Title: L.RRCAPS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE

Abstract: Human L.RRCAPS 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 regulate the activity of L.RRCAPS are provided.
LRRCAPs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. provisional patent applications 60/338,733 filed 10/22/2001, 60/357,600 filed 2/15/2002, and 60/361,196 filed 3/1/2002. The contents of the prior applications are hereby incorporated in their entirety.

BACKGROUND OF THE INVENTION


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).

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).

Leucine-rich repeats (LRRs) are short motifs of 22-28 residues in length and are found in various cytoplasmic, membrane, and extracellular proteins (Rothberg, J. et al. (1990) Genes Dev (12A): 2169-87). These proteins play diverse roles, with protein-protein

D2S448 is a melanoma associated gene, a tumor antigen, and possibly a peroxidase, which may be involved in p53-dependent apoptosis and immune responses. It also shows promise as a potential immunogenic peptide for cancer vaccination (Horikoshi, N. et al. (1999) Biochem Biophys Res Commun; 261(3): 864-9).

Glioma amplified on chromosome 1 protein (GAC1) is a member of the leucine rich repeat (LRR) superfamily and may play a role in signal transduction or cell adhesion. Gene amplification of this gene is seen in glioma and retinoblastoma tumors (Almeida, A. et al. (1998) Oncogene 16: 2997-3002). GAC1 contains 12 full-length LRR motifs, and its LRR block is flanked by cysteine-rich sequences. GAC1 is expressed in adult brain and at much lower levels in adult heart and kidney (Almeida, A. et al. (1998) supra).

Trophoblast glycoprotein (TPBG) is a protein that is expressed by all types of trophoblasts as early as 9 weeks of development, and was originally identified as a cell surface antigen defined by monoclonal antibody 5T4. TPBG plays a role in cell adhesion and motility and may be involved in metastasis, placentation, and trophoblast invasion. Expression of TPBG in gastric and colon cancers is associated with tumor metastasis and poor prognosis (Boyle, J. et al. (1990) Hum. Genet 84: 455-458). TPBG is expressed in several tumor cell lines (Myers, K. et al. (1994) Biol. Chem. 269: 9319-9324).

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, have 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,
74; Watson KL, et al., 1994 J Cell Sci. 18: 19-33; Miklos GL, and Rubin GM. 1996 Cell
86:521-529; Wassarman DA, et al., 1995 Curr Opin Gen Dev 5: 44-50; and Booth DR.
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.

All references cited herein, including patents, patent applications, publications, and
sequence information in referenced Genbank identifier numbers, are incorporated herein in
their entireties.

**SUMMARY OF THE INVENTION**

We have discovered genes that modify the p53 pathway in *Drosophila*, and identified
their human orthologs, hereinafter referred to as LRRCAPS (Leucine rich repeat,
capricious related). The invention provides methods for utilizing these p53 modifier genes
and polypeptides to identify LRRCAPS-modulating agents that are candidate therapeutic
agents that can be used in the treatment of disorders associated with defective or impaired
p53 function and/or LRRCAPS function. Preferred LRRCAPS-modulating agents
specifically bind to LRRCAPS polypeptides and restore p53 function. Other preferred
LRRCAPS-modulating agents are nucleic acid modulators such as antisense oligomers
and RNAi that repress LRRCAPS gene expression or product activity by, for example,

binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA).

LRRCAPS modulating agents may be evaluated by any convenient *in vitro* or *in vivo*
assay for molecular interaction with an LRRCAPS polypeptide or nucleic acid. In one
embodiment, candidate LRRCAPS modulating agents are tested with an assay system
comprising a LRRCAPS polypeptide or nucleic acid. 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. LRRCAPS-modulating agents include LRRCAPS related proteins (e.g. dominant negative mutants, and biotherapeutics); LRRCAPS-specific antibodies; LRRCAPS-specific antisense oligomers and other nucleic acid modulators; and chemical agents that specifically bind to or interact with LRRCAPS or compete with LRRCAPS binding partner (e.g. by binding to an LRRCAPS binding partner). In one specific embodiment, a small molecule modulator is identified using a binding assay. In specific embodiments, the screening assay system is selected from an apoptosis assay, a cell proliferation assay, an angiogenesis assay, and a hypoxic induction assay.

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).

The invention further provides methods for modulating the LRRCAPS function and/or the p53 pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a LRRCAPS 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

Genetic screens were designed to identify modifiers of the p53 pathway in Drosophila, where a genetic modifier screen was carried out in which p53 was overexpressed in the wing (Ollmann M, et al., Cell 2000 101: 91-101). The CAPS gene was identified as a modifier of the p53 pathway. Accordingly, vertebrate orthologs of these modifiers, and preferably the human orthologs, LRRCAPS (Leucine rich repeat, capricious related) 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.

In vitro and in vivo methods of assessing LRRCAPS function are provided herein. Modulation of the LRRCAPS 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. LRRCAPS-modulating agents that act by inhibiting or enhancing LRRCAPS expression, directly or indirectly, for example, by affecting an LRRCAPS function such as binding activity, can be identified using methods provided herein. LRRCAPS modulating agents are useful in diagnosis, therapy and pharmaceutical development.

Nucleic acids and polypeptides of the invention

Sequences related to LRRCAPS nucleic acids and polypeptides that can be used in the invention are disclosed in Genbank (referenced by Genbank identifier (GI) number) as GI#s 18073097 (SEQ ID NO:1), 16157510 (SEQ ID NO:3), 14758125 (SEQ ID NO:4), 14149931 (SEQ ID NO:5), 20547335 (SEQ ID NO:7), 5453655 (SEQ ID NO:9), 3253212 (SEQ ID NO:10), 21734210 (SEQ ID NO:12), 21706505 (SEQ ID NO:13), 14764197 (SEQ ID NO:14), 5729717 (SEQ ID NO:15), and 435654 (SEQ ID NO:18) for nucleic acid, and GI#s 11877257 (SEQ ID NO:19), 16157511 (SEQ ID NO:20), 14758126 (SEQ ID NO:21), 5453656 (SEQ ID NO:22), 14764198 (SEQ ID NO:23), and 5729718 (SEQ ID NO:24) for polypeptides. Further, nucleic acid sequences of SEQ ID NOs: 2, 6, 8, 11, 16, and 17 can also be used in the methods of the invention.

The term "LRRCAPS polypeptide" refers to a full-length LRRCAPS protein or a functionally active fragment or derivative thereof. A "functionally active" LRRCAPS fragment or derivative exhibits one or more functional activities associated with a full-length, wild-type LRRCAPS protein, such as antigenic or immunogenic activity, ability to bind natural cellular substrates, etc. The functional activity of LRRCAPS 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. In one embodiment, a functionally active LRRCAPS polypeptide is a LRRCAPS derivative capable of rescuing defective endogenous LRRCAPS activity, such as in cell based or animal assays; the rescuing derivative may be from the same or a different species. For purposes herein, functionally active fragments also include those fragments that comprise one or more structural domains of an LRRCAPS, such as a binding domain. Protein domains can be identified using the PFAM program (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2). For example, the approximate amino acid locations of LRR (leucine rich repeat) and IG domains of LRRCAPS from GI#s 11877257, 16157511, 14758126, 5453656,
14764198, and 5729718 (SEQ ID NOs:19, 20, 21, 22, 23, and 24, respectively) are listed in Table 1 further below in Example 1. Methods for obtaining LRRCAPS 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, and 24 (an LRRCAPS). In further preferred embodiments, the fragment comprises the entire functionally active domain.

The term "LRRCAPS nucleic acid" refers to a DNA or RNA molecule that encodes a LRRCAPS polypeptide. Preferably, the LRRCAPS 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 human LRRCAPS. Methods of identifying orthologs are known in the art. 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 MA and Bork P, Proc Natl Acad Sci (1998) 95:5849-5856; Huynen MA et al., Genome Research (2000) 10:1204-1210). Programs for multiple sequence alignment, such as CLUSTAL (Thompson JD 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 ProCerylion, 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) 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.

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.

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; 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."
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:1-18. 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:1-18 under high stringency hybridization conditions that comprise: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X 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 6X SSC, 1X Denhardt's solution, 100 μg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C for 1h in a solution containing 0.1X SSC and 0.1% SDS (sodium dodecyl sulfate).

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, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA; hybridization for 18-20h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5mM 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 2X SSC and 0.1% SDS.

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 x SSC, 50 mM sodium phosphate (pH 7.6), 5X 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 x SSC at about 37° C for 1 hour.

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

LRRCAPS nucleic acids and polypeptides, useful for identifying and testing agents that modulate LRRCAPS function and for other applications related to the involvement of
LRRCAPS in the p53 pathway. LRRCAPS 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 an LRRCAPS protein for assays used to assess LRRCAPS 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 SJ and Hames BD (eds.) Protein Expression: A Practical Approach, Oxford University Press Inc., New York 1999; Stanbury PF 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 JE et al, Current Protocols in Protein Science (eds.), 1999, John Wiley & Sons, New York). In particular embodiments, recombinant LRRCAPS 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.

The nucleotide sequence encoding an LRRCAPS polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native LRRCAPS 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. An isolated host cell strain that modulates the expression of, modifies, and/or specifically processes the gene product may be used.
To detect expression of the LRRCAPS gene product, the expression vector can comprise a promoter operably linked to an LRRCAPS 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 LRRCAPS gene product based on the physical or functional properties of the LRRCAPS protein in in vitro assay systems (e.g. immunoassays).

The LRRCAPS 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).

Once a recombinant cell that expresses the LRRCAPS 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). Alternatively, native LRRCAPS 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.

The methods of this invention may also use cells that have been engineered for altered expression (mis-expression) of LRRCAPS 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

Animal models that have been genetically modified to alter LRRCAPS 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 LRRCAPS in a p53 pathway process such as apoptosis or cell proliferation. Preferably, the altered LRRCAPS expression results in a detectable
phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal LRRCAPS 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 or rats), among others. 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.


In one embodiment, the transgenic animal is a “knock-out” animal having a heterozygous or homozygous alteration in the sequence of an endogenous LRRCAPS gene that results in a decrease of LRRCAPS function, preferably such that LRRCAPS 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 LRRCAPS gene is used to construct a homologous recombination vector suitable for altering an endogenous LRRCAPS 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 MH et al., (1994) Scan J Immunol 40:257-264; Declerck PJ et al., (1995) J Biol Chem. 270:8397-400).

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 LRRCAPS gene, e.g., by introduction of additional copies of LRRCAPS, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the LRRCAPS gene. Such regulatory sequences include inducible, tissue-specific, and constitutive promoters and enhancer elements. The knock-in can be homozygous or heterozygous.

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 P1 (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 (OGorman 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

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 LRRCAPS function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered LRRCAPS expression that receive candidate therapeutic agent.

In addition to the above-described genetically modified animals having altered LRRCAPS function, animal models having defective p53 function (and otherwise normal LRRCAPS 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**

The invention provides methods to identify agents that interact with and/or modulate the function of LRRCAPS and/or the p53 pathway. Modulating agents identified by the methods are also part of the invention. 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 LRRCAPS 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 LRRCAPS activity by administering a LRRCAPS-interacting or -modulating agent.

As used herein, an "LRRCAPS-modulating agent" is any agent that modulates LRRCAPS function, for example, an agent that interacts with LRRCAPS to inhibit or enhance LRRCAPS activity or otherwise affect normal LRRCAPS function. LRRCAPS function can be affected at any level, including transcription, protein expression, protein
localization, and cellular or extra-cellular activity. In a preferred embodiment, the LRRCAPS-modulating agent specifically modulates the function of the LRRCAPS. The phrases “specific modulating agent”, “specifically modulates”, etc., are used herein to refer to modulating agents that directly bind to the LRRCAPS polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the LRRCAPS. These phrases also encompasses modulating agents that alter the interaction of the LRRCAPS with a binding partner, substrate, or cofactor (e.g. by binding to a binding partner of an LRRCAPS, or to a protein/binding partner complex, and altering LRRCAPS function). In a further preferred embodiment, the LRRCAPS-modulating agent is a modulator of the p53 pathway (e.g. it restores and/or upregulates p53 function) and thus is also a p53-modulating agent.

Preferred LRRCAPS-modulating agents include small molecule compounds; LRRCAPS-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.

**Small molecule modulators**

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 LRRCAPS 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 LRRCAPS-modulating activity. Methods for generating and obtaining compounds are well known in the art (Schreiber SL,

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.

**Protein Modulators**

Specific LRRCAPS-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 LRRCAPS-modulating agents. In a preferred embodiment, LRRCAPS-interacting proteins affect normal LRRCAPS function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In another embodiment, LRRCAPS-interacting proteins are useful in detecting and providing information about the function of LRRCAPS proteins, as is relevant to p53 related disorders, such as cancer (e.g., for diagnostic means).


15


In preferred embodiments, an LRRCAPS-interacting protein specifically binds an LRRCAPS protein. In alternative preferred embodiments, an LRRCAPS-modulating agent binds an LRRCAPS substrate, binding partner, or cofactor.

**Antibodies**

In another embodiment, the protein modulator is an LRRCAPS specific antibody agonist or antagonist.

The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify LRRCAPS modulators. The antibodies can also be used in dissecting the portions of the LRRCAPS pathway responsible for various cellular responses and in the general processing and maturation of the LRRCAPS.

Antibodies that specifically bind LRRCAPS polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of LRRCAPS polypeptide, and more preferably, to human LRRCAPS. Antibodies may be polyclonal, monoclonal (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')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 LRRCAPS which are particularly antigenic can be selected, for example, by routine screening of LRRCAPS polypeptides for antigenicity or by applying a theoretical method for selecting antigenic regions of a protein (Hopp and Wood (1981), Proc. Nati. 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:19-24. 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 LRRCAPS or substantially
purified fragments thereof. If LRRCAPS fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of an LRRCAPS protein. In a particular embodiment, LRRCAPS-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.

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


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).

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).

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 (US Pat. No. 5,859,206; WO0073469).
Specific biotherapeutics

In a preferred embodiment, an LRRCAPS-interacting protein may have biotherapeutic applications. Biotherapeutic agents formulated in pharmaceutically acceptable carriers and dosages may be used to activate or inhibit signal transduction pathways. This modulation may be accomplished by binding a ligand, thus inhibiting the activity of the pathway; or by binding a receptor, either to inhibit activation of, or to activate, the receptor. Alternatively, the biotherapeutic may itself be a ligand capable of activating or inhibiting a receptor. Biotherapeutic agents and methods of producing them are described in detail in U.S. Pat. No. 6,146,628.

LRRCAPS, its ligand(s), antibodies to the ligand(s) or the LRRCAPS itself may be used as biotherapeutics to modulate the activity of LRRCAPS in the p53 pathway.

Nucleic Acid Modulators

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

In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to an LRRCAPS mRNA to bind to and prevent translation, preferably by binding to the 5’ untranslated region. LRRCAPS-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.

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 phosphodiimide 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 JC, 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; US Pat. No. 5,235,033; and US Pat No. 5,378,841).


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 JF, et al, Current Concepts in Antisense Drug Design, J Med Chem. (1993) 36:1923-1937; Tonkinson JL et al., Antisense Oligodeoxynucleotides as Clinical Therapeutic Agents, Cancer Invest. (1996) 14:54-65).

Accordingly, in one aspect of the invention, an LRRCAPS-specific nucleic acid modulator is used in an assay to further elucidate the role of the LRRCAPS in the p53 pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, an LRRCAPS-specific antisense oligomer is used as a therapeutic agent for treatment of p53-related disease states.
Assay Systems

The invention provides assay systems and screening methods for identifying specific modulators of LRRCAPS 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 LRRCAPS nucleic acid or protein. In general, secondary assays further assess the activity of a LRRCAPS modulating agent identified by a primary assay and may confirm that the modulating agent affects LRRCAPS in a manner relevant to the p53 pathway. In some cases, LRRCAPS modulators will be directly tested in a secondary assay.

In a preferred embodiment, the screening method comprises contacting a suitable assay system comprising an LRRCAPS polypeptide or nucleic acid with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. binding 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 LRRCAPS activity, and hence the p53 pathway. The LRRCAPS polypeptide or nucleic acid used in the assay may comprise any of the nucleic acids or polypeptides described above.

Primary Assays

The type of modulator tested generally determines the type of primary assay.

Primary assays for small molecule modulators

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 GS 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, immunogenicity 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.

Cell-based screening assays usually require systems for recombinant expression of LRRCAPS 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 LRRCAPS-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the LRRCAPS protein may be assayed by various known methods such as substrate processing (e.g. ability of the candidate LRRCAPS-specific binding agents to function as negative effectors in LRRCAPS-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 LRRCAPS 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.

The screening assay may measure a candidate agent’s ability to specifically bind to or modulate activity of a LRRCAPS polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The LRRCAPS polypeptide can be full length of a fragment thereof that retains functional LRRCAPS activity. The LRRCAPS polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The LRRCAPS polypeptide is preferably human LRRCAPS, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of LRRCAPS interaction with a binding target, such as an endogenous or exogenous protein or other substrate that has LRRCAPS–specific binding activity, and can be used to assess normal LRRCAPS gene function.
Suitable assay formats that may be adapted to screen for LRRCAPS 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 PB, Curr Opin Chem Biol (1998) 2:597-603; Sundberg SA, Curr Opin Biotechnol 2000, 11:47-53). In one preferred embodiment, screening assays use 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 PR, Nat Struct Biol (2000) 7:730-4; Fernandes PB, supra; Hertzberg RP and Pope AJ, Curr Opin Chem Biol (2000) 4:445-451).

A variety of suitable assay systems may be used to identify candidate LRRCAPS and p53 pathway modulators (e.g. U.S. Pat. Nos. 5,550,019 and 6,133,437 (apoptosis assays); U.S. Pat. No. 6,020,135 (p53 modulation), U.S. Pat. Nos. 5,976,782, 6,225,118 and 6,444,434 (angiogenesis assays), among others). Specific preferred assays are described in more detail below.

**Apoptosis assays.** Assays for apoptosis may be performed by terminal deoxynucleotidyl transferase-mediated 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 an LRRCAPS, 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 LRRCAPS function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express LRRCAPS relative to wild type cells. Differences in apoptotic response compared to wild
type cells suggests that the LRRCAPS plays a direct role in the apoptotic response. Apoptosis assays are described further in US Pat. No. 6,133,437.

**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.

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 LS 3800 Liquid Scintillation Counter). Another proliferation assay uses the dye Alamar Blue (available from Biosource International), which fluoresces when reduced in living cells and provides an indirect measurement of cell number (Voytk-Harbin SL *et al.*, 1998, In Vitro Cell Dev Biol Anim 34:239-46).

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 LRRCAPS are seeded in soft agar plates, and colonies are measured and counted after two weeks incubation.


Accordingly, a cell proliferation or cell cycle assay system may comprise a cell that expresses an LRRCAPS, 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 assay system. A cell proliferation assay may also be used to test whether LRRCAPS 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 LRRCAPS relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the LRRCAPS plays a direct role in cell proliferation or cell cycle.

**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 an LRRCAPS, 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 LRRCAPS function plays a direct role in cell proliferation. For example, an angiogenesis assay may be performed on cells that over- or under-express LRRCAPS relative to wild type cells. Differences in angiogenesis compared to wild type cells suggests that the LRRCAPS plays a direct role in angiogenesis. U.S. Pat. Nos. 5,976,782, 6,225,118 and 6,444,434, among others.

**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 glycolytic enzymes and VEGF. Induction of such genes by hypoxic conditions may be assayed by growing cells
transfected with LRRCAPS 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 an LRRCAPS, 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 LRRCAPS 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 LRRCAPS relative to wild type cells. Differences in hypoxic response compared to wild type cells suggests that the LRRCAPS plays a direct role in hypoxic induction.

**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 2.5g/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.

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.

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 JR et al., Bioconjug Chem. 2001 May-Jun;12(3):346-53).

**Cell Migration.** An invasion/migration assay (also called a migration assay) tests the ability of cells to overcome a physical barrier and to migrate towards pro-angiogenic signals. Migration assays are known in the art (e.g., Paik JH et al., 2001, J Biol Chem 276:11830-11837). In a typical experimental set-up, cultured endothelial cells are seeded onto a matrix-coated porous lamina, with pore sizes generally smaller than typical cell size. The matrix generally simulates the environment of the extracellular matrix, as described above. The lamina is typically a membrane, such as the transwell polycarbonate membrane (Corning Costar Corporation, Cambridge, MA), and is generally part of an upper chamber that is in fluid contact with a lower chamber containing pro-angiogenic stimuli. Migration is generally assayed after an overnight incubation with stimuli, but longer or shorter time frames may also be used. Migration is assessed as the number of cells that crossed the lamina, and may be detected by staining cells with hemotoxylin solution (VWR Scientific, South San Francisco, CA), or by any other method for determining cell number. In another exemplary set up, cells are fluorescently labeled and migration is detected using fluorescent readings, for instance using the Falcon HTS FluoroBlok (Becton Dickinson). While some migration is observed in the absence of stimulus, migration is greatly increased in response to pro-angiogenic factors. As described above, a preferred assay system for migration/invasion assays comprises testing an LRRCAPS's response to a variety of pro-angiogenic factors, including tumor angiogenic and inflammatory angiogenic agents, and culturing the cells in serum free medium.

**Sprouting assay.** A sprouting assay is a three-dimensional in vitro angiogenesis assay that uses a cell-number defined spheroid aggregation of endothelial cells
("spheroid"), embedded in a collagen gel-based matrix. The spheroid can serve as a starting point for the sprouting of capillary-like structures by invasion into the extracellular matrix (termed "cell sprouting") and the subsequent formation of complex anastomosing networks (Korff and Augustin, 1999, J Cell Sci 112:3249-58). In an exemplary experimental set-up, spheroids are prepared by pipetting 400 human umbilical vein endothelial cells into individual wells of a nonadhesive 96-well plates to allow overnight spheroidal aggregation (Korff and Augustin: J Cell Biol 143: 1341-52, 1998). Spheroids are harvested and seeded in 900μl of methocel-collagen solution and pipetted into individual wells of a 24 well plate to allow collagen gel polymerization. Test agents are added after 30 min by pipetting 100 μl of 10-fold concentrated working dilution of the test substances on top of the gel. Plates are incubated at 37°C for 24h. Dishes are fixed at the end of the experimental incubation period by addition of paraformaldehyde. Sprouting intensity of endothelial cells can be quantitated by an automated image analysis system to determine the cumulative sprout length per spheroid.

Primary assays for antibody modulators

For antibody modulators, appropriate primary assays test is a binding assay that tests the antibody's affinity to and specificity for the LRRCAPS 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 LRRCAPS-specific antibodies; others include FACS assays, radioimmunoassays, and fluorescent assays.

In some cases, screening assays described for small molecule modulators may also be used to test antibody modulators.

Primary assays for nucleic acid modulators

For nucleic acid modulators, primary assays may test the ability of the nucleic acid modulator to inhibit or enhance LRRCAPS gene expression, preferably mRNA expression. In general, expression analysis comprises comparing LRRCAPS expression in like populations of cells (e.g., two pools of cells that endogenously or recombinantly express LRRCAPS) 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 LRRCAPS mRNA expression is reduced in cells treated with the nucleic acid modulator (e.g., Current Protocols in Molecular Biology (1994) Ausubel FM et al., eds., John Wiley & Sons, Inc., chapter 4; Freeman WM et al., Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm DH 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 LRRCAPS 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).

In some cases, screening assays described for small molecule modulators, particularly in assay systems that involve LRRCAPS mRNA expression, may also be used to test nucleic acid modulators.

**Secondary Assays**

Secondary assays may be used to further assess the activity of LRRCAPS-modulating agent identified by any of the above methods to confirm that the modulating agent affects LRRCAPS in a manner relevant to the p53 pathway. As used herein, LRRCAPS-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 LRRCAPS.

Secondary assays generally compare like populations of cells or animals (e.g., two pools of cells or animals that endogenously or recombinantly express LRRCAPS) in the presence and absence of the candidate modulator. In general, such assays test whether treatment of cells or animals with a candidate LRRCAPS-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.

**Cell-based assays**

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.

Animal Assays

A variety of non-human animal models of normal or defective p53 pathway may be used to test candidate LRRCAPS 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.

In a preferred embodiment, p53 pathway activity is assessed by monitoring neovascularization and angiogenesis. Animal models with defective and normal p53 are used to test the candidate modulator’s affect on LRRCAPS 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 over-express the LRRCAPS. The mixture is then injected subcutaneously (SC) into female athymic nude mice (Taconic, Germantown, NY) 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.

In another preferred embodiment, the effect of the candidate modulator on LRRCAPS is assessed via tumorigenicity assays. In one example, a xenograft comprising human cells from a pre-existing tumor or a tumor cell line is used. Tumor xenograft assays are known in the art (see, e.g., Ogawa K et al., 2000, Oncogene 19:6043-6052). Xenografts are typically 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 LRRCAPS endogenously are injected in the flank, 1 x 10^5 to 1 x 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.

In another preferred embodiment, tumorogenicity is monitored using a hollow fiber
assay, which is described in U.S. Pat No. US 5,698,413. Briefly, the method comprises
implanting into a laboratory animal a biocompatible, semi-permeable encapsulation device
containing target cells, treating the laboratory animal with a candidate modulating agent,
and evaluating the target cells for reaction to the candidate modulator. Implanted cells are
generally human cells from a pre-existing tumor or a tumor cell line. After an appropriate
period of time, generally around six days, the implanted samples are harvested for
evaluation of the candidate modulator. Tumorogenicity and modulator efficacy may be
evaluated by assaying the quantity of viable cells present in the macrocapsule, which can
be determined by tests known in the art, for example, MTT dye conversion assay, neutral
red dye uptake, trypan blue staining, viable cell counts, the number of colonies formed in
soft agar, the capacity of the cells to recover and replicate in vitro, etc.

In another preferred embodiment, a tumorogenicity assay use a transgenic animal,
usually a mouse, carrying a dominant oncogene or tumor suppressor gene knockout under
the control of tissue specific regulatory sequences; these assays are generally referred to as
transgenic tumor assays. In a preferred application, tumor development in the transgenic
model is well characterized or is controlled. In an exemplary model, the "RIP1-Tag2"
transgene, comprising the SV40 large T-antigen oncogene under control of the insulin
gene regulatory regions is expressed in pancreatic beta cells and results in islet cell
switch," occurs at approximately five weeks, as normally quiescent capillaries in a subset
of hyperproliferative islets become angiogenic. The RIP1-TAG2 mice die by age 14
weeks. Candidate modulators may be administered at a variety of stages, including just
prior to the angiogenic switch (e.g., for a model of tumor prevention), during the growth of small tumors (e.g., for a model of intervention), or during the growth of large and/or invasive tumors (e.g., for a model of regression). Tumorogenicity and modulator efficacy can be evaluating life-span extension and/or tumor characteristics, including number of tumors, tumor size, tumor morphology, vessel density, apoptotic index, etc.

**Diagnostic and therapeutic uses**

Specific LRRCAPS-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 or impaired p53 function (e.g. due to overexpression, underexpression, or misexpression of p53, or due to gene mutations), comprising the step of administering an agent to the cell that specifically modulates LRRCAPS activity. Preferably, the modulating agent produces a detectable phenotypic change in the cell indicating that the p53 function is restored. The phrase “function is restored”, and equivalents, as used herein, means that the desired phenotype is achieved, or is brought closer to normal compared to untreated cells. For example, with restored p53 function, cell proliferation and/or progression through cell cycle may normalize, or be brought closer to normal relative to untreated cells. The invention also provides methods for treating disorders or disease associated with impaired p53 function by administering a therapeutically effective amount of an LRRCAPS-modulating agent that modulates the p53 pathway. The invention further provides methods for modulating LRRCAPS function in a cell, preferably a cell pre-determined to have defective or impaired LRRCAPS function, by administering an LRRCAPS-modulating agent. Additionally, the invention provides a method for treating disorders or disease associated with impaired LRRCAPS function by administering a therapeutically effective amount of an LRRCAPS-modulating agent.

The discovery that LRRCAPS 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.

Various expression analysis methods can be used to diagnose whether LRRCAPS expression occurs in a particular sample, including Northern blotting, slot blotting,
ribozyme protection, quantitative RT-PCR, and microarray analysis. (e.g., Current
chapter 4; Freeman WM et al., Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann

Tissues having a disease or disorder implicating defective p53 signaling that express an
LRRCAPS, are identified as amenable to treatment with an LRRCAPS modulating agent.
In a preferred application, the p53 defective tissue overexpresses an LRRCAPS 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 LRRCAPS cDNA sequences as probes, can determine whether particular
tumors express or overexpress LRRCAPS. Alternatively, the TaqMan® is used for
quantitative RT-PCR analysis of LRRCAPS expression in cell lines, normal tissues and
tumor samples (PE Applied Biosystems).

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

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

EXAMPLES

The following experimental section and examples are offered by way of illustration
and not by way of limitation.
I. Drosophila p53 screen

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 compared 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. CAPS was an enhancer of the wing phenotype. Human orthologs of the modifiers are referred to herein as LRRCAPS.


For example, PFAM was employed to determine approximate amino acid locations for the LRR and IG domains of GI#s 11877257, 16157511, 14758126, 5453656, 14764198, and 5729718 (SEQ ID NOs:19, 20, 21, 22, 23, and 24, respectively), as shown in Table 1.
Table 1. Approximate amino acid locations for various domains of LRRCAPS polypeptides

<table>
<thead>
<tr>
<th>LRRCAPS GI#</th>
<th>LRRCAPS SEQ ID NO</th>
<th>LRR domain (PFAM 01462, PFAM00560, PFAM 01463)</th>
<th>IG domain (PFAM00047)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11877257</td>
<td>19</td>
<td>28 to 59, 339 to 372, 63 to 86, 87 to 110, 111 to 134, 135 to 158, 159 to 181, 182 to 203, 376 to 399, 400 to 423, 424 to 447, 448 to 471, 472 to 494, 495 to 519, 216 to 264, 529 to 579</td>
<td>260 to 319, 356 to 414, 447 to 504, 539 to 596</td>
</tr>
<tr>
<td>16157511</td>
<td>20</td>
<td>63 to 86, 87 to 110, 111 to 134, 135 to 158, 159 to 182, 119 to 244</td>
<td></td>
</tr>
<tr>
<td>14758126</td>
<td>21</td>
<td>22 to 45, 46 to 69, 70 to 93, 94 to 117, 118 to 140, 141 to 162, 321 to 341, 345 to 368, 369 to 392, 393 to 416, 417 to 440, 175 to 225, 474 to 524</td>
<td></td>
</tr>
<tr>
<td>5453656</td>
<td>22</td>
<td>70 to 93, 94 to 117, 118 to 141, 142 to 165, 166 to 189, 190 to 213, 214 to 237, 238 to 261, 262 to 285, 286 to 310, 311 to 335, 336 to 359, 360 to 383, 369-421</td>
<td>438 to 499</td>
</tr>
<tr>
<td>14764198</td>
<td>23</td>
<td>3 to 26, 27 to 50, 51 to 74, 76 to 99, 100 to 123, 124 to 144, 165-210</td>
<td>227 to 285</td>
</tr>
<tr>
<td>5729718</td>
<td>24</td>
<td>61 to 90, 92 to 115, 119 to 142, 143 to 166, 211 to 234, 235 to 258, 259 to 282, 294 to 345</td>
<td></td>
</tr>
</tbody>
</table>

II. High-Throughput In Vitro Fluorescence Polarization Assay

Fluorescently-labeled LRRCAPS 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 LRRCAPS activity.

III. High-Throughput In Vitro Binding Assay

³²P-labeled LRRCAPS peptide is added in an assay buffer (100 mM KCl, 20 mM HEPES pH 7.6, 1 mM MgCl₂, 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.

IV. **Immunoprecipitations and Immunoblotting**

For coprecipitation of transfected proteins, 3 x 10^6 appropriate recombinant cells containing the LRRCAPS 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 x 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.

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).

V. **Expression analysis**

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.

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

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

Primers for expression analysis using TaqMan assay (Applied Biosystems, Foster City, CA) were prepared according to the TaqMan protocols, and the following criteria: a) each primer pair produced only one product.

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 300nM primer and 250 nM probe, and approximately 25ng 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).

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. In cases where normal tissue was not available, a universal pool of cDNA samples was used instead. In these cases, a gene was considered overexpressed in a tumor sample when the difference of expression levels between a tumor sample and the average of all normal samples from the same tissue type was greater than 2 times the standard deviation of all normal samples (i.e., Tumor - average(All normal samples) > 2 x STDEV(all normal samples)).

Results are shown in Table 2. Number of pairs of tumor samples and matched normal tissue from the same patient are shown for each tumor type. Percentage of the samples with at least two-fold overexpression for each tumor type is provided. “ND” means not done. 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>
<thead>
<tr>
<th>SEQ NO</th>
<th>Breast</th>
<th># of Pairs</th>
<th>Colo</th>
<th># of Head and Neck</th>
<th># of Pairs</th>
<th>Kidney</th>
<th># of Pairs</th>
<th>Lung</th>
<th># of Pairs</th>
<th>Ovary</th>
<th># of Pairs</th>
<th>Uterus</th>
<th># of Pairs</th>
<th>Prostate</th>
<th># of Pairs</th>
<th>Skin</th>
<th># of Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.3%</td>
<td>19</td>
<td>12.1%</td>
<td>33</td>
<td>12.5%</td>
<td>8</td>
<td>26.1%</td>
<td>23</td>
<td>10.0%</td>
<td>20</td>
<td>18.2%</td>
<td>11</td>
<td>10.5%</td>
<td>19</td>
<td>33.3%</td>
<td>12</td>
<td>66.7%</td>
</tr>
<tr>
<td>3</td>
<td>21.2%</td>
<td>33</td>
<td>30.3%</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
<td>70.0%</td>
<td>20</td>
<td>25.0%</td>
<td>32</td>
<td>30.0%</td>
<td>6</td>
<td>15.8%</td>
<td>19</td>
<td>0.0%</td>
<td>12</td>
<td>ND</td>
</tr>
<tr>
<td>4</td>
<td>36.8%</td>
<td>19</td>
<td>9.1%</td>
<td>33</td>
<td>37.5%</td>
<td>8</td>
<td>45.8%</td>
<td>24</td>
<td>45.0%</td>
<td>20</td>
<td>27.3%</td>
<td>11</td>
<td>26.3%</td>
<td>19</td>
<td>33.3%</td>
<td>12</td>
<td>0.0%</td>
</tr>
<tr>
<td>10</td>
<td>42.1%</td>
<td>19</td>
<td>9.1%</td>
<td>33</td>
<td>12.5%</td>
<td>8</td>
<td>21.7%</td>
<td>23</td>
<td>28.6%</td>
<td>21</td>
<td>66.7%</td>
<td>12</td>
<td>35.3%</td>
<td>17</td>
<td>0.0%</td>
<td>12</td>
<td>0.0%</td>
</tr>
<tr>
<td>14</td>
<td>31.6%</td>
<td>19</td>
<td>21.2%</td>
<td>33</td>
<td>50.0%</td>
<td>8</td>
<td>36.4%</td>
<td>22</td>
<td>30.0%</td>
<td>20</td>
<td>27.3%</td>
<td>11</td>
<td>10.5%</td>
<td>19</td>
<td>16.7%</td>
<td>12</td>
<td>33.3%</td>
</tr>
<tr>
<td>16</td>
<td>33.3%</td>
<td>18</td>
<td>32.3%</td>
<td>31</td>
<td>100.0%</td>
<td>8</td>
<td>26.1%</td>
<td>23</td>
<td>57.9%</td>
<td>19</td>
<td>36.4%</td>
<td>11</td>
<td>15.8%</td>
<td>19</td>
<td>8.3%</td>
<td>12</td>
<td>0.0%</td>
</tr>
<tr>
<td>15</td>
<td>50.0%</td>
<td>12</td>
<td>33.3%</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
<td>14</td>
<td>71.4%</td>
<td>14</td>
<td>42.9%</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>
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 LRRCAPS 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 LRRCAPS 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 LRRCAPS polypeptide, and the candidate test agent is a small molecule modulator.

5. The method of Claim 4 wherein the assay is a binding 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 LRRCAPS 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 LRRCAPS 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 LRRCAPS polypeptide comprising an amino acid sequence selected from group consisting of SEQ ID NOs:19, 20, 21, 22, 23, and 24, 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:
   e. providing a secondary assay system comprising cultured cells or a non-human animal expressing LRRCAPS,
f. 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

g. detecting an agent-biased activity of the second assay system,
h. 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,
i. 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.


20. A method of modulating p53 pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a LRRCAPS 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 LRRCAPS 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 2 as having >25% expression level.
SEQUENCE LISTING

<110>  Exelixis, Inc.

<120>  LLRCApS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

<130>  EX02-119C-PC

<150>  60/338,733
<151>  2001-10-22

<150>  60/357,600
<151>  2002-02-15

<150>  60/361,196
<151>  2002-03-01

<160>  24

<170>  PatentIn version 3.1

<210>  1
<211>  8580
<212>  DNA
<213>  Homo sapiens

<400>  1
ggtgtcagc tcggatggct gactgcttct tcactttctg ttaaccttcca cccagtaagct 60
ttagtctccc cttggacagc tgtgcctttc cctttcttc gacccgagcc 120
cgagcgtcc gcccaggttgg aagcgcagag ggtggcgcagag tttccagacg aagacggacc 180
ggagggctg tgtaagctctg cggacgcttg tgcagagaaag tgtgcaggagc getcttttgt 240
gggctcttatc atccccgtcct tgggtgcaca gccccgatgat tggccctcttg ggtggattga 300
taatgaggg agaggggggg cggagatccct ctttctccct ctttctccct ctttctccctc 360
tccctctctt cccctctctc ccctctctctc ccctctctcc ctctctctctt ctctctctcct 420
tccctctctc ctctctctctc tcttctctcc ctctctctctt ctctctctt ccctcagagaa 480
gctttgcccgg ttgctcagct ctctggcttgg gatctgctctg cctttgggggc cggccggccgg 540
tccctccggag tgctcctgcct ccctggcctgc ggcagctccga ggcctgtggc gctcttttgtc 600
tccctccgggt gcacagtgcct gcgggctgcct cccggaagag atcttgccggg cggtccctgc 660
cggtaagcactcgc ggcagcggctg ggcgcgcgc gcgtgaagcgc gcgtgcctctgc 720
tagcggctgccttt ccctccgggc caccctgtgtc tgacgggccgc acgcaggtgg gggctggcctg 780
cgccatccgg gcagccagctc tcgtgcgggc ccggccctctt ccctcaacct ccctccggccgag 840
tccctccacgg cgggcggctgc gcggccggcc gcggccgtgcg gcgcagcggggg 900
tggcacaggg gcgcaggggag gcggcgggcgg gcggcagc gagttgaaggc tgctgctgcgc 960
tggtgctgcag ccagtcccgag gcggagggcg gcggagtgcg ccgtgcgggct gtaggtgcctgg 1020
...
ttttgcgcct tttttttttt tcctaaagcg attgcgattt ctgtgggag ctcaagacgg 4680
gcgagctgcc cgagatctct tcgagatacc ccaggggaggg aggagatggg caggatttag 4740
taggacaact cgggttactaa tgaaccttggc gcctgctgct agccccccggg aatcaggttg 4800
caacatgttg tgttccccgg ggcccgtggtgt gggttggcctc ggcaaggggg aatagggagc 4860
ggagaaaga aacgcctttg gaaaaatgcaaa tgatcttctt tcgctgcttg tcacccccc 4920
tcttttccccc tccgcttcccctc ccggttttct tgaatattag ggcgctttcgc ggcttggagtt 4980
tttcctttcct ccctgccacc ccctcttccc cctgccccccca ccttaaggttt gcctctgtagt 5040
acgcgcgattgg aacacggaagg tgccctaaaga tgctgagccgg cgttggcttc ctctgctggt 5100
taaacgcaggg gacgctttta cagacagaga gtgcgaacaa agccaagacac ttggaagaag 5160
tccgctgtct cggcgaagaa aagaaaaacg tactgaaatat caactctgag aacaaagagt 5220
attacacagt tagcgctgtct cagccccccc agatagcaat ctactcagtttt tttctcaagt 5280
gaaacaccttt gacaagactg tatttggcaaa ttactcacaac gcggctgactc 5340
ttacacctgg taacaacgag tcagccaggg gcatttacgt gcctgtaaaa 5400
cctcaaaag cactctctcta aacaacaaca agcttgaatg atgtgaggg gacaccttcc 5460
taggcctgga gacgcttgag ctagcttccc aatacactga gctcaagcgg 5520
cctgacgat ttctctcagg cccacttacta tcagttacgt cgcactcgg 5580
ctgatcctgg aacgatgcct ctagcagctc aactgttcct cctgaatgac aacctcttg 5640
atagggcata agtaatgtcct ttgtgtggcgc tccctctgcaaa tattggaagg actcatggga 5700
ttcgcttggc gcggaaaata tgggaattac ctctctcctct cagcgcctggc 5760
tagacacact aactgtttttt ttgagcggaga ttgggtcttga gactcccttt aggttgtagt 5820
ggaaagactg gacccagctg accagggcaag acctctgctcc cagaaaaagtg gccagtgatt 5880
ccagctcaggg gggagcagcat gtgcacaccct acgttcacaa agctgtcactc acaatgaaac 5940
tcgcctctcc aaaaaaagg gcctcagggccc cagaaaaagt gaagatgtgc 6000
caccccctgg aagttcactgtg tcaagggaca gggaaaggttt tggacccatc atgtgtgatc 6060
agaccaagtc tctctgtgccct ctcacctgtc gcagcagctg tgtgctgcacc tctccagagct 6120
cagacacagt ttctctgtgta aacctgcaag aaggaaggtt cactaatctc tctctcctgcg 6180
agccccacccc gaccagtcaca aagaaactctc acctaaacagg gaactatatc cacaactgtct 6240
atatgaatga ccctctttaga ccagttctttt tggaacttaact gcacccgaga aacacagagsa 6300
ttggctgtct cagcgaggtt gcccttacaa acctgaacag tcctcgcaaa ctttatcttg 6360
atggcaatata ccgttaaggtg cctgtacctt ctacgttttga tggactgcag agcttgcaat 6420
atctctatgt aagatataat gtcattaaggg aaattaagcct tctgacccatt gatgcttttg 6480

ttaacctcaca gctacgtgttt ctaaacaaca accttcttcgt gcctttacct gataatatat 6540

ttgaggagggac ggcocctaaa dgggttaacc tggagaacaaac cctattttttc cacctgccccg 6600
tgaaaggggt tcttgacatt gcctccgcttt tcatcagatg aagatgtgcag gagaacccct 6660
ggacgtgtac ctgtgacatt atggggtgta aagactggtgac agaactgtgaa ccaatacccttg 6720
tcatcattaa tttggtgacct tggccagtctc ttgctaagcgc tgcaggggggag aataactaat 6780

tttgggggag gggaggttcct gcctccagac gcccccacactcc gctcagatgag accgtctttg 6840

caatgatttc caatacagac aacacetccgtg egcttagttg gcctccctag tctcatcctgg 6900

aactacacct gtaagtttcca ctgtctctct taattcttggg aattgcttggt gtttttctct 6960

tatctgcttg ttttggggtct gcgtttattccg tctttgtcttt gaacaacgccga aaggggtgcgc 7020
cagcgtctcc caggaatacc aacaacttag aacgtactgcct ctttcaatata cagatgtgggt 7080
cctacaacac ttagactcact gataaaacag acggccatgtg ctacaactat atccccccccac 7140
cgttgagccta gattgtgcacc aaccccatctt atcgtgcaaga ggaagggagc caagtagcct 7200

attaccgaaa ccctgcaagaag ttccatata cgaacctgga gggaaaaaa aaaagacgcag 7260

ccacacctgc ttacacataa agtgccactgg agctgtatag aagcagggcc acaacagag 7320

agcctgagct gctgtatcaaa aataattgctg aggcagctaag cgaacctccc aacgacggcc 7380

tagtcaacta taacctttttt accttcaactaa aacggcctgt tggccctctcc ctagaatcctc 7440

gacgcccaaa ccaagacaga atcaatata cagttttttata tggaaactccc agaataagct 7500

tttgtggggca gtcaaaaccc aacacccctt tactgcaagc taagccgcaaa gcaagccggg 7560

actacccgga agttctggga aaacaaactg caactgctca gctgtgaaag ggaaatcattt 7620

acacccctca ggacatcagag gtacgtgctc cgaactgttgga aacaaacgga ctagattttt 7680

ttgttttttt ttgttttctcc ttctccacgt gaataagggg gagtttttaa atgatgggga 7740

gataggatga agtcagttt gttgccttcttc aagttttttat ttaaattatt ttcttcttgc 7800

tctcctcccc ttcttttttt tttttttttt tctttttttc cttctctctt gagaaccatt 7860

cagtggcata gaatggtttct acaatgcatct cttctcataag ttttgtttttt gggttttttt 7920

cttctttttt cttttttttt cgatgtggga gttggagaag gagattatag tgaactggaag 7980

agagatagcc aacatctttca aatgaaaagt gatatttagt gtaatggtta gagaatctcc 8040

aaagatcttt ttgtgactaca accttttttggg taataaatga tatattgttg tttccatgctc 8100

agttacccag ttagccactc gggatctcct aacctgttgtt aaaggattgctt ccgacttttaa 8160

gtacattatt taatatgtgct ttttagcttt gataaattga aaataatatta atggtgtggta 8220
tttttgaaat tgaaaaacact gttaaataga ttgatgtgct agetatatga agtcacgta
8280
cagtttgtcct gagttataga aacccgctcg tctacaatag attctagttc taggacctttg
8340
taggaattacc tataaaatat ttctttttct ctgggttttaa gttgatattat ttaagtcac
8400
taaggggatt taacagttgga ctgaaggttaa taagcctact cagtcaggat taataattca
8460
ttaataaaat atataataacc caataatcaga tgaattgtag caattaatgc ctcttcgtaa
8520
atcattatgct tacactaaca tggtgaggtgt tttgattataa tttccataatt aaaaagaacgt
8580

<210> 2
<211> 2667
<212> DNA
<213> Homo sapiens

<400> 2
	ttccccctgc caccggctcc tttccccctgc ccccaacctaa gttttgccttg taggtaacctg
  60
	agttgacacc gaaggtgcct aaagatgcctg agcggcgttt gtttctccag tggtttaacc
  120

tggtgccgga tcttcacagac agagagcgcga aaaaacctgga aagacatttg caagatccgc
  180
tgtcttgctcg aagaaagga aacagtacctg aatacaacct gttgagacaa agagatttca
  240
	acagttagcc tgctccagcc cccgcagcatg cgaactcttc aaccccccttc caatggaacac
  300

tctttgacaa gactgtatcc aaaggaattt gtcactaact ccaacgcgggt gactctcacc
  360

tagttaacac acgggttaca ggagagccga acggggccat tcagttgctc gaaaacacctc
  420
	aaaagagcctc acttcaaaca caacaagttc gatatttga gggagcacac ctttctaggc
  480

tgaggagagcc tggagtattc ccagggccagc tacaatttca tcaagtgcctcagt gaggctggg
  540
gcatccagca aacccaacaa gctcagaagtc ctcatctctga atgcaaaacct tctgttttca
  600

tgcacccagca atgtggttccg tttgtctctg cggacccact tagacccctcag ggggaatagg
  660

tcaaaagtaaa tggctttttgc ttggcgttctt gaacatatgg gagggatcat ggagatcag
  720

tggtaggagaa atccatggaa ttgcaacttt gactcattcct ctctcaaggg ctgggtctacg
  780

tcataactg ttttttttgc ggagagcttc gtgtgagactc ccttttgagg gcattgggaa
  840

gacgtgaccc agctgaccag gcgaagactctc tggccagagg aaaaagtgccag tattttcagt
  900
cagaggggcaga cccgctgtgc caccaggctca ctaagggctgt caccctaatc gatactttgtc
  960

tcctaaacc cccaggttccc gaaagccagc cggccggcaca aatgagaaaa tctgctcaact
 1020
cccggagtga ctgtggcaaa ggacaggc aaagtttgga gcaaatgtgt ggacactgacc
 1080
aagctctctg ttgctctctac ctgctccagc agctgtgtct gcactcetca gaggtcagac
 1140
aatggtctgtg atgtaaaactg ccaagaaagg aagttaacta atatctctgta cctgacccgccc
 1200
aaacgacca gtccaaagaa actttacacta acagggact atcctcaaac tgtctataag 1260
aatgacctct tagaatcag ttctttggac ttacgcaact taggaacaa caggattgca 1320
gtcatcagg aaggtgcctt tacaaacactg accagtttac gcagaccttt tctgaatggc 1380
aattaccttg aagtgctgtga cccctctatg tttgatggac tgcagagctt gcaaatatctc 1440
tatttagagt ataatgtcat taaggaatatt aagcctcttg aaccttgatgc tttgattac 1500
c tacagctac tgttctggaa caaacaacctt cttcggtctct tacctgaaat tatatttggg 1560
gggacgcccc taacacgcttc gaatcgcgaa aacaaccatt ttttcacaact gcgccgtgaa 1620
ggggttctgg gttcagtttcc ggggttctatc cagatagatc tgcagggagaa cccctgggac 1680
tgtcactggt acatcagtttg gctgaaagac tggacagaaac atgccaattc cccctgcactc 1740
attaatgagg tgtacggtgca aagctctctgt aacgactgcag gggacatact aaatattctg 1800
ggaggagggt gttctgtgcc agaragccca aacctgctgag atggaaaccc tgttgtcaatg 1860
aatacacaata cagacacacc tgggtgcttcg attgttgcttc cttgcctcata tccgtgacta 1920
caacagtgaag ttccactgttc tgttcttttt ctgggatgctc tgttgttgtgtt cactttatct 1980
gtcgttgggtt gggcgtgtttt gttcttttgc ttcttgtgcac gggcgaaggg agtgcgcgagc 2040
gttcccaagga atacacacaa cttacagtct aaccttgtttc atatacagta tgggtttatc 2100
aacaactgaga ctacaggtaa aacagacggc catgctctaca acttatatcc cocacctgtg 2160
ggcagagtgt gcacaaaaacc catctcatact gagaagggag gagacccagt agccccattac 2220
cgaaacctgc aagagtttcag ctatatgccac ctggaggaga aaaaagaaga gcgacaccaca 2280
cctgctttaca caataagttgc cactgagctg ctgaaaaacg aggccacacc aagagacgct 2340
gagctgctgt agtcaaatat tgctgacgca gtcaggaac ttccccagcgcc aggccttagtc 2400
cactataact ttgtgtacctt acctaaaaag gatttggccct cttcttatga atctgcaagc 2460
caaaaccaag acagaatcaca taaaccctgg ttatatggaat tttccccagga atgtttttgt 2520
gggcagcag cgacaacac ccctttactg caagctcagc cgcacatcag acgcgactac 2580
c tgcaggtcgc tggaaaaaca aactgcacatc agtcagctgt gaagggaaat catttacaad 2640
cctaaggcat cagaggatgc tgctgccg 2667

<210> 3
<211> 6801
<212> DNA
<213> Homo sapiens

<400> 3
agccgccccgt ggtggctccg tgctgtgccg cgcctcagtc gcgcgctgcg catggacaaa 60
cgctgcagg gcgcgcgcgc gcgcgctgcg tgtgcgcttcg tgcgcgttctc gcgcgctgggg 120

7/56
acccacacg tggagccgat tgcagcgttt gacagacgta taaaactcaac cggagacacat 1980
tctgatgcca gccgcctctg ttctcaaat gatgctgtgg cctgttctgc gatcgcgag 2040
gatccatcag caggtgtagc ggtgggttcag ggacgcttcg ggttcagcgcg 2100
atctcaccacg atgtacagcag tggctctgtg gtgcacctca acgccacagg ttcaccacttc 2160
aacgcacttg gtcctcaccata cgtcactgaaag ctctctctctct cctgactcggctgtgcg 2220
cacggggcgc gcagaactcg ttcggcagct gtcctcaccac gagaattacc ggagcagcgc 2280
gagaagtagc taaaccctcag gcaccactcag tgtggcgctct cgctgaccgcg ctctgagcgc 2340
ctgctgagaat cccgtgtagca gaaatgtgctt cacaacccctcg ggggacatcaac cccactccag 2400
cgtgaccactgc gccaagcccttg ctccatgcctcg gcgcctgtgttg ccaccacactcg gatgggaag 2460
gagaccaagc caccgaagcg cgaattccact caacatcctgcag tgcagttggcg gcaagtctgt 2520
gacagacacc tgcactctac cgggttggtgcc ttgagccagcg cagctctctg cggagacagc 2580
caactcagca cacgcgtcag cacagacccc cccgtgctctc cccgtgctcag cccctctctgtgcg 2640
gacagocggcg cccagcagcgc gggggtctgc gttgctcag ttcggtctgctc cccttcgctgcgc 2700
ggcagccgcac tgccttcgtgctctagcgac gcgggaggca gatccacagcag 2760
tctcactctct acatacacgcgc atccaccagtg tacgggagca cggagcatga gcgcggcagc 2820
atccggccggc tggcgacagca cggggtcagc gctgctccctcg ggtgcgacagc gggggtccggc 2880
aagccgctgc tccctcttggc caccggcgcgc cccagggagc gcatgctggga cggagcgagc 2940
agggcacatc gctgctctgc gcggggagcc aacggcgcttg ccaccgagct gcctgcaagtcg 3000
agcatgcgca cgcgctgtgttc gcggggcacc aacggagatcg ccaccgagct gcctgcaagtcg 3060
aaccgcaacagtgccggcg caccctctac ttagagacca ggaagatcg ggtgctggag 3120
atccgacac tccactctca gcacgccgctg cgcacacgctc tgggggaggt gggcaggtgag 3180
acgctggag cgatcaccagc cgacccgcggt cgcagcattgc tggccacagc cctgttcttc cacaagcttc 3240
gccacgggg ctcctcaggtt tggcacacag ccctgctctca cactgctgtta ccgggctggag 3300
gagacactcgc agccacattgc aacagacatcc tcccctctctc aacaaagcttc tctctctctc 3360
ttcgggttgg gtaatcaggg cggcgctcag gggctttcttc gggggcttgtt cggggttgccg 3420
ggggatgac cggtgctccc gcacgtgtctg acacggagcg tcacggagcc gcgtgtctctc 3480
atggcaacac gcggctgtgtgc gcaggtcctgc ccacccgagc gcacgctccc gcagagctgcg 3540
ggatccagcg cctacacggtc tgcaggttggc tgcctactgc acacgtgcgtgc ccaccacagc 3600
gagacccctgc aaaaacgctg tggcagccgg cgcagacagt gaagctgtgtttgac 3660
ggctgcaacac tcdaacatcgc ccggttctgc gggcagctgag cggaggtctgg 3720
agccgctgg gcocacccc gatggctct tctagcaaac agttcaacgc cctgccagat 3780

ggggacagtgt tgggttatga gaacccctgga gtttctcccc cggcccaagt gactcagatc 3840

agccagacgt cgctggccag gatccatatgc gacaacgcgg acaacatcac cggggtgcag 3900

agccagcgtgt tcaggtggcc gcaggctccc caaagctcag gcagctgtga cagattcccc 3960

agggtagacc tcggggtgtgc gcagagactgc tgtagaagact gttagagacag ggggcaggtc 4020

aatggcttttt cctatcattt ccagggcaga cggttgtcttttg agtctcagcta ccaggagga 4080

aagccgacca agaaacaag accacggaaa atacccagtg ttgggagaca gggggagcatc 4140

ttcagcaaca gcacctcgac ctcagcaca cgctcaagatg catctggtggc aatgtgacttc 4200

agagatgcttg ttcctggaat gcagaagacc atcacagacct tcagaaacatc gataagaaa 4260

cgtgactacag ggtctagtag cagaggtgac ggtggtgcgg cgggcagatc tcagcggccac 4320

aagccagact gggaaaaaga tcagctgaccc attttgtgaat gcacacgaggg cagaggtcacc 4380

tgcggctgtgg aagcggttgcc ccctgccccct cgggtgctgtcc cccgtgaaccg cccaggggcc 4440

tgtggtcagc tgtgttcctca gaagagggcc gagggaaggct ccctagttgcc ttggggaggtc 4500

ttcagagttgg tgtctgctgtg ccaactgtag atcgggtggc cgtgtgcaggg cagctgcca 4560

ttcagccacc gcaccaagcc gagaacactgg acattttcatg aacaagccca gctgggtgctg 4620

tgagaacgcc caggtggcag ggtctccacc agagcacttc cggagagggc aggcagccaca 4680

ggtgccattg gggagcaggg cagggctcct gcagtcaggtc tagaatcttc acggttttatg 4740

ttaatttttt taaaaatgaa aattggtgctc acatatttatt gcaacagttg aatctattag 4800

aggctcaatagtgatatggcc tcccaacacc atttctttttt aatataagca ggataacctct 4860

atatgtcagc cttgtccttgt tcagatgcca ggagccggca gacgctgcaag ccgaggggtg 4920

gttgagtcttt ggagcgccca gaggggtgcc cgcggatcag gtttctcctca cagctgactg 4980

ttaaaagaa aattggtgtt tggcaaaaggg aacagaaacct ttgaatgagag ctttcatcag 5040

gcacagtgtc ggggtgtccag tcgaagcccc cggccttttc cctgggaacc tcagacctcc 5100

tttctctcttt gggctctctgt taacacttca ccacagctca gcatctatac ccaagacaaa 5160

cattccgctg gcctgaagagc gcggatgtcgg cggagtacgtc cgggtgtgctc gcctcttctc 5220

cagcctgatta gcacactcatg aagccacatt tagaaacagg tttggtcttca gctggcttaac 5280

gccacatacg tcgctagagg tgaaccaaaaa tggaaaaaac cttctctcatc ccatattgtgta 5340

tctgactctct gggcagggccc aaggggtgtgtg gttgcaaacag ccgctcccgag cggccctggg 5400

tgcgttcaac gcctctgaaac agagccgcttt cgggtggtgc ccctcccaaggg gaggagggac 5460

tcaaggttgtg ggaactggtct aacactgtgtt tgggtgagttg cgttaaaaaa aaaaaaaaaa 5520
aaagatctct tattaagttt ggcagaaaaa cccttccaac tattgcaggt ctttttagct
atccagtgaata atcaaccttc atacctctag gctctctcttt aattggcaaa aacccctcat
cttctctca ataacaagcgc atctgtgttct tgaagcacaact tcccccgcaac acacagcttc
agtgcacgacc gctgacgcttg gtatccaccact tgtgecnaggg agcaacgcagg
ccacaggtgc ggccgcaccc ccgcccagcag agggccagctt gaggctgtgctt agggggccc
aggaagttg gcgccggttcg gtccccgggttg gcaggtgtcct gcacactgggg tattggctctg
agggctccat gggtggtgggg aggcagcctt cctcttgaca ggtatgggcc atacagtgcc
ttggtgtgatt tgggtcacaac ccggttgttca gttgcatcact tctcccaagga gacaccagg
ccctccccgg acggggcggc caagttgcttg gcggggaggg agccatcttc ctgtaggggaa
gtcatttggg tccaactgttct tatcgcgttaa aatgggttgtt cttacttcttc taaaaccttta
ctggttaagtt attgataaaa attgggtgttg gtttcatcgac atgtggactt cttttagaat
agcaagtcctta atgtagtagc ccaatttttg gaggaccttc tggcaaatag gaaatgtgta
agttcgtctta aagagcttagt gttgtagtagt tggctcaggg acctcagcatc cagcaagttc
tggtggtgcgg gagggtctcttg tgcctctt ccagcctgcc ggcctcatgg tctcaggtct
aggaggggcac acctgtgacca cactggtgta atgtgccccggc tggggcagtt acaaggaatt
gtgttgtctta agcacaaggg gggagctgac tgggttttcc caaaaaatt tctgtataatc
aaaaaacat acctgcgggaa ttacccgtttt gaattctctt gtaatccagaa cagttgggacg
tgctggtgytg aacggtgtctta gcagccacttg gccggtgactc ctgtaacagg agtgaagttt
catatattat ttacatctctt aactgtccaa cacagcaaat gccccttttaa tgacattttg
attcagttcata ttcacaagaag aatggttgcgt ctctgtattg atgccttttca gtctaaaatct
attggtctcata ttccaatgtgt ctt cccttttg

<210> 4
<211> 2593
<212> DNA
<213> Homo sapiens

<400> 4
actcacaacaac tccagagtctc tcatccagag gctcttggta ttctcttggg aatgggtgaggg
aaaagaggtg ccacagtc tataaatgtg gaggctcaaggt tatcagagatg tattcagaa
ataacttgcc accatcagga ctttcacaac ttgctttattt aataaaccgg tttgacaggttggt
tcacacaaaaa tggcatttacc atcggtatctt aataaaccctt gggatattaca
atattgcaaa tagttgagata ggtgccatatta atggccttggg cctcctgaaa caacctcata 300
tcaatcacaa ttcttttagaa atctcttaaag aggatacttt ccatggaactg gaaraacctgg 360
aattctgtca agacctaac aaatttatca cagtgttcca acaaaagtcccc tttagcaagc 420
tcaacagact caaagtgtta acttttaaatg acaatgtctat tgagagtcttt cctccaaaca 480
tcttccgatt tgttctctttta acccatcttag atctctcttgga aatacaatta caaacaattgc 540
cattagttgg ttttctcgaa cacatggtgcc gaatattgga ttcttcagttg gaggcaaca 600
aatggggcttg caatgttgac tttttcatggt taaaaccttg gttggagaac atgcctccac 660
agtctataat tgggtatggt ttgctgcaaca gcccctcatt tttttaaggga agtatcactca 720
gtagactaag gaagagaatt aattgcctca ctctcaccagt gtatgagaa gaagagagtct 780
cctcaggact attacatctg gcaccacacat ctctcaataa tgaattctgc atgtcacaata 840
agacaccgct cattcttaaa ctaccacaca aagacccagg ttgataacct tatattacaa 900
agccatcacc tcaactcttca ggcacctttact ggcctattac tttgtaactgc aagacctctat 960
ccccatcagg acctctataa cattggtcagg aggcacaatc tgaagattct ctgagatctga 1020
gacctctccc gcaaaatcct agaaaagctca ttctagcggg aaatattatt ccaaatgttaa 1080
tgaagttctga tctagtgggaatt tatttacttt tgggaatgtct tcaacttggga aacatctgta 1140
ttggaagttct tgaagagaagat gctgtttatga actctaaagag attacaaaaa ctctatctaa 1200
agtggtaacc cctgacacaaa tttaagtaag gcagttctct tgggttcctactaatctcgaa 1260
acttatatact tgaataaaat gcccattaag aaataacgacc aggaaaccttt aatcacaatgc 1320
ctaaacctta aagtctctgtat ttaaaaaaca accctccccca agttttacca ccacatatit 1380
tttccggggt tctcttaact aaggttaatact tttaaaaaaa ccagttttacc cactctctgg 1440
tagtaatata ttctggttagat ccttgatttgc taacccagat tgcaccttgag gataacccct 1500
gggacgtgctc ctggtacccag gttggacactgc agcaaaatag ccaaaataatt aagacagaa 1560
cagtgacaga tgcacactctc tgcacctcccc cgggcatctc cgcacaaaaa ggataagaaa 1620
ccctaaatag tgaatctacc tgcctagggt tagtaattaata ccctactccatg ccacacaaga 1680
ctgattacct tatggtcacc actcttcgcaaa caaacaacaac taacggctgtactatattc 1740
 gaatctcttac ggacggctgtg ccaccttgtcg ttctaatatt gggactctctg attatgttca 1800
tcaactattgt tttctgctctg cgaagggtag tgggtttcttt ctcttcacccgg agggagaagat 1860
aacaaataaga caaagtagat gagcaaatga gagaacaacag tctgtgcatc ctctcagttaca 1920
gcatgtatgg ccataaaaccc actcatacaac actacataagg acocctctggc tcaacctctatg 1980
aacacgcaat ggtgagcacc gttgctctatg tctataagag tcacatctttt gttccaaagc 2040
agaagctcc gcaaaaatcct agaaagctca ttctagcggg aaatatattt cacagttaa 1080
tgaaagctga ttctagtgga tattttcaatt tggaaatgtct tcaacgttggga aacaatcgta 1140
ttgaaagttct tgaagaagga tctgattatga acctaaagag atttacaatata aataacaatgaa 1200
atggtaacca cctgacccaa ttaaagtaag gcatgttctct tgggtctccat aacatgttgaat 1260
acttataatct tgaataacaa gcctattagag aataaatgcg aggaaccttt aatcaaatgca 1320
tcaacatcct caaatctgtat ttataaataa ctcctctccaa gttttaaccac ccatatatct 1380
ttggagggtt ccctcttaacta aggttaaatct taaaacaacc agttttacca tctattctgtta 1440
agtaatatatttt ggtaatccttt gattttactaa cccagattgga ccctgaggtat aacccttggg 1500
ctgctctcgtt gagctgttggt gacctgcagca atggataccaa aagttaagcag aagacacagt 1560
gcacagtgcac atctctcgtca ctccctccgg gcctctcgcag aaaaaggaat tgaagcgctt 1620
aaatagtgaaga aatctctgtgc caggttttact aataaatccaa tccatgcccc caacagactag 1680
ttacctatgt gtcaccactc ctgcaacacc acaaaatacg gctgataacta tttttgcgatc 1740
tcttaacggac gctgtgccccat tgcgatttgttcc aatatgggga cttctgatttc gtttcactc 1800
tattttttttcc tttttgctgag ggatagtgggt ttttgcttttt cacccgagga gaagataccaa 1860
aaagaaaaaaga ttagatgacg aattgagaga caacagtctttttgatcacttcc aagtacagcat 1920
gtatggtcctttaa aaaaaacccta atccacactac tgaagagccc tctggcttac cttatatgaa 1980
gcatacttgtt acgccccctg ttagaaacatc tttttttttc ttaaatccagca aacagcccc 2040
ggaagagagga gaagagagga attgagagaga aggagagagta gcaaaaacct tcaaaagggag 2100
tttttggaa cagagataatc aatccacactc cacagggcctta aaatatgaat acaaaacccac 2160
gaaccaatca acagaatatttt tatctctcata gtagttccag tcattttcaca gaacaattttt 2220
agaaaaagaa aggaacacctc agcnnctcgag aatcacagaga taccaagaga aaaaactttgc 2280
tcagctccag cctgtataggg aggcacatttc tctctgagcc caggaagacgc tgaagttaat 2340
ggaacacatta atgtaactccg ctccagaggaag ggtatatagtg gaacagacaca aaaaatgata 2400
tttttgactttaa aagctatttat ctaactcgtga acctgacttat ttagaagtcct tggagccgca 2460
aacatagattttaggttttga gggctttctgc aagaaatgtcttg agattttttttccttaaagcttct 2520
acctgagtaaa taagtgattgtt acggctgatgt gcattcaactt aagaacatag ccagccagt 2580
aactttggaa aaaaaaaaaa aaaaaa

<210>  6
<211>  2574
<212>  DNA
<213> Homo sapiens
<d00> 6
tcatcactatg acaacagtaa gctggtgatta cactcttttt ttcatcctct cttgctctgt 60
atatctttac actcccacac tccagtgctc tcaatccagag gttccttgta ttccttttgc 120
aatgtgtgagg aaaaagatgg cacaattgcta ataaattgtg aagccaaagg tatcaagatg 180
gattcgaaa tagtgtgctc accattacga cctttccaaac taagctattt aataaacggc 240
ttgagagatgc ttccacaaaaag tggcttctac aatgcaatcc aatgccctttc 300
gatatttaaca aatagcgaga tattgagata gttgcaatctta atggccttgct cctcctgaaa 360
cacaacctata tcaatcacaag ttccttttaga aatctttctaat aggtggacctg 420
gaaaacctct gatccctgca agcagatcct aatctttatac cagctgattgca acaagctgsc 480
tttagcaagtc tcaacagctct acaatggtta attttaataag acaatgctat tggagctctt 540
cctccaaacat ttctcttgatt ttgttctctta atctccatctg ataactcctag aatcgctggt 600
caaacatgtgc cttatgtttgg ttctctgtaa cacatctggcc gaataatgtgg tctttcagtg 660
gagacaaaagct caggtgagtc caattggtacct tataagcggat taaaactcttg gttgagaac 720
atgctctacaac agtctataat gttggtatggt gttctggcaca gcctccctat tttttaagga 780
agtatacata gtagactaaa gaggaatact atttggctctca cttcccacaggt gtatgagac 840
catgagagac ttcctagagtc atatcatttg gcaagcaact cttctataaaa gatagctgac 900
atgtcaacta agaccacgctct cattctaaac ataccacccag aagcaccagg ttgataacct 960	
tatatccaa agcattcccc tcaacctcca ggcacccact gccctatctcc ttgtaacgct 1020
aaagtctctat ccccatcagg acttcttaata cattgtcagg agcgcaacat tgaagcgtta 1080
tcagatcctga gacctccctc gcggactctct agaaaatcctt ttctaggggg aatatattt 1140
cacagtttaaa tgaagctgctt tctagctgga tatctttcctt tgggaattgct tcacattggga 1200
aacaagtctga tttgagatctt tgaagagagga tctgctatgct aacctaacgag attaacaataa 1260
ttcctatctaa atgtgataaca cttgacaaaaa ttaataattg gcatgttctcttt ggtctctctat 1320
aatcttggaat acttatatact tagattacaa gcaattaggg gataaacctt gaaactcttt 1380
aatccttcagc atctctgtat taatctaaaca actcttcgctaa atgtccttcct 1440
cacataatttt ttctcagggt 6tcctctaat cagtaaatc taaaacacca ccggtttac 1500
catctacgtc taagataaat tttgtgatgt cttgttatatc taacccaggat tgcaccttgagg 1560
gataacccct gggagctgcct cttgtgaacctt gttggacgctg agcaatggat acaaaaattta 1620
agcaagaaca cagtgacaga tgcacctccct gcaactccct ccggccatct cgcacaaacag 1680
gaatggaag ccctaaatcag tgaacttcttc tgcctcaggttt tagtaataaa cccatcctagt 1740
ccaaacacaga ctagttacct tatgttcacc actcttgcaa caacaaca caa taccggtcgtat 1800
actatattc acatctttac ggaacctgttg ccaactgccttg ttctaatattggagcttctg 1860
attatgatca tcacactatgtg ttttcatgtct gcagggatagtg ttttaaatagtt tcatcagccc 1920
aggagaagat aacaaaagaa aacagtagat gagcaatgga gaagaacaag tctggtgcat 1980
cggactga cccgggtatgg ccataaaacc actctacaca ctactgaaag aacccctgtcc 2040
tcactctatg aacaggcaact gggagagccc atggtctcatg tctatagaaac tcacacctttt 2100
gttcacaagag acatgggagag ggaagagagag ggaatgaga aagagggagag agatgcaaaaa 2160
cattccccca gaagttttttt ggaagcacgaa aacatccatg cactcaccag gtcgaatattg 2220
aatataaaca caaaccaaca atcaacagaa tttttatctt tcacagatgc cagctctattg 2280
tacagaaaca tttttatgaa aagaaagggaa ctccagtaac atgggaatcag aagaatattca 2340
aggaaaaaa cctttatctgt gcaacctgtat atgggagcaca attaatccttg gcagccacgaa 2400
gagctgaagtt tatttgggaa ataatgatctc ctaacgctcaa ggaagtttatt atggaacag 2460
acaaaatagc agtatattttga acttaagatc aatattcatg ctagaactctga ctattaagaa 2520
gctctggagc agcaaaaactc gatggagagt tggaggtcttt tcgcgaatag gcgg 2574

<210> 7
<211> 1504
<212> DNA
<213> Homo sapiens

<400> 7
atggccttat gttggtttttc tcgaacaacat tggccgaata ttggtacttc agttggagga 60
caaacaattgg gctgcgaatt gtagctttatt gcagtttaaaa acttggttgg agaacatgc 120
tccacagtct ataattggtg atggtttcctg caacacgccc ccaacttttta aaggaagtat 180
actcagtga agaaagaaggg aacatcttttg ccctacctca ccagttgtag aagaaacatga 240
ggtttcttca ggatacttac atctggcagc aacatcttca aataagatga tggcagatgct 300
aaactaagcc acgtccttcttaaactcaca caccaagaag caagggtttga accttttatat 360
taacaaagca tcccactcaac tttcggacc tttcagctcct atctccttga actgcaaacg 420
cctatatccc acaattatct gcagagtcct aattcattgga aatattggca gccatcatcag 480
tctgagacct ctcctcgaaaa actcttagaaa gtcgatctct gcggagaaata ttatcacag 540
tttattgaat cccatcttttg gttcacaagca tctggaagag gaagaagaga ggaatgagaa 600
agaaggaagtt gatgcgaacag acatctcaaaag ttttagtggt tccacagag aatggtttatc 660
actcagggg tccatattga aataaacaac cacaagcaac tcaacagata ttttacacct 720
ccaaagatggc agtccattttt caacaaaaacttttaga gaaaaagggagag ggagatatc 780
gggaatcaca gaatacctaa ggaaaaacat tgctcagctc cagcctgata tggagggcaca 840
ttatctgga gccccacgaag agctgaaagt aatggaaaca ttaatgtact caacgtcacaag 900
gaaggtatta gttggacaga caaaaatga gtatatgtgaa ctttaagcta atttacatgc 960
tgaactctgc tatttagaag ttctcxagca gcaacacatag atgggagatttg tggaggtcttt 1020
cgacagaaatg ctggagttct gtatatagtc ctaacagtctg aataaaggtcg cttaacgctag 1080
tgtgctcaca atcagaaccct aagcagagca gttaacagtgg ggaaaaaaa aagaagaga 1140
aaaagaaact cagggatcac tg ggagagaagc cat ggccatta tcttcagggca atttagttg 1200
tcccaataaa aataaatctct tgcgtgtaaa tcattcaggg atttagttaa attttcatat 1260
actgaaaggt gtctctatagg agtcctctgtt ccacactctaa aaggtgtaac atattaagtat 1320
ccggaatttt ctttaaggggc ttttacctata gattaataac aattggatttt ccacatttaa 1380
aaaccatact tgtatagtga gttataaatg tgaaggaata cattgttttat aaccagatag 1440
tacctaaaa atgtgtattg tcaacacatc tcaacctttct tgcaataaat gcaaaagaaa 1500
cctgg 1504

<210> 8
<211> 3131
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<3131)
<223> "n" is A, C, G, or T

<400> 8
agctgcgaca acaagaataa ctgaaagagc aggaagggaa acagacagtgc acgttggttc 60
tacaacctaa gaagacagag gtcgcataaa tcaggtcctct taacatatct gcataacctt 120
tgctattcagc agcggcaaa attgtaataa tagcaaaaaa agtcagcata aatggtgaagt 180
atgccccgaa ttttaattga ggaaaaagg cacaatggt tccaggatgct ctagatgaca 240
cctggcgctga aatatatatta atgggaactg cagtttctta ctttttagcca gaggtttttaa 300
cattatgaag ctatgaggct tcgaaaaaag tggatattga gaagaaagta cgcaggggtt 360
ggtggtcttc tttttatatg aaaggatttg aattactttg aacacttcttc ccagcttgac 420
attacagata aagcagagga agttccctcttg tttacaggtta atcgggatttc atcacaatgc 480
aacatgaagc tgtggattca tcttctttat taactctctccc tngcctgtat atcttttac 540
tcccaacttc ccaggtccttc atcaagaggg ttcttggtatt ctctttgcaaa tttgtgaggaa 600
aaagatggca caatgctaat aaatgtgaa gcacaaaggtg tcaagatggt atctgaata 660
agtgtgccac catcagcaccc tttcacaacta acgttactaa ataaccgctt gacagtgcct 720
cacccaaatg acaatctatcga attcattccta gctacatcggc aatattacta 780
attgcagtata ttgagataagc tctatattact cctgctggggc tctgtaaaca acttcatatac 840
aatcacaatt ctctgtagaata atcttacaca tctagactgaa aacacccgta 900
 ttctctggcaag cagataaccac tttatccatt gccatgtaaac caagtctcctt tagacagctc 960
aacagactca aagttgtaaat ttttattgac aatgctattg agagtctcttcctc tccaaactc 1020
 ttccccattg ttctcttttaac caatctagat ccctctggaa atcaattaca aacattgcct 1080
tatgtggttt ttctccgcaaca catggccggc atatggcatc ttccagttgga ggacacaaa 1140
tgggctgctcg caaggtctattg gacgtcattg cagacacagt tgtccacag 1200
tcttacattc gtagttggtt ctgcacacagc ccctcatatttt ttaaaggagact tatactcagt 1260
agactaaga agaatctatc ttccttcctact ccaccagctgt atgaagaaaga tgaggactcct 1320
tcaggatcat tacatctgggc agcacaatcc tcaataaatag atagctgcat gctacaatag 1380
 accacgtcaca ctctttaaac ctccacacaa gcaccagatt ttgatcactta tattaaacag 1440
 ccaatcactc aaacctccaggg acccatctgg cctatttcctt gtaactgcga aagctctactc 1500
 ccaatcaggac ttctataatac ttgtcaggag cgcacactttg aagctttatct agatcttgaga 1560
 cccctccggg aaattcctagt aagctcattc ctagccggaa atattatact cagtttaaatg 1620
 aagtctgatct tagttgaata ttctcttttc gaaatctcttc acttgggaaa caaattgatt 1680
gaatctttcttgaagagattc gtttatggaacctc ctaacgagata ctaaaaaacct tcaattaaat 1740
 ggaagaacctc agacacaaa ttacagttctt gtttccataaa tcttgaatac 1800
 ttatatctttg ataataatgc cattaagggaa atactgcagc gacaccattaa tccaatgcttct 1860
 aaacctttaag ttctcattttt aaaaaaacaaac ccctctcaaag ttttacaccc acatataattttt 1920
 tccaggggttc ctctactttaa ggattaatctt aaaaaacaaac ctctcctacaag tttaacacct 1980
 agtatatttt ttggtgatcggc tgagttgcca acccagatttc agcctggagag ctaccctgga 2040
 gactgctctt gtaggacctggt tggacgacagc caattgataac aaagaattaag caagaacaca 2100
gtggacagatg acatcctctctg cacccctccc ccggcctctccg aacaaggaaga agtggagacc 2160
ctaaatagtg aaatttccttw ctccagtttta gtaataatcacc ctttctgcgcc aacacgacact 2220
agttaccccttggtgctccacac ccctctcccaaca caaacaata cgggtgataac tatttttacag 2280
tctcttaaggc acgtgtgtgcc actgtgcttta ctaaatattg gcattcttcatat tatttctatc 2340
actatttttt tttttggtgctgc agggatagttg gtttcttggct ttccacggcag gagaagacatac 2400
aaaagaaac aagtagatga gcacacgac acctggtcat ccagtacagc 2460
atgtatgccc ataaaaccc ctctcaact actgaaagac cctctgcctc actctatgaa 2520
cagcactacg tgtatcggtc tatagagct cacacctttgc tcccacagcct 2580
cagagagagc aagagaccg cagtagcggac ccaccaggtg atgcaatagc 2640
atctttgac aaccagaaaa ttctatgttc ccacagttcc cagactgcga atacaccttt 2700
agaacacagg ctactacaga ttcctcttct gcacagacac agcgttcttgta tcccagctg 2760
acagacactt ccacacagtt ttctcctctc aagagctgca gcgtcttcttct gagaacatcc 2820
atctcaagct cagcgttgtct ggggcagtac atctcttgc gcccacagag ggtcaatgta 2880
atgacagatg ggaagggagc ccccagcgttat cagccgactc cagccacgagc aaaaattgag 2940
attacatgcttt taagacgcac agttgagcag ctattttag tggcaagagc aaaaaattgag 3000
ttcaggtgtc ttaacagcct atctaagctt caggcttgct agaagcagcag ctggtgcagcag 3060
ctacaagagc atataaggtgt aggttgtctb gcacagagac tgggtatatgc ttcagcgttc 3120
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3180
ccagaggtgtc tgtccgtctg cagccggtat ttcagcctc cagctttgcgc cagcagcttgt 3240
ccagaggtgtg cgccggtgct ctccgtctcc gccgtctgtg cagctttgcgc cagcagctgtg 3300
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3360
ccagaggtgtc tgtccgtctg cagccggtat ttcagcctc cagctttgcgc cagcagctgtg 3420
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3480
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3540
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3600
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3660
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3720
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3780
ctagactag ggtccagcc ggtggccagt gggggccacc cgggtcagc 3840

<210> 9
<211> 3227
<212> DNA
<213> Homo sapiens

<400> 9
cagcagacgg cagacgggtgc cggcccgcct cgggaagctg aagatacagc ggtataaaaa 60
ccctccatgtc tcaagctgac ccacatgtga tctggacttg cccagcttgatt agctctgttc 120
ccaccaggct cgttcaagaag ccacagcttt gagggggccac cgggtcagc 180
ctggtcagc cggcgtcctg cggctgtggtg tgggttaatg cggcctctcg 240
gggacaagat gcacggggtt gtagaatgac tgcctccgttg ggggttctcc ttggggcctg 300
gcaggggtgac tgtcccttc agtccaaaac acacagcaac ggactccaga cttctctttgt 360
gaattgctcc ctgcccagcg ggtgcattct gcctgcccag cttcttgctc 420
tttggtgcttg tgtccagctt cccaggtcct cttggtatatc ctcggcccagc 480
tcagactagc ggtccagcc ggtggccagt gcggctcttgta ctcagctttct gcgtttcagc 540
caagtggcag tgcatacgcc tattctctgc ggcagtcctcg cggcagcctcc cggcagcctc 600
acacgctctg cttccgtaga gcacagcctg ttgctcggtg gaccagagtg agctggtggac 660
cctgcctaat cctccaagcgg tggacctgtgc ccagacagcg tttcggatgc cggagagctg 720
tagtttcagc gcctttccag agctggtgag cctgcacacta gaggagaacc agctggccccg 780
gctggagggc cagagtttgtt cggggctgc ggcagcttcag ggtcttctgc tccaacaaca 840
ccagctctac ccgatcggccc ccagggccttt ttcttgccctc agcaaccttc tcgggctgca 900
ctctaccc acccctctgta ggcccatttgga cagccgcttggt ttggaaatgcg tgcctcaacct 960
ggagatactc atgatttggcg gcaacaaggt agatgacatcc ctggacatgta aacctccggtc 1020
cctggcctaac ctgctgtagcc tgggtctacag aggacctgaac ctgctgggaga ttctcgcata 1080
tgccctgaggg ccggctgacaa ccctgcctcctc ttatgacaacc tcgctgggccc 1140
ggtgccccagg ccgcccagcttg aacagctgtgg ggcgctcaag tgccttagacct tccaatagga 1200
ccccctccag ccggttagggc cgggggacctt tgcctaaacag cttgaccaatcatg gaggctgta 1260
actgaacacac atggagggcgcc ggactctctatgg gacaaagtttc gcccctgtgaa accttccccga 1320
gctgaccaag gctggacatcc caataaaccc aagcgcgttcgc ttcactccac ccggtcctttc 1380
ccacccctcg ccggagctggtg aacgctctctcg ccctcaacaccc aacgctctca gctggctgca 1440
ccacgcagcgc tggagtgcgcc cggggcaacctg gcagccagttag ggctctccaggc tgcctctccag 1500
cgctggtgcac tggctgcatcc gcggggcctaa tggcgcagcgc aggcgtgtcc gcctctcatgcg 1560
gcggcaaatcc acccttgtgtgc ccaggagctccc gcctctccag ccggtcctcgg tcggctgaggt 1620
gcccttcgg ccagagtacagca accaatctttt gcccctctatt tccctccagac gcctcccccc 1680
aagcctccag tcagccagttt gcagagccatg ggtgtctgcatgc ccggggccac ccggtcaacc 1740
cgaaccggag atctactggg tcactccagcc tgcggtctctgc gctgacactca cccatgcag 1800
caggaggtgc cgggtgtaccc ccggagggac ccctggagctc cggagggtaaa ggtcgtcgatcc 1860
ggcaggggtct catcaacgtgtg ttggcggcagaa cctgggtgggg gctgtagcacta aggcgtagag 1920
tgtggttgttg gcggctgtgctc ccctctcagccc aggcagggag ccagagggag ggtcgtgagct 1980
cggggtgcag gcgacccacc ccttactcagc ttcgcttatct ttgggtccaccc caccccaacac 2040
agtctccacc accctctcactg ggtggctgtgc cttctccctct ccggggccaggg gcgcccaagc 2100
tctgggcccccgctggcttcggg gaacccacag ctacaacatt acgggcttcc ttcagggcagc 2160
ggagtaactg gcctgtctgc agtggcettt gcgtgtctgctg cccacccagtt ttggctttgtgg 2220
atggggccagg accaaagagg ccactctttag ccacagagcc ttgggggtac gcctgtggcct 2280
cattggccata cctggtctctcg ctgtctctctct cttggcagct gggtgagccgc ccacaccttg 2340
cacaggcgaca ccaggaagtt gtgtgggcttg gaggcgggctt ccctctccag ccggggctttt 2400
ctggggctgg agtgccctttc ctggccccggtg tggtgtctgtc cccccctgcgcc tgcctggga 2460
tccaggaggg ccacagcagc gaccccccgc cggggccagca ccgggggaca ctggtcctcacc catggcttca 2520
aatcctcgaag aagttcagcct gcctccgacca gtagagaaat cactaggact actttttactc 2580
aaaaagagaagt ggtctctggg cagatgccc ctgccagggaa ggtcactggacc cccagctgcc 2640
gaggcctggc agctgggcca agacagatgg ggttttgtgg cccctgggggt cctttctgcag 2700
ccttgaaaaa ggttgcctta cctctagggy tcaccttctgc tcgaccttcttg aggaacatct 2760
ccaaaggaacgg ggagggactt tgggtagagc ctcctgtcct tcctatctttct ctctgcccag 2820
aggtctctggg gctggtgccttg gctgttcccc acctgtgttccc cccggtgcga cccttttcctc 2880
ttcctttttct ctcgtacagtc tcagttgctt gctctttgtggt ctcctggtgcc aggggtcgaag 2940
gagaggccactc cactctcaatgt cggggggctg ccctccatgt tggagtgtgcc ccagccagat 3000
cggagacagc ttggtgagag ggtggccccc gagcgcctcga tttcagcagtc ctgggggtcgg 3060
catcctgaag ctggccttctc atagggcaat tttgacccct ttggagggat gtggccacctc 3120
cccaaccccg attcactctt tttccctgttt ttgttagaaaa taaaaataaa taataaactat 3180
aataggggag aaaggaacagc aagagactaa aaaaaaaaaa aaaaaa 3227

<210> 10
<211> DNA
<212> Homo sapiens
<400> 10
cagcagacgg cagacaggtgc cgggcggccca cgggagagtg aagataccgc ggtatacaaa 60
cctccatgtc tcaagctgac cacactttga tctggacttg ccacgtggatt agctctgtttc 120
ccaccagggct cgggcaagag cccacagctt gagggggcag agggcgtcc ccctgtagggc 180
cctgggcattg cttgagccgac ccagcccccag agggcagaggg agaggacagga tattggagag 240
gggcagaggt gcagcccggt ttagaagacg tgcctttttcgg gggtggtgcttc ttgggccttg 300
gcagggtgct gcagctttac cctgcaaaaa acaaaaggca ggactccaga ctctcttttg 360
gaatgtccgc ctgccccgca gccctcaccct gaggtcttttc cttggcccacc cttctgtgac 420
ttggtgtgct gggtgcacgtg ccgctggcccc ggtgtgccac ccgctgttac ctcgccccccc 480
tcagttgggc tcggcagatcc gcggcctgtta taccgcggcc tgcgctctacc gcaggcgtac 540
cactgtggac tgcgaatgcc tattcctgtac ggacgctcc ccggcaetcc ccgcaggcac 600
acagacccctg ctctctgcaga gcacagcag tgcctcgtggt gaccagaggt agcttgggta 660
ccttgcccaat ccacagcagc tggacctgtgc ccagacacgc tttcgggtga ccggagactg 720
tgattttctac gcccctcccc agctggctgg cctgaccccta gggggcgaac ccgtgacccc 780
gctggaggac cacagcttgg cagggccttc cagcctacag gaactcttac tcaaccacaa 840
ccagtcttac cgctcggccc ccagggcctcc tttggtbccct agcaacttgc tggggtgcga 900
cctcaacctc aacctctgtga gggccatttg cagcgcgtggt tttgaaattc tgccaacctt 960
ggagatacctc atgtatttggcg gcacaacaggt agatgcccac tcggacatga acttcgggcc
1020
cctgcccac ac tgtatgagcc tgggtctagc aggcagcaac ctgcccggaga tctccgacat
1080
tgccccctg ggtgtgcaaa gcctggagaag cctccctcttc tatgacaacc agctggcgcc
1140
ggtgccccaa gggcaccctgg aacaggtgacc cgggtctcaag ttccttagacc tcaacaagaa
1200
cctggtcgtgtc agggtggagcttg tgagcactttga cgcgtccatag ctcagccttt agagctgggg
1260
actgaacaac attgagaggc actgttcctcgc gcaacagtttt gctctgtgtgag acactccccga
1320
gctgacccaag cttgacatca ccaataaccc acggctgtcc ttcacaccccc cccgcccctt
1380
ccaccaacctc ccccagattgg agaccctcctg gctcaacaac aacgctctca gtgggctgtga
1440
ccagcaagagt ggtgcaacgc tggcacaacct gacaggagtt gaagtccctcag gcaaccccat
1500
cgcctgtgcc tgtgtccttcc ggtggcggcc cggccggccc aaccggtgctcc gttcctagcga
1560
gcgcctcaac cccctgtgtgg cggacccctcg ccgctcctccgg tcggctgagtt
1620
gcctctctgg gagatgacgg accactgtttt gcccctcactc tcccaacagaa gctcctccccc
1680
aagctctcag gtacccagttt gagacagcag cttgctgtcct tgcogggcagc tggccgaacc
1740
cgaaccgcag atctactctgg ttcacccagtc gggacttcagc actgacacctg cccatgcaagg
1800
caggaggtgc cgggtgtacc cccagggcgc cctggagcctg cggaggtgtgc aacgagagaa
1860
gccaggcota tacaccttggt tggcccaagaa cctggctgggg gctgacacta agacggttag
1920
tgtgtgttgtt gcgccgtgttc ttcctcctgc accggacagg gaagacaggg gcctggagct
1980
ccgggtgcag gagaccacc accatcacat cctgctatct tgggtcaccac ccccaaaacc
2040
agttgccacc aacacctcact cggccagtggc cttcctctccc cggggcggagc gggccacaccg
2100
tctggccgogc ctgctgcctgg gaaccccaacg ctacacacatt acctgcctcgc ttccagggcc
2160
ggatgactgg gcctgcctcg aagtgccctt tggctgtgcc caccaccagtg tggctgtgtgat
2220
atgggcccgg acccaagagg ccgagctttt ccacgaagcc ttaaggggac tctccttgcc
2280
cattgccatc ctggctctcg ctgtctctctct cctgccgacct gggctagcgg cccgccctgg
2340
cacaggccaa cccaggagag ggtggtggggtt gaggccgcccct ctccctccag cctggggcctt
2400
ctggggctgg agtgcctcccct ctgggtgctcc tttccctctgc tccgcctgaga
2460
tccaggagg aagtgccgca gatctctcagc agggtagaca ctgttcgccac ctatggctca
2520
aatactccta agctcagcct cttctcgcaga tcaagagaaat cactagagctt acttttttacc
2580
aagaagagac cagctctgggc cagatgcctc gccaggaaga ggcacatgac cacagtctc
2640
gaggcttgcc aagacagagctt gggttgtgtgg ctccggtggagt tccctgctcag
2700
ccttgaaaaa gttgcccttta cctccctagg ttcacctcctg tgcctattctg aggaacatct
2760
ccaggggaag ccgagccctt cctttcttct cccttttct ctctgcccag 2820
agggacactt gctgtggttg ggtgctcctt cctgtggtac cctgccgtga ccccttctcct 2880
tctctctctt gctagatcct cctgtccttt cctcttgggga ggggttgaag 2940
gagcccacact cctctctctt cccctccctg cctctctgctg ggggtgggtgcc 3000
tcttaaggggc cttaacctctt tctctctctt cggctttctgc gtctgatctcgt 3060
catttcccaag gcctttctct cccctctctt cggctttctgc gtctgatctcgt 3120
cctgaagccg cccctctctt cctctctctt cggctttctgc gtctgatctcgt 3180
aattggggg ccagggcag ccaatcctaa aaaaaa aaaaaa aaaaaa 3227

<210> 11
<211> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<592)
<223> "n" is A, C, G, or T

<400> 11
ttatcttctt ccctctctct cctctctctt cctctctctt ggtgaaccag cttagttgtt 60
tttcttctct cctctctctt cctctctctt cctctctctt cggggtttgcctt 120
tctctctctct cctctctctt cctctctctt cctctctctt cggggtttgcctt 180
ggctttctct cctctctctt cctctctctt cggggtttgcctt 240
ggctttctct cctctctctt cctctctctt cggggtttgcctt 300
tctctctctct cctctctctt cctctctctt cggggtttgcctt 360
aaaaatagat tctctctctt cctctctctt cctctctctt cggggtttgcctt 420
tctctctctct cctctctctt cctctctctt cggggtttgcctt 480
tctctctctct cctctctctt cctctctctt cggggtttgcctt 540
cccccccttt cctctctctt cctctctctt cggggtttgcctt 592

<210> 12
<211> DNA
<213> Homo sapiens

<400> 12
gagggagcag cccctctctt cctctctctt cctctctctt cggggtttgcctt 60
ggctttctct cctctctctt cctctctctt cggggtttgcctt 120
cccctggact tgggctgatt cttgtatctt gctgtctacc agctatgtga ctggagccaa
240
gttcttttacc ctcttccagga cctcatgttc ttaactctgca aacctggaggt attagtgttt
300
atctcatgtta cctgttatga gactcatatg aaaaaagaca ttcagagttct tcaagagatc
tgagccaatttg aggttgtgttc agtaagtttt tggcactattc ctttgacactt gtagctaaa
360
acttacctgg aatcagacctt ccactcagga gtcagaccag gctcagaaag gcaacacagt
tctggccccct tttgttttcca ttttggcttc taggccaagg cttgacctgtc cttcagcccc
420
tggcatctgta gagagaaact ggatagtgttc ctcatgctca ggtgagagga aacagcagtt
gctttgcatgg tttgattgag cgggtgtagtc taatctctctg aagaaagagaa ctcttgagac
480
acagagattg attcataaca tcaaacacaa ctcggcaccc atccccgact acagcaacaa
gaggcaagag aggtctcttg ttagagacgt ttctccacgg tgtggtcaga actgtgacag
540
taggaagcat ggtactgcga agacagataaa gcaagccccc ttagtcagcga ggtggggagg
ttttgagaa gactctctgc ggtggtgact gaaacttaag tgtgagcctg tgaacaagaagt
600
aggagttgct tggagacaaac tggagagag gctagggagc catcwctaatg tggtgccttc
cttacacccca aagatttttg tcagccgagat aagtggagt ggtgcaagag aagccaggtg
660
aagggggcgg gcaggtcacc agaaggtctt accgtagcac gttggactgt cagacccagg
cattctaaact ccactctgcct ctcttgccca acactgcccc gctgtggcga aagagaaac
720
actctctgcc ccagggctgg gaaagttgcca aacatcacaag caaatccccc acccagaga
tcctgaccgg cagaacaccc cagcctgcct ggtgtggtca ggtctgtgctg cttccacccc
780
agggtctcctt ggagcccccag ccgagttttt aagactttga tttatcttttc ccaacccttc
840
cccacccccc acactacactg ccagaggagat tttctggaaac tttatcttttt atctttcactt
900
tacccagttga gcagagggagc ctggagggaa aaggtgtaaa aaggttagag aactctacct
960
acatctctct cgggtccagaa ggactctgga gataacaca atttcagcccc atcactctttc
1020
cctccctccc aacacacact tccatgtacc acacacacat acacacacat acactctctcc
tcctctctact gcagactcag atgcacctcag tctgtgcagga ggtctcagaa aaggaacta
1080
aagctcttaag gacagggctgt gccatatta ctggtcctctt ctttggcgtgt tgtgcctaaa
1140
aacagttggag gcggccaggca cgggtgactca cccctgttaat cccagcattt tggagacccg
1200
aggagagagc atcagctttg cttcagagtt cgagacaccg ctggccacaa tggagaacc
1260
ccatctcttc ctcaaaaaac aaaaaattgc cagaggtggt gcgaggtgccc tgtaatccca
tgtactcagg tgtccagtgc gaggagactcg cttgaatccca ggagggggag gatgcaatca
1320
gctgagttgca cccctcgcact ccagccctgg tgaacagaatg aacactctgtc tcaacacaac
1380
aaacacggga  ggaggggtag  atactgcctc  tctgcaacct  ccttaaacct  gcactctcttt  1980
cttcacgggc  tgccctctgt  ggggccctggc  aatgactgag  caggcccaac  ccagagggac  2040
aaggaagaga  aggcataattg  agggagggacaa  gaagtgacgc  cgggtgtgaca  atgactgcc  2100
tggagggggtg  ttctccttggg  cctggcagg  gtgtgctgacc  cttaccctgc  aaaaccacaa  2160
gacagaggact  ccagactcttt  cttgtgaatg  gttcctctgcc  ctgcaagctcc  accatgagcgc  2220
tttctggtgc  cccacctcttg  ctatgcttggg  tggctgtgtgc  cactgcggct  gtgccctgtg  2280
tacccctggcg  tttctcctgc  ccccccactg  gttcctgtgca  gatccggcgc  tgtgatacg  2340
ccggctgttc  ctacccgcag  gctaccaactg  tggactgcaa  tgaactatct  ctagcggcag  2400
tcccccggggc  acctccggca  gcagcacaga  cccagccctt  gcagagcaac  agcattgtcc  2460
gtggagccca  gagctgagctg  ggttcaccagg  ccaatcctac  agagtccggac  cgtggcsga  2520
acagcttttc  gcagccgctg  gagtttgtgat  tccatcgcctt  gcccccgctg  ctgagcctgc  2580
acatagagga  gaaaccagctg  acgcccctgg  agggaccaac  ctgtgcaggg  ctggccagcc  2640	
tacaggaacct  cctactccac  cacaaccagc  tctacccgcat  cggccccagg  gctttttcttg  2700
gcctcagca  cctgctcgccg  ctgcaacctca  actccaaacct  cctgagggcc  attgcacagcc  2760
gcgtgtttttga  aatgctgcc  aacattggagaa  tactcatagatg  tggogggcaac  aaggtgtagatg  2820
ccatctgctga  catgaacctc  cggcctctgg  ccaaacctcgg  tagcctggtgtg  tctacgcaagc  2880
tgaaacttttc  gggatcctcc  gactatgccc  tggagggggtc  gcacaagcttc  gagaagctct  2940
cctctatga  caacagagctg  gccggcggtgc  cccggcgggccc  acctggaacag  gtgcggccgcc  3000
tcaagtttacct  agacactcaac  aagaaccgctc  tccagccgggt  agggccggccg  acctttgcaca  3060
acatgtgctca  ccttaaggag  ctgggactgga  acaacatgga  ggagcttggct  tccatgcaaca  3120
agtttcctct  ggtagacactc  cccgagctgta  ccaacgtgga  cattccacat  aaccccaacgc  3180
tgtctctactc  ccccccgcggc  gcttcaccacc  acctgccccca  gatggagacc  ctcatgtoctc  3240
acaacacgcc  ttctgctggc  ttgacaccac  agaaggtggga  gttccctgcc  aacctgacagg  3300
aggttagttctc  ccagggcaac  ccatccgctt  gtgaactggtg  cttaccgtgag  gcacatgccca  3360
cgggacacccg  tgcttgccttc  atcgagcgccg  aatccacacct  ctgctggcag  cttccggacc  3420
tccagccgcttc  ccggcttgcgtg  gattgctccct  tccgggagat  gcagcagcac  tgttgctcgc  3480
tcatctcccc  acgaagcttctc  cccccaagcgc  tccaggttagc  cagttggagag  agcatgcctgc  3540
tgcattgcgg  ggcacttgccg  gcacccgcacc  cccagatcta  ctgggtctact  cccagctgccgc  3600
ttcgactgac  acctccgcccctt  ccagggagaga  ggtaccgggt  gtacccggag  ggcaccctgg  3660
agctgccggag  ggtgacagca  ggaagggcag  ggctataaca  ctggtgctgcc  cagaaaccttg  3720
tgggggctga cactaagcac gttagtgtgag ttgtggccgaga gtctctctct cagccagggcag 3780
gggcggagac cagggggcttg gaagccggcgc tgccaggggac ccacccctctat ccaactctgc 3840	
tatctttgggt caccccaacc aaacagagttt cacaccaacct cacttggttcc agttctctctt 3900
ccttcctggg ccaggggtgac acagctctggg ccggcctgccc tggggaacc cacacgtcaca 3960
dacattaccgg cctcttcctcag gccaagggagt actgggctctgc cttgcaagagtg gcttttgcttg 4020
tttgcaccaccc ccaatttggtg tctgtttgtgg ccagggccacc aagggccact tttgcacaca 4080
gagccttagg ggaatctctct ggggctcatgg ccacaccctggc ttccctgctgc tctttctcggg 4140
cagctgggtactg ggggcccacc cttggccacag gccaacccgag gaaggggtgtg ggtggggagggc 4200
ggcctctccacctc gtttttctggg gctggaatgct cccctctctgtgc cggggtttgtgtg 4260
cctgccctcct gcgtctgctggtt ggaatccgag gggaggaagct gccgcaatcc tcaagaagggg 4320
dagaactgtttt gggcaccatttt ttttttaaatt tttggaatctgc gcaccttggtc tcaagctaga 4380
gaaatcactattg gaaacactcattt tttacaaagag aagagcactgc tggggccagat ggcctgtgcag 4440
gaaagggcaca tggccacaagctgctcggggctccttgccagctg ggcgaagacagttagggccctt 4500
tgtgccccctgtt ggggtgtctcc tgcagcctgtcga aataagttgc cccctacctctcctc taggggtccacc 4560
tctgtcgccaa tcctgttcagaa cattttaaacgc gacaagaggtg aacttttgtgtgc agagctctctc 4620
gcctccccat cttctctgcct cccagaggtct cctggggctgt gcctttctctgc cccctactcg 4680
tgtgcccc CGCC CGTC GCTC TTGT CTTG GTGG CCGG CCGG CTAGT GCCT GTGCT 4740
tgtgcctctctt gggcaggggtg tgaagagggct cactccattct cactcctgggg gggccctctgtc 4800
aattggtggag tggaccccaagc cagacagcagga gagacattttg gagagggagt cccagggagct 4860
cctcttcctcc gtcagcctgctt tggccccttgc ggaagttgctag aatttttctga gactccctg 4920
ccttttgctt gaaatgcctgct cttctcccaacc gacagtccctct ttttttctctgc ttttttctgt 4980
aaaaataaaa ataaaataa aacaataaaaa aaaaaa aaaaaaaaa aaaa 5036
accccttacce tgcgaacacac aaagagcagg actccagact cttctttgtgc atggcccotct 360
gccttgccag tccaccatgta ggtttcctgt ggcgccactc tgtgtagcttt ggtggtgccg 420
tgcaactgge gcgtgtgcggc tggtaacctg gcacgttcccc tggcccccccct acgtgtgctt 480
cagcagcgg cccctggtata gcgcggcggct gtctacctgce gcaggctaaaa cttggtgacgt 540
cattacatgct tgttctgcagc gactctccccc gcaggccacag agacccctgct 600
cctgccagac aacagcattg tccgtggtga gcagagttgag cttggtctcc tggcccaactct 660
cagcagcgtg gactgtgccc agaagagctt tttggtggtcc cggagacttggt atttcccattgc 720
cctgccggcg agcgtgtgagc tgacctagaga ggagaacccag ctgaccgcccc ggaggagcc 780
cagtcttcga gcctggccca gacctacagga acctcatctc aaccacaacc agctctacctg 840
catgcccccag ggccctcttt ttgcccctccg caacccctctg ccggctgcacc taacattcaca 900
cctccctgagg gcccgtgacct gcgcgaggtt cggaaagtgc ctacactctg ggcaacctgg agataactcat 960
gatttgccggc aacaaaggtag atggcactcct ggacatggaac tttcgccggccc tgcccaacctt 1020
gcgatggccgt gctgtcctacag gcacgaaacct gcggggagata tcggacatag cccctgaggag 1080
gctgcacaggc ctggagacgc tttcctcttct caggaactcag ctggccgggg gccgagggg 1140
gcggccgagac caggtgcgcc gcgtcagatt actagacattg aacaagaacc agctctgacgc 1200
ggggaggcgg cgggactttgg ccaacagtct gcacccctaa gacgctggac tggaaacaact 1260
ggggaggctg gctctccatcg aacaagttttgc cttggtgacat tccccgggc gcagccagct 1320
ggcacatccac ataattaccag ggcggctctctt catccaccccc cgcgcttccttc acacactggc 1380
ccagatggag accctcatgc tcaacacaac gcgtctctcag gcctttgcacc agcagagccgc 1440
ggagtcctgtt ccccaacctgcg agaggtaggg tctccacgcgg acccccatccc gctgtgacgt 1500
tgccatccgc ggcccacatgt ccaccgggac cgtgtgctgcc ttgctctgac gcgaatccac 1560
cctgttgcag gaggctgccg acctccagcc cctccccggtc cgctgtgacct cctctgttggg 1620
gatgcaggac caacctttgc ccatctctct cccacgaagg tcccccccaac gcctctcagtt 1680
gcggaggtgg aagagcagtgg tgtgtgctctg ggcggcactct gcggacaccc gacccgagat 1740
cctacttggac acctccacgctg gggctcgact gacacccgtgc catgcagcgc gggaggtacg 1800
gggtgacccc gaggccaccc ttgagctccg gagggtgaca gcagaagaggg cggggctata 1860
cacccctgctg cccggagacc tgtgggggcgc tggacatcaag aocgttagtg tgtttgtgga 1920
cggtgctctc ttcgggccgc caggggacag aaggccaggg cttggagctcg gggtggcgga 1980
ggcgccaccac taaactctac tgctctctcg ggctccaccc cccaaacacag tggcaccacc 2040
cctcaactgtt tccaggtgctt cctccctgcc gcggcaggcc gcagacagcgc tgcgcgcccc 2100
gcctcgggga acccacagct acaacattac ccggctctcctt cagggcaagg agtaetgggc 2160
ctgctcggca atggccctttg cttgatgcgca caccaggcttg gcttctgtat gggccaggac 2220
caaagagggg aaccttgcctgc aagggcgacag gctgggtgtaa cctggccattt 2280
ggcctcgcct gttccctctcc tggccagctgg gctagcgccgg caccctggca cagggcaacc 2340
caggaagggt gttggcgggga ggcggctcctt ccctecagcc tgggcttctct ggggcttggag 2400
tccctcctcttg tccgagggttg tgtcgtcgctc cctctgctcttg ctctggaatc cagggagga 2460
gctgcccaga tctccagaaag gggagacact gttgccacca aattccttgcag 2520
tcgcacgtgt gcctacgcgt atagagatcaca ctaggtaactacc ctttttacca aagagaagca 2580
gcttgccgca gatggccctgc cagggaggga acatggaccc acgtgccttgag ggcctggcag 2640
ctgccgcaagg acagatggggg ctggtggtgcc cttggggtgc ttctgcgagcc tcgaatttgg 2700
tgcccttacc tccataggtgc acctcgtgtgc acattctgtag gaaacattcc caggaacagg 2760
aggaccctttg gctagagcct cctg acctttct ccctgccccagag gctccggtgc 2820
cgtgtcttggc tgcctccctac ctggtctccc cggcctgacc ccctccctctt cctctctctc 2880
gtaagcttcc aggttgctttgc ctctctgtcc cctgggcaag ggcctggaag ggcacactcca 2940
tctccactcc gggcgcctgc ctcataattgg gactgacccc agccagatct gaggacattg 3000
tggcagcggg atggcccaag acggccctctc tcagcgtcct gggcctggca ttcgcgaagct 3060
gactcttttt atggcactttta gtcacgctttg ggcagaaaaa tgtgcctcctccc ccaacccgat 3120
tccactctttt ctctgttttt gtaaaaaata aaatatata ataacaatata tccgggggaa 3180
aggacgaaaa gggaaaaaaa aaaaaa 3207

<210> 14
<211> 3170
<212> DNA
<213> Homo sapiens

<400> 14
gctccacgga caacctcttc gcggcggctgc ggcgcgagaa cttcgccaaac atgacccagcc 60
ttgcgcaacct cacctcttcc cggacacacca tcggccagatt ggcagctggc gctttccggc 120
acctcgggtgc cttcgcggcc ccagacacctg cagcagccggg gctgcgggga gggcatttcctt 180
accagctcgg gcggcgctgggg aacctgcgggc acctgtctct tggaaacacac cagacgccggc 240
gggtgcagtc ggcggcgctttg gacgccttccc tgtccacgctt gaggacacct catgtcttctc 300
acaacaaacct ggcggcctttg cggctggagg ggtgggccaa gatgtatgac aataacaccc 360
tcagcgttggc cacaacacctct cgagaccaca tcggcggggg gaccttcgct cagcttcaca 420
agctggctcg tctggacatg acctccaaacc gctctgcataa actccgcccc gccgccccct 480
tctctaggct gcaggccacc gggcccaacgc cgcceacccc gctgacccgct agctcgggctg 540
gcaacccctt gcacccgcaac tgcagcgtgcc tctggtctgcg gggctgaccgcc cggagggaccg 600
acttagagac ctcgcggaagc cccgaacacc tcaacgcagcc ctaccttctgg tctccctccccg 660
aggaggattg cttgtgtgag ccccgccgtga tcacacgcca gcgggccccgg ggccgctctgg 720
tgttgggaagg ccagccgggtg agctgctcgct gcgcagccggtt ggttgacccccc gagccggctgg 780
tgcacctgggt gcgcctgctat ggccgctgtgc tcgggagaccc cagccgcacc gccgctccgg 840
gggacgaggac gctggtatgtg accatacacc ccttgagggc cagtggcacc ttcactttaga 900
tcgccctcaca tgctgctggg gaagcgcagcc gcgcctggtga ggttgctgtgg gcctctctgc 960
cctctgtgac accecggcggc gcgctgccccgc cgccctcacc cgaccgccccgc toctctgaca 1020
tcgccacgcc gcggcagacc cggccagaccg aggctgctgct cgacagcccggg cagaagcagcag 1080
ccagctctcc acacccctcct gcgctcgcttc cgctggccgcag cagcagctggt tgcgctgcgg 1140	tacgcattgta ccaggttccag tacaacagttt ccgctttgatga ctcctctgcac tccaggta 1200
tccgcttcac cagtcagacc ttctctgtgta atgacccgctg gcgcgcgcctg gccctacgac 1260
tgtgctgtgtc gcgcgcctcag gacgcgagggg ccacagcgctgc gcgcggcaccgc cagagttgtgg 1320
gctgtgtaca gcgctccacc ccttgaggggc gcggccctctgg cgcctgcgctgg aggccccatt 1380
tcctgcgggg cacatgtacgc atgctgcctcg gcgggtctcat gcgcgctgctgt gcctctctgtc 1440	tcacctgttt gcctcatctgc gctataaggg tgtatgctggg cggggacacgc gcgcgcgtcga 1500
agggcctccag gtctgctcccg cggggtacgcac acgtgtgtgtcc gcgccacacc cgcgcgcgc 1560
caggccgccg cagccgccccg gcgcctgcgctg cccagacacc ccacagggcgc ctgcgcgagg 1620
tggagtcctca ggctgtccccc gcgcggctgctg cccagccacc gcgcctacggc gcgcgcgcgg 1680
cattccgcggag gccgcttggtgc gcgtctgtgagg ctctctctctg gggctgccccg acctccgctgt 1740
gctgtgctgg aacgtgcaagc aacctccgcccc gggctggtggc gggccctctgg gcgcctgga 1800
gagagccgcatgc gcggccgctgcg gacccacacc cccagccggcc cccctatctcct gcctctgttc 1860
cggggggagcg cgcggccggcc gcggaggggc gcgcagctgtta tgcgctgtgg gcggactaccg 1920
ggagcccttt acagacgcacc atggaccccc gcgcgcgcgcgg gcggacacag gcgcaccgcgt 1980
ccagccgccga ccttgtggagc gcgcggagagc gcgcgccccg gcgcgtgacag gcacctgacccgc 2040
tgggtcctgac cggcgcctgcg gctggccctcag caccgacccga gccgtgtctgct gcagttaccg 2100
tgtgacgccc gcggccgggccc gcggagcgcct gcgtctgaccgc gccacccgcag ccagcgcgcac 2160
gagacgtcgg gcggagctgg gagaaagcgc agcgcaccaagc cattggacaca gatgtggagac 2220
gccccttcttg tcccgagggag ggccgggacc gctccgggtg gggctcggcg ggcacgcgcc 2280
cggcgcagg gcccgtggcg ggacggcgtg gcggctcgtc ttccctctat ggaactccctg 2340
acccccctcc taccctccct ctcggcgcct ccggacactc gctggagcgc gtcgcttcac 2400
ccagcagagc ccggaggggg caggcccccct tcacactgca gcactgagc acagagcccc 2460
tccccccagcg cgctacccgg ccggccgggc ccgggceaat ttcctgtatc ttgcctgact 2520
agattcccatt ctgtggtcgcc gcggctctcc aacgcctccca cccaccccccc gcagacttccc 2580
tgcgctgggc gggctctgct tggggctccc ctttctctgt ttccccgttt gttctctatc 2640
ccggcctctt ttgctcttctc tggagtgcct gttcttcctct atttgctctc cttttctctc 2700
tgcgcttgc ctggctctcc ctggcgcctc cggtgttttg tcagttcttc ctgtctctcc 2760
tgatctcttc ttctttacta ccggctcgcc cgggtgctac cggagcagc acacactcc 2820
acacactcc ccacacgcaac aacaatccaa accaaatcctc cctccctacc ccggcgcgggga 2880
cctccgagcg cagcagaaat aacacttttt tggctgtctga ggcctcgtgc acctgtgtgt 2940
gttgtctgtat gttgtgctgc gttgtgtgtg ttggtgtgtg ttggtgtgtgt ttggtgtgtgt 3000
ttgtgttggt cggagtttta ctcagattgc agcataagg cttcaagtga gactgaagga 3060
agatgggag atgactaact ggcccccggag gagaactggca gagagctttc tattctcttg 3120
gagaacctga ggtggcggat aattcacaata ataaatgat cataatgct 3170

<210>  15
<211>  2053
<212> DNA
<213> Homo sapiens

<400>  15
ccggctcggc ccctcgggac ccaagctccc cggccttccag acgccccccc gcgcctcggc 60
agctccgggg aaacgagac gcgcagatcc ggaggggctc cccgcccccc gcgcggggggg 120
gacgggggtct cggcgggtgc gcgcactgag agttgtactcc ttgggttggt ctctctctct 180
tctccacact cctgggcacat ctcctgctcct cccctggggtc ctggcgctgc 240
tccgcctccag ccggcgcgctc gcggagcagc ccggctcgtgc gcgcagctgc 300
cgcacagtca agtgccggtta ccggctatcg accgaggtgc ccaacggactc gcgcctcctc 360
gtggccaccc ttcctctctac ccggacagag ctcggcctgc ccctccctgg ccggctccgc 420
cgccggccgc gcgtgcggag gctggcgcgct ctcacccctc gcgcgccgcct ctcgggcgag 480
gtggccgggc gcgtgcctgc gcagctgccc gcgcctcgcgc gctgcgtcgc cgcggacacc 540
cactggccgc acctcaggtcc ttggcttctt ccggcagcag ctcgccgctg ctcgggccccc 600
agtccctttg tggacactgat ccgaaaccac atctgctgcc ctgaagatga gcgcagacac 660
cggagccttcg agggcatgtt ggtggcggcc ctgtctggcgc gcctgacact gcaggggctc
720
cggcgcttggt aagctggccag caccaacactt ctttcacctgc cgccgagatgt gctgccccaa
780
cgccgccgcc tcaggccact ggaatttaagtt aataatcctgc ttggtagcctgc gactctacagtg
840
tccttccgca acctgacaca ttctgataagcc tctcaccctgg aggacatgac ccctcaaggtc
900
tctccacatgg gcataagctgc tgagttgaaat ggtctacacc acatatggatt ttctctgagc
960
aacaatcctt ggtctcgtgca ctgccacatgg gcagacatgag tgcacctgctgca gcaggaacact
1020
saggtgcggc agggcgaaga ccggtcaccct tggccatatt ccggaaaaat gaggaactcgg
1080
gtctcctctgg aacctcaacag tgctgacacttg gactgtgacc gcattctccc cccatcctg
1140
caacatcctt gtcctctcctt ggtattgttg ttagccctga taggagctat tttctctctg
1200
gtctttgtatt ttgacccgca aagggataaac aagttgagagttg ataacatccag agatgcctgc
1260
aggaggatcaca tgggacggtta tctacacata gatgagccac ccgaggttaa aacaacacactg
1320
aaccctcagtct caaactccgga tgtctgagaa atattagagg acagaccaag gccacccctg
1380
cattgagatgt acacttaagcc tttatcccta ctaggctttgc tcctacttctc tctctccta
1440
tagatacaac cgaacctgtc taaaggcaat ccgagggatt tgctcctcttg ttatgtaaag
1500
tttctcgggttg ttgctctgtta atgtaagcag atgaaacagt ttggtatagt ttcttctctc
1560
ttttctctcttt ttggaactctt caaacgtatt ccgggatttt ttggagggttc agcatgaaaca
1620
ugggcctcttt gtgtctcgtgc tctctctcctg cagacattcag ggtgtaagcc ctgtaacccac
1680
acagatagcata ttttacaaaag tctgcctcgaan gaaaataactt tattcataa
1740
tacagttttt atctttcatgt acctaaggttg tggagaaat aattgcattc tataaactgtc
1800
tccagacagt tagcaggctc ttcaaatata cttcatggttg cacaggacag ccctgcatcc
1860
agaggcatgt ctacatcttct gtttctgcatt atataaaaaa ataactgtca actctcataac
1920
tcttttttga aatgaaatta ctctcttgatt tacagttttat atgaaattgt actgatttttt
1980
tttttaataaa ctgcctgcag acatcaggcag atccaaacgga ctgaatttgttt aaaaaaaa aaaaataaag
2040
attctttaaaa gaa
2053

<210> 16
<211> 973
<212> DNA
<213> Homo sapiens

<400> 16
cagcccccagct ccggggaaac gcagagcgcgc atggcctggg ggtgctcccg ggccccccgcc
60
gcgggggcag gcgcgtccgct gcctgcccgc ctagccctgg gttccttgg gtccttggc
tcgtcttttc ccaactcttc ggcatectcc ttctctctct cgggegcggtt cctggcttcc
240

gcgtgtgctg ccagccccc cgtgcggac ccagccgcc gcctgtgcga gtgtcgcgg
300
gcagccggca cagtcacgga cgattaaccgc aatctgacgg aggtggccac gcagctgccc
gcctacgtgc gcacaccttt ctttaacgcag aacacacctccc tttactctgcc gcgggggatgtg
cggtcccaac tcgcccagct caggcaacag tcacattaagta ataattcgtc ggtggagctg
tcagcgctgt ccttcgcgaa cctgacacat ctagaagaagc tccaccttggc gcacagatgc
540
tcaagctcc ttcaacaatgg cacaacctgcg cagttgcaag gtctaccccc cattaggggtt
600

ttcctggcaac acaatcccttg ggctgcggac tgcacacaggg cagacatggt gcagctggccc
720

aaggaacacag aggtagttgc gggaaagagc cggtgcaactt ggtcatactc gcagaaaaagtt
840

gagaatccgg tcctctttggga acactacaggt gcggacctgg gctgtgacccc gattctttccc
900

caatctctgg acaatctctta tgtctctctct ggttatgttt ttgcctcttg aggctgatttt
ttcctctttg tttttttttt gacccggagc gggataaaaa agtggagtgca taacacatcga
960

gatgccttcag cggatcagac gcagaggtta cattacagat atggtaatca tcgggacccc
acaactctggc atg
973

<210> 17
<211> 1331
<212> DNA
<213> Homo sapiens

<400> 17
cagcccaagct ccgagggaaac gcagagccggc atgcctggtgg ggtgcctcgg gcggccggcc
ggcgggacgc gcgtcttcgag cggcggtcag tcgctctgta tgtctcgcggc
tgcgttcttc ccagctctcgc ggcgtcctcc ttctctctct cgggcgcgtt gccgctcggc
ggcgttcgac ccagcgcgcgg cgtgcgcgagc cagtgcccccg cgtctgct public gcgctcggc
cgagcggcgcg cagtcacgga cgattaaccgc aatctgacgg agtgacccgac gcagctgccc
gcctacgtgc gcacaccttt ctttaacgcag aacacacctccc tttactctgcc gcgggggatgtg
cggtcccaac tcgcccagct caggcaacag tcacattaagta ataattcgtc ggtggagctg
tcagcgctgt ccttcgcgaa cctgacacat ctagaagaagc tccaccttggc gcacagatgc
540
tcaagctcc ttcaacaatgg cacaacctgcg cagttgcaag gtctaccccc cattaggggtt
600

ttcctggcaac acaatcccttg ggctgcggac tgcacacaggg cagacatggt gcagctggccc
720

aaggaacacag aggtagttgc gggaaagagc cggtgcaactt ggtcatactc gcagaaaaagtt
840

gagaatccgg tcctctttggga acactacaggt gcggacctgg gctgtgacccc gattctttccc
900

caatctctgg acaatctctta tgtctctctct ggttatgttt ttgcctcttg aggctgatttt
ttcctctttg tttttttttt gacccggagc gggataaaaa agtggagtgca taacacatcga
960

gatgccttcag cggatcagac gcagaggtta cattacagat atggtaatca tcgggacccc
acaactctggc atg
973

32/56
gccccactgc ccagcctcag gcacctggac ttaagtaata attogctgtt gacgccctgcc 780
taagttgctct ccgcgaacct gcacatctca gaagacctcc acctggaggga caatgccctc 840
aaggttcttc acaatggcgc cctggtctgag ttgcaaggtc taaccccacat tagggttttc 900
cctggacaaca ataacctgggt ctgcccacgct ccacatggcag acaatgtgct cctggctcaag 960
gaaacagagg tgtgcagagg caaagacccg ctcacatgctg cattccggga aaaaatgagg 1020
aatcgggcct tcctgccaaact caacaggtct gcacctggact gtggccgctat ctcctcccaca 1080
tccctgcaaa cctcttatgct cctcccttgtttag ctctgtatgcc cgtatattttt 1140
cctctgttttt tgtatttgaa ccgcaccgggg ataaadattg gatgcataa ctcagagat 1200
gccctcaggg atcaaatggtg aggattcact tgcacagatgg aatacatctgc gacaccaca 1260
ttaacaaaco tcagttctca ctcggtatgc tgagaaatat tagaggacag accaagcaca 1320
actctgcag a 1331

<210> 18
<211> 2053
<212> DNA
<213> Homo sapiens

<400> 18
ccgcctgcgg ccctccggcc ccagcctcag gcagcttggag agcgagggcgc gcctggccgcctg 60
agctccgggg aacagcggcgc cggagatcct ggggtgtgtgc cccggggcgc gcggggcgcc 120
gaacggcgtc tggcgcttgcc gcgcactcaga cttggtactcc tggggtgggtg ctcctcggtct 180
ttcctccactc cctgcgtccct ctcctcgtcc ctcctccctgc cgggtccgctgc ttcctccctgtg 240
tccgcgcagc ccocggcgcc gcgcagctgc cggccgcgtgt gcggaggtgct gcgggagccgcg 300
cgcacagtca bgctgcattaa ccggcattct gacgcagggc gcacgaggct cccggccttac 360
gtgcgccacc ttctcctttac gcggacaaag ctgggctgctgc cctctccgcgg gcctggcctgc 420
cgcggggcgc gcgttgcggga gtctgccggc gtcacattcgc gcgcaggccc gcgggagcag 480
gtgcggccgg ccggcctcga gcacatgccc agcctggccgg agcctgcagt gcggccacac 540
ccctcggcgg acctggagtc ctggcctttgc gcggcctgca gacgcagcgt gcggccccccc 600
agtcctcctg tgtgcacagt ttcgtggacc acjcgtgcct ccgtgacagtat ggcgcagaaac 660
cggacgcctcg ggacagctgg cttctgccc ctgggcctacc ggggcagat agcggggctgc 720
cgcgcctccgtc agtgagccag ccaccacacc cttgacctgc gcgggtgtgtg cgggtcaca 780
cgcgccaccgc tcagccacct gcgggttattgt tattacatgc gcggggctgt gcggcctaca 840
ttcctcggcgc acctgacaca tttcagaaagc ctcacctctgg ggagcaaatgc ctcacagttc 900

33/56
cttcacaatg gcaccctggc tgagggtgaa ggtctacccc acattagggt tttctggac 960
aacatcctt gggtctgcaag cgaccaacatg cgacatcttg tggcttctgt caaggaacaa 1020
gaggtatgct agggcaagaag cccgctcccc cctgctgacc cccgatgctgttg 1080
gctctcttggtaa ctcctcaacag tgcctgaccct gactgacccgcc cccatcccc 1140
cacactccctt gatgctctctt ttgctcgctga taggccgctatt tttctccttg 1200
gttttgtatt tcggcggacca aggcccaaaa aatctggagtt aatcatccag agatgctcgtg 1260
agggataca taagagggta tcattacaga tatgaaatca atgcggaccct cagattaacaa 1320
aacctcagtt ctaactcggagagttccttgaga atttagagg acagaccaag gcacaacctgtg 1380
catgagatgc agacttaagct ttattcctcatc tagctctggtc tccactttctca tcctccacat 1440
tagatacaac gcaccaagcc aaaaaacagt gaggggatttggtctctctg tatgtgaaag 1500
tttctctgttc gttgtctgtttatatgaagcc atggagacagtt atggatagtcgt cttatccctc 1560	ttttttttttt tgcacccagt gacgccagttt ttcacgttctcgct gcagactgacacta 1620
tgggtttctctttgct gtctctctcgctctctcag ttacacggctgag ggtgtagccgct atgggtccctc 1680
cagatagcatc tcaacaacaaat ggcggctctcag ctttttcgag aaaaaaatctttt tattcataaa 1740	tatcatatatattctctatgt accaaggttg tggagaataat atggccatcctataaaactgc 1800
tgcgagactgtg tagcaggtctctctaatctgccatctcagcacggagctactgcctctgcgctcttactgc 1860
agagcatgctactatctctctctgctctctagtgaaaataaataattctg 1920	tttttcaggtc aagtaataattcctttttgtatgcagtttatgagaaatgtactcgatttttt 1980
ttttaataaa tggcagctcggg atcaaccggcgtgaaatgcgtttgaattttaaaaattaaatcctctcttactgc 2040
attcattaaaa gaa 2053

<210> 19
<211> 845
<212> PRT
<213> Homo sapiens
<400> 19

Met Leu Ser Gly Val Trp Phe Leu Ser Val Leu Thr Val Ala Gly Ile 1 5 10 15

Leu Gln Thr Glu Ser Arg Lys Thr Ala Lys Asp Ile Cys Lys Ile Arg 20 25 30

Cys Leu Cys Glu Glu Lys Glu Asn Val Leu Asn Ile Asn Cys Glu Asn 35 40 45

---

34/56
Lys Gly Phe Thr Thr Val Ser Leu Leu Gln Pro Pro Gln Tyr Arg Ile
50  55  60

Tyr Gln Leu Phe Leu Asn Gly Asn Leu Thr Arg Leu Tyr Pro Asn
65  70  75  80

Glu Phe Val Asn Tyr Ser Asn Ala Val Thr Leu His Leu Gly Asn Asn
85  90  95

Gly Leu Gln Glu Ile Arg Thr Gly Ala Phe Ser Gly Leu Lys Thr Leu
100 105 110

Lys Arg Leu His Leu Asn Asn Lys Leu Glu Ile Leu Arg Glu Asp
115 120 125

Thr Phe Leu Gly Leu Glu Ser Leu Glu Tyr Leu Gln Ala Asp Tyr Asn
130 135 140

Tyr Ile Ser Ala Ile Glu Ala Gly Ala Phe Ser Lys Leu Asn Lys Leu
145 150 155 160

Lys Val Leu Ile Leu Asn Asp Asn Leu Leu Leu Ser Leu Pro Ser Asn
165 170 175

Val Phe Arg Phe Val Leu Leu Thr His Leu Asp Leu Arg Gly Asn Arg
180 185 190

Leu Lys Val Met Pro Phe Ala Gly Val Leu Glu His Ile Gly Gly Ile
195 200 205

Met Glu Ile Gln Leu Glu Glu Asn Pro Trp Asn Cys Thr Cys Asp Leu
210 215 220

Leu Pro Leu Lys Ala Trp Leu Asp Thr Ile Thr Val Phe Val Gly Glu
225 230 235 240

Ile Val Cys Glu Thr Pro Phe Arg Leu His Gly Lys Asp Val Thr Gln
245 250 255

Leu Thr Arg Gln Asp Leu Cys Pro Arg Lys Ser Ala Ser Asp Ser Ser
260 265 270

Gln Arg Gly Ser His Ala Asp Thr His Val Gln Arg Leu Ser Pro Thr
275 280 285
Met Asn Pro Ala Leu Asn Pro Thr Arg Ala Pro Lys Ala Ser Arg Pro
290 295 300
Pro Lys Met Arg Asn Arg Pro Thr Pro Arg Val Thr Val Ser Lys Asp
305 310 315 320
Arg Gln Ser Phe Gly Pro Ile Met Val Tyr Gln Thr Lys Ser Pro Val
325 330 335
Pro Leu Thr Cys Pro Ser Ser Cys Val Cys Thr Ser Gln Ser Ser Asp
340 345 350
Asn Gly Leu Asn Val Asn Cys Gln Glu Arg Lys Phe Thr Asn Ile Ser
355 360 365
Asp Leu Gln Pro Lys Pro Thr Ser Pro Lys Lys Tyr Leu Thr Gly
370 375 380
Asn Tyr Leu Gln Thr Val Tyr Lys Asn Asp Leu Leu Glu Tyr Ser Ser
385 390 395 400
Leu Asp Leu Leu His Leu Gly Asn Asn Arg Ile Ala Val Ile Gln Glu
405 410 415
Gly Ala Phe Thr Asn Leu Thr Ser Leu Arg Arg Leu Tyr Leu Asn Asp
420 425 430
Asn Tyr Leu Glu Val Leu Tyr Pro Ser Met Phe Asp Gly Leu Gln Ser
435 440 445
Leu Gln Tyr Leu Tyr Leu Glu Tyr Asn Val Ile Lys Glu Ile Lys Pro
450 455 460
Leu Thr Phe Asp Ala Leu Ile Asn Leu Gln Leu Leu Phe Leu Asn Asn
465 470 475 480
Asn Leu Leu Arg Ser Leu Pro Asp Asn Ile Phe Gly Gly Thr Ala Leu
485 490 495
Thr Arg Leu Asn Leu Arg Asn His Phe Ser His Leu Pro Val Lys
500 505 510
Gly Val Leu Asp Gln Leu Pro Ala Phe Ile Gln Ile Asp Leu Gln Glu
515 520 525
Ala Gly Leu Val His Tyr Asn Phe Cys Thr Leu Pro Lys Arg Gln Phe
    770  775  780

Ala Pro Ser Tyr Glu Ser Arg Arg Gln Asn Gln Asp Arg Ile Asn Lys
    785  790  795  800

Thr Val Leu Tyr Gly Thr Pro Arg Lys Cys Phe Val Gly Gln Ser Lys
    805  810  815

Pro Asn His Pro Leu Leu Gln Ala Pro Gln Ser Glu Pro Asp Tyr
    820  825  830

Leu Glu Val Leu Glu Lys Gln Thr Ala Ile Ser Gln Leu
    835  840  845

<210>  20
<211>  1477
<212> PRT
<213> Homo sapiens

<400>  20

Met Ala Lys Arg Ser Arg Gly Pro Gly Arg Arg Cys Leu Leu Ala Leu
    1    5    10   15

Val Leu Phe Cys Ala Trp Gly Thr Leu Ala Val Ala Gln Lys Pro
    20   25   30

Gly Ala Gly Cys Pro Ser Arg Cys Leu Cys Phe Arg Thr Thr Val Arg
    35   40   45

Cys Met His Leu Leu Leu Gln Ala Val Pro Ala Val Ala Pro Gln Thr
    50   55   60

Ser Ile Leu Asp Leu Arg Phe Asn Arg Ile Arg Glu Ile Gln Pro Gly
    65   70   75   80

Ala Phe Arg Arg Leu Arg Asn Leu Asn Thr Leu Leu Leu Asn Asn Asn
    85   90   95

Gln Ile Lys Arg Ile Pro Ser Gly Ala Phe Glu Asp Leu Glu Asn Leu
    100  105  110

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

Ala Phe Lys Gly Leu Ala Ser Leu Glu Gln Leu Tyr Leu His Phe Asn
Val Asp Pro Arg Val Asn Ile Thr Pro Ser Gly Gly Leu Tyr Ile Gln
Asn Val Val Gln Gly Asp Ser Gly Glu Tyr Ala Cys Ser Ala Thr Asn
Asn Ile Asp Ser Val His Ala Thr Ala Phe Ile Ile Val Gln Ala Leu
Pro Gln Phe Thr Val Thr Pro Gln Asp Arg Val Val Ile Glu Gly Gln
Thr Val Asp Phe Gln Cys Glu Ala Lys Gly Asn Pro Pro Pro Val Ile
Ala Trp Thr Lys Gly Gly Ser Gln Leu Ser Val Asp Arg Arg His Leu
Val Leu Ser Ser Gly Thr Leu Arg Ile Ser Gly Val Ala Leu His Asp
Gln Gly Gln Tyr Glu Cys Gln Ala Val Asn Ile Ile Gly Ser Gln Lys
Val Val Ala His Leu Thr Val Gln Pro Arg Val Thr Pro Val Phe Ala
Ser Ile Pro Ser Asp Thr Thr Val Glu Val Gly Ala Asn Val Gln Leu
Pro Cys Ser Ser Gln Gly Glu Pro Glu Pro Ala Ile Thr Trp Asn Lys
Asp Gly Val Gln Val Thr Glu Ser Gly Lys Phe His Ile Ser Pro Glu
Gly Phe Leu Thr Ile Asn Asp Val Gly Pro Ala Asp Ala Gly Arg Tyr
Glu Cys Val Ala Arg Asn Thr Ile Gly Ser Ala Ser Val Ser Met Val
Leu Ser Val Asn Asp Val Ser Arg Asn Gly Asp Pro Phe Val Ala Thr
Ser Ile Val Glu Ala Ile Ala Thr Val Asp Arg Ala Ile Asn Ser Thr
625 630 635 640

Arg Thr His Leu Phe Asp Ser Arg Pro Arg Ser Pro Asn Asp Leu Leu
645 650 655

Ala Leu Phe Arg Tyr Pro Arg Asp Pro Tyr Thr Val Glu Gln Ala Arg
660 665 670

Ala Gly Glu Ile Phe Glu Arg Thr Leu Gln Leu Ile Gln Glu His Val
675 680 685

Gln His Gly Leu Met Val Asp Leu Asn Gly Thr Ser Tyr His Tyr Asn
690 695 700

Asp Leu Val Ser Pro Gln Tyr Leu Asn Leu Ile Ala Asn Leu Ser Gly
705 710 715 720

Cys Thr Ala His Arg Arg Val Asn Asn Cys Ser Asp Met Cys Phe His
725 730 735

Gln Lys Tyr Arg Thr His Asp Gly Thr Cys Asn Asn Leu Gln His Pro
740 745 750

Met Trp Gly Ala Ser Leu Thr Ala Phe Glu Arg Leu Leu Lys Ser Val
755 760 765

Tyr Glu Asn Gly Phe Asn Thr Pro Arg Gly Ile Asn Pro His Arg Leu
770 775 780

Tyr Asn Gly His Ala Leu Pro Met Pro Arg Leu Val Ser Thr Thr Leu
785 790 795 800

Ile Gly Thr Glu Thr Val Thr Pro Asp Glu Gln Phe Thr His Met Leu
805 810 815

Met Gln Trp Gly Gln Phe Leu Asp His Asp Leu Asp Ser Thr Val Val
820 825 830

Ala Leu Ser Gln Ala Arg Phe Ser Asp Gly Gln His Cys Ser Asn Val
835 840 845

Cys Ser Asn Asp Pro Pro Cys Phe Ser Val Met Ile Pro Pro Asn Asp
Ala Phe Phe Ser Pro Phe Arg Ile Val Asn Glu Gly Gly Ile Asp
1100 1105 1110

Pro Leu Leu Arg Gly Leu Phe Gly Val Ala Gly Lys Met Arg Val
1115 1120 1125

Pro Ser Gln Leu Leu Asn Thr Glu Leu Thr Glu Arg Leu Phe Ser
1130 1135 1140

Met Ala His Thr Val Ala Leu Asp Leu Ala Ala Ile Asn Ile Gln
1145 1150 1155

Arg Gly Arg Asp His Gly Ile Pro Pro Tyr His Asp Tyr Arg Val
1160 1165 1170

Tyr Cys Asn Leu Ser Ala Ala His Thr Phe Glu Asp Leu Lys Asn
1175 1180 1185

Glu Ile Lys Asn Pro Glu Ile Arg Glu Lys Leu Lys Arg Leu Tyr
1190 1195 1200

Gly Ser Thr Leu Asn Ile Asp Leu Phe Pro Ala Leu Val Val Glu
1205 1210 1215

Asp Leu Val Pro Gly Ser Arg Leu Gly Pro Thr Leu Met Cys Leu
1220 1225 1230

Leu Ser Thr Gln Phe Lys Arg Leu Arg Asp Gly Asp Arg Leu Trp
1235 1240 1245

Tyr Glu Asn Pro Gly Val Phe Ser Pro Ala Gln Leu Thr Gln Ile
1250 1255 1260

Lys Gln Thr Ser Leu Ala Arg Ile Leu Cys Asp Asn Ala Asp Asn
1265 1270 1275

Ile Thr Arg Val Gln Ser Asp Val Phe Arg Val Ala Glu Phe Pro
1280 1285 1290

His Gly Tyr Gly Ser Cys Asp Glu Ile Pro Arg Val Asp Leu Arg
1295 1300 1305

Val Trp Gln Asp Cys Cys Glu Asp Cys Arg Thr Arg Gly Gln Phe
Asn Ala Phe Ser Tyr His Phe Arg Gly Arg Arg Ser Leu Glu Phe
1310 1315 1320
1325 1330 1335

Ser Tyr Gln Glu Asp Lys Pro Thr Lys Lys Thr Arg Pro Arg Lys
1340 1345 1350

Ile Pro Ser Val Gly Arg Gln Gly Glu His Leu Ser Asn Ser Thr
1355 1360 1365

Ser Ala Phe Ser Thr Arg Ser Asp Ala Ser Gly Thr Asn Asp Phe
1370 1375 1380

Arg Glu Phe Val Leu Glu Met Gln Lys Thr Ile Thr Asp Leu Arg
1385 1390 1395

Thr Gln Ile Lys Lys Leu Glu Ser Arg Leu Ser Thr Thr Glu Cys
1400 1405 1410

Val Asp Ala Gly Gly Glu Ser His Ala Asn Asn Thr Lys Trp Lys
1415 1420 1425

Lys Asp Ala Cys Thr Ile Cys Glu Cys Lys Asp Gly Gln Val Thr
1430 1435 1440

Cys Phe Val Glu Ala Cys Pro Pro Ala Thr Cys Ala Val Pro Val
1445 1450 1455

Asn Ile Pro Gly Ala Cys Cys Pro Val Cys Leu Gln Lys Arg Ala
1460 1465 1470

Glu Glu Lys Pro
1475

<210> 21
<211> 798
<212> PRT
<213> Homo sapiens
<400> 21

Met Leu Ile Asn Cys Glu Ala Lys Gly Ile Lys Met Val Ser Glu Ile
1  5  10  15

Ser Val Pro Pro Ser Arg Pro Phe Gln Leu Ser Leu Leu Asn Asn Gly
20  25  30

44/56
<table>
<thead>
<tr>
<th>Residue</th>
<th>Tyr</th>
<th>Ile</th>
<th>Thr</th>
<th>Lys</th>
<th>Pro</th>
<th>Ser</th>
<th>Thr</th>
<th>Gln</th>
<th>Leu</th>
<th>Pro</th>
<th>Gly</th>
<th>Pro</th>
<th>Tyr</th>
<th>Cys</th>
<th>Pro</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>275</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Pro</th>
<th>Cys</th>
<th>Asn</th>
<th>Cys</th>
<th>Lys</th>
<th>Val</th>
<th>Leu</th>
<th>Ser</th>
<th>Pro</th>
<th>Ser</th>
<th>Gly</th>
<th>Leu</th>
<th>Leu</th>
<th>Ile</th>
<th>His</th>
<th>Cys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>290</td>
<td>295</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Gln</th>
<th>Glu</th>
<th>Arg</th>
<th>Asn</th>
<th>Ile</th>
<th>Glu</th>
<th>Ser</th>
<th>Leu</th>
<th>Ser</th>
<th>Asp</th>
<th>Leu</th>
<th>Arg</th>
<th>Pro</th>
<th>Pro</th>
<th>Pro</th>
<th>Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>305</td>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Asn</th>
<th>Pro</th>
<th>Arg</th>
<th>Lys</th>
<th>Leu</th>
<th>Ile</th>
<th>Leu</th>
<th>Ala</th>
<th>Gly</th>
<th>Asn</th>
<th>Ile</th>
<th>Ile</th>
<th>His</th>
<th>Ser</th>
<th>Leu</th>
<th>Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>325</td>
<td>330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Lys</th>
<th>Ser</th>
<th>Asp</th>
<th>Leu</th>
<th>Val</th>
<th>Glu</th>
<th>Tyr</th>
<th>Phe</th>
<th>Thr</th>
<th>Leu</th>
<th>Glu</th>
<th>Met</th>
<th>Leu</th>
<th>His</th>
<th>Leu</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>340</td>
<td>345</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Asn</th>
<th>Asn</th>
<th>Arg</th>
<th>Ile</th>
<th>Glu</th>
<th>Val</th>
<th>Leu</th>
<th>Glu</th>
<th>Glu</th>
<th>Gly</th>
<th>Ser</th>
<th>Phe</th>
<th>Met</th>
<th>Asn</th>
<th>Leu</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>355</td>
<td>360</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Arg</th>
<th>Leu</th>
<th>Gln</th>
<th>Lys</th>
<th>Leu</th>
<th>Tyr</th>
<th>Leu</th>
<th>Asn</th>
<th>Gly</th>
<th>Asn</th>
<th>His</th>
<th>Thr</th>
<th>Lys</th>
<th>Leu</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>370</td>
<td>375</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Lys</th>
<th>Gly</th>
<th>Met</th>
<th>Phe</th>
<th>Leu</th>
<th>Gly</th>
<th>Leu</th>
<th>His</th>
<th>Asn</th>
<th>Leu</th>
<th>Glu</th>
<th>Tyr</th>
<th>Leu</th>
<th>Tyr</th>
<th>Leu</th>
<th>Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>385</td>
<td>390</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Tyr</th>
<th>Asn</th>
<th>Ala</th>
<th>Ile</th>
<th>Lys</th>
<th>Glu</th>
<th>Ile</th>
<th>Leu</th>
<th>Pro</th>
<th>Gly</th>
<th>Thr</th>
<th>Phe</th>
<th>Asn</th>
<th>Pro</th>
<th>Met</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>405</td>
<td>410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Lys</th>
<th>Leu</th>
<th>Lys</th>
<th>Val</th>
<th>Leu</th>
<th>Tyr</th>
<th>Leu</th>
<th>Asn</th>
<th>Asn</th>
<th>Leu</th>
<th>Leu</th>
<th>Gln</th>
<th>Val</th>
<th>Leu</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>420</td>
<td>425</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Pro</th>
<th>His</th>
<th>Phe</th>
<th>Ser</th>
<th>Gly</th>
<th>Val</th>
<th>Pro</th>
<th>Leu</th>
<th>Thr</th>
<th>Lys</th>
<th>Val</th>
<th>Asn</th>
<th>Leu</th>
<th>Lys</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>435</td>
<td>440</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Asn</th>
<th>Gln</th>
<th>Phe</th>
<th>Thr</th>
<th>His</th>
<th>Leu</th>
<th>Pro</th>
<th>Val</th>
<th>Ser</th>
<th>Asn</th>
<th>Ile</th>
<th>Leu</th>
<th>Asp</th>
<th>Asp</th>
<th>Leu</th>
<th>Asp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>450</td>
<td>455</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Leu</th>
<th>Leu</th>
<th>Thr</th>
<th>Gln</th>
<th>Ile</th>
<th>Asp</th>
<th>Leu</th>
<th>Glu</th>
<th>Asp</th>
<th>Pro</th>
<th>Trp</th>
<th>Asp</th>
<th>Cys</th>
<th>Ser</th>
<th>Cys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>465</td>
<td>470</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Asp</th>
<th>Leu</th>
<th>Val</th>
<th>Gly</th>
<th>Leu</th>
<th>Gln</th>
<th>Gln</th>
<th>Trp</th>
<th>Ile</th>
<th>Gln</th>
<th>Lys</th>
<th>Leu</th>
<th>Ser</th>
<th>Lys</th>
<th>Asn</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>485</td>
<td>490</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Val</th>
<th>Thr</th>
<th>Asp</th>
<th>Leu</th>
<th>Cys</th>
<th>Thr</th>
<th>Ser</th>
<th>Pro</th>
<th>Gly</th>
<th>His</th>
<th>Leu</th>
<th>Asp</th>
<th>Lys</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>500</td>
<td>505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

46/56
Glu Leu Lys Ala Leu Asn Ser Glu Ile Leu Cys Pro Gly Leu Val Asn
515 520 525
Asn Pro Ser Met Pro Thr Gln Thr Ser Tyr Leu Met Val Thr Thr Pro
530 535 540
Ala Thr Thr Thr Asn Thr Ala Asp Thr Ile Leu Arg Ser Leu Thr Asp
545 550 555 560
Ala Val Pro Leu Ser Val Leu Ile Leu Gly Leu Leu Ile Met Phe Ile
565 570 575
Thr Ile Val Phe Cys Ala Ala Gly Ile Val Val Leu Val Leu His Arg
580 585 590
Arg Arg Arg Tyr Lys Lys Lys Gln Val Asp Glu Gln Met Arg Asp Asn
595 600 605
Ser Pro Val His Leu Gln Tyr Ser Met Tyr Gly His Lys Thr Thr His
610 615 620
His Thr Thr Glu Arg Pro Ser Ala Ser Leu Tyr Glu Gln His Met Val
625 630 635 640
Ser Pro Met Val His Val Tyr Arg Ser Pro Ser Phe Gly Pro Lys His
645 650 655
Leu Glu Glu Glu Glu Glu Asn Glu Lys Gly Ser Asp Ala Lys
660 665 670
His Leu Gln Arg Ser Leu Leu Gln Glu Asn His Ser Pro Leu Thr
675 680 685
Gly Ser Asn Met Lys Tyr Lys Thr Thr Asn Gln Ser Thr Glu Phe Leu
690 695 700
Ser Phe Gln Asp Ala Ser Ser Leu Tyr Arg Asn Ile Leu Glu Lys Glu
705 710 715 720
Arg Glu Leu Gln Gln Leu Gly Ile Thr Glu Tyr Leu Arg Lys Asn Ile
725 730 735
Ala Gln Leu Gln Pro Asp Met Glu Ala His Tyr Pro Gly Ala His Glu
740 745 750
| Glu Leu Lys Leu Met Glu Thr Leu Met Tyr Ser Arg Pro Arg Lys Val |
|----------------------|-----------------|-----------------|-----------------|
| 755                  | 760             | 765             |
| Leu Val Glu Gln Thr Lys Asn Glu Tyr Phe Glu Leu Lys Ala Asn Leu |
| 770                  | 775             | 780             |
| His Ala Glu Pro Asp Tyr Leu Glu Val Leu Glu Gln Gln Thr |
| 785                  | 790             | 795             |

<210> 22
<211> 713
<212> PRT
<213> Homo sapiens

<400> 22

| Met Arg Leu Leu Val Ala Pro Leu Leu Ala Trp Val Ala Gly Ala |
|----------------------|-----------------|-----------------|
| 1                    | 5               | 10              | 15              |
| Thr Ala Ala Val Pro Val Val Pro Trp His Val Pro Cys Pro Pro Gln |
| 20                   | 25              | 30              |
| Cys Ala Cys Gln Ile Arg Pro Trp Tyr Thr Pro Arg Ser Ser Tyr Arg |
| 35                   | 40              | 45              |
| Glu Ala Thr Thr Val Asp Cys Asn Asp Leu Phe Leu Thr Ala Val Pro |
| 50                   | 55              | 60              |
| Pro Ala Leu Pro Ala Gly Thr Gln Thr Leu Leu Leu Gln Ser Asn Ser |
| 65                   | 70              | 75              | 80              |
| Ile Val Arg Val Asp Gln Ser Glu Leu Gly Tyr Leu Ala Asn Leu Thr |
| 85                   | 90              | 95              |
| Glu Leu Asp Leu Ser Gln Asn Ser Phe Ser Asp Ala Arg Asp Cys Asp |
| 100                  | 105             | 110             |
| Phe His Ala Leu Pro Gln Leu Leu Ser Leu His Leu Glu Glu Asn Gln |
| 115                  | 120             | 125             |
| Leu Thr Arg Leu Glu Asp His Ser Phe Ala Gly Leu Ala Ser Leu Gln |
| 130                  | 135             | 140             |
| Glu Leu Tyr Leu Asn His Asn Gln Leu Tyr Arg Ile Ala Pro Arg Ala |
| 145                  | 150             | 155             | 160             |

48/56
Phe Ser Gly Leu Ser Asn Leu Leu Arg Leu His Leu Asn Ser Asn Leu
165 170 175
Leu Arg Ala Ile Asp Ser Arg Trp Phe Glu Met Leu Pro Asn Leu Glu
180 185 190
Ile Leu Met Ile Gly Gly Asn Lys Val Asp Ala Ile Leu Asp Met Asn
195 200 205
Phe Arg Pro Leu Ala Asn Leu Arg Ser Leu Val Leu Ala Gly Met Asn
210 215 220
Leu Arg Glu Ile Ser Asp Tyr Ala Leu Glu Gly Leu Gln Ser Leu Glu
225 230 235 240
Ser Leu Ser Phe Tyr Asp Asn Gln Leu Ala Arg Val Pro Arg Arg Ala
245 250 255
Leu Glu Gln Val Pro Gly Leu Lys Phe Leu Asp Leu Asn Lys Asn Pro
260 265 270
Leu Gln Arg Val Gly Pro Gly Asp Phe Ala Asn Met Leu His Leu Lys
275 280 285
Glu Leu Gly Leu Asn Asn Met Glu Glu Leu Val Ser Ile Asp Lys Phe
290 295 300
Ala Leu Val Asn Leu Pro Glu Leu Thr Lys Leu Asp Ile Thr Asn Asn
305 310 315 320
Pro Arg Leu Ser Phe Ile His Pro Arg Ala Phe His His Leu Pro Gln
325 330 335
Met Glu Thr Leu Met Leu Asn Asn Ala Leu Ser Ala Leu His Gln
340 345 350
Gln Thr Val Glu Ser Leu Pro Asn Leu Gln Glu Val Gly Leu His Gly
355 360 365
Asn Pro Ile Arg Cys Asp Cys Val Ile Arg Trp Ala Asn Ala Thr Gly
370 375 380
Thr Arg Val Arg Phe Ile Glu Pro Gln Ser Thr Leu Cys Ala Glu Pro
385 390 395 400
Pro Asp Leu Gln Arg Leu Pro Val Arg Glu Val Pro Phe Arg Glu Met
405 410 415

Thr Asp His Cys Leu Pro Leu Ile Ser Pro Arg Ser Phe Pro Pro Ser
420 425 430

Leu Gln Val Ala Ser Gly Glu Ser Met Val Leu His Cys Arg Ala Leu
435 440 445

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

Leu Thr Pro Ala His Ala Gly Arg Arg Cys Arg Val Tyr Pro Glu Gly
465 470 475 480

Thr Leu Glu Leu Arg Arg Val Thr Ala Glu Glu Ala Gly Leu Tyr Thr
485 490 495

Cys Val Ala Gln Asn Leu Val Gly Ala Asp Thr Lys Thr Val Ser Val
500 505 510

Val Val Gly Arg Ala Leu Leu Gln Pro Gly Arg Asp Gly Glu Gln Gly
515 520 525

Leu Glu Leu Arg Val Gln Glu Thr His Pro Tyr His Ile Leu Leu Ser
530 535 540

Trp Val Thr Pro Pro Asn Thr Val Ser Thr Asn Leu Thr Trp Ser Ser
545 550 555 560

Ala Ser Ser Leu Arg Gly Gln Gly Ala Thr Ala Leu Ala Arg Leu Pro
565 570 575

Arg Gly Thr His Ser Tyr Asn Ile Thr Arg Leu Leu Gln Ala Thr Glu
580 585 590

Tyr Trp Ala Cys Leu Gln Val Ala Phe Ala Asp Ala His Thr Gln Leu
595 600 605

Ala Cys Val Trp Ala Arg Thr Lys Glu Ala Thr Ser Cys His Arg Ala
610 615 620

Leu Gly Asp Arg Pro Gly Leu Ile Ala Ile Leu Ala Leu Ala Val Leu
625 630 635 640
Leu Leu Ala Ala Gly Leu Ala Ala His Leu Gly Thr Gly Gln Pro Arg
645 650 655

Lys Gly Val Gly Gly Arg Arg Pro Leu Pro Pro Ala Trp Ala Phe Trp
660 665 670

Gly Trp Ser Ala Pro Ser Val Arg Val Val Ser Ala Pro Leu Val Leu
675 680 685

Pro Trp Asn Pro Gly Arg Lys Leu Pro Arg Ser Ser Glu Gly Glu Thr
690 695 700

Leu Leu Pro Pro Leu Ser Gln Asn Ser
705 710

<210> 23
<211> 684
<212> PRT
<213> Homo sapiens
<400> 23

Met Thr Ser Leu Val His Leu Thr Leu Ser Arg Asn Thr Ile Gly Gln
1   5  10  15

Val Ala Ala Gly Ala Phe Ala Asp Leu Arg Ala Leu Arg Ala Leu His
20  25  30

Leu Asp Ser Asn Arg Leu Ala Glu Val Arg Gly Asp Gln Leu Arg Gly
35  40  45

Leu Gly Asn Leu Arg His Leu Ile Leu Gly Asn Asn Gln Ile Arg Arg
50  55  60

Val Glu Ser Ala Ala Phe Asp Ala Phe Leu Ser Thr Val Glu Asp Leu
65  70  75  80

Asp Leu Ser Tyr Asn Asn Leu Glu Ala Leu Pro Trp Glu Ala Val Gly
85  90  95

Gln Met Val Asn Leu Asn Thr Leu Thr Leu Asp His Asn Leu Ile Asp
100 105 110

His Ile Ala Glu Gly Thr Phe Val Gln Leu His Lys Leu Val Arg Leu
115 120 125
Asp Met Thr Ser Asn Arg Leu His Lys Leu Pro Pro Asp Gly Leu Phe
130 135 140
Leu Arg Ser Gln Gly Thr Gly Pro Lys Pro Pro Thr Pro Leu Thr Val
145 150 155 160
Ser Phe Gly Gly Asn Pro Leu His Cys Asn Cys Glu Leu Leu Trp Leu
165 170 175
Arg Arg Leu Thr Arg Glu Asp Asp Leu Glu Thr Cys Ala Thr Pro Glu
180 185 190
His Leu Thr Asp Arg Tyr Phe Trp Ser Ile Pro Glu Glu Phe Leu
195 200 205
Cys Glu Pro Pro Leu Ile Thr Arg Gln Ala Gly Gly Arg Ala Leu Val
210 215 220
Val Glu Gly Gln Ala Val Ser Leu Arg Cys Arg Ala Val Gly Asp Pro
225 230 235 240
Glu Pro Val Val His Trp Val Ala Pro Asp Gly Arg Leu Leu Gly Asn
245 250 255
Ser Ser Arg Thr Arg Val Arg Gly Asp Gly Thr Leu Asp Val Thr Ile
260 265 270
Thr Thr Leu Arg Asp Ser Gly Thr Phe Thr Cys Ile Ala Ser Asn Ala
275 280 285
 Ala Gly Glu Ala Thr Ala Pro Val Glu Val Val Val Pro Leu Pro
290 295 300
Leu Met Ala Pro Pro Ala Ala Pro Pro Pro Leu Thr Glu Pro Gly
305 310 315 320
Ser Ser Asp Ile Ala Thr Pro Gly Arg Pro Gly Ala Asn Asp Ser Ala
325 330 335
 Ala Glu Arg Arg Leu Val Ala Ala Glu Leu Thr Ser Asn Ser Val Leu
340 345 350
Ile Arg Trp Pro Ala Gln Arg Pro Val Pro Gly Ile Arg Met Tyr Gln
355 360 365
Val Gln Tyr Asn Ser Ser Val Asp Asp Ser Leu Val Tyr Arg Met Ile 370 375 380

Pro Ser Thr Ser Glu Thr Phe Leu Val Asn Asp Leu Ala Ala Gly Arg 385 390 395 400

Ala Tyr Asp Leu Cys Val Leu Ala Val Tyr Asp Gly Ala Thr Ala 405 410 415

Leu Pro Ala Thr Arg Val Val Gly Cys Val Glu Phe Thr Thr Ala Gly 420 425 430

Asp Pro Ala Pro Cys Arg Pro Leu Arg Ala His Phe Leu Gly Gly Thr 435 440 445

Met Ile Ile Ala Ile Gly Gly Val Ile Val Ala Ser Val Leu Val Phe 450 455 460

Ile Val Leu Leu Met Ile Arg Tyr Lys Val Tyr Gly Asp Gly Asp Ser 465 470 475 480

Arg Arg Val Lys Gly Ser Arg Ser Leu Pro Arg Val Ser His Val Cys 485 490 495

Ser Gln Thr Asn Gly Ala Gly Thr Gly Ala Ala Gln Ala Pro Ala Leu 500 505 510

Pro Ala Gln Asp His Tyr Glu Ala Leu Arg Glu Val Glu Ser Gln Ala 515 520 525

Ala Pro Ala Val Ala Val Glu Ala Lys Ala Met Glu Ala Glu Thr Ala 530 535 540

Ser Ala Glu Pro Glu Val Val Leu Gly Arg Ser Leu Gly Gly Ser Ala 545 550 555 560

Thr Ser Leu Cys Leu Leu Pro Ser Glu Glu Thr Ser Gly Glu Ser 565 570 575

Arg Ala Ala Val Gly Pro Arg Arg Ser Arg Ser Gly Ala Leu Glu Pro 580 585 590

Pro Thr Ser Ala Pro Pro Thr Leu Ala Leu Val Pro Gly Gly Ala Ala 595 600 605

53/56
Met His Asn Ile Arg Asp Ala Cys Arg Asp His Met Glu Gly Tyr His

Tyr Arg Tyr Glu Ile Asn Ala Asp Pro Arg Leu Thr Asn Leu Ser Ser

Asn Ser Asp Val