SECRETED PROTEIN HFEAF41

Inventors: Paul Young, Gaithersburg, MD (US); John M. Greene, Gaithersburg, MD (US); Ann M. Ferrie, Tewksbury, MA (US); Steven M. Ruben, Olney, MD (US); Craig A. Rosen, Laytonsville, MD (US); Roxanne Duan, Bethesda, MD (US); Jing-Shan Hu, Sunnyvale, CA (US); Kimberly Florence, Rockville, MD (US); Henrik S. Olsen, Gaithersburg, MD (US); Reinhard Ebner, Gaithersburg, MD (US); Laurie A. Brewer, St. Paul, MN (US); Paul A. Moore, Germantown, MD (US); Yanggu Shi, Gaithersburg, MD (US); David W. Laffleur, Washington, DC (US); Jian Ni, Rockville, MD (US)

Correspondence Address:
Human Genome Sciences, Inc.
9410 Key West Avenue
Rockville, MD 20850 (US)

Applied No.: 09/983,966
Filed: Oct. 26, 2001

Publication Classification

International Cl. 7 ............................. C07H 21/04; C07K 16/46
U.S. Cl. ............... 536/23,53; 530/388.15; 530/391.1

ABSTRACT

The present invention relates to 87 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.
SECRETED PROTEIN HFEAF41

[0001] This application is a continuation-in-part of, and claims benefit under 35 U.S.C. §120 of copending United States patent application Serial No: PCT/US98/053 11, filed Mar. 19, 1998, which is hereby incorporated by reference, which claims benefit under 35 U.S.C. §119(e) based on U.S. Provisional Applications:

<table>
<thead>
<tr>
<th>Filing Date</th>
<th>Appln No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar 21, 1997</td>
<td>60/041,277</td>
</tr>
<tr>
<td>Mar 21, 1997</td>
<td>60/042,344</td>
</tr>
<tr>
<td>Mar 21, 1997</td>
<td>60/041,276</td>
</tr>
<tr>
<td>Mar 21, 1997</td>
<td>60/041,261</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,094</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,350</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,188</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,135</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/050,937</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,187</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,099</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,352</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,186</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,095</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,131</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,096</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,355</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,160</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,351</td>
</tr>
<tr>
<td>May 30, 1997</td>
<td>60/048,154</td>
</tr>
<tr>
<td>Aug 05, 1997</td>
<td>60/054,804</td>
</tr>
<tr>
<td>Aug 19, 1997</td>
<td>60/056,370</td>
</tr>
<tr>
<td>Oct 02, 1997</td>
<td>60/060,862</td>
</tr>
</tbody>
</table>

FIELD OF THE INVENTION

[0002] This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

BACKGROUND OF THE INVENTION

[0003] Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses “sorting signals,” which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

[0004] One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of protein to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, the proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

[0005] Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space—a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a “linker” holding the protein to the membrane.

[0006] Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

SUMMARY OF THE INVENTION

[0007] The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

DETAILED DESCRIPTION

Definitions

[0008] The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

[0009] In the present invention, “isolated” refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered “by the hand of man” from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be “isolated” because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

[0010] In the present invention, a “secreted” protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a “mature” protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

[0011] As used herein, a “polynucleotide” refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain
the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a “polypeptide” refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

[0012] In the present invention, the full length sequence identified as SEQ ID NO: X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO: X was deposited with the American Type Culture Collection (“ATCC”). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Va. 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

[0013] A “polynucleotide” of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO: X, the complement thereof, or the cDNA within the clone deposited with the ATCC. “Stringent hybridization conditions” refers to an overnight incubation at 42° C. in a solution comprising 50% formamide, 5×SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5×Denhardt’s solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1×SSC at about 65° C.

[0014] Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37° C. in a solution comprising 6×SSPE (20×SSPE 3 M NaCl; 0.2 M NaH2PO4; 0.02 M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 μg/ml salmon sperm blocking DNA; followed by washes at 50° C. with 1×SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5×SSC).

[0015] Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt’s reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[0016] Of course, a polynucleotide which hybridizes only to polyA sequences (such as any 3’ terminal polyA tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of “polynucleotide,” since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

[0017] The polynucleotide of the present invention can be composed of any polynucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. “Modified” bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, “polynucleotide” embraces chemically, enzymatically, or metabolically modified forms.

[0018] The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclical, with or without branching. Cyclical, branched, and branched cyclical polypeptides may result from posttranslational natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyrogallate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenylation, sullation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS—STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T.E. Creighton, W.H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B.C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifert et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)
“SEQ ID NO:X” refers to a polynucleotide sequence while “SEQ ID NO:Y” refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

“A polypeptide having biological activity” refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

Features of Protein Encoded by Gene No:1

The translation product of this gene shares sequence homology with nucleolin, which is thought to be important in macromolecule binding, as well as a number of membrane proteins. Preferred polypeptide fragments comprise the amino acid sequence:

-DPEAEDSGEPONKSTPDPEEEYY-
-KEEQNEEAVKMLVEATREFEEVVVDSES (SEQ ID NO:231);
-OQLKLKKAED-DPEAEDSGEPONKSTPDPEEEYYKEEQNEEAVKMLVEATREFEEVVVDSES (SEQ ID NO:232);
-KAMEKLQISTGISWLSLDDRLKHLRIQOE-HKYLGGAPVSSORLKKRAEDEP-EAEDSGEPONKSTPDPEEEYY-
-KEEQNEEAVKMLVEATREFEEVVVDSES (SEQ ID NO:233);

Also preferred are the polynucleotide fragments encoding these polypeptide fragments. This gene maps to chromosome 16, and therefore can be used as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in brain and kidney and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cell-cell interaction or cell-matrix interaction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and kidney, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and other tissue of the nervous system, and kidney, and cancers and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.121 as residues: Met-1 to Trp-10.

The tissue distribution in brain and kidney combined with the homology to nucleolin indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of diseases involving cell-cell interaction or cell-extracellular matrix interaction. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1665 of SEQ ID NO:11, b is an integer of 15 to 1679, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:2

The translation product of this gene shares sequence homology with a porcine zona pellucida protein ZPDS.1711. (See Accession No. R39356.) These two proteins have weak homology with Drosophila commissureless and metalloestrase proteins which are thought to be important in controlling growth cone guidance across the CNS midline and protecting cells against reactive oxygen toxicity. Thus, based on homology, it is likely that this gene may also be involved in development. Preferred polypeptide fragments comprise the amino acid sequence: LPSYDE-AERTKAEOATPLPVGREDDEF VGRDDFDDADQL-RIGNDGFMLTFMAFLWFWGFLSFC-LTTSAAATGTASGG FGLSLIKWILVRVSTFYPGFYGIDQYWL-WWVLYLGFLLRFOHNYKVRKMPETNSL- PIRIRVFLF (SEQ ID NO:234); and/or AGRYGAISGFLGSLIKWILVRFES (SEQ ID NO:235). Also preferred are polynucleotide fragments encoding these polypeptide fragments. The gene that encodes the disclosed cDNA is thought to reside on Chromosome 5. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 5.

This gene is expressed primarily in kidney, adrenal gland, brain, fetal and reproductive tissues, and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, fertilization control or tissue damage by metabolites or other toxic agents. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, urogenital or renal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. reproductive, kidney, adrenal gland, and brain and other tissue of the nervous
system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0029] The tissue distribution in reproductive tissues combined with the homology to zona pellucida protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for fertility control such as contraceptive development. The homology with metal homeostasis and commissureless genes indicates the gene’s function in spermatozoa guidance and protection. It would also be useful for the treatment/diagnosis of tissue damages caused by toxic metabolites and other agents since the gene product is also expressed in urosecretory tissues. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 1949 of SEQ ID NO:12, b is an integer of 15 to 1963, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:3

[0030] This gene is expressed primarily in liver and to a lesser extent in placenta. Preferred polypeptide fragments comprise the amino acid sequence:

MKIIKSA9FVT3LTQQLWEFEGSVEVC-QTLTVSSKL0R4YTFFS0GFTFY (SEQ ID NO:236).

[0031] Also preferred are polynucleotide fragments encoding these polypeptide fragments.

[0032] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, digestive, metabolic, developmental, and nutrient transport/utilization disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and circulatory system, expression of this gene at significantly lower or higher levels may be routinely detected in certain tissues or cell types (e.g., liver, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, bile, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0033] The tissue distribution in liver and placenta indicates that the protein product is either an extracellular enzyme or a molecular carrier. Therefore, polynucleotides and polypeptides corresponding to this gene are useful for diagnosis/treatment of digestive and nutrient transport/utilization disorders, including malabsorption and malnutrition. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1198 of SEQ ID NO:13, b is an integer of 15 to 1212, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:4

[0034] This gene shares homology with the sap47 gene of Drosophila melanogaster, a gene which codes for a conserved neuronal protein associated with synaptic terminals. (See Mol. Brain Res. 32:45-54 (1995); see also, Accession No. 929571.) Thus, based on homology, the gene of the present invention also should be associated with synaptic terminals. Preferred polypeptide fragments comprise the amino acid sequence:
Also preferred are polynucleotide fragments encoding these polypeptide fragments. Contact of cells with supernatant expressing the product of this gene increases the permeability of the plasma membrane of aortic smooth muscle cells to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the aortic smooth muscle cells. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating aortic smooth muscle cells.

This gene is expressed primarily in kidney pyramids and to a lesser extent in lung and other tissues of various types. This gene fluxes calcium in human aortic smooth muscle cells, and therefore is involved in signal transduction.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, renal, developmental, vascular, and nervous disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney and/or system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., kidney, lung, brain and other tissue of the nervous system, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney and lung and homology with sap47 indicates that the protein product has regulatory or direct functions in molecular exchange with body fluids and nervous system signaling. Polynucleotides and polypeptides corresponding to this gene are useful for treatment of disorders in kidney and nervous system. The activity of the translation product of this gene in activating aortic smooth muscle cells supports the notion that this protein is involved in regulatory or direct functions in molecular exchange with body fluids. This clone would be useful for the diagnosis and treatment of disorders in kidney and the nervous system. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2047 of SEQ ID NO:14, b is an integer of 15 to 2061, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:5

The translation product of this gene shares sequence homology with the mouse Ly-9.2 antigen which is thought to be an important cell surface marker in lymphoids, myeloids and hematopoietic progenitors. (See Accession No. gi|198932.) Preferred polypeptide fragments comprise the amino acid sequence:

```
PPCVARPNVSNESSPLARKICGAA (SEQ ID NO:243);
```

and/or

```
KEDPANYYSTVEIPKKMENPSHLLTMDTFLR (SEQ ID NO:244).
```

Also preferred are polynucleotide fragments encoding these polypeptide fragments. Based on homology, it is likely that this gene is also a cell surface marker, involved in hematopoiesis.

This gene is expressed primarily in activated macrophages, monocytes and T-cells and to a lesser extent in spleen and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, blood cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:125 as residues: Lys-26 to Tyr-33, Arg-44 to Ile-49, Ser-53 to Lys-71, Lys-86 to Pro-91.

The tissue distribution in immune tissue combined with the homology to a protein within the Ly-9.2 surface immunoglobulin family indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of immune and hematopoietic disorders. Polypeptides
and polynucleotides corresponding to this gene are also be used as a marker for leukemia or a modulator of the functions of the cells of macrophage/monocyte or T-cell types. Expression of this gene product in immune cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1398 of SEQ ID NO:15, b is an integer of 15 to 1412, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:6

[0045] The translation product of this gene shares sequence homology with the Drosophila glutactin gene which is thought to be important in cell-cell interaction or cell-extracellular matrix contact. The gene encoding the disclosed cDNA is thought to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

[0046] This gene is expressed primarily in colon tissue, aorta endothelial cells and to a lesser extent in skin, breast tissue and T-cells.

[0047] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of these tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the gastrointestinal tract, vascular system or T-cell development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, cardiovascular system, and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., colon, endothelial, cardiovascular tissue, skin, mammary tissue, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0048] The tissue distribution and homology to glutactin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the development and maintenance of the integrity of the basal membrane in the gastrointestinal tract, or vasculature in the cardiovascular system. The expression in T-cells also indicates the protein may be involved in T-cell adhesion, cell-cell interaction and development. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1058 of SEQ ID NO:16, b is an integer of 15 to 1052, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:7

[0049] The translation product of this gene shares sequence homology with MURF4 protein, an ATPase homolog, which is thought to be important in ATP hydrolysis.

[0050] This gene is expressed primarily in breast tissue.

[0051] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer and non-neoplastic breast diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., breast milk, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0052] The tissue distribution in breast tissue combined with the homology to the MURF4 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neoplastic or non-neoplastic breast diseases because ATPase like protein
may be involved in changed metabolic states of the breast. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO: 17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 669 of SEQ ID NO: 17, b is an integer of 15 to 683, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO: 17, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 8

[0053] This gene shares homology to the alcohol dehydrogenase gene. Preferred polypeptide fragments comprise comprise the amino acid sequence: ASAVLLDLPNSG

[0054] Polynucleotides encoding these fragments are also encompassed by the invention. Other groups have also recently cloned this gene, recognizing its homology to alcohol dehydrogenase. (See Accession No. 1778355.) Moreover, a second group recently cloned the mouse homologue of this gene. (See Accession No. 2078284.) They found that the mouse homologue binds to amyloid beta-peptide and mediates neurotoxicity in Alzheimer’s disease, calling the protein ERAB. This gene maps to chromosome X, and therefore can be used in linkage analysis as a marker for chromosome X. Therefore, mutations in the translated product of this gene may be involved in Alzheimer’s disease in humans, as well as other sex linked diseases. This gene can be used as a diagnostic marker for these diseases.

[0055] It has been discovered that this gene is expressed primarily in breast cancer tissue, infant brain, and to a lesser extent in fetal liver tissue.

[0056] Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neurodegenerative diseases, breast cancer, non-neoplastic breast diseases, or developmental disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunologic probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and CNS, and breast tissue, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., brain, breast, metabolic, developmental, immune, hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 128 as residues: Arg-45 to Ser-53.

[0057] The tissue distribution in neural tissue combined with the homology to the ERAB mouse gene suggests that the protein product of this clone would be useful for the diagnosis and treatment of Alzheimers and related neurodegenerative diseases. Mutations in the translated product of this gene may be involved in Alzheimer’s disease in humans, as well as other sex linked diseases. This gene can be used as a diagnostic marker for these diseases. Furthermore, the tissue distribution suggests that this gene may also be involved in neoplastic or non-neoplastic breast diseases in humans. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO: 18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1040 of SEQ ID NO: 18, b is an integer of 15 to 1054, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO: 18, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 9

[0058] The translation product of this gene shares week sequence homology with rat N-methyl-D-aspartate receptor subunit and other proline-rich proteins which are thought to be important in neurotransmission or protein-protein interaction.

[0059] This gene is expressed primarily in synovial hypoxia and to a lesser extent in ovary, senescent cells and brain.

[0060] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include,
but are not limited to, synovial hypoxia, reproductive, or neural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovia and brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., synovial tissue, ovary and other reproductive tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression-level in healthy tissue or bodily fluid from an individual not having the disorder.

[0061] The tissue distribution in synovial hypoxia and nerve tissues, and homology to N-methyl-D-aspartate receptor subunits and other proline-rich proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of synovial hypoxia and other synovial disorders, particularly disorders involving nitric oxide signaling. Protein, as well as antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1379 of SEQ ID NO:19, b is an integer of 15 to 1393, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:10

[0062] This gene is expressed primarily in prostate and keratinocytes, and to a lesser extent in placenta, ovary and primary dendritic cells.

[0063] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, male and female infertility, cancer, skin disorders, and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, skin, and neoplasia, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, skin, placenta, ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:130 as residues: Pro-17 to Met-23, Ala-30 to Trp-38, Ile-49 to Trp-54, Lys-68 to Gly-74, Thr-93 to Gly-99, Met-126 to Glu-132, Gly-173 to Ser-178, Lys-205 to Tyr-214.

[0064] The tissue distribution of this gene in the prostate, placenta and ovary indicates that this gene product is useful for treatment/diagnosis of male or female infertility, endocrine disorders, fetal deficiencies, ovarian failure, amenorrhea, ovarian cancer, benign prostate hyperplasia, prostate cancer, and other forms of cancer of the reproductive system. The tissue distribution also suggests that the protein product of this clone would be useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen’s disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget’s disease, mycosis fungoides, and Kaposis’s sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, stria, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete’s foot, and ringworm). Protein, as well as antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1201 of SEQ ID NO:20, b is an integer of 15 to 1215, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:11

[0065] This gene is expressed primarily in the thyroid and to a lesser extent in the pineal gland. The gene encoding the disclosed cDNA is thought to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10. Preferred polypeptide fragments comprise the amino acid sequences:

HPIEWSAINAATLQPY (SEQ ID NO:248);  
CWYKCLTMQQAQLMDQNG (SEQ ID NO:249);
Also preferred are polynucleotides encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, thyroid and pineal gland disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, thyroid and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.132 as residues: Ser-2 to Ser-8, Thr-38 to Arg-44.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating/detecting immune disorders such as arthritis, asthma, immune deficiency diseases (e.g., AIDS), and leukemia, as well as treating/detecting thymus disorders (e.g., Graves Disease, lymphohcytoid thyroiditis, hyperthyroidism, and hypothyroidism), and treating/detecting pineal gland disorders (e.g., circadian rhythm disturbances associated with shift work, jet lag, blindness, insomnia and old age). Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of \( a \cdot b \), where \( a \) is any integer between \( 1 \) to \( 2028 \) of SEQ ID NO:21, \( b \) is an integer of \( 15 \) to \( 2042 \), where both \( a \) and \( b \) correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where the \( b \) is greater than or equal to \( a+0.14 \).

Features of Protein Encoded by Gene No:13

This gene is expressed primarily in progenitor cells (CD34 cells) of lymphoid, myeloid and erythroid cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, blood cells, myeloid cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0075] The predominant tissue distribution of this gene in hematopoietic cell types indicates that the gene could be important for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Preferred embodiments of the present invention are polypeptide fragments comprising the amino acid sequence:

FTHLSTCLLSLLLVRMSGFLLLARASPSICALDSSCFVEYCSSYSSSCFLHQHFPSLLDKLCQ (SEQ ID NO:253); or
FLLLARASPSICALDSSCFVQSY (SEQ ID NO:254).

[0076] Also preferred are polynucleotide fragments encoding these polypeptide fragments. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 275 of SEQ ID NO:23; b is an integer of 15 to 289, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23; and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:14

[0077] This gene is homologous to the Drosophila Regina (Rga) gene. (See Accession No. 1658504.) This Drosophila gene is thought to be a homolog of the global negative transcriptional regulator NOT2 (CDC36) from yeast, which modifies gene expression and suppresses position effect variegation. Preferred polypeptide fragments comprise the amino acid sequence:

PDGRVTNTGQVTQGFDGMIGLLTFIRAETFDGMVYL (SEQ ID NO:255); and/or
AGSDLTLTTGGLNLS

VHLGALGDLTTLGLNLEPENLVP (SEQ ID NO:257); or
EDLLFYLYNNGDVQILLLAVEELPFRDWRHKEERVWITR (SEQ ID NO:256); and/or
EDLLFYLYNNGDVQILLLAVEELPFRDWRHKEERVWITR (SEQ ID NO:259); or (SEQ ID NO:256); and/or
HNEDFPPALPS (SEQ ID NO:259).

[0078] This gene is expressed primarily in placenta and to a lesser extent in infant brain.

[0079] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurodegenerative and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., placenta, and brain and other tissue of the nervous system, reproductive, developmental tissues, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid,
publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1134 of SEQ ID NO:25, b is an integer of 15 to 1148, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:15

This gene is expressed primarily in adrenal gland tumor and osteoclastoma.

GRIIDTSLRFLPVLVIRLGOKVQVIPGLEQSLLDMSVGEKRRAIIPSH (SEQ ID NO:249); and/or LAVGKRGFPSPVPSUADVYGVDVIVLALIR (SEQ ID NO:261). Also preferred are the polynucleotide fragments encoding these polypeptides.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, endocrine and bone disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system and in bone, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adrenal gland, and bone, skeletal tissues, and cancers and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:135 as residues: Ilc-52 to Trp-57.

The tissue distribution of this gene in endocrine tissue indicates that it may be involved in the treatment and/or detection of adrenal gland tumors, osteosarcomas, endocrine disorders and bone disorders, particularly osteoporosis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b,

Features of Protein Encoded by Gene No:16

The translation product of this gene shares sequence homology with the FK506 binding protein, a protein which plays an important role in immunosuppression. (See Accession No. M75099.) Specifically, a 12-kDa FK506-binding protein (FKBP-12) is a cytosolic receptor for the immunosuppressants FK506 and rapamycin. (See, Proc. Natl. Acad. Sci. 88: 6677-6681 (1991).) Thus, based on homology, it is likely that this gene also has immunosuppression activity or may be involved in other activities related to calcium dependent regulation. Preferred polypeptides comprise the amino acid sequence:

GRIDTSLRFLPVLVIRLGOKVQVIPGLEQSLLDMSVGEKRRAIIPSH (SEQ ID NO:249); and/or LAVGKRGFPSPVPSUADVYGVDVIVLALIR (SEQ ID NO:261). Also preferred are the polynucleotide fragments encoding these polypeptides.

This gene is expressed primarily in melanocytes. Furthermore, northern analysis demonstrated that this gene is also abundant in fetal liver and kidney. In adult tissues, it is expressed relatively highly in spleen, placenta, and thymus, and at a low level in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental conditions, or cancer and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and cancer, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, melanocytes, developmental, intumetary, hepatic, renal, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, bile, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:136 as residues: Ala-118 to Phe-124, Arg-178 to Lys-201.

The tissue distribution in developing tissues combined with the homology to the FK506 binding proteins which are believed to a role in immunosuppression mediated by the immunosuppressant drugs rapamycin and cyclosporin, indicates that this gene could serve as a novel target for the identification of novel immunosuppressant drugs. Protein, as well as, antibodies directed against the protein may show...
utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 703 of SEQ ID NO:26, b is an integer of 15 to 717, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:17

[0089] The translation product of this gene shares sequence homology with the rat calcium-activated potassium channel rSK3, which is thought to be important in regulating vascular tone. (See Accession No. gill2564072, gill1575663, and gill1575661.) Although homologous to these proteins, this gene contains an 18 amino acid insert, not previously identified in the homologs. Preferred polypeptide fragments comprise the amino acid sequence:

CESPEPSAAPGSSLPAWHY (SEQ ID NO:262).

[0090] Also preferred are the polynucleotide fragments encoding these polypeptides.

[0091] This gene is expressed primarily in B-cells, frontal cortex and endothelial cells.

[0092] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, cardiovascular (hypertension, heart disease, restenosis, atherosclerosis, stroke, angina and thrombosis) or neurologic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, smooth muscle, vascular, and brain and other tissues of the nervous system, and endothelium, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO 138 as residues: Lys-43 to Arg-49, Tyr-58 to Glu-65.

[0093] The tissue distribution in endothelial cells combined with the homology to calcium-activated potassium channels indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of vascular disorders (hyper/hypotension, atherosclerosis, stroke, angina and thrombosis). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1085 of SEQ ID NO:27, b is an integer of 15 to 1099, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:18

[0094] This gene is expressed primarily in smooth muscle and hematopoietic cells and to a lesser extent in brain (amygdala, corpus colosum, hippocampus).

[0095] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cardiovascular (hypertension, heart disease, atherosclerosis, stroke, angina and thrombosis, and wound healing), immune, or neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and neurological systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, smooth muscle, vascular, and brain and other tissue of the nervous system, and endolheliun, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO 138 as residues: Lys-43 to Arg-49, Tyr-58 to Glu-65.

[0096] The tissue distribution in smooth muscle indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cardiovascular disorders (hypertension, heart disease, atherosclerosis, stroke, angina, thrombosis, and wound healing). Expression in brain indicates a role in the treatment and diagnosis of behavioral or neurological disorders, such as depression, schizophrenia, Alzheimer’s disease, mania, dementia, paranoia, and addictive behavior. Expression of this gene product in hematopoietic cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or
other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, as well as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b where a is any integer between 1 to 927 of SEQ ID NO:28, b is an integer of 15 to 941, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:19

[0097] This gene is expressed primarily in T-cells (Jurkats, resting, activated, and anergic T-cells), endothelial cells, pineal gland, and to a lesser extent in a variety of other tissues and cell types. Preferred polypeptide fragments comprise the amino acid sequence:

```
EEAGAGRRCSHGGARPAGLGNEGLGGDPDHTDTGSRSKQRINNWKESKHKVIMASASARGNQDKDAHFPPPSKQSLLFCPKSKLHIHRAEISKSKQRINNWKESKHKVIMAS

(SEQ ID NO:263); and/or

SKQRINNWKESKHKVIMAS

(SEQ ID NO:264).
```

[0098] Also preferred are the polynucleotide fragments encoding these polypeptides.

[0099] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to immune disorders, such as, inflammation, immunodeficiencies, or cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s) For a number of disorders of the above tissues or cells, particularly of the immune, neurological and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., T-cells and other blood cells, endothelial cells, and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Phe-71 to Arg-76, Pro-82 to His-87, Glu-103 to Ala-111.

[0100] The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immune, immuno-suppressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. In addition, expression in the pineal gland might suggest a role in the diagnosis specific brain tumors and treatment of neurological disorders. Endothelial cell expression might suggest a role in cardiovascular or respiratory/pulmonary disorders or infections (asthma, pulmonary edema, pneumonia). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, as well as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 742 of SEQ ID NO:29, b is an integer of 15 to 756, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:20

[0101] The gene encoding the disclosed cDNA is thought to reside on chromosome 15. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis form chromosome 15.

[0102] This gene is expressed primarily in brain and embryo and to a lesser extent in leukocytes.

[0103] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental, immune, and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or
cell types (e.g., brain, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.140 as residues: Met-1 to Gly-8.

[0110] The tissue distribution in immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders including: leukemias, lymphomas, autoimmune, immuno-suppressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain—and in particular the fetal brain—would suggest a possible role in the treatment and diagnosis of developmental and neurogenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer’s disease, mania, dementia, paranoia, and addictive behavior). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2086 of SEQ ID NO:30, b is an integer of 15 to 2100, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:21

[0115] The gene encoding the disclosed cDNA is thought to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

[0116] This gene is expressed primarily in brain, kidney, lung, liver, spleen, and a variety of leukocytes (especially T-cells) and to a lesser extent in a variety of other tissues and cell types.

[0117] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, autoimmune, immuno-suppressive, and immunodeficiencies, hematopoietic disorders, as well as renal disorders, and neoplasms. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal, pulmonary, immune, and central nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and other tissue of the nervous system, renal, pulmonary tissue, liver, spleen, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0118] The tissue distribution in immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of renal conditions, such as acute renal failure, kidney fibrosis, and kidney tubule regeneration. The expression in leukocytes and other immune tissues indicates a role in immune disorders including: leukemias, lymphomas, autoimmune, immuno-suppressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain—and in particular the fetal brain—indicates a possible role in the treatment and diagnosis of developmental and neurogenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer’s disease, mania, dementia, paranoia, and addictive behavior). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1434 of SEQ ID NO:31, b is an integer of 15 to 1448, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:22

[0120] The gene encoding the disclosed cDNA is thought to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

[0121] This gene is expressed primarily in skin (fetal epithelium, keratinocytes and skin).

[0122] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skin cancers (e.g., melanomas), eczema, psoriasis or other disorders of the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., keratinocytes, epithelium, integumentary, endothelial and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.142 as residues: Pro-28 to Gln-35, Ser-39 to Phe-44, Ala-94 to Gln-99.

[0112] The tissue distribution in intestinometry tissue, suggests that the protein product of this clone would be useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, portwine syndrome), intestinometry tumors (i.e. keratosis, Bowen’s disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget’s disease, mycosis fungoides, and Kaposi’s sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), attherosclerosis, ulceria, eczema, photosensitivity, and immune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, streias, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete’s foot, and ringworm). Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula a-b, where a is any integer between 1 to 442 of SEQ ID NO:32, b is an integer of 15 to 456, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:23

[0113] This gene maps to chromosome 11. Another group recently isolated this same gene, associating the sequence to the region thought to harbor the gene involved in Multiple Endocrine Neoplasia Type 1, or MEN 1. (See Accesion No. 2529721 and Genome Res. 7(7), 725-735 (1997), incorporated herein by reference in its entirety.) Preferred polypeptide fragments comprise the amino acid sequence: LEHWACLNERAQLPRTAXAGYQQFSCNGFS (SEQ ID NO:265).

[0114] This gene is expressed primarily in epididymus, pineal gland, T-cells, as well as fetal epithelium, lung and kidney.

[0115] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, metabolic mediated disorders, reproductive, endocrine, and MEN. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, renal, neurological and pulmonary systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, epididymus and other reproductive tissue, pineal gland, T-cells and other blood cells, epithelium, lung, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0116] The tissue distribution in fetal tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental deficiencies or abnormalities as well as a host of different disorders which arise as a result of conditions in the indicated tissues or cell types. An area of particular interest is in the treatment and diagnosis of immune disorders including: leukemia, lymphomas, auto-immune, immuno-suppressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoetic disorders. The expression in the brain, and in particular the fetal brain, would suggest a possible role in the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer’s disease, mania, dementia, paranoia, and addictive behavior). Respiratory/pulmonary disorders, such as ashtma, pulmonary edema are also potential therapeutic areas, as well as renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration. Moreover, this gene can be used in the treatment and/or detection of MEN 1. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula a-b, where a is any integer between 1 to 1312 of SEQ ID NO:33, b is an integer of 15 to 1326, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:24

[0117] This gene is expressed primarily in fetal spleen.

[0118] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to developmental, leukemia, lymphoma, AIDS, hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, spleen, developmental, hepatic, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue

or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0119] The tissue distribution in fetal spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-suppressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. Expression of this gene product in fetal spleen suggests a role in the regulation of the proliferation, survival, differentiation and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer of 15 to 710, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where the b is greater than or equal to a+14.

[0121] This gene is expressed primarily in fetal lung and kidney, human embryo and osteoclastoma stromal cells and to a lesser extent in a variety of other tissues and cell types.

[0122] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental disorders and cancers, as well as pulmonary and renal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory/pulmonary, skeletal and renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., lung, kidney, embryonic tissue, and bone cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:145 as residues: Thr-5 to Pro-18, Ala-76 to Thr-84.

[0123] The tissue distribution in fetal tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of osteoporosis, fractures, osteosarcoma, ossification, and osteonecrosis, as well as respiratory/pulmonary disorders, such as asthema, pulmonary cлема, and renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration. Alternatively, this gene may function in a tumor suppression capacity, and it may be down-regulated by tumor cells or proto-oncogenes. Expression of this gene may be important in the prevention of tumor growth or metastasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1174 of SEQ ID NO:35, b is an integer of 15 to 1188, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:26

[0124] This gene is homologous to the HIV envelope glycoprotein. (See Accession No. 2641463.) Preferred polypeptide fragments comprise the amino acid sequence:

NVRALLHRMPEPPRKINTAKSNKXKKMNLSL  
(SEQ ID NO:288).

[0125] This gene is expressed primarily in pinclal gland and skin, and to a lesser extent in lung.

[0126] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological and behavioral disorders; respiratory/pulmonary disorders, such as atchasea, pulmonary edema; skin conditions such as eczema, psoriasis, acne and skin cancer, as well as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and respiratory systems, as well as skin and AIDS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., blood cells, pinclal gland, integumentary, endocrine, epidermis, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Gln-15 to Gln-20.

[0127] The tissue distribution in integumentary tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions which affect the above tissues, such as skin cancer, eczema, psoriasis, acne, athesma, pulmonary edema, neuro-degenerative or developmental disorders such as Alzheimer's, depression, schizophrenia, dementia, and AIDS. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 942 of SEQ ID NO:36, b is an integer of 15 to 956, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:27

[0128] Preferred polypeptide encoded by this gene comprise the following amino acid sequence:

NVRALLHRMPEPPRKINTAKSNKXKKMNLSL
(SEQ ID NO:288).

[0129] Polynucleotides encoding such polypeptides are also provided as are complementary polynucleotides thereto. The gene encoding the disclosed cDNA is thought to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2. Contact of cells with a supernatant expressing the product of this gene increases the permeability of the plasma membranes of both astrocytes and monocytes to calcium. Thus, it is likely that the product of this gene is involved in signal transduction pathway(s) which are initiated when the product binds a receptor(s) on the surface of both astrocytes and monocytes. Thus, polynucleotides have uses which include, but are not limited to, activating astrocytes and monocytes.

[0130] This gene is expressed primarily in liver (adult and fetal) and spleen tissue, and to a lesser extent in placenta, T helper cells, kidney tumor, ovarian tumor, melanocytes and fetal heart.

[0131] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and developmental diseases and disorders and liver diseases such as liver cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, circulatory and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, spleen, placenta, blood cells, developmental, kidney, ovary and other reproductive tissue, melanocytes, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, bile, serum-plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0132] The tissue distribution in immune cells indicates that the protein products of this gene are useful for study, diagnosis and treatment of growth, hematopoietic and immune system disorders particularly related to the liver.
Expression of this gene product in hematopoietic cells suggests a role in the regulation of the proliferation, survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1589 of SEQ ID NO:37, b is an integer of 15 to 1603, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:28

The translation product of this gene shares sequence homology with prostatigrandin transporter which is thought to be important in metabolic and endocrine disorders. See, for example, Gastroenterology Oct: 109(4):1274-1282 (1995). Preferred polypeptides encoded by this gene comprise the following amino acid sequence:  

SYLSACFACGALTCACTTVPAEANTVPGKDSPCGQQSLFLPLLCW (SEQ ID NO:270); and/or

PSVILIILRTSPESKLXVLOVFLLRL, LQFIPPLIFGPAGIDSTCLPNSTPCSEQGACVLYDNAVYLVL (SEQ ID NO:271).

This gene is expressed primarily in hematopoietic and brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic, immune and endocrine diseases and disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic, immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, endocrine tissue, hematopoietic tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developmental and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary tissue, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, pulmonary, surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:149 as residues: Val-50 to Trp-55.

The tissue distribution in fetal lung indicates that the protein products of this gene are useful for study, diagnosis and treatment of respiratory and growth diseases and disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 615 of SEQ ID NO:39, b is an integer of 15 to 629, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:30

The translation product of this gene shares sequence homology with human DNA helicase which is thought to be important in accurate and complete DNA replication in creation of new cells. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

QSLFTRFVKVGYVTVLQDAQQRARA
SLCXXYMWYKLGNHPWQLLEPSFSTAMAGLLLYDFQQLNESFDQQWGESEP
PYFQNGLGMAGYVVALFMYCCGLGYPNAKDNLSTLSSYQQHLNLDVIRIBEGHN
PLRRMNKTPTVRFPQQNDYDILLYVRTHAHLGIDVWRLLVAVSRRAR
LVKXAKIANTCTHAALKDOLVLQGKFYIKILMKE
AAQILEITEFPILLLQNPNQDFSRLEMKWIMGDHHQLPPVI
(SEQ ID NO:1272); and/or

Features of Protein Encoded by Gene No:31

This gene encoding the disclosed cDNA is thought to reside on chromosome 15. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 15.

This gene is expressed primarily testes tumor and to a lesser extent in adrenal gland tumor and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, reproductive disorders, cancers and endocrine/growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, developmental, and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., testes and other reproductive tissue, adrenal gland, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testes combined with the homology to a DNA helicase indicates that the protein products of this gene are useful for study, treatment, and diagnosis of many cancer types, including testicular cancer; as well as disorders involving endocrine function and normal growth and development. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1950 of SEQ ID NO:40, b is an integer of 15 to 1964, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:31

The translation product of this gene shares sequence homology with BID-apoptotic death gene (mouse), Genbank accession no. PID g1669514, which is thought to be important in programmed cell death.

This gene is expressed primarily in jurkat membrane bound polysomes and activated neutrophils and to a lesser extent in endothelial cells and human cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene
The distribution in immune cells combined with the homology to the BID-apoptotic death gene indicates that the protein products of this gene are useful for study of cell death, and treatment and diagnosis of proliferative disorders and cancers. Apoptosis—programmed cell death—is a physiological mechanism involved in the deletion of peripheral T lymphocytes of the immune system, and its dysregulation can lead to a number of different pathogenic processes. Diseases associated with increased cell survival, or the inhibition of apoptosis, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, such as breast cancer, prostate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune disorders (such as systemic lupus erythematosus and immune-related glomerulonephritis rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation; graft vs. host disease, acute graft rejection, and chronic graft rejection. Diseases associated with increased apoptosis include AIDS; neurodegenerative disorders (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration); myelodysplastic syndromes (such as aplastic anemia), ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia. Thus, the invention provides a method of enhancing apoptosis in an individual by treating the individual with a polypeptide encoded by this gene. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1508 of SEQ ID NO:41; b is an integer of 1 to 1522, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:32

The translation product of this gene shares sequence homology with human fructose transporter which is thought to be important in normal metabolic function and activity.

This gene is expressed primarily in T-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemia and other cancers, and metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, lymph and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, brain, T-cells and other blood cells, metabolic tissues, and cancers and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:152 as residues: Pro-22 to Gly-46, Ser-54 to Pro-61.

The tissue distribution in T-cell lymphoma indicates that the protein products of this gene are useful for study of mechanisms leading to cancer, treatment and diagnosis of cancers and pre-cancerous conditions; as well as the study and treatment of various metabolic diseases and disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 861 of SEQ ID NO:42, b is an integer of 15 to 875, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where the b is greater than or equal to a+14.

This gene is expressed primarily in human meningioma and placental tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and other disorders of the CNS. Similarly, polypeptides and antibodies directed to
these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, meningeal, developmental, proliferating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.153 as residues: Asn-23 to Pro-31.

[0155] The tissue distribution in neural tissue indicates that the protein products of this gene are useful for study, diagnosis and treatment of disorders of the CNS and inflammatory responses. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 829 of SEQ ID NO:43, b is an integer of 15 to 843, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:34

[0156] This gene is expressed primarily in activated monocytes and wound healing tissues and to a lesser extent in fetal epithelium.

[0157] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and inflammatory disorders and wound healing and tissue repair dysfunctions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, epithelial and gastrointestinal systems, and healing wounds, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, keratinocytes, monocytes, intestinal, developmental, and other blood cells, and epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.154 as residues: Ala-28 to Ala-33, Gly-35 to Glu-45.

[0158] The tissue distribution in immune cells indicates that the protein products of this gene are useful for diagnosis, study and treatment of immune and inflammatory disorders and wound healing dysfunctions. Expression of this gene product in immune cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leuke-

Features of Protein Encoded by Gene No:35

[0159] This gene is expressed primarily in human osteosarcoma and prostate cancer.

[0160] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skeletal and neoplastic conditions such as bone and prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and skeletal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, bone, prostate, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:155 as residues: Ser-14 to Gly-22, Leu-37 to Gln-43.

[0161] The tissue distribution in skeletal cells indicates that the protein products of this gene are useful for diagnosis and treatment of skeletal disorders and cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 520 of SEQ ID NO:45, b is an integer of 15 to 534, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where the b is greater than or equal to a +14.

Features of Protein Encoded by Gene No:36

[0162] This gene encodes a protein which is highly homologous to a protein called congenital heart disease protein 5, presumably implicated in congenital heart disease (see Genbank PID g2810996).

[0163] This gene is expressed primarily in Hodgkin's lymphoma, erythroleukemia cells, and TNF activated synovial fibroblasts, to a lesser extent in ovarian cancer, cerebellum, spleen, fetal liver and placenta and finally to a lesser extent in various mesenchymal tissues.

[0164] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer, immune, hematopoietic and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and cardiovascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., heart and other cardiovascular tissue, immune, lymphoid tissue, blood cells, bone marrow, ovary and other reproductive tissue, brain and other tissue of the nervous system, spleen, liver, and mesenchymal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ED NO.156 as residues: Lys-41 to Met-49, Gln-54 to Glu-59, Glu-76 to Thr-88.

[0165] The homology of this gene and translation product to congenital heart disease protein 5 indicates a role for this protein in the diagnosis, prognosis and/or treatment of heart disease or other cardiovascular related disorders. In addition, predominant expression in cells associated with the immune and hematopoietic system indicates a role for this protein in the treatment, diagnosis and/or prognosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, AIDS, thymus disorders such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism and hypothyroidism, Craft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia bone marrow fibrosis, and myeloproliferative disease. The protein could also be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic disorders including immunodeficiency diseases, leukemia, and septicemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1360 of SEQ ID NO:46, b is an integer of 15 to 1374, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where the b is greater than or equal to a +14.

Features of Protein Encoded by Gene No:37

[0166] This gene is expressed primarily in ovarian cancer.

[0167] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, urogenital neoplasias, reproductive, or endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ED NO.157 as residues: Asn-22 to Asn-27.

[0168] The tissue distribution in ovarian tissues indicates that polynucleotides and polypeptides corresponding to this
gene are useful for study, diagnosis and treatment of ovarian and other tumors. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 857 of SEQ ID NO:48, b is an integer of 15 to 851, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:38

[0169] The translation product of this gene shares sequence homology with zinc finger proteins, which are small DNA-binding molecules noted for their occurrence in a large number of eukaryotic transcription factors.

[0170] This gene is expressed primarily in fetal, cancer, and endothelial lines.

[0171] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and other tissue of the nervous system, endocrine tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, bile, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0172] The tissue distribution in fetal tissue indicates that the protein products of this gene are useful for study, diagnosis and treatment of immune and developmental conditions and cancer. The homology to zinc finger proteins suggests that this protein may play a role in the transcriptional regulation of certain cancer genes. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2006 of SEQ ID NO:49, b is an integer of 15 to 2020, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where the b is greater than or equal to a+14.

[0173] This gene is expressed primarily in fetal, infant, and adult brain and to a lesser extent in other brain and endocrine organs and blastomas.

[0174] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, brain tumors and neurodegenerative conditions, in addition to developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and other tissue of the nervous system, endocrine tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0175] The tissue distribution in neural tissue indicates that the protein products of this gene are useful for the study, diagnosis and treatment of brain cancer and other neurological disorders such as Alzheimers Disease, Parkinsons Disease, Huntington Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psycho- ses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protect as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2006 of SEQ ID NO:49, b is an integer of 15 to 2020, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where the b is greater than or equal to a+14.
Features of Protein Encoded by Gene No:40

[0176] The translation product of this gene shares sequence homology with vesicular glycoproteins and lectins. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

```
DTYPNEEKQQERVFPXXSAMWNNGSLSYDHER DGRPTELGGCXAIWRNLHYDTFLWIRYWKRHLTIMMDIDGKHEWRDCIEWPGW RLPGYYF7SSGTGDLSDNVDVSLXKLFELTVERTPEKE LKREHSLEDKXYQVGTGGSSLWNLMSNNAMWTVQYIRLTPDMGSQSA
```

[0177] The gene encoding the disclosed cDNA is thought to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2. When tested against U937 myeloid cell lines and Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS pathway. Thus, it is likely that this gene activates myeloid cells and T-cells through the Jaks-STAT signal transduction pathway. The Gamma Activating Sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. When tested against sensory neuron cell lines, supernatants removed from cells containing this gene activated the EGR1 pathway. Thus, it is likely that this gene activates sensory neuron cells through a signal transduction pathway induced by the EGR1 promoter. The Early Growth Response Gene 1 (EGR1) is a separate signal transduction pathway in which the EGR1 promoter induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

[0178] This gene is expressed primarily in infant brain and to a lesser extent in various normal and transformed neural, endocrine, and immune organs.

[0179] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological and neurodevelopmental conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and hormonal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, endocrine tissue, and tissues of the immune system, developmental disorders, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:160 as residues: Pro-64 to Gly-71, Gly-94 to Leu-100, Thr-110 to Pro-116, Thr-135 to Arg-145, Glu-164 to Glu-171, Asp-204 to Asp-211, Arg-253 to His-261, Asn-312 to Tyr-323.

[0180] The tissue distribution in neural tissue indicates that the protein products of this gene are useful for the study, diagnosis and treatment of mental retardation and other neurological disorders and neoplasias. The activity of this gene seen in various biological assays indicates that this gene is involved in a number of signal transduction assays, which further suggests that this gene could be important in cell proliferation and differentiation. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2418 of SEQ ID NO:50, b is an integer of 15 to 2432, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:41

[0181] This gene displays homology to the glycosyltransferase family, which catalyze the addition of sialic acids to carbohydrate groups which are present on glycoproteins and glycolipids.

[0182] This gene is expressed primarily in smooth muscle and to a lesser extent in pineal gland, fetal liver, and infant brain.

[0183] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, gastrointestinal injury, inflammatory and neurodegenerative conditions, endocrine, hematopoietic, hepatic or developmental disorders. Similarly, polypeptides
and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., smooth muscle, pinéal gland, liver, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:161 as residues: Ser-12 to Trp-21, Arg-24 to Pro-32, Asp-73 to Lys-82, Lys-90 to Ala-97.

[0184] The tissue distribution in neural and fetal tissues indicates that the protein products of this gene are useful for the study, diagnosis and treatment of neurodegenerative and growth disorders and gastrointestinal repair. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapeutic targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2336 of SEQ ID NO:51, b is an integer of 15 to 2340, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:42

[0185] The translation product of this gene shares sequence similarity with metallothionein polypeptides. See, for example, Proc. Natl. Acad. Sci. USA 1992 July 15;89(14):6333-6337. Metallothioneins are believed to inhibit neuronal survival among other biological functions. Based on the sequence similarity (especially the conserved cysteine motifs characteristic of the metallothionein family) the translation product of this gene is expected to share certain biological activities with other members of the metallothionein polypeptide family. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

PQTQLQCSALHIDPOCANCSCRPRCD CSPACOC (SEQ ID NO:276).

[0186] This gene is expressed exclusively in placenta and fetal liver, and to a lesser extent in osteoblast and bone marrow cells.

[0187] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders and hepatic or skeletal system conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, placenta, liver, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0188] The tissue distribution in immune cells and homology to metallothioneins indicates that the protein products of this gene are useful for diagnosis and treatment of immune and hematopoietic system disorders and neurological diseases, especially in fetal development. Expression of this gene product in hematopoietic cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also-used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapeutic targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 587 of SEQ ID NO:52, b is an integer of 15 to 601, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:43

[0189] Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

FLYDVLMLXHEAVMRTHQHLPDFEPS (SEQ ID NO:277).
This gene is expressed primarily in T-cells and synovial tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., synovial tissue, and T-cells and other blood cells, and cancers and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells indicates that the protein products of this gene are useful for treatment and diagnosis of disorders of the immune system. Expression of this gene product in immune cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 345 of SEQ ID NO:53, b is an integer of 15 to 359, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:44

The translation product of this gene shares sequence similarity with several methyltransferases (e.g., see Genbank gi1065505) which suggests this protein would be important in normal developmental and cellular processes.

This gene is expressed primarily in ovary, thymus, infant adrenal gland, tissues of the nervous system and the hematopoietic tissue, and to a lesser extent in adipose tissue and other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the reproductive system, the endocrine-system, the hematopoietic system and the CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, endocrine, CNS and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., ovary and other reproductive tissue, thymus, adrenal gland, brain and other tissue of the nervous system, hematopoietic tissue, and adipose tissue, and cancersous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Ser-3 to Gly-12, Asp-19 to Arg-31, Tyr-70 to Tyr-77, Asn-130 to Lys-140, Pro-165 to Gln-170, Pro-192 to Lys-199, Leu-216 to Glu-227, Glu-254 to Phe-261.

The tissue distribution in hematopoietic cells and homology to methyltransferase indicates that the protein products of this gene are useful for diagnosis and treatment of disorders of the CNS, the hematopoietic system and reproductive organs and tissues. For example, the abundant expression in the ovary may indicate that the gene product can be used as a hormone with either systemic or reproductive functions; as growth factors for germ cell maintenance and in vitro culture; as a fertility control agent; remedy for sexual dysfunction or sex development disorders; diagnosties/treatment for ovarian tumors, such as serous adenocarcinoma, dysgerminoma, embryonal carcinoma, choriocarcinoma, teratoma, etc.; The expression in thymus may indicate its utility in T-cell development and thus its applications in immune related medical conditions, such as infection, allergy, immune deficiency, tissue/organ transplantation, etc. Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more-polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a
is any integer between 1 to 1127 of SEQ ID NO:54, b is an integer of 15 to 1141, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:45

[0197] The translation product of this gene shares sequence homology with cytochrome C oxidase which is thought to be important in the metabolic function of cells. This gene has now recently been published as estrogen response gene. See Genbank accession no. AB007618 and Mol. Cell. Biol. 18 (1), 442-449 (1998). See also J Immunol. March 1:154(5): 2384-2392 (1995), where the mouse homologue was published and implicated in silicos. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

PADXKPWWSTEAPP11PAFAPPTLTSQTVY
DYGTKKPY3PELQKFQAD7VFYLRKGLDQLYRTM1LVGGTYICLIAL
YNASQ PK
SFGAVALAADDASRTLGVNTFFSFPTQ
KLAGAWASEAYSPQ1SNLFPPQKhLSYLPWFQN

[0198] Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

[0199] This gene is expressed primarily in adipose tissue, kidney and fetal brain and to a lesser extent in other tissues and organs.

[0200] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic diseases involving especially adipose tissue, brain and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adipose tissue, kidney, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:165 as residues: Thr-8 to Ser-13, Ser-29 to Ala-34, Pro-64 to Lys-77.

[0201] The tissue distribution and homology to cytochrome C oxidase, estrogen response gene product and silicos related gene product indicates that the protein products of this gene are useful for diagnosis and treatment of metabolic disorders in the CNS, adipose tissue and kidney, particularly silicos. Expression within fetal suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a→b, where a is any integer between 1 to 1546 of SEQ ID NO:55, b is an integer of 15 to 1560, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:46

[0202] The translation product of this gene shares sequence homology with reticulocein. See, for example, J. Biochim. 117 (5), 1113-1119 (1995). Based on the sequence similarity, the translation product of this gene is expected to share certain biological activities with reticulocein, e.g., Ca++ binding activities. This gene product is sometimes hereinafter referred to as “Reticulocein”. When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS pathway. Thus, it is likely that this gene activates T-cells through the Jaks-STAT signal transduction pathway. The Gamma Activating Sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jas-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. When tested against K562 leukemia cell lines, supernatants removed from cells containing this gene activated the ISRE pathway. Thus, it is likely that this gene activates leukemia cells through a signal transduction pathway induced by the ISRE promoter. The Interferon-Sensitive Responsive Element (ISRE) is a promoter element found upstream in many genes which are involved in the
Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in breast, endothe-lial cells, synovial, heart and smooth muscle cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the breast, vascular, skeletal/cardiac muscular system as well as the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast, vascular and skeletal-muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, endothelial cells, synovial tissue, heart and other cardiovascular tissue, smooth muscle, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, breast milk, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:166 as residues:

```
SEQ ID NO. 166 as residues:
```

The tissue distribution in smooth muscle cells indicates that the protein products of this gene are useful for diagnosis and treatment of diseases of the vascular and skeletal-cardiac muscular system. The homology of the gene with reticulocalbin indicates its biological function in regulating calcium store, a particularly important function in muscular cell types. The gene expression in the heart may indicate its utilities in diagnosis and remedy in heart failure, ischemic heart diseases, cardiomyopathy, hypertension, arrhythmia, etc. The abundant expression in the breast may indicate its applications in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Paget disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma; juvenile hypertrophy and gynecomastria, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases, etc. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1493 of SEQ ID NO:56, b is an integer of 15 to 1507, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:47

The translation product of this gene shares weak sequence homology with H+-transporting ATP synthase which is thought to be important in cell metabolism or signal transduction.

This gene is expressed only in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of some types of diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and hematopoietic tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Since only one out of about a million expressed sequence tags are found in testes, it is reasonable to suggest that the expression of this gene is selective for testes. Since some of the genes only expressed in testes are usually expressed in brain or in certain induced hematopoietic cells/tissues, it is speculated that this gene will be expressed in brain or hematopoietic cells/tissues and is useful for diagnosis and treatment of disorders of these systems. Similarly, the secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It
may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombotic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory, activity (e.g. for treating septic shock, Crohn’s disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a,b, where a is an integer between 1 to 436 of SEQ ID NO:57, b is an integer of 15 to 456, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:48

[0210] The translation product of this gene shares sequence homology with human polymeric immunoglobulin receptor (accession No. X73079) which is thought to be important in antibody recognition and immune defenses. In one embodiment, polypeptides of the invention comprise the sequence

GWWNGC (SEQ ID NO:230).

[0211] Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

[0212] This gene is expressed primarily in placenta and to a lesser extent in corpus callosum and fetal liver and spleen.

[0213] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the immune system, e.g. autoimmune diseases and immunodeficiency, in addition to developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, liver, and spleen, developmental tissues, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 168 as residues: Tyr-37 to Cys-49, Gly-51 to Tyr-56, Lys-88 to Thr-93, Leu-130 to Glu-136.

[0214] The tissue distribution in fetal liver and spleen combined with the homology to human polymeric immunoglobulin receptor indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, e.g. autoimmune diseases and immunodeficiencies. Expression within fetal tissues and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a,b, where a is an integer between 1 to 1133 of SEQ ID NO:58, b is an integer of 15 to 1147, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:49

[0215] This gene is expressed in thymus.

[0216] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, such as inflammation or immunodeficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., immune, hematopoietic, thymus and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0217] The tissue distribution in thymus indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, such as autoimmunity and immunodeficiency disorders. Similarly, the gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore, it may also be used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 763 of SEQ ID NO:59, b is an integer of 15 to 777, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:50

[0218] Preferred polypeptide encoded by this gene comprise the following amino acid sequence:

```
MKVGAR1VENVISVKKHSPV5E7HNNKFAENPQ6MHFLAQEPRT
VKKRPGLAG5F1Q6S5TRKPLQET1M8AD7A1WVFY85HRHEQHELQY
LQYKMDHLDSSQGLIGHTLLQ65NLALTNI
```

[0219] Polynucleotides encoding this polypeptide are also provided as are complementary polynucleotides thereto.

[0220] This gene is expressed primarily in adrenal gland, pituitary, T-helper cells, and breast cells and to a lesser extent in a wide variety of tissues.

[0221] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the same diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely determined in certain tissues and cell types (e.g., adrenal gland, pituitary, T-cells and other blood cells, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, breast milk, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:170 as residues: Ghn-39 to Ser-47, Arg-57 to Gln-67, Tyr-82 to Gln-95.

[0222] The tissue distribution in immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of a wide range of disorders, such as immune and endocrine disorders. Similarly, the secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating cancer or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokine activities (e.g. for treating infections, tumors); hemostatic or thrombotic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Pref-
erably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1177 of SEQ-ID NO:60, b is an integer of 15 to 1191, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where the b is greater than or equal to 14.

Features of Protein Encoded by Gene No:51

The translation product of this gene shares sequence homology with human Sop2p-like protein which is important in cytoskeleton structure. In one embodiment, polypeptides of the invention comprise the sequence

SLHRNSVSQYSVLSGKRKQSFCTTG- MDGGMNWVDSKSLALKR (SEQ ID NO:282).

Polynucleotides encoding this polypeptide are also encompassed by the invention. This gene maps to chromosome 7. Therefore, polynucleotides of the invention can be used in linkage analysis as a marker for chromosome 7.

This gene is expressed primarily in immune and hematopoietic tissues/cells and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immunological and hematopoietic disorders and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., immune and hematopoietic tissue/cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in

SEQ ID NO. 171 as residues: Lys-49 to Gln-54, Ala-61 to Arg-66,
Lys-82 to Lys-87, Glu-126 to Val-133, His-136 to Ile-141, Glu-175 to Ser-137, Asp-286
to Leu-296, Ala-298 to Ser-310.

The tissue distribution in immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immunological, hematopoietic, and inflammatory disorders, e.g., immunodeficiency, autoimmunity, inflammation. Protein, as well as, antibodies directed against the protain may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1566 of SEQ ID NO:61, b is an integer of 15 to 1580, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where the b is greater than or equal to 14.

Features of Protein Encoded by Gene No:52

The translation product of this gene shares sequence homology with Caenorhabditis elegans R05.5 gene encoding a putative secreted protein.

This gene is expressed primarily in endothelial cells, brain and several highly vascularized, and tumor tissues and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, aberrant angiogenesis and tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and neural systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, brain and other tissue of the nervous system, and vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.172 as residues: Thr-43 to Asn-60, Thr-106 to Phe-115, Asp-122 to Arg-133, Arg-180 to Asp-192, Leu-211 to Lys-216.

The tissue distribution in vascular tissue combined with the homology to a C. elegans secreted protein indicates...
through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1103 of SEQ ID NO:62, b is an integer of 15 to 1117, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:53

[0232] In one embodiment, polypeptides of the invention comprise the sequence

EASKSSHAGDLDFSVAACHR (SEQ ID NO:263).

[0233] Polynucleotides encoding this polypeptide are also encompassed by the invention. When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS pathway. Thus, it is likely that this gene activates T-cells through the Jaks-STAT signal transduction pathway. The Gamma Activating Sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

[0234] This gene is expressed primarily in T-cells and to a lesser extent in brain.

[0235] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, lymphocytic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., immune, T-cells, or other blood cells, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:173 as residues: Pro-3 to Thr-8, Arg-37 to Asp-46.

[0236] The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of lymphocytic disorders. Alternatively, expression within neural tissue suggests that the protein product of this clone would be useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 347 of SEQ-ID NO:63, b is an integer of 15 to 361, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No:54

[0237] The translation product of this gene shares sequence homology with secreted cartilage matrix protein, a major component of the extracellular matrix of nonarticular cartilage which is thought to be important in cartilage structure. In specific embodiments, polypeptides of the invention comprise the sequence:

RCKKCTEAP  (SEQ ID NO:292);
DLVIPWIDGSKRGHERFRLVWQP  (SEQ ID NO:285);
VTGIIDSLTISPKAARRWGL  (SEQ ID NO:285);
LQYGTQYK  (SEQ ID NO:286);
TEFTLRHPNSAKMDKKAQVAHMKYM  (SEQ ID NO:286);
GKGMONTLALKMPFRSPTQEQGARPP  (SEQ ID NO:287);
[0238] Poly nucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

[0239] This gene is expressed primarily in placenta, infant brain, prostate, fetal lung and to a lesser extent in endometrium and fetal tissues.

[0240] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormal placenta and pregnancy, disorder and injury in brain, prostate, and vasculature. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproduction, neuronal, and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developing, placenta, brain and other tissue of the nervous system, prostate, lung and endometrium, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, seminal fluid, pulmonary surfactant, or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0241] The tissue distribution in placental tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of abnormalities in placenta and pregnancy, disorder and injury in brain, prostate, and vasculature. Similarly, the homology to the cartilage matrix protein suggests that the protein product of this clone would be useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary disorders (i.e. keratoses, Bowen’s disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget’s disease, mycosis fungoides, and Kaposi’s sarcoma), infections and inflammation of the skin (i.e. wounds, rashes, prickle heat disorder, psoriasis, dermatitis), attherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasmas. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Moreover, the protein product of this clone may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendinitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior
to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1654 of SEQ ID NO:64, b is an integer of 15 to 1668, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 55

[0242] The translation product of this gene is the human ortholog of bovine and hamster CII-3, a succinate-ubiquinone oxidoreductase complex II membrane-intrinsic subunit, which is thought to be important in mitochondrial electron transport chain during metabolism. In specific embodiments, the polypeptides of the invention comprise

MAALLLRHVSRCRANASFGQCRINNAVLOTTAEMERFWMNKHIGNRPLQFHITTSY (SEQ ID NO:295);
VFPMYHTWNGIHLMILGKGLKIPOLYQSG (SEQ ID NO:296);
MAALLLRHVSRCRANAH (SEQ ID NO:297);
V3I1CLOPALINTAKPAL (SEQ ID NO:298);
VFPMYHTWNGIHLMILGKGL (SEQ ID NO:299).

[0243] This gene is expressed in 8-week old early stage human.

[0244] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, metabolic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0245] The tissue distribution in fetal tissue combined with the homology to a metabolic protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of a metabolic disorder. Similarly, expression within embryonic tissue and other cellular sources marked by proliferating cells suggest that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may be publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1339 of SEQ ID NO:65, b is an integer of 15 to 1353, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 56

[0246] This gene is expressed primarily in umbilical vein endothelial cells, human ovarian tumor cells, human meningioma cells, and human Jurkat membrane-bound poly-somes. In specific embodiments, polypeptides of the invention comprise the amino acid sequence:

RVWDRVRFAPKRCVRKIFQGNV (SEQ ID NO:300);
HNFEKLRHCSWGEPGAKIAAGSADRFVV (SEQ ID NO:301);
WDTTSRILXLYLQHAGSINEFAPGDEF1 (SEQ ID NO:302),
YQGLGLQHKLTTYNMGRADEVTG (SEQ ID NO:303), or
LSLSETSYLSSAMOUNTVRVNWDRVRFAPKRCVRKIFQGNVHNFEKLRHCS.

Mar. 27, 2003
Polynucleotides encoding these polypeptides are also encompassed by the invention.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation, immune and cardiovascular disorders and urogenital neoplasias, and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neurological, urogenital, reproductive system and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, cells, endothelial cells, ovary and other reproductive tissue, developmental, meningina, and cancerous and wounded tissues) or bodily fluids (e.g. amniotic fluid, seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard.

Features of Protein Encoded by Gene No: 57

The translation product of this gene shares sequence homology with type I collagen. In specific embodiments, the polypeptides of the invention comprise the sequence:

\[
\text{GRIPAPASVFAQGKQSR} \quad \text{(SEQ ID NO:308)};
\]
\[
\text{VRGRTVLPGOLDAPELGEPE} \quad \text{(SEQ ID NO:305)};
\]
\[
\text{EGRVLERKKEREKRKEEQ} \quad \text{(SEQ ID NO:306)};
\]
\[
\text{ARRSGAELAWDYLCRWAQKHKNWRFOKTROTWLLLHMYDSDKWPDEHFSTLLAYLE} \quad \text{(SEQ ID NO:309)};
\]
\[
\text{and/or}
\]
\[
\text{GLQQR} \quad \text{(SEQ ID NO:307)}.
\]

Polynucleotides encoding these polypeptides are also encompassed by the invention. Polynucleotides of the invention do not comprise the nucleic acid sequence shown as Genbank Accession No. gb[J07392]HUMRETPIGA, which is hereby incorporated herein by reference.

This gene is expressed primarily in epididymis, prostate cell line (LNCAP), and pituitary gland; and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymis, prostate (especially prostate cancer), pituitary gland, or other reproductive, urogenital, or endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system and neuroendocrine system, expression of this gene at significantly higher or lower levels...
may be routinely detected in certain tissues and cell types (e.g., epididymus and other reproductive tissue, prostate, and pituitary gland, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0254] The tissue distribution and homology to type I collagen, indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, prostate (especially prostate cancer), and pituitary gland. Similarly, the protein product of this clone may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenital, familial osteoarthris, Atelectogenesis type II, metaphysial chondrodysplasia type Schmid). Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 1179 of SEQ ID NO:67, b is an integer of 15 to 1193, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 58

[0255] This gene is expressed primarily in the frontal cortex of the brain from a schizophrenic individual.

[0256] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neural disorders, particularly neurodegenerative disorders such as schizophrenia. Similarly, polynucleotides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0257] The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 546 of SEQ ID NO:68, b is an integer of 15 to 560, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 59

[0258] The polypeptide encoded by Gene 59 is homologous to human surface 4 integral membrane protein. In specific embodiments, the polypeptides of the invention comprise the sequence:

TCVVLVLSNRFVQYACPLGLGIIALQTVIAYSILWDKLFLMRRN (SEQ ID NO:310);
SREKGKMPAGQTPNRESSFPQKWQLGVPGLVWLVKPMFLLL (SEQ ID NO:311);
PDGFGPIVQWGNV (SEQ ID NO:312);
GTAEDGPQQLAVQQLLP HVARCLIST (SEQ ID NO:313);
PLEDGIPMPQWSEQQDRYTDITNCOYLLAS (SEQ ID NO:314);
including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may also be used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, poririasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmune disorders, such as autoimmune infertility, tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren’s disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publically available and accessible through sequence databases. Some of these sequences are related to SEID NO:69 and may have been publically available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1643 of SEQ ID NO:69, b is an integer of 15 to 1657, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where the b is greater than or equal to a+14.

**[0259]** NO:320. Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 9.

**[0260]** This gene is expressed primarily in Hodgkin’s lymphoma and lung; and to a lesser extent in many other human tissues.

**[0261]** Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, particularly Hodgkin’s lymphoma, tumors or other abnormalities of the lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. hematopoietic, lymphoid tissue, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:179 as residues: Met-20 to Thr-27.

**[0262]** The tissue distribution in immune tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of Hodgkin’s lymphoma, tumors or other abnormalities of the lung. Similarly, expression of this clone within immune tissues, particularly Hodgkin’s lymphoma, suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages,
Features of Protein Encoded by Gene No: 60

[0263] The gene encoding the disclosed cDNA is believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

[0264] This gene is expressed primarily in bone cancer and stomach cancer, and to a lesser extent in many other tissues.

[0265] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bone cancer and stomach cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone, and the stomach, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and stomach, skeletal, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, chyme, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0266] The tissue distribution in skeletal tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of skeletal or gastrointestinal disorders, particularly cancer. Similarly, the expression of this gene in skeletal tissue would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues (e.g., arthritis, trauma, tendinitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 697 of SEQ ID NO:70, b is an integer of 15 to 711, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 61

[0267] The gene encoding the disclosed cDNA is believed to reside on the X chromosome. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for the X chromosome.

[0268] This gene is expressed primarily in epididymus, and lymph node of breast cancer, and to a lesser extent in many other tissues.

[0269] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymus, and breast cancer or other reproductive conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the epididymus and breast, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., epididymus and other reproductive tissue, lymphoid tissue, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:181 as residues: Arg-57 to Ser-65.

[0270] The tissue distribution in reproductive tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, breast cancer, or other reproductive disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 921 of SEQ ID NO:71, b is an integer of 15 to 935, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 62

[0271] The translation product of this gene appears to be the human homolog of bovine NADH dehydrogenase which is thought to be important in cellular metabolism. In specific embodiments, the polypeptides of the invention comprise the amino acid sequence:
These polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neural disorders, particularly neurodegenerative disorders or abnormalities of the amygdala. Similarly, polynucleotides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the amygdala, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g. neural, amygdala, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.183 as residues: Gln-17 to Gln-29, Pro-41 to Phe-46, Ser-59 to Ile-70, Thr-97 to Leu-105.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of amygdala. Similarly, expression within neural tissues suggests that the protein product of this clone would be useful for the detection/treatment of neurodegenerative disease states, behavior disorders, or inflammatory conditions such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations's, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 490 of SEQ ID NO:72, b is an integer of 15 to 504, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 63

This gene is expressed primarily in amygdala, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neural disorders, particularly neurodegenerative disorders or abnormalities of the amygdala. Similarly, polynucleotides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the amygdala, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g. neural, amygdala, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.183 as residues: Gln-17 to Gln-29, Pro-41 to Phe-46, Ser-59 to Ile-70, Thr-97 to Leu-105.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of amygdala. Similarly, expression within neural tissues suggests that the protein product of this clone would be useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations's, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 606 of SEQ ID NO:73, b is an integer of 15 to 620, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 64

This gene is expressed primarily in female bladder, and to a lesser extent in chronic synovitis and hemangiopericytoma.
Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, urogenital or skeletal disorders, particularly bladder cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., bladder, synovial tissue, and vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.184 as residues: Pro-2 to Glu-7, Pro-27 to Phe-34.

The tissue distribution in urogenital tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatments of defects of the urinary tract, especially bladder cancer. Alternatively, expression within synovitis tissue suggests a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues such as arthritis, trauma, tendonitis, chronobondomalacia, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodyplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteurthratis, Taulostogedosis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 567 of SEQ ID NO:74, b is an integer of 15 to 581, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 65

This gene is expressed primarily in fetal spleen, and to a lesser extent in hemangiopericytoma, thymus, and synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, defects of immune of hematopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune of hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, spleen, vascular tissue, thymus, blood cells, and synovial tissue, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The protein product of this gene is useful for treatment of defects of the immune or hematopoietic systems, because of the gene’s expression in thymus and spleen. Similarly, the secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn’s disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1829 of SEQ ID NO:75, b is an integer of 15 to 1843, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 66

This gene is expressed primarily in human pituitary and to a lesser extent in placenta and fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification
of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, endocrine growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pituitary and other endocrine tissue, placenta, developmental and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.186 as residues: Val-38 to Asn-44, Gly-53 to Ser-65.

**[0288]** The tissue distribution in fetal tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of disorders related to endocrine or pituitary dysfunction, particularly growth disorders. Similarly, expression within fetal tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1427 of SEQ ID NO:76, b is an integer of 15 to 1441, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 67

**[0289]** The translation product of this gene shares sequence homology with a *Caenorhabditis elegans* gene. In specific embodiments, the polypeptides of the invention comprise the sequence:

```
<table>
<thead>
<tr>
<th>Sequence</th>
<th>(SEQ ID NO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPRPKNVSLYPFPSE CNPALDDPFP (SEQ ID NO:1323);</td>
<td></td>
</tr>
<tr>
<td>DYNLLGIPSNMCGLNKLKCCAWA VYCS (SEQ ID NO:1324);</td>
<td></td>
</tr>
<tr>
<td>PIFANSRSSEPETQGMGPS (SEQ ID NO:1322);</td>
<td>and/or</td>
</tr>
<tr>
<td>and/or</td>
<td></td>
</tr>
<tr>
<td>MLSISAVHVTQLQFQPMTPPW (SEQ ID NO:1325).</td>
<td></td>
</tr>
</tbody>
</table>
```

Features of Protein Encoded by Gene No: 68

**[0290]** Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

**[0291]** This gene is expressed primarily in primary breast cancer and lymph node breast cancer and to a lesser extent in adult brain, lung cancer, colon cancer, epithelioid sarcoma, and Caco-2 cell line.

**[0292]** Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, reproductive, neural, or endothelial disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer and tumor tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., mammary tissue, lymphoid tissue, brain and other tissue of the nervous system, lung, colon, and epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:187 as residues: Asn-34 to Lys-42.

**[0293]** The tissue distribution in a variety of cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of a variety of cancer and tumor types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 896 of SEQ ID NO:77, b is an integer of 15 to 910, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 68

**[0294]** The translation product of this gene shares sequence homology with steroid membrane binding protein. The translation product of this gene has recently been published as progesterone binding protein. See Genbank AJ002030. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:
The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polymucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in breast, and to a lesser extent in placenta and fetal tissue.

Therefore, polymucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of breast or fetal tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, mammary tissue, placenta, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, breast milk, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:188 as residues Pro-43 to Asp-49, Glu-54 to Pro-64, Asp-110 to Asp-118, Lys-138 to Tyr-143, Pro-150 to Asp-170.

The tissue distribution in reproductive tissues combined with the homology to a steroid membrane binding protein and to progesterone binding protein indicates that the protein products of this gene are useful for treatment of breast cancers, especially those caused by estrogen and progesterone binding. Similarly, expression within fetal tissues and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polymucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polymucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polymucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2762 of SEQ ID NO:78, b is an integer of 15 to 2776, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 69

It is likely that the open reading frame containing the predicted signal peptide continues in the 3' direction. Therefore, preferred polypeptides encoded by this gene comprise the following amino acid sequence:

AADDNYGIPRACRNSARSYGAAWLLLXPAGSSRWEPTQDISISDDQLGGQDV PFRN8ELLN8VGVQNVFSLFRGTSREVPPAEVAT6TLLAFATAQPL

The present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula

of a-b, where a is any integer between 1 to 2762 of SEQ ID NO:78, b is an integer of 15 to 2776, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where the b is greater than or equal to a+14.

This gene is expressed primarily in macrophage (GM-CSF treated), and to a lesser extent in monocytes and dendritic cells.

Therefore, polymucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune or hematopoietic disorders, particularly inflammation and infection. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g. immune, macrophages and other blood cells, and dendritic cells, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune tissue indicates that the protein products of this gene are useful for treatment of infection or inflammation or other events or defects involving the immune system. Similarly, the tissue distribution suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used

AADNYGIPRACRNSARSYGAAWLLLXPAGSSRWEPTQDISISDQLGGQDV PFRN8ELLN8VGVQNVFSLFRGTSREVPPAEVAT6TLLAFATAQPL

IPLWVLYQPLSFPLMKPINCISHN (SEQ ID NO:128).
as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren’s disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1473 of SEQ ID NO:79, b is an integer of 15 to 1487, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 70

[0303] This gene was found to have homology to a conserved human 15 kDa selenoprotein (See Genbank Accession No. gi|305111 (AF051894)) which may be involved in the regulation of important cellular functions such as metabolism or cell cycle regulation.

[0304] This gene is expressed primarily in adult brain and to a lesser extent in thyroid, 12 week old early stage human, and stromal cell TFE274.

[0305] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological or neuro-endocrine diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous or endocrine systems, expression of tiss gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, developmental, immune, thyroid, endocrine, and stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:190 as residues: Pro-65 to Cys-71.

[0306] The tissue distribution in neural tissue indicates that the protein products of this gene are useful for treatment and diagnosis of neurological diseases or metabolic conditions involving the neuro-endocrine system. Similarly, the protein product of this clone would be useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimer’s Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, myeloplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1549 of SEQ ID NO:80, b is an integer of 15 to 1563, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 71

[0307] In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

CTLA6XLHGDGRCTGKLRGVLGLGHRPGVFLSPQYAYS
PADQQLQAQSVNVDGWMLDQETTPGK (SEQ ID NO:329).

[0308] Polynucleotides encoding these polypeptides are also encompassed by the invention.

[0309] This gene is expressed in helper T-cells and, to a lesser extent, in adult brain and adult testes.

[0310] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, meningitis or reproductive problems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the
above tissues or cells, particularly of the immune, neural and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, brain and other tissue of the nervous system, testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g. seminal fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.191 as residues: Val-18 to Tyr-24, Ala-89 to Asp-99, Asp-104 to Ala-117, Leu-121 to Pro-136.

[0311] The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune and reproductive disorders. Similarly, the secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppresive activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); modulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein. As well as antibodies, directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1006 of SEQ ID NO:81, b is an integer of 15 to 1020, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:8, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 72

[0312] The translated polypeptide of this contig has a high degree of identity with the Ob Receptor-Associated Protein deposited as GenBank Accession No. 2266638. No function has been determined for the Ob Receptor-Associated Protein, however it is expressed upon stimulation of the Ob Receptor by Leptin. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

```
SGRGARIMTAMSISFSFG-
GAGLMIMMEGLCALPVNEVPLFVYPI
```

[0313] Polynucleotides encoding these polypeptides are also encompassed by the invention.

[0314] This gene is expressed in T-cells and to a lesser extent in endothelial and bone marrow cells.

[0315] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, acute lymphoblastic leukemia, hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g. immune, T-cells and other blood cells, endothelial cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.192 as residues: Ser-61 to Thr-70.

[0316] The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of leukemia and other disorders of the primary immune system. In addition, since this gene appears to be related to the Ob Receptor-Related Protein, it is likely that this polypeptide is also involved in the Ob/Leptin signal transduction cascade. As a result, this protein may be of use in the molecular diagnosis and therapeutic intervention of obesity and related disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these polynucleotides are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 756 of SEQ ID NO:82, b is an integer of 15 to 770, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 73

[0317] The translation product of this contig has homology with furin, a protein thought to be a key endopeptidase
in the constitutive secretory pathway. The identification and initial characterization of Furin was reported by Takahasi and colleagues (Biochem Biophys Res Commun 1993 September 15;195(2):1019-1026).

[0318] This gene is expressed primarily in neutrophils.

[0319] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the immune system such as allergies, wound healing and antigen recognition. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g. immune tissues, neutrophils and other blood cells, and cancers and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0320] The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of allergies or other immune disorders since neutrophils are an important part of an allergic response. Further, since this protein appears to be related to furin, it can be used diagnostically and therapeutically to treat secretory protein processing disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 467 of SEQ ID NO:83, b is an integer of 15 to 481, where both a and b correspond to the positions of nucleotide residues shown in SEQ D NO:83, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 74

[0321] This gene is expressed in the frontal cortex.

[0322] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, of the motor activity and sensory functions that involve the central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and other tissue of the nervous system, and cancers and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

[0323] The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of neural disorders that affect cognitive functions. Similarly, the protein product of this clone would be useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimer Disease, Parkinsons Disease, Huntington's Disease, Tourette Syndrome, menigitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 630 of SEQ ID NO:84, b is an integer of 15 to 644, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO: 84, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 75

[0324] The translation product of this gene shares sequence homology with inorganic pyrophosphatase which is thought to be important in the catalysis the hydrolysis of diphosphate bonds, chiefly in nucleoside di- and triphosphates and essential enzymes that are important for controlling the cellular levels of inorganic pyrophosphate (PiP). The bovine homolog of this gene has been identified by Yang and Wensel (J. Biol. Chem. 267:24641-24647 (1992)). In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

ABVRXGALSLSVGAACGEWALWQR3RQDSGT
(SEQ ID NO:331)
Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed in osteoclastoma cells and to a lesser extent in epithelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoporosis and other skeletal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, and epithelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:195 as residues: Tyr-22 to Tyr-28, Asp-64 to Lys-77, Pro-86 to Ile-91, Gln-99 to Pro-119, Tyr-169 to Asp-174, Lys-176 to Gly-181, Thr-189 to Asn-202, Lys-233 to Gly-239, Ser-250 to Asp-257.

The tissue distribution in osteoclastoma cells and homology to inorganic pyrophosphatase indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of osteoporosis through the removal of bone by demineralization. Similarly, the polypeptides as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1337 of SEQ ID NO:85, b is an integer of 15 to 1351, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 76

The translation product of this gene shares exact sequence homology with ATP sulfurylase/APS kinase (GenBank Accession No. 2673862) which is thought to be important in biosynthesis of the activated sulfate donor, adenosine 3’-phosphate 5’-phosphosulfate, involves the sequential action of two enzyme activities: ATP sulfurylase, which catalyzes the formation of adenosine 3’-phosphosulfate (APS) from ATP and free sulfate, and APS kinase, which subsequently phosphorylates APS to produce adenosine 3’-phosphate 5’-phosphosulfate. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

```
LSNNAQQWQQRATINVTQAHVSRSNKRQVUGFTGSCTYWL (SEQ ID NO:332),
VNALEKYLVCGERPCITTLGDGNIQIQLNAGQFSPEQ (SEQ ID NO:333),
TCQRNHQIHEGAPLPPFPHYVDALPVLCEQVRDGKLY (SEQ ID NO:334),
FTGIDGYKPEAPELVKTDSDCVDNCQVVQVQKQRD (SEQ ID NO:335),
AETLPAIHKVQNQVQLNEWGLATPLNHRERELQCL (SEQ ID NO:336),
VFVILVATRTIEDKHLDEGTAFALNYEGRB (SEQ ID NO:337),
IGGDLQVLRQRYVWNGLGOLDQYLISPTELEKQFKRMPNADAV (SEQ ID NO:338),
GHALLNQOTHQEGLERGRYRVRLHHFLGSGTRXDGDV (SEQ ID NO:339),
MYAGPETVQWNCRAMWAGANFTIVURGDPANGHPEFTKDDL (SEQ ID NO:340),
LMAPQLLILEIVFPPRVAAYNNKKEEHRMDYDSEH (SEQ ID NO:341) or,
GMAPKAWTVLHTEYKLE (SEQ ID NO:342).
```

expression of this gene product in osteoclastoma cells would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues such as arthritis, trauma, tendinitis, chondromalacia, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormali-
and for diagnosis of diseases and conditions, which include, but are not limited to, antibiotic resistant bacterial infections, osteoarthritis and other autoimmune diseases, or skeletal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or skeletal structure expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., bone, and developmental tissues, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in

SEQ ID NO. 196 as residues: Asn-15 to Trp-20,
Ser-36 to Gly-41, Pro-103 to Val-110, Pro-134 to Arg-143, Leu-173 to Arg-176,
Ser-190 to Ala-197, His-314 to Arg-319, Arg-354 to Asn-362, Aas-391 to Arg-397, Glu-402 to Asp-409, Asp-434 to Leu-439, Glu-441 to Arg-446, Gly-455 to Asp-462, Pro-528 to His-541, Asn-566 to Arg-571, Tyr-574 to Glu-581, Thr-589 to Glu-603.

Features of Protein Encoded by Gene No: 77

[0033] This polypeptide is identical to the SLP-76-associated protein reported by Musci and colleagues (J. Biol. Chem. 272 (18), 11674-11677 (1997)) and to the FYB protein reported by da Silva and coworkers (Proc. Natl. Acad. Sci. U.S.A. (1997) In press). These proteins have been reported to be novel T-cell Proteins which bind FYN and SLP-76 and regulate IL-2 production. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

RITDWFQKLNLSTGSGSYGTTEKTTAVIEKYLSDKLKEDSLGAPSKPIED
DQVYDVARQQDSHQSGQGQIPPPPDYDDFYQEEADGFPPAP
FRQLMDFVYIQTSDFVVNPSSQDONGKANTEEDKELKELKEKQK

[0033] The tissue distribution and homology to A TP sulfurylase/APS kinase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment or detection of autoimmune diseases. Similarly, the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues such as arthritis, trauma, tendinitis, chondromalacia, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenital familial osteoarthrisis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2513 of SEQ ID NO:86, b is an integer of 15 to 2527, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where the b is greater than or equal to a+14.

-continued

KEKDFKKKPKYDGEIRVLYSTKVGTTEITTETKQMTSTLQDVHESLVLVQT
TGDTKLCLHRHETGQTVYRLGLACNDQGYETTDDAGCISYMD (SEQ ID NO:343).

[0035] This gene is expressed in CD34 positive cells (hematopoietic progenitor cells) and to a lesser extent in adult spleen derived from a chronic lymphocytic leukemia patient.

[0036] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, chronic lymphocytic leukemia; hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., T-cells and other blood cells, bone marrow, hematopoietic cells, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Further, nucleic acids and polypeptides of the present invention are useful both diagnostically and therapeutically in the intervention of immune and other disorders in which the
ability to alter IL-2 expression is desired. Preferred epitopes include those comprising a sequence shown in
above tissues or cells, particularly of the immune system—most notably the T-cell compartment, expression of this
gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and
other blood cells, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e.,
the expression level in healthy tissue or bodily fluid from an
individual not having the disorder.

The tissue distribution in immune cells indicates that
polynucleotides and polypeptides corresponding to this
gene are useful for the treatment of a variety of hematopoietic
disorders. The noted expression of this gene in hematopoietic
progenitor cell—as determined by its expression on
CD34 positive hematopoietic stem and progenitor cells—
indicates that it plays a critical role in the expansion or
proliferation of hematopoietic stem/progenitor cells, as well
as in the differentiation of the various blood cell lineages.
Thus it could be useful in the reconstitution of the hematopoietic
system of patients with leukemias and other hematopoietic
diseases. Protein, as well as, antibodies directed
against the protein may show utility as a tissue-specific
marker and/or immunotherapy target for the above-listed
tissues. Many polynucleotide sequences, such as EST
sequences, are publicly available and accessible through
database searches. Some of these sequences are related to
SEQ ID NO:87 and may have been publicly available prior
to conception of the present invention. Preferably, such
related polynucleotides are specifically excluded from the
scope of the present invention. To list every related sequence
would be cumbersome. Accordingly, preferably excluded
from the present invention are one or more polynucleotides
comprising a nucleotide sequence described by the general
formula of a-b, where a is any integer between 1 and 2552 of
SEQ ID NO:87, b is an integer of 15 to 2566, where both a
and b correspond to the positions of nucleotide residues
shown in SEQ ID NO:87, and where the b is greater than or
equal to a+14.

Features of Protein Encoded by Gene No: 78

This gene is homologous to heparin cofactor II (HCII) which is a 66-kDa plasma glycoprotein that inhibits
thrombin rapidly in the presence of dermatan sulfate or heparin.

This gene is expressed in apoptotic and anergic
T-cells.

Therefore, polynucleotides and polypeptides of the
invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample
and for diagnosis of diseases and conditions, which include,
but are not limited to, thrombopenia T-cell lymphomas;
Hodgkin’s lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the
gene at significantly higher or lower levels may be routinely
detected in certain tissues and cell types (e.g., T-cells and
other blood cells, and lymphoid tissue, and cancerous
and wounded tissues) or bodily fluids (e.g., lymph, serum,
plasma, urine, synovial fluid or spinal fluid) or another tissue
or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e.,
the expression level in healthy tissue or bodily fluid from an
individual not having the disorder.

The homology to heparin cofactor II (HCII) and the
tissue distribution indicates that polynucleotides and
polypeptides corresponding to this gene are useful for the
treatment and diagnosis of hematopoietic disorders
particularly in thrombopoeisis, most notably of the T-cell
compartment. This could include immune modulation,
inflammation, immune surveillance, graft rejection, and
autoimmunity. Protein, as well as, antibodies directed
against the protein may show utility as a tissue-specific
marker and/or immunotherapy target for the above-listed
tissues. Many polynucleotide sequences, such as EST
sequences, are publicly available and accessible through
database searches. Some of these sequences are related to
SEQ ID NO:88 and may have been publicly available prior
to conception of the present invention. Preferably, such
related polynucleotides are specifically excluded from the
scope of the present invention. To list every related sequence
would be cumbersome. Accordingly, preferably excluded
from the present invention are one or more polynucleotides
comprising a nucleotide sequence described by the general
formula of a-b, where a is any integer between 1 and 526 of
SEQ ID NO:88, b is an integer of 15 to 540, where both a
and b correspond to the positions of nucleotide residues
shown in SEQ ID NO:88, and where the b is greater than or
equal to a+14.

Features of Protein Encoded by Gene No: 79

The translation product of this gene shares sequence homology with a mouse protein believed to rep-
 resent an integral membrane protein.

This gene is expressed in fetal cochlea and epid-
 idymus and to a lesser extent in adult spleen and osteoclas-
toma.

Therefore, polynucleotides and polypeptides of the
invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample
and for diagnosis of diseases and conditions, which include,
but are not limited to, osteoclastoma; disorders of the inner ear; male fertility disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the inner ear; male reproductive tract; bone; and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cochlea, epididymus and other reproductive tissue, spleen, immune tissue, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level; i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 199 as residues: Lys-13 to Gly-23, Cys-38 to Asp-43, Gly-48 to Trp-53, Cys-223 to Ile-237, Ile-240 to Ser-246.

[0345] The tissue distribution in reproductive tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of hearing and fertility disorders. Likewise, it may have a role in the modulation of immune function and in the treatment of osteoporosis. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1849 of SEQ ID NO:89, b is an integer of 15 to 1863, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 80

[0346] The translation product of this gene shares sequence homology with reticulocabin which is thought to be important in the binding of calcium, particularly within the endoplasmic reticulum.

[0347] This gene is expressed in endothelial cells and stromal cells and to a lesser extent in osteoblasts, osteoclasts, and T-cells.

[0348] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoporosis; osteoelasticomas; T-cell lymphomas; Hodgkin’s disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vasculature, bone, and immune systems - particularly the T-cell compartments, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, stromal cells, bone, T-cells and other blood cells, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 200 as residues: Lys-20 to Arg-27, Pro-32 to Asp-48, Leu-64 to Arg-72, Asp-108 to Lys-114, Glu-120 to Thr-133, Asp-139 to Phe-147, Thr-196 to Ala-204, Tyr-218 to Glu-228, Val-230 to Gln-236, Arg-241 to Lys-255, Glu-276 to Lys-287.

[0349] The tissue distribution and homology to reticulocabin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of bone disorders such as osteoporosis; the diagnosis and treatment of T-cell lymphomas and Hodgkin’s lymphoma; and the treatment of diseases and defects of the vasculature, such as vascular leak syndrome and aberrant angiogenesis that accompanies tumor growth. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2464 of SEQ ID NO:90, b is an integer of 15 to 2478, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 81

[0350] The translation product of this gene shares sequence homology with a family of peptide transport genes—particularly the AIPTR2-B gene from Arabidopsis— which are thought to be important in the uptake of small peptides.

[0351] This gene is expressed in a number of fetal tissues, most notably lung, brain, cochlea, and liver/spleen, and to a lesser extent in osteoclastoma and endometrial tumors.

[0352] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoclastoma; endometrial tumors;
cancer; leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone and endometrium, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., fetal tissue, pulmonary tissue, bone, brain and other tissue of the nervous system, coagula, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.201 as residues: Lys-186 to Asn-199, Pro-202 to Ala-207.

[0353] The tissue distribution in fetal tissues combined with the homology to peptide transport proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for the control of cell proliferation, owing to its strong expression in fetal tissues undergoing active cell division, as well as its expression in a variety of tumors or cancers of adult tissues. Potentially, it may regulate the uptake of peptides that stimulate cell proliferation. This gene product may also be useful in stimulating the uptake of a variety of peptide-based drug compounds. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is an integer between 1 to 2044 of SEQ ID NO:91, b is an integer of 15 to 2058, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 82

[0354] This gene is expressed in fetal liver and spleen and to a lesser extent in endothelial cells.

[0355] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer and tumors of a hematopoietic and/or endothelial cell origin; leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and/or vasculature, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, endothelial cells, vascular tissue, and tissue and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, bile, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.202 as residues: Met-1 to Asp-9, Arg-66 to Gly-76, Asp-164 to Arg-171.

[0356] The tissue distribution in immune tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of disorders of the immune system. Expression of this gene product in both fetal liver/ spleen and endothelial cells indicates that it may be expressed in the hemagloblast, the progenitor cell for both the immune system and the vasculature. Thus, it is most likely expressed in hematopoietic stem cells, and may be useful for the expansion of hematopoietic stem and progenitor cells in conjunction with cancer treatment for a variety of leukemias. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1397 of SEQ ID NO:92, b is an integer of 15 to 1411, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 84

[0357] The translation product of this gene shares sequence homology with NADH dehydrogenase which is thought to be important in cellular metabolism.

[0358] This gene is expressed in fetal duodenum and to a lesser extent in T-cells and hypothalamus.

[0359] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases affecting cellular metabolism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissue, T-cells and other blood cells, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,
relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO.204 as residues: Pro-27 to Gln-32, Arg-42 to Gln-51.

[0360] The tissue distribution and homology to NADH dehydrogenase indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases involving cellular metabolism. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 743 of SEQ ID NO:94, b is an integer of 15 to 757, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:94, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 85

[0361] The translation product of this gene shares sequence homology with I-TRAF, a novel TNF receptor associated factor (TRAF)-interacting protein that regulates TNF receptor-mediated signal transduction. This protein is thought to be important in regulating the cellular response to tumor necrosis factor (TNF), which is an important mediator of inflammation.

[0362] This gene is expressed in endothelial cells and to a lesser extent in glioblastoma and osteoblastoma.

[0363] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation; glioblastoma and osteoblastoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, bone, and glial cells and tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissues or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in

SEQ ID NO. 205 as residues: Glu-15 to Thr-22,
Glu-46 to Leu-62, Arg-103 to Glu-119, Gln-127 to

-Glu-132, Asn-152 to Trp-158, Gln-191 to Gln-210,
Glu-264 to Thr-271, Tyr-282 to Leu-288, Trp-319 to
Thr-331, Glu-335 to Ser-348, Ser-353 to Ser-358,
Asp-382 to Asn-392.

[0364] The tissue distribution and homology to the I-TRAF protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of inflammatory diseases, including rheumatoid arthritis, sepsis, inflammatory bowel disease, and psoriasis, particularly where tumor necrosis factor is known to be involved. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:95 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2380 of SEQ ID NO:95, b is an integer of 15 to 2394, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 86

[0365] This gene has homology with a candidate gene involved in X-linked Retinopathy reported by Wong and colleagues (Genomics 15:467-471 (1993)).

[0366] This gene is expressed in a T-cell line.

[0367] Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and autoimmune diseases; T-cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissues or bodily fluid from an individual not having the disorder.

[0368] The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of inflammatory disorders such as sepsis, inflammatory bowel disease, psoriasis, and arthritis.
riasis, and rheumatoid arthritis as well as autoimmune disease such as lupus. It could also be useful in immune modulation and in the process of immune surveillance. The present invention can be used diagnostically and therapeutically to treat X-linked Retinopathy. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-mentioned tissues. Many nucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related nucleotide sequences are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more nucleotide sequences comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 658 of SEQ ID NO:96, b is an integer of 15 to 672, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where the b is greater than or equal to a+14.

Features of Protein Encoded by Gene No: 87

This gene is expressed in human brain tissue. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, brain disorders; neurodegenerative disorders; tumors of a brain origin. Similarly, polynucleotides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and cancers and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:211 as residues: Cys-32 to Tyr-38. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:207 as residues: Cys-32 to Tyr-38.

The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of CNS disorders such as epilepsy, paranoia, depression, Alzheimer’s disease, and schizophrenia. It could be useful in the survival and/or proliferation of neurons and could effect neuronal regeneration. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above-listed tissues. Many nucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related nucleotide sequences are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more nucleotide sequences comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1665 of SEQ ID NO:11, b is an integer of 15 to 1679, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where the b is greater than or equal to a+14.

Many nucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related nucleotide sequences are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more nucleotide sequences comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1405 of SEQ ID NO:97, b is an integer of 15 to 1419, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:97, and where the b is greater than or equal to a+14.
<table>
<thead>
<tr>
<th>Gene No.</th>
<th>cDNA Clone ID</th>
<th>ATCC Deposit No and Date</th>
<th>Vector</th>
<th>NT Seq NO: X</th>
<th>Total NT Seq 1</th>
<th>5' NT of Clone Seq 1</th>
<th>3' NT of Clone Seq 1</th>
<th>5' NT of Start Condon</th>
<th>5' NT of First AA of Signal Pep</th>
<th>AA Seq NO: Y</th>
<th>First AA of Sig Pep</th>
<th>Last AA of Sig Pep</th>
<th>First AA of Secreted Portion</th>
<th>Last AA of ORF</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>HAFBL78</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>14</td>
<td>2061</td>
<td>882</td>
<td>2061</td>
<td>900</td>
<td>900</td>
<td>124</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>HASAV70</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>15</td>
<td>1412</td>
<td>10</td>
<td>733</td>
<td>103</td>
<td>103</td>
<td>125</td>
<td>1</td>
<td>20</td>
<td>21</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>HBNAF22</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>16</td>
<td>1052</td>
<td>276</td>
<td>880</td>
<td>538</td>
<td>538</td>
<td>126</td>
<td>1</td>
<td>23</td>
<td>24</td>
<td>63</td>
</tr>
<tr>
<td>7</td>
<td>HBNBL77</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>17</td>
<td>683</td>
<td>683</td>
<td>181</td>
<td>181</td>
<td>127</td>
<td>1</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>HCDDR890</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>18</td>
<td>1054</td>
<td>86</td>
<td>1007</td>
<td>86</td>
<td>86</td>
<td>128</td>
<td>1</td>
<td>23</td>
<td>24</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>HCEEF50</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>19</td>
<td>1393</td>
<td>132</td>
<td>1393</td>
<td>192</td>
<td>192</td>
<td>129</td>
<td>1</td>
<td>17</td>
<td>18</td>
<td>57</td>
</tr>
<tr>
<td>10</td>
<td>HCEMU42</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>20</td>
<td>1215</td>
<td>277</td>
<td>1070</td>
<td>401</td>
<td>401</td>
<td>130</td>
<td>1</td>
<td>18</td>
<td>19</td>
<td>216</td>
</tr>
<tr>
<td>11</td>
<td>HCENE16</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>21</td>
<td>2042</td>
<td>614</td>
<td>2011</td>
<td>793</td>
<td>793</td>
<td>131</td>
<td>1</td>
<td>26</td>
<td>27</td>
<td>49</td>
</tr>
<tr>
<td>12</td>
<td>HMSIJ74</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>22</td>
<td>1872</td>
<td>21</td>
<td>1872</td>
<td>69</td>
<td>69</td>
<td>132</td>
<td>1</td>
<td>23</td>
<td>24</td>
<td>68</td>
</tr>
<tr>
<td>13</td>
<td>HCUH515</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>ZAP Express</td>
<td>23</td>
<td>289</td>
<td>1</td>
<td>289</td>
<td>89</td>
<td>89</td>
<td>133</td>
<td>1</td>
<td>29</td>
<td>30</td>
<td>52</td>
</tr>
<tr>
<td>14</td>
<td>HE2DE47</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>24</td>
<td>3533</td>
<td>2821</td>
<td>3532</td>
<td>808</td>
<td>808</td>
<td>134</td>
<td>1</td>
<td>30</td>
<td>31</td>
<td>540</td>
</tr>
<tr>
<td>14</td>
<td>HE2DE47</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>25</td>
<td>1154</td>
<td>435</td>
<td>115</td>
<td>515</td>
<td>515</td>
<td>209</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>81</td>
</tr>
<tr>
<td>15</td>
<td>HKMLH01</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>pBluescript</td>
<td>25</td>
<td>1348</td>
<td>171</td>
<td>907</td>
<td>196</td>
<td>196</td>
<td>135</td>
<td>1</td>
<td>26</td>
<td>27</td>
<td>57</td>
</tr>
<tr>
<td>15</td>
<td>HE5DG34</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>100</td>
<td>734</td>
<td>25</td>
<td>734</td>
<td>295</td>
<td>295</td>
<td>210</td>
<td>1</td>
<td>36</td>
<td>37</td>
<td>49</td>
</tr>
<tr>
<td>16</td>
<td>HE9DG49</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>26</td>
<td>717</td>
<td>1</td>
<td>717</td>
<td>70</td>
<td>70</td>
<td>136</td>
<td>1</td>
<td>27</td>
<td>28</td>
<td>201</td>
</tr>
<tr>
<td>16</td>
<td>HE9DG49</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>101</td>
<td>713</td>
<td>17</td>
<td>713</td>
<td>78</td>
<td>78</td>
<td>211</td>
<td>1</td>
<td>28</td>
<td>29</td>
<td>203</td>
</tr>
<tr>
<td>17</td>
<td>HELBA06</td>
<td>79723 03/07/97 200071 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>27</td>
<td>1099</td>
<td>1</td>
<td>1099</td>
<td>38</td>
<td>38</td>
<td>137</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>216</td>
</tr>
<tr>
<td>Gene No.</td>
<td>ATCC No.</td>
<td>cDNA Clone ID</td>
<td>ATCC Deposit Nr and Date</td>
<td>NT SEQ NO: X</td>
<td>Total NT Seq.</td>
<td>5' UT of Clone Seq.</td>
<td>3' UT of Clone Seq.</td>
<td>5' UT of Stret Condon</td>
<td>First AA of Signal Pep</td>
<td>AA SEQ NO: Y</td>
<td>First AA of Sig Pep</td>
<td>Last AA of Sig Pep</td>
<td>First AA of Secreted Portion</td>
<td>Last AA of ORF</td>
</tr>
<tr>
<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>---------------</td>
</tr>
<tr>
<td>17</td>
<td>97923</td>
<td>HELBA06</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>102</td>
<td>1080</td>
<td>1</td>
<td>1080</td>
<td>149</td>
<td>149</td>
<td>212</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>97923</td>
<td>HELBW28</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>29</td>
<td>756</td>
<td>62</td>
<td>756</td>
<td>294</td>
<td>294</td>
<td>139</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>97923</td>
<td>HETHN28</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>30</td>
<td>2100</td>
<td>408</td>
<td>2093</td>
<td>496</td>
<td>496</td>
<td>140</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>97923</td>
<td>HFCDK17</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>31</td>
<td>1448</td>
<td>475</td>
<td>1392</td>
<td>567</td>
<td>567</td>
<td>141</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>97923</td>
<td>HFEAF41</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>32</td>
<td>456</td>
<td>1</td>
<td>409</td>
<td>21</td>
<td>21</td>
<td>142</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>97923</td>
<td>HFKFL13</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>33</td>
<td>1326</td>
<td>1</td>
<td>1322</td>
<td>210</td>
<td>210</td>
<td>143</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>97923</td>
<td>HFSBG13</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>34</td>
<td>710</td>
<td>1</td>
<td>710</td>
<td>242</td>
<td>242</td>
<td>144</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>97923</td>
<td>HFTBE43</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>35</td>
<td>1188</td>
<td>110</td>
<td>1161</td>
<td>178</td>
<td>178</td>
<td>145</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>97923</td>
<td>HFTDJ36</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>36</td>
<td>956</td>
<td>1</td>
<td>938</td>
<td>144</td>
<td>144</td>
<td>146</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>97924</td>
<td>HKTAC77</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>37</td>
<td>1603</td>
<td>974</td>
<td>1581</td>
<td>1104</td>
<td>1104</td>
<td>147</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>97924</td>
<td>HLHSH66</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>pBluescript</td>
<td>38</td>
<td>1089</td>
<td>55</td>
<td>1057</td>
<td>209</td>
<td>209</td>
<td>148</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>97924</td>
<td>HLHSH66</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>pBluescript</td>
<td>39</td>
<td>629</td>
<td>1</td>
<td>629</td>
<td>119</td>
<td>119</td>
<td>149</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>97924</td>
<td>HLQPO98</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Lambda ZAP II</td>
<td>40</td>
<td>1964</td>
<td>408</td>
<td>1793</td>
<td>581</td>
<td>581</td>
<td>150</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>97924</td>
<td>HLTXB31</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>41</td>
<td>1522</td>
<td>13</td>
<td>1123</td>
<td>126</td>
<td>126</td>
<td>151</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>97924</td>
<td>HLTC63</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>42</td>
<td>875</td>
<td>1</td>
<td>875</td>
<td>43</td>
<td>43</td>
<td>152</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>97924</td>
<td>HMKAH44</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>pSport I</td>
<td>43</td>
<td>843</td>
<td>1</td>
<td>843</td>
<td>171</td>
<td>171</td>
<td>153</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>97924</td>
<td>HMQA64</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>44</td>
<td>489</td>
<td>3</td>
<td>489</td>
<td>55</td>
<td>55</td>
<td>154</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>97924</td>
<td>HMQA64</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>45</td>
<td>534</td>
<td>1</td>
<td>534</td>
<td>17</td>
<td>17</td>
<td>155</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>97924</td>
<td>HOABJ65</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>46</td>
<td>1374</td>
<td>1</td>
<td>1374</td>
<td>15</td>
<td>15</td>
<td>156</td>
<td>1</td>
</tr>
<tr>
<td>37</td>
<td>97924</td>
<td>HODCL38</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>47</td>
<td>996</td>
<td>1</td>
<td>996</td>
<td>269</td>
<td>269</td>
<td>157</td>
<td>1</td>
</tr>
<tr>
<td>38</td>
<td>97924</td>
<td>HODCV74</td>
<td>03/07/97</td>
<td>03/07/97</td>
<td>55229</td>
<td>Uni-ZAP XR</td>
<td>48</td>
<td>851</td>
<td>99</td>
<td>822</td>
<td>170</td>
<td>170</td>
<td>158</td>
<td>1</td>
</tr>
<tr>
<td>Gene No.</td>
<td>cDNA Clone ID</td>
<td>ATCC Deposit No. and Date</td>
<td>NT SEQ ID NO.:</td>
<td>NT of First AA of Signal Pep</td>
<td>AA SEQ ID NO.:</td>
<td>First AA of Sig Pep</td>
<td>Last AA of Sig Pep</td>
<td>First AA of Secreted Portion</td>
<td>Last AA of ORF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<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></td>
</tr>
<tr>
<td>39</td>
<td>HODCZ16</td>
<td>97924</td>
<td>03/07/97</td>
<td>49 2020 569 2020 638 638</td>
<td>159 1 17 18 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>HTOEU03</td>
<td>97924</td>
<td>03/07/97</td>
<td>50 2432 848 2432 99 99</td>
<td>160 1 19 20 323</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>HTOEU03</td>
<td>97924</td>
<td>03/07/97</td>
<td>105 2435 849 2435 928 928</td>
<td>215 1 31 32 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>HPBCF74</td>
<td>pBluescript SK</td>
<td>03/07/97</td>
<td>51 2340 1627 2340 150 150</td>
<td>161 1 60 61 320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>HPBCF74</td>
<td>97924</td>
<td>03/07/97</td>
<td>106 805 92 791 239 239</td>
<td>216 1 21 22 83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>HPMBU33</td>
<td>97924</td>
<td>03/07/97</td>
<td>52 601 188 601 432 432</td>
<td>162 1 31 31 72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>HSAUL66</td>
<td>97924</td>
<td>03/07/97</td>
<td>53 359 1 337 142 142</td>
<td>163 1 18 19 72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>HSIDQ018</td>
<td>97924</td>
<td>03/07/97</td>
<td>54 1144 1 1141 25 25</td>
<td>164 1 30 31 281</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>HSIDQ018</td>
<td>97924</td>
<td>03/07/97</td>
<td>107 1166 21 1166 433 433</td>
<td>217 1 30 31 43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>HSJBB37</td>
<td>97924</td>
<td>03/07/97</td>
<td>55 1560 413 1498 714 714</td>
<td>165 1 31 32 81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>HSJBO79</td>
<td>97924</td>
<td>03/07/97</td>
<td>56 1507 164 608 57 57</td>
<td>166 1 19 20 327</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>HSJBO79</td>
<td>97924</td>
<td>03/07/97</td>
<td>108 586 4 586 35 35</td>
<td>218 1 23 24 184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>HTEGAIL6</td>
<td>97958</td>
<td>03/13/97</td>
<td>57 450 1 450 90 90</td>
<td>167 1 43 44 65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>HTEJN13</td>
<td>97958</td>
<td>03/13/97</td>
<td>58 1147 1 1147 163 163</td>
<td>168 1 15 16 159</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>HTEJN13</td>
<td>97958</td>
<td>03/13/97</td>
<td>59 1134 1 1134 155 155</td>
<td>219 1 19 20 71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>HTHBL86</td>
<td>97958</td>
<td>03/13/97</td>
<td>59 777 1 777 115 115</td>
<td>169 1 18 19 123</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>HTSFO71</td>
<td>97958</td>
<td>03/13/97</td>
<td>60 1991 48 598 52 52</td>
<td>170 1 30 31 129</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>HTSFO71</td>
<td>97958</td>
<td>03/13/97</td>
<td>110 1333 594 1333 829 829</td>
<td>220 1 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>HAPNO80</td>
<td>97958</td>
<td>03/13/97</td>
<td>51 1580 443 1554 114 114</td>
<td>171 1 1 2 372</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>HAUCC47</td>
<td>97958</td>
<td>03/13/97</td>
<td>111 1015 249 708 244 244</td>
<td>221 1 28 29 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>HBMCCL41</td>
<td>97958</td>
<td>03/13/97</td>
<td>62 1117 105 1034 182 182</td>
<td>172 1 28 29 216</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>HCFLD84</td>
<td>pSPORT1</td>
<td>03/13/97</td>
<td>63 361 1 361 97 97</td>
<td>173 1 32 33 55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>HEBEM69</td>
<td>97958</td>
<td>03/13/97</td>
<td>64 1668 1 1638 150 150</td>
<td>174 1 20 21 23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>HE8EZ48</td>
<td>97958</td>
<td>03/13/97</td>
<td>65 1353 35 1303 231 231</td>
<td>175 1 33 34 103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene No.</td>
<td>cDNA Clone ID</td>
<td>ATCC Deposit Nr and Date</td>
<td>Vector</td>
<td>NT SEQ ID NO: X</td>
<td>NT Total Seq: 1</td>
<td>5' NT of Clone Seq: 2</td>
<td>3' NT of Clone Seq: 3</td>
<td>5' NT of Start Condon: 4</td>
<td>5' NT of First AA of Signal Pep: 5</td>
<td>AA SEQ ID NO: Y</td>
<td>First AA of Sig Pep: 6</td>
<td>Last AA of Sig Pep: 7</td>
<td>First AA of Secreted Portion: 8</td>
<td>Last AA of ORF: 9</td>
</tr>
<tr>
<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>------------------</td>
</tr>
<tr>
<td>56</td>
<td>HEBGF33</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>66</td>
<td>1011</td>
<td>655</td>
<td>1011</td>
<td>703</td>
<td>703</td>
<td>176</td>
<td>1</td>
<td>38</td>
<td>39</td>
<td>48</td>
</tr>
<tr>
<td>57</td>
<td>HFEBF41</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>67</td>
<td>1193</td>
<td>267</td>
<td>1090</td>
<td>459</td>
<td>459</td>
<td>177</td>
<td>1</td>
<td>35</td>
<td>36</td>
<td>96</td>
</tr>
<tr>
<td>58</td>
<td>HFRBNU4</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>68</td>
<td>560</td>
<td>1</td>
<td>560</td>
<td>63</td>
<td>63</td>
<td>178</td>
<td>1</td>
<td>29</td>
<td>30</td>
<td>95</td>
</tr>
<tr>
<td>59</td>
<td>HFVGVZ96</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Bluescript</td>
<td>69</td>
<td>1657</td>
<td>765</td>
<td>1581</td>
<td>839</td>
<td>839</td>
<td>179</td>
<td>1</td>
<td>21</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>60</td>
<td>HHGCM76</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>70</td>
<td>711</td>
<td>8</td>
<td>711</td>
<td>270</td>
<td>270</td>
<td>180</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>89</td>
</tr>
<tr>
<td>61</td>
<td>HHGCM76</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>112</td>
<td>711</td>
<td>8</td>
<td>711</td>
<td>270</td>
<td>270</td>
<td>222</td>
<td>1</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>HHGCQ058</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>71</td>
<td>935</td>
<td>111</td>
<td>935</td>
<td>272</td>
<td>272</td>
<td>181</td>
<td>1</td>
<td>19</td>
<td>20</td>
<td>65</td>
</tr>
<tr>
<td>63</td>
<td>HHGCP52</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>72</td>
<td>504</td>
<td>113</td>
<td>484</td>
<td>45</td>
<td>45</td>
<td>182</td>
<td>1</td>
<td>15</td>
<td>16</td>
<td>105</td>
</tr>
<tr>
<td>64</td>
<td>HHGDB72</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>73</td>
<td>620</td>
<td>1</td>
<td>620</td>
<td>96</td>
<td>96</td>
<td>183</td>
<td>1</td>
<td>18</td>
<td>19</td>
<td>132</td>
</tr>
<tr>
<td>65</td>
<td>HHGDI71</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Lambda ZAP II</td>
<td>74</td>
<td>581</td>
<td>156</td>
<td>581</td>
<td>248</td>
<td>248</td>
<td>184</td>
<td>1</td>
<td>32</td>
<td>33</td>
<td>69</td>
</tr>
<tr>
<td>66</td>
<td>HHSDF45</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>75</td>
<td>1843</td>
<td>537</td>
<td>1786</td>
<td>630</td>
<td>630</td>
<td>185</td>
<td>1</td>
<td>27</td>
<td>28</td>
<td>45</td>
</tr>
<tr>
<td>67</td>
<td>HHSEB86</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>76</td>
<td>1441</td>
<td>116</td>
<td>800</td>
<td>167</td>
<td>167</td>
<td>186</td>
<td>1</td>
<td>36</td>
<td>37</td>
<td>65</td>
</tr>
<tr>
<td>68</td>
<td>HAUA83</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>77</td>
<td>910</td>
<td>1</td>
<td>886</td>
<td>253</td>
<td>253</td>
<td>187</td>
<td>1</td>
<td>37</td>
<td>38</td>
<td>49</td>
</tr>
<tr>
<td>69</td>
<td>HJPAZ83</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>78</td>
<td>1076</td>
<td>398</td>
<td>1076</td>
<td>575</td>
<td>223</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>HLDBO49</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>pCMVSport 3.0</td>
<td>79</td>
<td>2776</td>
<td>18</td>
<td>1888</td>
<td>187</td>
<td>187</td>
<td>188</td>
<td>1</td>
<td>14</td>
<td>15</td>
<td>170</td>
</tr>
<tr>
<td>71</td>
<td>HLDBQ19</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>pCMVSport 3.0</td>
<td>79</td>
<td>1487</td>
<td>401</td>
<td>1487</td>
<td>534</td>
<td>534</td>
<td>189</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>132</td>
</tr>
<tr>
<td>72</td>
<td>HLDBQ19</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>pCMVSport 3.0</td>
<td>79</td>
<td>1525</td>
<td>401</td>
<td>1480</td>
<td>534</td>
<td>534</td>
<td>224</td>
<td>1</td>
<td>22</td>
<td>23</td>
<td>66</td>
</tr>
<tr>
<td>Gene No.</td>
<td>cDNA Clone ID</td>
<td>ATCC Deposit No and Date</td>
<td>Vector</td>
<td>NT SEQ ID NO: X</td>
<td>NT Seq.</td>
<td>Clone Seq.</td>
<td>First AA of Signal Pep</td>
<td>AA SEQ ID NO: Y</td>
<td>First AA of Sig Pep</td>
<td>Last AA of Sig Pep</td>
<td>First AA of Secreted Portion</td>
<td>Last AA of ORF</td>
<td></td>
<td></td>
</tr>
<tr>
<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></td>
</tr>
<tr>
<td>70</td>
<td>HMSGT42</td>
<td>97958 03/13/97 200702 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>80 1563 33 1077 40 40 190 1 32 33 92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>HMWTC78</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-Zap XR</td>
<td>81 1020 18 780 238 238 191 1 23 24 176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>HTTCT79</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>82 770 101 770 286 286 192 1 26 27 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>HNGRU84</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>83 481 1 481 58 58 193 1 20 21 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>HNTAC73</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>pCMV5Sport 3.0</td>
<td>84 644 1 623 14 14 194 1 25 26 73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>HOSEI45</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>85 1351 435 1284 98 98 195 1 12 13 289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>HOSEI45</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>85 1350 428 1283 545 225 1 28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>HOSFD58</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>86 2527 290 1747 56 56 196 1 30 31 624</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>HOSFD58</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>86 2527 288 1747 477 477 226 1 32 33 61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>HSAUM95</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>87 2566 1843 2566 251 251 197 1 30 31 649</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>HSAUM95</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>87 1098 375 1098 677 677 227 1 21 22 29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>HSAUR67</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>88 540 1 540 83 83 198 1 32 33 55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>HSKD81</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>89 1863 152 1165 188 188 199 1 11 12 266</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>HSKD81</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>89 1679 152 1166 315 315 228 1 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>HSKDW91</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>90 2478 1149 2449 92 92 200 1 19 20 315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>HTEX50</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>Uni-ZAP XR</td>
<td>91 2088 476 2088 414 414 201 1 20 21 207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>HSKH356</td>
<td>97957 03/13/97 200703 05/22/97</td>
<td>pBluescript</td>
<td>92 1411 345 1411 157 157 202 1 69 70 195</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 1 summarizes the information corresponding to each “Gene No.” described above. The nucleotide sequence identified as “NT SEQ ID NO:X” was assembled from partially homologous (“overlapping”) sequences obtained from the “cDNA clone ID” identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in “ATCC Deposit No:Z and Date.” Some of the deposits contain multiple different clones corresponding to the same gene. “Vector” refers to the type of vector contained in the cDNA Clone ID.

“Total NT Seq.” refers to the total number of nucleotides in the contig identified by “Gene No.” The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as “5’ NT of Clone Seq.” and the “3’ NT of Clone Seq.” of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as “5’ NT of Start Codon.” Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as “5’ NT of First AA of Signal Pep.”

The translated amino acid sequence, beginning with the methionine, is identified as “AA SEQ ID NO:Y,” although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as “First AA of Sig Pep” and “Last AA of Sig Pep.” The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as “Predicted First AA of Secreted Portion.” Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as “Last AA of ORF.”

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence
diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

[0378] Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

[0379] The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

[0380] Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

[0381] The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

[0382] The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequence, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[0383] The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

[0384] Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heijne, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -3 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heijne, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

[0385] In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heijne are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

[0386] As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or -5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

[0387] Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polyonucleotide and Polypeptide Variants

[0388] “Variant” refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

[0389] By a polynucleotide having a nucleotide sequence at least, for example, 95% “identical” to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each
100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence, may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

**[0390]** As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

**[0391]** If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

**[0392]** For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

**[0393]** By a polypeptide having an amino acid sequence at least, for example, 95% “identical” to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

**[0394]** As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

**[0395]** If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the
above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

[0396] For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of the present invention.

[0397] The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

[0398] Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985)). These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

[0399] Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988)).

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine II-1a. They used random mutagenesis to generate over 3,500 individual II-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract). In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989)). The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid
substitutions. The authors further indicate which amino acid changes are likely to be permissible at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp; and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

[0046] Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

[0047] For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

[0048] In the present invention, a “polynucleotide fragment” refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment “at least 20 nt in length,” for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

[0049] Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context “about” includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

[0040] In the present invention, a “polypeptide fragment” refers to a short amino acid sequence contained in SEQ ID NO:x or encoded by the cDNA contained in the deposited clone. Protein fragments may be “free-standing,” or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context “about” includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

[0041] Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

[0042] Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within consensus domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

[0043] Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.
In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3908-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Pat. No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Buttle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.) As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab)2 fragments) which are capable of specifically binding to protein. Fab and F(ab)2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., 3. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Examples of domains that can be fused to polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827, Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-0 464 533 (Canadian counterpart 2043869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A-0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, Calif., 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984)).
Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention. Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UAG or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera SF9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.


Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an internal modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ D NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using
a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

[0437] Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosome spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verna et al., “Human Chromosomes: a Manual of Basic Techniques,” Pergamon Press, New York (1988).

[0438] For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

[0439] Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

[0440] Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

[0441] Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

[0442] In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix—see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1560 (1991)) or to the mRNA itself (antisense—Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, Fla. (1988)). Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

[0443] Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

[0444] The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual’s genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of “Dog Tags” which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

[0445] The polynucleotides of the present invention can also be used as an alternative to RFLP by determining the actual base-by-base DNA sequence of selected portions of an individual’s genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

[0446] Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlith, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

[0447] There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for
example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

[0448] In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to “subtract-out” known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a “gene chip” or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

[0449] Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

[0450] A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunosassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunosassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0451] In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

[0452] A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S. W. Burchiel et al., “Immunopharmaceuticals of Radiolabeled Antibodies and Their Fragments.” (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S. W. Burchiel and B. Rhodes, eds., Masson Publishing Inc. (1982)).

[0453] Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

[0454] Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

[0455] Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

[0456] At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

[0457] The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

[0458] A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be
genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, DiGeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison’s Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic eczephalomyelitis, glomerulonephritis, Goodpasture’s Syndrome, Graves’ Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter’s Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn’s disease, or resulting from over production of cytokines (e.g., TNF or Interleukin).

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvis, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hyper-
gammaglobulinemia, lymphoproliferative disorders, paraproteinemia, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

**Infectious Disease**

[0470] A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

[0471] Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegaviruses (e.g. Paramyxoviridae, Morbillivirus, Rhabdoviridae, Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's vari, and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

[0472] Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergilllosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatococyes, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysiplothrix, Helicobacter, Legionelliosis, Leptospirosis, Listeria, Mycoplasma, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menococccal, Pasteurellocella Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Richettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cut-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatococces), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

[0473] Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthesis, Leishmaniasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Tromboculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

[0474] Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

**Regeneration**

[0475] A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997)). The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

[0476] Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

[0477] Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage.
Specific diseases that could be treated include tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial, and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic (See, Coligan et al., Current Protocols in Immunology (2):Chapter 5 (1991)). Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeasts, Drosophila, or E. coli. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

Preferably, a polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian char-
acteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

[0492] A polypeptide or polynucleotide of the present invention may be used to change a mammal’s mental state or physical state by influencing biorhythms, cardiaic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilites (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

[0493] A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

[0494] Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

[0495] Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5’ Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3’ Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

[0496] Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5’ Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3’ Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

[0497] Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5’ Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3’ Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

[0498] Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

[0499] Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

[0500] A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5’ Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3’ Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

[0501] A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

[0502] Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

[0503] Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a CDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

[0504] Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

[0505] Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

[0506] Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

[0507] A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

[0508] A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

[0509] A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a CDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in
Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

[0510] Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

[0511] A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO: X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0512] The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

[0513] Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO: X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0514] The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

[0515] Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO: X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

[0516] Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO: Y wherein Y is any integer as defined in Table 1.

[0517] Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO: Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO: Y in Table 1.

[0518] Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO: Y.

[0519] Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO: Y.

[0520] Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO: Y.

[0521] Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0522] Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0523] Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0524] Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA
Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0525] Also preferred is an isolated polypeptide comprising an amino acid sequence that is at least 90% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0526] Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0527] Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

[0528] Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0529] Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

[0530] Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0531] Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

[0532] Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0533] In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

[0534] Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0535] Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

[0536] Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

[0537] Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated
nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

[0538] Also preferred is a method of making an isolated polypeptide comprising cultivating this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

[0539] Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

[0540] Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

EXAMPLES

Example 1
Isolation of a Selected cDNA Clone From the Deposited Sample

[0541] Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector “Lambda Zap,” the corresponding deposited clone is in “pBluescript.”

<table>
<thead>
<tr>
<th>Vector Used to Construct Library</th>
<th>Corresponding Deposited Plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>lambda Zap</td>
<td>pBluescript (pBS)</td>
</tr>
<tr>
<td>Uni-Zap XR</td>
<td>pBluescript (pBS)</td>
</tr>
<tr>
<td>Zap Express</td>
<td>pBK</td>
</tr>
<tr>
<td>lfnmid BA</td>
<td>pBlasmid BA</td>
</tr>
<tr>
<td>pSport1</td>
<td>pSport1</td>
</tr>
</tbody>
</table>

[0542] Vectors Lambda Zap (U.S. Pat. Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Pat. Nos. 5,128,256 and 5,286,636), Zap Express (U.S. Pat. Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, Calif., 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polynucleotides to the T7 and T3 primer sequences which flank the polynucleotides ("S" is for Sac and “K” is for Kpn which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue is initiated from the f1 ori generates sense strand DNA and in the other, antisense.

[0543] Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, Md. 20879. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafnmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, Calif. 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

[0544] The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNA's, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

[0545] Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNA's cited for
that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

[0546] Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with \( ^{32} \text{P} \)-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, N.Y. (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edt., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104, or other techniques known to those of skill in the art.

[0547] Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under reaction conditions, for instance, in 25 \( \mu \)l of reaction mixture with 0.5 \( \mu \)g of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl\(_2\), 0.01% (w/v) gelatin, 20 \( \mu \)M each of dATP, dCTP, dGTP, dTMP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94° C. for 1 min; annealing at 55° C. for 1 min; elongation at 72° C. for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

[0548] Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993)).

[0549] Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

[0550] This above method starts with total RNA isolated from the desired source, although poly-A+RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

[0551] This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2

Isolation of Genomic Clones Corresponding to a Polynucleotide

[0552] A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X, according to the method described in Example 1. (See also, Sambrook.)

Example 3

Tissue Distribution of Polypeptide

[0553] Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with \( ^{32} \text{P} \) using the rediprime\textsuperscript{TM} DNA labeling system (Amersham Life Sciences), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100\textsuperscript{TM} column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

[0554] Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb\textsuperscript{TM} hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization a \( \alpha \)-washing, the blots are mounted and exposed to film at -70° C. overnight, and the films developed according to standard procedures.

Example 4

Chromosomal Mapping of the Polynucleotides

[0555] An oligonucleotide primer set is designed according to the sequence at the 5' and of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95° C.; 1 minute, 56° C.; 1 minute, 70° C. This cycle is repeated 32 times
followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5% agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5

**Bacterial Expression of a Polypeptide**

[0556] A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5’ and 3’ ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5’ end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, Calif.) This plasmid vector encodes antibiotic resistance (Amp’), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

[0557] The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the *E. coli* strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lac repressor and also confers kanamycin resistance (Kan’). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

[0558] Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 μg/ml) and Kan (25 μg/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D. 600) of between 0.4 and 0.6. IPTG (Isopropyl-β-D-thiogalactopyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lac repressor, clearing the P/O leading to increased gene expression.

[0559] Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000×g). The cell pellet is solubilized in the chaotropic agent 6 M guanidine-HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilotri-acetic acid (“Ni-NTA”) affinity resin column (available from QIAGEN, Inc., supra). Proteins with a 6×His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QiAexpressionist (1995) QIAGEN, Inc., supra).

[0560] Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl, pH 5, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

[0561] The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: reconstitution using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at –80°C.

[0562] In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on Feb. 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Dalgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, Md.). The promoter sequence and operator sequences are made synthetically.

[0563] DNA can be inserted into the pHEs by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5’ primer) and XbaI, BamHI, XhoI, or Asp718 (3’ primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

[0564] The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

**Example 6**

**Purification of a Polypeptide from an Inclusion Body**

[0565] The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

[0566] Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Hercules Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

[0567] The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or AVP
The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

[0568] The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000g centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

[0569] Following high speed centrifugation (30,000g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium phosphate, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

[0570] To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 μm membrane filter with appropriate surface area (e.g., Filtrol), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded on to a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the eluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

[0571] Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros IQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A280 monitoring of the eluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

[0572] The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Comassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7
Cloning and Expression of a Polypeptide in a Baculovirus Expression System

[0573] In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 (“SV40”) is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from E. coli under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

[0574] Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcD1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

[0575] Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., “A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures,” Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

[0576] The amplified fragment is isolated from a 1% agarose gel using a commercially available kit (“GeneClean,” BIO 101 Inc., La Jolla, Calif.). The fragment is then digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

[0577] The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit (“GeneClean” BIO 101 Inc., La Jolla, Calif.).

[0578] The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. E. coli HB101 or other suitable E. coli hosts such as XL-1 Blue (Strategene Cloning Systems, La Jolla, Calif.) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

[0579] Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA (“BaculoGold™ baculovirus DNA,” Pharmingen, San Diego, Calif.), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace’s medium (Life Technologies Inc., Gaithersburg, Md.). Afterwards, 10 μl Lipofectin plus 90 μl Grace’s medium are added, mixed and incubated for 15 minutes at room temp-
perature. Then the transfection mixture is added drop-wise to S9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace’s medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace’s insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C. for four days.

[0580] After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, supra. An agarose gel with “Blue Gal” (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a “plaque assay” of this type can also be found in the user’s guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipetter (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 μl of Grace’s medium and the suspension containing the recombinant baculovirus is used to infect S9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

[0581] To verify the expression of the polypeptide, S9 cells are grown in Grace’s medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection (“MOI”) of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, Md.). After 42 hours, 5 μCi of 35S-methionine and 5 μCi 35S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

[0582] Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8
Expression of a Polypeptide in Mammalian Cells

[0583] The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV, and the early promoter of the cyto-megalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

[0584] Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H4 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

[0585] Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosone. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

[0586] The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991)). Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991), Begg et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

[0587] Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV enhancer (Boshart et al., Cell 41:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3′ intron, the polyadenylation and termination signal of the rat preproen- sin gene, and the mouse DHFR gene under control of the SV40 early promoter.

[0588] Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphatases by procedures known in the art. The vector is then isolated from a 1% agarose gel.

[0589] A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g.,WO 96/34891.)

[0590] The amplified fragment is isolated from a 1% agarose gel using a commercially available kit (“Geneclean”, BIO 101 Inc., La Jolla, Calif.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

[0591] The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel.
The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB-101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

[0592] Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μg of the expression plasmid pC6 is cotransfected with 0.5 μg of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 μg/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μM, 2 μM, 5 μM, 10 μM, 20 μM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-200 μM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9

Protein Fusions

[0593] The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptide to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein provides purification. (See Example 5; see also EP 394,827; Trauernicht et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

[0594] Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5’ and 3’ ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

[0595] For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3’ BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

[0596] If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

[0597] Human IgG Fc region:

(SEQ ID NO:1)

GCGATCCGGAGCAGCCATTTCTGACCAACTCACTACATGCCACCCAGTGC
CCACACACTTTAACACCACTTACTGAGACCTAGCTTCTCAGCACTCCCTCA
ACCCAAAGCTACCTCTCAGCTTCTCAGCTTCTCAGCTTCTCAGCTTCATG
TGCTGAGCTCTGAGGCAACACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTC
GACACGTCCATAGCTGAAGTAACACCTACATCCATCAGAAGTACAGAAG
CACAGACTTCTCCATCCATCTCCATCCATCCATCCATCCATCCATCCATC
GCTCAGATTGCAAGAAGTTACAAGTACATCAAGTCATCCATCCATCCATC
ACCCCATGTGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA
ACAGCTGCTGCACTGAGTTCCCCAACTGAGTTCCCCAACTGAGTTCCCC
TCAAGTGGGACAAGCATGCGAACAGGGAGAACAGAACAGAACAGAACAG
GAATCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CCTCGCTGCTGCACTGAGTTCCCCAACTGAGTTCCCCAACTGAGTTCCCC
ACAGAAGTACAGAAGTACAGAAGTACAGAAGTACAGAAGTACAGAAGT
GAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGT
ATAATGAGTGCGACGGCCGCGACTCTAGAGGATG

Example 10

Production of an Antibody from a Polypeptide

[0598] The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

[0599] In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in
any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

[0600] The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

[0601] Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

[0602] It will be appreciated that Fab and F(ab)2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab)2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

[0603] For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., Bio-Techniques 4:214 (1986); Cabbily et al., U.S. Pat. No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601333; Robinson et al., WO 8702671; Boulainne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11
Production Of Secreted Protein For High-Throughput Screening Assays

[0604] The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

[0605] First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1 mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50 µg/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipettor may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1 ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

[0606] Plate 293T cells (do not carry cells past P20) at 2x10^6 cells/well in 0.5 ml DMEM(Dulbecco's Modified Eagle Medium) (with 45 g/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

[0607] The next day, mix together in a sterile solution container: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5 µl OptiMEM I (31985070 Gibco/BRL). 96-well plate with a small volume multi-channel pipette, aliquot approximately 2 µg of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipette, add 50 ul of the Lipofectamine/ OptiMEM I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipette to add 150 ul OptiMEM I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

[0608] Preferably, the transfection should be performed by tag-teaming the following tasks. By tag- teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with 0.5-1 ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipettor with tips on every other channel, adds the 200 ul of DNA/Lipofectamine/ OptiMEM I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

[0609] While cells are incubating, prepare appropriate media, either 1% BSA in DMEM with 1x penstrep, or CHO-5 media (0.116.6 mg/L of CaCl2 (anhyd); 0.00130 mg/L CuSO4·5H2O; 0.050 mg/L of Fe(NO3)3·9H2O; 0.417 mg/L of FeSO4·7H2O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl2·6H2O; 48.84 mg/L of MgSO4·7H2O; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO3; 62.50 mg/L of NaH2PO4·H2O; 71.02 mg/L of Na2HPO4; 0.4320 mg/L of ZnSO4·7H2O; 0.002 mg/L of Arachidonic Acid; 1.022 mg/L of Cholesterol; 0.070 mg/L of DL- alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L-Alanine; 147.50 mg/ml of L-Arginine-HCl; 7.50 mg/ml of L-Asparagine-H2O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCl-H2O; 31.29 mg/ml of L-Cystine-2HCl; 7.35 mg/ml of L-Glutamic Acid; 365.0
mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCl-HO; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCl; 52.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.00 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2HhlO; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Panthenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of l-Ascorbic; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCl; 0.031 mg/L of Pyridoxine HCl; 0.319 mg/L of Riboflavin; 5.17 mg/L of Thiamine HCl; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B12; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipic Acid; 0.081 mg/L of Sodium Pyruvate-2HCl; 5.50 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20 uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexes with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2 mm glutamine and 1xpensstrept. (BSA (81-068-3 Bayer) 100 gm dissolved in IL DMEM for a 1% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15 ml polystyrene conical.

[0610] The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5 ml appropriate media to each well. Incubate at 37° C. for 45 or 72 hours depending on the media used: 1% BSA for 45 hours or CHO-5 for 72 hours.

[0611] On day four, using a 300 ul multichannel pipetter, aliquot 600 ul in one 1 ml deep well plate and the remaining supernatant into a 2 ml deep well. The supernatant from each well can then be used in the assays described in Examples 13-20.

[0612] It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12

Construction of GAS Reporter Construct

[0613] One signal transduction pathway involved in the differentiation and proliferation of cells is called the JakSTAS pathways. Activated proteins in the Jaks-STATS pathway bind to gamma activation site “GAS” elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

[0614] GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or “STATs.” There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

[0615] The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase (“Jaks”) family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

[0616] The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schaller and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jak5, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, II-7, II-9, II-11, II-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thymopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cytokine motif (a set of four conserved cytokines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Tri-Ser-Xxx-Tri-Ser (SEQ ID NO:2)).

[0617] Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jak5-STATs signal transduction pathway.

[0618] Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>tyk2</th>
<th>Jak1</th>
<th>Jak2</th>
<th>Jak3</th>
<th>STATS</th>
<th>GAS (elements) or ISRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFN-a/B</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>1,3</td>
<td>ISRE</td>
<td></td>
</tr>
<tr>
<td>IFN-g</td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>1</td>
<td>GAS (RF3 &gt; Lys6 &gt; IFP)</td>
</tr>
<tr>
<td>IL-10</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>1,3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18 bp of sequence complementary to the SV40 early promoter sequence and is flanked with an Hind III site. The sequence of the 5' primer is:

```
5' - CGCCGATTTTCGGGAAATCTAGGATTTTCGGGAAATGATTTTCGGGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCG...
```

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site:

```
3' - GCGGGAAGGCTTTTTGCAGAGCCGGC3'  (SEQ ID NO:4)
```

PCR amplification is performed using the SV40 promoter template present in the B-gal-promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2- (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

```
5' - CCCGGATCCATTTGGCCGGAAATCTAGGATTTTCGGGAAATGATTTTCGGGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCG...
```

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or “SEAP.” Clearly, however, any reporter molecule can be included instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS-SV40 promoter element to create the GAS:SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.
Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAI, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAI, or NF-KB/Osteocalcin). Similarly, other cell lines can be used to test reporter construct activity, such as HEK293 (epithelial), HUVEC (endothelial), Rhl (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13
High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jak-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) can also be used.

Jurkat T-cells are lymphoblastic CD4+Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1 ml of 1x10^7 cells in OPTI-MEM to T25 flask and incubate at 37° C. for 6 hrs. After the incubation, add 10 ml of RPMI+15% serum.

The Jurkat GAS-SEAP stable reporter lines are maintained in RPMI+10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI+10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophane covers) and stored at ~20° C. until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4° C. and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 50 fold induction is typically observed in the positive control wells.

High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jak-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TH-1, HL-60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2x10^5 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inacti-
vated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

[0638] Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 mM NaH2PO4·7H2O, 1 mM MgCl2, and 675 mM CaCl2. Incubate at 37°C for 45 min.

[0639] Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

[0640] The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

[0641] These cells are tested by harvesting 1×10^6 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^6 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

[0642] Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15

High-Throughput Screening Assay Identifying Neuronal Activity

[0643] When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

[0644] Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

[0645] The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (~633 to +1) (Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCACGCGAGCGAGGATACACCTGATAGACCC (SEQ ID NO:16) GCG-3'

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC (SEQ ID NO:7)

[0646] Using the GAS-SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS-SEAP/Neo vector using restriction enzymes Xhol/HindIII, removing the GAS/neo400 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

[0647] To prepare 96-well-plates for cell culture, two ml's of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

[0648] PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

[0649] Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

[0650] To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

[0651] The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

[0652] Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^4 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ml of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16

High-Throughput Screening Assay for T-cell Activity

[0653] NF-kB (Nuclear Factor kB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotxin-alpha and lymphotxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-kB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-kB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.
In non-stimulated conditions, NF-κB is retained in the cytoplasm with IκB (Inhibitor κB). However, upon stimulation, IκB is phosphorylated and degraded, causing NF-κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF-κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-κB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-κB would be useful in treating diseases. For example, inhibitors of NF-κB could be used to treat those diseases related to the acute or chronic activation of NF-κB, such as rheumatoid arthritis.

To construct a vector containing the NF-κB promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF-κB binding site

\[
\text{GGGACCTTCCC} \quad (\text{SEQ ID NO:8}),
\]

18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

\[
\text{GGGACCTTCCCATCGGCGCGCT} \quad (\text{SEQ ID NO:19})
\]

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

\[
\text{GGGCGCGGCGGCGCT} \quad (\text{SEQ ID NO:4})
\]

PCR amplification is performed using the SV40 promoter template present in the pB-gal promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLS2. (Stratagene) Sequencing with the T7 and T3 primers confirms the following contains the sequence:

\[
\text{CTCGAGGCGGGGGCGCTTCCCGGCGGCTTCCC} \quad (\text{SEQ ID NO:10})
\]

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and Not.

Once NF-κB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 0.1, 0.1 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17

Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BI400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μl of 2.5x dilution buffer into Optiplates containing 35 μl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

<table>
<thead>
<tr>
<th>Reaction Buffer Formulation</th>
<th># of plates</th>
<th>Ass buffer diluent (ml)</th>
<th>CSPD (ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>60</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>65</td>
<td>3.25</td>
<td>3.25</td>
</tr>
<tr>
<td>12</td>
<td>70</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>75</td>
<td>3.75</td>
<td>3.75</td>
</tr>
<tr>
<td>14</td>
<td>80</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>85</td>
<td>4.25</td>
<td>4.25</td>
</tr>
<tr>
<td>16</td>
<td>90</td>
<td>4.5</td>
<td>4.5</td>
</tr>
<tr>
<td>17</td>
<td>95</td>
<td>4.75</td>
<td>4.75</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>105</td>
<td>5.25</td>
<td>5.25</td>
</tr>
<tr>
<td>20</td>
<td>110</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>21</td>
<td>115</td>
<td>5.75</td>
<td>5.75</td>
</tr>
<tr>
<td>22</td>
<td>120</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>23</td>
<td>125</td>
<td>6.25</td>
<td>6.25</td>
</tr>
<tr>
<td>24</td>
<td>130</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>25</td>
<td>135</td>
<td>6.75</td>
<td>6.75</td>
</tr>
<tr>
<td>26</td>
<td>140</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>27</td>
<td>145</td>
<td>7.25</td>
<td>7.25</td>
</tr>
<tr>
<td>28</td>
<td>150</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>29</td>
<td>155</td>
<td>7.75</td>
<td>7.75</td>
</tr>
<tr>
<td>30</td>
<td>160</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>31</td>
<td>165</td>
<td>8.25</td>
<td>8.25</td>
</tr>
<tr>
<td>32</td>
<td>170</td>
<td>8.5</td>
<td>8.5</td>
</tr>
<tr>
<td>33</td>
<td>175</td>
<td>8.75</td>
<td>8.75</td>
</tr>
</tbody>
</table>
### Reaction Buffer Formulation

<table>
<thead>
<tr>
<th># of plates</th>
<th>Run buffer diluent (ml)</th>
<th>CSPD (ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>180</td>
<td>9</td>
</tr>
<tr>
<td>35</td>
<td>185</td>
<td>9.25</td>
</tr>
<tr>
<td>36</td>
<td>190</td>
<td>9.5</td>
</tr>
<tr>
<td>37</td>
<td>195</td>
<td>9.75</td>
</tr>
<tr>
<td>38</td>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>39</td>
<td>205</td>
<td>10.25</td>
</tr>
<tr>
<td>40</td>
<td>210</td>
<td>10.5</td>
</tr>
<tr>
<td>41</td>
<td>215</td>
<td>10.75</td>
</tr>
<tr>
<td>42</td>
<td>220</td>
<td>11</td>
</tr>
<tr>
<td>43</td>
<td>225</td>
<td>11.25</td>
</tr>
<tr>
<td>44</td>
<td>230</td>
<td>11.5</td>
</tr>
<tr>
<td>45</td>
<td>235</td>
<td>11.75</td>
</tr>
<tr>
<td>46</td>
<td>240</td>
<td>12</td>
</tr>
<tr>
<td>47</td>
<td>245</td>
<td>12.25</td>
</tr>
<tr>
<td>48</td>
<td>250</td>
<td>12.5</td>
</tr>
<tr>
<td>49</td>
<td>255</td>
<td>12.75</td>
</tr>
<tr>
<td>50</td>
<td>260</td>
<td>13</td>
</tr>
</tbody>
</table>

---

Example 18

High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

[0672] For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

[0673] To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca**2+** concentration.

Example 19

High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

[0674] The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

[0675] Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g. src, yes, fak, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

[0676] Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

[0677] Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Lopodryne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, Mo.) or 10% Matrigel purchased from Becton Dickinson (Bedford, Mass.), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar.
To prepare extracts, A431 cells are seeded onto the nylon membranes of Lopodryne plates (20,000/200 ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60 ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer (20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, Ind.) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000xg.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10 ul of 5 um Biotinylated Peptide, then 10 ul ATP/Mg2+ (5 mM ATP/50 mM MgCl2), then 10 ul of 5x Assay Buffer (40 mM imidazole hydrochloride, pH 7.3, 40 mM beta-glycerophosphate, 1 mM EGTA, 100 mM MgCl2, 5 mM MnCl2, 0.5 mg/ml BSA), then 5 ul of Sodium Vanadate (1 mM), and then 5 ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10 ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120 mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated well plate to associate with the biotinylated peptide. Wash the MTP module with 300 ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD (0.5 u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above. After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10 ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1 ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELTIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a
disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95° C. for 30 seconds; 60-120 seconds at 52-58° C.; and 60-120 seconds at 70° C., using buffer solutions described in Sambrook, J., et al., Science 252:706 (1991).

[0690] PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies.) The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

[0691] PCR products is cloned into T-tailed vectors as described in Holton, T. A. and Graham, M. W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

[0692] Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenin-5'-dUTP and anti-digoxigenin-conjugated alkaline phosphatase (Boehringer Mannheim), and hybridized as described in Johnson, G. C. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

[0693] Chromosomes are counterstained with 4,6-diamidino-2-phenylindole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, Vt.) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, Ariz.) and variable excitation wavelength filters. (Johnson, C. et al., Genet. Anal. Tech. Appl., 8:75 (1991)). Image collection, analysis and chromosomal fractional length measurements are performed using the Isee Graphical Program System. (Invision Corporation, Durham, N.C.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22
Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

[0694] A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

[0695] For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ng/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

[0696] The coated wells are then incubated for 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide.

[0697] Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate.

[0698] Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (PNP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23
Formulating a Polypeptide

[0699] The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

[0700] As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 mg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 mg/kg/hour to about 50 mg/kg/hour, either by I-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be used. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

[0701] Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracutaneously, intravaginally, intraepithelially, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term
“parenteral” as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intrarterial injection and infusion.


[0703] For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

[0704] Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer’s solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

[0705] The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or triptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

[0706] The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

[0707] Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

[0708] Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as a aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-mI vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

[0709] The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24

Method of Treating Decreased Levels of the Polypeptide

[0710] It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

[0711] For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25

Method of Treating Increased Levels of the Polypeptide

[0712] Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is
one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

[0713] For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26

Method of Treatment Using Gene Therapy

[0714] One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham’s F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

[0715] At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

[0716] pMV-7 (Kirschmeier, P. T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

[0717] The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5’ and 3’ end sequences respectively as set forth in Example 1. Preferably, the 5’ primer contains an EcoRI site and the 3’ primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

[0718] The amphotropic PA317 or GP-+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco’s Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

[0719] Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

[0720] The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodes 3 microcarrier beads.

[0721] It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

[0722] The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entirety.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 343
<210> SEQ ID NO 1
<211> LENGTH: 733
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 1

ggatcggag caccaaac tctgacaca ctcacacact tcgacgctg ccgaacatcg 60
<210> SEQ ID NO 2
<211> LENGTH: 5
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> Feature: MISC_FEATURE
<221> NAME/KEY: LOCATION; (3)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins.

<200> SEQUENCE: 2
Trp Ser Xaa Trp Ser
1 5

<210> SEQ ID NO 3
<211> LENGTH: 86
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 3
ggcgatcag attcccccca aatctagatt toccogaat gattttcccc aatgatttc 60
cocgaatat cttccatatc aatag 86

<210> SEQ ID NO 4
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 4
ggcgcaagtt tttgcaagag ctaggc 27

<210> SEQ ID NO 5
<211> LENGTH: 271
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 5
cctagattt cocgaaatcc tagatttccc cggagatgatt toccogaat gattttcccc 60
aaatatacc cttcaataa aagcagaaac cgatctcccg cccotaactc cgoccatcocc 120
gcocaactt cccoccaagtt cccocaatct ccocccoccat ggtgcactct ttttcttat 180
ttatgcagag ggcagggcgg cctgggcttc tgagttttc cagagtgtgt gaggaggttt 240

 ttgagggc ctaggctttt gcaaaaagct t 271

<210> SEQ_ID: NO 6
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 6

gcgtcgag ggagcagcgg atagaaccoc cc gg 32

<210> SEQ_ID: NO 7
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 7

gcagagcttc gcgactccgc gcgtcgcctc c 31

<210> SEQ_ID: NO 8
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 8

gggagctttt cc 12

<210> SEQ_ID: NO 9
<211> LENGTH: 73
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 9

gggtctcgc gggtggacct tcggggact ttcgcccact ttcctctcctg 73

coscotcoast tag

<210> SEQ_ID: NO 10
<211> LENGTH: 256
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 10

cctcgagggga ctttccogcg gacattccog gcagttccog ggaacattcc occtggagg aocctgggoc ccctgcggccc cacgctgccc 256

cacattgcag gcacccatttc cggcggcgcc cggcggccgc ccctgcggccc cacgctgccc 180

gggcgcgtcc gcgtcgagtc tttgagagt gcgtgccttg ggcgggtcgg 240

cgggaggg ggtgggagct ttcggcccact ttcctatcctt tag 60

coscoastc actcagct cgagcttctg gggcggcggg ggtgggagct ttcggcccact ttcctatcctt tag 60

<210> SEQ_ID: NO 11
<211> LENGTH: 1679
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:  
<222> LOCATION: (1656)
<223> OTHER INFORMATION: n equals a, t, g, or c
<222> LOCATION: (1654)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 11

gacgcgcacc ccggccgcat cttcaaggtg gcggacgaac tagccatctg ttaactacgtg 60
gagaaaat gcggcgccgc cagctctcct acgcagcagc cttgtgagaa aggagattcg 120
agagctgcc tcagccgac a cttggtgagc tctcctgaggg agcagctact ccagcacttg 180
cgggagccgag acaatagtaa cttgtgctgg gacgcgcgcc tgagctcctc ctgccagca 240
cctcaaggg aggcggcggag gcggccgata cgggggaacc acaagataaag 300
agaacctcag atgttctgtg ctaaggcccag ccggacatct tggagtcttg gtttggtcag 360
gcgtccaaa agatgctctt cgaaggccac cgggagtttg aggaagygttg gtttggtcag 420
agcctctctc attttaaatt atgttaaatc atgtgtgatt atgtctccac ccctcctctg 480
gagacgtcaaac aacaagaccc tgtgagggag gaagagagaa aagaaaagaa agttttcccag 540
cgaagatgg gcggctcgcc taagactcatt gcagcatcag ttggagaatt taaacttttgat 600
cctcaagcc ctgcctactaa aatagctggtg agcttgagcgc tccctccggc 660
ttcctcatt cttgctcaag agctggttga tatttctcatt gttcgcggca agatgacatta 720
gatggcaca aagatacgtga ggttccgtag gaacatcgg ctaaaaaatt tgggtctcag 780
aagtgtgctc gcggagttgg atctgcagaa aactaaatcg ccagataatc ggaaagaaaa 840
aagactccg tagctgtattgt gttttaaaaat atttgctgctc ataagactctt aagagctct 900
tctcttcggcct cttttgctcaggt atcgtcttcg ttgctctgct aagtctctgct 960
ttttgctgc acagacgtct gtagaggggt ataataaagaa aacaaatgtg atgttttttttc 1020
agctgttct gccaaagttat catttttgtg agtaaaggg gttctttataa atgacagagct 1080
tgctctcatt cattttgtga atggtcctca cttgctgatt gttctttatgta gtagcaatgtg 1140
gaagagactt attgattgagc tggccgcttt caggtatcct aagtaaattc ctgctcttg 1200
gcagacgcgt cttttagatg gtagatccga caaataaaga aactataagat gcggaagatc 1260
tataattgag gctctcagat aggctttcct cccaaaccct tcgcttatttc ccacgtgagc 1320
tcgatttct gacagcttct gasatagcttg tattctttttttt gttcactttag tattttttct 1380
ccctcctccttgttttgcgg gttttgctacg cttgttcgct cgggtttctt 1440
tctcttctc atcttcctactttttagata cttgctgctag atgtttttttc 1500
tataatattc cttgacttttag atttatggtt gttccttttagatatagctatgctatgttg 1560
attaaaaagtattgtgatt gactctcatttt tgttggagt gttccttaagag ttaaaaaaat 1620
tataactttg tatttttttaaaa aaaaaaaataa aaaaatcagtt cgggctccagc agaagatctc 1679

<210> SEQ ID NO: 12
<211> LENGTH: 1963
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (335)
<223> OTHER INFORMATION: n equals a, t, g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1959)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 12

gctttccgg gcggccgccc gttctctctc gtgtgctgctgtgactcag cttgagaagagcgctgctgc cgggagcgcc 60
ggcggcacta gcagccccac ccagggcgtc ccctcctggc ctccacagcg tcccaagcgc
120
cagcggcgc gcocctcctag ctacagctgt ccgctgctct gttctcctgac tcgctggtgc
180
gatgctggat gggcgctggc gcctgcgtgg cgcgcagcgc gcctcgcggc cggtctcggc
240
cagtcgcaac atggcagacccc gctgagagcg gctgagcttc gcgttcttca
300
cottacagca gcttgcttgc ccagagcgcca ccctttgctt gggccacagt gcgtgtgtgg
360
gtctccacag ccoccttcct ccacagtacct tgcacacttgc ccacagttatg ctaagagcga
420
gagggccag gctgagacca ccctctctctt gttctcctggcc agagagcagct attttcgagg
480
tggtgattct tttgatgctg ccagccagct gaggatcag taagagtagcta tttcaagttg
540
aactctctct ctggactcct tcacattcct gccagcattc ggtgctgcttc ttggctgctc
600
cacgcaagc gtcagcagcgt atgggccact gccgatgatt ccgctcctcc ttaatattag
660
gatctgatt gtcaagctttt ccacactattt ccctctctgt cctggtgactc agtacgctgtc
720
tgtgacgtggt gccgtcctctt gtcctccttc cctgcttcatg cctgctgttc gcgtgtgctc
780
tttccggctgg atggcttctct gccatccttc tcaagatcattc aggagtattt ctcagcccctctc
840
tataaggccc tgtttttcttg cccagccctg ccctcctcatg tggaggtcctg cccatggtcc
900
tagagacaac ctcagccagg tgaatcncaga ttgacaaacc aggacattta ccaactccag
960	tagaaaaaatg aagtctggct gcagagatc ttaaaatcttc ttgtagccta ctgttactat
1020
ttaaagctg ttcagcagat cctgcaacct tttgtaagtt cagttttatt gcccttttct
1080
ttgtgacgca ctaagctctct ggtgtattac cttgcttttt ctgctagttt ctggcgccat
1140
taaggagct gcacagcttg ctggcccttg gcgtttctgg gtttgccttt ctcctttcat
1200
cctctcagca gccttcttatt gcctatattg cttatagcttg gccttatttcc
1260
tggtgttct gcaacacttg acgtcctagtgtc tctttgctgc cctctccattt acccatattt
1320
tgcggctcag gtcctttctcc cctctcctct ctttctttcg tttgctgttc gacacccgtaa
1380
tgatcggcaa gtaaaggctg agatctgctg cctcctcttt ctcttttttt tttctttttc
1440
gcatttatt ttttactcgc gccttctctt ctttctttc tttctttttt ctctttttttt
1500
tattggagct gtgctctgct gtaaaacgct gtaaaaagtg agaactcttg agaactcttg gtaaaatcctg
1560
tatcaaaat cgattctttct gctgcttcctg tggctgctgct ctttctttct gctgctttct
1620
tacaagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct
1680
tggtggtcatt ctgctccttg gccatctgtg cttgctgctgct cttgctgctgct cttgctgctgct
1740
tcacaagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct
1800
tggtggtcatt ctgctccttg gccatctgtg cttgctgctgct cttgctgctgct cttgctgctgct
1860
tacaagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct
1920
tacaagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct gcaagagcct
1980

<210> SEQ ID NO 13
<211> LENGTH: 1212
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 13

tggtggtcatt ctgctccttg gccagcagc acgtcctttc ttcacagca aagtctggtg cttgctggca tagtgagcgg 60
tagacgtag tttttcttac ctcagcagc agtctggtt gcacagcagc acgtccttt 120
<210> SEQ ID NO 14
<211> LENGTH: 2061
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1763)
<223> OTHER INFORMATION: n equals a,t,g, or c
<210> SEQ ID NO 14
<211> LENGTH: 2061
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1763)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 14

```
gtttctcttc cccttttacctg tatacctttg 180
tagtagtagac atacagcagat acgtagtagat tttttaataat tataatatgat ttttgcaggt tagatacataa 240
taataaatc ttaatgtatgc 300
tttttattt attttgattt attttgattt attttgattt attttgattt 360
tatataata atataata atataata atataata atataata atataata 420
tagagacgtc atataata atataata atataata atataata atataata 480
ttagtagtagac atacagcagat acgtagtagat tttttaataat tataatatgat ttttgcaggt tagatacataa 540
tttaataata atataata atataata atataata atataata atataata 600
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 660
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 720
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 780
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 840
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 900
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 960
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1020
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1080
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1140
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1200
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1260
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1320
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1380
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1440
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1500
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1560
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1620
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1680
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1740
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1800
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1860
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1920
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 1980
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2040
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2100
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2160
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2220
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2280
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2340
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2400
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2460
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2520
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2580
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2640
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2700
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2760
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2820
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2880
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 2940
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3000
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3060
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3120
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3180
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3240
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3300
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3360
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3420
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3480
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3540
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3600
tttttttttt attttttttt attttttttt attttttttt attttttttt attttttttt 3660
```
cagatgaga cttgccctct tgtctaaacct tgtgaaggga gaaagttct gagggaacta 720
cattcctgcc gctctcctg ttaagaagtc agcctaagtct aggccctg gcctccaaac 780
gcaggcgcga gcggggagt gacgtggaga gcagaattgtc cggctggaga 840
gcgaggtgg ccacccacgc cacccctgttg aactaaacct caggttaaaa ctgacaagga 900
tggagagag aattggctgt gccttggttt ttcctgtggac tccttgagtgc 960
cgctacact aacacgaagat cttctaaaag aaaaatggg acacatttgct gccttaaaa 1020
gcagaaggag acagcgcgca tggagaagtg tctctgagat tyggtaaaac aacgagtga 1080
gcgacctca gcatacagag gggctgacga aacctgaata aagagggcata ccgttgatat 1140
ggaatagag aacatgcctc aagggtaaaa ttaagttgct ctggaaatga aagatataccc 1200
ttaactgcct gcagaacctgc attaaacctc agatgttgac aatctgtgaa tcaagaagga 1260
tgagagatga taatttttaat aatactaaaa ttatcttttt ttcaggattga aagttgctc 1320
ttaactctat aaaaaggttata taaggacgtt attacttaa atcagaaaaga gtgttttttt 1380
gcagtttccc tttttaataa agtttagat gttctctag gcaagtctggc taatcttgcc 1440
gcaggttagaatgc acctaatctt ttagattcct gttctctaga aggtccaaaa tttatatttt 1500
tctcctgctgt gaaatttttgc aatctgtta atctcatctct acacccatct cttttattt 1560
tgttaasgta ctaattcat cgggtgattata cttcaaaata gcacaaatata tcaaaagaa 1620
cggatgagagc gacaccattca gacaattcgg cttggcagtt atttctttaa attccccctg 1680
aacggccaaa gtaatatata atmggtttga tttatattat atagattacat cccactttta 1740
cattaagag tctagaggct gccatgcgtag tggctctag ggctctactc gcgaaggctg 1800
gcagggcag gttggaaggt caagggaggt cagggtaggt agcaagactg tggcaaaata 1860
gtggacacct gcattttacta aatactaaaa aatactgaggg caagctggtgg aggcaacctt 1920
aatccgact actagggaggt ctttggaaac cggagggaag agtggtggcg gaagctgagat 1980
cgctgcaacct ctaacgccag tgggtgagat aagtggtagc actctcaca aaaaaaaaaa 2040
aaaaaaaaaa atagctggc a 2061
agacacac ccnaaaggaaa tcccacacac actccccctt actctctct tcccccttca actaccttcc tccacactctctc atagggtgct 370
agacccgaaa caaagcactt atctcct actcgacgc acccgactgt tccacacctctctc atagggtgct 420
gacacccg tccacaccgc actgcacccg cttcctac ccctctcct tccacacttca aaaaacacaa aaaaacacaa 480
cactcaagc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 540
tgacacccg tccacaccgc actgcacccg cttcctac ccctctcct tccacacttca aaaaacacaa aaaaacacaa 600
gtccacacccg tccacaccgc atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 660
tccacacccg ggcgcggtgct tccacaccgc atctctcct tcccccttca actaccttcc tccacactctctc 720
tacccacccg aataactact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc 780
cgcacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 840
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 900
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 960
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1020
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1080
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1140
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1200
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1260
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1320
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1380
tgacacccg tccacaccgc accgacac ccaggactt accttgcttt acctgaacacc atgggtgct 1440

<210> SEQ ID NO: 16
<211> LENGTH: 1052
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 16

tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 60
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 120
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 180
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 240
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 300
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 360
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 420
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 480
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 540
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 600
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 660
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 720
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 780
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 840
tacttact ccacacccg atctctcct tcccccttca actaccttcc tccacactctctc atagggtgct 900
-continued

tctccaccc tccccggtgg ccaagaacca gctagtctct tccaggtgtgc tctttgattt tctt
960
tgtttaggty acagagaamt ctttttagc ttttcattta ttttcacaa ataaaatgtt
1020
catttttat tttggaanaa aaaaaaaaaa aa
1052

<210> SEQ ID NO 17
<211> LENGTH: 683
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17
aattcggagya aggtactatt cattgtcata taseotgatttt tttacbcttg ttagcgaag  60
taggatatc ccttcctcct caagacacca taaaattgta atgtgatgtg gctgtgatgct 120
cattcgccact acacagcgtt cctacaccaac ctctttgtaa aaccacccac ccatatgaa 180
tagcctttctct catttccact ttggttactt tttcctaacaa tttctatcttt ggttttctgt 240
ttcacagaac taaaacacac tctttttgtta tgatctttct tttttttttt tttttttttt 300
gtgggtccta ccaaggggtc ccacactgatgt ggagaggttg ttgcatgtctt ctgagcctca 360
cccacctctt attttggtt agaaaaaaa ttggagtcctt aataaatcattt 420
atcatagttw atgtgtctat cttccacatc agaatcagtgc actaannacc accaggtgat 480
waccaatag atgataaaw tw ttatctattt ttttataatw ttatctatta atgatgctgc 540
tcttcccata accttatttt ccaagagctg ttaatctttta aataatgttt caaacagctga 600
gcaacctac ccagttcct ggataaata cagttctctt gcaaaaaaa aaaaaaaaaa aaaaaaaa 660
acttggaggg ggcctgctacc ctt
683

<210> SEQ ID NO 18
<211> LENGTH: 1054
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<222> NAME/KEY: misc_feature
<222> LOCATION: (74)
<223> OTHER INFORMATION: n equals a, t, g, or c
<222> NAME/KEY: misc_feature
<222> LOCATION: (1014)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 18
aactctattc atgtagaact atagaaggtta cgctgcaggg taccggctcc ggatcocaag  60
gtgcagcaac gmnncsgggg acagagtggc agcgcggtgt cgcagcgtga aaggcctcgt 120
ggccttaata acgcgagac gctcgggctt ggctggctgc aagccgagac attgtgaggg 180
cagggagcct ctgctgtcct cctgcctcttg ccaaacttgcgg ggccggggc ccacggcaag 240
aagtggatcc caacaagtgc tttccgcccag gcgcagctgta cctcttgagaa ggagctggca 300
acagctotgg ctctagacaa agggaagttc ccagctgttg atgtagctgt caaatgctca 360
ggcagctccgg ttgtagtaaa gacgtccaca tcaagaaaag ggcaagacca taacttgaga 420
gactcgcagc gatgtctttg atgtgacact tcacaggattgt cagcgcggttg 480
gtggctgtga tgggctgagaa taaccacagc cagggagc acogtggggt catcataaac 540
agcttcagct gtgcgtccct cggaggttgat ggtagcaaac ctgcatacct ctgctcctag 600
ggggagaac agttggctgctg acgtcactt gtagcggact gtgcctcgatag aaactttccc 660
gttgcgcaaa ctcgccccag ttctgttgcag acoccdactgc ttacccagct ccacggaaaaa 720
gttgcaacactttgggccagcacaagggcccctccttcagcggactgtgtggaccccctttatagcgaataatcagtctggtctaccagtttcagtctgggtataggtgagttcttccctcccgagcttggagttgcctggtagctcttggggtttgagaagaagtttggagagcagagattagggatcgtttggggaggaagagggcagagagttgagtttacgttttcatgtttcctcgccccctagtcctttttttaattaagttgctgacacttcttggctttaaaaaaaataaaaaaaaataaaaaataaataaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaaaaaaataaaa
atgttttga gacacaccttc tagtccacgga ctctcgtttct tccagcactgt gacactatttt  
1080
tgtgatgagc tgtgtgtgtg gatgtgatgt gccgtgaggg ccgtgagcc gcgggttccc  
1140
tgctggaagt ctgttccttgg gatgtgatgtt tctcttactt ctctcgtatt gaatcttccc  
1200
tctccagcag tgtctgctgcgct ccacaagggc tcctgaggg tccaaggaag  
1260
tctctttgct caaagagcaag tagtacacggc tatagcatgg tctctttgag gcocaatacg  
1320
atgttcctg tcaagtaacta aaataataaa acaacccaaact ggaaggtgtg gaaaaaana  
1380
annnnnnnnnc tgc  
1393

<210> SEQ ID NO 20
<211> LENGTH: 1215
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (61)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (65)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (104)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (180)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 20

agqasagtt tccacaaatg gaagcgggcc agttgagcgcg acaacacttaa tggagatgag  
60
ntcctcaatt agagacacccca ggttttacacc ttattgcttc ggntctctat tgtgtgagga  
120
atctgagcgc gatcataaac tccacacgga aaccgtcactt acacacta ta gccacagctn  
180
tatcagact caaatcaggg aagctgtgtta gctggtcaag taccgctcgg gacatccogg  
240
tgctcaccac gcctgctcgc cgcctgctcttt gatggctcgg gcacatgctgg  
300
tgacggcgcg ccgggtccaga gttgctccttc gcagggacag gctgacacag gaaggttgcg  
360
agcgttttttc gcgacagagtg attctttctac aacctccaggc atggctgtcttt tctggtctgt  
420
gctctcggcg tccaggcgctc cttgctcgcgg cccattcaag tcatggatg gcgcoccagg  
480
ctctgctcga gcaagacagag acaacccccc atggatcact ccgocgcgcnc cgggtgtygg  
540
accatatcct tgttctctgt ttgctctcggc acctgtcagc ctcttagagc  
600
caaggatgcg ggacaagcgc gcnaaagttct tcaagttact cggccacgcn caactgggtg  
660
cgtggtcgagctg acacacacac attacagttca cttgctcagc accatggtt acocggcggc  
720
catgttccct actgagaccc cctttgcttc acggocacgc aaccattggt aacotatggg  
780
cagggaccc accatcagtc agttgacatc tacgttaagga gggagcggc gacatgctctc  
840
cacacgacta gcggagatatt tcccacaccc gcgatttccc agacgtgtag gcactgctcc  
900
tgaaactctg attgtggtggg cacaaaccag caactgtgga cgtcttcttg agaaactcota  
960
tgtgctctct ctaaacaccc tcctgggaggg gtgtgtcgg gcctgcttgg gcacagacc  
1020
cgggatcagcc aagagacgtg atggagtagc agctcgggac agacacagtct ccoccaacct  
1080
ctctgctcagc tggcagaccc ctctccatct ctcttcctcag tcccccttcag gacactat  
1140
<210> SEQ ID NO: 21
<211> LENGTH: 2042
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 21

tgctcttcg aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa 1200
aaaaaggggc gcgcgc 1215

<210> SEQ ID NO: 21
<211> LENGTH: 2042
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 21

tgctcttcg aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa 1200
aaaaaggggc gcgcgc 1215
ctaaatgtag ttaatataaa gtgttaaattg gttgctctta attatagag aaaaaagct 1980
tgttagaat aanaataact gcctttatttc actatagaa aaaaaaaa aaaaaaaa 2040
tt 2042

<210> SEQ ID NO 22
<211> LENGTH: 1872
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1871)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 22

ggtgcaccc acgcgtcoga ttggcttaga gttotgta ga cogaagagcg caoggaagcc 60
tggggatatgt tggggaggag tttttttcttt ggtaattgg tcggccctctc 120
aagacctcag tgggtctcttt gcgtttttag aacatcaga acggaagttc attagaatcc 180
aaagggcaga ggcggtttaa gacggttcc caagggattgg gttgtgcttta ctggcggaga 240
cogaagaaatgt gcctctcota gcaaaccagag gttgaccaca ccccccctt cccgagcag 300
tcaactagcc ctgcgggagc raaaccaca aacgcaaga aagctggtgc caggcaagt 360
aagagagag agagtaggttt attttatgta ggtgagcgtg acctctctga ctgctctotc 420
cocccgggg acgctaccac gcctggctgt gcggctagac cttgggtcttct tcggctgcc 480
aagcggctac taagcgggtgc ttcacccttc agtacgcoggc gggcgctggtc tgcagacggg 540
aaacccagcg catgatcaga caggtgtga agaagcaccg cggccacctt gccctgcaaa 600
acccacgct gcagttcgag tgggacgct ccccacgct tccctgggcc tggcggcc 660
aaggggcagc cagccgggtt ccgcaacctc tagcctcacc aagtcacagt gccacaccct 720
cactgcacag ggtatgcgaa cgtagttttta aaattggatt ttagacgctgcc ggcgaagcat 780
gtgccgctac tctctctat ttcggtcag gagagttcota ttcccttga ccatctggga 840
aagatcagtt ccgtctgtatt tttggactgc cagncagctc aagcggcttgt ggtgcttgct 900
tggacccg cgggtacaag tgggtcgttt tttttaatct tccctgaccc cgggttttta 960
aagaggtaga taggtgctgc tcttaangt tgcncctggg tgtgtgccttc gcggccagct 1020
ggtagctcc attgagaggt cactgctgc cacggccggg cctccaggg agggctttgc 1080
tctgaggtat cccactttca tctggtgatt tctacctgac atcagctctg tcacagctctg 1140
aacagctgtt ccacgtcagg atagttcaggt tccgctggta cgtctctctc ccggccgctt 1200
tccggctct tccctcctgc cccgctgctc gaatactgtt tggtgctag atgcgaattt 1260
ctacaaccgc cagctttgctt ccagcagcat ccgggagga cctccggctt ccggcccttt 1320
taccctgttt tcccttctttg ggtgagggaa gtaaaaaa gaacagaggg ttgctctggt 1380
tctgctccttg ggggtttaa tctgcgcac ctggggaga aagcgtcgagc ctggctcaac 1440
aagcgcggga aaggtcgctca atgccttttgg ctaaacgcctg ctgctggaga agaagttta 1500
caccgcctca aaaaagctgg cgcaagcccc taagtttaag aggcttctgtg 1560
ttctattgg gcctgctcttc tcttccttat cctccaaaaa cttttttaaa agggctccag 1620
ttccttttga ctgctggtt gcggcccttt ttcagctcctt ccagagcttg 1680
taacctctgaagctctttttg ctgtgctggg cctgctgctt ccagagcttg 1740
-continued

cotgctggct ctctgagagt gacatttacc aactaagag cttccacttg gcggctgtc 1800
tgacocgtga acaatacttt actactggtt taacatatag tataactggtg gaaaaaaa 1960
aaaaaaac nt 1872

<210> SEQ ID NO 23
<211> LENGTH: 289
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (204)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 23

cattctccac octaaccaca tggctgtca ctctcttgtt gggtgaaagt agtgctctct 60
tgctctgag tagagaaagt cttccattag tggtttctga tcgtccgtag tttgtcagag 120
aatattgac aagctatattct ctctctgttt tcctgatgca gcaattcccc ttctatctag 180
atcctctg tcaatctatg aacctcttgg tgggttcctg agaagacact gcctctatct 240
tagagatcc tacaggtctga gttaaaaaaa aaaaaaaaag cctngggggg 289

<210> SEQ ID NO 24
<211> LENGTH: 3533
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (44)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 24

ctttatatcc ttcacatatta attcgttaa tcaactagta aacgagatag ttttcacctt 60
tagagcgtacc acaattccca actacactgt ctggtgctct tttaattagcc gaaatgtota 120
gttaacacaa aaaaagtgaa agttttacag tataaagga aaaaagctg aattttgca 180
aagaaaaga gtttattccc actttgaagt aacagacttt ttcagagag cagatctttc 240
actcttttaas gttttttacc aagccacatt atttttcacc gttcttggag atggacagag 300
tgctctcatgt ttgttatggt ggacctggtg gttgagagcg cccaaaaag agctatagagc 360
ctaatattgg tagagaaagt ttcacagagca cagagctgct tagaagctgg ttgaaacctttат gctctctgtt 420
atgtagatcc ttcatctcag tggggaaaa aaaaaagtgtg gggatgytga taaaagacac 480
cagttgtttc atacagaggg aacttcatac tcgagagtct aagtgacagt ggtgccgctc 540
cctagagata ggcacaccca ttctggagtt gggtaggtt ccggctgacc ctgagggga 600
gccagagca ggggctctac gcgagcgtgg ggggtgaaag ttgctgctct ccagcagct 660
caggggtggt cagtcgccct tcactgggct cggcgacca aaacctatcg ogagggagac 720
gtgggtgggg gcctgttatctg tcacagagcc ctttaggtac agttcttattt gtggacgctc 780
ggagtgtgca gaaaaagac caactctgtg ggtaggactgt aagcagactac ttatctttag 840
aagaaaact aacgctggtcg aacacactgt ttttgtgttt caagaaagaa gtttttagag 900
gggtctctga ctgtagacaa agttactaca gocacttcttc ttatgctc 960
catagcctc aaaaaagata gttgggctca ctatctcag cagtgagagct gtcgcatgtttt 1020
gggagacatt tataagggca acaagctgca atagggctcc caatgagaagg gttgagacaco 1080
-continued

aattccctcg agtttaacag cagttttaca cagggcactc agttaccgag caacgtcagt 1140
coaacaacag ggagtccacaa atgttcactt cacaocctcgc catctttcgag caggggtatt 1200
tgtcttaga atcttargaa tattgtacac cctctctacgg tgtctcagag cattgyaatt 1260
cotagcagga caattgacgta ggaagaagta gggagaacct gcocccaaag aggtctgcca 1320
agctaaaaat gtattccaaaa gcaagcaggt tttgacacgc cttttacgtg gaacagtagt 1380
tctggatttg gataggaacg gaaagcggca tttggaaatga ataaactctt ttaaagtaaac 1440
aatcttaagc gnaacagaggg aagtaaanaat gttggaggtac ttcacatgctg atgatcttcca 1500
gcaatcggcc cagccaaacgc ggagggaga agtgtaacc ccaacctcatt ataatatacc 1560
ttggccgg gatgacgctt gttgtgaggt gtcaacaaacc cggacatagc acaacttccag 1620
gactctcacc taaactatga agatttttcaca gatattcagc gttccagctg taanaactcca 1680
acatacgta atgtgacagc taaattcaaat tttgacatagc ttcagaaagc acaatcttacg 1740
acagatggac caaattcccc tttgagataaa agttccacaa cacaaaatata taacccacag 1800
aaaaaagaatc tccgaggttc acctgtatgt cgggtaatct caaacattctgc aaggatgttg 1860
aaggacacat cttgtatctgt cgggctgtata ccttcttcat gggagcaagc gcaccccaac 1920
ggatgttaat ctctccatgc aggaactgac cttcactatc aagtcttata tccacagccttg 1980
ctggaaacat cttttaacccaa attggtagcact gttgaccgtt accttcctacg ttcacccaatc 2040
gcacatcctgacet cttcttctgac cttcagaaacct gtaaacttctg gcattgaagt 2100
gtctgacaaa acctgtgacag ctaagctgata taattcttctt tttacttctta ttaatgatag 2160
ggagaggagtt catattctac ttttagctgtc gttgactgctt tttaaccgtgtg ttgaggtatac 2220
cacacacacc agagctatgg cttcagcaggg gcccacgacc tgtttaacccacta 2280
aatatacttg aagggagaac atattctacct ttggagctgtc ttaacttgagg gaaaatagct 2340
aaggacttgccc atctgtaatcg tcaaaatata gaaaacgagc cttcactgagc acaactttcc 2400
aatctacacat cgtctgacgc aagcttttta aaaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaatc 2460
cocctccccat ggggtatgggg gttgtctgata ctaacctaaa cataactgcga acatgtgtatg 2520
ggctgctgag gctggttcct tttctctctgag ggagttgagc atgggtctag gcacaagctc 2580
acatatctg gctggtcttc atcaattttag tcggacacgc ccaaccttata ttttatatttg 2640
tttttttttc agttgccaggg gttgactaatcttttttt ttttttttttt ggtgccccagc 2700
agactttgtg cttgtaacag aaccaatatta ctgagcaga gcctacacta gttcttaaaa 2760
tgtatataag aacacactaaca aagcttttaac gaaaatctgac cttaatatgt ggaagtacctt 2820
atttgagttc aataaaaactt aataataataat ctatatgttg acatatgtgtg gagaatatcctt 2880
cctgatcttc atttgagaggtt ggtgctttcct ccttttcctct ttccttccttc 2940
tttttttttg agagaggttgc gcaagcttat atcgctgtgtc ttgagagaaa 3000
atatagccag caaactctggt atctttttgcc aagggtgttgat tattactgagctctttcttt 3060
cagatgtaga caaactctggt tctgtttttcgtgtgatttg cttataaaa atgtcacttc tttttttttt 3120
aaaaaaaatgctggttcta gactttgtgat gatagtttctgatttattagtc gacccctttttt 3180
atatatatgc atatatgtttgc gaaatgttct gcgtcagcttt tgaaccttttg gaggagttgt 3240
aggtgagttc aatatagcaaa aagggagattc aattgagactg gttgagaggtttgtcttttgg 3300
cocctgctac taaaagggcc aagttttttag cagttttttcag cactttcaca caactgtgaag 3360
-continued

tgatgggaact taaagaagat gaacgagac aatctcagt ggtggttttg aacagcaggt 3420

<210> SEQ ID NO: 25
<211> LENGTH: 1148
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

acccacagct cccagaaatta tacticctca ttcnattat gttgataaca aagacctttg 60
cpaccattacc cccacagct ttttaagagat gaacaggga ttctctgca aacatcaga 120
acccctttt aaaaattttag ggttctttac tggcccatac atggagatag ccaggttattg 180
aatnattaata cttattattcg cttcctaaac ggtcctttgt caatctttgg tctctatttc 240
toccttttct cctattcaac ccaaaatacg ctgaaaccaa taagagtatt tatttattty 300
atattgttg gcataattcc tagttttctc tttccattca acatattatg aacctcag 360
acattgtaa atgttttta ataaacccoca gacccaatag aaaaattcna agtcaacact 420
agcagtctca tgaatttaaa tttgtcctta taattttcag ctaattttata acacaaagga 480
canattcagt gacgggacg tatttacact cttttctagc anaacatcgg ttactgctct 540
taatcact caccttcata gcacccctac cttctctact aagtcttcac cacactctgt 600
cgtctcctca gtaattgagc ctttatagc aanatttttt ttaaatattag gcaaatattg 660
tcttttttact aagttttctt ctcctattta ttgctttgt ttcttttaatt gggtcttttt 720
tttgtcctca gttttttttt ttgtttttttttttttttgg gatacacta attttgtcct 780
tattttct tttttttgtg ttttatactgg gccctctgtt ggacatcttta gggataagtt 840
tatnattaatt ctttatttttag taataaactc tatnattctag tgtttatatt 900
tttgacaac cttcgatggtt aanntttgca ttcttttatt cttgtatatttt ctctctcctta 960
cctattttg cttttctttt ctatatttatt ggaattata ctgctctgct ctaaaattaca 1020
gacgttaata ttattttctc gtaattttg aaacgttggc ccaggttttt ttcattataa 1080
taaaatgtg ggtaaatttg taanaaaataaaa aacatcgggg ggcccggatga 1140
cctattta 1148

<210> SEQ ID NO: 26
<211> LENGTH: 717
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

ggcacagcgt agctgcagcgc accggaacag cccggcctgg tggcccgctt cccggcgc 60
cggccgctca gagaacggcg cccggcgcct cccggcgcct aagttttctg cttgcttctg 120
cctagcggg cgcgtgcttg gggctgaggt tggccgcaaa cgccagctcc gcggcggcggc 180
cctcaggctg gcggcctggt gggctgcca gggcggatag ctgctgctgat 240
gacgctccc cccgggagcc cggctagtct cccccagttt gcggctgcaaac gctgctgctgct 300
cgtgccccag accgctctgt tataaggatt ggcacaaaaa agtggctttg ccgctctctttg 360
cagagcctccc tccacctccgt tgtgggagag aacgggaagg cccactccttc ctctctctct 420
gctatggaas aacggggttc tacaccatct gtcocaggg gtagcagtcgt gcaggtatgac 480
gttgggtgga ttgcacttaa cgcagccaca tacggtcgtaa agcttggtgna ggcccattttg 540
cctctcgtag gcattgacct ggtgcacgccctcctggcct tcactggtata caactctatcc 600
gaacaagccc atgcacccca agttccttca aagactccta agyaagagña agcacaaccag 660
agcaaaaaaga ataatataa ataatatatc aaaaaaaaaaaaa aaaaaaaaaaa aaaaaaa 717

<210> SEQ ID NO 27
<211> LENGTH: 1099
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1030)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 27

ggcacagcgc ggttgggaactctcgtgc tctcccctcct gtctgtcgccgt tgtacctgtat 60
cgccggatc atcgggtggtg agggacaggtt ccacccgct acgcgaacg gcgcacggttg 120
gggtccttac aatccactac actuacgccctt gtctgcagtaa tgcaggtcctg ggtcctttcgtg 180
cctccgact ggaatccatgt tcctacggct ttcacctgagt ccggttcatgtc gtttctgttg 240
cgggtgagg acctacccggtt gcccgctg tccgtctgtg atccgctggct cccgctgagt 300
gccctcagc gacgccagc gccgctctgtg tccgctgctg ggtgctctgtg ggtgctctgtg 360
gctctctttcttccttgtc ttcgggctgactgtctgccgcttgcactgtcgc ggtgctctgtg 420
gctctctttcttccttcctgc ggtgctctgtc gccgctctgtg ggtgctctgtg 480
cgcgagctgc gacgccagc gccgctctgtc gccgctctgtc gcgcctgtgcc gcgcctgtgccc 540
gctctctttcttcctgc ggtgctctgtc gccgctctgtc gcgcctgtgcc gcgcctgtgccc 600
taaccaacc gtcgctggtac gccctgctgtca cgggtacatct gcggtatgtca 660
taaccaacc gtcgctggtac gccctgctgtca cgggtacatct gcggtatgtca 720

gttctccggtac gacggtacag gcgtgtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc 780
tacacacacc gcgtgtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc 840
agtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc 900
gggttcctgc ggtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc ggtggtggtgc 960
tgcggccca cccacccca atccgccgtc ggcgccttgtgc gcgcctgtgcc ggcgcctgtgcc 1020
tgcggccca cccacccca atccgccgtc ggcgccttgtgc gcgcctgtgcc ggcgcctgtgcc 1080
aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa 1099
ncacagcaga gaaactctcg cttctggggc gtaacotaca ggtatcctt ggacagaag
atcttatgt gaagatcagc tccaaaggtg ctgtgaggct tgttctgcct gccccattgtg
30
ataaatactta tccagcagac atcaacatat tggtagggcc aaaaacatttt cttgtcaactg
cagagatgc aagatctgtc caagatcattc tggtaggc aacaaatcg aaaaaaaggt
120
gctggtagag tgtctctcca tggcttccac tggaaaaagc atggttacacy gatggtgatga
ttttgtcttg acaagactgtg ctctctatgc gtaaaccgcc aaaaaagctg caagatctgtg
180
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
gcttctaga tagtgaag acagctcacc gatcctcagtt caaatcagcttg aatcactttttt
360
gcgcacacatgaagctcag cttcagagc aatcactttttt aatcactttttt aatcactttttt
420
agtgcaccag gttggtctagc tgggtcaactg ccctgtagactg cttcagagc aatcactttttt
540
gagtctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
600
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
720
cagctttaca acctgaggag gatggtgagag cttcttcaga aatcactttttt aatcactttttt
840
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
960
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1080
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1200
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1320
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1440
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1560
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1680
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1800
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
1920
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
2040
gatcctcacc gcatctcagc aagcttccac gcatctcagtt caaatcagcttg aatcactttttt
2160
<210> SEQ ID NO: 31
<211> LENGTH: 1448
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<401> SEQUENCE: 31

aaanaaaaaa anacccccc acgctgcctg tcctccttc gcttgcactg cctccagtg
   60
ggacctaca gcagtggtgt ttcttctagt gacccccccc ctaattcaag ggtgtcctc
   120
gagtgtgttt ttaattttg tgggatataa atacatgca gtaagtcccc
   180
agcttctcag ccacattgct tcattctctg tggtaaatatt ttctccctcg ggacggggag
   240
cacgggtgag gcacagggaa caacgagggag tcggagagtg gcgggtcttc tgggtctact
   300
aacggccggag attaaggagt cacgtttaag ggtgtgcagc cccgggaggt tcgggaacct
   360
ggtgtcaccg gggagcagtc gcgtgtggat ggcggcggcg gcagagagga ggtggagcgg
   420
cacgcctcgc cacccccctt cgatctggcc ttccaggggg gtcggagggct gtgcotcacc
   480
gttcggggct ttggtgccgc ggccccctcc ctaaatccac cctcgaacc gcgaagctcc
   540
gacagagag aacgccaggg gaattccattg gacggcagcc ccactggtta gcagagttca
   600
catctcctat gtgggcgcag tactcacttt tactcgtgct tgcctcctga
   660
tggtgccctg gttttcctcc tgtccctcc gacccgctga ggtgcagctg gggcagcac
   720
tggggagctg gcgtctgctg tgggctccct ttcgacgggg gcggctgctg ggtcggtcgtg
   780
cacctcctac gcgtgcgctg acacggaggg gaggctggcc gacccgagcc tggactccag
   840
agagacccg aacggtggcc aacggtccgt gtcgtctctc ctcgtctgggt ccccgtctgt
   900
gcgtggtcgt taaatttttt ctccacccca acctcccccc aacgacaaca caagcgctcc
   960
tcgtctctcc cccggcgctg gtcggagcct gcggctccttc ttcgacccac gcgcgatgag
  1020
ggtgcctcc cccgtccgct atacgaccag gcggcagggc tggatggttcgtggtcgacag
  1080
gaggcggcct ctctctcgag cccggccttc cccctccttt cccatcctct cccctctctg
  1140
tcgtctctct ctgtgacagat gtaaaagga atctggcgatt ctaacacgctg acaccttggat
  1200
tgttattatt gccaaatggtg atacagtcg aacggctttc ggtcggtgtctg
  1260
atattttaaa aaaaaacctcg tgttttaaaa aacgaccttg tattcactgt tttacatcgg
  1320
cgatgtgctg ttggttacgt gtcgctgctgc tgtttacta aaatccagact aatccctgattccttaca
  1380
aaaaaanaaaaaaaa aaaaaaaa aacccccccgg gggggccccc gggcccagtt
  1440
coccccaa
  1448

<210> SEQ ID NO: 32
<211> LENGTH: 456
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> LOCATION: (444)
<223> OTHER INFORMATION: n equals a,t,g, or c

<401> SEQUENCE: 32

ggggcacagcc attagaagcgc ataaagatocg cggcttctcc tgcggctgtctg ctctctctcc
   60
toocggcctg ccaactctggcc cacggagcca cggggagggg toocgggagga gaaaagggga
   120
ctggcacttg aatggctcga cccgagctcc ttaacaccoc gcggctgac ccacgaaaaa
   180
tggagtctcg gttgaaagct gattgagccc tgcaccttgcc gcgcctctttt gcgctctaca
   240
gagaagaga anaaaggtgg agctcggcc cccggggttg tcggatatgtg tgggtttctg 60
gcgccgatt ggtcgcggc gcggcagaa tcacaagcac acacgactat tttatgtgtc 120
ttagcaacag tcctacaaccag tagagagaac accacaggtag gacaccacagt ttgatgtaccc 180
tagaggctgg ggcaggcctct gcagagagca accttctgcgc tcggatcaact tcgttacag 240
aatgtcggcc ctctttttttttc tcctttttc caactgtcgtt ctttttggtg 300
gagttcttcc cgccgacaca cggggcgcc gcggcagca cggacatcttg ccctggtgta 360
gaccgcggg gacagtgwact atatattttta gcgcctaaac atattttgca aatcttcga 420
ganagacct tcagagccg ctttgccctca agctgacaca acatgcygga rgtggcggcc 480
gtctccacag gccgtcgcaca aataagggcag ctggctgtgat artgcggccg tgtgtgcacg 540
atcccccccc ccctcttttct ttctttttt ttactacaaggg aagacttagg ctcttttgag 600
gacaacgccg tcgctttag ttacaccagaa tgcggaattta agagacggac ccaggccttt 660
ggaaaacgccggtc aatcttttta acatcgcgtc gcccccagccc gccttcttcc 710

<210> SEQ_ID NO 35
<211> LENGTH: 1188
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 35

gatgcggcttt attatactttta tcgaccccaac gcgcgctgac atgatagtt atcagacag 60
gatctgtgg aggagggcaca gtaactggcc taccaagccgc accttcagggc agagaccttc 120
tacaacattta agatggcagtg tttcaatgaa ggggagggaga gcgggtcgag cagaagtgatg 180
atctgctgaga cccaaagtgcttc gcgcctcgtggt ctgcagctgc acgcccaacact 240
cctggccccac ccagagccgccc ccctcttggaa aacataagagg gcggctgggg gccttggggcc 300
attgtgcgcct gcgcctcctat cgtttgctgtgg ggcggcttctc ggcggctcctc 360
gttccaca ctcgtttcttt catcccccct tcgctctgtt ggggcttggtgta taagccaaaa 420
catacccaag acctttggttt tcctcgagat gcgttctcaca ctctctccgg cttatctagt 480
gtgcctccct gccgggaggg gcggccggtc ggagatcctttc accttccttc agtgggctatc 540
gggagagggc cttgctgtat gggccacaca ctgataggg cttgctctccct gcgtgctgag 600
gtaaaaggg cttgagagtt gcagagagga gctgccgtgag gtggagagtt gcggagagtt 660
cgagagagtt gcggagagtt gcggagagtt gcggagagtt gcggagagtt gcggagagtt 720
tacaagggc gcagaagcttgc agctcttcagct agatgcgtttt ctctatacaca ctgccccagcc 780
aacctccac ccagctgtcg ggcggccaca acgcagcctg ccaagggcgac gggagagactg 840
tcaggctgag gcggctctgg gtggagagtt gcggagagtt gcggagagtt gcggagagtt 900
gggacccgt attttttcct gccggggtct gctgtttgct gcggagagtt gcggagagtt 960
tggcaagctc tgctctccgg caagtttagtg ggagagactg gtggagaggt gcggagagtt 1020
gggctctag gcggagagtt cttgctttcc ccgccatgcg gggcagcttg tggacagtgtg 1080
cctttgagcc aaccctttttt actactaggc aagacttagg tgttccttccc ggacagactg 1140
ttgtttttttttttt gggggggg ggggccc 1188
ggccagacgc tgaattcga tctttaaaat tcctaatttt aacgctggtc atgcacccaa 60
gatgtgacat ctggacagca ggggtaagcc acgtggtata ctaaaccctca gaattgtatc 120
cgagaagct ccatacgctt ttaatttatttg ttcgttggatt gttgtggcct cctttattttttt 180
aatggtgc ctggacaagt ggggtagggc cccactcctc ctgggttgc cccgctgggttc 240
aatgcagag ggccagtggcg ttaaaatatt gcaccgtgag tttctccttt 300
atgaaacgg gccagcagtctca cctgattttct cttgctcgag tcacaaattgg 360
tacacagc gggcctcgcc caaaccctgag ctaatcttctt ctttaattttttttt 420
ctggacagca cccgctctgc ctgccggttt ccctgtgtttc ctggtaagtcgttt 480
gcagagcg gccggcgagc 540
tttataaata taaggaaccag cccattctcgg gcctgttgggg gggcattcctttt 600
gccgagagc ccgcgtagc agagcttcgg cggagtcagc aagccggttc aagatagtata 660
ccttacagcgc cggccggtgc cggggttgcct cctgcctgca gaaagcggag 720
aggggtgagg gcggaggtgg tagagctctgg caagagtcac gctgggatgc 780
gtcgctggag atgatgttctgtggattt cattcctggag cggcctggag 840
aacctactg aaccgctga cggctgaacgt cgtggttcgt ttggtaatccttttctg 900
ttttttattttttt cttgctgtgcg acgctgtgtg cttggtttttttttttttttttttttttttt 956
agagataact ttagtttgcc aattgttaag gacacttgaa ctgtggggtc a ttcacaaga 840
gaacgttcccc ttagttgaca aactagaaaaa actaactggtg ttagttgtcct ttaactaagc 900
gacaacttcc caacagcag gctggtgtt ggtggtgtgt gcggggttct ggtggtggaa 960
cataaggttgt ttttttttta ttttt ttgttacgaa ttttttaaca aggtctccttc cagtcgtatt 1020
tactttaactg ctgtgggttt ttagttatc acagttacac ttttacttta taaagcacg 1080
aaaagcaaat gtagctcccc ggtggttgtt ttagttttta cttttttttt tttttttt 1140
tttttttttt ttggtggaac ttggtggaac ttggtggaac ttggtggaac ttggtggaac ttggtggaac 1200	tagttttttt ttagtttttt ttagtttttt ttagttttttt ttagttttttt ttagttttttt ttagttttttt 1260
tacttttttttttt aggtagttttttt aggtagttttttt aggtagttttttt aggtagttttttt aggtagttttttt 1320
tatggtttttt ttagttttttt ttagttttttt ttagttttttt ttagttttttt ttagttttttt ttagttttttt 1380
cagcaagtgggc cgttgggttt tttttttttttttt tttttttttttttt tttttttttttttt tttttttttttttt 1440
tatagttttttt tttttttttttttt tttttttttttttt tttttttttttttt tttttttttttttt tttttttttttttt 1500
tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
agctcagttc ccttgaatt gaaatattaa atgccactac caggtaacg actgagccca 60
gtgaggtgta tcagcttga caatagcgat taagatcatt ttaaggtgag aagagatatc 120
gaatcaagt tcaacctctaa gatctcgtg agtagtgag tttgtcagtt atctctcatl 180
tctctcttc ttaagttttt ccttgaagg tcctctatgt gctatcttgg gggaccgtgc 240
tccaaataa tggaaaggct ggtttggtcg gttcagagag ttggactttc cccactcag 300
agttttgcgc caagctcttc acataatgaa gaaatcacca gttgagctga cagatttggc 360
aataacttag ctctaggaga tctactgagc tttcaactcc ttgtcctact cgagacctct 420
gtattttctc gttttttgct tttgtgctcc tattggcag gttatatttt gtatattctg 480
ttattttacte cactaatctg cctgttctcc cctctctctc ttttcctttc 540
cocgcggctc gctctgcgac atagctctga gacttttctt gcccaactct gtaaaaatcc 600
ttggactctg accccactct caggtgtcg 629

<210> SEQ ID NO 40
<211> LENGTH: 1964
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (353)
<223> OTHER INFORMATION: n equals a,t,g, or c
<211> LENGTH: 1964
<220> SEQ ID NO 40
<222> LOCATION: (476)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 40
aagaagagcat ggaasattgct gaaagagtt tcaagctat taaagaaatc tttacgcaac 60
ttgagaatt cagagctttc gaaagtggct gaaagacgct aataaatcttt 120
tagtttaaag aagcactctct tagctgatgtccttactc tttggctt 180
actctctc aagtgtgact aacctttact gaaagagct gttcagttcct 240
tggagatgct aacctattct cccctctctc taccgacccgactgtcttg gattaggg 300
tsacccattg gattatttt cggctatatt accctttacc ttaagatattt aagaagatct 360
gctctttcaag acattacttc catgacctttc ctgctttctg ttcgctttgaa 420
gtcttaacgctctgtccttgcgcttttg tagcctgctttaagcagagcagagctgtcgtgag 480
aactctctctctacagactctctcgccactctagc agatctggct gatgtttattg 540
aagaagactt cggctttcatctgcttttactgcttaagttagttcgctgctt 600
gttggagatgct tcaagctttc tcctactactta tctagagctc tggag agaatgtta 660
ttagctctctt tagctctact ctgctctttc ttgagacttgcctcattattactgta 720
aacagcttct ctaaggacttct 780
aatctctgtgcttctacta agcgagcttcg tctagagctc tggagag 840
aatatctgctctctctc tctgttgagc acacggctcag tcggcctactcttgagct 900
cctgctttcttgagctcagc agaagagctgttatattct accacccgta 960
tctctctccttctaagcttct ccacagttgctgctctcc acagctgagcctgcagcagc 1020
cttcatttgc atataacct aacagaacct ttccccata ctgaaagaa tcggagaga
1080
cctcctagt aagctaatct ctaaaanat atggcagaga tcgcaaaacct tgtatacaac
1140
tagtcagc atttgataca gactcaacat cattatatc agactttta caaacttaca
1200
cottcatag tggagaggg tgtaggagtt ccaactcaag aacactaaatt ggaacagaga
1260
gagagcggcc tgagcttccc aagctgagtc atacccagtc caacagaccc cagctgcccgt
1320
cgaanacact cagccttctca aactgacacc acccaccagt agaacggagc cacttctact
1380
cgaanacact tctgctgctt atctgagacc acccactact tygtagagc tgtatctgca
1440
cggcagag ccactccacc tccaggtgccc aacatgccc cagagagccc caagtagagca
1500
aactgagc cttcctcaagg agccatgagc agtcnaanag tcggtagaaa gcttgttttttt
1560
gttatatata ttttctttct ccaccttact acgttaaat attgctttat tgtttatatatt
1620
gttaaaccttt cttgcttct tgtaatatttt ttttttaaat atgtgtatgaa aacacttcat
1680
cttccattg tttcaacagaa gcgtaaatct cttaccttctt ctgatttttaa tcagctagtt
1740
tgcgttctt tcatatctca ctaaactcaca tcatttctctaa tcattttctttg ttcatagc
1800
ttttttagc aactccttca aaagagggac agaatatact gtgaatttca aacgatacc
1860
tgtagtgcc aatagcggag ggaactctgtt ttcctgyttta actacgacaa tggtaaatatt
1920
ckgttttctt ttttcttttctt ttcttgttgc gcttttaagtg aaag
1984

<210> SEQ ID NO 41
<211> LENGTH: 1522
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1282)
<223> OTHER INFORMATION: n equals a,t,g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1376)
<223> OTHER INFORMATION: n equals a,t,g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1482)
<223> OTHER INFORMATION: n equals a,t,g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1492)
<223> OTHER INFORMATION: n equals a,t,g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1501)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 41
cgtgtcccg cgtctggaga acqctgtctc gcqccggagc ccqccggtg
60
gaggttgtc gcacqctggga caqctgtaac cagagttrag tcggactgc cgcqctggc
120
aggcctgtg cttgctgtgtc aacacgtcttt ccaagctcg ggtagtggtc atcacaacc
180	accggtgtg gtctgttttc caaqctgttt tttgcaacag cttcggcag cagqcgtgag
240
cattggccqa cagqcgtgagc gtctgtgtgc cccagtggga ggtatagcat gcgtgcag
300
tctaatgcgg cccagacccg acactctgtjac ccgag tgtaatctct gtagctgacg
360
tggaagcaat cactcggacc aatgctgagc caccctgacc gaqcatgaca caccagcgtg
420
cgcagcagc ggcagcagcgc acacctgtgac tgcggtagag cccgcagagc gacqggtggc
480
cgagagggc ccqgaaagag cactcgtgag ctcggctgag ccgcttgtag gcgcctggag
540
cgtgggtgtc atagctgagc gacqggtgag cctgggttggc gctgtgtgtc gcggagaggg
600
-continued

tgcccagtc cacgccgtcc tggctcgggt atgcttttca cacacaagtg aaratttatta 660
accagaacct acgcacccat ggacagacct taagggatgt ccaggttcnn ccagggagc 220
gttccagcaag gttgcacgta taaggttcaggt ctggttcacn ctggttcaggt ctggttcagt 780
tgacgagag cctgctgcat tccacagcgt tctagacacg gcagacgctgtc aagagagcag 940
tgactctctat tttataagcga atgatttcac tataacccat ctttaatatact ctacactaat 900
actacatgt gaaataacgct attacatcgg atttgatgag ctttcatata atacacacat 960
ggaatactc atctgtgaat ctaacacagg tgcctttatt ttcacaggtc aggttccac 1020
ttaaaaatc atggagaca cagttttcaga ggaagctgaa acaaaatca caatatttacatg 1080
taaaaaaat atgtgtttata ttaaaagcga aacctagagc atgcttttcag cagacgcatg 1140
actggtttg aggagatcagc gaatcagcaga agggatcatac cctttcataca tgaataaatc 1200
caactaaaaa atgggtacat yatattataa ctaaggaacc acaacatcgt atgtttaaaa 1260
aacacatgg ccagccagtt cngacatcaaa gtaagcactg caaatgatcgg gaaatcccg 1320
cocactgca ccaaggggtct ggatcctttt actaatctgc tctctctgtgg tggantcoca 1380
aagacaccag ctttctgtcc ttaaaaggtgt gtgaagcctg gttanngggg ccaggtcagc 1440
cactggtttctg ttaaaatnct ctggttaacact aacgctcata aacgttttccc ctttttcatg 1500
noatcagacagcttttcagagc 1522

<210> SEQ ID NO 42
<211> LENGTH  875
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

tggtgattc ctatttactg gaggctttgt ccacattcct ctatgtccct ttcctttgtg 60
tctggtctg tgggaggccc toacttgcgoc tcggctttcc ttggacacan ggcaagacct 120
tocaagagatatc ctcaggagaa ttacacagc taaccttccc ccagggggc cagggcoca 180
cgctgagag cctgagggtt aatccctcaa caacatctta gggcccaggg gtggggtgta 240
gcacaacgcatctgcagtc tggcgcagc gcggctggcag ccctctggcag ctcagmcyct 300
gctccatgct ctattttacag aggttttaatt gacacatcct tgggtctgagct cttcacgta 360
ggtggtcagc atcactgtggc gacgctgtca taacgcgacta aggtaatccta ctcagagrrc 420
atacaggtaca aacataatcct ccgtgagtctgctgg cagctctgat cccacacttctc 480
tgggagccag cagctttcag actaatggtg gcaagagttt aaamcagctt aggtaacatt 540
tgacgacccc ctactctcact aaaaaaactt aacattaccg tggcactttg ggtatctgct 600
cacactgatg cgtactgcag atcagaggtg ccaggttcaacctgtaacta aagagcttcag 660
gacgctgatg ccgctgatg gttgtgctcag ttcggcagcg acacgtcacta ctggttcac 720	tctcttttaa acaataact aacatttaca gctcagggaa atcagctgga agaaatcatac 780
aggtggagcgg ggtgaggctg gttgatgcctg ttaaacccgacttggggag gccagatcgg 840
gaaagatgtc tggagacccag aatgttggcag acagc 875

<210> SEQ ID NO 43
<211> LENGTH: 843
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:  
<221> NAME/KEY: misc_feature  
<222> LOCATION: (44)  
<223> OTHER INFORMATION: n equals a,t,g, or c  

<400> SEQUENCE: 43  
cocacgcgct ccganatgct ccctcctcgc ttccagagagt ggccagagct gataaccacc  
agagggacaa gtaagggtcc agttccaaaa cattcactgag aggtatcttc ccactctgct  
caacacagac ttacagagga aacgccnncc ccagnatgcac gigtctttct aagggacaaa  
tcaacaggaa gtggctaggt ccggccacca tcggctgcgtg cttgctgcgtg cccatgagat  
cotacagact ttctcagcgc ccctaaatgc acagacccct ttacagcagac aacattacgc  
tttgtctgaa ccagacccca ccggctgcgtg cttcagagga ccctcctcgc ttccagagagt  
tocacatact ccctcctcgc ccctcctgcc ccctcctcgc ccctcctgcc ccctcctcgc  
ttttcttgtt ccagatggga ttccacactc cctctatcgc cggctgccct cggctgccct  
ggctctctgc ccctcctgcc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
cctcagagag ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
tttttcttgtt ccagatggga ttccacactc ccctcctcgc ccctcctcgc ccctcctcgc  
ggctctctgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
gacggtgtgc gcagcctgtg cccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  

<410> SEQ ID NO 44  
<411> LENGTH: 489  
<412> TYPE: DNA  
<413> ORGANISM: Homo sapiens  

<400> SEQUENCE: 44  
ccttcagagct ttgagagcct tttgctgtct cttcctggtc cttcctggtc cttcctggtc  
tttcagagct ttgagagcct tttgctgtct cttcctggtc cttcctggtc cttcctggtc  
acacacagcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
ggctctctgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
acacacagcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
ttttcttgtt ccagatggga ttccacactc ccctcctcgc ccctcctcgc ccctcctcgc  
ggctctctgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
ggctctctgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  
gacggtgtgc gcagcctgtg cccctcctcgc ccctcctcgc ccctcctcgc ccctcctcgc  

<410> SEQ ID NO 45  
<411> LENGTH: 534  
<412> TYPE: DNA  
<413> ORGANISM: Homo sapiens  
<220> FEATURE:  
<221> NAME/KEY: misc_feature  
<222> LOCATION: (470)  
<223> OTHER INFORMATION: n equals a,t,g, or c  
<221> NAME/KEY: misc_feature  
<222> LOCATION: (477)  
<223> OTHER INFORMATION: n equals a,t,g, or c  

<aaa>

<aaa>
<400> SEQUENCE: 45

gagctctgct gtaactatga ttatatctct gttcattctat attttttttga ctagtagcnn 60
caacccctac ctttatacag gccactctcct cgcggcgggc gtctgttggtg cccagtgyttg 120
gcttttggac gctgtgagga tcgatgtgctc gacgtatatt tctttttgtgc tttaaaacgt 180
ttttttttct ttttttttct ctccttctct ctctctctct cctcttctct cctcttctct 240
tctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 300
ttttccttcc actcttccac cccatcctcc cccatcctcc cccatcctcc cccatcctcc cccatcctcc 360
ttttccttcc actcttccac cccatcctcc cccatcctcc cccatcctcc cccatcctcc cccatcctcc 420
aaaaa cccctcccccccc ggcccccccc ccccc ccccc ccccc ccccc ccccc ccccc ccccc 480
tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<211> LENGTH: 596
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 47

gaattcgca cagagttact tcgccctgaa aagactcagg ttcaagtta ttcatttact 60
aagttgtta atacagggc ctccactgag cctctatttg gtaaattgaa aaatgaaat 120
ataacacta gtaacctata ttctcaacctg ctccacactag gcaccaggg ggctactaat 190
acacaccaca tccctccctg tgcacaggtatat atagagaaca attccatcga aagaagaactcc 240
ttcacttat ccttttctctc atatatggtc ggtttagaact ataataatattg tcctagaaaa 300
acccctatga tattattttag ttcctgttgca gaaactgcgtc agaaaaaata tgcatttataa 360	tatatttcct atatasacata atatactaaa atttccat aatggaaact tcatctaaaaa 420
tgagaaaaag tcggcattttc ttctcattata acagttacaa attctcatgt cacaatcttt 480	tttttttttt ttggttttagg ttggaggtcct ccctgggtcgc cagggctgggg aacagagcag 540
gacgctgtc ttatattaaa aaaaattaata aacactgagg gggggcgggt aacota 596

<210> SEQ ID NO: 48
<211> LENGTH: 851
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

caacagagga cacoactgg tagaagcacc ttcctgcgtgc cccgctgcttc tataagcctct 60
cctaatctg gtaatagcag acgcacggag cggactgctta agaagagcag gctttgg 120
tgctgttctt gctgggtgcc cttggcacc taaagcctctc ggctgcttcagcctctctg 180
caccttgcg acaacctcct ctgggtcact ctttcggcata cacaagctgac ccagctctctg 240
aacatcaacc gtcacagcag ggccagagcag ccctttcggct ggcacgctctctg 300
catatctac ggcacagtgc agacactaacc acgctcaacc ggggttctgt gcggccagcctgt 360
ggcgaggg gcctgctcttc ctgggtcctctc aggggcgggg cgggcaacgcat ggggccccttt 420
cctggtttcag cctggtgccc cccgaggggctc ccggaggtga cctggaggtgc ccctgttc 480
cagagcctgg caagactgctt tctgttcacctc caggtgtgctg ccagccacag 540
catatctacc cccgtctggtgc cgggcaacg cggaggtga cacagtata cctgggacc 600
agccagctgag ggggggtgcag ggccagcata cggagcccagcc tggcccggct caggtgtgtcctg 660
cctccgagc ctggagctggg agcggcaggc taaagggggtc ccgaggtggc gccaggtggg 720
cacaaaggg ccggggtcctgc ccggagtagc cccggttatt gttcatttaaa tttcagttgcc 780
ttcacagatt gnaacactagc ttcctgtctga atggcaatg ggggcccatt 840
tgcatctgcc c 851
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1567)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 49
gtgaatgag aacagccttt ttatatcttt tagctttgag gtttggaggt ttgggggggc
60
tatctggttt ttgctttctt cttgcttttg gaggagagtgt gagggttttc ttaagggtcat
120
aacacaacc aggtgaacac gctgaactgtg aacgctgctt gtatcgggag cttggtggtg
180
caagagggg atgcaacaaa attggagag gttcctttggt cccaagctct aacctctcctta
240
cctaatagcg ggctcaaggt gcagtgcaac aagctctctt gctcgtggtg ccctgggggt
300
gatggtgct aagctgctct gagggggggag cttatcgtat ccatacttgcc aagctctttc
360
cactcagct cgctctcagg aagatgagcg ggagattctgag tgaagctgctg actgaagttt
420
ttgcttattg gttgtgtaag cgctcatgag gtcgaataag aataatggggt gatggagggg
480
ggagggcaca ggcaaggtgg aatgcaagct gcctgtgggggt tttccaaca ggaattccaa
540
gacaatcctg agtgaatggt aacgccctccag agtattaact tgtttccttt gtaaaagagc
600
ccttcatcgc cggtcagcgc cagcttggga gcggcaagtatt ctcgacctgg cttgggtctg
660
tttcctcttt acagcaggag aacgaggtctt cctctgagct gggagacggc taaasgtggt
720
attgaaagg ggaagggccc tttgctgtct gttgactcgt gtcactact ctcctgtgca
780
cacctgggt gcagtgcaac acctctcttt gcctcggaga agacaaggagc aagtgagcag
840
tccataagc tagagctggct cacgggctaa cccctgggcgg cccgaagggc acaggtgcaag
900
cacctagata gctctcatgg caagggctca gtcaggggca agctatggct aagatggtcct
960
tgtagacayt ctgtctgacc gtcacatcacc caagtttcttc cttatgattc gttacatcgc
1020
tgtagacayt atcataaatc gccacattct ctacatcctt ccacatcacc atagtttataa
1080
tctgtgctt ttacaggcag gtaagtttttg gggatctttgct tgtgttaagg
1140
aacatcact gggagagctt actcttccgt gcataaaatg ttggacatgg ttcttagat
1200
tgacacac acaacacacg ggtgctcttc tagcttttac gcgggaagag acaagacca
1260
ggaagactt ggtgcttccc ccactctggg cgcacatctta gcacggcctg aagtatggcg
1320
agaatctag gcgtcagcgc gttctgctac gttgtgaccc taacacatata accagagaga
1380
gcagagacg ggaacgctg tggagataac aagctgacta aagatggcttt ttatatttaa
1440
cagaaattt ggcctgttttg taacctgtgt ttcttttac tggcggggttc attccacacc
1500
aacacacc ttcacatcag ccctagagcg ccctttattg tttacattg tgcctgtttt
1560
gggtctcgg gcaacattaa gcagcccctt ccgctttttgt gcaacattc gttgctcttg
1620
gggtctcgg gcaacattaa gcagcccctt ccgctttttgt gcaacattc gttgctcttg
1680
ggggcattga ccttgaaagt gcacaaagag aggcttctcag gcgtggctgg gcagcagacg
1740
gttgctctca ttctcaacct ttgggctttt ttaaaaag aaaaaagga aaaaaatttcc
1800
taagcaaggg ctctggagca gtttaagcag ttaaagtctt cttttctttc ttatgagat
1860
gaaagtgtt gttgaaatct cattgtagt tttgtaagtt ttctatgac taacagtat
1920
taagcaaggg ctctggagca gtttaagcag ttaaagtctt cttttctttc ttatgagat
1980
aagtatgtt gttgaaatct cattgtagt tttgtaagtt ttctatgac taacagatat
2020
<210> SEQ ID NO 50
<211> LENGTH: 2432
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<200> SEQUENCE: 50

atgagaggtc gttggtgga aagtgagccc gacacctggg acccatctggg tccttgagcgtcgc
150
agtgggcccag atggtttgtcg gcttgggtgag gtctcaaggt atcatcctcttttcgttctctcag
225
tggtggtggt gcgtgagcgtg cgcagcgtg ctgctgctg ccgtgagcc
290
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
355
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
420
gttggatacct gcacggtcact gggttcagct gacatctgccctggctctg gggctgggctgtgggttct
485
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
550
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
615
tggctgagcgtg cgcagcgtg ctgctgctg ccgtgagcc
680
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
745
tggctgagcgtg cgcagcgtg ctgctgctg ccgtgagcc
810
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
875
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
940
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1005
tggctgagcgtg cgcagcgtg ctgctgctg ccgtgagcc
1070
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1135
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1200
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1265
tggctgagcgtg cgcagcgtg ctgctgctg ccgtgagcc
1330
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1395
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1460
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1525
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1590
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1655
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1720
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1785
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1850
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
1915
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
1980
ggcggcagctg cggcggtcg aaggcgcgtgg ccggctggtg ctggtctctc
2045
atcgggatcg cggccattcct gcttgggtgt gctgactggt ccggctggtg ctggtctctc
2110
-continued

ttttctataaaa gcaaacagag caaatctgctt ttcacatgtc caagtcctgag agcacaagag 2160
cocctacgt gtgcagtcgg aagagctcct gtctctgacg agcacagcgct gaggctggcg 2220
cctgtcaac cctatttcca cctatttcca ttaaaggagag ttcagctgtc gcctacattc 2280
tgcccctgag tttaaactcag tctcagccac agttctcttg gaggcccttg aactcttgctg 2340
cocctatctg aacctctgtag cctaaatgaa atttataaata tccagccgacg aaacaaaaaa 2400

aaaaaaanaaaaaa aaaaaaaaaaaaaaaa aa 2432

<210> SEQ ID NO: 51
<211> LENGTH: 2340
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (96)
<223> OTHER INFORMATION: n equals a, t, g, or c
<400> SEQUENCE: 51

gacgtctgggg gcgggtgggg gcagcgggta ccgcgctggg gcgcgaggggc gcgcoccttc 60
aatagntgca gcacggagag cgcgtcggcg ccgagacgcgc tccctctgctg cgccgagcgc 120
agccagcgg gcacccctgcg gcgccaggcca tgccagtgta gcctacctg tgaacactga 180
gaagtgattg ctgagcgtctgcc ttcgctgtat gcgtcaagcgc agcaagacgc aagatgagaa 240
ggcgcgcttg ctgataaaaa gaactcctca aagtctctat gcgtgctgat gcaagtctggag 300
tcctttttat ccctctgaat cctctttact ctgagaaagtt tcgtcatgaaa aaaaagattt 360
atgtagcacg tgaacatctg aagacagcgc agaacttctg tcgaagagctgca ttcagagggc 420
aagtgtgcttc cgaactctca ggcacactca cggagtctgt aatgtggccc aggttctcgag 480
tgacactact ctcttttctg ctcacagcgc ccacgacgac tgcagatgctg tcaaaagtgc 540
atctctctt tgtgctgctgg aagtctctca gctaaacgctt gcaagtcttg ccgtctgtcg 600
cagaacagc ccocctgctga cactgagac cggaccgcgct cgccgctcgct gccgcgcttg 660
 gaacgagcttg aaccttctacc gcctctctgc tgaagctctgg gtggctgttg 720
tgataaggtt aacttcgcttc ccacgtcgag gatattcctat aactttgtaga aaaaaactt 780
tctataagct gagctctcttc ggcgcgagct cctgctctgg tcttctcttt tctttctctt 840
aotattttg tgtgcttttta ttcagactgtg tgtgttcctta ctggtttctca gcaagtagttta 900
anaagacagc cccgctcctg tgggtacgct tcttttcttg gccgaagctg cgggaaaaa 960
ttoctctttgc cgcacgtctgg tcagagttgt gcaacgagctgc tagacaaaaac 1020
tgtgactctg ccagcactcg tccagctgt tggcttgtc gggggcgcgag aataagacgt 1080
cocccacact gtgtcctgctt gctgctttcttc aagcacaact tgcgtcgagtg atctagctttt 1140
ggctttttttt ggttagctgg tcaatcaacc cagacactgt tgcgtctattg tgcagcagca 1200
atacgcttt gccatgctt tgcagagctg gcataatggtg aacaagaaaa acaagccagct 1260
ccttaagcttg ttcacgagag gacgctgaag gctagctctg gggggcgttg gatgtgatt 1320
ttgacaacag aaaaactctgg ttaaaatgcg aactctacgt cggagacgtg ttttggccag 1380
cotctctgtg tattttttcc atctcgcagag taccttgaag tgcagctgtg ctgaaaaactt 1440
ctttttttttt ttttttttttttctgctacat tttttttttttt ttttttttttttt 1500
gagccgctgtg tttgttggc cacccacctg tgaataatgca acctagactg tgaatgcgaa 1560
agccctccta aangagaattg tatgtaacga tgttgttwtg attttaaga asgtaatatta 1620
attgttaaa attctgttgc tttcaartgc acaatggaats cagtaacta attggaagga 1680
gagggsggt cactcttnttg atgtgggccc tgaacootat tctcgttccc tctcgcttgtg 1740
cctgtgtgac cccagcgaggg atccactccgc aggtacaggt tctcgttcgag tcctcgtgtg 1900
atatcggctc tcgctatccac cgggctggagg cctggcctgg tggagaagg ttccaaacctt 1960
ttcgctggt cctcgccttt tcgtaanagc tcggaaaaco accaggggag cttctcgcga 1920
ggtcctctct cggagagggga ctagaatctg gtaatctctg aaacacttga agccacccctg 1980
ggtcacgac acaccgatct tcoccgccat tattcaaat tcatgaacctt tggggaatttt 2040
tttacctcgc ttttaacagca ctaatgttgtt acattcggcag cttccctagtgc tcgtcatcag 2100
gtacatcaca aanaaacagc aaacatttca ttaaacagag tcggagaaag tcgtaaacct 2160
tocagaaata atggtttggag aaacactgaa atggaaccttt ccgctatatta taatctgtgt 2220
attnaaaxx aangaaacctt ctctatgagc tcctgyggct gtaaatcccg cagctggccat 2280
gttttaaag atggaaaaaa ataanaaactt ttgggaanam aaannnaaa aaaaactcga 2340

<210> SEQ ID NO 52
<211> LENGTH: 601
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (115)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (184)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (539)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 52
agtaggggag aagcagtatg cagcggggag ccaagcgcag cccggacagc 60
tgatcgaggg cccggagacg ccctagcggc ttcoccttgc tgggcggctg 120
ccttccttt tccctgaggg cccggacagc aagtgctgac tgcctgagcct tcccgccgaa 180
taagggggct tccagaggg gaggaaagcc tccgaggata cctgggagggc gggcaagcag 240
gagccaggg ccctcctccc tccctgctgtg cagcttatac ggggcaacggc accgcagct 300
tgtctttaacc ccaatcttcg tctctctct ttcagcatgg tgtgttgatg gacaggtgga 360
cagacagag aagccagccct tccctactgc cggcccccgc tccggacacc tacaaggtcct 420
tgtagggggc cggacacctg gatgtgggac cttgatgtgct tttgctgctg atgagctgccc 480
tctctcttc caagtttata tccagtgtttct tccgtgggac gcgtctatag gatgcttgag 540
aagctctctg tctgagtcct tcctctctgc agcctccgga aagagaaaaa 600

<210> SEQ ID NO 53
<211> LENGTH: 359
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (343)
<223> OTHER INFORMATION: n equals a,t,g, or c
ctctgtgccaga atctgagcag agagctggtta ctctttaaagag taatt aggctg gtaaagagt gtaaatagag tgcct gctagatg 60
gatttcatct tttcttagc agaagcgaac ggtttctocc ggtttaagtt ggtttaaagag 120
agaagcgtt gttttaaaac aaatgctgctc tttctatat ttttctttttag tttttaaa 180
ctctgtgccc tttcttcttc ttcagatgct ttgctgcrcc actgagcgeact ctgagaancc 240
cacccagata acgtgctgta ttcattgt ttcacgccaact cccgcccata gtaaagagtt gtcct ccattataa 300
actcccttaa ataaatccaga aaaaaaaaaaaa aaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa a 359

<210> SEQ ID NO 54
<211> LENGTH: 1141
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54
ggacagagct gccagagct gagaagcgcc ttcgagcgcch gcgagcgcg gaagcgcg 60
ccccccagac gttgggtatat cgcggagcga gctctggtgacta ctcagcagtg 120
atattatac actacagccact gctttgtgact ctctttacct gcaagcagat 180
aacagtcgctt aacctggctt ggcacagcct ctccttgagaact tttctttaca 240
gattcagccg taactggt ggcctggctact caagcgtctgct ctgtctctgctctgctctgctc 300
ggcagagcata ggaagagcg ccagacggct actcattgtgact ctgttgctcctc ctcttttctctgctctgctc 420
aacagcgccctt tttctttttt ttcagtcgct ccagcgccctt tttctttttt tttctttttt 540
cagcctgctgctgg atgtccgcgg ctacccgctt gtaaagagtt gtaaagagtgg 600
cttacccaa cccccctcggc aacagcgcctt cccccttcggc ccagcgcttg gtagctgctgtg 660
aacagctcca cggagccaa ctttaccttct ctgtttttttct gttgcgccctgc gttttttata 720
caccaacagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct
```
<222> LOCATION: (8)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1428)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 55

gagagagag gagaggtatc actgtcaggc tctatagtct atttttaaat caaccaatct 60
tatctgagc asqaqcacgct tggtagttgtc tggtaggtcag cacaactctat ctctctat 120
ataataaaat tccacactttt tccaaattta atataggtta aagtagctct cccacactt 180
aagcagagag aataagatgct ataactgtcgt tcaggtctac cactagctctt cactagggg 240
cagagcttt tttataaata accatgtgct tgtacatctaat tgcatacctgctctactct 300
tagcattta taagattatt ctggtagcct atcaccaaga aqaacaaact 360
actgagagttt cagatgctgc tgtggtctctg cattactctt qaaagctct ctaaaagact 420
tcacactaas atttttctct tcagttcattt atccagctgg ccacaaatca caaatattta 480
agtataaca ttttaaag ataaaggaga ttttyggtgca cgctgtccatt aacaacageg 540
acacagcgtc aagaacacta acactaaacgg tacctctct gcacccggtt gataattac 600
acacataac acataatcagg ctggagctag gcacacataac cctttgttcaatt 660
ttcctggtca ccaaaatagag gcagaggcata tcctataaac atttttagctg aaaaagaaaa 720
acacccccga acacacactgc acaacctgac acatgcacgct acataactgc cagctttctg 780
tttttcta acacatacata aatccacactgc aagctctcag aacccacactc caaanatgaa 840
acacatctg cctgaagctgc tagcagagctc tagccacaga cttgaggtttcct cttagtttc 900
atcagagactc tccagggaaa tctcgggacc aacaagcgg cagagggcct ctttttttaaa 960
cataccttgc cactgtcagc tggagcattg caagcactac gctttctaa aaccaacacaa 1020
aacacactgc tataatccta ctatttttct gcagaaactg ttaaaggcata cagttataaa 1080
aaaaaattt accattacgact ctaaaagagc actctaaaggg gattagcgcc aaaaacacaa 1140
cagagctctg cagacattgct ctttgggtgg aagcatactg aagccagtac 1200
gcagtagtcg gctctcccc gctctcagc gcctgtcggg catttaaact cattgtcag 1260
gagcgctcg tctgctctcg acggcagcgc ctagcctcct ccggaaaccc ttggcctgt 1320
cctgttcctc gttctttcacc gctatatatc ctaaggctgg aagcttttgag 1380
gttgtggccaa ttgtagtctgt gtaagagcct acggctttgtagcgtggct 1440
ataggctcc gacggtcctgct cccgggagtta ctctctggtg tacaacacta aactgttagt 1500
catgacgc acgtctcagc tcctcgactgc ccgtgaccaac gagagcggccc gagagagaca 1560
```
<400> SEQUENCE: 56

ggaacgcaga ggcggagctg ggagcggagc ggaagctgga taaaccgggga cggtgatggtg 60
ggacgacatc ggctgctgt gctcttggctg tacagggac caagggcctg gggacgcacat 120
cocccagaac ggggcatctac gggcggaggg gggcggacca gggcggcccc ctgagagcag 180
cctccccacatac ggccggccac ggggactccc agtcggacca ttgactttttcttgagcaggg 240
aacgagcccg gcaatgtcag cacccccccc cagggagaaa ccaagggcctg tctggaggaga 300
tgcggagcg cctggagcgcc cggcggggcag cggcatctgt ggtgtcgtg ctcgcgttc 360
gcggtggtatt gcgccgacag cccacccgag acataacgggt ccctcggtgg ggcggctgg 420
acaggcaga cccacggacag gccgggaggt gctgggtggg gggggcggac ccacagacacct 480
atggcagctc cgggctgcgtc gaaagacccc atggctggg caaggtggag cactcacaaca 540
agatggtgctc cgggctgcgtc cggggtgtaga cccggtgccg ggcggaggg gccgtcagtt 600
cacgtgaggg ggggtgacca gctccgtcgc cccggtgccg ggcgtgacca taacgccgca 660
tgcgtgttgc ggaaaccacc cgagacgctg agtgggttgcg gtcgggttgg 720
agggatcagc ggcggagctg tctccgaccc agcggaggg ggcggaggg gcgtggtgctc 780
agagagggag cggcagtacgc cggggagctcc ggggtgatga ccacgaggg gcaacctgtc 840
ggaggtggtt gggcagcact gctgtgctcc ctcggagaga cccggtgcct cttggagcca 900
acacgtctct gcaagccgcca gcaacgaggg gccggcagcc carccggcaca gccaacatcc 960
tgggtacttc gcacgttact ggcggcagcc cccggtgcct cttggagcaca gcaacatcc 1020
ggcggcagcc tggctgtgct gc dccgccccga cggcggacag gctgtgacct gcgtgacca 1080
acaggcaga cggccggtgt ggtctggccc cctcctgtgc cggcggccag ggcggaggg 1140
tgcgatccca ggcggcagct gctccctggg ctcgggcgtc ccgcgtctgc gggtgctgct 1200
cctggtacac cccgacccca cagggaggg ggctaatagc cccagggag cacgaatcc 1260	taattctgc gcacggccac gggcggcccc ggccgcaggg ctgctctgta ccgtcgggcc 1320
agaacccgcc cccacccccc acggtgagatgc gcttgcggag cagccgacaag cccggtgcct 1380
cctggtacac cccgtttcct gcgggtgctg gcgggtgcct cctgtcctc cccggtgcct 1440
aatcagaac cggcggcagct gggggggggg gggggggggg gggggggggg gggggggggg 1500

<210> SEQ ID NO 57
<211> LENGTH: 450
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57

ttttttttact gqaagaaatg tttaataagaa tttaaatttt ttaatctcagg gaattttgaa 60
gttcaactct tttcagagag ggtgtaagga tgatataaat cctgagagga gcaatttgaag 120
acacccctt ctgtcatcttc tgtgagactct ttcttcatct attatctat ctagatatttc 180	tttctctct gctggtccttg cctgaagtttt ctgctctttc atttcctctcattcctct 240
cattggtcttc gattgttctgc gggggttgag aaaaaact attttgagaa atgatgtaag 300
aatgggtagc gnaagagagg gargatagag gccaggagta cctgccgagc caggctgcgg 360
cctccactct ctgctctcagctgctgcctc cggcagctgg gcacggatgg gcgggctgg 420
acgtgttgag acactgcctcg tgcgaattc
ggcatcgcg ccatggagca gaggagggcg ggtggtggag gttcgtcggg 450
<210> SEQ ID NO: 58
<211> LENGTH: 1147
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 58

<210> SEQ ID NO: 59
<211> LENGTH: 777
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 59
cgaaggggtaaagaattggtg aagacttgca ggcgttgttt cgccctgggcc gggctgggcc 600
agctgtcagggccgggtcttc ttctgacatt tcagctcaat ggtggtgact ggagcccaggg 660
agagcctccgagccagtgatt caccatatgc ggcctctggt gcggctggga gaattacca 720
tgagatcca ggtttaaaaac aaccacacag gcttccttgag cctacccttg gtcagga 777

<210> SEQ ID NO: 61
<211> LENGTH: 1350
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KAS: misc_feature
<222> LOCATION: (1350)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 61

cocccoccccgccccacccg ggaagtgagct gctgctcggcg ccggagcccc sgagcgtgcc 60
---continued

gagcggtcga ctgccccagc tccggcgcgg ggcgggggag ggcgcagggc gcgtgggctt
120
gactacgcttt cctggtggag ccctcagcgt gcgctgccttg ggcgacgggc gcgacacaga
180
tgctctctct cccccacact ccggggctgc atatctatgt ccagggcggt gcggaaatgga
240
cgaaagttgc gagctagcagc agcgccacccg gcgcagggag ttgggggggg ctgctgtggtc
300
agagaggcag ttactgtgacc tgcggccacag ccgcaaaaag ctaaggttaag agctggaggg
360
ggcgcacag gacgcaaccg ctggtgactc tgcggagaaa cgggggtgct gcgtgggctc
420
gtggggggcc caacagaaac aagttgttgct ttggggaggg cttcctgtgct atctcaactc
480
gttaagggag gcagggagat gactgctgggt ttcggcaagc cttgcaagag ccacattcgc
540
cgcagttgct ggcgagggag tggaccccaactaaagctgtgc ggtggctggc ggcctgtcg
600
dttccacagtg tcggagttct tcaagctcaca tcaagggcct gggagacgcgg ccggaccaca
660
gcggctgggg ctgccccactg cccctggggg aaccagtggct ccaacccgag cgtatgctcc
720
gctggcactag cggccctgct gttggtgatgt gcaagaggg cgggaggttg gaagacaccg
780
gacgcaagtct gcctcctggt gtcggccaga agagagggac cgcggcgcac gcgtggggtc
840
aagactccg acgtggtggc gtcagacttc ccaagcaaca cagcgtttggc gcggacggcc
900
gacgatgctt cccggtgttc tctactatag acgcgcacgg gggatctgctg aagctctgagc
960
gagctgctgag ggcggctcag cggccgtggc gcggcaggtc gcggcctggc ggcgctgcct
1020
aagactccg ccagacggcc agtcctggag gtcggccagc tggcgcctgg gcgctgtgcct
1080
gctgcaccc gcagacgttc agccagatctcg gctggccagc cgggacggac gcagctgtct
1140
gcggctggct ccacaacgtg atgggtgggag gcagtgtatt gctgggttgct agagagtggg
1200
gtcagcgtgc gcggagacct acagaaactt gacccgttga gaattgagct ttctcactcc
1260
gagctgctgcc gcggacggag gcgggctgcag gcgggtttctg gctgttggagt cggctgtttc
1320
ttggcttcgcc actgagggag cccatagggg ctcggccttc cccaaagggg gggacagatc
1380
gggagttttt ctctactatt ccagagatcc gcgttttttt tctaaaaagt tcctttttacc
1440
ccaaaaaaaa aaaaagggg ccacgagac cttccttctg cgggtttact ccacaaagttg
1500
aggtatataa atggcaacttgt ttcataaaaa aaaaaaaaaa aatagcccct cgggtctttg
1560
aactgncgag acgctggtttg
1580

<210> SEQ ID NO 62
<211> LENGTH: 1117
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 62

gccacaggg ccgggggac ccagggctgag ggtcgtgctgaa ggcggggtggt ggcgggggc ggcggggtggt 60
agccctgggg ccgggtggttg tggccggtctga ggcgaggggt gcgggggtggt tggaggggctg 120
gcaaggttgg cagaggtttttt cctcagggcg ctcggcgtgtgg cttcggggtggt ggctggggtggt 180
agtgtgggct tgtggggttgg gcgggggac cggggggttgg gcgggggggt ggcggggttgg 240
cgcgcgggtg ctcggggtttgg gcggggggttgg gcgggggggt ggcggggttgg ctcggggtttgg 300
ggggggggtg ctcggggtttgg gcggggggttgg gcgggggggt ggcggggttgg ctcggggtttgg 360
tggggttgtt gtcggggtttgg gcggggggttgg gcgggggggt ggcggggttgg ctcggggtttgg 420
gtcggggtttgg gcggggggttgg gcgggggggt ggcggggttgg ctcggggtttgg 480
| acgcyacagcc gcctccagg aagcagcttc gcctggygca gtaaagcaca ggcgaagtgt | 660 |
| atccatagt atgctgtgag gttgggaaaa gcccctgagc agagaatcaca agagagctc | 720 |
| tggagcaca caacagcagc tcttctctat gcgaagact tcaagacact gattagata | 780 |
| agtggaaac tcaacaagag catcttggaag gttctagaaag acctgctaggg aagagggcag | 840 |
| ttctcagcag gggaactgtcc aaaacagctc caaagggaca cagtgacacc ctgctatcgt | 900 |
| tttgaaag acatcttttc agggtctaca caaagccttc cccatccaca aaaaacctca | 960 |
| gggagccttt tggagggaaa acagagctca tgcgaatgtg aaaaacttat aagttgacag | 1020 |
| aacctggcaac acgaggaagt aagaaataa aacacagcct tgaagaatgt gcacagcag | 1080 |
| atggagcagc tggaaatctcg ccggagatac agatgaagat tgaagaatcg gcacacattg | 1140 |
| tattacctgt atccaggtggt acaatgacag caagcagag cococaagct caggtctatg | 1200 |
| ttaaatcaat aagttgtgta agtaaaggaa ccgatctact gcggacoatt cctggcagac | 1260 |
| acaagagcgg gaaatctaca atcaactgtg taanattatc ttggaaannaa atctctctga | 1320 |
| atctcagat gataatcaccg ggtgacatng ataaagctatt gcgaagcttt tgggaatata | 1380 |
| ctgttgacac acactggttc tgccttcatc ttgctttactg tcgaactttg tttgatata | 1440 |
| cggatagca tgcgaaatcg tcagttctgt gcagctgtgg ccatggaact ctgtcttttt | 1500 |
| tttgatgagc tttttcttgt ttaagatgtga tttggcatgt ccatcaacac ttgggaatata | 1560 |
| tttgctctgg gccaacggtg atttttatca caatacttac cttaaatcact tcaagagaga | 1620 |
| aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa | 1668 |

---

<210> SEQ ID NO 65
<211> LENGTH: 1353
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1222)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1341)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 65

| gggtcgaccc acgctgcccc cccagctgccc ggtacctgc gctgtgctgt agacaaggtg | 60 |
| gtctgcttcc ctgcgacctc caacacacct gcacagcttg tataagacct ggtgctctct | 120 |
| ttggaaacac ggccacagaa gaggtgggagc gcgtctggaa tgaagaatata gttcaaccc | 180 |
| gtctgctgcc ccacagacct acctacactc ggtgctctct tccacaggcc gttcaacctc | 240 |
| gcacagcttg acctgctttt gctttgagtg caggggtcct ctcttctgctg actgcccctc | 300 |
| ttgtaaccttc ggccgacattt gaaattcatttt ttaagctggt gattcctctgt gcgtgtggtg | 360 |
| cagccgtctgg ccacacagcct gcctgctcttc tctacatctt gatagctgga | 420 |
| atggagtctac acaacttggtg tggacactcg gaaagctctc agaattcctc cctgtatacc | 480 |
| agttctgctct ggtgtgactg tgcctgtctg ttcggtcttg gcaagtgtctg agctggtctg | 540 |
| gaagagagcg gtctcggacgcc atctactctc taaacattt taaactcact ccatcttttg | 600 |
| tttgcttctg ctcaactgtc cgctgctctcg gacgaagtttt gttgcttttc ttcattgatt | 660 |
| ttttttctgt cctgccgtgc gattagactc cggtgctac ataatagttg gaaaaagcct | 720 |
-continued

agttttcgcc tttgcttcta aagatgggtg gctgcaaaaas tctocctctt cttgccacacg 780
cctgcctct cctcgccctag agaaggttat tttctcttca tattgggttg tggattttgtc 940
tgaagggctcg ttattttgtct tttttttgct agaaggtgga aaaaagacca gctcctgyggg 900
tttttgctc ggagagggcc ggattggggg gtgggttttg ggattttttg gttgtccttg 960
cctgtggtgtg ctttgctttta gacaaattct ctattttgtg gagaagccag gccttacaaa 1020
cctacacgtg ttattggaaa gaagagagtg ggattttggag gggatatagtt cttgcttccag 1080
tagagggagaa taaagagggc tagttagtct tcctgagcag tgcttttttg gagaanatata 1140
atagctttgy ctaaggggaa gaatctgaaa attttatgta acataattaa tttttatttt 1200
tttttttctg atttttccaga aagctttttta attttatagt ttctctccag aagtattgac 1260
ccttttttttt gtacacgtggt tttatatttct tttatatgtc gaaaaaaata aaaaaaacc 1320
tngggggggg cccgggaccc naattggccct tat 1353

<210> SEQ ID NO 66
<211> LENGTH: 1011
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KSD: misc_feature
<222> LOCATION: (951)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KSD: misc_feature
<222> LOCATION: (952)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 66
cggagaagag cacaccttca cacatatcag aagttgatgc tttgaacttc 60
aagtacacaat gttgttgtg tttttttt ggtatatcaag agttgatgc 120
tggagccagaa ccccttttccat ccttcttctag aagttgatgc actttgttga 180
gttttttttcccccaaaattgaaa gatattttt ccagaaaaa ggtttttttttttt 240

<210> SEQ ID NO 67
<211> LENGTH: 1193
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (512)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature
LOCATION: (1167)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature
LOCATION: (1169)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature
LOCATION: (1171)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature
LOCATION: (1185)
OTHER INFORMATION: n equals a,t,g, or c

SEQUENCE: 67

```
ggccccgggc tggcactgc ggggcatcgc ctggcgcggc gcgccgccgc ccgccgggcc 60
gttcagccgc ggtccagggc cagactgtgc tcagccggcc ggtgcacgcag gacccagcag 120
ttgcccaacgg ggccgaggg ggcttcggaa ggacctgaa aagacagggc agaagagggc 180
gggccagcc tggcagcggg gcaagccctg tggcagcagc ggcgcgtgcag 240
ggccccact gcgcgtgagc tacctctgcga gcggccgccaa aagacagccag acctggaggt 300
ttcagcagac ggcagccgcag tgtgccgtgcg tcgcaagtgtc aagagctgcgg aagggctccg 360
taggcacct ttctccacgct gcctggcacc cggagggagt gcggaggggct gcggagagggc 420
tgcgggctca gcagcgggaa gcctgagtc ggcctgtagct ggcggggcct gcctggggc 480
cgcgggggct gcgagcgcgag gctgcgtgag cctgctcttc gttggctgca 540
ggcgggggc gcggccgctc ccaagcgaggt ggcctgccag agccacacag ctgcctgctc 600
tcgggggctt ggcgccggtg ctacccgcac gcggagggct gcggagagggc ttctagcttc 660
ttcgctctgct gggcgggggc tcacccgcac cagagctgag caggccccctcccc tcctccctc 720
cagccccctc tgcgggagtg cgagctgagc agtgagcgcag gttcagctgc ctccgagcgc 780
gagcggctgct ggtctgctag cccccgacgc ccctgccctg cggagcagct ccagccgcag 840
gatctgctgt ggtgcagcgc ccctccggag cattttagc gttggggcgc 900
tgctggccct gccacccggc cccctcgctg cccctcgcct ccagcgtgctgc tcgtgccattg 960
tgcggcgcgt cttggcagcgc ctcgctggcctt gcggccctgc ctggtctgctg 1020
tccggcagc ccaacccagc ccctggcggg cccctccggc gcggctgctg gcggctgctg 1080
tagctctgtac gctacgtctgc ccacacccggc cccctccggc gcggctgctg gcggctgctg 1140
cgagcgcagc ccggccccctcgccagcc cggagggcccctggctcagc ctccgctccc 1193
```

SEQ ID NO 68
LENGTH: 560
ORGANISM: Homo sapiens

SEQUENCE: 68

```
gaattgcgca cgaagttgca ccgattccaa aatggatatc tcagatgcga ttcgtcgcag 60
aattgtgca cgaattcagc cctggcagac cgtcagctcg gtctgcggct gacccagcag 120
ttttccct ccctctctct ccctctctct ccctctctct ccctctctct ccctctctct 180
cattttttttttt tcctctctct ccctctctct ccctctctct ccctctctct ccctctctct 240
aagtactac aagccgctgc gcgcctctct gcttcgtgcag acgttgtgct 300
```
-continued

ggatctgcc caactctactc aaggctcttg tgcctctgtgc ttggcattaa ttctctaggc 360
taagcacac accttttcct tatactactc tttctgtgtgc taagcacagc ttccaaagtg 420
cctctttat aatttttcta tttatgcca aatgttaac tgaatagggc atttgttaggt 480
cataatttct tctgatttt ttttttcaca catttaaaata tcagtaggtg gaaagagggc 540
cgcgacaaaa aaagaaaaaa 560

<210> SEQ ID NO 69
<211> LENGTH: 1657
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6)
<223> OTHER INFORMATION: n equals a,t,g, or c
<222> LOCATION: (343)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 69
cgcaagcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 60
ggagagagag cgccgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 120
gagagagag cgccgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 180
gctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg 240
tatgtgtgtc tctgagagg gcagagagcg tctgagcgac aaccctggga actgaggtca 300
cctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg 360
tgtgagagc gaaacatctcg tctgtagggc ttttggaat cttgagcgc 420
cagacactgc cctcagctgc tttagggac ttgagagagc ctctgtagggc 480
ggagagagag cgccgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 540
gctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg cgctgctgctg 600
tatgtgtgtc tctgagagg gcagagagcg tctgagcgac aaccctggga actgaggtca 660
cagacactgc tggcgcggcc ctctgtgatt tattgtggcc attggtttta aaaccaagct 720
gggtctgttg acctgtctgg tttcgctgtg ttcacatcag gattatrrca cacoctctcg 780
gacaattcgc tctcagggc cattcagcag tttcagcttc ttgacatcccc 840
gtgatgttt ggcctgcttg gcctgttggt gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 900
ttagagagc aaggagtgtt aaagtcgtca gatctcactc tttctgtgtgc agaagcgttg 960
tcgtgtagc tgggattcagt gaaataaagc actatatttg ttatcctcctt 1020
ccttcccttc ctctgatgaa ggcacagcct tttcgagaca ttttttctgc gacacagtg 1080
gagaagacact gcgctgcttt cttctgcttt cttctgcttt ctctgcttt cttctgcttt 1140
atgaagactg cgtgctgctgc gttgtctttca cttccacag cggggtcttt gcataacactg 1200
cgctctactg tctgctgtga tgcacagcag ccccataaa tcttcgcgc gcacgccggtt 1260
tttctggtgc aaggagtgtt aaagtcgtca gatctcactc tttctgtgtgc agaagcgttg 1320
gtacagcgc gcacgagcct tcgctcagcag aggcggcgt gcgtgtgtgt tcagctgttg 1380
ttcctgctgc ctgggtgctgg tggggtgctt cttgggagct gtcctctttg gggcggggcg 1440
gagaagacact gcgctgcttt cttctgcttt cttctgcttt cttctgcttt cttctgcttt 1500
gttggaaca gatcaaatc tcatttttag ctgttatatta taccttaggt gtggttggaa 1560
agtttttgt ttttggaaac atcataaat aataatgcgg gttggtgtag aaaaaaaa 1620
aaaaaactc ggggggggcc cccgttacca aacctgcc 1657

<210> SEQ ID NO: 70
<211> LENGTH: 711
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

ggccgagcg aagacocgtg tcggacoccttg ccoccgattc ccagactcaggt agatcgtgg 60
cataccctct acgtcgacgc ccggggtcga tgtggccctg cggagagaga tcagagatcc 120
cocaggaggt ggggtcrtc gcgtggggag attgacactc cccctcctcc attgctaacct 180
tggaaaaatt cagagacaag gggtgagagg atattaagtga gctggggaag gggccttcct 240
tctttacgg cacagactcc cttttggggca tgtcgggggg ttcctctttt ctcctgtttg 300
tgctcagttg tctacgctcc ccccttcctcg tgggaggag cttgccccgga cggacactc 360
tcagccctcc cttggtgagcc gatcgggagc gcagctctca gcagaacagg cagcagctgg 420
ggcaagcag octcagcgac acacgggggc ctggcgcgctg cgagtgccgg ccocctctct 480
ggrctgctgg cagcctcctc cagcgctctcg tgaagggaga gcagcagcct cttgggttag 540
tccatagct tcggctctgg ccagagaccc tgggctctcg gttgacctag 600
tggctgttaa gttggacagt cctccagcc tccttgggca tgttttgtgac gcggacggtg 660
gcctttgct taataatgt gctttatattt aaananaaaaa anananaanc t 711

<210> SEQ ID NO: 71
<211> LENGTH: 935
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<221> LOCATION: (510)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 71

ggcacaggg ggaaagccgt taaaacccccca tcggagaaaq gaaagacagt gttgctccca 60
taagtttatr gcacacattc tcggagaggg ccagactcaggt agatcgtgg 120
gttaggggaa ccacagagca agatcttttt tcggagaggg cggagcgcat cggctacggag 180
cgcgggcgcc ccggggtcga tgtggccctg ccagacagct cccctcctcc attgctaacct 240
cagctacgtc ccagctcagct ccagagaccc tgggctctcg gttgacctag 300
gacagttcgc ccagactcaggt tggctctcc cttgggctag aatacttaga cggagacgac 360
ccttctata ccaggtcagtt cagcctccttc ctggctagcc cagctcctcg cccctcctcc 420
gctttgaggg tcggagaggg ccagagaccc tgggctctcg gttgacctag 480
tgctttgac ggagagctcg ccagcctcct cggagaggg cggagcgcat cggctacggag 540
cgcgggcgcc ccggggtcga tgtggccctg ccagacagct cccctcctcc attgctaacct 600
cagctacgtc ccagctcagct ccaagagaccc aaatacagtt ctagagaggg 660
cgcgggcgcc ccggggtcga tgtggccctg ccagacagct cccctcctcc attgctaacct 720
gacagttcgc ccagactcaggt tggctctcc cttgggctag aatacttaga cggagacgac 780
cagccgacga cgtctgatttg agtgcaattg gcacatgaag ttatttataca cctggttttct 840
tttcatgaag ttccttagact aagtagaatt gttctttaaa tatttggtca aagccattaa 900
tatattcatt ttgaaaaaa aaaaaaaaaa aaact 935

<210> SEQ ID NO: 72
<211> LENGTH: 594
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEN: misc_feature
<222> LOCATION: (504)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 72

gcggggccga ggggygygggg accgcggggc gcaccggggg ggaattttgg cgtctgaccc 60
gggtggtgct ttcggctgctc gttgggagc cgctttaaag aaggggtgtgc gcacggactg 120
tcgagagatt tcggcttgctg gtggcggtca cattgaccc ggtttagac 180
ggttttcggc gattggagag ccccgcggct cgtctgactg gcgcctttct ggaattctgt 240
ggctcggcag ttcgggacct ggcgtgctag catttacctt gttgtacgtg 300
atctctgatc ttcggggctg acagtttaga aatttagcctt ccctctcgtc gattgagact 360
gggaatgtag actaaccgcc actctgtcag aatgtacagtg gaaatttttaa aggattactg 420
aattattcatt gcacatataa gtataacattt aagttggttc ggtcagacag garagaaaaa 480
aaaaaaaatt gggggggggc ccom 504

<210> SEQ ID NO: 73
<211> LENGTH: 620
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73

gattttgctg cggagggaggg gcggggccgg ggattttagtt gggaaactct tcggtaattttc 60
wtttttaacta tcacacccgatt gcaagttggc ofgacgtcttg gtagaatctcctc 120
ttggctggc gtttttaggt ggcacgagat gtgggaggtc agaaagactg gcggggaggg 180
aatattcctc ggtctgtcatt tacacgcctt gtagatgggtg tttacgcagcattttggg 240
atccacccct ttctagaaa atcttggtagg cggagcagag gcgtgacagcct ttgggtaagg 300
acatacacta acogacttagg aatccaccct cttttagaacc acogacttagg gacccattcga 360
agcctttgac gttgagaaaaa ctttctcaag gccgacagag tcggatctttg gctgcttgcc 420
ttgctgctgctg cggagggactc atttacaccc tcctctagttg tgcgttgcag cggggtgag 480
tgtattttacctattctgtg agaaggaagctactgattc ctttggagtgg agctcggcc 540
gagaaatttg gagaatttttaa aatattcattt gttgaaaattttaa aaaaaaa caaaaattct 600
gggggggggc ggtacccaat 620

<210> SEQ ID NO: 74
<211> LENGTH: 581
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74

daagagttcact tgtaaaaagttt tgggtaggctt actcttactt aaatttttttaa caaaaagac 60
-continued

tgagtctga ttcatctttac aagcagcccgct gagtcgaact tttctgc cagcagct cgcacatct
300
ttctatctc ctccgcttac cctccacgct ccacagcctt ccagccactg tttccagtct
400
tgtccactg ctcctaa cctctccgc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
500
tttctgtgct ccacagcagc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
600
ttgtctgct ccacagcagc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
700
tgtctgct ccacagcagc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
800
tgtctgct ccacagcagc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
900
tgtctgct ccacagcagc cttgagcgtg tctttctcct ccagcagcgc ctgctgctg
1000
<210> SEQ ID NO 75
<211> LENGTH: 1843
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (10)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (24)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (91)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (213)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1633)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 75

aaacccacac gcttcgtcct cccaaagaaa cagcagcccc aatacctaaag cgcgtcagta 60
ggcccggagr agacgcttctc gacaagcccc gacccgaggg cgcggggagg gacccgaggg 120
gagagacctc tggggggctcg aaggggaggg cttctcgagcg gcgaacagcc gcagcccagc 180
gcgcagcg cagcagcatt tattctcgcc gtcgcccag ccagcagtc ctgcagctca 240
gacggagcag cacagtgctt actccacatgct gttgtttttg aagagcgagc 300
tttcttcac aacgggtcgct cttttctcgct ctgctgctgctt ctgctgctgct 360
cgcacagcc gtcgctcgc tgggtcttttt tgggtcttcgt ccagcagcgc ccagcagcgc 420
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 480
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 540
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 600
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 660
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 720
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 780
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 840
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 900
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 960
tttcttcac ccacagcgc cttttctctcc cttttcttttt ccagcagcgc ccagcagcgc 1020
coctctcct ttcagcctaa tactggacct ccgagtgtgc ttccagggggc gtctactctc 1080
cagcgtcgcc tgtgctgctg gcccaggggg ccggctaggc tggctagtgtg tattgtgtac 1140
tatgtacta gcactagcgg tgttattgttta aacaactgcc atctgtgtac ccacgcccac 1200
ccgccgct cgcggtcgct tcctgggaga ctcctgagga gcgggggttg ctgatactgt 1260
atatatttac atgtactata tactggtcttg tgtacaagtg tgtgtaaaattatatcctt 1320
gtggtgaag acgccttttttttttttttgc ttccacccgcacccccccgc cgcgactct 1380
aagggcact ctgcccagcct tactgactttt ctgtactttt tttttttttttaac ccocaattt 1440
cocctctct ctcttcgcccc cgcacactcc acocagggttg tccagagccct tgcgctgca 1500
tggccgcggc cgctcagggcg ggcttggggaa ggcgcaggtg ccagcggcagc gagtcctcag 1560
tactgagg gcgcttctct tatgtcttact gcctgctcttg gcctcgagac ctctcccttt 1620
gtgagctgttct ctcgcttgggg acaacccccc cgcctccaggg tgcacaaggg gtctctttgg 1680
gcgcocctt ccgtccggtc cgcgctgtgt gcctcgtggg agtcgctttg acctccctgtt 1740
gcctcaactt cctctctgct acaacggggga ctcgctctgc ccctccctac ctccagggca 1800
tggtgctgaga ataactaggg ttcgctgtaa aaaaagaaaaa att 1843

<210> SEQ ID NO: 76
<211> LENGTH: 1441
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1056)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1081)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1109)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1328)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1362)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1419)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 76
tgcacacacg ctcggcagtc cccagccctg gcacacatctg gigaacttgg gtctgtcrocg 60
ggtggagacgc gccctgtgcgg cccagccacgc ggagttccgga ccagcactgc gcctgagctc 120
acacgccttc atgaacggggt ttttccacctt cgcacagcgc acocggacgg ggctttggctt 180
gactgatgg cctgctgtagc gtctagggcg gttctgggcga agtctgcaga gtaacgacaa 240
cctggcctg ctcgctgtcag gcatggggtc gcagagtagtg gttcgcggaga atacgacaca 300
cctggtgct tctgctgcat gcctgctgtg gactaggggt gtaagccgag gtaacgacaa 360
cctagagct gcacagcagct tggggggcag gaggactctg gcacagcagcg 420
ttcctggcgcctct ctcgagccag gcacagcctg ccacacgacct ggtctcagcc agtcgacttc 480
tctggcagc tctgctgcat gcctgctgtg agtctgctgc ccacacgctg 540
gaacgctgc acracttgct gcacagacag gttctgcttg ggtctgcttg ttcctgcttg 600
ggactggcgt tctctgtcct gcctgctgct gcctgctgtg ggtctgcttg gcacacatctg 660
atgggacaag ggctgaaaggg gcgccgctaca ctctgtcatct tctctcttag ccctcgtcaco 720
tocaaacaggt cccaattggac agctggtgtct aaggcgc ggttaataaaa tggttacctc 780
ttctctctca ccaatttttc acctgaccag ttgaacaaca actacattaggt cttaaatggat 840
gacagctctac atgaaccctcat tattacccctactgatgtcat tctgctcaagt 900
dactacgtgacacactacagg gaagttcaca gagctgtgaaacc cccaaattacg agctcgcac 960
cgcgcgggtgc ggcacagtcg gaccaacaccag agcctctcgt ttagggggct gtcctctgtga 1020
gcaacctggcctgttgcccga gaggcaaatgg ggcctcggctc gcacgaccctc ggtgtgtgg 1080
nagargagg acctgiggsgc acaacgcacnt gctatattc ctccctcctctt ctatttocca 1140
aatctacttcc acaagggctg acgctagcgcc tagtacgtct gctagctgtgct gaggaagagt 1200
tgtcctggtgc gacacgctaca cccaccacccc gcgacagggg cggcagagac gtgacgac 1260
qacgacaccc acacactcctc atcattggtctt ctgctgtgctt ccccacacgcc gggagagcc 1320
aacacattg aacctggtgc cccaccacccct gtcgtgcgc acggcctgaca ttcagcagaa 1380
ttgagacaa gtatccgtgct cccacaccgc cccgtgcccct ttcagcagac 1440
g 1441

<210> SEQ ID NO 77
<211> LENGTH: 910
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78
ctgagcaag acgaagagcg gggcagtgatg ggggctgttg aatgggaagt gaagctaggc 60
aacgtgggga tggcgaagga gacagcaac gacggcggca gggagacttc aggggacgcg 120
gggcggcag cgaaggggag gggtgctggg tgggctgttg tggcggtggc gaaaaaagggc 180
GGGGAAGAAC ggtggaagcgc agcggtggtg agcggtggtg agcggtggtg 240
tgggagac gccggcggc ggtgctgggg ggggggggag gggaggggga ggggggggga 300
gcggggggc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 360
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 420
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 480
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 540
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 600
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 660
gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 720
agggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 780
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 840
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 900
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 960
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1020
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1080
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1140
tgagccag gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1200

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1260

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1320

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1380

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1440

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1500

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1560

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1620

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1680

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1740

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1800

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1860

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1920

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 1980

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 2040

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 2100

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 2160

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 2220

gggtgaggg aggacgcaac aagggcggc gcggggtgcgc gcggggtgcgc gcggggtgcgc 2280
-continued

atctcatatatcaactcttaggcttgctgctgctgggagaagtaataattttctagaagtttaaatgttagagaatcagaagtt1530
atctgcttgcttgctggctggcttgctggctggcttgctggcttgctggcttgctggcttgctggcttgctggcttg1600
gctggctgctggcttgctggctggctggcttgctggcttgctggcttgctggcttgctggcttgctggctggct1680
gctggctgctggcttgctggctggctggcttgctggcttgctggcttgctggcttgctggcttgctggctggct1760
gctggctgctggcttgctggctggctggcttgctggcttgctggcttgctggcttgctggcttgctggctggct1840
gctggctgctggcttgctggctggctggcttgctggcttgctggcttgctggcttgctggcttgctggctggct

<210> SEQ ID NO: 79
<211> LENGTH: 1487
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (78)
<223> OTHER INFORMATION: n equals a,t,g, or c
<224> NAME/KEY: misc_feature
<222> LOCATION: (948)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 79
cctgcctgtggctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt60
cctgcctgtggctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt120
gatatgatggctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt180
tgttgggtgaagcctgcttcattcactgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt240
ggacctgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt300
cctgcctgtggctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt360
ggacctgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt420
ggacctgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt480
ggacctgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt540
ggacctgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt600
tgttgggtgaagcctgcttcattcactgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatcagaagtt660
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc720
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc780
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc840
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc900
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc960
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc1020
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc1080
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc1140
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc1200
taatctgtggcagcactgagctacatctgcaccctccctctctgctgctgctgctgggagaagtaatctttttctagaagtttaaatgttagagaatc1260
<210> SEQ ID NO: 80
<211> LENGTH: 1563
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<210> FEATURE:
<211> NAME/KEY: misc.feature
<222> LOCATION: (14)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 80

aattcgcgac gacncagaaaa cttgcgagaaaa atgtagcgcg cgctgctggct ggcgtagctgg
60
tgtctgtgc gcgccttgg gctacgttgg tcgttggag gtagttgcgt tctgcgcttgc
120
gttttgtgg cagatgttttc atggagggca tgcagaggt tagcatttttc tagcaacctgg
180
tttgcgacgc cttgcgcatc tctgctgacgc ttcacgtgaa cctgataattgc
240
aggatgcg tctgcagagac acncaacttt gcaccaccacaa agcgtgatagc aggcgtatatt
300
tttgcagattttgtcaggttacaa atggcctcag ccaatggtttt cttggagagt
360
gatcagcgg acagcagcga ccagcctgtgc atcggctgctg gcgcgggtg
420
tttttatgtttttgtacctatataa atgaaaaattt acacgaacota gaaatatatt
540

tttgtccta tctttttgtt acctatataa atgaaaaattt acacgaacota gaaatatatt
600
tttaggtgtgt tgtgctctct gctagcata tctgcgcttgc
660
cgtgaaatgg cagatagtgc ctgatatttct aagggatttc gcgtattaca acaacacatt
720
ttttataaat gttatatcctct cttgctattttgc tttgtcgttggtg
780
acctcgctttt cctcccattttt tggagatttttt tcagcggaga acaacagaga
840
gttcataata ttagaggtt ttaacccc tctgtactt cttgctgctg
900
tttttttttt gaaaatgaca agacatcaaca gcacattttt taaacacaat
960
ttttttttttt cctgctctct ctctctctttttttt gttttttttttttttt tatttttttttt
1020
gttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt...
<211> LENGTH: 1020
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (20)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 81

tgcaagtctg ccatgctggg gtgggcccac tgcgacccccc gcgcgcggcg gcggccggcg 60
tggcccccgc tggggcctgtg gcgcgctctg cgccgcccgc agacattcgg cggtcttggtg 120
tgaccgccgc tgcgggacgct gtaacgctgag ccgccacgag gcgcgtggtgc gcgcgcgctct 190
ggggtcgcc gcgcgcgccag ctgcgtgcgcc gccgcgcgcgc gcgcgcgcgc gcgcgcgcgc 240
cgggggcgg ctcctcgctac gtcgtgccgc ccgccgcgcgc gcgcgcgcgc gcgcgcgcgc 300
cgcgccatgc aacctctctg gtggagagcg ttttgctaccg ctctctctct ctctctctct 360
cgcgactctc ctggcatttgc gcgcgaggag tttaaatgag ctgacaggctt ggagaggcgcgg 420
aacccgacgc taacctgacaa tgacagcggcc tgggggacgc ggagaggcgc ggagaggcgc 480
gagcaggag ttcgtggcaca tgcacagcggc agccgcagag agagagagag cgcacccctc 540
gatgtagaag ccgcggagac gtggcgggaaac cccacgccgc ccgccgcgag ccccgccgctg 600
cgcctgacgc ctgacgccag ctgacagcggc gcgcgacggag ggagagagag cgcacccctc 660
ggcgcgacgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc 720
gccgcgacgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc 780
cattacatttt tggcgcggag gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 840
tgcagcagc ggcctgttgag gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc gcgcgacgcgc 900
tccgcaaag ctgcgcgtct gcgcgcgag gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 960
aaagagacag aaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa
cttgggggtga ttttaggtgc tcocctttctca ttttatgtg aagcatacta ttccacaca 660
gacttggtga agattaaaaa ggatatttttc ttttggaas aaasaaaaaa aaasacacyga 720
gggggggcct gctwccacac aggcyatatc aatctecnit ttacacaccc 770

<210> SEQ ID NO 83
<211> TYPE: DNA
<212> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (322)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (365)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (379)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (390)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 83

gacttggcga cgacctgagc tttaaccocat agacctcaact gcoctttcota tccaaaaagt 60
acacactcga tcattttttcct tctttttctct ttttacacaat tmaacaaatc agtgygtctc 120
tccactacc gcggagctgga ttctgtatggag tcctttcatgga cccacoaccc 180
acagatcttggctgctact cctggctgcat ccaacggcag caaggctgta cctgggaacaa 240
cacggacactt aggctgaygg ccaactgtct cctggctgccc ggaggagamm cggggggtct 300
ttcactctcc caaggactcag cngggggggc eggtacccaa attccogccccc tatagtgat 360
cgctggattc aataaaactt gggcggtcnn tttttaaann cgtctcgttgg aataagaaaa 420
aaacccttgcc cttgttttcc caaacatctt ccgcttttgc aagacactc ccocccotttt 480
c 481

<210> SEQ ID NO 84
<211> TYPE: DNA
<212> ORGANISM: Homo sapiens

<400> SEQUENCE: 84

gtcgggttag acagacagag gagaacagct cocctttcctctt ttttcacag tttgttattag 60
gtttatataa cttttagttgtg gtaaagccgc cagttaconc antcgctatt gcocaaaotct 120
cccttttttct gcgcagctggc agacctttag agacacgaga tttctagcag gcatttattt 180
catattctgg caaacttggta taatttttatttaaatcaactgct actactgggc agtagattcota 240
atttttttt gagacatatat tattatatct atttttactagcttacgacgatcttaaaa 300
gaatattttaa gaagatattct cttgaggactt ctttttgtgaggcagacacag agcttttcca 360
gacacactt taatctttcct caactagaggg aagctgttcg tggctttctct cttttgagaa 420
tagctgcttt gcctttttatc ttaattttgta aggctggaat aagacttattt cttccaaattc 480
cctattatt attatatattt tccatttttt actaaagggct aagttatatta aagttgtta 540
atgtgaccc tacattttt cttgctatcc tccatttttctt taaatatttctt aataggactt 600
taattgaan aaaaaaan aaaaaaan aaaaaaan aaaaaaan aaaaaa 644
The provided text is a sequence of nucleotides, specifically a DNA sequence, as indicated by the sequence and gene model annotation tags. Here is the sequence in a readable format:

```
<210> SEQ ID NO: 85
<211> LENGTH: 1351
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (133)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1305)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (1344)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 85

ggcaacgagt gcgaacgagg gggctctctc ctggtcaagt gggcgcggt gggcggtgt 60
ggctctgtgg cagcgccggg gcggagacac cggcaacttg agggcttaca gcagcgcgga 120
gcgcggccgc cccctcccgc tggggtcagg agcttctgct aaaaactgaga aggacaata 180
tatatccca tttcatgata tttcaattta tgcgataag gatgtgtttc acatggtagt 240
tgaagaac ccgtgtctta tttgaaamas ggtgttagct acaagggacc tttaaancc 300
tattaaacac gatgggaaca aaggaanacot tcgcatatgg ggaatattgt ttocgtaata 360
agggataa ccgtgcaacttg gtccatccc tccacattgg gaagaccacg gcaggaccag 420
taatcactg ggtctgtgtgg tgtgctcattg cccacatttg gttgctgana ttggaagcna 480
gtattgctga aaggttagaa tattcggctg gaaatcctta gcatctattg ctatgattga 540
goagggaga aacgtccttg aaccttcttg ctgctgactg atgcaagcna 600
tttaatgat atcagactag tcaacggcgt gaacagcttg tcattttaga ctactgtgga 660
tggctttga gatgtagttt ttcctcgttgg aaaaaacagaa atcagtttgg ctgtttattg 720
agaatattaa gtaaagacct ttcacattga tattatttta aacactcatc agcatttgc 780
aggttattg actaagaaag cgaatgtaaa agggatcatg tcaagtaaat caacttttgtc 840
tgagacccc tttaatgtctg atgcagggc cttggtgattg ctttaaacoc 900
accccttga tggctctggca cagcaccacg agcgtatgtt gattgtttccc atccacagaa 960
aaactcattg gtttttttgg gaatatacctg tgcatttctg acattctggtt cctattgtg 1020
tatattgat aaaaattcta gtttccccaa gtttaaaatt tgcattttct gtttaactaa 1080
agttaacctg gttgtgacta atccactata tttgagattt tggcactata aacgactattt 1140
catactctaa tcaagactac ttttttgacg tgcattataata ctaacgagt tgcattttgg 1200
aagttcattg tgaagggata gcaacatttg gatgtatattg ttcaacattg 1260
tgtgaana aatttatttt gttgaanana aaaaaaaan aacncggcgg gggcgcggcgg 1320
tccacttatt gcagctttcgg ggaggttttt a 1351
```

The sequence appears to be a DNA sequence with annotations indicating its features and characteristics, such as the presence of certain nucleotide sequences and their locations within the DNA strand.
ggtggtgagg accagaggtg gtttcggtgg tggctaaccg gttgtcctgg 240
agccggaag actacttgtga gcagtgccct gaggagctac ctgttttgct acgattaccc 300
atgctcact ctggagttgg aaactattgg tcaggtcttc aataaaatc ttgcttttag 360
tcctgacag acagagagaa agttgcaggg atctcgcaaa ggtgctaatc tgttccgaga 420
tcgtgacct gttggtcata caagttctcc atcagctac actcaggact gaaaaactgg 480
aagggcata ccaggaaggt ccaagtttcc atgtttggcag gattttgcaag atttttgcct 540
gcgtttgct cagacagggg aagcagaaag actataacaas aagaccgggg cagaggaat 600
taaagtttcc aagggagtgc attctgctta tgaanaagca gaggccctgt agttggcgtct 660
gaaacacg cttctctgacct ttaatgattc ttgctcagcgg atctcgacag tcctctagaga 720
agggatatg ctacctctgg agttgtccatt tggagattaa gaactattga tgcagaaaaa 780
taatccttc cctaccaaaa ccagttcggga acacactaca gaaactaca aactactaaat 840
ggtatcgag tcggtcggac ttgatgcgcga aagttcgccca aaocccatcga atggctttat 900
gagagagaag gtagctcttg cggttcggcca gtttcagatct ggtgctctgtag ggtgccct 960
taatctyctc gtaacattag ttctgacgc gcacatgact gataagagca ggtgtgcagg 1020
cgtacactga tttgctctga tgtatgagg ccccaagtct gcaacatatc gcgattaca 1080
gtttttttag ccacagagag aaggagcttg gcgcagcagc tggggaagca ccttcagagaa 1140
ccccccattt aatagattg ctgtggagcc aggaggattg cacatgctgc ccgacatttca 1200
agttctgtat ccagttttac ggaaggttg tttgttgat ctcgctttta ctcctctgta 1260
ggtaaaaacaaatattcag ataaagctcg ttagtctgct tttgcatttc aacataaggc 1320
cctcagtag acatggacag cccctgttaat gaacgctacc atcctagccacc ttcacagat 1380
ggtaccaggg cccccctgctc tctctctcgg gtttgcacag aaaacccgaag 1440
tgctctcttg acggtggtgca tgggacacga ggcggagttg ggatgccgaag 1500
ttctgagac gcagagtggcc cccctctccct actcctcatg atgtatgtgc gaaacagtga 1560
gttgctcgct tctgacgagc ccagctgtttg ctcgacagcc accttttaca tgcgtggacc 1620
agccctgtgt ggcctgttcc atccggaacc aggaggact cttttcgct caagctatgg 1680
tgcacgatag ctgagctggt cccccgggttt aactaattgg gaaataattt ccttctcagtt 1740
tgcagcttc acacaaagga agaggattat gcacatctacat gctctactaa accagagga 1800
c tgtggattt attctcagga caagcagact cagcttttct cagagagggc gaaaaaccacc 1860
tgacacgttt atcccttccc aggttgggac ggtgtgctgc gaaattatac aactcttgg 1920
gaaatccag ctcgctctg ctctctcttt ggttgtcttg gtcgasatgc ttcactaaaa 1980
cacaccattt cttccactt cgaactgcct tggctgcttg tatgatgtct gtggctacaa 2040
aggtgaaca acgtgagggt ttaatgtctt ttcacacttt atatatgat tattctctac 2100
aatccacattt aataatctct cttctatatt tattctctct ctggcgtcgc gattttttca 2160
agoctgtata ctcgctctgc acaacgactc aacatctata gattctaaaaa ttcctatttc 2220
aanaaaga aacatcaca accaaataac tggcttagc cttgcttcca gaaatctaca 2280
agocttttg agcagagatata ttttttcttt acatgtgaaa tagaaactcg ttcctttcct 2340
cttccacttc agatcattgt cttccacttt gttttctatct aagttatctt ttcgctctgt 2400
| aaaaaaaaaa | 2527 |
| gtaagctctg aatgaacctc tttactcaat aasattaatt ttttggcttc ttaaasaaaaa | 2520 |

Sequence 87:
```
cocaagaaatt cggcagagc gggcgcaagag tgggttttct gaaaccgta ggcggcaagc 60
cactctcaac gcccagaaa ctatctttcc ccctgtcaat cggaaagagc 120
cactctcga cagcagatcc caatgacagc acacttccac aacggtggcc 180
tactctccaa acctggagaa ttcgcaagaa ggcgcttatt ccaatactga cctgggttaa 240
gagggatttt agttctagct cccaaagcctc tggcccaacag cccttttccc ccaacccggc 300
ccttggcgcag aacgcgcccc taagacccag gcaccccocagt gacagagaa ggcgcgtgaa 360
gtaatgtct tcaatccaaag ggcgccagc tccctggga gtaaaggcaaa aaacgcggcc 420
ttcacaaag cgcagagagc actcgagaaa ttcagcaaat ggcggagcag cttaaaggttt 480
gcccttcttc ggctgggttc ttagaatgcg ggaggacggc gtttccaaac 540
aataagcgtgc gaaaaaaaact cacagatgg cacagatgct gctcagacca ccttcagag 600
caataataat cgagagagaat tggctgtcag gactctctcc gccaagtttt ctaagccc 660
ttctagcgtc aagcattgagg ggcgttgggg cccaaagctg ccaacccgaa aacggacaa 720
gactacccgc aaccccgaag caagacccatt gcccctttgc tttaatttgcc ttgcaacctc 780
accaaaaaacc aacgaaacc ccaatgtgtga accctgagaa ttcocacaan ttcctcttgg 840
aacaagactg acaagagcgc acagagcttta ctcacaactt tccctggccac caactcacc 900
tacaaactcc cggccagcgc caagcctgagc acatcagcct cccccacca cccccgctt 960
aagccatct cccaaacaaatt ttaacctccc gttttaccta aaaaaagctg ttcaggtaga 1020
cactacaat ggcttccagc actctgtgag tggcgaata aatgatggag aacagacag 1080
tgacggcag acaatcgagc actcgagaaa actccagaaag gacagagaa aacggaaaaa 1140
ggaagaaacg aagaggttgcag ctggggagaa aagaggcaag aagagagag aaaaagaag 1200
dagcaaatgta aagccgtata ttaacataac ggcccgtatt caggtcttcc atctcagaa 1260
agtctgtgg gtttgcggag gggaggaagc ttagcagag cccacgaagc ggcgcgataat 1320
tgacattcct acagccgca gaaaagggg tggcgagctg cagcagaggg 1380
ctctacagctc tatactaaaa caactggtttc agagattgtgct catgatcttt taaacagtga 1440
aagaactctg cttggcgcag cttcagaccc tatggaagat gccagaaaaa tatagatgag 1500
tggtctcgag cagagattga ttaagcgocca cagtcagact ggaaatgggg ggtatttccc 1560
tgccccacaag cattagacga tttatgaatt gattaagagc gaaataagtgc atgatgcgct 1620
cacacatcgc gtcacagaga aagcataaca gtcgtccttg ggtgttttga agatgatgtta 1680
gggagaagcag ccaagagaa aagatcagag agagagaccag aagctccttg aatccgacaa 1740
taaagagatt ctctctcccc ctgtctccac taacaaatgg goacacgaggc atgagatgta 1800
cgtgatggt gatacctctg atttccccgt ttctacaaca gagatgatgc aaggaactaa 1860
tgttggaaaa gcaagacag aagaaanga cottaagaag ctaaaaaca agraaaaaa 1920
araaagaac ttccagaaaa aatataataa tgaatggaas attagatcct tatattacac 1980
taaagtcac accctccatc cttcaaaaaa gggggaacc aagatctcnc agttaaaacc 2040
tggtcactct ctagaagttcta tcaaaaaacct aagtaaacca aaagttctctg gcggaaatga 2100
agaaagagaaa atgtgccctg ttctcaggg ttaaacaggg gcaaaggatgc ggaagagocaa 2160
tgatgatatt ggtcagactc gcctctctag caatgacact acctccctct ttgctcattc 2220
gctggtcttc ctattgatcag atgtgtagtc tggctttctta aatggagatcg attggttac 2280
agaaaaaaa atgcacaaa ccatactttta cttaatggta tgaataatca cctataatctta 2340
canagtttt canagtttta cggagaaac aatctctcgct titaattgcg aatctcagag 2400
actacattag ttgagatgtas gattatass aatattcttc tgcgttggg cctacatctt 2460
gcaagatgg aggtgctactc ttttgaacc ccaagaaaaa aatctctcttc aaaaaatc 2520
aattaaaaa aaaaaaata cttgagggg gggcaggtta ccacat 2566

<210> SEQ ID NO: 88
<211> LENGTH: 1640
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 88

gaatagcga caagaccttct cttctctctct tttctctctct ttgggctttt taggctgacc ccagggggag 60
atcagctcgg ggatgagtct ctacatggtt ccaggttgga ccagaggttac gctttctctct 120
gccatgtc tttatccttct gccctgagtt gccatccgtt ctacatggtt ccagaaagct 180
aaacatagta cttaaacaggt gcatagctgg cctctgacgt gctttcctctct gtaatttcttca 240
gaatgtgaa cgaggtggtg ccagatattt gcagtggagat tggctttctct ttctctctct 300
gaggaggg gcagatggagat tggctttctct ccagaaagct gcatagctgg cctctgacgt 360
acagctttgt cttctctctct ttctctctct gcttttctctct gtaatttcttca 420
gatcctgcg ccagaaagctt ccagaaagctt ccagaaagctt ccagaaagctt ccagaaagctt 480
ttcagcttt ccagaaagctt ccagaaagctt ccagaaagctt ccagaaagctt ccagaaagctt 540

<210> SEQ ID NO: 89
<211> LENGTH: 1643
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> LOCATION: (1836)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 89

tgaaac cacg cgtcagcgcc gccaactac ccaggtgg gctctgctgg cggaggtt 60
cogaaa cttcaagag ccaggtttg gctctgctgg cggaggtt 120
cacgcggg ccaggtttg gctctgctgg cggaggtt 180
cogacaag gctagagagc gctctgctgg cggaggtt 240
cgacgcggag gctagagagc gctctgctgg cggaggtt 300
cacgcggag gctagagagc gctctgctgg cggaggtt 360
cttggacct gcatttatgc ttsgcgggtg tattttagga ggcaatatact tgtacaAata 420
tttgcacct ccacacagat agctgttacta ctgtaggata aagtacatca aagatgttgt 490
catcttaaat gacgcctcg ccgatggcgc agctgctttc taacgacagc ttaagaagaa 540
tatatataac ttgagagaga agaagttgta attatactcg tggctgctgc ccgaagtttg 600
gatagctgtat cgtggccaca ttgctcata gtttacacag aaacctacag ccatcttaca 660
ttttagcct gtaaagtgct attgtgtcacc tctgacactG tocatctgtta tggcaacoag 720
aaacacactc gtagtatctt ttaacataca ggctgaacac atattgctct aggctaatct 780
gactagtaac ccatgctgct ttaactgtagc cattgaaac aagtatccacc ttgggttttt 840
tatttctgga cttgctcagt ccaaaagaca ttaacactg caaccycagag aacatctaaa 900
agtagttcag caagctgaag ccaocatattg ttggcataat cggctattgga aacaaacatt 960
tgcggtaga ctaacttatct ggcttacag acgtaagaaaa acaatttatg aggaaatatt 1020
ataccacagt atacattccag ctttccagtt ttggcagttg atttatttttt aagagctttc 1080
ttcatcatacg tggcagacag ggtgtaggtca tcttattcat tcattagctt tagnaataac 1140
attacactaa aattttttttttt ttggaggttg tggtgtcttt tattatttta ttagatagt 1200
tatggaatgt cacagtaag cctatgtgcac cttgaagttg aggaagaatt acaatttttt 1260
taaatctttttt atctgttgtt tttgattaatt cttattttttt caaagagagct gttatcttaag 1320
agatctagta atttaatttgc tagtttaaaa tggacaaaaag tggctattttt tgaagagtt 1380
gttacagct taccagagtt ttgtaggtagtt ctttacagac cttatatattt cttcatattt 1440
aasattctgca gattttttcc tcgtgataat ctttacaatta ttaagcaatgtt attcacctt 1500
ttagttatgt ttaaatcgcgg cattttatatg ctaaatatttt taggaattttt 1560
agatatattttt ctttgtacact ttttatactg tcgttggtg tcttctgtaga attatggatg 1620
acaatgtt gtaacatag tgcctttttc ttctcttagta taatanaaac aacagatagtt 1680
ttggttttta ccttattcag aagactcaca taanaagaga aacatttttt aaaaaataca 1740
tcttcatata ttaaactgag ctaaattccac atttttttta taatanaaaa aagtttatttt 1800	taatactaa aacagtagctt ttggtagcact attaangtaaa cttttacgcca caaaaaaa 1860
aan

<210> SEQ ID NO 90
<211> LENGTH: 2478
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 90
ggacacagagg cacaggtttta cttgagccgg cggggtcgag cggccacagc cattctgtag 60
tgctgggtcg acagagaaag atctataatt tatgagctttc gcaactttc tttgtggctt 120
tgctctcgtc cttgacagaa cccgcagaa aagagggccc cttctaatgc 180
ttagccctac cttagagcag aaggcttacc ttagtttattt aacaagtatt 240
tgctcttctg tgggttagag aagagggagc cttgtgacag cttctaatgc 300
ggaaaggttc ggaagactttagtaaagttttt cgggacggtt ggttgctgtg 360
ggtagatgg aagacagttca ttaaattttgt cacaacagcg cggattttaag aggacttata 420
ggacagctgg acggggtcct gctccttcggc gagaatgttaa 480
-continued

aatcg caccc tacggctacg ttttagata gtcgatgtcct taaactat gatggatgat ttaactatgaa  540
acagatgcag ggattagacg agccgagggt tcacagatttc taacttgccg tacagagatg cagagactct  600
tgcaccagcag caggggttag gctgtaaggc gtagatgct ctaatggagg  660
tatagctacg caccaacacc cggagagatg agatgacgat agttagtgct  720
agagagt atctagcactat ttgtagca tcgaatgcgtc ctagtactc aactatggcgt  780
aacagacag tagagagcaag tcgtagtgac acatggaatg aggagactct  840
cagcagag cacggagacttt cggacaccc cttagcagcag agcagagactt  900
cctgctggctgtcactattc acatgtagcgtc ctagaacttg gtagagctgt  960
cgctgacacag cagcagag ccagcgaggtc gttttttggtg ggtatagct 1020
agtagctgtaggtcctgtgctctgtagg aggctgtcactttttaaccttagg 1080
tactagcttc ggcttactgtc gaattgtggtg cgcagttgac gcagtctgtctg 1140
taagctgtc cagagtgcgtc cctggacgcc ctaacacctt ctggttagctg 1200
agtactgccag ccagacaccc cgcaggtcctg ccctggcgct gcagactctct 1260
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 1320
agatgcagtg acaactatgc gatccgtgtg ctaaactgtc aacatggtcttct 1380
cctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 1440
agatgcagtg acaactatgc gatccgtgtg ctaaactgtc aacatggtcttct 1500
attttagag gcagccacctg ctcctgagctg cccttgactctgc acatggtactg 1560
ggctgcaacta gctctgata ggctttccttg cttcctgtc gcagactctct 1620
aacagatgg cctggagcc acctgtggtc tggctctctct cggacctgggct 1680
cgtagctgtaggtcctgtgctctgtagg aggctgtcactttttaaccttagg 1740
caagcgctc cagagagcctg ctttgttgcct cagagagcctg ctttgttgcct 1800
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 1860
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 1920
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 1980
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2040
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2100
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2160
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2220
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2280
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2340
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2400
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2460
ccctgtatc accagagggc tattcatgtg cggcttacttt gtaggttgcttg 2478

<210> SEQ ID NO 91
<211> LENGTH 2058
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
-continued

<222> LOCATION: (69)
<223> OTHER INFORMATION: n equals a, t, g, or c
<221> NAMS/KET: misc_feature
<223> LOCATION: (161)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 91

toggcttgc ttattttggc ttctcttggt gcccagacgt tttcatcaacc aagctcttcg tttcgttgtg 60
atggtctggtc ttctcccgatt ttgcgtatgc ttctcgcttgt tcctcagag tggtaaaac 120
gaagttggact cggcccagagt aagtttgg aatggcgtag gatttcttgtg tttcatcaacc 180
aagctcttcg tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<210> SEQ ID NO: 92
<211> LENGTH: 1411
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1391)
<223> OTHER INFORMATION: n equals a,t,g, or c
<222> LOCATION: (1403)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 92

ggcagccggc cagccgaggc cagggagggc cggagggagc agcggagggc gtcggcaggg 60
gaccgagggc cagcatcggc cagccccctg tttcggaggc tttcagatct atcgatccaga 120
cagcagcag ccgagcgaag agaagtcacc atctcttgaa gatctgcaag cgggggttttc 180
gacgctcag ccatcacttg atccttgctgc aactctcata tcgagctcatc aaaactgtct 240
gagaataat tcatgcaagt ttctgtccaa agaagagtag atacactctctt gagaagatgg 300
gttggctgg ggccttgcgct cctctgcttg acaagcagctt taatgctgca aagagactct 360
gcataagtg aaaaaatggtg agggcccaaa ttcgttggag tggtttttttc tggcagctttt 420
gtggagcatt ccatctcctcc caacaaaaat ccctttggtg gagaacatac tttttttttct 480
gagctctctg tcgctggcgtt cttcctctcg ccctgacag tagcaagctg gtgtggcgctg 540
tggcctctctg tggcagctgg accgagctga aactctcagtg ttggtgtttt ttgggtgttg 600
tggatgttttt cctgtctcat atgtgcgtcc aacgctttct tcgctgatag cagcagcttc 660
aacgagcaac gcctgctctgatt tctctctctt cttcatcttt cttctgctgt cttgctgtct 720
attcctctcc tctctcctca gtacatcagg aatgggaat taaaaccagc tgaattggaa 780
gcaaatgca aagatgcaat ccacatggg acctttgggcttgctctag tggcctcttc tgcgctcttc 840
ttgacctgtt ggtctgagtt ggtctgagtt ggtctgagtt ggtctgagtt ggtctgagtt ggtctgagtt 900
aacoccttgg acggttgttt ggtctgagtt ggtctgagtt ggtctgagtt ggtctgagtt ggtctgagtt 960
ttttaaaata cagctgtctat tcacacacag tatataacgt ctctcttaagc ctagctgtcgg 1020
tcccgctgt ctatttttgg gacaacagctgacctggag gcaagcatg acatctacata cagctgtcgg 1080
agtggagct ggtgaccttt tcgctgctctg tcgctgctctc tcgctgctctgc ggtgaccttt 1140
aacacattg ccatgctgca ctagatgaac ctctctctctc tcgctgctctg tcgctgctctc 1200
tactatctac attttttttttt tcgctgctctc tcgctgctctgc ggtgaccttt tcgctgctctc 1260
gggtctgtt ctagaagttg gatacttgta gatctctctc gcacatgtac acatctacata cagctgtcgg 1320
caggggaaat acttctgcct tttctaaaaa aaaaaaaaaa aacgcttgg cggacccggtg 1380
acacatctgc ntatactgt gtaaacaagt c 1411

<210> SEQ ID NO: 93
<211> LENGTH: 2187
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 93

ggttgcttt ttctggggc gccgtggggc cccgaggcag gcggctcacc tcgcaaaag 60
tttcgccag gcggccggag gggtggggag cccccctcag tcgcaaaag 120
-continued

ggggtcaag aagtagaactcg tgctgctgctc gagaagcagga gtcagactgcc ggcctaga 180
cagcggaca cagggccacc gtggtgcaag caaaaccacttt ctctgcagcg cgtccttccc 240
tgtttctcag cgctttctgc ctcctggagg ggccctgggc cggcgaccaag aagaaaaacttgg 300
cagcgcgaag aggttccgctg acctgacagg aagtgacgcc caagggcgcc acgtcgtgcc 360
agctgacagc gcctggccctt cagctgttagc cagccctactg ccaaggggcc ggcagtgagg 420
aacactccttg ctcgacacctc ggtggctgcct tcgtgcttgc ggctctgtgct gctgagggcgg 480
aagggccaca cgccctgcaag ggcaaacgagc gttgtgcagcc cggagcactg ggcgcaag 540
gagggctgaa ccggctttggg cagctgtctg cgtcactcca gcaactctcc cgccgctaac 600
gtacgcagaag aggttccgctg acctgacagg aagtgacgcc caagggcgcc acgtcgtgcc 660
tttgctgagc gcaagcagcc gcaactaacc tggcagacat ccaagatccaa ctctgctgac 720
aagggccaca cgccctgcaag ggcaaacgagc gttgtgcagcc cggagcactg ggcgcaag 780
cogaggtcca ccaagagcgct ggggctgccag cacagcagctgc tttgatgttggcg cgtgagggcgg 840
tctacgagctt acctgggtatg ccaacttaacc ccaagatccag ggtcagagcc ctctgctgac 900
gtactcagct octataacttg ggggctgtcag tttgctgagc ggaacacttg ctctgctgac 960
tacgacgcag aagagccataa gttgcaaaaa ggtctgcctgc gtttaactgc ccaagccatcc 1020
agcctacatct ctctgcaccc ccaactataag cggagcactg aagctgctg ttaatatagtca 1080
accaacgagc aagagccataa gttgcaaaaa ggtctgcctgc gtttaactgc ccaagccatcc 1140
ttggccacctgt gttgcttgga ggtgacactg gacgctgcag tggagctggc ctctgctgac 1200
tctagctgagc ccattgacac ggaacacctgc gcaactagcc gcaactacctt ccaagatccag 1260
tctacgagctt acctgggtatg ccaacttaacc ccaagatccag ggtcagagcc ctctgctgac 1320
ttggccacctgt gttgcttgga ggtgacactg gacgctgcag tggagctggc ctctgctgac 1380
acatctaatg ggcgtcagcacc caagagccataa gttgcaaaaa ggtctgcctgc gtttaactgc 1440
taaggtcagc aagagccataa gttgcaaaaa ggtctgcctgc gtttaactgc ccaagccatcc 1500
gcgtcaacag cagctgcact ccctttggtgt gttggggtggt ggtgagctgct gtttaactgc 1560
agcttcgctg acctgcctgg ccctttggtgt gttggggtggt ggtgagctgct gtttaactgc 1620
gcgacctcc acagtaacct ttcacacttc gcaacactgag cccagctactg ctctgctgac 1680
gcggccacct gttgcttgga ggtgacactg gacgctgcag tggagctggc ctctgctgac 1740
cccgctctg acgcgtcctca atcgacacttg atatatgtag cgctctacct ccaagatccag 1800
qagaacatag cagcagctagc cgggtcagcacc caagctcttttt gcacatataa ccaagatccag 1860
gcgtcaacag cagctgcact ccctttggtgt gttggggtggt ggtgagctgct gtttaactgc 1920
acgacctcc acagtaacct ttcacacttc gcaacactgag cccagctactg ctctgctgac 1980
qgggtctctg gcggctgactc cgggtcagcacc caagctcttttt gcacatataa ccaagatccag 2040
ctttctctcttt cttccctctcttt ccaagatccag cgggtacagt gcgtctggg ggtgagctgct 2100
cctttctcttt cttccctctctcttt ccaagatccag cgggtacagt gcgtctggg ggtgagctgct 2160
catagccac aaaaaaaaaaa aaaaaaaaaa 2187
<220>  FEATURE:
<221>  NAME/KEY:  misc_feature
<222>  LOCATION:  (756)
<223>  OTHER INFORMATION: n equals a,t,g, or c

<220>  FEATURE:
<221>  NAME/KEY:  misc_feature
<222>  LOCATION:  (757)
<223>  OTHER INFORMATION: n equals a,t,g, or c

<400>  SEQUENCE:  94

gacgtaacg tcggattccc gggtagaccc acgcgtcgc cgacqgtgna gaagqtaag  60
atggcgcttg ccaagggccg ggtcttggga gtctcagtggc tgcaaaaggg atcocgggaac  120
gtactagccc gcagacgtcc gagcagctcgc acactgacca aggcgtgttt cgccggccgc  180
tatcctaggc cccccacagaa acgcggcggc gcggccagaa agttataatt gcttgctgaa  240
gactacgac cttaaccgga tgatggtgct gggtatggtg gtatcccgaa gtcctctgac  300
cgctcaacgc atgagagaga tcccatggta gctgggcaac gacgcggccct ggcggttaac  360
tgggggtgat cgatctgact gcacactagac atgtacagaa gacacagtgt ggcatctcc  420
cocactctgc ttctctggca tgcctgatgt atgcagcctct tgcgctctct ggtgctcttg  480
atattcagt gctgctgcttg gacgcgtacg ccctcctacac agctgtgagg accagacgac  540
tatcctacg atatcaataca ctggaagagc gcgcgtctgg tctcctaaaga accagacggg  600
gtgctctact atgagatgcc agaggtcttc gttggcttttt ggtgctttaa atctagactc  660
cctctccgtt acgaatttaa ctttaaagaa atcctcttaa aatcactagt ctgtgttaaa  720
saaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa  757

<210>  SEQ ID NO: 95
<211>  LENGTH: 2394
<212>  TYPE: DNA
<213>  ORGANISM: Homo sapiens

<220>  FEATURE:
<221>  NAME/KEY:  misc_feature
<222>  LOCATION:  (1783)
<223>  OTHER INFORMATION: n equals a,t,g, or c

<400>  SEQUENCE:  95

ggcagcggac ccctctgcact tcccccccacc ccaagccgaa cctgggttct gcgagaaccc  60
cacagctcc ccggtctcga ctctcgggttt ccctcggcgt ccctcgccggg gcagagctgt  120
cacagcgccccc tttcctgggt gacgagctcc ccccccaacc cactcctgcc ccgagatcct  180
cctccctccgg gcgcggccct gcgcggtgccg ggtctctgcc ggcggttct ccgggttct  240
ggtcaggg cccccctgtg cctggcctgg gctgccctct ggctcctctc tctctctcctcc  300
cagagctcc ccggcgggcg gcagagagcc gcagagctgg cctctctctc cgcctctctcct  360
gacagcggcc ccctcggcgt gcagagcggc tctctctcctc gctctctctc tctctctctcc  420
acccctcgtc acgcggccct cttcctcccgt tttcctgggt gacgagctcc ccccccaacc  480
cctccgcct ccgagctggg ccctcggcgt gcagagagcc ccggggctgg ctctctctctctc  540
ctgatattcc gcctcggcgc gacgagctcc ccccccaacc cctctctctct cgcctctctctc  600
tccctctctc gctctctctc gctctctctc gctctctctc gctctctctc gctctctctc  660
gacacatggg atagagacttc cctgctcctga tggagttaaaa gaagagatag  720
agatccgctt acagagagcc ccccccaacc cctctctctc gctctctctc gctctctctc  780
gacagagac ccagagagcc ccccccaacc cctctctctc gctctctctc gctctctctc  840
aagagcacc tggcaaaat gntataaac acctctgaat ctttgaagtt attgactgag 900
cagtacact cttaaaagtt atgactcccc cagcttgagga cagaggtggaa acctcagcag 960
gttgcaaga aatttaaacct acctctgca aacctggaggg ttgaaaagtt gacctttcagc 1020
tgaagacgg atcgtccggc acaacaggttg gaaagagtgat tagctacgccc 1080
aacattacg tcacagaaac ccacacacccac gatactaccgcagcgtgacacacacccctctgg 1140
agactacgcc aaaaaacaacg cttttcaagt gattatagcag aacctcagccttccccagtt 1200
aagagagaga tggctcactt acctcttgagt acatcagttc aagctgaact cactagaaaaa 1260
ctgaaaaacct caattcaggct ccaaaacaacg cttcctttcct gcagaggtcta gcaaggtttt 1320
ggaagagacac ccaaaacacc ctaagcagttt gaaatctacaat ccaactaccaac aagacaccctt 1380
cctctctaca caaattgcagc agctctttgag cctaccccaag ctttcctctctc aaccagagac 1440
gtaaagtttt tatcaagagata aagactctcc caattaagcag cagacactga ggtatcatct 1500
coaattctg atctactttttg tcaagacaccagc agctctgtttg cagaaaaattcttcttgagc 1560
aatctctg gttctccagc tcacactttaaatacgtgagac cgacatttttc ggaacactaa 1600
actaaaacct tcgctcctac ccacacacaac caaccttcagc attctactcagac atcatactca 1650
aactctcttt ataagacttta atcgggaga cttcactcat gtcacatcaag gagacacttt 1700
cctttcaagt gtctgctcagc agaatttattaa ctaactcacttg ttaataattcaaatg 1750
ttttcgagc gttctctcagc tttatatatttttgtac ccctactgttttg gaaaaacactttc 1800
atagggcttt ttacagcttg aagacacggc ccctacttct cttccactcagc taattaattttct 1850
atgctcact atcagatttaa ctaattcactt aaggtgcttc aaaaactttgct ctctctttctat 1900
taatagttta attttgctttc tcaagagatat ggaatggagag acctgtaag tgggttttaca 2050
aagtctgga ataattgagcc ctgccttataa aaaaattttac gaaaacactgt tctttttattct 2100
tgcgtcact ccttgcttaca aaaaaagcaac gacaaagatg ttcacactat 2150
gattagag ataacttactca ctgatcataaagactcag ttcagtttgattt tttctttttggt 2200
cgattccataactttaagtttcttattccttctctggatacagttgacttatttttctttt 2250
tttttagact gaaattttgaa tttaagtttctatcacttgaa acacactatttctttttaaacttc 2300
gtgcacacac acttccattg tcttttttttgg gaaaaaaaccttttttcagtaa 2340

<210> SEQ ID NO 96
<211> LENGTH: 672
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 96

agtgctcctgt tggccaggct ggtgctgttt ggtgtatcgcg cagctctacgt caacctcacc 60
cocoaagttc aagaacattc atctctcctag ccccaagagt aagtagaaaa tgcctggagt 120
acccacacccc ccagctctttg ttttatattt ttttttagataa ttaagttttctttt 180
tttttttttttt tttttttttttt ttgttgacctagc tttttttttttttttttttttttttttttttt 240
acacacacacac acacagctcgatt cttataatattc aagtttctgtcgt ttttcttcacat 300
totaagtgct caagcttctgg gataatctgag ggaattttttc ggtgcagggg actcctctaaac 360
aagagagata ttcggcttccttttctttttt aaggggtttt cttcttctac tttttttttttttttt 420
aatttttta atctcatctaat ccaagggacc cctgataaca aaggctaaag tcgtctttttg 480
gaggttggag rgccacccgt gggtctctcct tccaaanatgt gaaagaaga gaacgcgtcgt
540
aastmcgag aacgccacaa aackttctcc ttgctasaag gaaagaatgtccacmaaatg
600
cocataanac gacacgtcggc ataagagga raattgctct ctyttgccacc ccacaacaaa
660
cgtaattac cc
672

<110> SEQ ID NO: 97
<111> LENGTH: 1419
<210> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (517)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (539)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (604)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (676)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (912)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 97

taaagaacag aacacagtga taaaccacat ggcacctaaa acaatatgggt gtgaagtcc
60
cttatagac caaaccattat aacelcaca caccggtctc atcaggtggc ttcaggtgctc
120
tctcgtcacc gtcactccac acatactccct atttaggtct attataacta agaagaaca
180
tgggaagcat tagttgagc tagaacaatga acggagatat attggtatat tggataacta
240
caactcttc aataaaggtgt attcactactg ctataaactag ataaaaatcat ataaagaatc
300
gtaacacaa tagtatttcct tgaataatgt acatcatgta gcaaatatt
360
tgtgatgata cctctagagac gctatcataa aagaggctcc taataactct cttataaaca
420
attataactc cttcctacc tgcacagag ttttagagat aataaattat gcaagaagta
480	tttggagc ctcccyctttg tagctagag ccaccagaa ttttaataat tagttgycnca
540
gaaatcaccg aagtcagcgct ctaaacacact ttaaatactc acagacatana cctctgttagc
600
tatgctgccgataaagctctaaggtct gccacctaaa atccactctttc cattgaccgc
660
toaanagat gcataatccat tggatatpgg aacacactgtg tttttcttca gccaacaattt
720
ttttatataa tggctctgctg tggcttgatt ggtctctgtct atccttaaca ggtagagcca
780
tataaagcg acaacaagac atcctcttaca caccatctg cttgtagat tcttgtatag
840
aataaaata aataaggtt acataactctg tggctggctt tagcttggtat tttttgagc
900
tatataacta gntntnatag aacacatttc aacctccgta cttaactcttg
960
gtctgatatt tctctctagct atgkagatatt tcttttagct tctgtcgtct gcwacgaat
1020
gactcgatt cttatctccc tcaagttgtt gtaaatcttg gataatgctt actgtatatn
1080
atgattgcac agtttcttac caaaacaattt tttatgtag ctctcttacta cttctctgtct
1140
aataacagta gatatactctgc taataataagttaa acatacttt atagaaggaca
1200
gtctctgtc gctccctagga aacctttaatc aagagctgc agtctgatttt attgtcacttt
1260
agtctactga agtggcgata aaacactta caagaacagg ttgcccttag ccagatgatgy
1320
---continued---

```
aatgaacct cttgacactct ttctatatct aacgttcttc aaccaccccc ccacccccc
cccoccttatt tctcttcttc ctctctggtga ttgagccaa ggctgggagt aanggagga

ctagtgacct agggacaaag aaagagtag ctgtgacact cttgagagtct cccacata

cctactact tgggttttaca tggggttac gagaaccagt gggttatatt ttaggagtgt

tggctttcctt cccaggtgtt aacctact tggggcctag tggaaactcg taataccaca

agaaccttct tataaactctc tgtgtgacca ccttttaaas ananaanaa ananaanaa


<210> SEQ ID NO 99
<211> LENGTH: 1145
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 99

```
tttttttttttttttttttt tttgactgaac taagtggttt ttattttaga gaaagccaga
    60
attacaanag actctctctc tcttggygta tggctgtctc agccaaactc tcaacatac
    120
tgccaagacct atgtgggctga gcaccccctg tttaaacttc ttggagctct ggaacctggc
    180
tgacacca gctgttctct tgtgcttctc cctcttctca ttagtgcttc gcacacaactc
    240
aataggctact tttggttctct ataggtttag gagggctga aattttcttt attttttcttt
    300
tttggccagc agacgacactt ggtctcttgta agacacgacg atcaacatgtcc
    360
catgggtctt aaaaatgaa aacacacacc ccaccacagc tcaagaacat tagacacaaca
    420
atttggctt ggttagtga gcaatattgg taataaata acaaatgtttt tggcatttttt
    480
atgtgagat cttctctgta tttcatattgg aaagatggco aagaggtctg cttcctcaat
    540
ttactctcc cttctcaattg tgaagccag ttgctgcaag ctttaatgca gaaatctctga
    600
tggattgaa gaaagatatt gcccacactct tggagatggt itcaaggggtt ggtagaatac
    660
tgagctctat tggcaccagc tgaacanaac actgicactgt cttttgtaag atttgtaatt
    720
aatttgtagc attttttactttttaa gactaatcga ataaaggtat gtagtaacga
    780
atgtgtatat atatatata taagacactt tttgaatttt gttgttccctg tttacctccc
    840
tttttagagt gttgtcactgt atcaatata ttgaaattgg acatccttgg aactgctggc
    900
aggggttccctt gttcaccactt gccacacctt ttcaccagca aggacatcaca
    960
taaccacagt aatgtagtgg gactaaaag aaagtaacct agcaaatcctt cttggtggtt
   1020
tggaacagat cttctttcata ggaagagaa aaagaccaaa ttcggatttt tttgacagaca
   1080
taaaggtctc ttttttattg aaaagttgag aagacgcttc cttaaananaaaaaaaa
   1140

```
-continued-

cctgctgta actcagcgt aagtaactgt tactagaana ttitttatcna ttaactcgca 240
aataattactttc atttttaagta gtttcttcca tttttattct gctagtctcc cnaaatgtgt 300
ctccttttgg atcggcgttt tttttttgttt gttggtttgtt ctgaaanaat ctatcacttt 360
tgtgcttctta ttgctttttt gttttttgttt acctagttgcc cttgcgcccc aaggaagg 420
aataattcct tactaattcct tagttttata taattttgaa taattttat catttatatt actcagytt 480
accaatattg tgcocctttg gtagttatan gttgcattat ttaattttga gtttttttttc 540
cctcactac tatcttttttt gtaacttttaa acttggggyg aaatatcact gttctgctca 600
gaaagccgc ttataatcct tgcagactt ccgaaaccc gcgtatttttttta 660
tataataaaa atyttttgtaa ctataattgaa aaaaaaaa aaaa aaacacct cyaggggggyc 720
cogttcccc tatta 734

<210> SEQ ID NO 101
<211> LENGTH: 713
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURES:
<221> NAME/KEY: misc_feature
<222> LOCATION: (27)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 101

cogcggggcgc gctgtcctcg ctgcggcncac ccgaaacgcc tgcocctgtgt gcccgcgttcc 60
cctgcoccccg cccagctcgt agcctgcgcc cctcactatt cctgcctccat ctgctgtgcc 120
tgtgtgtcgtt cagcggcggg gttgccccgg cttgcgcttg gccgaaaccc gaaagtcccg 180
ttoagaccc ccaactggg gacctgtttg agccgccaag aaactgtgca gacoccgtgc 240
ccttctgga ccccgctttc ctcacactac cgggaagct ttgatggctta gccatattttg 300
acccactcc gagacccagc cttgtgttta tataacttgg caaaaaagcg gtgttccacg 360
gttctggcgc ggtacactcg gcacactgtg tggaagagcg gacaggggca atctctccc 420
cctcactgcag tttgagaatc cccacatat gcccaggtat gcaaggtgca 480
agtatagct gacgcgtcct gcacataact cgcagcaact ctgctctatg tctgctgagg 540
gatcttcccc gctgggtttag ggccacttc gtcggcctc agccgcttac ttggtgatca 600
cotatacag aagcccaatt cttccaaagcat aagcccaag ggcaagaggg 660
aacaagact aaaaagat aataaataat aatatttttaa aaaaactttaaa aaa 713

<210> SEQ ID NO 102
<211> LENGTH: 1080
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURES:
<221> NAME/KEY: misc_feature
<222> LOCATION: (514)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (721)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 102
cogatgtgga ccacactctcg ttttacccca gtttcttgcc cctgacactcgt gtcggatcag 60
taatgctgt gcaacagagc ttttacccca gtttcttcc cgaacactcg gggcctttca 120
cagacacac cttacaccc ctgctttcag gcacactcttg gcctcctctg gccgctcag 180
-continued

tgctgctgct gctgctgctg gagaattct gagaattct gagaattct gagaattct 240
tgctgctgct gagaattct gagaattct gagaattct gagaattct 300
cagccagcag cagccagcag cagccagcag cagccagcag cagccagcag 360
ctcactctc tctcactctc tctcactctc tctcactctc tctcactctc 420
cagccagcag cagccagcag cagccagcag cagccagcag cagccagcag 480
gagaattct gagaattct gagaattct gagaattct gagaattct 540
gagaattct gagaattct gagaattct gagaattct gagaattct 600
gagaattct gagaattct gagaattct gagaattct gagaattct 660
gagaattct gagaattct gagaattct gagaattct gagaattct 720
ntcctcctcata cgcagctctg tctctctctct tctctctctct tctctctctct 780
acgacggagc acgacggagc acgacggagc acgacggagc acgacggagc 840
cagccagcag cagccagcag cagccagcag cagccagcag cagccagcag 900
gagaattct gagaattct gagaattct gagaattct gagaattct 960
cagccagcag cagccagcag cagccagcag cagccagcag cagccagcag 1020
gagaattct gagaattct gagaattct gagaattct gagaattct 1080

<210> SEQ ID NO 103
<211> LENGTH: 489
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 103
gagaattct gagaattct gagaattct gagaattct gagaattct 60
gagaattct gagaattct gagaattct gagaattct gagaattct 120
gagaattct gagaattct gagaattct gagaattct gagaattct 180
gagaattct gagaattct gagaattct gagaattct gagaattct 240
gagaattct gagaattct gagaattct gagaattct gagaattct 300
gagaattct gagaattct gagaattct gagaattct gagaattct 360
gagaattct gagaattct gagaattct gagaattct gagaattct 420
gagaattct gagaattct gagaattct gagaattct gagaattct 480

<210> SEQ ID NO 104
<211> LENGTH: 1529
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 104
gagaattct gagaattct gagaattct gagaattct gagaattct 60
gagaattct gagaattct gagaattct gagaattct gagaattct 120
gagaattct gagaattct gagaattct gagaattct gagaattct 180
gagaattct gagaattct gagaattct gagaattct gagaattct 240
cncatagtc gcagagtttg ccagatagc caggtggaag agaaagatca acaagatgc 300
gttagctct aaacctcttg tcgaacctcg gcagctcaaa ttgaccalga taaattggtt 360
gtagaagtc gcttcctag tattcagac tcotacactca tttygaagta 420
ttatgtgct cctctggtt ctttgtggag taatcggta aaccctctag acgccgggtt 480
ageccctct cctagcagtc cagttgggtt ttgactttac ttgtaagttt taacctgttaa 540
caaagttgct gcctcttttg ttccttctgt cagctgaaca gcagagtttg taacoagcog 600
agtaaaaaaa cgggttctcc ctctctact ttaaatcgtga ttttacactgt ttgctttttt 660
aagaaaaa ggctgctcag aacctgattg ttagttgtg gatactgtctct tttggaacct 720
tatgagatac tttaataaa gaccaagttt ttcacatcttg tcatttgagcc aaaaaagttc 780
gttgattgac acctagtacat tgttgaacac cctctctagt gctctgaaaat tgtgaactgt 840
ttagggcact ttcactgttg tcaccttgta ttgctctctc tataatctttt ggtctatatg 900
caactgagaa agtctacctg tttaataacac tttaataaag tttcttgagaacctcagt 960
ttttttaaaa gctgtagagag aagcttactaa ttttactactc tttactacta aagggaaaaa 1020
ttcctctcttt cttacacttt ttatgtaaagttc tccttttgtag cttggatctct 1080
tttctttctc tttgggttatg ttgtagctagc agctcttactaatcacttact 1140
taccggcg gctgtaaatg ttcttcatgt gttcgtaa gttggtacac aagcttactat 1200
gcctttagtt tttggtcact tggactttgt attttaagca ttttgtagta aggtgatatttt 1260
ataggctttt atctttatat ctctttattgag tcgacagac aaaaaacagtt aatggacttt 1320
tgtttactactt gaaatcacat gccttgtggtt gctcctacat aagctattttaatttttctct 1380	taaacaagtta ttgctttctta aaaaacttgg caatactttg ttgacatctg ttggttattct 1440
ctactttgct cctgtaggc gctcagctacct gactagttgt gcgttaacact gcacttactgacg 1500
ccagctaag gcagagtttg tcctctact ttaaatcgtga ttttacactgt ttgctttttt 1529

<210> SEQ ID NO 105
<211> LENGTH: 2435
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (455)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (2107)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (2435)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 105
tagagggctc gttgggtgga aagatggcggc agacttccgg ttacccctttgt ctgggaagca 60
gtaggattgc tcggatgggg ttcagaggttc ttacctctttct ttcctttgtt 120
ggagctggcg cgggggccc agaaagttgg gcgctgccaag acgttgtagt actttgaacg 180
gagacatctg cttctgacag tcctacaggg tttggtgcaca ggcagttcctt caagtgggaa 240
tatctgagcc aattcctatg tggattaccc actcattcctg atatgaaagag 300
tacagccttt gcctcttttg ttccttctgt cagctgaaca gcagagtttg taacoagcog 360
gcacctaaat acatctggac aaggaagaa gaatggtacg ccagatagc caggtggaag 420
gtacacaaag grwtcgagag cagcccgaggc ctgtnnttgg gaaacatgg acaaatgttg 480
ggctgggag tattttaga caactacccc aatgaggga aagcaagaga gogggatct 540
coctcccttc cagcaaggt gtcaacacct atgaatcatga gogggatggt 600
cggtcatag acgtggagg cttgcaagcc attgtcagca aatccctaa attgaactcttg 660
ctgggtatct ctgcagctca gaggcaatta agcataaag aagcataagt ggcgaagact 720
gagttgaggg actgacatga atgcacccaga gttcatgcttg ccgcggtgtg ctaacgctgg 780
acocctctca tcacggtgga tctctccatg aatcngatgt tcatttccct gaagttggttt 840
 gaactcaagct gcaaggacgcc ggcaagggag aaaaaacgtc attcrgagct gctcttgccc 900
tcattgagac aatgaagagct goctgagatg agacctccca tggcggcctct ggtgaggctg 960
 gcocctcctc tacatttctt tttctcccct ggtyttcttt tattttggca taagcgcttg 1020
tacaaaactc taaaacaatt ggccgaaggca gaggcaagac agcgctttatc ggcaccctct 1080
gtcgacccct cttggttgac ggatagagga agagagcgcc actggctggt 1140
gcatctgcac tggagagctt cttgtgtctct atgcagctggt ttgggactat attctgtcc 1200
tgaggtttgg aatgcagggga cccggcatct ccaggttttg gctagtgggg agaactctct 1260
ggtctggggc gcgacggccc gggaggcaaat ggtctgtgta tggctgctctt cttgctgctc 1320
ttcacatgg gggagagatt ggagagggga tttctggtct gtgtgaccc aatcagggca 1380
acacannttc atacgctggc ctgctgctgt ttttgcacct caaggccttc ctctctgctg 1440
tttgacattc cagaacacca aaaaaactgta aacacccaggg cttcttgaccc atocctctgt 1500
tggytttggc atgggccgca gggctctgcc aacggcagga gtctttcgttg gaaacaggg 1560
tgaaaccttc ttcocgctct tacocctttc ttacocccat cattgtcttc tcttgagcna 1620
aacagctgag ggttgagctgct tttggagctct tctgttggag ggtgtgggcct gcaagaacca 1680
ccttggttcc atcggccttc aatgggtgag acctgggaga tggcttgctcg ctttggtcag 1740
cctgtctctc gccttggtttg cttcggcttct gcggcctcct cttcgggatt actacgtctct 1800
cttgaaccct acctgggaag tttggtatat gggttggttg aaaaaacagct gggccggaga 1860
gacactctag atgcacatgga ttagcgtgac aacgaccaag ccctgttcttt gtcnaacacc 1920
aaacacactg ttgctgtgcct ggtctgagct gttgctgtct tttgctgcct ctctcagctg 1980
gcttgagac atgtgctttga gtttcgatcttt tggatctccct cttggtgtgctg ctaagagct 2040
aaagaggttc ttcttccttc cttgcgtgctttt ggtatttcgc cagagaagaa aatggcttctg 2100
 ttttttnctt aatggcaag cagctgctgct tttctcagag ttcagacgta gcaagaacca 2160
 gaacccacact tttgctgtgcct ggaagctctg acgctgtcgtt gcaagaccaag cttgcgtctgct 2220
ggcttgcct acctctccct tttgcaagag cggcttctcat cttgctcaccct 2280
 tattgcatctc gatcagcctg cagctgctgct cttgagcggc gaaactctgt 2340
 agcttcttct gacacgctgtctgagttttc gaaactctta aatcagggca cttcttgccc 2400
 aaaaaaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaa aaaa

<210> SEQ ID NO 106
<211> LENGTH: 805
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 106
atgaactctta agaattgcast tgggaagact tctcaagag gattgtatgt aacgagtcttg 60
tattgttttt taagaaagta attttaattg taanaacatt gctcggttta ac tctcgaatt 120
gaaatacgct acacttaattg aagagaggg gaggcacttc tttttaggt ggcctgtaac 180
tctatcctgg tgtc tgtcctgg tgtagcccaac ggaagataca ctcacagctt 240
gaggtgcttc tgtgacatct tgtgacatct tgtgacatct gtaaggcagcg tgaaggcttg 300
gaggtggatg gaaagctcaaa aaccttctct gtttctcttg ctttctctta aagactcagag 360
aaccaccaag ggaagctgtgc ctggagtctc cttcgctggag agygacattag aatcctgtag 420
cctgcaacc tggaggaagc cccctgggtca cgaannaccag atcgctccacc gcgattatta 480
cactctctgga aatctctggg gatcttattac tcgcttcttca aagacactttaa gttgtagactc 540
taacgtgctcc aagctgtgctg cttgagtggac tttanacactt aagcaacacac ttcattattc 600
cagagcattga ggagctatca cccctttccag gaataagtgt tgtggaacaa cttgaattga 660
atctctcccc tattataaat ttgtaattta anataanagac atttctcttg aatggctactc 720
gcggtgtata ccagggcagtgt tcggcgatgtta aaagaggtga aaagagtaaa aaccttttgcag 780
gaacacaaaaa aaaacaaaan aaaaat 805
<210> SEQ ID NO 107
<211> LENGTH: 1166
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> LOCATION: (1039)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 107
ggcagcagag gcgcacagtc cgaggtgtcg gctgagcggtg gagaatggcg ttccccggycc 60
ggcgcgcagaga gcgcacagtc cccgccccgg cttcttttttct cggagaaagtt ggcgagcttt 120
gctgagcagc gctgagcagc cggaggcgaa ggcgagagac ggccgacactg gctgagcagc 180
tttctttttg cgcagaggtc aagctctgtg ttcgatgttg gccgctacgtg 240
ttgagtgagc tccttctcttc gttggaggtcc gatgctgttg gccgtgcttc 300
caggtgtttag ggcgcgagc ggcgagagac ggcgagagac ggcgtgagtt 360
gcgcgcacagaga gcgcagcttt gctgaggtttg ctcgcttgagct ctcgcccttg 420
gagtgttgact taattgagct aaaaaaccttt gcagctttttt ctgctgtttt 480
tttctttttg ttctgctgccgc gctgagagac gctgagagac gctgagagac 540
gatcgcgcgc gcgcgcgcgc gacgagcttc cccgccccgg cttcttttttct 600
gctgagcagc gctgagcagc cggagacgtaa gggagggagtc aacctctccc tgcgcttttt 660
tctggtcgtcc gactttttct cccgccccgg ctcgaggttc ctcgctttttct 720
gggtgagttt gtcagcagttt ctcgaggttc ctcgctttttct 780
ggcagcagag gcgcacagtc cgaggtgtcg gctgagcggtg gagaatggcg ttccccggycc 840
tcgagctgctgc gcgcacagtc cccgccccgg cttcttttttct cggagaaagtt ggcgagcttt 900
tgcgagtcgac gcgcagcttt gctgaggtttg ctcgcttgagct ctcgcccttg 960
ttcgagcttgc ctcgaggttc ctcgctttttct aaaaaaccttt gcagctttttt ctgctgtttt 1020
cgcgcgcacagaga gcgcagcttt gctgaggtttg ctcgcttgagct ctcgcccttg 1080
tttgagcga ggagtttggg acctgctctg gcaacataaat gaaacttctct ttccaggag
1140

<210> SEQ ID NO: 108
<211> LENGTH: 596
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 109

agcgccgagc aagctggata acaaggggcc gatgatgtgg gcacacatctg tttgctgtct 60
tgtgtgtat ctggacacgg cggccacagg gaggccatcc ccagacgctg cgcctctatgg 120
cagcggagag tgcaacacag cggcccctct gaggcacagt cccctcatgt ccacccacgg 180
gaaacctag tgcgacccat aggcttttctc ggccgaggag cgtggcagag aatctgacca 240
actcaccgaa gacgaaacgc aagggcgtctg ggggggacgt ctggacagca ttcacgagggc 300
ggggggacgc gcgctctggc tttgctgtgg cgagttctgc cgtggagtgc gcgacagccg 360
gcgccgagc atagctggct cctgtgacgc ggcctgggac acggccagca ccacacacggc 420
cgggggttg ggtggtgagg aactgcaacag gggccactat ccgcccaggg gggggtggtga 490
agaattctg aagctggagg atgcaagac gttcagagct ctggacagcg ggcacagcg 540
ggttcgctg gtggccgagc agggagggga tocgatgcggt ctcagaagcc 586

<210> SEQ ID NO: 109
<211> LENGTH: 1134
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<220> FEATURE: misc.feature
<221> NAME/KEN: (418)
<222> LOCATION: (803)
<223> OTHER INFORMATION: n equals a, t, g, or c
<221> NAME/KEN: misc.feature
<222> LOCATION: (816)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 109

acggactgac ccaggaaggag ccagttggga gaggctctgg gaagagccag cagactggag 60
acggctgtgc ttgactcttg agtcatcaatt cagatcttcg ggcctctcct tggcattctct 120
atcagtgac cggccctttt cttttcctctt cagatctggc cctctctc ctacctcttttt 180
 CGTTCTCT aagcagagctg tggccctttt ccctctctttt gcccttttcttt ttctccttttt 240
acggcctttt cctcatttt ctttcttctt ctttcttctt ctttcttctt ctttcttctt 300
cggtttactctt ctgctcttctt ctgctcttctt ctgctcttctt ctgctcttctt 360
gtggggctca gaggacacag gacgactctggt attgctctgc acacacacacacacacac 420
gacgactctggt attgctctgc acacacacacacacacac 480
acagtagagct gtagatagctg ggtcatcttcc cagggctttt cagttctttt cagttcttttt 540
gcctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt 600
gccctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt 660
tctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt 720
tggctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt ttgctctcttttt 780
<210> SEQ ID NO 110
<211> LENGTH: 1333
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 110

caacctggcgg ccaaatctccg agaaggtgag tcgaacatcgc ggaattctga aagacctctc
tgagtttcga tttcataaag tttgctcatt gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc
tttctcaaatccccg ccacgacctg cagcaggttcc ggaaaattct gcagacgttcc
ttgccatatgcagaggttgaga gattctgcat gcagacgttcc ggaattctga aagacctctc

<210> SEQ ID NO 111
<211> LENGTH: 1015
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<220> FEATURE: misc_feature
<223> OTHER INFORMATION: n equals a,t,g, or c
-continued

gcagcagcgc gcacagcagc gcacagcagc gcacagcagc cttaagctgt cgtatcttttt cagctatcct 60
cacagcagc gccacagcagc gcacagcagc cccgctggttt ccaagatgct cttgggaga 120
cgtatcttttt cagctatcct cagctatcct cagctatcct cagctatcct cagctatcct 180
ggaggcgggg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc 240
acacagcagc cacacagcagc cacacagcagc cacacagcagc cacacagcagc cacacagcagc 300
acacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 360
gcagcagcgc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 420
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 480
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 540
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 600
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 660
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 720
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 780
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 840
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 900
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 960

<210> SEQ ID NO 112
<211> LENGTH: 711
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (345)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 112

gcagcagcgc gcacagcagc gcacagcagc gcacagcagc cttaagctgt cgtatcttttt cagctatcct 60
cacagcagc gccacagcagc gcacagcagc cccgctggttt ccaagatgct cttgggaga 120
cgtatcttttt cagctatcct cagctatcct cagctatcct cagctatcct cagctatcct 180
ggaggcgggg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc 240
acacagcagc cacacagcagc cacacagcagc cacacagcagc cacacagcagc cacacagcagc 300
acacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 360
gcagcagcgc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 420
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 480
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 540
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 600
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 660
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 720
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 780
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 840
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 900
ggagggcgg ggggctgcttt cagcagcagc gcacagcagc gcacagcagc gcacagcagc gcacagcagc 960

<210> SEQ ID NO 113
<211> LENGTH: 1076
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1029)
<223> OTHER INFORMATION: n equals a, t, g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1037)
<223> OTHER INFORMATION: n equals a, t, g, or c
<211> NAME/KEY: misc_feature
<222> LOCATION: (1040)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 113

ggcaccaggg gnaagcagcc ccccccggac tcctctcttg caggttcttta tacgctctgta 60
cgggaactt ccgctcgctt aaacccacgc tgtgctgtaa ctctatttgg ttctctcctga 120
cotaatctct ttgctcacct tcgctgctca cttctttttc ctgcttcgaa atccgctggc 180
catattgag ccgcttcttc ggtcttcttc cccctttttc cttctgaaggg acgcctccttg 240
gtcttctctt atctctctgt ctctggttta ggggatttccg agaccccttc tgcggcgaagc 300
gacacacg ggctcgcttg ttctctcttt aclctctctcc cctgggagcc gcgcctcctgcg 360
gggtctccct ccgggaggtt gcgtctctctc ggcgcttgcac tgtgctcgctt caaactcactc 420
tagctttgcc gcaagcaggg gnaagcagcc ccccccggac tcctctcttg caggttcttta 480
atactggttc ccagggcagc gacgagcttg ccctgtcact tgcgtctctt ctgctctctt 540
tgcgtctcttt gcctctggtg gtcgctgctt gcgctctctt ctgctctctt 600
tactcacttt ccgggctttt gtggtgccttg gcgtctgctt gcgctctctt 660
tcgctcctct gagaaggctt gcgctctctt gcgctctctt gcgctctctt 720
tggtgctcct acagctagcc ggtctcctct tgcggtcctt gccctcctct gccctcctct 780
tggtgctcct acagctagcc ggtctcctct tgcggtcctt gccctcctct gccctcctct 840
cggttggcctt acagctagcc ggtctcctct tgcggtcctt gccctcctct gccctcctct 900
tggtgctcct acagctagcc ggtctcctct tgcggtcctt gccctcctct gccctcctct 960
tattctcttt gttctctcttt gccctcctct gattaatag ttttctctct tgaaaaaaa 1020
anananana nanaanannag gggggggggc gcagccccat ccgctccata gtaggt 1076

<210> SEQ ID NO 114
<211> LENGTH: 1525
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (78)
<223> OTHER INFORMATION: n equals a, t, g, or c

<400> SEQUENCE: 114

cgcgatgta taactatgac ccccccggac tcctctcttg caggttcttta 60
cgctctgctt cctctctcttt gccctcctct gcgctctctt gcgctctctt gcgctctctt 120
gccgctcttt ccccccggac tcctctcttg caggttcttta 180
tggtgctcct acagctagcc ggtctcctct tgcggtcctt gccctcctct gccctcctct 240

ggcgctcttt ccccccggac tcctctcttg caggttcttta 300
cgcgctcttt ccccccggac tcctctcttg caggttcttta 360
-continued

cacctccca cctggcccaag aagtttcatcg cggcacttccc cctggtgatg taoctcnagcg 480
gctcttgctc cttcctcttc atgaagcoca tcasacagtga cattggaggg aaactgacct 540
actctccag cctctctctg atcctcgtctgt ttgcgcgttt gcggcagggg tacgcttggc 600
tgggttggc cgtgctacga ggcgtcttgtc ttggttgctgc tgggtcttgtgc acaactcttg 660
taactcctgt gcgcattagc gcggacacta tcggcactcc tccgacacac gcattaatcg 720
gtaacgctct atgaaccttc tggataagtgg gcgcaatgag cttgcagctca tggcactca 780
gagctgctac cctggccccc ttagcctctg cttgagggct tcggctgacct ttacacactg 840
gggctagtg tcggcttgagc gggcggctgg gcggcggctg gtcgcgtcgc gcccgtgcct 900
cctgcttgg gcgccttgcgc tgcggagcct gcggcctgggt gcccgtgcct gcggcctgggt 960
gagaacctgc acgcctccgc cggcgactgg gcggcctgct gcggcctgct gcggcctgct 1020
ggggaagaag cccacactgac cctccacctc ctctggaccc cccttccactc tccctcactc 1080
gtccctgcgg tggggctgggt ggggcgaggc cggctggctgc gcggcagccc tggcagggac 1140
cctcttggtt tgagggttgt ccccttcttc acctcaacac atccacgccc tcggaggtct 1200
ttggttgtgc ctcctggcca ggacacagga agtggagact gcacacgagct cttggggagc 1260
cgtacatct gcggcagcct gcggcagcct gcggcagcct gcggcagcct gcggcagcct 1320
gcgcgctgg ccctgctggct gcggcagcct gcggcagcct gcggcagcct gcggcagcct 1380
gctgctgggc ccctgctggct gcggcagcct gcggcagcct gcggcagcct gcggcagcct 1440
cgacacacac agatggagag gcacaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa 1500
aaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa 1525

<210> SEQ ID NO 115
<211> LENGTH: 1350
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<222> LOCATION: (15)
<223> OTHER INFORMATION: n equals a,t,g, or c
<212> LOCATION: (1343)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE:

ggcaccagtct gcaccagtct gcgctctctc ttggctagtgc gcgccttgct gcggccttgct 60
ggcccttgct cgcgcctggc gcgcctgcct gcgccttgct gcgcctggc gcgcctgcct 120
ggcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct 180
gatctcctat tctcttattc gcacatgaa aggcttcagcg gatctcctat tctcttattc 240
gagccgccct gcggctccacct gcggcactttg gcagtgctctgctgc gcggcactttg gcggcactttg 300
gatctaaccg atcgagaaaa aggggctaag ggtagtcttg cggagtctctt aaggggctatg 360
ggatctctc ggtcctttctg ttggctagtgc gcgccttgct gcgccttgct gcgccttgct 420
ggcgccttgct cgcgcctggc gcgcctgcct gcgccttgct gcgcctggc gcgcctgcct 480
ggcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct gcgcctgcct 540
gggcgggaa agcccctgtgc agcggctcag ttatgtggag ttgactgtgc tggcagcag 600
ggctgcttc acacacacag gactctgtgc tttgactgct gcgcctttct gcggcactttc 660
tatatgct tcatgtgc tcaagcctgt gcgcctttct gcggcactttc gcggcactttc 720
gaatttacaag ataaggactt tgccatgttg atatatataa gcaccatcag caaatggaata
780
gcattttgaa ctaaagaaac gauagaaaaa ggacgagttt gcaagcatct acatagctct
940
ggagccct ctaagtctga tcttagttgc ggcaccagac tgggtgatgt tttcaaccac
960
cocctgatct ctgctggcag acctacacaa cgacgtgtaa aagtttctc cctccagaaa
960
acataaatg aggctctgggg aatcaacagct gataatagtgta cactctgttgc atctggagtt
1020
atgacgaaga anagctagaag ctttccaaag cttgaatatt gigaacgctca tctatatttta
1080
gtaatctgtg ctagcactaa tcaaatatac tcagactgtt aacatctcaag agacctttttt
1140
atatatcaac taagatcaact tttgacatc gtttaaatat caacagatgt gcataatggga
1200
agcccttctt gatataagtg gcaacggggg gacactatggg aatgttatgtc tctaatattttt
1260
taggaactaas aatttttggg ctgaanaaaaa aaanaaaaaa aactcgaggg ggacacocggg
1320
cccactctgg cccctttgggg ggaagttttta
1350
<210> SEQ ID NO: 116
<211> LENGTH: 2527
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 116
ctttcgtac ccctcggcgg cagagaaaaa cagcttgccca cggctagtgc ggtgcctggg
60
gtggcggg ccagcttgccca aagaaatcaaa gttgacaaat aacgccgaag actgggaatt
120
goagagac gccagaaaaa cctaacagac ccacagatgc aacagaaaaa aagaggtcc
180
agtccaggg cccagaggtg cgttccctgg tttgacagtt tcgtaacagtc gtttgtctgg
240
agcggaagac aactaagtgag gctagccctt ggagagatgc ccgctttttc atgtattttc
300
ctgtaatct ctgatgtgtg acaatatgcg cactgtctct astaaataac tcgctttttag
360
tctgagac ccagagagga atgttctcgg cagcctcgaa gttgctataac tgctctggga
420
tggcctttc gttgcttttc cagagatctc atacccattc actcagggc gcaacattgt
480
aagccaaaat cagagatctt ccccatttcc gtttgcctaa gttttttgaa gtttttttgat gtttctcttt
540
gtaatctttc gaaagagaggg attgtaaagg actctaaacaa aagacccggg cagagaaat
600
taaagttct actgaggccg attctgaata tgaanagaaaaa gggcccctcg aagttggtctg
660
gaaaaagac tctgtaagtt taaatggct gtccagccag gtttggagaa ttctagagga
720
aagggatatg tgaattcctggg atgcattttc taagataaa taatataatgt ttggcagaaaaa
780
taaacttcct tggcagaaaaa cagagatcga aacacattca goaagtgaana taataaatgt
840
gatatttggt gaggcttggc gtttgcttga aagtttgccg accctattga atgctttttat
900
gagagaggg gatattttgc agcttctccat ttttttttcgt tttctttgat gggctgtttc
960
aactactcag gcacactagct tctggacgtg gactctgag aagaagagaa ggtgtagcgg
1020
tgtagatac cttttctgct gttatgaggg ccagctcttg gcaatttttc gcataagaga
1080
gtttttttt gaagagagga aggctgttgg tggctcagag tagaaaaagc gcttggaaaa
1140
cgctttcat atattagttg tggcagacaag aggattttgc cgtattttgg gagatctttca
1200
agtctgtgg cagatgtttt gaaatgaggg tttgtgttag tttcatatcgt cttctaatgt
1260
gtaaagggc aaattttaag atataataag tcgtgagttgc tttccttttc aataatagaa
1320
cctgcagcac agaagcatct cctgtagata gacctagcata gatactggagct cctagagag
1380
-continued

ggctacccg ccgccctgcc tcctctcctaa ccctcttggt ggtgccacaa agagatgacga 1440
tgtttcttgg attgtgggta tgaacgaacga tcgtgcaagtg tggagagag aagttgtgaa 1500
ttcggagcc cccttgctgc ctccttccc axttcctgct atgctgtgg caaacaagta 1560
ggtacctcgg cctggccgag ccgcctggtg ctcgacggcc accttttaca ttggttgagc 1620
agaacctgtgc gccatgcctc atccgagaca aagggaagat cttttagcc caagctcagtg 1680
tgcaaaagct ctgcaagatcc ccctctgttta acaccccttg gaaaattgctt cccttctggt 1740
tgcgcttac gccagaaaaa aagaggttac ggactacatg gatctgtgaca acatagaga 1800
ccttgaattt attttcgagaa ccagatagcc caaaccctgt caagggacgcc aagaaaccacc 1860
ttgatgtgct cttgtgctctc ccggattgag cctgtgctgca gcctcattaca cacttcgctg 1920
gaaaccttct ggtgttccaca cctctactcct cctcctgaca cttcctgata tccctatcatg 1980
gacccacgt atctctctcttg gtcgcttccgt cttgagctgtg ccctcttcctaatcctaa 2040
cagactcattt ccotctatctg gatctagtt cttgctgctt tacatgctct ctctgcttac 2100
agttgcaac acagactgtt ttattgtacttt ttcctactctt ttagtatattaccatcct 2160
atacactttt aaaaattctc ttttacttat tattagtctg tctctctcgt gatttttctca 2220
agcgtttata tcattgttac ccaaggatt tctaatattatct tgggttttttttttttttctctta 2280
aaaaaagaaa aaaaaacata aacaataaatt tggctgaccc cttggtttttgg gattttttctaca 2340
agcctttactg agcctagtaca tttttttttt acctgaataa taagactgtgctttctttctctct 2400
ccttccagct agcctttgtg ttcctcagct ttatcatct atagctact catgctctgt 2460
gtataagctg ctaaatctc ttttctcctt aatattaattt tttttgcttt ctaaaaaat 2520
aaaaaaa 2527

<210> SEQ ID NO 117
<211> LENGTH: 1998
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: 
<221> NAME/KEY: misc_feature
<222> LOCATION: (88)
<223> OTHER INFORMATION: n equals s,t,g, or c
<221> NAME/KEY: misc_feature
<222> LOCATION: (89)
<223> OTHER INFORMATION: n equals s,t,g, or c
<400> SEQUENCE: 117

cgcatccag acaacccaga agaasaattt ggttggcagaa cagcagaggg tttctatggtc 60
tatataaa caacccaggt aggatttncat tgtattttc tcagccgtta aacaggctcc 120
cctttctgcc cttccagcagttcttgagat gaaccaagag tatactgtga tgtgtgacag 180
cagagagaa ttcagcagca ccagctcagat ggataaggg gattatccttc ttcacccca 240
gagctgtcct tttattgattt gattagacag gattagcgtg atgyggtttt cctgtctctc 300
cotacaaatg tgaacagtctttag tgaagttgcc tgcagctcct ctttttcccct 360
gtattcgtgctcag cagctcagtc tcaagccact aatgtttggaa aagctcaagct aagaagaaaag 420
gacccaaaga agattaaaaa gcacaaaaa gaaaaaaa aactcagaa aaaaatatgaa 480
tatgatgctg anattagatt cctatatcct aactagttct caacaccctc aacccctcatctaatcct 540
aagttgccaagc agagagatgt cagctttaaa cctgtggtat cttotagact tgtataaaatc 600
acagatgcac caaagtcttc ctgcagaaaat gargaagggga aatactggtt ttcgcctcttgcg 660
agttacattcg caccaaatgactg gagaagatc ttagatgtata ttagtctggg ctcctcctat 720
gacactgcgt acctccaccc tttcgttatt ctgtcgttctt cattgggtc caatggaagg 780
ttagttttt atggcaact tttattggtta tcmagaataat tgaagcatca accacatttt 840
tacatcatttg ttagaatcact caattacatct tcaanaaagtt ttaatgttgg accatagaaa 900
ataatcctcttg ttagttcata gaagactacat agtgaagtgt agaacatttt 960
aaatattca ttcogcttct ggttacattc atgaaagtt tgtgaaatct ttttttagac 1020
cacagtaga taatccctctg tcaaaataa aaaaataaaaa aaaaataaaa aaaaataaaa 1080
aggggocggc taccacat 1080

<210> SEQ ID NO: 118
<211> LENGTH: 1679
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KET: misc_feature
<222> LOCATION: (1679)
<223> OTHER INFORMATION: n equals a,t,g, or c

<400> SEQUENCE: 118
tgcaccacag cgcccgggca gacccctaac gcagttactgc cctctgcocg ccgctccgtc 60
cocggactcc ttcggcggcc gggccgggct gggccggggcc ggctggagcg gccatccccg 120
cacgcccgag cgcgggcccc ggcgggggag ccggtcggcg cgcctggggg gcgcctccgg 180
cgcctccag tgcgcgaagtg cgttcacctc ctgcttggc cgcagggagg ccagagagg 240
cgcctcggca aaggggaggg agggctctcg cctccccccc gacggcctcg ggtgcgtcgt 300
caccggccga gttcttggtg taccgctttg ccaagagaga cgcctggttt tgtgcgtctg 360
cctttgcct gactttttgc ttcctgtggt tattttagga gacgctact tgtcataataa 420
cttctttcat ccacagctag acctgtacta ctgcttgaata aatactaca aatacatgtt 480
catatttaa aaccagcctg cagcagcggc aacctgctct taagacaga aagaaagaaa 540
tataaaaa tgtggaagag aagagttgga attatacagt gttgtgctcg cgcattttgc 600
agataagtt gcttcggaaca ttcgttcgtag cttttaaacag aaactctacag cctttaagga 660
ttttcacct gcgactttg aacgctggctc acctgtgcct tcaaatcactgt tgaccaaccag 720
aactacactg gatgctactta taacacttaa ggtgtgacag ctattctgcto agtocctct 780
gattctcg actcaagtta attctgatct gctgaaaccc attattcactt cttcttccttt 840	tattttactg aacagcagctc ctttacactg aacacactct aataattcctaa 900
agttattc cagccggaatt cggagctctg tcttggctcctt ggacccatt 1020
tgtgggatgaa aatactctctt gcgttgcgac agtctccaaa aataattccttct 1080
ttcattattg tagaacaact gttgactact ttacattact taatactacca aataaaccacc 1140
attactataa aacaaaaaaa actgcgtgactt ttgacctagtt attatattccttgaaaattag 1200
ttagaacta ctagaatctg aatactcaacct tagaatctag aggaaatttctt aatactttttct 1260
taattacttt attcattgttt tattttctgc ctaacctactttt aacacactcttctt 1320
agaataata ccataactctg cagataacag tgaattttt gtagtttt 1380
gtgccctgtgc taccctagttt gttagtgcat tgtgacacac attttaatatttt toctcataattt 1440

aaatgtgca gatattttcg tgcataaat atttactat attagagatg atttcaacct 1500

ttagtttat atatcatcag cactgtgctg attaatattt tagaatagtg tgggaattta 1560

agaataact tgggctacta atttgtaata ccctatacttg tcgaactgga tataaatctc 1620

acaacagtgttt ttaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaa 1679

<210> SEQ ID NO: 119
<211> LENGTH: 1411
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> LOCATION: (1391)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc.feature
<222> LOCATION: (1403)
<223> OTHER INFORMATION: n equals a,t,g, or c
<400> SEQUENCE: 119

ggccagcagc cgaccggcgggg caagccgagg aacgccgggg gctcgggggg 60

gaccggg gacgcatgcc caggcccgcttg tggagggcc ttgcagatc atccattgca 120
cagcaccatc cgttagagag aagatctgcc attcatactg gagttgagtt gggagagaaa 180
gagctcctc ctagaatgta atccttctgca atatccatcag tcggtgctct aaaagctgtt 240
gggaaaaat tcagctgatgt ttttacgcc agaaaagta taatctttttttt gagagaagtt 300

gattttggg gccttttgat tcttttttgct gacactgccat taatgtgcag aagagctctcag 360

gagctagtg ctagacaggg aggccccata tttgcagagg tgggtctttttt tttcggtttttt 420
ggttcagtt gatctacacgc ccaacaaaaa ctctttttgag ggacactata tttttttccag 480

gctctttttg tggctgagca caggttttgta aactctagttg ttcgggagtt gatttttgcc 540
cagtctgcttt ggcctttgcc gcttgagttt ttcggtttcctt ttgctgtat 600
gatgcctggttt cctgtcttgca agtttctcca taagtttctg tagaactgtaa gggggagaga 660

<210> SEQ ID NO: 120
<211> LENGTH: 2223
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> LOCATION: (1243)
<223> OTHER INFORMATION: n equals a,t,g, or c
<221> NAME/KEY: misc.feature
<222> LOCATION: (1253)
<223> OTHER INFORMATION: n equals a,t,g, or c
cctccaggac gtttctcacc ttttcagagc tttctcagga cgggcaggg gcgggtgggga
60
cctccagag tggctggag ggcaggggca ggggtcagag aacatgactg tggcgttgt
120
cggactcag aatccagtcgc cggtgctagc cagcggcagc cggcggcacc ctggtgccgc
180
caccaccttt ctggctcaag tcgcttggttg ctccttttcgc ttattggagt
240
cggcgcgttgcc gcggcaggttg aaagaacctg cagcggcagc agctctctgc acgctcggaga
300
gtggacgcaac cagcggccgc aggcttgcgcc agggcagncg ggtgctgcctt cattttttac
360
caggcatcgc ccaaggcaac gcgggtgccg acagcggcagc ggtgctgcctt cattttttac
420
tgtgcgtgacctc gctggggtgc aagagcaggc cggcggcagc ggtgctgcctt cattttttac
480
gctgcgtgacctc gctggggtgc aagagcaggc cggcggcagc ggtgctgcctt cattttttac
540
cggcgcgttgcc gcggcaggttg aaagaacctg cagcggcagc agctctctgc acgctcggaga
600
ggcccacgt cagctgtgctg tcgctgcag cagcggcagc cagcggcagc cctcctacac
660
cggcgcgttgcc gcggcaggttg aaagaacctg cagcggcagc agctctctgc acgctcggaga
720
ggcccacctc tttccgcaac gcggcgcgttg ccacggtgacctt cctcctgatt tattttttac
780
gggcgccctt caggtgaacac ccacacacac acacacacac acacacacac acacacacac
840
cacagtggtac ctgctacgctg gcgtcgccttc tctcctcctg tggcgtgtgtc ggtcgtgtgtc
900
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
960
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1020
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1080
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1140
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1200
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1260
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1320
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1380
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1440
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1500
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1560
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1620
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1680
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1740
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1800
gacctgccgg ccaacctccg ccacccctcg ccacccctcg ccacccctcg ccacccctcg
1860
```
-continued

aacctctccc acataaaaa ctagtggtg cagcccccttg ggcagggcag cccgammaag 
1920
aacctggsccg aatggagggct attgagaaggg agctccccaggg acggacctc 
1980
tgtaaagaa gcatgtggtg tggcttggtg gggctaagat ttcttttttt gttcttcttttt 
2040
ttttagtct tcaagagaa gaggggaggg ggcagacaga gctctttgtg ctataatacc 
2100
aagaacctat ttgtaacctt ttttttttca taaaactttt ccaatgaca saaaaaaaaaaa 
2160
aaaaaaaaaa mmgagggagg gccgctccaa gaggcatcccc tgggagnggg ccacacagaa 
2220
aast 
2223

<210> SEQ ID NO 121
<211> LENGTH: 198
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 121

Met Lys Lys Gin Ser Lys Arg Cys Leu Trp Lys Pro Pro Gly Ser Leu 
1    5     10    15
Arg Arg Leu Trp Trp Met Arg Ala Leu Leu Ile Leu Lys Tyr Ile 
20   25   30

<210> SEQ ID NO 122
<211> LENGTH: 348
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 122

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gin Val Met Leu Leu 
1    5     10    15
His Leu Thr Ala Ala Phe Leu Gin Arg Ala Gin His Xaa Phe Asp Tyr 
20   25   30  35
Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr 
40   45  50  55
Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr 
60  65  70  75  80
Ile Pro Leu Val Pro Gly Arg Asp Glu Phe Val Gly Arg Asp Asp 
85  90  95 100
Phe Asp Asp Ala Asp Gin Leu Arg Ile Gly Asn Asp Gly Ile Phe Met 
120 125
Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu 
140 145 150
Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser 
165 170 175
Gly Phe Gly Leu Ser Leu Ile Lys Try Leu Leu Ile Val Arg Phe Ser 
190 195 200
Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gin Tyr Trp Leu Trp Trp Val 
210 215 220
Phe Leu Val Leu Gly Phe Leu Phe Leu Arg Gly Phe Ile Asn Tyr 
230 235 240
```

OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 122

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gin Val Met Leu Leu 
1    5     10    15
His Leu Thr Ala Ala Phe Leu Gin Arg Ala Gin His Xaa Phe Asp Tyr 
20   25   30  35
Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr 
40   45  50  55
Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr 
60  65  70  75  80
Ile Pro Leu Val Pro Gly Arg Asp Glu Phe Val Gly Arg Asp Asp 
85  90  95 100
Phe Asp Asp Ala Asp Gin Leu Arg Ile Gly Asn Asp Gly Ile Phe Met 
120 125
Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu 
140 145 150
Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser 
165 170 175
Gly Phe Gly Leu Ser Leu Ile Lys Try Leu Leu Ile Val Arg Phe Ser 
190 195 200
Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gin Tyr Trp Leu Trp Trp Val 
210 215 220
Phe Leu Val Leu Gly Phe Leu Phe Leu Arg Gly Phe Ile Asn Tyr 
230 235 240
```
Arg Val Leu Phe Ile Tyr
195

180

<210> SEQ ID NO 123
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 123
Met His Asn Gln Arg Gln Val Phe Leu Phe His Leu Phe Ser Asn Tyr
1    5    10   15
Leu Leu Ser Ile Asn Ser Val Pro Gly Thr Leu Leu Ala Ala Thr Tyr
20   25   30
Cys Leu Asn Met Thr Thr Tyr Gly
35

<210> SEQ ID NO 124
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 124
Met Arg Lys Phe Leu Leu Ala Gln Val Phe Leu Ser Leu Ser Val
1    5    10   15
Met Pro Ser Met Pro Val Thr
20

<210> SEQ ID NO 125
<211> LENGTH: 110
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 125
Met Val Leu Leu Cys Leu Leu Leu Val Pro Leu Leu Leu Ser Leu Phe
1    5    10   15
Val Leu Gly Leu Phe Leu Trp Phe Leu Lys Arg Gln Arg Gln Glu Glu
20   25   30
Tyr Ile Glu Glu Lys Arg Val Asp Ile Cys Arg Glu Thr Pro Asn
35   40   45
Ile Cys Pro His Ser Gly Glu Thr Glu Tyr Asp Thr Ile Pro His
50   55   60
Thr Asn Arg Thr Ile Leu Lys Glu Asp Pro Ala Asn Thr Val Tyr Ser
65   70   75   80
Thr Val Glu Ile Pro Lys Lys Met Glu Asn Pro His Ser Leu Leu Thr
85   90   95
Met Pro Asp Thr Pro Arg Leu Phe Ala Tyr Glu Asn Val Ile
100  105  110

<210> SEQ ID NO 126
<211> LENGTH: 63
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 126
Met Leu Leu Leu Phe Ile Tyr Phe Tyr Ser His Pro Ala Pro Val Pro
1    5    10   15
-continued

Ala Gly Ala Thr Ser Lys Pro Arg Tyr Arg Val Ile Thr Cys Gly Pro
   20          25          30

Ala Ser Val Phe Ser Thr Ser Phe Ser His Ser Pro Pro Ala Arg Cys
   35          40          45

Leu Gly Arg Leu Glu Gln Met Phe His Phe Gly Leu Ala Ser Gly
   50          55          60

<210> SEQ ID NO 127  
<211> LENGTH: 30  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens  

<400> SEQUENCE: 127  
Met Pro Phe Pro Ile Ser Ile Leu Gln Leu Cys Leu Gln Ile Ser Arg
   1      5     10          15
Leu Ser Phe Cys Leu Gln Lys Ile Tyr Lys Ile Pro Phe Val
   20     25          30

<210> SEQ ID NO 128  
<211> LENGTH: 53  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens  

<400> SEQUENCE: 128  
Met Ala Ala Ala Cys Arg Ser Val Lys Gly Leu Val Ala Val Ile Thr
   1      5     10          15
Gly Gly Ala Ser Gly Leu Gly Leu Ala Thr Ala Asp Leu Trp Gly
   20     25          30
Arg Glu Pro Leu Leu Cys Phe Trp Thr Cys Pro Thr Arg Val Gly Arg
   35     40          45
Pro Lys Pro Arg Ser
   50

<210> SEQ ID NO 129  
<211> LENGTH: 57  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens  

<400> SEQUENCE: 129  
Met Leu Leu Val Tyr Asp Leu Tyr Leu Xaa Pro Lys Leu Trp Ala Leu
   1      5     10          15
Ala Thr Pro Glu Lys Asn Gly Lys Gly Ala Arg Xaa Gly Asp Gly Thr
   20     25          30
Pro Ala Glu Ala Phe Trp Asp Phe Trp Ser His Leu Ile Ser Ala Asp
   35     40          45
Pro Glu Thr Trp Glu Arg Ala Ala Pro
   50     55

<210> SEQ ID NO 130  
<211> LENGTH: 216  
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 130

Met Arg Leu Ser Ala Leu Leu Ala Leu Ala Ser Lys Val Thr Leu Pro
  1  5  10  15

Pro His Tyr Arg Tyr Gly Met Ser Pro Pro Gly Ser Val Ala Asp Lys
  20  25  30

Arg Lys Asn Pro Pro Trp Ile Arg Arg Arg Pro Val Val Val Glu Pro
  35  40  45

Ile Ser Asp Glu Asp Trp Tyr Leu Phe Cys Gly Asp Thr Val Glu Ile
  50  55  60

Leu Glu Gly Lys Asp Ala Gly Lys Gly Lys Val Val Gin Val Ile
  65  70  75  80

Arg Gln Arg Asn Trp Val Val Gly Leu Asn Thr His Tyr Arg
  85  90  95

Tyr Ile Gly Lys Thr Met Asp Tyr Arg Gly Thr Met Ile Pro Ser Glu
 100 105 110

Ala Pro Leu Leu His Arg Gln Val Lys Leu Val Asp Pro Met Asp Arg
 115 120 125

Lys Pro Thr Glu Ile Glu Trp Arg Phe Thr Glu Ala Gly Glu Arg Val
 130 135 140

Arg Val Ser Thr Arg Ser Gln Arg Ile Pro Lys Pro Glu Phe Pro
 145 150 155 160

Arg Ala Asp Gly Ile Val Pro Glu Thr Trp Ile Asp Gly Pro Lys Asp
 165 170 175

Thr Ser Val Glu Asp Ala Leu Glu Arg Thr Tyr Val Pro Cys Leu Lys
 180 185 190

Thr Leu Gln Glu Glu Val Met Glu Ala Met Gly Ile Lys Glu Thr Arg
 195 200 205

Lys Tyr Lys Lys Val Tyr Trp Tyr
 210 215

<210> SEQ ID NO 131
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 131

Met Ser Leu Arg Gin Lys Ser Ser Ser Phe Arg Leu Met Val Met Ser Leu
  1  5  10  15

Thr Ile Leu Lys Leu Ser Lys Thr Thr Val Leu Cys Leu Arg Cys Leu
  20  25  30

His Ser Ser Leu Thr Trp Arg Asp Gly Ala Arg Cys Ile Asn Ala
  35  40  45

Glu

<210> SEQ ID NO 132
<211> LENGTH: 58
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 132

Met Ser Gly Ser Phe Ile Leu Cys Leu Ala Leu Val Thr Arg Trp Ser
  1  5  10  15
<210> SEQ ID NO: 133
<211> LENGTH: 52
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 133

Met Cys Phe Arg Phe Phe Leu Phe Cys Ser Arg Ile Leu Leu Lys Leu
1     5     10    15

Phe Phe Leu Leu Phe Pro Ala Ser Ala Phe Pro Leu Ser Thr Arg Ser
20    25    30

Ser Leu Ser Val Asn Glu His Val Val Val Ser Pro Arg Ser Thr Val
35    40    45

Ser Ile Ser Arg
50

<210> SEQ ID NO: 134
<211> LENGTH: 540
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (337)

<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins.

<400> SEQUENCE: 134

Met Val Arg Thr Asp Gly His Thr Leu Ser Glu Arg Asn Tyr Gln
1     5     10    15

Val Thr Asn Ser Met Phe Gly Ala Ser Arg Lys Phe Val Glu Gly
20    25    30

Val Asp Ser Asp Tyr His Asp Glu Asn Met Tyr Ser Gin Ser Ser
35    40    45

Met Phe Pro His Arg Ser Gin Leu Met Leu Ala Ser Pro Ser Thr
50    55    60

Ser Gly Gin Leu Ser Gin Phe Gly Ala Ser Leu Tyr Gly Gin Gin Ser
65    70    75    80

Ala Leu Gly Leu Pro Met Arg Gly Met Ser Asn Thr Pro Gin Leu
85    90    95

Asn Arg Ser Leu Ser Gin Thr Gin Leu Pro Ser His Val Thr Pro
100   105   110

Thr Thr Gly Val Pro Thr Ser Leu His Thr Pro Pro Ser Pro Ser
115   120   125

Arg Gly Ile Leu Pro Met Asn Pro Xaa Asn Met Met Asn His Ser Gin
130   135   140

Val Gly Gin Gly Ile Gly Ile Pro Ser Arg Thr Asn Ser Met Ser Ser
145   150   155   160

Ser Gly Leu Gly Ser Pro Asn Arg Ser Ser Pro Ser Ile Ile Cys Met
Pro Lys Gln Gln Pro Ser Arg Gln Pro Phe Thr Val Asn Ser Met Ser
165 170 175
180 185
Gly Phe Gly Met Asn Arg Asn Gln Ala Phe Gly Met Asn Asn Ser Leu
190 195 200 205
Ser Ser Asn Ile Phe Asn Thr Gly Thr Asp Gly Ser Glu Val Thr Gly
210 215 220
Leu Asp Leu Ser Asp Phe Pro Ala Leu Ala Asp Arg Asn Arg Arg Glu
225 230 235 240
245 250 255
Gly Ser Gly Asn Pro Thr Pro Leu Ile Asn Pro Leu Ala Gly Arg Ala
255
Pro Tyr Val Gly Met Val Thr Lys Pro Ala Asn Glu Gln Ser Gln Asp
260 265 270
Phe Ser Ile His Asn Glu Asp Phe Pro Ala Leu Pro Gly Ser Ser Tyr
275 280 285
Lys Asp Pro Thr Ser Ser Asn Asp Ser Ser Asn Leu Asn Thr
290 295 300
Ser Gly Lys Thr Thr Ser Ser Thr Asp Gly Pro Lys Phe Pro Gly Asp
305 310 315 320
Lys Ser Ser Thr Gln Asn Asn Asn Gln Gln Lys Gly Ile Gln
325 330 335
Val Leu Pro Asp Gly Arg Val Thr Asn Ile Pro Gln Gly Met Val Thr
340 345 350
Asp Gln Phe Gly Met Ile Gly Leu Thr Phe Ile Arg Ala Ala Glu
355 360 365
Thr Asp Pro Gly Met Val His Leu Ala Leu Gly Ser Asp Leu Thr
370 375 380
Leu Gly Leu Asn Leu Asn Ser Pro Glu Asn Leu Tyr Pro Lys Phe Ala
385 390 395 400
Ser Pro Trp Ala Ser Ser Pro Glu Pro Glu Asp Ile Asp Phe His
405 410 415
Val Pro Ser Glu Tyr Leu Thr Asn Ile His Ile Arg Asp Leu Leu Ala
420 425 430
Ala Ile Lys Leu Gly Arg Tyr Gly Glu Asp Leu Leu Phe Tyr Leu Tyr
435 440 445
450 455 460
Tyr Met Asn Gly Gly Asp Val Leu Gln Leu Ala Ala Val Glu Leu
465 470 475 480
Phe Asn Arg Asp Trp Arg Tyr His Lys Glu Glu Arg Val Trp Ile Thr
485 490 495
Arg Ala Pro Gly Met Glu Pro Thr Met Lys Thr Asn Thr Tyr Glu Arg
500 505 510
Gly Thr Tyr Tyr Phe Phe Asp Cys Leu Asn Trp Arg Lys Val Ala Lys
515 520 525
Glu Phe His Leu Gly Tyr Asp Leu Glu Glu Arg Pro His Leu Pro
530 535 540
545

<210> SEQ ID NO 135
<211> LENGTH: 57
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 135
Met Ile Cys Pro Gin Cys Pro Leu Ser Leu Leu Leu Cys Leu Ile Ser Ser
  1    5    10    15
Leu Cys Ser Leu Val Ile Gin Ile Ser Leu Lys Thr Ile Arg Asp Ile
  20   25    30
Thr Leu Leu Asn Met Val Gly Ile Lys Phe Ser Ile Ser Leu Ser Asn
  35   40    45
Lys Ile Asn Ile Asn Ser Arg Thr Trp
  50   55

<210> SEQ ID NO 136
<211> LENGTH: 221
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 136
Met Thr Leu Arg Pro Ser Leu Leu Pro Leu His Leu Leu Leu Leu Leu
  1    5    10    15
Leu Leu Ser Ala Ala Val Cys Arg Ala Glu Ala Gly Leu Glu Thr Glu
  20   25    30
Ser Pro Val Arg Thr Leu Gln Val Glu Thr Leu Val Glu Pro Pro Glu
  35   40    45
Pro Cys Ala Glu Pro Ala Ala Phe Gly Asp Thr Leu His Ile His Tyr
  50   55    60
Thr Gly Ser Leu Val Asp Arg Ile Ile Asp Thr Ser Leu Thr Arg
  65   70    75    80
Asp Pro Leu Val Ile Leu Gly Glu Gin Val Ile Pro Gly Leu
  85    90    95
Glu Gin Ser Leu Leu Asp Met Cys Val Gly Glu Lys Arg Arg Ala Ile
 100  105   110
Ile Pro Ser His Leu Ala Tyr Gly Lys Arg Gly Phe Pro Pro Ser Val
 115  120   125
Pro Ala Asp Ala Val Val Tyr Asp Val Glu Leu Ile Ala Leu Leu
 130  135   140
Arg Ala Asn Tyr Trp Leu Lys Leu Val Lys Gly Ile Leu Pro Leu Val
 145  150   155   160
Gly Met Ala Met Val Pro Ala Leu Leu Gly Ile Gly Tyr His Leu
 165   170   175
Tyr Arg Lys Ala Asn Arg Pro Lys Val Ser Lys Lys Leu Lys Glu
 180  185   190
Glu Lys Arg Asn Lys Ser Lys Lys Lys
 195  200

<210> SEQ ID NO 137
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 137
Met Phe Leu Arg Leu Tyr Leu Ile Ala Arg Val Met Leu Leu His Ser
  1    5    10    15
Lys Leu Phe Thr Asp Ala Ser Ser Arg Ser Ile Gly Ala Leu Asn Lys
  20   25    30
Ile Asn Phe Asn Thr Arg Phe Val Met Lys Thr Leu Met Thr Ile Cys
35 40
Pro Gly Thr Val Leu Leu Val Phe Ser Ile Ser Leu Trp Ile Ile Ala
50 55 60
Ala Trp Thr Val Arg Val Cys Gln Ser Pro Glu Ser Pro Ala Gln Pro
65 70 75 80
Ser Gly Ser Ser Leu Pro Ala Trp Tyr His Asp Gln Gln Asp Val Thr
85 90 95
Ser Asn Phe Leu Gln Ala Met Trp Leu Ile Ser Ile Thr Phe Leu Ser
100 105 110
Ile Gly Tyr Gly Asp Met Val Pro His Thr Tyr Cys Gly Lys Gly Val
115 120 125
Cys Leu Leu Thr Gly Ile Met Gly Ala Gly Cys Thr Ala Leu Val Val
130 135 140
Ala Val Val Ala Arg Lys Leu Glu Thr Lys Ala Glu Lys His Val
145 150 155 160
His Asn Phe Met Met Asp Thr Glu Leu Thr Lys Arg Ile Lys Asn Ala
165 170 175
Ala Ala Asn Val Leu Arg Thr Trp Leu Ile Tyr Lys His Thr Lys
180 185 190
Leu Leu Lys Lys Ile Asp His Ala Lys Val Arg Lys His Gin Arg Lys
195 200 205
Phe Leu Pro Ser Tyr Pro Pro Val
210 215

<210> SEQ ID NO 138
<211> LENGTH: 102
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 138
Met Ser Asn Thr Thr Val Pro Asn Ala Pro Gln Ala Asn Ser Asp Ser
1 5 10 15
Met Val Gly Tyr Val Leu Gly Pro Phe Phe Leu Ile Thr Leu Val Gly
20 25 30
Val Val Val Ala Val Met Tyr Val Gin Lys Lys Arg Val Asp
35 40 45
Arg Leu Arg His Leu Leu Pro Met Tyr Ser Tyr Asp Pro Ala Glu
50 55 60
Glu Leu His Glu Ala Gln Glu Leu Leu Ser Asp Met Gly Asp Pro
65 70 75 80
Lys Val Val His Gly Trp Gin Ser Gly Tyr Gin His Lys Arg Met Pro
85 90 95
Leu Leu Asp Val Lys Thr
100

<210> SEQ ID NO 139
<211> LENGTH: 112
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 139
Met Arg Glu Cys Gin Glu Glu Ser Phe Trp Lys Arg Ala Leu Pro Phe
1 5 10 15
-continued

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

<210> SEQ ID NO 140
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 140
Met Lys Asn Asp Arg Asn Gln Gly Phe Ser Leu Leu Gln Leu Ile Asp
1 5 10 15
Trp Asn Lys Pro
20

<210> SEQ ID NO 141
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 141
Met Gly Thr Gln Pro Pro Val Val Ala Gly Phe Thr Ile Pro Met Leu
1 5 10 15
Gly Tyr Thr Val Arg Val Leu Thr Phe His Leu Ser Cys Ser
20 25 30

<210> SEQ ID NO 142
<211> LENGTH: 99
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 142
Met Lys Ile Pro Val Leu Pro Ala Val Val Leu Leu Ser Leu Leu Val
1 5 10 15
Leu His Ser Ala Gln Gly Ala Thr Leu Gly Gly Pro Glu Glu Glu Ser
20 25 30
Thr Ile Glu Asn Tyr Ala Ser Arg Pro Glu Ala Phe Asn Thr Pro Phe
35 40 45
Leu Asn Ile Asp Lys Leu Arg Ser Ala Phe Lys Ala Asp Glu Phe Leu
50 55 60
Anh Trp His Ala Leu Phe Glu Ser Ile Lys Arg Lys Leu Pro Phe Leu
65 70 75 80
Anh Trp Asp Ala Phe Pro Lys Leu Lys Gly Leu Arg Ser Ala Thr Pro
85 90 95
Asp Ala Gln

<210> SEQ ID NO 143
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 143

Met Val Trp Gly Leu Leu Leu Gly
  1  5

<210> SEQ ID NO 144
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1-30)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 144

Met Leu Pro Leu Ser Leu Leu Phe Leu Phe Phe Ser Thr Val Ser
  1  5  10  15
Ser Phe Cys Gly Met Pro Leu Ala His Thr Arg Ala Xaa Ala His
   20  25  30  35
Thr Arg Thr Phe Ala Ser Arg

<210> SEQ ID NO 145
<211> LENGTH: 131
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 145

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

<210> SEQ ID NO 146
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 146

Met Gly Ala Pro Ser Leu Thr Met Leu Leu Leu Lys Val Gln Pro
  1  5  10  15
-continued

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

<210> SEQ ID NO 147
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 147
Met Cys Leu Ile Phe Leu Leu Leu Leu Leu Leu Ser Phe Ser
1 5 10

<210> SEQ ID NO 148
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 148
His Pro His Gin Asp Ser Gin Pro
1 5

<210> SEQ ID NO 149
<211> LENGTH: 68
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 149
Met Asn Thr Ser Tyr Ile Leu Arg Leu Thr Val Val Ser Val Val
1 5 10 15
Ile Tyr Leu Ala Ile His Pro Leu Leu Ser Phe Ser Leu Gin Ser Pro
20 25 30
Leu Leu Val Pro Trp Arg Asp Cys Gin Asn Ile Trp Lys Ser Gly
35 40 45
Ser Val Trp Tyr Lys Arg Trp Thr Leu Pro His Met Gin Val Cys
50 55 60
Gln Asp Leu His
65

<210> SEQ ID NO 150
<211> LENGTH: 26
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 150
Met Leu Lys Ile Phe Lys Gin Asn Leu Gin Ile Leu Thr
1 5 10 15
Ser Ile Arg Ile Leu Gin Arg Gin Asn Met
20 25

<210> SEQ ID NO 151
<211> LENGTH: 195
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 151
Met Asp Cys Gin Val Asn Gin Ser Ser Leu Arg Asp Glo Cys Ile
1 5 10 15
Thr Asn Leu Leu Vel Phe Gly Phe Leu Gin Gin Ser Cys Ser Gin Gin Ser
20 25 30
Phe Arg Arg Glu Leu Asp Ala Leu Gly His Glu Leu Pro Val Leu Ala
35  40  45
Pro Gln Trp Gly Tyr Asp Glu Leu Gln Thr Asp Gly Asn Arg Ser
50  55  60
Ser His Ser Arg Leu Gly Arg Ile Ala Asp Ser Glu Ser Gln Glu
65  70  75  80
Asp Ile Ile Arg Asn Ile Ala Arg His Leu Ala Gln Val Gly Asp Ser
85  90  95
Met Asp Arg Ser Ile Pro Pro Gly Leu Val Asn Gly Leu Ala Leu Gln
100 105 110
Leu Arg Asn Thr Ser Arg Ser Glu Asp Arg Asn Arg Asp Leu Ala
115 120 125
Thr Ala Leu Glu Gln Leu Gln Ala Tyr Pro Arg Asp Met Glu Lys
130 135 140
Glu Lys Thr Met Leu Val Leu Ala Leu Leu Ala Lys Lys Val Ala
145 150 155 160
Ser His Thr Pro Ser Leu Leu Arg Asp Val Phe His Thr Thr Val Asn
165 170 175
Phe Ile Asn Gln Asn Leu Arg Thr Tyr Val Arg Ser Leu Ala Arg Asn
180 185 190
Gly Met Asp
195

<210> SEQ ID NO 152
<211> LENGTH: 91
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: 85
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: 87
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 152
Met Ser Leu Ser Leu Val Ser Val Ser Val Val Gly Pro Ser Thr Leu Ala
1  5  10  15
Cys Ser Phe Leu Arg Pro Lys Ala Arg Pro Ser Lys Arg Ser Pro Arg
20  25  30
Asn Tyr Thr Asp Thr Ser Ser Gln Gly Pro Arg Ala Pro Arg Gly
35  40  45
Gly Ala Trp Arg Ser Leu Ser Gln Gln Asn Ser Ser Pro Lys Gln Val
50  55  60
Ala Val Ala Lys Ala Ser Tyr Arg Pro Val Leu Cys Phe Leu Pro Gly
65  70  75  80
Pro Trp Ser Ser Xaa Pro Xaa Ala Phe Leu Ile
85  90

<210> SEQ ID NO 153
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 153
Met Gly Thr Leu Ser Ala Glu Cys Ser Gly Pro Ala Thr Leu Gly Leu
1                  5                  10                 15
Cys Leu Val Val Pro Trp Asn Ser Ser Gly Leu Ser Gln Pro Pro
20                  25                 30

<210> SEQ ID NO 154
<211> LENGTH: 90
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 154
Met Lys Phe Leu Ala Val Leu Val Leu Gly Val Ser Ile Phe Leu
1                  5                  10                 15
Val Ser Ala Gin Asn Pro Thr Thr Ala Ala Pro Ala Asp Thr Tyr Pro
20                 25                 30
Ala Thr Gly Pro Ala Asp Asp Ala Pro Ala Gly Thr Thr Ala
35                40                45
Ala Ala Thr Thr Ala Thr Ala Ala Pro Thr Ala Ala Thr Thr Ala
50                55                60
Ala Ser Thr Thr Ala Arg Lys Asp Ile Pro Val Leu Pro Lys Trp Val
65                70                75                80
Gly Asp Leu Pro Asn Gly Arg Val Cys Pro
85                90

<210> SEQ ID NO 155
<211> LENGTH: 89
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 155
Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr Cys Ser Asn Thr
1                  5                  10                 15
Ser Pro Ser Tyr Gin Gly Thr Gin Gly Leu Leu Gly Leu Pro Ser Ala
20                 25                 30
Gln Trp Trp Pro Leu Thr Gly Arg Met Gin Cys Arg Leu Phe
35                40
Cys Phe Leu Leu Gin Asn Cys Leu Phe Pro Phe Pro Leu His Leu Ile
50                55                60
Gln His Asp Pro Cys Glu Leu Val Thr Ile Ser Trp Asp Trp Ala
65                70                75                80
Glu Ala Gly Ala Ser Leu Tyr Ser Pro
85

<210> SEQ ID NO 156
<211> LENGTH: 174
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 156
Met Ser Ser Ala Ala Ala Asp His Trp Ala Trp Leu Leu Val Leu Ser
1                  5                  10                 15
Phe Val Phe Gly Cys Asn Val Leu Arg Ile Leu Leu Pro Ser Phe Ser
20                 25                 30
Ser Phe Met Ser Arg Val Leu Gin Lys Asp Ala Glu Gin Glu Ser Gin
35                40                45
Met Arg Ala Glu Ile Gln Asp Met Lys Gln Glu Leu Ser Thr Val Asn
50  55  60
Met Met Asp Glu Phe Ala Arg Tyr Ala Arg Leu Glu Arg Lys Ile Asn
65    70  75   80
Lys Met Thr Asp Lys Leu Lys Thr His Val Lys Ala Arg Thr Ala Gin
85    90  95
Leu Ala Lys Ile Lys Trp Val Ile Ser Val Ala Phe Tyr Val Leu Gin
100  105 110
Ala Ala Leu Met Ile Ser Leu Ile Trp Lys Tyr Tyr Ser Val Pro Val
115  120 125
Ala Val Val Pro Ser Lys Trp Ile Thr Pro Leu Asp Arg Leu Val Ala
130  135 140
Phe Phe Thr Arg Val Ala Gly Gly Val Gly Ile Thr Cys Trp Ile Leu
145  150 155 160
Val Cys Asn Lys Val Val Ala Ile Val Leu His Pro Phe Ser
165  170

<210> SEQ ID NO: 157
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 157
Met Gly Lys Leu Ile Asn Ile Val Ile Arg Lys Pro Leu Leu Leu Leu
1   5  10  15
Leu Val Gln Cys Glu Asn Cys Cys Arg Lys Asn Met Leu Thr Tyr Asn Ile
20  25  30
Phe Leu Asn Ile His Asn Ile His Lys Phe Ser Asn His
35  40  45

<210> SEQ ID NO: 158
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 158
Met Val Ala Ser Thr Leu Val Thr Asn Leu Phe Gly Val Ala Phe Ala
1   5  10  15
Thr Thr Ala Ala Thr Arg Ala
20

<210> SEQ ID NO: 159
<211> LENGTH: 70
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (33)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<400> SEQUENCE: 159
Met Leu Met Ala Pro Val Val Cys Leu Ser Phe Ser Pro Cys Pro Ala
1   5  10  15
Asp Thr Ser Leu Thr Gly Asp Gly Leu Lys Ala Gly Leu Glu Arg Gly
20  25  30
Xaa Ala Leu Val Thr Leu Phe Asp Ser Val Thr His Phe Leu Ala His
35  40  45
Thr Leu Phe Glu Leu Leu Asp Phe Gln Leu Ala Phe Leu Arg Ser Gly 59 55 60
Lys Gln Thr Ala Pro His 65 70

<210> SEQ ID NO 160
<211> LENGTH: 323
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 160
Met Leu Leu Leu Leu Leu Leu Leu Gly Ser Gly Gln Gly Gln Gly Pro Gin Gln 1 5 10 15
Val Gly Ala Gly Gin Thr Phe Glu Tyr Leu Lys Arg Glu His Ser Leu 20 25 30
Ser Lys Pro Tyr Gin Gly Val Gly Thr Gin Gly Ser Ser Ser Leu Trp Asn 40 45
Leu Met Gly Asn Ala Met Val Met Thr Gin Tyr Ile Arg Leu Thr Pro 55 60
50
Asp Met Gin Ser Lys Gin Gly Ala Leu Trp Asn Arg Val Pro Cys Phe 80
65 70 75
Leu Arg Asp Trp Glu Leu Gin Val His Phe Lys Ile His Gly Gin Gly 90 95
85
Lys Lys Asn Leu His Gly Asp Gly Leu Ala Ile Trp Tyr Thr Arg Asn 110
105 110
Arg Met Gin Pro Gly Pro Val Phe Gly Asn Met Asp Lys Phe Val Gly 115 120 125
Leu Gly Val Phe Val Asp Thr Tyr Pro Asn Glu Lys Gin Gin Glu 130 135 140
Arg Val Phe Pro Tyr Ile Ser Ala Met Val Asn Gin Ser Gin Ser Leu Ser 145 150 155 160
Tyr Asp His Glu Arg Asp Gly Arg Pro Thr Glu Leu Gly Gin Gly Cys Thr 145 170 175
Ala Ile Val Arg Asn Leu His Tyr Asp Thr Phe Leu Val Ile Arg Tyr 180 185 190
Val Lys Arg His Leu Thr Ile Met Met Asp Ile Asp Gly Lys His Glu 195 200 205
Trp Arg Asp Cys Ile Glu Val Pro Gly Val Arg Leu Pro Arg Asp Arg Tyr 210 215 220
Tyr Phe Gly Thr Ser Ser Ile Thr Gly Asp Leu Ser Asp Asn His Asp 225 230 235 240
Val Ile Ser Leu Lys Leu Phe Leu Thr Val Glu Arg Thr Pro Glu 245 250 255
Glu Glu Lys Leu His Arg Asp Val Phe Leu Pro Ser Val Asp Asn Met 260 265 270
Lys Leu Pro Glu Met Thr Ala Pro Leu Pro Leu Ser Gly Leu Ala 275 280 285
Leu Phe Leu Ile Val Phe Phe Ser Leu Val Phe Ser Val Phe Ala Ile 290 295 300
Val Ile Gly Ile Ile Leu Tyr Asn Lys Trp Gin Glu Gin Ser Arg Lys 305 310 315 320
Arg Phe Tyr
<210> SEQ ID NO 161
<211> LENGTH: 120
<212> TYPE: PRO
<213> ORGANISM: Homo sapiens
<220> FEATURE: MISC_FEATURE
<222> LOCATIONS: (120)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 161

Met Pro Ser Glu Tyr Thr Tyr Val Lys Leu Arg Ser Asp Cys Ser Arg
  1  5 10  15

Pro Ser Leu Gln Trp Tyr Thr Arg Ala Gln Ser Lys Met Arg Arg Pro
  20 25 30

Ser Leu Leu Leu Lys Asp Ile Leu Lys Cys Thr Leu Leu Val Phe Gly
  35  40  45

Val Trp Ile Leu Tyr Ile Leu Lys Leu Asn Tyr Thr Thr Glu Glu Cys
  55  60

Asp Met Lys Lys Met His Tyr Val Asp Pro Asp His Val Lys Arg Ala
  65  70  75  80

Gln Lys Tyr Ala Gln Gln Val Leu Gln Lys Cys Arg Pro Lys Phe
  85  90

Ala Lys Thr Ser Met Ala Leu Leu Phe Glu His Arg Tyr Ser Val Asp
 100 105 110

Leu Leu Pro Phe Val Glu Lys Xaa Pro Lys Asp Ser Glu Ala Glu Ser
 115 120 125

Lys Tyr Asp Pro Pro Phe Gly Phe Arg Lys Phe Ser Ser Lys Val Gln
 130 135 140

Thr Leu Leu Glu Leu Leu Pro Glu His Asp Leu Pro Glu His Leu Lys
 145 150 155 160

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

His Gly Leu Glu Leu Gly His Thr Leu Asn Glu Phe Asp Val Val Ile
 180 185

Arg Leu Asn Ser Ala Pro Val Glu Gly Tyr Ser Glu His Val Gly Asn
 190 200 205

Lys Thr Thr Ile Arg Met Thr Tyr Pro Glu Gly Ala Pro Leu Ser Asp
 210 215 220

Leu Glu Tyr Ser Asn Asp Leu Phe Val Ala Val Leu Phe Lys Ser
 225 230 235 240

Val Asp Phe Asn Trp Leu Gln Ala Met Val Lys Glu Thr Leu Pro
 245 250 255

Phe Trp Val Arg Leu Phe Phe Trp Lys Gin Val Ala Glu Lys Ile Pro
 260 265 270

Leu Gin Pro Lys His Phe Arg Ile Leu Asn Pro Val Ile Ile Lys Glu
 275 280 285

Thr Ala Phe Xaa His Pro Ser Val Leu Arg Ala Ser Val Lys Val Leu
 290 295 300

Gly Ala Glu Ile Arg Thr Ser Pro Gin Ser Val Ser Leu Pro Leu Ser
<table>
<thead>
<tr>
<th>305</th>
<th>310</th>
<th>315</th>
<th>320</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Thr Leu Asp Val Gin Thr Val Val Val Phe Ala Val Ile Val Val</td>
<td>1</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Leu Leu Leu Val Asn Val Ile Leu Met Phe Phe Leu Gly Thr Arg</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>310</th>
<th>315</th>
<th>320</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Leu Pro Leu Leu Phe Cys Ala Phe Cys Leu His Lys Leu Gly Pro</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Leu Leu Phe Leu Tyr Asp Val Leu Met Xaa His Glu Ala Val Met Arg</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Thr His Gin Ile Gin Leu Pro Asp Pro Glu Phe Pro Ser Gin Gin Asn</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>Gin Val Leu Asn Lys Thr Leu Phe Asn Lys Leu Lys Lys Lys Lys Lys</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>Lys Lys Lys Xaa Xaa Xaa Lys Lys</td>
<td>65</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>320</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Ala Ser Arg Gly Arg Arg Pro Glu His Gly Gly Pro Pro Glu Leu</td>
</tr>
<tr>
<td>Phe Tyr Asp Glu Thr Glu Ala Arg Lys Tyr Val Arg Asn Ser Arg Met</td>
</tr>
<tr>
<td>Ile Asp Ile Gin Thr Arg Met Ala Gly Arg Ala Leu Glu Leu Leu Tyr</td>
</tr>
<tr>
<td>Leu Pro Glu Asn Lys Pro Cys Tyr Leu Leu Aap Ile Gly Cys Gly Thr</td>
</tr>
</tbody>
</table>
-continued

Gly Leu Ser Gly Ser Tyr Leu Ser Asp Gly Gly His Tyr Trp Val Gly
65 70 75 80

Leu Asp Ile Ser Pro Ala Met Leu Asp Gly Ala Val Asp Arg Glu Ile
85 90 95

Glu Gly Asp Leu Leu Gly Asp Met Gly Gln Gly Ile Pro Phe Lys
100 105 110

Pro Gly Thr Phe Asp Gly Cys Ile Ser Ile Ser Ala Val Gln Trp Leu
115 120 125

Cys Asn Ala Asn Lys Lys Ser Glu Asn Pro Ala Lys Arg Leu Tyr Cys
130 135 140

Phe Phe Ala Ser Leu Phe Ser Val Leu Val Arg Gly Ser Arg Ala Val
145 150 155 160

Leu Gln Leu Tyr Pro Glu Asn Ser Glu Leu Glu Leu Ile Thr Thr
165 170 175

Gln Ala Thr Lys Ala Gly Phe Ser Gly Gly Met Val Val Asp Tyr Pro
180 185 190

Asn Ser Ala Lys Ala Lys Phe Tyr Leu Cys Leu Phe Ser Gly Pro
195 200 205

Ser Thr Phe Ile Pro Glu Gly Leu Ser Glu Asn Gln Asp Glu Val Glu
210 215 220

Pro Arg Glu Ser Val Phe Thr Asn Gly Arg Phe Pro Leu Arg Met Ser
225 230 235 240

Arg Arg Gly Met Val Arg Lys Ser Arg Ala Trp Val Leu Gly Lys Lys
245 250 255

Glu Arg His Arg Arg Gln Gly Arg Glu Val Arg Pro Asp Thr Gln Tyr
260 265 270

Thr Gly Arg Lys Arg Lys Pro Arg Phe
275 280

<210> SEQ_ID NO 165
<211> LENGTH: 81
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 165

Met Glu Lys Ile Pro Glu Val Thr Asn Ser Asn Ser Ser Phe His Ala
1  5 10 15

His Asp Leu Gly Phe Cys Val Leu Ser Ile Ala Thr Ser Lys Ser Arg
20 25

Lys Ala Pro Ala Pro His Ala Gln Lys Cys Asn Leu Lys Ser Leu Arg
35 40

Ser Ser Ala Gin Thr Asp Ile Asn Pro Val Phe Ser Leu His Pro
55 55 60

Glu Pro Pro Gly Lys Ser Gly Ala Gin Thr Gin Ser Lys Ala Pro Phe
65 70 75 80

Leu

<210> SEQ_ID NO 166
<211> LENGTH: 327
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (308)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly
<400> SEQUENCE: 166

Met Trp Arg Pro Ser Val Leu Leu Leu Leu Leu Leu Arg His Gly
     1   5    10   15
Ala Gin Gly Lys Pro Ser Pro Asp Ala Gly Pro His Gly Gin Gly Arg
     20  25    30
Val His Gin Ala Ala Pro Leu Ser Asp Ala Pro His Asp Ala His
     35   40   45
Gly Asn Phe Gin Tyr Asp His Glu Ala Phe Leu Gly Arg Glu Val Ala
     50   55   60
Lys Glu Phe Asp Gin Leu Thr Pro Glu Gin Ala Arg Leu Gly
     65   70   75   80
Arg Ile Val Asp Arg Met Asp Arg Ala Gly Asp Gly Asp Gly Trp Val
     85   90   95
Ser Leu Ala Glu Leu Arg Ala Trp Ile Ala His Thr Gin Gin Arg His
    100  105   110
Ile Arg Asp Ser Val Ser Ala Ala Trp Asp Thr Tyr Asp Thr Asp Arg
    115  120  125
Asp Gly Arg Val Gly Trp Glu Gin Leu Arg Asn Ala Thr Tyr Gly His
    130  135   140
Tyr Ala Pro Gly Glu Phe His Tyr Val Gin Asp Ala Glu Thr Tyr
    145  150  155   160
Lys Lys Met Leu Ala Arg Asp Gin Arg Phe Arg Val Ala Asp Gin
    165   170   175
Asp Gly Asp Ser Met Ala Thr Arg Glu Leu Thr Ala Phe Leu His
    180  185   190
Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    195  200   205
Glu Asp Leu Gin Gin Asp Gin Tyr Gin Val Gin Val Glu Gin Tyr
    210  215   220
Ile Ala Asp Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    225  230  235   240
Val Gin Thr Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    245  250   255
Asp Gly Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    260  265   270
Asp Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    275  280   285
Asp Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    290  295   300
Trp Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    305  310  315   320
Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
    325

<210> SEQ ID NO 167
<211> LENGTH: 65
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 167

Met Ile Lys Ile Leu Lys Glu Ala Ile Glu Glu Thr Ser Phe Cys Ser
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Phe</td>
<td>Trp</td>
<td>Arg</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Cys</td>
<td>Ala</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ile</td>
<td>Ser</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 168
<211> LENGTH: 159
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 168

Met Trp Leu Phe Ile Leu Leu Ser Leu Ala Leu Leu Ser Asp Ala Met
1    5    10   15

Val Met Asp Glu Lys Val Lys Arg Ser Phe Val Leu Asp Thr Ala Ser
20   25   30   35

Ala Ile Cys Asn Tyr Asn Ala His Tyr Lys Asn His Pro Lys Tyr Trp
35   40   45   50

Cys Arg Gly Tyr Phe Arg Asp Tyr Cys Asn Ile Ala Phe Ser Pro
55   60

Asn Ser Thr Asn His Val Ala Leu Lys Asp Thr Gly Asn Gin Leu Ile
65   70   75   80

Val Thr Met Ser Cys Leu Asn Lys Glu Asp Thr Gly Trp Tyr Trp Cys
85   90   95

Gly Ile Gin Arg Asp Phe Ala Arg Asp Met Asp Phe Thr Glu Leu
100  105  110

Ile Val Thr Asp Asp Lys Gly Thr Trp Pro Met Thr Leu Val Trp Glu
115  120  125

Arg Leu Ser Gly Thr Lys Pro Glu Ala Ala Arg Leu Pro Lys Leu Ser
130  135  140

Ala Arg Leu Thr Ala Pro Gly Arg Pro Ser Ser Phe Ala Tyr
145  150  155
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (101)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 169

Met Ala Xaa His Phe Leu Leu Val Ala Leu Gln Ser Val Pro His Cys
1     5      10  15

Pro His Leu Leu Glu Glu Glu His Lys Leu Cys Lys Val Ser His Phe
20    25     30

Ser Gly Val Thr Leu Val Thr Ser Arg Gin Asp Ser Ser Tyr Val
35    40     45

Pro Val Gin Thr Leu Phe Ile His Leu Gly Pro Trp Ala Trp Asp Leu
50    55     60

Xaa Pro Cys Thr Ala Glu Asp Pro Glu Ala Glu Arg Ser Leu Arg Leu
65    70     75  80

Cys His Ser His Leu Ala Arg Xaa Asn Val Ser Pro Ser Gin Ala Ala
85    90     95

Glu Gly Xaa Xaa Xaa Arg Gly Cys Gin His Arg Gly Ser Arg Glu Leu
100   105   110

Thr Phe Leu Ser Ala Glu Asn Glu Ala Gly Ile
115   120

<410> SEQ ID NO: 170
<411> LENGTH: 129
<412> TYPE: PRT
<413> ORGANISM: Homo sapiens

<400> SEQUENCE: 170

Met Lys Val Gin Ala Arg Ile Arg Val Lys Met Ser Val Asn Lys Ala
1     5      10  15

His Pro Val Val Ser Thr His Trp Arg Trp Pro Ala Glu Trp Pro Gin
20    25     30

Met Phe Leu His Leu Ala Gin Gin Pro Arg Thr Glu Val Lys Ser Arg
35    40     45

Pro Leu Gin Leu Ala Gly Phe Ile Gin Gin Gin Gin Gin Gin Gin Gin Gin
50    55     60

Pro Leu Gin Gin Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
65    70     75  80

Pro Tyr Gin His Gin Gin Arg Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
85    90     95

Tyr Leu Gin Tyr Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
100   105   110

Gly His Thr His Thr Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
115   120   125

Ile

<210> SEQ ID NO: 171
<211> LENGTH: 372
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 171

Met Ala Tyr His Ser Phe Leu Val Glu Pro Ile Ser Cys His Ala Trp
---continued---

1  5  10  15

Asn Lys Asp Arg Thr Gln Ile Ala Ile Cys Pro Asn Asn His Glu Val
20  25  30

His Ile Tyr Glu Lys Ser Gly Ala Trp Thr Lys Val Val His Glu Leu
35  40  45

Lys Glu His Asn Gly Gln Val Thr Gly Ile Asp Trp Ala Pro Glu Ser
50  55  60

Asn Arg Ile Val Thr Cys Gly Thr Asp Arg Asn Ala Tyr Val Trp Thr
65  70  75  80

Leu Lys Gly Arg Thr Trp Lys Pro Thr Leu Val Ile Leu Arg Ile Asn
85  90  95

Arg Ala Ala Arg Cys Val Arg Trp Ala Pro Asn Glu Asn Lys Phe Ala
100 105 110

Val Gly Ser Gly Ser Arg Val Ile Ser Ile Cys Tyr Phe Glu Gln Glu
115 120 125

Asn Asp Trp Trp Val Cys Lys His Ile Lys Lys Pro Ile Arg Ser Thr
130 135 140

Val Leu Ser Leu Asp Trp His Pro Asn Asn Val Leu Leu Ala Ala Gly
145 150 155 160

Ser Cys Asp Phe Lys Cys Arg Ile Phe Ser Ala Tyr Ile Lys Glu Val
165 170 175

Glu Glu Arg Pro Ala Pro Thr Pro Trp Gly Ser Lys Met Pro Phe Gly
180 185 190

Glu Leu Met Phe Glu Ser Ser Ser Cys Gly Trp Val His Gly Val
195 200 205

Cys Phe Ser Ala Ser Gly Ser Arg Val Ala Trp Val Ser His Asp Ser
210 215 220

Thr Val Cys Leu Ala Asp Ala Asp Lys Met Ala Val Ala Thr Leu
225 230 235 240

 Ala Ser Glu Thr Leu Pro Leu Leu Ala Leu Thr Phe Ile Thr Asp Asn
245 250 255

Ser Leu Val Ala Ala Gly His Asp Cys Phe Pro Val Leu Phe Thr Tyr
260 265 270

Asp Ala Ala Ala Gly Met Leu Ser Phe Gly Gly Arg Leu Asp Val Pro
275 280 285

Lys Gln Ser Ser Gln Arg Gly Leu Thr Ala Arg Glu Arg Phe Glu Asn
290 295 300

Leu Asp Lys Lys Ala Ser Ser Glu Gly Gly Thr Ala Ala Gly Ala Gly
305 310 315 320

Leu Asp Ser Leu His Lys Asn Ser Val Ser Gln Ile Ser Val Leu Ser
325 330 335

Gly Gly Lys Ala Lys Cys Ser Gin Phe Cys Thr Thr Gly Met Asp Gly
340 345 350

Gly Met Ser Ile Trp Asp Val Lys Ser Leu Glu Ser Ala Leu Lys Asp
355 360 365

Leu Lys Ile Lys
370

<210> SEQ ID NO 172
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 172
Met Trp Ser Ile Gly Ala Gly Ala Leu Gly Ala Ala Ala Ala Leu Ala Leu
1 5 10 15
Leu Leu Ala Asn Thr Asp Val Phe Leu Ser Lys Pro Gln Lys Ala Ala
20 25 30
Leu Glu Tyr Leu Glu Asp Ile Asp Leu Lys Thr Leu Glu Lys Glu Pro
35 40 45
Arg Thr Phe Lys Ala Lys Glu Leu Trp Glu Lys Asn Gly Ala Val Ile
50 55 60
Met Ala Val Arg Arg Pro Gly Cys Phe Leu Cys Arg Glu Ala Ala Ala
65 70 75 80
Asp Leu Ser Ser Leu Lys Ser Met Leu Asp Glu Leu Gly Val Pro Leu
85 90 95
Tyr Ala Val Val Lys Glu His Ile Arg Thr Glu Val Lys Asp Phe Gln
100 105 110
Pro Tyr Phe Lys Gly Glu Ile Phe Leu Asp Glu Lys Lys Phe Tyr
115 120 125
Gly Pro Gln Arg Arg Lys Met Met Phe Met Gly Phe Ile Arg Leu Gly
130 135 140
Val Trp Tyr Asn Phe Phe Arg Ala Trp Asn Gly Glu Gly Phe Ser Gly Asn
145 150 155 160
Leu Glu Gly Glu Phe Ile Leu Gly Gly Val Phe Val Val Gly Ser
165 170 175
Gly Lys Gln Gly Ile Leu Leu Glu His Arg Glu Gly Phe Gly Asp
180 185 190
Lys Val Asn Leu Leu Ser Val Leu Glu Ala Ala Lys Met Ile Lys Pro
195 200 205
Gln Thr Leu Ala Ser Glu Lys Lys
210 215

<210> SEQ ID NO 173
<211> LENGTH: 55
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 173
Met Lys Pro Val Ser Arg Arg Thr Leu Asp Trp Ile Tyr Ser Val Leu
1 5 10 15
Leu Leu Ala Ile Val Leu Ile Ser Trp Gly Cys Ile Ile Tyr Ala Ser
20 25 30
Met Val Ser Ala Arg Arg Glu Leu Arg Lys Lys Tyr Pro Asp Lys Ile
35 40 45
Phe Gly Thr Asn Glu Asn Leu
50 55

<210> SEQ ID NO 174
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: Misc Feature
<222> LOCATION: (19)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly
found in naturally occurring proteins
Met Ala Ala Asn Thr Phe Val Leu Ile Met Gly Ile Pro Thr Ser Ala
  1    5  10  15
Asn Ala Xaa Arg Asp Leu Phe
  20

<400> SEQ ID NO: 175
<110> LENGTH: 193
<120> TYPE: PRT
<210> ORGANISM: Homo sapiens
<400> SEQUENCE: 175
Met Ser Ile Cys His Arg Gly Thr Gly Ile Ala Leu Ser Ala Gly Val
  1    5  10  15
Ser Leu Phe Gly Met Ser Ala Leu Leu Leu Pro Gly Asn Phe Glu Ser
  20   25  30
Tyr Leu Glu Leu Val Lys Ser Leu Cys Leu Gly Pro Ala Leu Ile His
  35   40  45
Thr Ala Lys Phe Ala Leu Val Phe Pro Leu Met Tyr His Thr Trp Asn
  50   55  60
Gly Ile Arg His Leu Met Trp Asp Leu Gly Lys Gly Leu Lys Ile Pro
  65   70  75  80
Gln Leu Tyr Gln Ser Gly Val Val Val Leu Val Val Thr Val Leu Ser
  85   90  95
Ser Met Gly Leu Ala Ala Met
 100

<410> SEQ ID NO: 176
<110> LENGTH: 48
<120> TYPE: PRT
<210> ORGANISM: Homo sapiens
<400> SEQUENCE: 176
Met Thr Lys Ala Ser Ser Leu Trp Pro Leu Lys Thr Thr Cys Gln Ile
  1    5  10  15
Ser Gly Thr Val Phe Phe Phe Leu Phe Phe Leu Ser Cys Phe Leu Met
  20   25  30
Gln Ala Gln Cys Asp Lys Phe Val Gly Trp Asp Phe Phe Phe Phe Leu
  35   40  45

<210> SEQ ID NO: 177
<110> LENGTH: 96
<120> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (18)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<400> SEQUENCE: 177
Met Arg Arg Ala Leu Ile Pro Pro Cys Arg Gly Gly Pro Ser Ala Ser
  1    5  10  15
Asp Xaa Cys Ser Cys Ser Ser Pro Ser Gly Phe Ser Ala Gly Arg Gly
  20   25  30
Arg Cys Pro Val Gln Gly Cys Leu Arg Pro His Arg Val Gin Leu Leu
  35   40  45
Arg Arg Trp Gly Pro Gly Ser Pro Ala Gly Gln Arg Leu Ser Lys Gly
50  55  60
Arg Lys Gly Pro Phe Pro Pro Pro Asp Pro Pro Trp Pro Val Thr Leu
85  90  95

<210> SEQ ID NO 178
<211> LENGTH: 95
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (78)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 178
Met Leu Glu Thr Thr Lys His Val Gln Ile Ala Cys Met Leu Leu Leu
1   5   10   15
Thr Cys Gin Ile Phe Leu Pro Ser Ser Leu Ser Pro Ser Phe Ile His
20  25   30
Ser Leu Thr Asp Ser Phe Ile Pro Leu Lys Leu Tyr Val Cys Phe
35  40   45
Val Gin Ser Thr Leu Leu Lys Ala Ala Gly Tyr Lys Ser Ile Ser Glu
50  55   60
Ala Leu Gly Phe Asp Xaa Leu Leu Cys Ser Ser Ser Ala Arg Phe Val Trp
65  70   75   80
Ile Cys His Thr Tyr Ser Arg Pro Leu Val Thr Cys Ala Leu His
85  90   95

<210> SEQ ID NO 179
<211> LENGTH: 27
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 179

Met Ser Val Ile Gly Gly Leu Leu Val Val Ala Gly Pro Gly
1   5   10   15
Gly Val Ser Met Asp Glu Lys Lys Gly Trp
20  25

<210> SEQ ID NO 180
<211> LENGTH: 99
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (12)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (13)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (72)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 180
Met Ser Gly Gly Leu Ser Phe Leu Leu Leu Val Xaa Xaa Gly Thr Gin
<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser Pro Leu His Leu Ala Gly Ser Cys Pro Gly Gln Thr His Leu Ser</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Phe Pro Leu Gly Gln Asp Arg Gly Gln Leu Gln Gln Lys Gin Gin</td>
<td>35</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>Asp Leu Gln Gin Gin Gly Leu Gin Gly Leu Leu Ala Gly</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Glu Trp Ala Pro Pro Leu Trp Xaa Leu Gly Ser Leu Phe Gin Ala Phe</td>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Val Lys Arg Glu Ser Gin Gin Ala Tyr Ala</td>
<td>85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-<210> SEQ ID NO 191
-<211> LENGTH: 65
-<212> TYPE: PRT
-<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 181

Met Phe Ala Asp Phe Ile Val Val Thr Ala Thr Val Gin Arg Cys Pro
1  5  10  15
Gly Ser Pro Pro Leu Ser Gin Leu Gin Gin Gin Gin Gin Gin Gin Gin
20 | 25 | 30 |
Ile Ser Ser His Ala Gly Leu Pro Trp Leu Ser Ser Trp Pro Ala Pro
35 | 40 | 45 |
Pro Trp Thr Trp Ser Trp Ile Ser Arg Arg Arg Gin Gin Gin Gin Gin
50 | 55 | 60 |
Ser
65

-<210> SEQ ID NO 182
-<211> LENGTH: 105
-<212> TYPE: PRT
-<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 182

Met Ser Ala Leu Thr Arg Leu Ala Ser Phe Ala Arg Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Met Asp Val Leu Phe Val Ala Ile Phe Ala Val Pro Leu Ile Leu Gly
1 5 10 15

Gln Glu Tyr Glu Asp Glu Glu Arg Leu Gly Glu Asp Glu Tyr Tyr Glu
20 25 30

Val Val Tyr Tyr Thr Val Thr Pro Ser Tyr Asp Phe Ser Ala
35 40 45

Asp Phe Thr Ile Asp Tyr Ser Ile Phe Ser Glu Ser Glu Asp Arg Leu Asn
50 55 60

Arg Leu Asp Lys Asp Ile Thr Glu Ala Ile Glu Thr Thr Ile Ser Leu
65 70 75 80

Glu Thr Ala Arg Ala Asp His Pro Lys Pro Val Thr Val Lys Pro Val
85 90 95

Thr Thr Glu Pro Gln Ser Pro Asp Leu Asn Asp Ala Val Ser Ser Leu
100 105 110

Arg Ser Pro Ile Pro Leu Leu Leu Ser Cys Ala Phe Val Gln Val Gly
115 120 125

Met Tyr Phe Met
130

<210> SEQ ID NO 184
<211> LENGTH: 69
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 184

Met Pro Cys Gln Pro Gly Gln Val Pro Ser Cys Gln Cys Thr Phe Gly
1 5 10 15

Leu Leu Leu Met Leu Pro Ser Leu Pro Ser Pro Ala Ser Gln Pro Arg
20 25 30

Pro Phe Cys Ser Ser Met Glu Tyr Phe His Gly Cys Ala Ser Pro Ser
35 40 45

Gln Ala Ile Ile Gly Gly Phe Pro Phe Ala Ser Val Ala Leu Ala Asp
50 55 60

Ile Leu Cys Leu Gln
65

<210> SEQ ID NO 185
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 185

Met Ser Leu Leu Ser Pro Ala Ile Pro Ala Leu Thr Leu Ile Phe Ile
1 5 10 15

Leu Met Phe Phe Ser Phe Pro Phe Arg Ala His Thr Val Thr Ile
20 25 30

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

<210> SEQ ID NO 186
<211> LENGTH: 65
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 186
Met Ala Phe Gly Leu Gln Met Phe Ile Gln Arg Lys Phe Pro Tyr Pro
  1    5    10    15
Leu Gln Trp Ser Leu Leu Val Ala Val Val Val Ala Gly Ser Val Val Ser
  20    25    30
Tyr Gly Val Thr Arg Val Glu Ser Glu Lys Cys Asn Asn Leu Trp Leu
  35    40    45
Phe Leu Glu Thr Gly Glu Leu Pro Lys Asp Arg Ser Thr Asp Gln Arg
  59    55    60
Ser
  65

<210> SEQ ID NO 187
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 187
Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly Leu Met Leu Lys
  1    5    10    15
Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe
  20    25    30
Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe
  35    40    45
Met

<210> SEQ ID NO 188
<211> LENGTH: 170
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 188
Met Leu Leu Asn Val Ala Leu Val Ala Leu Leu Leu Leu Gly Ala Tyr
  1    5    10    15
Arg Leu Trp Val Arg Trp Gly Arg Arg Gly Leu Gly Ala Gly Ala Gly
  20    25    30
Ala Gly Glu Glu Ser Pro Ala Thr Ser Leu Pro Arg Met Lys Lys Arg
  35    40    45
Asp Phe Ser Leu Glu Gin Leu Arg Gin Tyr Asp Glu Ser Ser Arg Asn Pro
  50    55    60
Arg Ile Leu Leu Ala Val Asn Gly Lys Val Phe Asp Val Thr Lys Gly
  65    70    75    80
Ser Lys Phe Tyr Gly Pro Ala Gly Pro Tyr Gly Ile Phe Ala Gly Arg
  85    90    95
Asp Ala Ser Arg Gly Leu Ala Thr Phe Cys Leu Asp Lys Asp Ala Leu
 100   105   110
Arg Asp Glu Tyr Asp Asp Leu Ser Asp Leu Asn Ala Val Gin Met Glu
 115   120   125
Ser Val Arg Glu Trp Glu Met Gin Phe Lys Glu Tyr Asp Tyr Val
 130   135   140
Gly Arg Leu Leu Lys Pro Gly Glu Glu Pro Ser Glu Tyr Thr Asp Glu
 145   150   155   160
Glu Asp Thr Lys Asp His Asn Lys Gin Asp
 165   170
<210> SEQ ID NO: 189
<211> LENGTH: 132
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 189

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

<210> SEQ ID NO: 190
<211> LENGTH: 92
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 190

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

<210> SEQ ID NO: 191
<211> LENGTH: 176
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURES:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (137)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly
found in naturally occurring proteins

<400> SEQUENCE: 191

Met Arg Gly Ser His Leu Arg Leu Leu Pro Tyr Leu Val Ala Ala Asn
1    5    10   15
---continued

<table>
<thead>
<tr>
<th>Pro Val Asn Tyr Gly Arg Pro Tyr Arg Leu Ser Cys Val Glu Ala Phe</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Ala Thr Phe Cys Ile Val Gly Phe Pro Asp Leu Ala Val Ile Leu</td>
<td>35</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>Leu Arg Lys Phe Lys Trp Gly Lys Gly Phe Leu Asp Leu Asn Arg Gin</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Leu Leu Asp Lys Tyr Ala Ala Cys Gly Ser Pro Glu Val Leu Gin</td>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Ala Glu Gin Glu Phe Leu Ala Asn Ala Lys Gin Ser Pro Gin Glu Glu</td>
<td>85</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Glu Ile Asp Pro Phe Asp Val Ser Gly Arg Glu Phe Gly Asn Pro</td>
<td>100</td>
<td>105</td>
<td>110</td>
</tr>
<tr>
<td>Asn Arg Pro Val Ala Ser Thr Arg Leu Pro Ser Thr Asp Asp Ser</td>
<td>115</td>
<td>120</td>
<td>125</td>
</tr>
<tr>
<td>Asp Ala Ser Glu Asp Pro Gly Pro Xaa Ala Glu Arg Gly Glu Ala Ser</td>
<td>130</td>
<td>135</td>
<td>140</td>
</tr>
<tr>
<td>Ser Ser Cys Gly Glu Glu Glu Thr Gin Gly Arg Gly Ala Glu Ala</td>
<td>145</td>
<td>150</td>
<td>155</td>
</tr>
<tr>
<td>Arg Ala Pro Ala Glu Val Trp Lys Gly Ile Lys Lys Arg Gin Gin Gin Gin</td>
<td>165</td>
<td>170</td>
<td>175</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 192
<211> LENGTH: 70
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 192

<table>
<thead>
<tr>
<th>Met Ser Asn Ala Cys Lys Glu Leu Ala Ile Phe Leu Thr Thr Gly Ile</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Val Ser Ala Phe Gly Leu Pro Ile Val Phe Ala Arg Ala His Leu</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Ile Glu Trp Gly Ala Cys Ala Leu Val Leu Thr Gly Asn Thr Val Ile</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Phe Ala Thr Ile Leu Gly Phe Leu Val Phe Gly Ser Asn Asp Asp</td>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Phe Ser Trp Gin Gin Trp</td>
<td>65</td>
<td>70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 193
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
| <221> NAME/KEY: MISC_FEATURE |
| <222> LOCATION: (31) |
| <223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins |

<220> FEATURE:
| <221> NAME/KEY: MISC_FEATURE |
| <222> LOCATION: (35) |
| <223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins |

<400> SEQUENCE: 193

<table>
<thead>
<tr>
<th>Met Thr Leu Leu Ile Ile Phe Leu Pro Phe Xaa Phe Thr Thr Xaa Thr</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn Ser Gly Gly Ser Phe Pro Val Arg</td>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 194
<211> LENGTH: 73
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (21)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 194

Met Lys Gly Leu Leu Pro Phe Leu Phe Leu Thr Val Trp Leu Trp
1    5   10   15
Leu Tyr Lys Leu Xaa Phe Gly Glu Ser Pro Arg Tyr Pro Asn Val Ile
20   25  30
Gly Lys Thr Tyr Phe Phe Phe Trp Thr Asp Gln Ile Ser Arg Glu Ser
35   40  45
Arg Phe Leu Glu Arg Leu Ala Phe Ile Val Ser Glu Asn Cys Leu Ile
50   55  60
Phe Leu Ile His Ala Ile Thr Gly Gln
65   70

<210> SEQ ID NO 195
<211> LENGTH: 289
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 195

Met Ser Gly Phe Ser Thr Glu Glu Arg Ala Ala Pro Phe Ser Leu Glu
1    5   10   15
Tyr Arg Val Phe Leu Lys Asn Glu Lys Gly Gln Tyr Ile Ser Pro Phe
20   25  30
His Asp Ile Pro Ile Tyr Ala Asp Lys Asp Val Phe His Met Val Val
35   40  45
Glu Val Pro Arg Trp Ser Asn Ala Lys Met Glu Ile Ala Thr Lys Asp
50   55  60
Pro Leu Asn Pro Ile Lys Glu Asp Val Lys Lys Gly Leu Arg Tyr
65   70  75  80
Val Ala Asn Leu Phe Pro Tyr Lys Gly Tyr Ile Trp Asn Tyr Gly Ala
85  90  95
Ile Pro Gin Thr Trp Glu Asp Pro Gin Gly His Asn Asp Lys His Thr Gly
100 105 110
Cys Cys Gly Asp Asn Asp Pro Ile Asp Val Cys Glu Ile Gly Ser Lys
115 120 125
Val Cys Ala Arg Gly Glu Ile Gly Val Lys Val Leu Gly Ile Leu
130 135 140
Ala Met Ile Asp Glu Gly Thr Asp Trp Lys Val Ile Ala Ile Asn
145 150 155 160
Val Asp Asp Pro Ala Ala Asn Tyr Asn Asp Ala Asn Val Asp Lys
165 170 175
Arg Leu Lys Pro Gly Tyr Leu Glu Ala Thr Val Asp Trp Phe Arg Arg
180 185 190
Tyr Lys Val Pro Asp Gly Lys Pro Gly Asn Glu Phe Ala Phe Asn Ala
195 200 205
Glu Phe Lys Asp Asp Phe Ala Ile Asp Ile Lys Ser Thr His
Aap His Trp Lys Ala Leu Val Thr Lys Lys Thr Asn Gly Lys Gly Ile
225 230 235 240
Ser Cys Met Asn Thr Thr Leu Ser Glu Ser Pro Phe Lys Cys Asp Pro
245 250 255
Asp Ala Ala Arg Ala Ile Val Asp Ala Leu Pro Pro Pro Cys Glu Ser
260 265 270
Ala Cys Thr Val Pro Thr Asp Val Asp Lys Trp Phe His His Gin Lys
275 280 285
Asc

<210> SEQ ID NO 196
<211> LENGTH: 624
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 196
Met Glu Ile Pro Gly Ser Leu Cys Lys Lys Val Lys Leu Ser Asn Asc
1 5 10 15
Ala Gin Asn Trp Gly Met Gln Arg Ala Thr Asn Val Thr Tyr Gin Ala
20 25 30
His His Val Ser Arg Asc Lys Arg Gin Gin Val Val Gin Thr Arg Gly
35 40 45
Gly Phe Arg Gly Cys Thr Val Trp Leu Thr Lys Leu Ser Gly Ala Gly
50 55 60
Lys Thr Thr Val Ser Met Ala Leu Glu Tyr Leu Val Cys His Gly
65 70 75 80
Ile Pro Cys Tyr Thr Leu Asp Gly Asp Asc Ile Arg Gin Gly Leu Asc
85 90 95
Lys Asc Leu Gin Phe Ser Pro Glu Asp Arg Glu Asc Val Arg Asc
100 105 110
Ile Ala Gin Val Ala Lys Leu Phe Ala Asc Ala Gin Leu Val Cys Ile
115 120 125
Thr Ser Phe Ile Ser Pro Tyr Thr Gin Asc Arg Asc Asc Ala Asc Gin
130 135 140
Ile His Glu Gin Ala Ser Leu Pro Phe Gin Val Gin Thr Gin Val Gin
145 150 155 160
Pro Leu His Cys Glu Gin Asc Val Lys Gin Lys Gin Thr Gin Lys
165 170 175
Ala Asc Ala Gin Lys Gin Gin Phe Gin Gin Gin Gin Gin Gin Gin Gin
180 185 190
Glu Gin Pro Asc Gin Gin Gin Gin Gin Gin Gin Lys Gin Gin Asc Asc Asc
195 200 205
Val Asc Asp Cys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
210 215 220
Ile Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
225 230 235 240
Glu Asc Lys Leu His Leu Ala Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
245 250 255
Leu Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
260 265 270
Gly Gin Asc Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
275 280 285
Thr Pro Lys Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
290 295 300
Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
305 310 315
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
320 325 330
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
335 340 345
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
350 355 360
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
365 370 375
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
380 385 390
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
395 400 405
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
410 415 420
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
425 430 435
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
440 445 450
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
455 460 465
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
470 475 480
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
485 490 495
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
500 505 510
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
515 520 525
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
530 535 540
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
545 550 555
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
560 565 570
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
575 580 585
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
590 595 600
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
605 610 615
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
620 625 630
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>275</td>
<td>Gln</td>
<td>Cys</td>
<td>Leu</td>
</tr>
<tr>
<td>280</td>
<td>His</td>
<td>Phe</td>
<td>Asp</td>
</tr>
<tr>
<td>285</td>
<td>Cys</td>
<td>Leu</td>
<td>Asp</td>
</tr>
</tbody>
</table>

Gly Gly Val Ile Asn Leu

Ser Val Pro Ile Val Leu Thr Ala Thr His Glu Asp Gly Glu Arg Leu

Asp Gly Cys Thr Ala Phe Ala Leu Met Tyr Glu Gly Arg Arg Val Ala

Ile Leu Arg Asn Pro Glu Phe Phe Glu His Arg Gly Glu Arg Cys

Ala Arg Gin Trp Gly Thr Thr Cys Lys Asn His Pro Tyr Ile Lys Met

Val Met Glu Gin Gly Asp Trp Leu Ile Gly Gly Asp Leu Glu Val Leu

Asp Arg Val Tyr Trp Asn Asp Gly Leu Asp Gin Tyr Arg Leu Thr Pro

Thr Glu Leu Lys Gin Lys Phe Lys Asp Met Asn Ala Asp Ala Val Phe

Ala Phe Gin Leu Arg Asn Pro Val His Asn Gly His Ala Leu Leu Met

Gln Asp Thr His Lys Gin Leu Leu Glu Arg Gly Tyr Arg Arg Pro Val

Leu Leu Leu His Pro Leu Gly Glu Trp Thr Lys Asp Asp Asp Val Pro

Leu Met Trp Arg Met Lys Gin His Ala Ala Val Leu Glu Gly Val

Leu Asn Pro Glu Thr Thr Val Val Ala Ile Phe Pro Ser Pro Met Met

Tyr Ala Gly Pro Thr Glu Val Gin Trp His Cys Arg Ala Arg Met Val

Ala Gly Ala Asn Phe Tyr Ile Val Val Gly Arg Asp Pro Ala Gly Met Pro

His Pro Glu Thr Gly Lys Asp Leu Tyr Glu Pro Ser His Gly Ala Lys

Val Leu Thr Met Ala Pro Gly Leu Ile Thr Leu Glu Ile Val Pro Phe

Arg Val Ala Ala Tyr Asn Lys Lys Lys Arg Met Asp Tyr Tyr Asp

Ser Glu His His Glu Asp Phe Glu Phe Ile Ser Gly Thr Arg Met Arg

Lys Leu Ala Arg Glu Gly Gin Lys Pro Pro Gly Gly Phe Met Ala Pro

Lys Ala Thr Thr Val Leu Thr Glu Tyr Tyr Lys Ser Leu Glu Lys Ala

610 |   |   |   |

615 |   |   |   |

620 |   |   |   |

<210> SEQ ID NO 197
<211> LENGTH: 649
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (555)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
-continued

<222> LOCATION: (557)
<223> OTHER INFORMATION: \( \alpha \) equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC FEATURE
<222> LOCATION: (558)
<223> OTHER INFORMATION: \( \alpha \) equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 197

Met Ser Ala Ser Gin Asp Leu Glu Pro Lys Pro Leu Phe Pro Lys Pro
1   5    10  15

Ala Phe Gly Gin Lys Pro Pro Leu Ser Thr Glu Asn Ser His Glu Asp
20  25   30

Glu Ser Pro Met Lys Asn Val Ser Ser Ser Lys Gly Ser Pro Ala Pro
35  40   45

Leu Val Arg Ser Lys Ser Gin Pro Leu Pro Ala Arg Glu Asp
55  55   60

Ser Glu Asn Lys Asp His Ala Gly Glu Ile Ser Ser Leu Pro Phe Pro
65  70   75  80

Gly Val Val Leu Lys Pro Ala Ala Ser Arg Gly Gin Gly Pro Gly Leu Ser
95  90   95

Lys Asn Gly Gin Gin Lys Lys Gin Gin Asp Asp Ala Ala Lys
100 105  110

Asn Thr Phe Gin Ser Lys Ile Asn Gin Glu Leu Ala Ser Gly Thr
115 120  125

Pro Pro Ala Arg Phe Pro Lys Ala Pro Ser Lys Leu Thr Val Gly Gin
139 135  140

Pro Trp Gin Gin Ser Gin Glu Lys Gin Gin Gin Gin Gin Gin Gin
145 150  155  160

Thr Pro Lys Gin Gin Gin Gin Gin Gin Pro Thr Gin Pro Pro Gin Gin Gin
165 170  175

Pro Pro Lys Pro Gin Gin Gin Gin Gin Gin Pro Gin Gin Gin Gin Gin Gin
180 185  190

Lys Thr Ser Ser Gin Ser Thr Ser Lys Gin Thr Ser Tyr Ser
195 200  205

Thr Thr Ser Leu Pro Pro Pro Pro Pro Ser His Pro Ala Ser Gin Pro
210 215  220

Pro Leu Pro Ala Ser His Pro Ser Gin Pro Pro Val Pro Ser Leu Pro
225 230  235  240

Pro Arg Gin Ile Lys Pro Pro Gin Gin Pro Gin Gin Gin Gin Gin Gin Gin
245 250  255

Asp Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
260 265  270

Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
275 280  285

Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
295  300

Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
305  310  315  320

Lys Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
325  330  335

Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
340  345  350
Gln Gly Glu Glu Ile Glu Ile Ile Arg Ile Thr Asp Asn Pro Glu Gly 355 360 365
Lys Trp Leu Gly Arg Thr Ala Arg Gly Ser Tyr Gly Tyr Ile Lys Thr 370 375 380
Thr Ala Val Glu Ile Asp Tyr Asp Leu Lys Leu Lys Lys Asp Ser 385 390 395 400
Leu Gly Ala Pro Ser Arg Pro Ile Glu Asp Asp Gln Glu Val Tyr Asp 405 410 415
Asp Val Ala Glu Asp Asp Ser Ser His Ser Gln Ser Gly Ser 420 425 430 435
Gly Gly Ile Phe Pro Pro Pro Asp Asp Ile Tyr Asp Gly Ile 435 440 445
Glu Glu Glu Asp Ala Asp Gly Ser Thr Leu Gln Val Gln Glu Lys 450 455 460 465
Ser Asn Thr Trp Ser Trp Gly Ile Leu Lys Met Leu Lys Gly Lys Asp 470 475 480
Asp Arg Lys Ser Ile Arg Gly Lys Pro Lys Val Ser Asp Ser Asp 485 490 495 499
Asn Asn Glu Gly Ser Ser Phe Pro Ala Pro Pro Lys Gln Leu Asp Met 500 505 510 515
Gly Asp Glu Val Tyr Asp Val Asp Thr Ser Asp Phe Pro Val Ser 515 520 525
Ser Ala Glu Met Ser Gln Gly Thr Asn Val Gly Lys Ala Lys Thr Glu 530 535 540 545
Glu Lys Asp Leu Lys Leu Lys Lys Gln Xaa Lys Xaa Xaa Lys Asp 550 555 560 565
Phe Arg Lys Lys Phe Lys Tyr Asp Gly Glu Ile Arg Val Leu Tyr Ser 570 575 580 585
Thr Lys Val Thr Thr Ser Ile Thr Ser Lys Trp Gly Thr Arg Asp 590 595 600 605
Leu Gln Val Lys Pro Gly Glu Ser Leu Gln Val Ile Gln Thr Thr Asp 610 615 620 625
Asp Thr Lys Val Leu Cys Arg Asn Glu Gly Lys Tyr Gly Tyr Val 630 635 640 645 650
Ala Asp Gly Cys Ile Tyr Asp Asn Asp 655

<210> SEQ ID NO 198
<211> LENGTH: 55
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 198
Met Ala Trp Pro Ser Arg Ser Lys Met Phe Thr Leu Leu Pro Val Leu 1 5 10 15
Cys Tyr Leu Trp Ser Leu Trp Leu Pro Gln Phe Ser Trp Ile Gln Glu 20 25 30
Leu Lys Ala Val Leu Arg Asp Gly Leu Ile Ser Ala Val Ala Trp 35 40 45 49
Asn Ala Glu Phe Gln Thr Cys 54
<210> SEQ ID NO 199
<211> LENGTH: 266
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 199

Met Val Lys Val Thr Phe Asn Ser Ala Leu Ala Gln Lys Glu Ala Lys
1   5   10   15
Lys Asp Glu Pro Lys Ser Gly Glu Ala Leu Ile Ile Pro Pro Asp
20  25  30
Ala Val Ala Val Asp Cys Lys Asp Pro Asp Asp Val Val Pro Val Gly
35  40  45
Gln Arg Arg Ala Trp Cys Trp Cys Met Cys Phe Gly Leu Ala Phe Met
50  55  60
Leu Ala Gly Val Ile Leu Gly Gly Ala Tyr Leu Tyr Lys Tyr Phe Ala
65  70  75  80
Leu Gln Pro Asp Asp Val Tyr Tyr Cys Gly Ile Lys Tyr Ile Lys Asp
85  90  95
Asp Val Ile Leu Asn Glu Pro Ser Ala Asp Ala Pro Ala Ala Leu Tyr
100 105 110
Gln Thr Ile Glu Glu Asn Ile Lys Ile Phe Glu Glu Glu Glu Val Glu
115 120 125
Phe Ile Ser Val Pro Val Pro Glu Ala Asp Ser Asp Pro Ala Asn
130 135 140
Ile Val His Asp Phe Asn Lys Leu Thr Ala Tyr Leu Asp Leu Asp
145 150 155 160
Leu Asp Lys Cys Tyr Val Ile Pro Leu Asn Thr Ser Ile Val Met Pro
165 170 175
Pro Arg Asn Leu Leu Leu Leu Ile Asn Ile Lys Ala Gly Thr Tyr
180 185 190
Leu Pro Gln Ser Tyr Leu Ile His Glu His Met Val Ile Thr Asp Arg
195 200 205
Ile Glu Asn Ile Asp His Leu Gly Phe Phe Ile Tyr Arg Leu Cys His
210 215 220
Asp Lys Glu Thr Tyr Lys Leu Gin Arg Arg Glu Thr Ile Lys Gly Ile
225 230 235 240
Gln Lys Arg Glu Ala Ser Asn Cys Phe Ala Ile Arg His Phe Glu Asn
245 250 255
Lys Phe Ala Val Glu Thr Leu Ile Cys Ser
260 265

<210> SEQ ID NO 200
<211> LENGTH: 315
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 200

Met Asp Leu Arg Gin Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe
1   5   10   15
Ala Leu Ser Lys Pro Thr Glu Lys Leu Asp Arg Val His His Glu Pro
20  25  30
Gln Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Asp Tyr Asp
35  40  45  45
His Asp Ala Phe Leu Gly Ala Glu Ala Lys Thr Phe Asp Gin Leu
50      55        60
Thr Pro Glu Glu Ser Lys Glu Arg Leu Gly Lys Ile Val Ser Lys Ile
65      70        75        80
Asp Gly Asp Lys Asp Gly Phe Val Thr Val Asp Glu Lys Asp Trp
85      90        95
Ile Lys Phe Ala Gin Lys Arg Trp Ile Tyr Glu Asp Val Glu Arg Gin
100     105       110
Trp Lys Gly His Asp Leu Asn Glu Asp Gly Leu Val Ser Trp Glu Glu
115     120       125
Tyr Lys Asn Ala Thr Tyr Gly Tyr Val Leu Asp Pro Asp Pro Asp
130     135       140
Asp Gly Phe Asn Tyr Lys Gin Met Met Val Arg Asp Gin Arg Arg Phe
145     150       155       160
Lys Met Ala Asp Lys Asp Gly Leu Ile Ala Thr Lys Glu Glu Phe
165     170       175
Thr Ala Phe Leu His Pro Glu Tyr Asp Tyr Met Lys Asp Ile Val
180     185       190
Val Gin Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile
195     200       205
Asp Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gin Thr
210     215       220
Asp Glu Pro Glu Trp Val Lys Thr Glu Arg Glu Gin Phe Val Glu Phe
225     230       235       240
Arg Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Thr Lys Asp
245     250       255
Trp Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His
260     265       270
Leu Val Tyr Glu Ser Asp Gin Asn Lys Asp Gly Lys Leu Thr Lys Glu
275     280       285
Glu Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gin Ala Thr Asp
290     295       300
Phe Gly Glu Ala Leu Val Arg His Asp Gin Phe
305     310       315

<210> SEQ ID NO 201
<211> LENGTH: 207
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 201
Met Phe Asp Ala Val Leu Ile Leu Leu Leu Ile Pro Leu Lys Asp Lys
1       5        10        15
Leu Val Asp Pro Ile Leu Arg Gin Gin His Leu Gin Leu Pro Ser Ser Leu
20      25        30
Lys Arg Ile Ala Val Gly Met Phe Phe Val Met Cys Ser Ala Phe Ala
35      40        45
Ala Gly Ile Leu Glu Ser Lys Arg Leu Asn Leu Val Lys Glu Lys Thr
50      55        60
Ile Asn Gin Thr Ile Gly Asn Val Val Tyr His Ala Ala Asp Leu Ser
65      70        75        80
Leu Trp Trp Gin Val Pro Gin Tyr Leu Leu Ile Gly Ile Ser Glu Ile
85  90  95
Phe Ala Ser Ile Ala Gly Leu Glu Phe Ala Tyr Ser Ala Ala Pro Lys 100  105  110
Ser Met Gin Ser Ala Ile Met Gly Leu Phe Phe Phe Ser Gly Val 115  123  125
Gly Ser Phe Val Gly Ser Gly Leu Ala Leu Val Ser Ile Lys Ala 130  135  140
Ile Gly Trp Met Ser Ser His Thr Asp Phe Gly Asn Ile Asn Gly Cys 145  150  155  160
Tyr Leu Asn Tyr Tyr Phe Phe Leu Leu Ala Ala Ile Gln Gly Ala Thr 165  170  175
Leu Leu Leu Phe Leu Ile Ile Ser Val Gly Tyr Asp His His Arg Asp 180  185  190
His Gln Arg Ser Arg Ala Asn Gly Val Pro Thr Ser Arg Arg Ala 195  200  205

<210> SEQ ID NO 202
<211> LENGTH: 195
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 202
Met Arg Ser Arg Ile Arg Glu Phe Asp Ser Ser Thr Leu Asn Glu Ser 1  5  10  15
Val Arg Asn Thr Ile Met Arg Leu Lys Ala Val Gly Lys Lys Phe 20  25  30
Met His Val Leu Tyr Pro Arg Lys Ser Asn Thr Leu Leu Arg Asp Trp 35  40  45
Asp Leu Trp Gly Pro Leu Ile Leu Cys Val Thr Leu Ala Leu Met Leu 50  55  60
Gln Arg Asp Ser Ala Asp Ser Glu Asp Gly Gly Pro Gin Phe Ala 65  70  75  80
Glu Val Phe Val Ile Val Trp Phe Gly Ala Val Thr Ile Thr Leu Asn 85  90  95
Ser Lys Leu Leu Gly Gly Asn Ile Ser Phe Phe Gln Ser Leu Cys Val 100  105  110
Leu Gly Tyr Cys Ile Leu Pro Leu Thr Val Ala Met Leu Ile Cys Arg 115  120  125
Leu Val Leu Leu Ala Asp Pro Gly Pro Val Ala Phe Met Val Arg Leu 130  135  140
Phe Val Val Ile Val Met Phe Ala Trp Ser Ile Val Ala Ser Thr Ala 145  150  155  160
Phe Leu Ala Asp Ser Gin Pro Pro Asn Arg Arg Ala Leu Ala Val Tyr 165  170  175
Pro Val Phe Leu Phe Tyr Phe Val Ile Ser Trp Met Ile Leu Thr Phe 180  185  190
Thr Pro Gln
195

<210> SEQ ID NO 203
<211> LENGTH: 330
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 203

Met Ala Lys Asp Gin Ala Val Gin Asn Ile Leu Val Val Ser Pro Val Val
  1  5  10  15  
Val Ala Ser Ser Leu Gly Leu Val Ser Leu Gly Gly Lys Ala Thr Thr
  20  25  30  
Ala Ser Gin Ala Val Leu Ser Ala Gin Gin Gin Lys Leu Arg Asp Glu
  35  40  45  
Glu Val His Ala Gin Leu Gly Leu Leu Arg Ser Leu Ser Gin Gin Gin
  50  55  60  
Thr Ala Arg Asn Val Thr Trp Lys Leu Gly Ser Arg Leu Tyr Gly Pro
  65  70  75  80  
Ser Ser Val Ser Phe Ala Asp Phe Val Arg Ser Ser Lys Gin His
  85  90  95  
Tyr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  100 105 110  
Leu Gin Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  115 120 125  
Pro Gin Val Thr Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  130 135 140  
Asn Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  145 150 155 160  
Met Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  165 170 175  
Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  180 185 190  
Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  195 200 205  
Ser Gin Leu Ile Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  210 215 220  
Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  225 230 235 240  
Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  245 250 255  
Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  260 265 270  
Ile Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  275 280 285  
Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  290 295 300  
Asp Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  305 310 315 320  
Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
  325 330 335  

<210> SEQ ID NO 204
<211> LENGTH: 58
<212> TYPE: SWT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 204

Met Cys Met Gin Leu Phe Gly Phe Leu Ala Phe Met Ile Phe Met Cys
  1  5  10  15  

Trp Val Gly Asp Val Tyr Pro Val Tyr Gln Pro Val Gly Pro Lys Gln
20         25         30
Tyr Pro Tyr Asn Asn Leu Tyr Leu Glu Arg Gly Gly Asp Pro Ser Lys
35         40         45
Glu Pro Glu Arg Val Val His Tyr Glu Ile
50         55

<210> SEQ ID NO 205
<211> LENGTH: 392
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 205
Met Asp Ala Leu Val Glu Asp Ile Cys Ile Leu Asn His Glu Lys
1          5         10         15
Ala His Lys Arg Asp Thr Val Thr Pro Val Ser Ile Tyr Ser Gly Asp
20        25         30
Glu Ser Val Ala Ser His Phe Ala Leu Val Thr Ala Tyr Glu Asp Ile
35        40         45
Lys Lys Arg Leu Lys Asp Ser Glu Lys Glu Asn Ser Leu Leu Lys Lys
50         55         60
Arg Ile Arg Phe Leu Glu Glu Lys Leu Ile Ala Arg Phe Glu Glu Glu
65         70         75         80
Thr Ser Ser Val Gly Arg Glu Gin Val Asn Lys Ala Tyr His Ala Tyr
85         90         95
Arg Glu Val Cys Ile Asp Arg Asp Asn Leu Lys Ser Lys Leu Asp Lys
100       105        110
Met Asn Lys Asp Asn Ser Glu Ser Leu Lys Val Leu Asn Glu Gin Leu
115        120        125
Gln Ser Lys Glu Val Glu Leu Leu Leu Arg Thr Glu Val Glu Thr
130        135        140
Gln Gin Val Met Arg Asn Leu Asn Pro Pro Ser Ser Asn Trp Glu Val
145        150        155        160
Glu Lys Leu Ser Cys Asp Leu Lys Ile His Gly Leu Glu Gin Glu Leu
165        170        175
Glu Leu Met Arg Lys Glu Cys Ser Asp Leu Lys Ile Glu Leu Gin Lys
180        185        190
Ala Lys Gin Thr Asp Pro Tyr Gin Glu Asp Asn Leu Lys Ser Arg Asp
195       200        205
Leu Gin Lys Leu Ser Ile Ser Ser Asp Asn Met Gin His Ala Tyr Trp
210      215        220
Glu Leu Lys Arg Glu Met Ser Asn Leu His Leu Val Thr Gin Val Gin
225       230        235        240
Ala Glu Leu Leu Arg Lys Leu Lys Thr Ser Thr Ala Ile Lys Lys Ala
245       250        255
Cys Ala Pro Val Gly Cys Ser Glu Asp Leu Gly Arg Asp Ser Thr Lys
260       265        270
Leu His Leu Met Asn Phe Thr Ala Thr Tyr Thr Arg His Pro Pro Leu
275      280        285
Leu Pro Asn Gly Lys Ala Leu Cys His Thr Thr Ser Ser Pro Leu Pro
290      295        300
Gly Asp Val Lys Val Leu Ser Glu Lys Ala Ile Leu Gin Ser Trp Thr
305       310        315        320
Asp Asn Glu Arg Ser Ile Pro Asn Asp Gly Thr Cys Phe Gln Glu His
325 330 335
Ser Ser Tyr Gly Arg Asn Ser Leu Glu Asp Asn Ser Trp Val Phe Pro
340 345 350
Ser Pro Pro Lys Ser Ser Glu Thr Ala Phe Gly Glu Thr Lys Thr Lys
355 360 365
Thr Leu Pro Leu Pro Asn Leu Pro Pro Leu His Tyr Leu Asp Gln His
370 375 380
Asn Gln Asn Cys Leu Tyr Lys Asn
385 390

<210> SEQ ID NO 206
<211> LENGTH: 26
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 206
Met His His Thr Gln Leu Met Phe Ile Tyr Leu Phe Ile Tyr Leu
1 5 10 15
Phe Ile Leu Gly Val Phe Phe Phe Phe Phe
20 25

<210> SEQ ID NO 207
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 207
Met Asn Cys Ile Leu Leu Leu Tyr Leu Leu Ile Pro Thr Ile Ser Ile
1 5 10 15
Ser Val Val Pro Tyr Val Ala Leu Asn Ile Lys Tyr Ile Lys Glu Cys
20 25 30
Thr Glu Asn Ser Phe Tyr
35

<210> SEQ ID NO 208
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 208
Met Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
1 5 10 15
His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr
20 25 30
Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
35 40 45

<210> SEQ ID NO 209
<211> LENGTH: 81
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 209
-continued

Met Ser Lys Arg Ser Ala Ser Phe Ile Leu Leu Pro Leu Leu Phe Leu
  1  5  10  15
Lys Gly Ser Phe Ala Lys Leu Asn Ala Arg Ile Ser Asp Cys Leu Glu
  20  25  30
Glu Arg Tyr Cys His Asn Leu Trp Met Val Phe Gln Gly Cys Val Ile
  35  40  45
Thr Glu Leu His Leu Ser Arg Met Ser Lys Thr Leu Ser Ser Leu Cys
  50  55  60
Tyr Asp Phe Val Ile Asn Val Tyr Ile Phe Phe Lys Phe Leu Asp Ile
  65  70  75  80
Thr

<210> SEQ ID NO 210
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 210

Met Cys Ser Leu Phe Glu Ser Arg Phe Phe Cys Phe Val Leu Phe Ser
  1  5  10  15
Glu Lys Ile Ile Gln Leu Cys Ala Ser Ile Ala Phe Leu Cys Phe Val
  20  25  30
Lys His Val Pro Trp Pro Lys Trp Lys Arg Lys Cys Leu Ile Asn Ala
  35  40  45
Phe

<210> SEQ ID NO 211
<211> LENGTH: 203
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 211

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

<210> SEQ ID NO 212
<211> LENGTH: 116
<212> TYPE: PRO
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<223> NAME/KEY: MISC_FEATURE
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<223> NAME/KEY: MISC_FEATURE
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<223> NAME/KEY: MISC_FEATURE
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 212
Met Lys Thr Leu Met Thr Ile Cys Pro Gly Thr Val Leu Leu Val Phe
1 5 10 15
Ser Ile Ser Leu Trp Ile Ile Ala Ala Trp Thr Val Arg Val Cys Glu
20 25 30
Ser Pro Glu Ser Pro Ala Gln Pro Ser Gly Ser Ser Leu Pro Ala Trp
35 40 45
Tyr His Asp Gln Gln Asp Val Thr Ser Asn Phe Leu Gly Ala Met Trp
50 55 60
Leu Ile Ser Ile Thr Phe Leu Ser Ile Gly Tyr Gly Asp Met Val Pro
65 70 75 80
His Thr Tyr Cys Gly Lys Gly Val Cys Leu Leu Thr Thr Gly Ile Met Gly
85 90 95
 Ala Gly Cys Thr Ala Leu Val Ala Val Val Ala Arg Lys Leu Glu
100 105 110
Leu Thr Lys Ala Glu Lys His Val His Xaa Phe Met Met Asp Thr Gln
115 120 125
Leu Thr Lys Arg Ile Lys Asn Xaa Ala Ala Asn Val Leu Xaa Glu Thr
130 135 140
Trp Leu Ile Tyr Lys Thr Leu Leu Leu Lys Ile Asp His Ala
145 150 155 160
Lys Val Arg Asn Thr Arg Gly Ser Ser Ser Lys Tyr Pro Pro Val Glu
165 170 175
Glu Arg Gln Asp Gly Thr Glu Glu Ala Glu
180 185

<210> SEQ ID NO 213
<211> LENGTH: 90
<212> TYPE: PRO
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 213
Met Lys Phe Leu Ala Val Leu Val Leu Leu Leu Gly Val Ser Ile Phe Leu
1 5 10 15
Val Ser Ala Gln Asn Pro Thr Thr Ala Ala Pro Ala Asp Thr Tyr Pro
   20  25  30

Ala Thr Gly Pro Ala Asp Glu Ala Pro Ala Glu Thr Thr Ala
   35  40  45

Ala Ala Thr Thr Ala Ala Ala Pro Thr Ala Thr Ala Thr Ala
   50  55  60

Ala Ser Thr Thr Ala Arg Lys Asp Ile Pro Val Leu Pro Lys Trp Val
   65  70  75  80

Gly Asp Leu Pro Asn Gly Arg Val Cys Pro
   85  90

<210> SEQ ID NO: 214
<211> LENGTH: 48
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 214

Met Ser Ser Ala Ala Ala Asp His Trp Ala Trp Leu Leu Val Leu Ser
   1  5 10 15
Phe Val Phe Gly Cys Asn Val Leu Arg Ile Leu Leu Pro Ser Phe Ser
   20 25 30
Ser Phe Met Ser Arg Val Leu Gln Lys Asp Ala Asp Arg Ser His Arg
   35 40 45

<210> SEQ ID NO: 215
<211> LENGTH: 70
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 215

Met Thr Ala Pro Leu Pro Pro Leu Ser Gly Leu Ala Leu Phe Leu Ile
   1 5 10 15
Val Phe Phe Ser Leu Gly Val Phe Cys Ile Cys His Ser His Thr Tyr
   20 25 30
His Thr Leu Gin Gin Met Ala Gly Thr Glu Pro Lys Ala Leu Leu Leu
   35 40 45
Ser Pro Ala Ala Thr Thr Phe Val Thr Val Thr His Glu Val Trp
   50 55 60
Lys Gin Gin Ala Leu Ala
   65 70

<210> SEQ ID NO: 216
<211> LENGTH: 83
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 216

Met Thr Cys Ser Val Ala Leu Leu Leu Ile Leu Gly Leu Arg Cys Ser
   1 5 10 15
Gly Val Arg Pro Gly Leu Val Gly Glu Gly His Asn Pro Ser Leu Leu
   20 25 30
Val Cys Leu Leu Leu Lys Asp Ser Arg Thr Asn Gin Gly Ser Cys Pro
   35 40 45
Gly Gly Pro Trp Ser Glu Arg Asp Ile Glu Ser Val Thr Ser Asp Asn
   50 55 60
Cys Glu Ala Thr Leu Gly Tyr Arg Asn His Ser Leu Pro Ser Asn Tyr
Tyr Asn Ser

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
</tbody>
</table>

SEC ID NO 210
LENGTH: 43
TYPE: PRO
ORGANISM: Homo sapiens

SEQUENCE: 217

Met Leu Thr Arg Ser Leu Lys Thr Leu Pro Ser Ala Cys Thr Ala Phe
1    5    10   15
Leu Leu Leu Phe Phe Leu Leu Ser Ser Gly Asp Pro Glu Leu Ser Cys
20   25   30
Ser Cys Thr Leu Arg Thr Gln Ser Ser Ser Trp Ser
35   40

SEC ID NO 218
LENGTH: 184
TYPE: PRO
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (140)
OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
NAME/KEY: MISC_FEATURE
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
NAME/KEY: MISC_FEATURE
LOCATION: (146)
OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
NAME/KEY: MISC_FEATURE
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
NAME/KEY: MISC_FEATURE
LOCATION: (165)
OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

SEQUENCE: 218

Met Trp Arg Pro Ser Val Leu Leu Leu Leu Leu Leu Arg His Gly
1    5    10   15
Ala Gln Gly Lys Pro Ser Pro Asp Ala Gly Pro His Gly Gin Gly Arg
20   25   30
Val His Gln Ala Ala Pro Leu Ser Asp Ala Pro His Asp Ala His
35   40
Gly Asn Phe Gin Tyr Asp His Gin Ala Phe Leu Gly Arg Gin Val Ala
50   55   60
Lys Glu Phe Asp Gin Leu Thr Pro Glu Glu Ser Gin Ala Arg Leu Gly
65   70   75   80
Arg Ile Val Asp Arg Met Asp Arg Ala Gly Asp Gly Arg Trp Val
85   90
Ser Leu Ala Glu Leu Arg Ala Trp Ile Ala His Thr Gin Gin Arg His
100  105  110
Ile Arg Ser Val Ser Ala Ala Trp Asp Thr Tyr Asp Thr Asp Arg
115  120  125
Asp Gly Arg Val Gly Trp Glu Glu Leu Arg Asn Xaa Thr Tyr Gly His
130  135  140
Xaa Xaa Pro Xaa Glu Glu Phe His Asp Val Glu Asp Ala Glu Thr Tyr
145 150 155 160
Lys Lys Met Leu Xaa Arg Asp Glu Arg Phe Arg Val Ala Asp Gln
165 170 175
Asp Gly Asp Ser Met Ala Thr Arg
180

<210> SEQ ID NO 219
<211> LENGTH: 71
<212> TYPE: PWT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (40)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (51)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (55)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 219
Met Trp Leu Phe Ile Leu Leu Ser Leu Ala Leu Ile Ser Ser Asp Ala Met
1  5   10  15
Val Met Asp Glu Lys Lys Arg Ser Leu Cys Trp Thr Arg Leu Leu
20  25  30
Pro Ser Ala Thr Thr Met Pro Xaa Thr Arg Ile Thr Pro Asn Thr Gly
35  40  45
Ala Glu Xaa Xaa Ile Ser Val Xaa Thr Ala Thr Ser Ser Pro Ser Ser Pro Leu
50  55  60
Thr Ala Pro Ile Met Trp Pro
65  70

<210> SEQ ID NO 220
<211> LENGTH: 10
<212> TYPE: PWT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 220
Met His Val Phe Val Leu Glu Ile Phe Leu
1  5   10

<210> SEQ ID NO 221
<211> LENGTH: 138
<212> TYPE: PWT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 221
Met Ala Val Ala Thr Leu Ala Ser Glu Thr Leu Pro Leu Leu Ala Leu
1  5   10  15
Thr Phe Ile Thr Asp Asn Ser Leu Val Ala Ala Gly His Asp Cys Phe
20  25  30
Pro Val Leu Phe Thr Tyr Asp Ala Ala Ala Gly Met Leu Ser Phe Gly
35  40  45
Gly Arg Leu Asp Val Pro Lys Gln Ser Ser Gln Arg Gly Leu Thr Ala
59  55  60
Arg Glu Arg Phe Glu Asn Asp Leu Asp Gly Asp Lys Ala Ser Ser Ser Glu Gly Gly
65 70 75 80
Thr Ala Ala Gly Ala Gly Leu Asp Ser Leu His Lys Asn Ser Val Ser
85 90 95
Gln Ile Ser Val Leu Ser Gly Lys Ala Lys Cys Ser Glu Phe Cys
100 105 110
Thr Thr Gly Met Asp Gly Gly Met Ser Ile Trp Asp Val Lys Ser Leu
115 120 125
Glu Ser Ala Leu Lys Asp Leu Lys Ile Lys
130 135

<210> SEQ ID NO 222
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 222

Met Ser Gly Gly Leu Ser Phe Leu Leu Leu Val
1 5 10

<210> SEQ ID NO 223
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 223

Leu Gly Ser Leu Ser Thr Ala Pro Ser Ser Ala Leu Pro Thr Leu Gly
1 5 10 15
Ala Arg Arg Thr Arg Ser Lys
20

<210> SEQ ID NO 224
<211> LENGTH: 66
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 224

Met Thr Tyr Phe Ser Gly Leu Leu Leu Val Ile Leu Ala Phe Ala Ala Trp
1 5 10 15
Val Ala Leu Glu Gly Leu Gly Val Ala Val Tyr Ala Ala Ala Val
20 25 30
Leu Leu Gly Ala Gly Cys Ala Thr Ile Leu Val Thr Ser Leu Ala Met
35 40 45
Thr Ala Asp Leu Ile Gly Pro His Thr Asn Ser Gly Leu Ser Cys Thr
50 55 60
Ala Pro
65

<210> SEQ ID NO 225
<211> LENGTH: 28
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 225

Gly Lys Pro Thr Gly Lys Ser Leu Pro Leu Met Trp Met Ile Leu Met
1 5 10 15
Gln Pro Ile Ile Met Ile Ser Met Ser Asn Gly
<210> SEQ ID NO 226
<211> LENGTH: 61
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 226
Met Gln Gly Lys Phe Met Lys Val Gln Val Tyr Arg Phe Leu Lys Tyr
1  5 10 15
Leu Leu Met Leu Leu Cys Met Phe Val Asn Arg Gly Met Ser Lys Asp
20 25 30
Ser Thr Lys Lys Pro Gly Glu Glu Leu Lys Val Ser Leu Gly Ser
35 40 45
Ile Leu Asn Met Lys Ser Gln Arg Pro Leu Ser Trp Cys
55 55 60

<210> SEQ ID NO 227
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 227
Met Met Glu Arg Ser Met Met Met Leu Leu Met Ala Ala Ser Met Thr
1  5 10 15
Met Thr Ser Thr Gln Leu Trp Ser Phe Cys Cys Val His
20 25

<210> SEQ ID NO 228
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 228
Met Trp Tyr Gin Leu Ala Lys Glu Glu Pro Gly Val Gly Ala Cys Ala
1  5 10 15
Leu Asp

<210> SEQ ID NO 229
<211> LENGTH: 72
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 229
Met Leu Ile Cys Arg Leu Val Leu Leu Ala Asp Pro Gly Pro Val Asn
1  5 10 15
Phe Met Val Arg Leu Phe Val Ile Val Met Phe Ala Trp Ser Ile
20 25
Val Ala Ser Thr Ala Phe Leu Ala Asp Ser Gln Pro Pro Asn Arg
35 40
Ala Leu Ala Val Tyr Pro Val Phe Leu Phe Tyr Phe Val Ile Ser Trp
55 60
Met Ile Leu Thr Phe Thr Pro Gin
65 70

<210> SEQ ID NO 230
<211> LENGTH: 142
<212> TYPE: PRT
**ORGANISM:** Homo sapiens  
**FEATURE:**  
**NAME/KEY:** MISC_FEATURE  
**LOCATION:** (47)  
**OTHER INFORMATION:** Xaa equals any of the L-amino acids commonly found in naturally occurring proteins  
**NAME/KEY:** MISC_FEATURE  
**LOCATION:** (121)  
**OTHER INFORMATION:** Xaa equals any of the L-amino acids commonly found in naturally occurring proteins  

### SEQUENCE: 230

| Met | Arg | Ser | Leu | Leu | Leu | Ser | Ala | Phe | Cys | Leu | Leu | Glu | Ala | Ala | Ala | Ala | Ala | Ala | Ala | Ala | Ala | Ala |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     | 1   | 5   | 10  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Leu | Ala | Glu | Val | Lys | Lys | Pro | Ala | Ala | Ala | Ala | Ala | Pro | Gly | Thr |     |     |     |     |     |     |     |     |     |     |     |
|     | 20  |     | 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ala | Glu | Lys | Leu | Ser | Pro | Lys | Ala | Ala | Thr | Leu | Ala | Glu | Arg | Xaa | Arg |     |     |     |     |     |     |     |     |     |     |     |
| 35  |     | 40  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Pro | Gly | Leu | Gin | Leu | Val | Pro | Gly | His | Gly | Gin | Gly | Pro | Gly | Ser | Gly |     |     |     |     |     |     |     |     |     |     |     |     |
| 50  |     | 55  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | His | Pro | Gly | Val | Thr | Arg | Gly | Gly | Gly | Leu | Val | Ala | Gly | Ala | Arg |     |     |     |     |     |     |     |     |     |     |     |     |
| 65  |     | 70  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Val | Ala | Gly | Arg | Glu | Gin | Gly | Val | Ala | Glu | Gin | Gly | Glu | Gly | Ser | Ala |     |     |     |     |     |     |     |     |     |     |     |
| 85  |     | 90  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | Arg | Arg | Ala | Ala | Ala | Ala | Arg | Arg | Arg | Glu | Arg | Pro | Gly | Arg |     |     |     |     |     |     |     |     |     |     |     |     |
| 100 |     | 105 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ala | Ala | Ala | Leu | Thr | Gin | Gin | Leu | Xaa | Gin | Gin | Arg | Gin | Gin | Arg | Asp | Leu | Glu |     |     |     |     |     |     |     |     |
| 115 |     | 120 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ala | Gly | Gin | Pro | Thr | Val | Arg | Thr | Gin | Leu | Ser | Glu | Leu | Arg |     |     |     |     |     |     |     |     |     |     |     |     |
| 130 |     | 135 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

### SEQUENCE: 231

| Asp | Pro | Glu | Ala | Ala | Asp | Ser | Gly | Glu | Pro | Gin | Asn | Lys | Arg | Thr | Pro |     |     |     |     |     |     |     |     |     |     |     |
| 1   | 5   | 10  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

| Asp | Leu | Pro | Glu | Glu | Tyr | Val | Lys | Glu | Glu | Ile | Gin | Gin | Asn | Glu |     |     |     |     |     |     |     |     |     |     |     |     |
| 20  |     | 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

| Glu | Ala | Val | Lys | Lys | Met | Leu | Val | Glu | Ala | Thr | Arg | Glu | Phe | Glu | Glu |     |     |     |     |     |     |     |     |     |     |     |
| 35  |     | 40  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

### SEQUENCE: 232

| Gln | Lys | Leu | Lys | Arg | Lys | Ala | Glu | Asp | Pro | Glu | Ala | Ala | Asp | Ser |     |     |     |     |     |     |     |     |     |     |     |
| 1   | 5   | 10  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

| Gly | Glu | Pro | Gin | Asn | Lys | Arg | Thr | Pro | Asp | Leu | Pro | Glu | Glu | Gly | Tyr |     |     |     |     |     |     |     |     |     |     |     |
| 20  |     | 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

| Val | Lys | Glu | Glu | Ile | Gin | Glu | Asn | Glu | Ala | Val | Lys | Lys | Met | Leu |     |     |     |     |     |     |     |     |     |     |     |
| 35  |     | 40  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
Val Glu Ala Thr Arg Glu Phe Glu Glu Val Val Val Asp Glu Ser
50  55       60

<210> SEQ ID NO 233
<211> LENGTH: 111
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 233

Lys Ala Met Glu Lys Ser Ser Leu Thr Gin His Ser Trp Gin Ser Leu
1   5  10  15

Lys Asp Arg Tyr Leu Lys His Leu Arg Gly Gin Glu His Lys Tyr Leu
20  25    30

Leu Gly Asp Ala Pro Val Ser Pro Ser Gin Gin Gin Gin Gin Gin Gin Gin
35  40  45

Lys Val Asp Pro Glu Ala Ala Asp Ser Gin Glu Gin Gin Gin Gin Gin Gin Gin
50  55    60

Arg Thr Pro Asp Leu Pro Glu Glu Tyr Val Lys Gin Glu Gin Ile Gin
65  70  75  80

Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
85  90  95

Phe Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
100 105 110

Ile

<210> SEQ ID NO 234
<211> LENGTH: 148
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 234

Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr Ile
1   5  10  15

Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe
20  25    30

Asp Asp Ala Asp Gin Leu Arg Ile Gin Gin Asp Gin Ile Gin Gin Gin Gin
35  40  45

Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Glu Gin Phe Phe Leu Ser
50  55    60

Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser Gin
65  70  75  80

Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser Thr
85  90

Tyr Phe Pro Gly Tyr Phe Asp Gly Gin Tyr Trp Leu Thr Trp Val Phe
95

Leu Val Leu Gly Phe Leu Leu Leu Arg Gly Phe Ile Asn Tyr Ala
100 105 110

Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr Arg
115 120 125

Val Leu Phe Ile
130 135 140

<210> SEQ ID NO 235
<211> LENGTH: 24
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 235

Ala Gly Arg Tyr Gly Ala Ile Ser Gly Phe Gly Leu Ser Leu Ile Lys
  1    5       10        15
Trp Ile Leu Ile Val Arg Phe Ser
  20

<210> SEQ ID NO 236
<211> LENGTH: 51
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 236

Met Lys His Leu Ser Ala Trp Asn Phe Thr Lys Leu Thr Phe Leu Gln
  1    5       10        15
Leu Trp Glu Ile Phe Glu Gly Ser Val Glu Asn Cys Gln Thr Leu Thr
  20   25       30
Ser Tyr Ser Lys Leu Gln Ile Lys Tyr Thr Phe Ser Arg Gly Ser Thr
  35   40       45
Phe Tyr Ile
  50

<210> SEQ ID NO 237
<211> LENGTH: 213
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 237

Phe Ser Ser Asp Phe Arg Thr Ser Pro Trp Glu Ser Arg Arg Val Glu
  1    5       10        15
Ser Lys Ala Thr Ser Ala Arg Cys Gly Leu Trp Gly Ser Gly Pro Arg
  20   25       30
Arg Arg Pro Ala Ser Gly Met Phe Arg Gly Leu Ser Ser Trp Leu Gly
  35   40       45
Leu Gln Gln Pro Val Ala Gly Gly Gln Gly Gln Pro Asn Gly Asp Ala Pro
  50   55       60
Pro Glu Gln Pro Ser Glu Thr Val Ala Glu Ser Ala Glu Glu Gln Leu
  65   70       75       80
Gln Gln Ala Gly Asp Gin Glu Leu His Gin Ala Lys Asp Phe Gly
  85    90
Asn Tyr Leu Phe Asn Phe Ala Ser Ala Thr Lys Lys Ile Thr Glu
  100  105      110
Ser Val Ala Glu Thr Ala Gin Thr Ile Lys Ser Val Glu Glu Gly
  115  120     125
Lys Ile Asp Gly Ile Ile Asp Lys Thr Ile Ile Gly Asp Phe Gin Lys
  130  135     140
Glu Gin Lys Lys Phe Val Glu Gin His Thr Lys Lys Ser Glu Ala
  145  150     155     160
Ala Val Pro Pro Trp Val Asp Thr Asn Asp Glu Gly Thr Ile Gin Gin
  165  170     175
Gln Ile Leu Ala Leu Ser Ala Asp Lys Arg Asn Phe Leu Arg Asp Pro
  180  185     190
Pro Ala Gly Val Gin Phe Asn Phe Asp Phe Gin Gin Met Tyr Pro Val
<continued>

<table>
<thead>
<tr>
<th></th>
<th>195</th>
<th>200</th>
<th>205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Val</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 238
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 238

Met Arg Phe Ala Leu Val Pro Lys Leu Val Lys Glu Glu Val Phe Trp
1    5    10   15

Arg Asn Tyr Phe Tyr Arg Val Ser Leu Ile Lys Gin Ser Ala Gin Leu
20   25   30

Thr Ala Leu Ala Ala Gin Gin Ala Ala Gly Lys Gly Gly Glu Glu
35   40   45

Gln

<210> SEQ ID NO 239
<211> LENGTH: 76
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 239

Ser Thr Ser Pro Gly Val Ser Glu Val Lys Phe Val Ser Asp Ala Phe Asp Ala
1    5    10   15

Cys Asn Leu Asn Gin Glu Asp Leu Arg Lys Glu Met Glu Gin Leu Val
20   25   30

Leu Asp Lys Lys Gin Glu Glu Thr Ala Val Leu Glu Glu Asp Ser Ala
35   40   45

Asp Trp Glu Lys Leu Gin Gin Glu Leu Gin Glu Tyr Glu Val Val
55   55   60

Thr Glu Ser Glu Lys Arg Asp Glu Asn Trp Asp Lys
65   70   75

<210> SEQ ID NO 240
<211> LENGTH: 62
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 240

Ser Pro Trp Glu Ser Arg Arg Arg Val Glu Ser Lys Ala Thr Ser Ala Arg
1    5    10   15

Cys Gly Leu Trp Gly Ser Gly Pro Arg Arg Arg Arg Pro Ala Ser Gly Met
20   25   30

Phe Arg Gly Leu Ser Ser Trp Leu Gly Leu Gin Gin Pro Val Ala Gly
35   40   45

Gly Gly Gin Pro Asn Gly Asp Ala Pro Pro Glu Gin Pro Ser
50   55   60

<210> SEQ ID NO 241
<211> LENGTH: 65
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 241

Pro Val Ala Gly Gly Gly Gin Pro Asn Gly Asp Ala Pro Pro Glu Gin
1 5 10 15

Pro Ser Glu Thr Val Ala Glu Ser Ala Glu Glu Glu Leu Gin Gin Ala
20 25 30

Gly Asp Gin Glu Leu Leu His Gin Ala Lys Asp Phe Gly Asn Tyr Leu
35 40 45

Phe Asn Phe Ala Ser Ala Ala Thr Lys Ile Thr Glu Ser Val Ala
50 55 60

Glu
65

<210> SEQ ID NO 242  
<211> LENGTH: 72  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 242

Phe Gln Lys Glu Gin Lys Lys Phe Val Glu Glu Gin His Thr Lys Lys
1 5 10 15

Ser Glu Ala Ala Val Pro Pro Trp Val Asp Thr Asp Ala Glu Glu Thr
20 25 30

Ile Gin Gin Gin Ile Leu Ala Leu Ser Ala Asp Arg Asn Phe Leu
35 40 45

Arg Asp Pro Pro Ala Gly Val Glu Phe Asn Phe Asp Asp Gin Met
50 55 60

Tyr Pro Val Ala Leu Val Met Leu
65 70

<210> SEQ ID NO 243  
<211> LENGTH: 28  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 243

Pro Phe Ile Cys Val Ala Arg Asn Pro Val Ser Arg Asn Phe Ser Ser
1 5 10 15

Pro Ile Leu Ala Arg Lys Leu Cys Glu Gly Ala Ala
20 25

<210> SEQ ID NO 244  
<211> LENGTH: 33  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 244

Lys Glu Asp Pro Ala Asn Thr Val Tyr Ser Thr Val Glu Ile Pro Lys
1 5 10 15

Lys Met Glu Asn Pro His Ser Leu Leu Thr Met Pro Asp Thr Pro Arg
20 25 30

Leu

<210> SEQ ID NO 245  
<211> LENGTH: 227  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 245

Ala Ser Ala Val Leu Leu Asp Leu Pro Asn Ser Gly Gly Glu Ala Gin
-continued

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala Lys Leu Gly Asn Ann Cys Val Phe Ala Pro Ala Asp Val Thr</td>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser Glu Lys Asp Val Gln Thr Ala Leu Ala Leu Ala Lys Gly Lys Phe</td>
<td>35</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly Arg Val Asp Val Ala Val Asn Cys Ala Gly Ile Ala Val Ala Ser</td>
<td>50</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys Thr Tyr Asn Leu Lys Gly Gin Thr His Thr Leu Glu Asp Phe</td>
<td>65</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln Arg Val Leu Asp Val Asn Leu Met Gin Thr Phe Asn Val Ile Arg</td>
<td>85</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu Val Ala Gly Gin Met Gin Asn Glu Pro Asp Gin Gly Gin</td>
<td>100</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg Gin Val Ile Ile Asn Thr Ala Ser Val Ala Ala Phe Gin Gly Gin</td>
<td>115</td>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Gin Ala Ala Tyr Ser Ala Ser Lys Gin Gin Ile Val Gin Met</td>
<td>130</td>
<td>135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Leu Pro Ile Ala Arg Asp Leu Ala Pro Ile Gin Gin Arg Val Met</td>
<td>145</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Ile Ala Pro Gin Gin Leu Phe Gin Thr Gin Thr Gin Ser Gin Thr</td>
<td>160</td>
<td>165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gin Lys Gin Cys Gin Phe Gin Ala Gin Gin Gin Gin Gin Gin Gin</td>
<td>180</td>
<td>185</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin G
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 249

His Pro Ile Glu Trp Ala Ile Asn Ala Ala Thr Leu Ser Gln Phe Tyr

Ile Asn Lys Leu Cys Phe

<210> SEQ ID NO: 249
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 249

Cys Trp Ile Lys Tyr Cys Leu Thr Leu Met Gln Asn Ala Gln Leu Ser

Met Gln Asp Asn Ile Gly

<210> SEQ ID NO: 250
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 250

Lys Val Ser Tyr Leu Arg Pro Leu Asp Phe Glu Glu Ala Arg Glu Leu

Phe Leu Leu Gly Gln His Tyr Val Phe

<210> SEQ ID NO: 251
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (11)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 251

Met Glu Arg Arg Cys Lys Met His Lys Arg Xaa Ile Ala Met Leu Glu

Pro Leu Thr Val Asp Leu Asn Pro Gin

<210> SEQ ID NO: 252
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 252

Ser His Ile Val Lys Ile Asn Ala Leu Asn Lys Ser Ala Leu Lys

Tyr Tyr Gln Leu Phe Leu Asp

<210> SEQ ID NO: 253
<211> LENGTH: 64
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 253
Phe Thr His Leu Ser Thr Cys Leu Leu Ser Leu Leu Leu Val Arg Met
1      5      10     15
Ser Gly Phe Leu Leu Leu Ala Arg Ala Ser Ser Pro Ser Ile Cys Ala Leu
20     25     30
Asp Ser Ser Cys Phe Val Gln Glu Tyr Cys Ser Ser Tyr Ser Ser Ser
35     40     45
Cys Phe Leu His Gln His Phe Pro Ser Leu Leu Asp His Leu Cys Gln
50     55     60

<210> SEQ ID NO 254
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 254
Phe Leu Leu Leu Ala Arg Ala Ser Ser Ile Cys Ala Leu Asp Ser
1      5      10     15
Ser Cys Phe Val Gln Glu Tyr
20

<210> SEQ ID NO 255
<211> LENGTH: 53
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 255
Pro Asp Gly Arg Val Thr Asn Ile Pro Gln Gly Met Val Thr Asp Gln
1      5      10     15
Phe Gly Met Ile Gly Leu Leu Thr Phe Ile Arg Ala Ala Glu Thr Asp
20     25     30
Pro Gly Met Val His Leu Ala Ala Gln Ser Leu Thr Thr Leu Gly
35     40     45
Leu Asn Leu Asn Ser
50

<210> SEQ ID NO 256
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 256
Glu Asp Leu Leu Phe Tyr Leu Tyr Tyr Met Asn Gly Asp Val Leu
1      5      10     15
Gln Leu Leu Ala Ala Val Glu Leu Phe Asn Arg Asp Trp Arg Tyr His
20     25     30
Lys Glu Glu Arg Val Thr Ile Thr Arg
35     40

<210> SEQ ID NO 257
<211> LENGTH: 24
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 257
Val His Leu Ala Leu Gly Ser Asp Leu Thr Thr Leu Gly Leu Asn Leu
1      5      10     15
Asn Ser Pro Glu Asn Leu Tyr Pro
20

<210> SEQ ID NO 259
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 259

Glu Asp Leu Leu Phe Tyr Tyr Tyr Met Asn Gly Gly Asp Val Leu
1  5  10  15

Gln Leu Leu Ala Ala Val Glu Leu Phe Asn Arg Asp Trp Arg Tyr His
20  25  30

Lys Glu Glu Arg Val Trp Ile Thr Arg
35  40

<210> SEQ ID NO 259
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 259

His Asn Glu Asp Phe Pro Ala Leu Pro Gly Ser
1  5  10

<210> SEQ ID NO 260
<211> LENGTH: 75
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 260

Gly Arg Ile Ile Asp Thr Ser Leu Thr Arg Asp Pro Leu Val Ile Glu
1  5  10  15

Leu Gly Gin Lys Gin Val Ile Pro Gly Leu Glu Gln Ser Leu Leu Asp
20  25  30

Met Cys Val Gly Lys Arg Arg Ala Ile Ile Pro Ser His Leu Ala
35  40  45

Tyr Gly Lys Arg Gly Phe Pro Pro Ser Val Pro Ala Asp Ala Val
50  55  60

Gln Tyr Asp Val Glu Leu Ile Ala Leu Ile Arg
65  70  75

<210> SEQ ID NO 261
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 261

Ile His Tyr Thr Gly Ser Leu Val Asp Gly Arg Ile Ile Asp Thr Ser
1  5  10  15

<210> SEQ ID NO 262
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 262

Cys Glu Ser Pro Glu Ser Pro Ala Gln Pro Ser Gly Ser Ser Leu Pro
1  5  10  15
Ala Trp Tyr His
20

<210> SEQ ID NO: 263
<211> LENGTH: 95
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 263

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

<210> SEQ ID NO: 264
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 264

Ser Lys Gln Arg Ile Asn Asn Trp Lys Glu Ser Lys His Lys Val Ile
  1  5  10  15
Met Ala Ser Ala Ser Ala Arg
  20

<210> SEQ ID NO: 265
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 265

Leu Phe His Trp Ala Cys Leu Asn Glu Arg Ala Ala Glu Leu Pro Arg
  1  5  10  15
Asn Thr Ala Lys Ala Gly Tyr Gln Cys Pro Ser Cys Asn Gly Pro Ser
  20  25  30

<210> SEQ ID NO: 266
<211> LENGTH: 185
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 266

Phe Tyr Ile Tyr Tyr Arg Pro Thr Asp Ser Asp Asp Ser Asp Tyr
  1  5  10  15
Lys Lys Asp Met Val Glu Gly Asp Lys Tyr Trp His Ser Ile Ser His
  20  25  30
Leu Gln Pro Glu Thr Ser Tyr Asp Ile Lys Met Gln Cys Phe Asn Glu
35  40  45
Gly Gly Glu Ser Glu Phe Ser Ser Asn Val Met Ile Cys Glu Thr Lys Ala
50  55  60
Arg Lys Ser Ser Gly Gin Pro Gly Leu Pro Pro Pro Thr Leu Ala
65  70  75  80
Pro Pro Gin Pro Pro Leu Pro Glu Thr Ile Glu Arg Pro Val Gly Thr
85  90  95
Gly Ala Met Val Ala Arg Ser Ser Asp Leu Pro Tyr Leu Ile Val Gly
100 105 110
Val Val Leu Gly Ser Ile Val Leu Ile Val Thr Phe Ile Pro Phe
115 120 125
Cys Leu Trp Arg Ala Trp Ser Lys Gin His Thr Thr Asp Leu Gly
130 135 140
Phe Pro Arg Ser Ala Leu Pro Pro Ser Cys Pro Tyr Thr Met Val Pro
145 150 155 160
Leu Gly Gly Leu Pro Gly His Gin Ala Val Asp Ser Pro Thr Ser Val
165 170 175
Ala Ser Val Asp Gly Pro Val Leu Met
180 185

<210> SEQ ID NO 267
<211> LENGTH: 66
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 267

Tyr Ile Tyr Tyr Arg Pro Thr Asp Ser Asp Asn Ser Asp Tyr Lys
1    5    10   15
Lys Asp Met Val Glu Gly Asp Tyr Trp His Ser Ile Ser His Leu
20   25   30
Gln Pro Glu Thr Ser Tyr Asp Ile Lys Met Gln Cys Phe Asn Glu Gly
35   40   45
Gly Glu Ser Glu Phe Ser Ser Val Met Ile Cys Glu Thr Lys Ala Arg
50  55  60

Lys Ser
65

<210> SEQ ID NO 268
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 268

Asn Val Arg Ala Leu Leu His Arg Met Pro Glu Pro Pro Lys Ile Asn
1    5    10   15
Thr Ala Lys Phe Asn Asn Lys Arg Lys Asn Leu Ser Leu
20   25   30

<210> SEQ ID NO 269
<211> LENGTH: 185
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 269
<210> SEQ ID NO: 270
<211> LENGTH: 84
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 271

Pro Ser Val Ile Ile Leu Ile Arg Thr Val Ser Pro Glu Leu Lys Ser
  1    5    10    15

Tyr Ala Leu Gly Val Leu Phe Leu Leu Leu Arg Leu Gly Phe Ile
  20   25    30

Pro Pro Pro Leu Ile Phe Gly Ala Gly Ile Asp Ser Thr Cys Leu Phe
  35   40    45

<210> SEQ ID NO: 270
<211> LENGTH: 65
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 270

Ser Tyr Leu Ser Ala Cys Phe Ala Gly Cys Aan Ser Thr Asn Leu Thr
  1    5    10    15

Gly Cys Ala Cys Leu Thr Thr Pro Ala Glu Aan Ala Thr Val Val
  20   25    30

Pro Gly Lys Cys Pro Ser Pro Gly Cys Gin Glu Ala Phe Leu Thr Phe
  35   40    45

Leu Cys Val Met Cys Ile Cys Ser Leu Ile Gly Ala Met Ala Arg His
  50   55    60

<210> SEQ ID NO: 270
<211> LENGTH: 85
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 270

Asn Thr Asn Gin Arg Glu Ala Leu Gin Tyr Ala Lys Asn Phe Gin Pro
  1    5    10    15

Phe Ala Leu Aan His Gin Lys Asp Ile Gin Val Leu Met Gly Ser Leu
  20   25    30

Val Tyr Leu Arg Gin Gly Ile Glu Aan Ser Pro Tyr Val His Leu Leu
  35   40    45

Asp Ala Asn Gin Trp Ala Asp Ile Cys Asp Ile Phe Thr Arg Asp Ala
  50   55    60

Cys Ala Leu Leu Gly Leu Ser Val Glu Ser Pro Leu Ser Val Ser Phe
  65   70    75    80

Ser Ala Gly Cys Val Ala Leu Pro Ala Leu Ile Aan Ile Lys Ala Val
  85   90    95

Ile Glu Gin Arg Gin Cys Thr Gly Val Trp Aan Gin Lys Asp Glu Leu
 100 105    110

Pro Ile Glu Val Asp Leu Gly Lys Cys Trp Tyr His Ser Ile Phe
 115 120   125

Ala Cys Pro Ile Leu Arg Gin Gin Thr Thr Asp Aan Aan Pro Pro Met
 130 135   140

Lys Leu Val Cys Gly His Ile Ile Ser Arg Asp Ala Leu Aan Lys Met
 145 150   155   160

Phe Asn Gly Ser Lys Leu Lys Cys Pro Tyr Cys Pro Met Glu Gin Ser
 165 170   175

Pro Gly Asp Ala Lys Gin Ile Phe Phe
 180 185
-continued

Trp Ser Thr Phe Cys Gly Glu Gln Ala Cys Val Leu Tyr Asp Asn
50
55
60
Val Val Tyr Arg Tyr Leu Tyr Val Ser Ile Ala Ala Leu Lys Ser
65
70
75
80
Phe Ala Phe Ile

<210> SEQ ID NO: 272
<211> LENGTH: 192
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (29)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (30)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 272

Gln Ser Leu Phe Thr Arg Phe Val Arg Val Gly Val Pro Thr Val Asp
1
5
10
15
Leu Asp Ala Gin Gly Arg Ala Ser Leu Cys Xaa Xaa Tyr Asn
20
25
30
Trp Arg Tyr Lys Asn Leu Gly Asn Leu Pro His Val Gin Leu Leu Pro
35
40
45
Glu Phe Ser Thr Ala Asn Ala Gin Gly Leu Leu Tyr Asp Gin Leu Ile
50
55
60
Asn Val Glu Asp Phe Gln Val Gly Val Gly Ser Gly Pro Asn Pro Tyr
65
70
75
80
Phe Tyr Gin Asn Leu Gly Glu Ala Glu Tyr Val Val Ala Leu Phe Met
85
90
95
Tyr Met Cys Leu Leu Gly Tyr Pro Ala Asp Lys Ile Ser Ile Leu Thr
100
105
110
Thr Tyr Asn Gin Lys His Leu Ile Arg Asp Ile Ile Asn Arg Arg
115
120
125
Cys Gly Asn Asn Pro Leu Ile Gly Gly Pro Asn Lys Val Thr Thr Val
130
135
140
Asp Arg Phe Gin Gly Gin Gin Asn Asp Tyr Ile Leu Leu Ser Leu Val
145
150
155
160
Arg Thr Arg Ala Val Gly His Leu Arg Asp Val Arg Arg Leu Val Val
165
170
175
Ala Met Ser Arg Ala Arg
180

<210> SEQ ID NO: 273
<211> LENGTH: 77
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 273

Leu Val Lys Glu Ala Lys Ile Ile Ala Met Thr Cys Thr His Ala Ala
1
5
10
15
Leu Lys Arg His Asp Leu Val Lys Leu Gly Phe Lys Tyr Asp Asn Ile
20
25
30
Leu Met Glu Glu Ala Ala Gln Ile Leu Glu Ile Glu Thr Phe Ile Pro
Leu Leu Leu Gln Asn Pro Gln Asp Gly Phe Ser Arg Leu Lys Arg Trp 50 55 60
Ile Met Ile Gly Asp His His Gln Leu Pro Pro Pro Val Ile 65 70 75

<table>
<thead>
<tr>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
</table>

<210> SEQ ID NO 274
<211> LENGTH: 125
<212> TYPE: PROT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (16)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (17)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (43)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp Thr Tyr Pro Asn Glu Lys Gln Gln Glu Arg Val Phe Pro Xaa 20 25 30 35 40 45 50 55 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser Ala Met Val Asn Asn Gly Ser Leu Ser Tyr: Asp His Glu Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp Gly Arg Pro Thr Glu Leu Gly Gly Cys Xaa Ala Ile Val Arg Asn 35 40 45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu His Tyr Asp Thr Phe Leu Val Ile Arg Tyr Val Lys Arg His Leu 50 55 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Ile Met Met Asp Ile Asp Gly Lys His Glu Trp Arg Asp Cys Ile 65 70 75 80 85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu Val Pro Gly Val Arg Leu Pro Arg Gly Tyr Phe Gly Thr Ser 90 95 100 105 110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser Ile Thr Gly Asp Leu Ser Asp Asn Asp Val Ile Ser Leu Lys 115 120 125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu Phe Glu Leu Thr Val Glu Arg Thr Pro Glu Glu Glu</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 275
<211> LENGTH: 85
<212> TYPE: PROT
<213> ORGANISM: Homo sapiens

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu Lys Arg Glu His Ser Leu Ser Lys Pro Tyr Gln Gly Val Gly Thr 20 25 30 35 40 45 50 55 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly Ser Ser Ser Leu Trp Asn Leu Met Gly Asn Ala Met Val Met Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln Tyr Ile Arg Leu Thr Pro Asp Met Gln Ser Lys Glu Gly Ala Leu 50 55 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp Asn Arg Val Pro Cys Phe Leu Arg Asp Trp Glu Leu Gln Val His</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe Lys Ile His Gly Gln Gly Lys Lys Asn Leu His Gly Asp Gly Leu 65 70 75 80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly Ser Ser Ser Leu Trp Asn Leu Met Gly Asn Ala Met Val Met Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
-continued

Ala Ile Trp Tyr Thr
85

<210> SEQ ID NO 276
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 276
Pro Gly Thr Leu Gin Cys Ser Ala Leu His His Asp Pro Gly Cys Ala
1 5 10 15
Asn Cys Ser Arg Phe Cys Arg Asp Cys Ser Pro Pro Ala Cys Gin Cys
20 25 30

<210> SEQ ID NO 277
<211> LENGTH: 27
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (8)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 277
Phe Leu Tyr Asp Val Leu Met Xaa His Gin Ala Val Met Arg Thr His
1 5 10 15
Gln Ile Gin Leu Pro Asp Pro Gin Phe Pro Ser
20 25

<210> SEQ ID NO 278
<211> LENGTH: 92
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 278
Pro Ala Asp Xaa Lys Pro Val Val Ser Thr Glu Ala Pro Pro Ile Ile
1 5 10 15
Phe Ala Thr Pro Thr Lys Leu Thr Ser Asp Ser Thr Val Tyr Asp Tyr
20 25 30
Ala Gly Lys Asn Lys Val Pro Glu Leu Gin Lys Phe Phe Gin Lys Ala
35 40
Asp Gly Val Pro Val Tyr Leu Lys Arg Gly Leu Pro Asp Gin Met Leu
50 55 60
Tyr Arg Thr Thr Met Ala Leu Thr Val Gly Arg Thr Ile Tyr Cys Leu
65 70 75 80
Ile Ala Leu Tyr Met Ala Ser Gin Pro Lys Asn Lys
85 90

<210> SEQ ID NO 279
<211> LENGTH: 63
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (45)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly
found in naturally occurring proteins

<400> SEQUENCE: 279
Ser Phe Ser Gly Ala Val Ala Leu Ala Ala Asp Ala Gly Ser Arg Thr
1      5     10       15
Leu Gly Val Met Tyr Tyr Lys Phe Ser Gly Phe Thr Gln Lys Leu Ala
20     25     30
Gly Ala Trp Ala Ser Glu Ala Tyr Ser Pro Glu Ile Xaa Ser Leu Trp
35     40     45
Phe Pro Gln Lys His His Leu Ser Tyr Leu Pro His Gln Leu Asn
50     55     60

<210> SEQ ID NO 280
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 280
Gly Trp Tyr Trp Cys Gly
1      5

<210> SEQ ID NO 281
<211> LENGTH: 129
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 281
Met Lys Val Gly Ala Arg Ile Arg Val Lys Met Ser Val Asn Lys Ala
1      5     10       15
His Pro Val Val Ser Thr His Trp Arg Trp Pro Ala Glu Trp Pro Gln
20     25     30
Met Phe Leu His Leu Ala Gln Gln Pro Arg Thr Glu Val Lys Ser Arg
35     40     45
Pro Leu Gly Leu Ala Gly Phe Ile Arg Glu Asp Ser Lys Thr Arg Lys
50     55     60
Pro Leu Glu Gln Glu Thr Ile Met Ser Ala Ala Asp Thr Ala Leu Trp
65     70     75     80
Pro Tyr Gly His Gly Asn Arg Glu His Gln Glu Asn Glu Leu Gln Lys
85     90     95
Tyr Leu Gln Tyr Lys Asp Met His Leu Leu Asp Gln Ser Leu
100    105    110
Gly His Thr His Thr Leu Gln Gly Ser His Asn Leu Thr Ala Leu Asn
115    120    125
Ile

<210> SEQ ID NO 282
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 282
Ser Leu His Lys Asn Ser Val Ser Glu Ile Ser Ser Val Leu Ser Gly Gly
1      5     10       15
Lys Ala Lys Cys Ser Gln Phe Cys Thr Thr Gly Met Asp Gly Gly Met
20     25     30
Ser Ile Trp Asp Val Lys Ser Leu Glu Ser Ala Leu Lys Asp Leu Lys
Ile

35

<210> SEQ ID NO 283
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 283

Glu Ala Ser Lys Ser Ser His Ala Gly Leu Asp Leu Phe Ser Val Ala
1    5     10    15

Ala Cys His Arg Phe
20

<210> SEQ ID NO 284
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 284

Tyr Met Gly Lys Gly Ser Met Thr Gly Leu Ala Leu Lys His Met Phe
1    5     10    15

Glu Arg Ser Phe Thr
20

<210> SEQ ID NO 285
<211> LENGTH: 27
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 285

Val Thr Gly Ile Ile Asp Ser Leu Thr Ile Ser Pro Lys Ala Ala Arg
1    5     10    15

Val Gly Leu Leu Gin Tyr Ser Thr Gin Val His
20  25

<210> SEQ ID NO 286
<211> LENGTH: 24
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 286

Thr Glu Phe Thr Leu Arg Asn Phe Asn Ser Ala Lys Asp Met Lys Lys
1    5     10    15

Ala Val Ala His Met Lys Tyr Met
20

<210> SEQ ID NO 287
<211> LENGTH: 27
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 287

Gly Lys Gly Ser Met Thr Gly Leu Ala Leu Lys His Met Phe Glu Arg
1    5     10    15

Ser Phe Thr Gin Gly Glu Gly Ala Arg Pro Phe
20  25

<210> SEQ ID NO 288
<211> LENGTH: 44
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 288

Ser Thr Arg Val Pro Arg Ala Ala Ile Val Phe Thr Asp Gly Arg Ala
1     5     10     15

Gln Asp Asp Val Ser Glu Trp Ala Ser Lys Ala Lys Ala Asn Gly Ile
20    25    30

Thr Met Tyr Ala Val Gly Val Gly Lys Ala Ile Glu  
35    40

<210> SEQ ID NO 289
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 289

Glu Glu Leu Gin Glu Ile Ala Ser Glu Pro Thr Asn Lys His Leu Phe
1     5     10     15

Tyr Ala Glu Asp Phe Ser Thr Met Asp Glu Ile Ser Glu Lys Leu Lys
20    25    30

Lys Gly Ile Cys Glu Ala Leu Glu Asp Ser  
35    40

<210> SEQ ID NO 290
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 290

Thr Gln Arg Leu Glu Glu Met Thr Gln Arg Met
1     5     10

<210> SEQ ID NO 291
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 291

Pro Gln Gly Cys Pro Glu Gln Pro Leu His
1     5     10

<210> SEQ ID NO 292
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 292

Arg Cys Lys Lys Cys Thr Glu Gly Pro Ile Asp Leu Val Phe Val Ile
1     5     10     15

Asp Gly Ser Lys Ser Leu Gly Glu Glu Asn Phe Glu Val Val Lys Gin
20    25    30

Phe

<210> SEQ ID NO 293
<211> LENGTH: 193
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (35)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 293

Gly Trp Glu Thr Leu Pro Lys Asp Val Cys Lys Ser Thr His His
1    5      10     15
Gly Cys Glu His Ile Cys Val Asn Gly Asn Ser Tyr Ile Cys Lys
20   25     30
Cys Ser Xaa Gly Phe Val Leu Ala Glu Asp Gly Arg Arg Cys Lys Lys
35   40     45
Cys Thr Glu Gly Pro Ile Asp Leu Val Phe Val Ile Asp Gly Ser Lys
50   55     60
Ser Leu Gly Glu Glu Asn Phe Glu Val Val Lys Gln Phe Val Thr Gly
65   70     75     80
Ile Ile Asp Ser Leu Thr Ile Ser Pro Lys Ala Ala Arg Val Gly Leu
85   90     95
Leu Gln Tyr Ser Thr Glu Val His Thr Glu Phe Thr Leu Arg Asn Phe
100  105    110
Asn Ser Ala Lys Asp Met Lys Ala Val Ala Met Ser Lys Tyr Met
115  120    125
Gly Lys Gly Ser Met Thr Gly Leu Ala Leu Lys His Met Phe Glu Arg
130  135    140
Ser Phe Thr Gin Gly Glu Gly Ala Arg Pro Phe Pro Gin Gly Cys Pro
145  150    155    160
Glu Gln Pro Leu Cys Ser Pro Thr Asp Gly Leu Arg Met Thr Ser Pro
165  170    175
Ser Gly Pro Val Lys ProArg Pro Met Val Ser Leu Cys Ser Met Leu Leu
180  185    190
Gly

<210> SEQ ID NO: 294
<211> LENGTH: 193
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 294

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

Pro Leu Glu Glu Lys His Asp Glu Cys Lys Glu Asn Leu Ile Met

Phe Glu Asn Leu Ala Asn Glu Glu Val Arg Lys Leu Thr Gin Arg Leu

Glu Glu Met Thr Gin Arg Met Glu Ala Leu Glu Asn Arg Leu Arg Tyr

Arg

<210> SEQ ID NO 295
<211> LENGTH: 60
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 295

Met Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Ala

His Phe Ser Pro Gin Leu Cys Ile Arg Asn Ala Val Pro Leu Gly Thr

Thr Ala Lys Glu Glu Met Glu Arg Phe Trp Asn Lys Asn Ile Gly Ser

Asn Arg Pro Leu Ser Ser Pro His Ile Thr Ile Tyr Ser

<210> SEQ ID NO 296
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 296

Val Phe Pro Leu Met Tyr His Thr Trp Asn Gly Ile Arg His Leu Met

Trp Asp Leu Gly Lys Gly Leu Lys Ile Pro Gin Leu Tyr Gin Ser Gly

<210> SEQ ID NO 297
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 297

Met Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Ala

His

<210> SEQ ID NO 298
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 298

Val Lys Ser Leu Cys Leu Gly Pro Ala Leu Ile His Thr Ala Lys Phe

Ala Leu
<210> SEQ ID NO 299
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

400> SEQUENCE: 299

Val Phe Pro Leu Met Tyr His Thr Trp Asn Gly Ile Arg His Leu Met
1    5    10    15
Trp Asp Leu Gly Lys Gly Leu
20

<210> SEQ ID NO 300
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

400> SEQUENCE: 300

Arg Val Trp Asp Val Arg Pro Phe Ala Pro Lys Glu Arg Cys Val Lys
1    5    10    15
Ile Phe Gln Gly Asn Val
20

<210> SEQ ID NO 301
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

400> SEQUENCE: 301

His Asn Phe Glu Lys Asn Leu Leu Arg Cys Ser Trp Ser Pro Asp Gly
1    5    10    15
Ser Lys Ile Ala Ala Gly Ser Ala Asp Arg Phe Val Tyr Val
20    25    30

<210> SEQ ID NO 302
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

400> SEQUENCE: 302

Trp Asp Thr Thr Ser Arg Arg Ile Leu Tyr Lys Leu Pro Gly His Ala
1    5    10    15
Gly Ser Ile Asn Glu Val Ala Phe His Pro Asp Glu Pro Ile
20    25    30

<210> SEQ ID NO 303
<211> LENGTH: 141
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

400> SEQUENCE: 303

Tyr Gln Gly Leu Gly Leu Arg Gln Asn Lys Leu Thr Tyr Thr Met Arg
1    5    10    15
Gly His Ala Asp Ser Val Thr Gly Leu Ser Leu Ser Leu Ser Glu Gly Ser
20    25    30
Tyr Leu Leu Ser Asn Ala Met Asp Thr Val Arg Val Trp Asp Val
35    40
Arg Pro Phe Ala Pro Lys Glu Arg Cys Val Lys Ile Phe Gln Gly Asn
50    55    60
Val His Asn Phe Glu Lys Asn Leu Leu Arg Cys Ser Trp Ser Pro Asp
Gly Ser Lys Ile Ala Ala Gly Ser Ala Asp Arg Phe Val Tyr Val Trp
85 90 95
Asp Thr Thr Ser Arg Arg Ile Leu Tyr Lys Leu Pro Gly His Ala Gly
100 105 110
Ser Ile Asn Glu Val Ala Phe His Pro Asp Glu Pro Ile Ile Ile Ser
115 120 125
 Ala Ser Ser Asp Lys Arg Leu Tyr Met Gly Glu Ile Gln
130 135 140

<210> SEQ ID NO 304
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 304
Arg Lys Lys Ala Ala Ile Gln Thr Phe Gln Asn Thr Tyr Glu Val Leu
1 5 10 15
 Ala Val Thr Phe Asn Arg Thr Ser Asp Glu Ile Ile Ser Gly Gly Ile
20 25 30
 Asp Asn Asp Ile Lys Val Trp Asp Cys Ala Arg Thr Ser
35 40 45

<210> SEQ ID NO 305
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 305
Val Arg Gly Arg Thr Val Leu Arg Pro Gly Leu Asp Ala Glu Pro Glu
1 5 10 15
 Leu Ser Pro Glu
20

<210> SEQ ID NO 306
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 306
Glu Gln Arg Val Leu Glu Arg Lys Leu Lys Gly Glu Arg Lys Lys Glu
1 5 10 15
 Glu Arg Gln

<210> SEQ ID NO 307
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 307
Arg Leu Arg Glu Ala Gly Leu Val Ala Gln His Pro Pro
1 5 10
Gly Arg Ile Pro Ala Pro Ala Pro Ser Val Pro Ala Gly Pro Asp Ser
1 5 10 15

Arg

 Ala Arg Arg Ser Gly Ala Glu Ala Trp Asp Tyr: Leu Cys Arg Trp
1 5 10 15
 Ala Gln Lys His Lys Asn Trp Arg Phe Glu Lys Thr Arg Gln Thr Trp
20 25 30
 Leu Leu Leu His Met Tyr Asp Ser Asp Lys Val Pro Asp Glu His Phe
35 40 45
 Ser Thr Leu Leu Ala Tyr Leu G1y Leu G1y Leu Gly Arg
50 55 60

Thr Gly Cys Val Leu Val Leu Ser Arg Asn Phe Val Gln Tyr Ala Cys
1 5 10 15
 Phe Gly Leu Phe Gly Ile Ile Ala Leu G1n Thr Ile Ala Tyr Ser Ile
20 25 30
 Leu Trp Asp Leu Lys Phe Leu Met Arg Asn
35 40

Ser Arg Ser Glu Gly Lys Ser Met Phe Ala Gly Val Pro Thr Met Arg
1 5 10 15
 Glu Ser Ser Pro Lys Gln Tyr Met Gin Leu Gly G1y Arg Val Leu Leu
20 25 30
 Val Leu Met Phe Met Thr Leu Leu His Phe Asp Ala Ser Phe Phe Ser
35 40 45
 Ile Val Gln Asn Ile Val Gly
50 55

Gly Thr Ala Glu Asp Phe Ala Asp Gin Phe Leu Arg Val Thr Lys Gin
1 5 10 15
 Tyr Leu Pro His Val Ala Arg Leu Cys Leu Ile Ser Thr Phe Leu Glu
20 25 30
Asp Gly Ile Arg Met Trp Phe Gln Trp Ser Glu Gln Arg Asp Tyr Ile
35 40 45
Asp Thr Thr Trp Asn Cys Gly Tyr Leu Leu Ala Ser
50 55 60

<210> SEQ ID NO 313
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 313

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

<210> SEQ ID NO 314
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 314

Leu Met Arg Asn Glu Ser Arg Ser
1 5

<210> SEQ ID NO 315
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 315

Ala Ser Phe Leu Leu Ser Arg Thr Ser Trp Gly Thr Ala
1 5 10

<210> SEQ ID NO 316
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 316

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

<210> SEQ ID NO 317
<211> LENGTH: 72
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 317

Pro Ser Phe Thr Leu Thr Pro Ala Ser Phe Leu Leu Ser Arg Thr Ser
1 5 10 15
Trp Gly Thr Ala Leu Met Ile Leu Val Ala Ile Gly Phe Lys Thr Lys
20 25 30
Leu Ala Ala Leu Thr Leu Val Val Leu Phe Ala Ile Asn Val Tyr
35 40
Phe Asn Ala Phe Thr Thr Phe Thr Ile Pro Val Tyr Lys Pro Met His Asp Phe
50 55 60
Leu Lys Tyr Asp Phe Phe Gln Thr
<210> SEQ ID NO 319
<211> LENGTH: 236
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (115)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 319

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

Tyr Ala Cys Phe Gly Leu Phe Gly Ile Ile Ala Leu Gln Thr Ile Ala
 130    135    140
Tyr Ser Ile Leu Trp Asp Leu Lys Phe Leu Met Arg Asn Leu Ala Leu
 145    150    155    160
Gly Gly Gly Leu Leu Leu Leu Ala Glu Ser Arg Ser Gly Gly Lys
 165    170    175
Ser Met Phe Ala Gly Val Pro Thr Met Arg Glu Ser Ser Pro Lys Gln
 180    185    190
Tyr Met Gin Leu Gly Gly Arg Val Leu Leu Val Leu Met Phe Met Thr
 195    200    205
Leu Leu His Phe Asp Ala Ser Phe Phe Ser Ile Val Gin Asn Ile Val
 210    215    220
Gly His Ser Ser Asp Asp Phe Ser Gly His Trp Phe
 225    230    235

<210> SEQ ID NO 319
<211> LENGTH: 114
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (2)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 319
Gly Xaa Ser Arg Arg Arg Ala Leu Pro Val Glu Ala Ala Ala Ala Gly Ala
1 5 10 15
Gly Ala Asp Gly Arg Glu Pro Ala Ser Glu Arg Arg Ser Arg Ala Glu
20 25 30
Pro Pro Ala Val Ala Met Gly Gin Asn Asp Leu Met Gly Thr Ala Glu
35 40 45
Asp Phe Ala Asp Gin Phe Leu Arg Val Thr Lys Gin Tyr Leu Pro His
50 55 60
Val Ala Arg Leu Cys Leu Ile Ser Thr Phe Leu Glu Asp Gly Ile Arg
65 70 75 80
Met Trp Phe Gin Trp Ser Glu Gin Arg Asp Tyr Ile Asp Thr Thr Trp
85 90 95
Asn Cys Gly Tyr Leu Leu Ala Ser Ser Phe Val Phe Leu Asn Leu Leu
100 105 110

Gly Xaa

<210> SEQ ID NO 320
<211> LENGTH: 63
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 320
Trp Val Phe Leu Phe Leu Leu Ala Leu Gly Gly Leu Gly Pro Asp Ser
1 5 10 15
Gly Arg Cys Leu Cys Arg Gly Arg Ile Ser Gly Ile Tyr Gin Leu
20 25 30
Ile Leu Ala Lys Gin Phe Leu Arg Phe Phe Cys Phe Met Trp Glu Thr
35 40 45
Asp Leu Asn Leu Ile Leu Cys Cys Ile Leu Tyr Leu Ser Cys Val
50 55 60

<210> SEQ ID NO 321
<211> LENGTH: 154
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 321
Ser Met Ser Ala Leu Thr Arg Leu Ala Ser Phe Ala Arg Val Gly Gly
1 5 10 15
Arg Leu Phe Arg Ser Gly Cys Ala Arg Thr Ala Gly Asp Gly Gly Val
20 25 30
Arg His Ala Gly Gly Gly Val His Ile Glu Pro Arg Tyr Arg Gin Phe
35 40 45
Pro Gin Leu Thr Arg Ser Gin Val Phe Gin Ser Glu Phe Ser Gly
50 55 60
Leu Met Trp Phe Trp Ile Leu Trp Arg Phe Trp His Asp Ser Gin Leu
65 70 75 80
Val Leu Gly His Phe Pro Tyr Pro Asp Pro Ser Gin Trp Thr Asp Glu
85 90 95
Glu Leu Gly Ile Pro Pro Asp Asp Glu Asp
100 105

<210> SEQ ID NO 322
<211> LENGTH: 20
<table>
<thead>
<tr>
<th>Phe</th>
<th>Ile</th>
<th>Ser</th>
<th>Phe</th>
<th>Ala</th>
<th>Asn</th>
<th>Ser</th>
<th>Arg</th>
<th>Ser</th>
<th>Ser</th>
<th>Glu</th>
<th>Asp</th>
<th>Thr</th>
<th>Lys</th>
<th>Gln</th>
<th>Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td>10</td>
<td>15</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>Met</th>
<th>Ser</th>
<th>Ser</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asp</th>
<th>Pro</th>
<th>Arg</th>
<th>Arg</th>
<th>Pro</th>
<th>Asn</th>
<th>Lys</th>
<th>Val</th>
<th>Leu</th>
<th>Arg</th>
<th>Tyr</th>
<th>Lys</th>
<th>Pro</th>
<th>Pro</th>
<th>Pro</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td>10</td>
<td>15</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>Glu</th>
<th>Cys</th>
<th>Asn</th>
<th>Pro</th>
<th>Ala</th>
<th>Leu</th>
<th>Asp</th>
<th>Pro</th>
<th>Thr</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</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>Asp</th>
<th>Tyr</th>
<th>Met</th>
<th>Asn</th>
<th>Leu</th>
<th>Leu</th>
<th>Gly</th>
<th>Met</th>
<th>Ile</th>
<th>Phe</th>
<th>Ser</th>
<th>Met</th>
<th>Cys</th>
<th>Gly</th>
<th>Leu</th>
<th>Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td>10</td>
<td></td>
<td>15</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>Leu</th>
<th>Lys</th>
<th>Leu</th>
<th>Lys</th>
<th>Trp</th>
<th>Cys</th>
<th>Ala</th>
<th>Trp</th>
<th>Val</th>
<th>Ala</th>
<th>Val</th>
<th>Tyr</th>
<th>Cys</th>
<th>Ser</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</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>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asp</th>
<th>Gly</th>
<th>Asp</th>
<th>Gly</th>
<th>Asp</th>
<th>Val</th>
<th>Lys</th>
<th>Leu</th>
<th>Gly</th>
<th>Thr</th>
<th>Leu</th>
<th>Gly</th>
<th>Ser</th>
<th>Gly</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td>10</td>
<td>15</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>Glu</th>
<th>Ser</th>
<th>Ser</th>
<th>Asn</th>
<th>Gly</th>
<th>Gly</th>
<th>Ser</th>
<th>Gly</th>
<th>Ser</th>
<th>Pro</th>
<th>Gly</th>
<th>Asp</th>
<th>Ala</th>
<th>Gly</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</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>Ala</th>
<th>Ala</th>
<th>Xaa</th>
<th>Gly</th>
<th>Gly</th>
<th>Gly</th>
<th>Gly</th>
<th>Trp</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Leu</th>
<th>Ala</th>
<th>Leu</th>
<th>Leu</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</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>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Gly</th>
<th>Gly</th>
<th>Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 327
<211> LENGTH: 62
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (45)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 327
Ser Thr His Ala Ser Gly Ala Val Met Ala Ala Gly Asp Gly Asp
1 5 10 15
Val Lys Leu Gly Thr Leu Gly Ser Gly Ser Glu Ser Ser Asn Asp Gly
20 25 30
Gly Ser Glu Ser Pro Gly Asp Ala Gly Ala Ala Ala Xaa Gly Gly Gly
35 40 45
Trp Ala Ala Ala Leu Ala Leu Leu Thr Gly Gly Gly Gly Glu
50 55 60

<210> SEQ ID NO 328
<211> LENGTH: 177
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (26)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (84)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (111)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 328
Ala Ala Asp Asn Tyr Gly Ile Pro Arg Ala Cys Arg Asn Ser Ala Arg
1 5 10 15
Ser Tyr Gly Ala Ala Trp Leu Leu Leu Xaa Pro Ala Gly Ser Ser Arg
20 25 30
Val Glu Pro Thr Gin Asp Ile Ser Ile Ser Asp Gin Leu Gly Gly Gin
35 40 45
Asp Val Pro Val Phe Arg Asn Leu Ser Leu Val Val Gly Val Gly
50 55 60
Ala Val Phe Ser Leu Phe His Leu Gly Thr Arg Gly Arg Arg Arg
65 70 75 80
Pro His Ala Xaa Glu Pro Gly Glu His Thr Pro Leu Leu Ala Pro Ala
85 90 95
Thr Ala Gin Pro Leu Leu Leu Thr Lys His Thr Leu Arg Glu Xaa Ala
100 105 110
Phe Tyr Gin Val Gly Ile Tyr Met Thr Thr Arg Leu Ile Val Asn
115 120 125
Leu Ser Gin Thr Tyr Met Ala Met Tyr Leu Thr Ser Leu His Leu
130 135 140
Pro Lys Lys Phe Ile Ala Thr Ile Pro Leu Val Met Tyr Leu Ser Gly
Phe Leu Ser Ser Phe Leu Met Lys Pro Ile Asn Lys Cys Ile Gly Arg
Asn

<210> SEQ ID NO 329
<211> LENGTH: 79
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (7)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 329

Cys Thr Leu Ala Met Trp Xaa Leu Gly His Cys Asp Arg Arg Cys
1  5  10  15
Thr Gly Arg Lys Leu Ala Arg Leu Gly Leu Val Arg Cys Leu Arg Leu
20 25  30
Gly His Arg Phe Gly Gly Leu Val Leu Ser Pro Val Gly Lys Gin Tyr
35  40  45
Ala Ser Pro Ala Asp Arg Gin Leu Val Ala Gin Ser Gly Val Ala Val
50  55  60
Ile Asp Cys Ser Trp Ala Arg Leu Asp Glu Thr Pro Phe Gly Lys
65  70  75

<210> SEQ ID NO 330
<211> LENGTH: 72
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 330

Ser Gly Arg Gly Ala Arg Ser Asp Val Thr Ala Met Ala Gly Ile Lys
1  5  10  15
Ala Leu Ile Ser Leu Ser Phe Gly Gly Ala Ile Gly Leu Met Phe Leu
20 25  30
Met Leu Gly Cys Ala Leu Pro Ile Tyr Asn Lys Tyr Trp Pro Leu Phe
35  40  45
Val Leu Phe Phe Tyr Ile Leu Ser Pro Ile Pro Tyr Cys Ile Ala Arg
50  55  60
Arg Leu Val Asp Asp Thr Asp Ala
65  70

<210> SEQ ID NO 331
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (5)
<223> OTHER INFORMATION: Xaa equals any of the L-amino acids commonly found in naturally occurring proteins

<400> SEQUENCE: 331

Ala Arg Val Arg Xaa Arg Gly Ala Leu Ser Leu Ser Val Gly Ala Ala
1  5  10  15
Cys Gly Leu Val Ala Leu Trp Gln Arg Arg Gln Asp Ser Gly Thr
20 25  30
<table>
<thead>
<tr>
<th>&lt;210&gt; SEQ ID NO 332</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;211&gt; LENGTH: 45</td>
</tr>
<tr>
<td>&lt;212&gt; TYPE: PRT</td>
</tr>
<tr>
<td>&lt;213&gt; ORGANISM: Homo sapiens</td>
</tr>
<tr>
<td>&lt;400&gt; SEQUENCE: 332</td>
</tr>
<tr>
<td>Leu Ser Asn Ala Gln Asn Trp Gly Met Gln Arg Ala Thr Asn Val</td>
</tr>
<tr>
<td>Thr Tyr Gln Ala His His Val Ser Arg Asn Lys Arg Gly Gin Val</td>
</tr>
<tr>
<td>Gly Thr Arg Gly Gly Phe Arg Gly Cys Thr Val Trp Leu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;210&gt; SEQ ID NO 333</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;211&gt; LENGTH: 38</td>
</tr>
<tr>
<td>&lt;212&gt; TYPE: PRT</td>
</tr>
<tr>
<td>&lt;213&gt; ORGANISM: Homo sapiens</td>
</tr>
<tr>
<td>&lt;400&gt; SEQUENCE: 333</td>
</tr>
<tr>
<td>Val Ser Met Ala Leu Glu Glu Tyr Leu Val Cys His Gly Ile Pro Cys</td>
</tr>
<tr>
<td>Tyr Thr Leu Asp Gly Asp Ile Arg Gin Gly Leu Asn Lys Asn Leu</td>
</tr>
<tr>
<td>Gly Phe Ser Pro Glu Asp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;210&gt; SEQ ID NO 334</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;211&gt; LENGTH: 39</td>
</tr>
<tr>
<td>&lt;212&gt; TYPE: PRT</td>
</tr>
<tr>
<td>&lt;213&gt; ORGANISM: Homo sapiens</td>
</tr>
<tr>
<td>&lt;400&gt; SEQUENCE: 334</td>
</tr>
<tr>
<td>Thr Gin Asp Arg Asn Ala Arg Gin Ile His Glu Gly Ala Ser Leu</td>
</tr>
<tr>
<td>Pro Phe Phe Glu Val Phe Val Asp Ala Pro Leu His Val Cys Glu Gin</td>
</tr>
<tr>
<td>Arg Asp Val Lys Gly Leu Tyr</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;210&gt; SEQ ID NO 335</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;211&gt; LENGTH: 40</td>
</tr>
<tr>
<td>&lt;212&gt; TYPE: PRT</td>
</tr>
<tr>
<td>&lt;213&gt; ORGANISM: Homo sapiens</td>
</tr>
<tr>
<td>&lt;400&gt; SEQUENCE: 335</td>
</tr>
<tr>
<td>Phe Thr Gly Ile Asp Ser Glu Tyr Glu Lys Pro Glu Ala Pro Glu Leu</td>
</tr>
<tr>
<td>Val Leu Lys Thr Ser Asp Cys Asp Val Asn Asp Cys Val Gin Gin Val</td>
</tr>
<tr>
<td>Val Glu Leu Leu Gin Glu Arg Asp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;210&gt; SEQ ID NO 336</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;211&gt; LENGTH: 41</td>
</tr>
<tr>
<td>&lt;212&gt; TYPE: PRT</td>
</tr>
<tr>
<td>&lt;213&gt; ORGANISM: Homo sapiens</td>
</tr>
<tr>
<td>&lt;400&gt; SEQUENCE: 336</td>
</tr>
<tr>
<td>Ala</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Val</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 337
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

Val Pro Ile Val Leu Thr Ala Thr His Glu Asp Lys Glu Arg Leu Asp Gly Cys Thr Ala Phe Ala Leu Met Tyr Gly Glu Arg Val

<210> SEQ ID NO 338
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 338

Ile Gly Gly Asp Leu Gln Val Leu Asp Arg Val Tyr Trp Asn Asp Gly Leu Asp Gly Tyr Arg Leu Thr Pro Cys Thr Glu Leu Lys Gly Lys Phe Lys Asp Met Asn Ala Asp Ala Val

<210> SEQ ID NO 339
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 339

Gly His Ala Leu Leu Met Gln Thr His Lys Gln Leu Leu Glu Arg Gly Tyr Arg Pro Val Leu Leu Leu His Pro Leu Gly Gly Trp Thr Lys Asp Arg Asp Val

<210> SEQ ID NO 340
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 340

Met Tyr Ala Gly Pro Thr Glu Val Gln Trp His Cys Arg Ala Arg Met Val Ala Gly Ala Asn Phe Tyr Ile Val Gly Arg Asp Pro Ala Gly Met Pro His Pro Glu Thr Gly Lys Asp Leu
Leu Thr Met Ala Pro Gly Leu Ile Thr Leu Glu Ile Val Pro Phe Arg
1  5  10  15

Val Ala Ala Tyr Asn Lys Lys Lys Arg Met Asp Tyr Tyr Asp Ser
20 25 30

Glu His

Gly Phe Met Ala Pro Lys Ala Trp Thr Val Leu Thr Glu Tyr Tyr Lys
1  5  10  15

Ser Leu Glu

Arg Ile Thr Asp Asn Pro Gly Gly Tyr Ile Lys Thr Thr Ala Val Glu Ile Xaa Tyr Asp
1  5  10  15

Gly Ser Tyr Gly Tyr Ile Lys Thr Thr Ala Val Glu Ile Xaa Tyr Asp
20 25 30

Ser Leu Lys Leu Lys Asp Ser Ser Leu Gly Ala Pro Ser Arg Pro Ile
35 40

Glu Asp Asp Gin Glu Val Tyr Asp Val Ala Glu Gin Asp Asp Ile
55 60

Ser Ser His Ser Gin Ser Gin Ser Gin Asp Gly Ile Gly Ile Phe Pro Pro Pro Pro Pro
65 70 75 80

Asp Asp Asp Ile Tyr Asp Gly Ile Glu Glu Glu Glu Asp Ala Asp Asp Gly
85 90 95

Phe Pro Ala Pro Pro Lys Glu Leu Asp Met Gly Asp Glu Val Tyr Asp
100 105 110

Asp Val Asp Thr Ser Asp Phe Pro Val Ser Ser Ala Glu Met Ser Gin
115 120 125

Gly Thr Asn Val Gly Lys Ala Lys Thr Glu Glu Lys Asp Leu Lys Lys
130 135 140
1. An isolated antibody or portion thereof that specifically binds to a protein whose sequence consists of amino acid residues 29 to 99 of SEQ ID NO:142.

2. The isolated antibody or portion thereof of claim 1, wherein said protein specifically bound by said antibody or portion thereof is glycosylated.

3. The isolated antibody or portion thereof of claim 1 which is a monoclonal antibody.

4. The isolated antibody or portion thereof of claim 1 which is a polyclonal antibody.

5. The isolated antibody or portion thereof of claim 1 which is a chimeric antibody.

6. The isolated antibody or portion thereof of claim 1 which is a humanized antibody.

7. The isolated antibody or portion thereof of claim 1 which is a Fab fragment.

8. The isolated antibody or portion thereof of claim 1 which is labeled.

9. The isolated antibody or portion thereof of claim 1 which is labeled.

10. The isolated antibody of claim 9, wherein the label is selected from the group consisting of:
    (a) an enzyme label;
    (b) a radioisotope; and
    (c) a fluorescent label.

11. An isolated cell that produces the isolated antibody of claim 1.

12. A hybridoma that produces the isolated antibody of claim 1.

13. A hybridoma that produces the isolated antibody of claim 3.

14. An isolated antibody produced by immunizing an animal with a protein whose sequence comprises of amino acid residues 29 to 99 of SEQ ID NO:142, wherein said antibody or portion thereof specifically binds to the protein of SEQ ID NO:142.

15. The isolated antibody of claim 14, wherein said protein specifically bound by said antibody is glycosylated.

16. The isolated antibody of claim 14 which is a monoclonal antibody.

17. The isolated antibody of claim 14 which is a polyclonal antibody.

18. The isolated antibody of claim 14 which is labeled.

19. The isolated antibody of claim 18, wherein the label is selected from the group consisting of:
    (d) an enzyme label;
    (e) a radioisotope; and
    (f) a fluorescent label.

20. An isolated cell that produces the isolated antibody of claim 14.

21. A hybridoma that produces the isolated antibody of claim 14.

22. A hybridoma that produces the isolated antibody of claim 16.

23. An isolated antibody or portion thereof that specifically binds to a protein whose sequence consists of amino acid residues 1 to 99 of SEQ ID NO:142.

24. The isolated antibody or portion thereof of claim 23 which is a polyclonal antibody.

25. The isolated antibody or portion thereof of claim 23 which is a monoclonal antibody.

26. The isolated antibody or portion thereof of claim 23 which is a chimeric antibody.

27. The isolated antibody or portion thereof of claim 23 which is a humanized antibody.

28. The isolated antibody or portion thereof of claim 23 which is a Fab fragment.

29. The isolated antibody or portion thereof of claim 23 which is labeled.

30. The isolated antibody or portion thereof of claim 23 which is labeled.

31. The isolated antibody or portion thereof of claim 23 which is labeled.

32. The isolated antibody of claim 31, wherein the label is selected from the group consisting of:
    (a) an enzyme label;
    (b) a radioisotope; and
    (c) a fluorescent label.
33. An isolated cell that produces the isolated antibody of claim 23.
34. A hybridoma that produces the isolated antibody of claim 23.
35. A hybridoma that produces the isolated antibody of claim 25.
36. An isolated antibody or portion thereof that specifically binds to a protein whose sequence consists of the amino acid sequence of the secreted polypeptide encoded by the HFEA41 cDNA contained in ATCC Deposit Number 97923.
37. The isolated antibody or portion thereof of claim 36, wherein said protein specifically bound by said antibody or portion thereof is glycosylated.
38. The isolated antibody or portion thereof of claim 36 which is a monoclonal antibody.
39. The isolated antibody or portion thereof of claim 36 which is a polyclonal antibody.
40. The isolated antibody or portion thereof of claim 36 which is a chimeric antibody.
41. The isolated antibody or portion thereof of claim 36 which is a humanized antibody.
42. The isolated antibody or portion thereof of claim 36 which is a Fab fragment.
43. The isolated antibody or portion thereof of claim 36 which is labeled.
44. The isolated antibody or portion thereof of claim 36 which is labeled.
45. The isolated antibody of claim 44, wherein the label is selected from the group consisting of:

(a) an enzyme label;
(b) a radioisotope; and
(c) a fluorescent label.
46. An isolated cell that produces the isolated antibody of claim 36.
47. A hybridoma that produces the isolated antibody of claim 36.
48. A hybridoma that produces the isolated antibody of claim 38.
49. An isolated antibody produced by immunizing an animal with a protein whose sequence comprises the amino acid sequence of the secreted polypeptide encoded by the HFEA41 cDNA contained in ATCC Deposit Number 97923, wherein said antibody or portion thereof specifically binds to the polypeptide encoded by the HFEA41 cDNA contained in ATCC Deposit Number 97923.
50. The isolated antibody of claim 49, wherein said protein specifically bound by said antibody is glycosylated.
51. The isolated antibody of claim 49 which is a monoclonal antibody.
52. The isolated antibody of claim 49 which is a polyclonal antibody.
53. The isolated antibody of claim 49 which is labeled.
54. The isolated antibody of claim 53, wherein the label is selected from the group consisting of:

(g) an enzyme label;
(h) a radioisotope; and
(i) a fluorescent label.
55. An isolated cell that produces the isolated antibody of claim 49.
56. A hybridoma that produces the isolated antibody of claim 49.
57. A hybridoma that produces the isolated antibody of claim 51.
58. An isolated antibody or portion thereof that specifically binds to a protein whose sequence consists of the amino acid sequence of the full-length polypeptide encoded by the HFEA41 cDNA contained in ATCC Deposit Number 97923.
59. The isolated antibody or portion thereof of claim 58 wherein said protein specifically bound by said antibody or portion thereof is glycosylated.
60. The isolated antibody or portion thereof of claim 58 which is a monoclonal antibody.
61. The isolated antibody or portion thereof of claim 58 which is a polyclonal antibody.
62. The isolated antibody or portion thereof of claim 58 which is a chimeric antibody.
63. The isolated antibody or portion thereof of claim 58 which is a humanized antibody.
64. The isolated antibody or portion thereof of claim 58 which is a human antibody.
65. The isolated antibody or portion thereof of claim 58 which is a Fab fragment.
66. The isolated antibody or portion thereof of claim 58 which is labeled.
67. The isolated antibody of claim 66 wherein the label is selected from the group consisting of:

(a) an enzyme label;
(a radioisotope; and
(c) a fluorescent label.
68. An isolated cell that produces the isolated antibody of claim 58.
69. A hybridoma that produces the isolated antibody of claim 58.
70. A hybridoma that produces the isolated antibody of claim 60.

* * * * *