser Glu Glu Met Glu Val
        515                 520                 525

Glu Glu Arg Pro Ala Gly Ser Thr Pro Ala Thr Phe Ser Thr Ser Gly
    530                 535                 540

Ile Gly Leu Gln Thr Gln Gly Lys Gln Asp Gly
545                 550                 555
```

What is claimed is:

1. A method of identifying a candidate p53 pathway modulating agent, said method comprising the steps of:

(a) providing an assay system comprising a purified LIMK polypeptide or nucleic acid or a functionally active fragment or derivative thereof;

(b) contacting the assay system with a test agent under conditions whereby, but for the presence of the test agent, the system provides a reference activity; and

(c) detecting a test agent-biased activity of the assay system, wherein a difference between the test agent-biased activity and the reference activity identifies the test agent as a candidate p53 pathway modulating agent.

2. The method of claim 1 wherein the assay system comprises cultured cells that express the LIMK polypeptide.

3. The method of claim 2 wherein the cultured cells additionally have defective p53 function.

4. The method of claim 1 wherein the assay system includes a screening assay comprising a LIMK polypeptide, and the candidate test agent is a small molecule modulator.

5. The method of claim 4 wherein the assay is a kinase assay.

6. The method of claim 1 wherein the assay system is selected from the group consisting of an apoptosis assay system, a cell proliferation assay system, an angiogenesis assay system, and a hypoxic induction assay system.

7. The method of claim 1 wherein the assay system includes a binding assay comprising a LIMK polypeptide and the candidate test agent is an antibody.

8. The method of claim 1 wherein the assay system includes an expression assay comprising a LIMK nucleic acid and the candidate test agent is a nucleic acid modulator.

9. The method of claim 8 wherein the nucleic acid modulator is an antisense oligomer.

10. The method of claim 8 wherein the nucleic acid modulator is a PMO.

11. The method of claim 1 additionally comprising:

(d) administering the candidate p53 pathway modulating agent identified in (c) to a model system comprising cells defective in p53 function and, detecting a phenotypic change in the model system that indicates that the p53 function is restored.

12. The method of claim 11 wherein the model system is a mouse model with defective p53 function.

13. A method for modulating a p53 pathway of a cell comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a LIMK polypeptide comprising an amino acid sequence selected from group consisting of SEQ ID NOs:19, 20, 21, 22, 23, 24, and 25, whereby p53 function is restored.

14. The method of claim 13 wherein the candidate modulator is administered to a vertebrate animal predetermined to have a disease or disorder resulting from a defect in p53 function.

15. The method of claim 13 wherein the candidate modulator is selected from the group consisting of an antibody and a small molecule.

16. The method of claim 1, comprising the additional steps of:

(d) providing a secondary assay system comprising cultured cells or a non-human animal expressing LIMK,

(e) contacting the secondary assay system with the test agent of (b) or an agent derived therefrom under conditions whereby, but for the presence of the test agent or agent derived therefrom, the system provides a reference activity; and

(f) detecting an agent-biased activity of the second assay system,

wherein a difference between the agent-biased activity and the reference activity of the second assay system confirms the test agent or agent derived therefrom as a candidate p53 pathway modulating agent,

and wherein the second assay detects an agent-biased change in the p53 pathway.

17. The method of claim 16 wherein the secondary assay system comprises cultured cells.

18. The method of claim 16 wherein the secondary assay system comprises a non-human animal.

19. The method of claim 18 wherein the non-human animal mis-expresses a p53 pathway gene.

20. A method of modulating p53 pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a LIMK polypeptide or nucleic acid.

21. The method of claim 20 wherein the agent is administered to a mammalian animal predetermined to have a pathology associated with the p53 pathway.

22. The method of claim 20 wherein the agent is a small molecule modulator, a nucleic acid modulator, or an antibody.

23. A method for diagnosing a disease in a patient comprising:

(a) obtaining a biological sample from the patient;

(b) contacting the sample with a probe for LIMK expression;

(c) comparing results from step (b) with a control;

(d) determining whether step (c) indicates a likelihood of disease.

24. The method of claim 23 wherein said disease is cancer.

25. The method according to claim 24, wherein said cancer is a cancer as shown in Table 1 as having >25% expression level.

* * * * *