

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
6 January 2005 (06.01.2005)

PCT

(10) International Publication Number
WO 2005/000099 A2

(51) International Patent Classification⁷: **A61B**

(21) International Application Number:
PCT/US2004/018461

(22) International Filing Date: 9 June 2004 (09.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/477,291 9 June 2003 (09.06.2003) US

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): **GENZYME CORPORATION** [US/US]; 500 Kendall Street, Cambridge, MA 02142 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): **ROBERTS, Bruce, L.** [US/US]; 108 Pinehill Road, Southboro, MA 01772 (US).

(74) Agents: **KONSKI, Antoinette, F.** et al.; Bingham McCutchen LLP, Three Embarcadero Center, Suite 1800, San Francisco, CA 94111-4067 (US).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BLOOD FACTOR DOMAINS (BFDs)

(57) Abstract: The present invention provides methods and compositions for detecting, diagnosing, prognosing and monitoring the progress of diseases. Further provided are methods for screening to identify agonists and antagonists of antigens associated with these diseases.

WO 2005/000099 A2

BLOOD FACTOR DOMAINS (BFDs)

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit under 35 U.S.C. § 119(e) to the following U.S. Provisional Application No. 60/477,291, filed June 9, 2003, the contents of which hereby incorporated by reference into the present disclosure.

TECHNICAL FIELD

This invention is in the field of immunology and immunotherapy. In particular, the present invention provides compositions that act as cytokines and/or chemokines which useful in modulating the immune response.

10

BACKGROUND

The mammalian immune system comprises two types of antigen-specific cells: B cells and T cells. B cells synthesize both membrane-bound and secreted antibody. T cells can be characterized phenotypically by the manner in which they recognize antigen, by their cell surface markers, and by their secreted products. T cells express distinctive membrane molecules. Included among 15 these are the T cell antigen receptor (TCR), which appears on the cell surface in association with CD3; and accessory molecules such as CD5, CD28 and CD45R. Subpopulations of T cells can be distinguished by the presence of additional membrane molecules. Thus, for example, T cells that express CD4 recognize 20 antigen associated with Class II MHC molecules and generally function as helper cells whose roles include enhancement of antibody production by B cells, while T cells that express CD8 recognize antigen associated with Class I MHC molecules and generally function as cytotoxic cells.

Immune cells recognize discrete sites, known as epitopes or antigenic determinants, on the antigen. Epitopes are regions of an immunogen or antigen that bind to antigen-specific membrane-bound receptors on immune cells or to their soluble counterparts, such as antibodies. Both membrane-bound antibody
5 on the surface of a B lymphocyte and secreted antibody recognize soluble antigen. Unlike B cells, which recognize soluble antigen, T cells recognize antigen only when the antigen is associated with self major histocompatibility complex (MHC) gene products on the surface of an antigen presenting cell. This antigen can be displayed together with MHC molecules on the surface of antigen-
10 presenting cells or on virus-infected cells, cancer cells, and grafts.

Disease states can result from invasion by a pathogenic organisms, including bacterial, viral, and protozoan pathogens, and subsequent inefficient or ineffective immune response to the invader. Disease states can also result from the activation of self-reactive T lymphocytes, from the activation of T
15 lymphocytes that provoke allergic reactions, or from the activation of autoreactive T lymphocytes following certain bacterial and parasitic infections, which can produce antigens that mimic human protein, rendering these protein "autoantigens." These diseases include, but are not limited to, the autoimmune diseases, autoimmune disorders that occur as a secondary event to infection with
20 certain bacteria or parasites, T cell mediated allergies, and certain skin diseases such as psoriasis and vasculitis. Furthermore, undesired rejection of a foreign antigen can result in graft rejection or even infertility, and such rejection may be due to activation of specific T lymphocyte populations.

Small, soluble proteins have also been shown to critical to the immune
25 response. Although lymphocytes, macrophages, and granulocytes play a major role in the response, the soluble proteins, known as lymphokines, cytokines, or monokines, and secondly, chemokines, are now known to play a critical role.

Cytokines exhibit a wide variety of functions. A hallmark feature is their ability to elicit chemotactic migration of distinct cell types, including
30 polymorphonuclear cells and macrophages. Many cytokines have pro-inflammatory activity and are involved in multiple steps during inflammatory reactions. They also have been implicated in a number of physiological and

disease conditions, including lymphocyte trafficking, wound healing, hematopoietic regulation and immunological disorders such as allergy, asthma and arthritis.

Chemokines are a family of small cytokines thought to mediate the directional migration of specific target populations of leukocytes along concentration gradients through the endothelial cell layer to the site of lesion. Baggiolini *et al.* (1994) *Adv. Immunol.* **55**: 97-179. This cascade of events, however can display a high degree of specificity in relation to the inflammatory stimulus, the stage of the inflammatory response and the tissue or organ involved. Butcher *et al.* (1991) *Cell* **67**:1033-1036.

DISCLOSURE OF THE INVENTION

This invention provides compositions and methods for modulating an immune response. In one aspect, the immune response is a disease or condition related to a population of CD4⁺ or CD8⁺ T cells. The compositions identified in Table 1, *infra*, are immune response modulators exhibit chemokine and cytokine biological activity and therefore are useful to up-regulate immune response and alternatively, to down-regulate the same. Also provided are therapeutic agents that enhance or diminish the native activity of the immune modulators.

This invention provides isolated proteins and/or polypeptides useful in the methods identified herein. Further provided are polynucleotides encoding the proteins, fragments thereof, or polypeptides, (also referred to herein as gene expression product), gene delivery vehicles comprising these polynucleotides and host cells comprising these polynucleotides. Polynucleotides of the invention are intended to include DNA, cDNA, RNA and genomic DNA. Expression systems, including gene delivery vehicles such as liposomes and vectors, and host cells containing the polynucleotides are further provided by this invention. The proteins, polypeptides or fragments thereof are also useful to generate antibodies that specifically recognize and bind to these molecules. The antibodies can be polyclonal or monoclonal. These antibodies can be used to isolate protein or polypeptides expressed from the genes identified in Table 1.

The invention also provides isolated host cells and recombinant host cells that contain a gene of Table 1 or its expression product and/or fragments thereof. The cells can be prokaryotic or eukaryotic and by way of example only, can be any one or more of bacterial, yeast, animal, mammalian, human, and particular subtypes thereof, *e.g.*, stem cells, antigen presenting cells (APCs) such as dendritic cells (DCs) or T cells.

The present invention also provides proteins encoded by the polynucleotides.

Additionally, nucleic acid probes and primers that hybridize to invention polynucleotides are provided, as well as isolated nucleic acids comprising unique, expressed gene sequences.

The present invention further includes antisense oligonucleotides, antibodies, hybridoma cell lines and compositions containing same.

The present invention also provides methods of monitoring gene expression using invention polynucleotides. Also provided are compositions and methods to monitor expression of the polynucleotides and expression of the polynucleotides by detecting the expression products such as mRNA and/or polypeptides.

This invention further provides localized and systemic methods for: modulating the expression of the immunomodulatory polynucleotides and expression products, altering the activity of the proteins encoded by the polynucleotides, and treating symptoms of cancer, viral infection and autoimmune disorders.

This invention also provides a method for screening for candidate agents that modulate the expression of a polynucleotide or its complement and the expression products of the polynucleotide. The present invention also provides assays for the isolation of the ligand or ligands capable of modulating the activity of the invention polynucleotides and/or proteins.

This invention further provides assays for the identification, assessment and development of candidate agents capable of modulating the activity of the of the invention polynucleotides and/or proteins.

Further provided by this invention is a method for monitoring an immune response in a subject by assaying, at different times, the expression level of at least one gene identified in Table 1 and comparing the expression levels of the gene (transcript or expression product) to determine if expression has increased or decreased, thereby monitoring the disease or condition in the subject. A kit for
5 use in a diagnostic method or drug screen is further provided herein. The kit comprises at least one agent (*e.g.*, probe, primer or antibody) that detects expression of at least one gene identified in Table 1 and instructions for use.

In addition, the invention provides methods for active immunotherapy, such as, inducing an immune response in a subject by delivering the proteins,
10 polypeptides and fragments thereof, as described herein, to the subject.

The polynucleotides and expression products having immunomodulatory activity are set forth in Table 1, below.

TABLE 1

Description	Probe or Primer	Unigene Ref.	LocusLink Id.	GenBank References	Name
206 aa 4 cysteine	GTGGTATGGC SEQ ID NO: 1	Hs 25199	51537	NM_016498.2 NP_057582	HSPC242 Hypothetical SEQ ID NOS 2 & 3
283 aa 5 cysteine	CTAATAAACT SEQ ID NO: 4	Hs 279583	51108	NM_057109 NP_057109	CGI-81 Protein SEQ ID NOS 5 & 6
228 aa 4 cysteine	GCTCCCTACG SEQ ID NO: 7	Hs 185705	10148	NM_005755.2 NP_005746	EBV Induced Gene 3 SEQ ID NOS 8 & 9
227 aa 5 cysteines predicted N-terminal signal peptide	TGTATTTTGA SEQ ID NO: 10	Hs 29882	10447	NM_014888.1 NP_055703	FAM3C Family with sequence similarity 3, member C SEQ ID NOS 11 & 12
86 aa 5 cysteine	GTCCCCCACC SEQ ID NO: 13	Hs 380427	7940	NM_007161 NP_009092	LST-1 membrane receptor alternative spliced soluble form SEQ ID NOS 14 & 15
277 aa 8 cysteine	ACCTTGAAAGT SEQ ID NO: 16	Hs 29352	7130	NM_007115.2 NP_009046	TNFA induced protein 6 SEQ ID NO: 17
194 aa 9 cysteine	GGACCGAGGG SEQ ID NO: 18	Hs 13885	84246	NP_115662	MGC5309 Protein SEQ ID NOS 19 & 20
439 aa 11 cysteine	GCCGGCTTGA SEQ ID NO: 21	Hs 113277	11026	NM_006865.1 NP_006856	ILT6 soluble inhibitory receptor SEQ ID NOS 22 & 23
298 aa 6 cysteine	TTCACGTGTA SEQ ID NO: 24	Hs 109439	4969	NM_014057.2 NP_054776	Osteoglycin- induces ectopic bone formation SEQ ID NO: 25
305 aa 7 Cysteine	AATTCTGTAA SEQ ID NO: 26	Hs 241575	84572	NM_032520.2 NP_115909	CAB56184 Hypothetical SEQ ID NOS 27 & 28
293 aa 9 cysteine	GTCTACAATT SEQ ID NO: 29	Hs 289008	116150	NM_138459 NP_612468	Chromosome 6 ORF 68 SEQ ID NOS 30 & 31
323 aa 6 Cysteine	GGGCTCGGGG SEQ ID NO: 32	Hs 105927	None	NM_002975 NP_002966	Stem cell growth factor SEQ ID NOS 33 & 34

TABLE 1 (CONTINUED)

Description	Probe or Primer	Unigene Ref.	LocusLink Id.	GenBank References	Name
270 aa 1 cysteine predicted N-terminal signal peptide	CTGCTAACCC SEQ ID NO: 38	Hs.170310	51816	NM_016498.2 NP_059120 (Isoform a 511 aa) NM_016498.2 NP_803124 (Isoform b 270 aa)	CECR1 Cat eye syndrome chromosome region, candidate 1 SEQ ID NO: 39 SEQ ID NOS 40 & 41
173 aa 2 cysteines predicted N-terminal signal peptide	AGCCTGCAGA SEQ ID NO: 42	Hs.10927	56005	XP_035638	IL27w; interleukin 27 working designation SEQ ID NO: 43
142 aa 8 cysteines predicted N-terminal signal peptide	AAAAGAAATG SEQ ID NO: 44	Hs.41185	157869	NM_153225 NP_694957	FLJ40021 Hypothetical protein SEQ ID NOS 45 & 46
332 aa 7 cysteines	ATGACCTGAA SEQ ID NO: 47	Hs. 74050	2531	NM 002035	Follicular lymphoma variant SEQ ID NOS 48 & 49
160 aa 5 cysteines	TTTAATACAT SEQ ID NO: 50	Hs. 406548	84191	NM 032231	FLJ22875 Hypothetical SEQ ID NOS 51 & 52
375 aa 8 cysteines	TGTAACAAT SEQ ID NO: 53	Hs. 170040	5157	NM 006207	Platelet Derived Growth Factor Receptor Like SEQ ID NOS 54 & 55
194 aa 9 cysteines	GGACCGAGGG SEQ ID NO: 56	Hs. 13885	84246	NM 032286	MGC5309 Hypothetical SEQ ID NOS 57 & 58
355 aa 22 cysteines	TCCAACCCAG SEQ ID NO: 59	Hs. 302163	7478	NM 031933	WNT8A SEQ ID NOS 60 & 61
248 aa 5 cysteines predicted N-terminal signal peptide + transmembrane	ACTCGAATAT SEQ ID NO: 62	Hs.95655	6398	NM_003004.1 NP_002995	SECTM1; secreted and transmembrane 1; also known as K12 protein precursor type 1a SEQ ID NOS 63 & 64

MODES FOR CARRYING OUT THE INVENTION

Throughout this disclosure, various publications, patents and published patent specifications are referenced by an identifying citation. The disclosures of these publications, patents and published patent specifications are hereby
5 incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

Definitions

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of immunology, molecular biology,
10 microbiology, cell biology and recombinant DNA, which are within the skill of the art. *See, e.g.*, Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, *et al.* eds., (1987)); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A
15 PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R.I. Freshney, ed. (1987)).

As used herein, certain terms have the following defined meanings.

As used in the specification and claims, the singular form “a”, “an” and
20 “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes a plurality of cells, including mixtures thereof.

The terms “polynucleotide” and “oligonucleotide” are used interchangeably, and refer to a polymeric form of nucleotides of any length, either
25 deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides can have any three-dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: a gene or gene fragment (for example, a probe, primer, EST or SAGE tag), exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA,
30 recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated

DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide can comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure can be imparted before or after assembly of the polymer. The sequence of nucleotides can be interrupted by non-nucleotide components. A polynucleotide can be further modified after polymerization, such as by conjugation with a labeling component. The term also refers to both double- and single-stranded molecules. Unless otherwise specified or required, any embodiment of this invention that is a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double-stranded form.

A polynucleotide is composed of a specific sequence of four nucleotide bases: adenine (A); cytosine (C); guanine (G); thymine (T); and uracil (U) for guanine when the polynucleotide is RNA. Thus, the term "polynucleotide sequence" is the alphabetical representation of a polynucleotide molecule. This alphabetical representation can be input into databases in a computer having a central processing unit and used for bioinformatics applications such as functional genomics and homology searching.

A "gene" refers to a polynucleotide containing at least one open reading frame (ORF) that is capable of encoding a particular polypeptide or protein after being transcribed and translated. Any of the polynucleotides sequences described herein may be used to identify larger fragments or full-length coding sequences of the gene with which they are associated. Methods of isolating larger fragment sequences are known to those of skill in the art.

A "gene product" or alternatively a "gene expression product" refers to the amino acid (*e.g.*, peptide or polypeptide) generated when a gene is transcribed and translated.

The term "polypeptide" is used interchangeably with the term "protein" and in its broadest sense refers to a compound of two or more subunit amino acids, amino acid analogs, or peptidomimetics. The subunits may be linked by peptide bonds. In another embodiment, the subunit may be linked by other bonds, *e.g.*, ester, ether, *etc.* As used herein the term "amino acid" refers to either

natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. A peptide of three or more amino acids is commonly called an oligopeptide if the peptide chain is short. If the peptide chain is long, the peptide is commonly called a polypeptide or a protein.

“Under transcriptional control” is a term well understood in the art and indicates that transcription of a polynucleotide sequence, usually a DNA sequence, depends on its being operatively linked to an element which contributes to the initiation of, or promotes, transcription. “Operatively linked” refers to a juxtaposition wherein the elements are in an arrangement allowing them to function.

The following terms define the proteins and polynucleotides identified in Table 1, *infra*.

As used herein, the term “HSPC242 gene” refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 3 is one example of an HSPC242 gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession No: NM_016498.2 and the sequences that encode HSPC242 gene expression products as defined herein. See Zhang *et al.* (2000) Genome Res. 10(10):1546-1560. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO.: 2 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 3, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 25199.

As used herein, the term “HSPC242 gene expression product, protein or polypeptide” includes the amino acid sequence of SEQ ID NO.: 2 (see Zhang

et al. (2000) *supra* and GenBank Accession No. NP_057582) as well as the amino acid sequences transcribed and translated from the HSPC242 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 2 and which have the biological activity as set forth herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 2 or other HSPC242 gene expression product that has been modified by conservative amino acid substitutions.

Zhang *et al.* (2000) *supra*, reported isolation of the HSPC242 sequence from three hundred cDNAs containing putatively entire open reading frames (ORFs) obtained from CD34⁺ hematopoietic stem/progenitor cells (HSPCs), based on EST cataloging, clone sequencing, *in silico* cloning, and rapid amplification of cDNA ends (RACE).

As used herein, the term "IGSF6 (DORA) or "CGI-81" gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 6 is one example of an CGI-81 gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession No: NM_016025.2 (*see* Bates *et al.* (2000) Immunogenetics 52(1-2):112-120) and the sequences that encode CGI-81 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO.: 5 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 6, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency

to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 279583.

5 As used herein, the term "CGI-81 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 5 (*see also* GenBank Accession No. NP_057582) as well as the amino acid sequences transcribed and translated from the CGI-81 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term
10 includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 5 and which have the biological activity of this sequence as set forth herein. Examples of homologous
15 amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 5 or other CGI-81 gene expression product that has been modified by conservative amino acid substitutions.

Bates *et al.* (2000) *supra*, report that IGSF6 (DORA), is a novel member of the immunoglobulin superfamily (IGSF) from human and rat expressed in
20 dendritic and myeloid cells. Using a probe from the open reading frame of the rat cDNA, the authors isolated a cosmid which contains the entire mouse gene. By comparative analysis and reverse transcriptase polymerase chain reaction, the intron/exon structure and the mRNA of the mouse gene and, with respect to human BAC clones, the human gene, were defined by the authors. The genes
25 span 10 kb (mouse) and 12 kb (human), with six exons arranged in a manner similar to other members of the IGSF. All intron/exon boundaries follow the GT-AG rule. The authors report that expression of the mouse IGSF6 gene is restricted to cells of the immune system, particularly macrophages. The human and mouse genes were localized to a locus associated with inflammatory bowel
30 disease. The authors report that this gene is transcribed and processed, and is reported to contain homologues in *Caenorhabditis elegans* and prokaryotes, and is expressed in most organs in the mouse.

As used herein, the term “Epstein-Barr virus (EBV) Induced Gene 3” (“EBI3”) gene” refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID No.: 9 is one example of an EBI3 gene, and
5 others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No.: NM_005755.2 and the sequences that encode EBV-3 gene expression products as defined herein. (*See also* Devergne, *et al.* (1996) PNAS 94(22):12041-12046 and (2000) J. Virol. 70(2):1143-1153). Also included within this definition are biologically
10 equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 8 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 9, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides
15 that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs. 185705.

As used herein, the term “EBI3 gene expression product, protein or
20 polypeptide” includes the amino acid sequence of SEQ ID NO.: 8 (*see also* GenBank Accession No. NP_005746) as well as the amino acid sequences transcribed and translated from the EBI3 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes
25 isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 8 and which have the biological activity as set forth herein herein. Examples of homologous amino
30 acid sequences include, but are not limited to polypeptides have the amino acid sequence of SEQ ID NO.: 8 or other EBI3 gene expression product that has been modified by conservative amino acid substitutions.

Devergne (1996) *supra* and (2000) *supra* originally identified this gene and protein product by the induction of its expression in B lymphocytes by Epstein-Barr virus infection. The protein product is a secreted glycoprotein which is a member of the hematopoietic receptor family related to the p40 subunit of interleukin 12 (IL-12).

As used herein, the term "FAM3C family with sequence similarity 3, member C ("FAM3C") gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 12 is one example of an FAM3C gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession NO.: NM_014888.1 and the sequences that encode FAM3C gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 11 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 11, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs. 29882.

As used herein, the term "FAM3C gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 11 (see also GenBank Accession No.: NP_055703.1) as well as the amino acid sequences transcribed and translated from the FAM3C genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 11 and which have the

biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 11 or other FAM3C gene expression product that has been modified by conservative amino acid substitutions.

5 Zhu *et al.* (2002) *Genomics* **80**(2):144-150, reported the identification and characterization of a novel cytokine-like gene family using structure-based methods to search for novel four-helix-bundle cytokines in genomics databases. Four genes were reported in this family, FAM3A, FAM3B, FAM3C, and FAM3D, each encoding a protein (224-235 amino acids) with a hydrophobic
10 leader sequence. Northern analysis indicates that FAM3B is highly expressed in pancreas, FAM3D in placenta, and FAM3A and FAM3C in almost all tissues.

 As used herein, the term leukocyte specific transcript 1 (“LST-1”) gene” refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein.
15 Sequence ID NO.: 15 is one example of an LST-1 gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession No.: NM_007161 and the sequences that encode LST-1 gene expression products as defined herein. (*See* Holzinger *et al.* (1995) *Immunogenetics* **42**:315-322; Spies *et al.* (1989) *Science* **243**:214-217.) Also
20 included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 14 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 15, and as determined by percent identity sequence analysis run under default parameters. Also within this
25 definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 380427.

30 As used herein, the term “LST-1 gene expression product, protein or polypeptide” includes the amino acid sequence of SEQ ID NO.: 14 as well as the amino acid sequences transcribed and translated from the LST-1 genes identified

above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 14 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 14 or other LST-1S gene expression product that has been modified by conservative amino acid substitutions.

Holzinger *et al.* (1995) *supra*, reported that the LST-1 (LST-1) gene is constitutively expressed in leukocytes and dendritic cells, and it is characterized by extensive alternative splicing. The authors reported seven (7) different LST-1 splice variants in PBMC; thus, 14 LST-1 splice variants (LST-1/A-LST-1/N) have been detected in various cell types. The authors also reported that these isoforms code for transmembrane as well as soluble LST-1 proteins characterized by two alternative open reading frames at their 3' end. The authors further reported that the presence of the transmembrane variant LST-1/C on the cell surface of the monocytic cell lines U937 and THP1. Recombinant expression of LST-1/C permitted its profound inhibitory effect on lymphocyte proliferation to be observed. In contrast, the alternative transmembrane variant LST-1/A, the extracellular domain of which shows no amino acid sequence homology to LST-1/C exerted a weaker but similar inhibitory effect on PBMC. The authors concluded that this data demonstrate the protein expression of LST-1 on the cell surface of mononuclear cells, and they show an inhibitory effect on lymphocyte proliferation of two LST-1 proteins although they have only a very short amino acid homology.

As used herein, the term "TNF α IP6 tumor necrosis factor, alpha-induced protein 6 gene" ("TNF α -IP6 gene") refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequences that encode Seq. ID NO.: 17 are examples of TNF α -IP6 genes, *see, e.g.*, the sequences that encode the protein set

forth under GenBank Accession No.: NM_007115 and the sequences that encode TNF α -IP6 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 17 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as GenBank Accession No.: NM_007115, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 29352.

As used herein, the term "TNF α -IP6 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 17 as well as the amino acid sequences transcribed and translated from the TNF α -IP6 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 17 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 17 or other TNF α -IP6 gene expression product that has been modified by conservative amino acid substitutions.

Lee *et al.* (1992) J. Cell Biol. 116(2):545-557, reported that the protein encoded by the TNF α -IP6 gene is a secretory protein that contains a hyaluronan-binding domain, and thus is a member of the hyaluronan-binding protein family. The hyaluronan-binding domain is known to be involved in extracellular matrix stability and cell migration. This protein has been shown to form a stable complex with inter-alpha-inhibitor (I alpha I), and thus enhances the serine protease inhibitory activity of I alpha I, which is important in the protease

network associated with inflammation. The authors also reported that expression of this gene can be induced by tumor necrosis factor alpha and interleukin-1.

As used herein, the term "MGC5309 gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 20 is one example of an MGC5309 gene, and others are known in the art examples of which include, but are not limited to the sequences that encode proteins set forth under GenBank Accession NO: NP_115662 and the sequences that encode MGC5309 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 19 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 20, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 13885.

As used herein, the term "MGC5309 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 19 as well as the amino acid sequences transcribed and translated from the MGC5309 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 19 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 19 or other MGC5309 gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term leukocyte immunoglobulin-like receptor ("ILT-6 gene") refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 23 is one example of an ILT-6 gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession NO: NM_006865 and the sequences that encode ILT-6 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 22 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 23, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 113277.

As used herein, the term "ILT-6 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 22 (see also GenBank Accession No. NP_006856) as well as the amino acid sequences transcribed and translated from the ILT-6 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 22 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 22 or other ILT-6 gene expression product that has been modified by conservative amino acid substitutions.

Borges *et al.* (1997) J. Immunol. **159**(11):5192-5196, reported that leukocyte Ig-like receptors (LIRs) are a family of immunoreceptors expressed on

monocytes and B cells and at lower levels on dendritic cells and NK cells and that the amino acid sequences in the extracellular regions of eight of these receptors show between 63 and 84% identity to a LIR-1 sequence. LIRs contain either two or four Ig domains and fall into three classes: those with cytoplasmic domains containing two, three, or four immunoreceptor tyrosine-based inhibitory motif-like motifs; those with a short cytoplasmic domain and no immunoreceptor tyrosine-based inhibitory motif-like motifs; and those with no transmembrane domain represented by a single LIR molecule that is presumably secreted. The authors reported that the LIRs are structurally related to the human Fc(alpha)R and the killer inhibitory receptors and map to the same region of chromosome 19 as these genes.

As used herein, the term "Osteoglycin (osteoinductive factor) ("OGN") gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Examples of OGN genes, known in the art, examples of which include, but are not limited to the sequences set forth under GenBank Accession No: NM_014057.2 and the sequences that encode OGN gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 25 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as that deposited under GenBank Accession No: NM_014057.2, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 109439.

As used herein, the term "OGN gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 25 (*e.g.*, *see* GenBank Accession No.: NP_054476.1) as well as the amino acid sequences transcribed and translated from the OSC genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell,

mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or
5 alternatively at least 95% homologous to SEQ ID NO.: 25 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 25 or other OGN gene expression product that has been modified by conservative amino acid substitutions.

10 Pellagata *et al.* (2000) Nature Genetics **25**(1):91-95, report that OGN was highly expressed in the cornea.

As used herein, the term "CAB56184 gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. The gene was isolated from
15 chromosome 6. Sequence ID NO.: 28 is one example of an CAB56184 gene, and others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No.: NM_032520.2 and the sequences that encode CAB56184 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as
20 those sequences that code for the polypeptide of SEQ ID NO: 27 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 28, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified
25 by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 241575.

As used herein, the term "CAB56184 gene expression product, protein or
30 polypeptide" includes the amino acid sequence of SEQ ID NO.: 27 (see also GenBank Accession No. NP_115909) as well as the amino acid sequences transcribed and translated from the CAB56184 genes identified above, without

regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also
5 includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 27 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 27 or other CAB56184 gene expression product that
10 has been modified by conservative amino acid substitutions.

As used herein, the term "Chromosome 6 ORF68 ("C6-68") gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 31 is one example of an C6-68 gene, and others are known in the art
15 examples of which include, but are not limited to the sequence set forth under GenBank Accession No: NM_138459 and the sequences that encode C6-68 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 30 and those having at least 90% or alternatively, at
20 least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 31, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human
25 genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 289008.

As used herein, the term "C6-68 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 30 (*see also* GenBank Accession No.: NP_612468) as well as the amino acid sequences
30 transcribed and translated from the C6-68 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes

isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 30 and which have the biological activity as set forth herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 30 or other C6-68 gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term "Stem Cell Growth Factor ("SCGF") gene" refers to at least the ORF of a contiguous polynucleotide sequence that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 34 is one example of an SCGF gene, and others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No: NM_002975 and the sequences that encode SCGF gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 33 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 34, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 105927.

As used herein, the term "SCGF gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 33 (see also the protein described under GenBank Accession No.: NP_002966) as well as the amino acid sequences transcribed and translated from the SCGF genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The

term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 33 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 33 or other SCGF gene expression product that has been modified by conservative amino acid substitutions.

Akin *et al.* (2002) Am. J. Clin. Pathol. **118**(2):242-247, reported the presence of stem cell factor in lesional mast cells isolated from bone marrow biopsies. Meade *et al.* (2002) Biochem. Pharm. **64**(2):317-325, suggested that stem cell factor may play a role, within adenosine, in asthma.

As used herein, the term "HTO36 gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 37 is one example of an HTO36 gene, and others are known in the art examples of which include, but are not limited to the sequences that encode the protein set forth under GenBank Accession No: NP_112484 and the sequences that encode HTO36 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 36 (see also GenBank Accession No. NM_031207) and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 37, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 321669.

As used herein, the term "HTO36 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 36 as well as the amino acid sequences transcribed and translated from the HTO36 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or

human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 36 (*see also* GenBank Accession No. NP_112484) and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 36 or other HTO36 gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term “cat eye syndrome chromosome region, candidate 1” (“CECR1 gene”) refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 41 (*see also* GenBank Accession No. NM_177405.1) is one example of an CECR1 gene, and others are known in the art examples of which include, but are not limited to the sequences that encode the protein sequence set forth under GenBank Accession Nos: NP_059120 and NP_803124 and the sequences that encode CECR1 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NOS: 39 and 40 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 41, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 170310.

As used herein, the term “CECR1 gene expression product, protein or polypeptide” includes the amino acid sequence of SEQ ID NOS.: 39 and 40 (*see also* proteins described under GenBank Accession Nos.: NP_059120 and NP_803124) as well as the amino acid sequences transcribed and translated from the CECR1 genes identified above, without regard to the gene expression system,

e.g., bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NOS.: 39 and 40 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NOS.: 39 and 40 or other CECR1 gene expression product that has been modified by conservative amino acid substitutions.

Riazi *et al.* (2000) Genomics 64(3):277-285 reported that the gene is alternatively spliced and expressed in numerous tissues, with most abundant expression in human adult heart, lung, lymphoblasts, and placenta as well as fetal lung, liver, and kidney. The authors report that *in situ* hybridization of a human embryo shows specific expression in the outflow tract and atrium of the developing heart, the VII/VIII cranial nerve ganglion, and the notochord. The authors also reported that location of this gene in the CES critical region and its embryonic expression suggest that the overexpression of CECR1 may be responsible for at least some features of CES, particularly the heart defects.

Footz *et al.* (2001) Genome Res. 11(6):1053-1070, sequenced and characterized a 1.1-Mb region of human chromosome 22g believed to be the dosage-sensitive gene responsible for cat eye syndrome (CES) as well as a putative 450-kb homologous region on mouse chromosome 6. The authors also reported that the gene encodes a member of a subfamily of the adenosine deaminase protein family. Two transcript variants encoding distinct isoforms were identified for this gene. Transcript variant (2) contains a distinct 5' UTR and lacks an in-frame portion of the 5' coding region, compared to variant 1. The resulting isoform (b) has a shorter N-terminus compared to isoform a.

As used herein, the term "interleukin 27 working designation" ("IL27W gene") refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. The sequence deposited under GenBank Accession No: XP_035638 is

one example of an IL27W gene. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 43 and those having at least 90% or alternatively, at least 95% sequence homology the sequence deposited under GenBank Accession No: XP_035638, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 10927.

As used herein, the term "IL27W gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 43 as well as the amino acid sequences transcribed and translated from the IL27W genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 43 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 43 or other IL27W gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term "FLJ40021 gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 46 is one example of an FLJ40021 gene, and others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No: NM_153225 and the sequences that encode FLJ40021 gene expression products as defined herein (*see, e.g.* GenBank Accession No. NP_694957). Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 45 and those

having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 46, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 41185.

As used herein, the term "FLJ40021 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 45 as well as the amino acid sequences transcribed and translated from the FLJ40021 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 45 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 45 or other FLJ40021 gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term "Follicular lymphoma variant ("FLV") gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 49 is one example of an FLV gene, and others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No: NM_002035 and the sequences that encode FLV gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 48 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID

NO.: 49, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human
5 genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 74050.

As used herein, the term "FLV gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 48 as well as the amino acid sequences transcribed and translated from the FLV genes identified
10 above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The expression product is reported to be a secreted protein weakly expressed in hematopoietic tissue. It has shown a high rate of transcription in some T-cell malignancies and in phytohemagglutinin-stimulated lymphocytes.
15 The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 48 and which have the biological activity as described herein. Examples of homologous amino acid
20 sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 48 or other FLV gene expression product that has been modified by conservative amino acid substitutions.

As used herein, the term "FLJ22875 ("FLJ22875") gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or
25 polypeptide having the biological activity as set forth herein. Sequence ID NO.: 52 is one example of an FLJ gene, and others are known in the art examples of which include, but are not limited to the sequences set forth under GenBank Accession No: NM_032231 and the sequences that encode FLJ22875 gene expression products as defined herein. Also included within this definition are
30 biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 51 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID

NO.: 52, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human
5 genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 406548.

As used herein, the term "FLJ22875 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 51 as well as the amino acid sequences transcribed and translated from the FLJ22875 genes
10 identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid
15 sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 51 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 51 or other FLJ22875 gene expression product that has been modified by conservative amino acid
20 substitutions.

As used herein, the term "platelet-derived growth factor receptor-like gene" "PDGF gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 55 is one example of an PDGF gene, and
25 others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No: NM_006207 and the sequences that encode PDGF gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 54 and those having at
30 least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 55, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are

biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 170040.

5 As used herein, the term "PDGF gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 54 as well as the amino acid sequences transcribed and translated from the PDGF genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other
10 prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 54 and which have the biological activity as shown herein. Examples of homologous amino
15 acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 54 or other PDGF gene expression product that has been modified by conservative amino acid substitutions.

 This gene is reported to encode a protein with significant sequence similarity to the ligand binding domain of platelet-derived growth factor receptor
20 beta. Mutations in this gene, or deletion of a chromosomal segment containing this gene, are associated with sporadic hepatocellular carcinomas, colorectal cancers, and non-small cell lung cancers.

 As used herein, the term "WNT8A gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide
25 having the biological activity as set forth herein. Sequence ID NO.: 61 is one example of an WNT8A gene, and others are known in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No.: NM_031933 and the sequences that encode WNT8A gene expression products as defined herein. Also included within this definition are biologically
30 equivalent sequences such as those sequences that code for the polypeptide of SEQ ID NO: 60 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 61, and as

determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 302163.

As used herein, the term "WNT8A gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 60 as well as the amino acid sequences transcribed and translated from the WNT8A genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 60 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 60 or other WNT8A gene expression product that has been modified by conservative amino acid substitutions.

Saitoh and Katoh (2001) *Int. J. Oncol.* **19**:123-127, searched human genome draft sequences for homologs of the *Xenopus* wnt8 gene and identified WNT8A. The authors assembled a full-length WNT8A cDNA sequence. The authors predicted that the 6 exons of the WNT8A gene encode a 351-amino acid protein with an N-terminal signal peptide, 3 N-linked glycosylation sites, and residues conserved among members of the WNT family. The WNT8A protein shares 63.2% sequence identity with human WNT8B and homology with the mouse Wnt8d protein.

As used herein, the term "secreted and transmembrane 1 gene" "SECTM1 gene" refers to at least the ORF of a contiguous polynucleotide sequence and that encodes a protein or polypeptide having the biological activity as set forth herein. Sequence ID NO.: 64 is one example of an SECTM1 gene, and others are known

in the art examples of which include, but are not limited to the sequence set forth under GenBank Accession No.: NM_003004.1 and the sequences that encode SECTM1 gene expression products as defined herein. Also included within this definition are biologically equivalent sequences such as those sequences that
5 code for the polypeptide of SEQ ID NO: 63 and those having at least 90% or alternatively, at least 95% sequence homology to an exemplary sequence, such as SEQ ID NO.: 64, and as determined by percent identity sequence analysis run under default parameters. Also within this definition are biologically equivalent genes or polynucleotides that are identified by the ability to hybridize under
10 conditions of high stringency to the minus strand. It may be desirable to use non-human genes, the polynucleotide sequences of which are known in the art, see for example, UniGene Cluster Hs 95655.

As used herein, the term "SECTM1 gene expression product, protein or polypeptide" includes the amino acid sequence of SEQ ID NO.: 63 as well as the
15 amino acid sequences transcribed and translated from the SECTM1 genes identified above, without regard to the gene expression system, *e.g.*, bacterial or other prokaryotic cell, yeast cell, mammalian cell such as a simian, bovine or human cell. The term includes isolated, naturally occurring polypeptides isolated from tissue samples as well as recombinantly produced proteins and
20 polypeptides. The term also includes polypeptides having the amino acid sequences that are at least 90 % or alternatively at least 95% homologous to SEQ ID NO.: 63 and which have the biological activity as described herein. Examples of homologous amino acid sequences include, but are not limited to polypeptides having the amino acid sequence of SEQ ID NO.: 63 or other SECTM1 gene
25 expression product that has been modified by conservative amino acid substitutions.

Slentz-Kesler *et al.* (1998) Genomics 47(3):327-340, reported that the investigation of a DNase-hypersensitive site upstream of the CD7 gene on chromosome 17q25 led to the discovery of a novel human gene designated K12
30 (SECTM1). This gene spans approximately 14 kb and encodes a 1.8-kb mRNA detected at the highest levels in peripheral blood leukocytes and breast cancer cell lines. The open reading frame predicts a 248-amino-acid protein with the

hydropathic characteristics of a type 1a membrane protein. Western blots showed that the K12 protein exists as a cluster of bands around 27 kDa, and extractions using nonionic detergents or high pH conditions demonstrate that it behaves as an integral membrane protein. Immunofluorescence localization studies reveal that K12 is not detectable on the cell surface, but instead is found in a perinuclear Golgi-like pattern and colocalizes with a well-known Golgi marker. The authors report that in addition, an approximately 20-kDa soluble form of the K12 protein derived from the N-terminal domain is specifically secreted by cells into the culture medium. Immunohistochemical analysis of peripheral blood cells shows that K12 is found in leukocytes of the myeloid lineage, with the strongest staining observed in granulocytes and no detectable expression in lymphocytes.

As used herein, the term “comprising” is intended to mean that the compositions and methods include the recited elements, but not excluding others. “Consisting essentially of” when used to define compositions and methods, shall mean excluding other elements of any essential significance to the combination. Thus, a composition consisting essentially of the elements as defined herein would not exclude trace contaminants from the isolation and purification method and pharmaceutically acceptable carriers, such as phosphate buffered saline, preservatives, and the like. “Consisting of” shall mean excluding more than trace elements of other ingredients and substantial method steps for administering the compositions of this invention. Embodiments defined by each of these transition terms are within the scope of this invention.

The term “isolated” means separated from constituents, cellular and otherwise, in which the polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, are normally associated with in nature. In one aspect of this invention, an isolated polynucleotide is separated from the 3' and 5' contiguous nucleotides with which it is normally associated with in its native or natural environment, *e.g.*, on the chromosome. As is apparent to those of skill in the art, a non-naturally occurring polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, does not require “isolation” to distinguish it from its naturally occurring counterpart. In addition, a “concentrated”, “separated” or

“diluted” polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, is distinguishable from its naturally occurring counterpart in that the concentration or number of molecules per volume is greater than “concentrated” or less than “separated” than that of its naturally occurring counterpart. A
5 polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, which differs from the naturally occurring counterpart in its primary sequence or for example, by its glycosylation pattern, need not be present in its isolated form since it is distinguishable from its naturally occurring counterpart by its primary sequence, or alternatively, by another characteristic such as glycosylation pattern.
10 Thus, a non-naturally occurring polynucleotide is provided as a separate embodiment from the isolated naturally occurring polynucleotide. A protein produced in a bacterial cell is provided as a separate embodiment from the naturally occurring protein isolated from a eukaryotic cell in which it is produced in nature.

15 “Gene delivery,” “gene transfer,” and the like as used herein, are terms referring to the introduction of an exogenous polynucleotide (sometimes referred to as a “transgene”) into a host cell, irrespective of the method used for the introduction. Such methods include a variety of well-known techniques such as
20 vector-mediated gene transfer (by, *e.g.*, viral infection/transfection, or various other protein-based or lipid-based gene delivery complexes) as well as techniques facilitating the delivery of “naked” polynucleotides (such as electroporation, “gene gun” delivery and various other techniques used for the introduction of polynucleotides). The introduced polynucleotide may be stably or transiently maintained in the host cell. Stable maintenance typically requires that the
25 introduced polynucleotide either contains an origin of replication compatible with the host cell or integrates into a replicon of the host cell such as an extrachromosomal replicon (*e.g.*, a plasmid) or a nuclear or mitochondrial chromosome. A number of vectors are known in the art to be capable of mediating transfer of genes to mammalian cells.

30 A “gene delivery vehicle” is defined as any molecule that can carry inserted polynucleotides into a host cell. Examples of gene delivery vehicles are liposomes, biocompatible polymers, including natural polymers and synthetic

polymers; lipoproteins; polypeptides; polysaccharides; lipopolysaccharides; artificial viral envelopes; recombinant yeast cells, metal particles; and bacteria, or viruses, such as baculovirus, adenovirus and retrovirus, bacteriophage, cosmid, plasmid, fungal vectors and other recombination vehicles typically used in the art which have been described for expression in a variety of eukaryotic and prokaryotic hosts, and may be used for gene therapy as well as for simple protein expression.

A "viral vector" is defined as a recombinantly produced virus or viral particle that comprises a polynucleotide to be delivered into a host cell, either *in vivo*, *ex vivo* or *in vitro*. Examples of viral vectors include retroviral vectors, adenovirus vectors, adeno-associated virus vectors, alphavirus vectors and the like. Alphavirus vectors, such as Semliki Forest virus-based vectors and Sindbis virus-based vectors, have also been developed for use in gene therapy and immunotherapy. See, Schlesinger and Dubensky (1999) Curr. Opin. Biotechnol. 5:434-439 and Ying *et al.* (1999) Nat. Med. 5(7):823-827. In aspects where gene transfer is mediated by a retroviral vector, a vector construct refers to the polynucleotide comprising the retroviral genome or part thereof, and a therapeutic gene. As used herein, "retroviral mediated gene transfer" or "retroviral transduction" carries the same meaning and refers to the process by which a gene or nucleic acid sequences are stably transferred into the host cell by virtue of the virus entering the cell and integrating its genome into the host cell genome. The virus can enter the host cell via its normal mechanism of infection or be modified such that it binds to a different host cell surface receptor or ligand to enter the cell. As used herein, retroviral vector refers to a viral particle capable of introducing exogenous nucleic acid into a cell through a viral or viral-like entry mechanism.

Retroviruses carry their genetic information in the form of RNA; however, once the virus infects a cell, the RNA is reverse-transcribed into the DNA form which integrates into the genomic DNA of the infected cell. The integrated DNA form is called a provirus.

In aspects where gene transfer is mediated by a DNA viral vector, such as an adenovirus (Ad) or adeno-associated virus (AAV), a vector construct refers to

the polynucleotide comprising the viral genome or part thereof, and a transgene. Adenoviruses (Ads) are a relatively well characterized, homogenous group of viruses, including over 50 serotypes. *See, e.g.*, WO 95/27071. Ads are easy to grow and do not require integration into the host cell genome. Recombinant Ad
5 derived vectors, particularly those that reduce the potential for recombination and generation of wild-type virus, have also been constructed. *See*, WO 95/00655 and WO 95/11984. Wild-type AAV has high infectivity and specificity integrating into the host cell's genome. *See*, Hermonat and Muzyczka (1984) Proc. Natl. Acad. Sci. USA **81**:6466-6470 and Lebkowski *et al.* (1988) Mol. Cell.
10 Biol. **8**:3988-3996.

Vectors that contain both a promoter and a cloning site into which a polynucleotide can be operatively linked are known in the art. Such vectors are capable of transcribing RNA *in vitro* or *in vivo*, and are commercially available from sources such as Stratagene (La Jolla, CA) and Promega Biotech (Madison,
15 WI). In order to optimize expression and/or *in vitro* transcription, it may be necessary to remove, add or alter 5' and/or 3' untranslated portions of the clones to eliminate extra, potential inappropriate alternative translation initiation codons or other sequences that may interfere with or reduce expression, either at the level of transcription or translation. Alternatively, consensus ribosome binding sites
20 can be inserted immediately 5' of the start codon to enhance expression.

Gene delivery vehicles also include several non-viral vectors, including DNA/liposome complexes, recombinant yeast cells, and targeted viral protein-DNA complexes. Liposomes that also comprise a targeting antibody or fragment thereof can be used in the methods of this invention. To enhance delivery to a
25 cell, the nucleic acid or proteins of this invention can be conjugated to antibodies or binding fragments thereof which bind cell surface antigens, *e.g.*, TCR, CD3 or CD4.

A "probe" when used in the context of polynucleotide manipulation refers to an oligonucleotide that is provided as a reagent to detect a target potentially
30 present in a sample of interest by hybridizing with the target. Usually, a probe will comprise a label or a means by which a label can be attached, either before or subsequent to the hybridization reaction. Suitable labels include, but are not

limited to radioisotopes, fluorochromes, chemiluminescent compounds, dyes, and proteins, including enzymes.

5 A “primer” is a short polynucleotide, generally with a free 3’ -OH group that binds to a target or “template” potentially present in a sample of interest by hybridizing with the target, and thereafter promoting polymerization of a polynucleotide complementary to the target. A “polymerase chain reaction” (“PCR”) is a reaction in which replicate copies are made of a target polynucleotide using a “pair of primers” or a “set of primers” consisting of an “upstream” and a “downstream” primer, and a catalyst of polymerization, such as
10 a DNA polymerase, and typically a thermally-stable polymerase enzyme. Methods for PCR are known in the art, and taught, for example in “PCR: A PRACTICAL APPROACH” (M. MacPherson *et al.*, IRL Press at Oxford University Press (1991)). All processes of producing replicate copies of a polynucleotide, such as PCR or gene cloning, are collectively referred to herein
15 as “replication.” A primer can also be used as a probe in hybridization reactions, such as Southern or Northern blot analyses. Sambrook *et al.*, *supra*.

An expression “database” denotes a set of stored data that represent a collection of sequences, which in turn represent a collection of biological reference materials.

20 The term “cDNAs” refers to complementary DNA that is mRNA molecules present in a cell or organism made into cDNA with an enzyme such as reverse transcriptase. A “cDNA library” is a collection of all of the mRNA molecules present in a cell or organism, all turned into cDNA molecules with the enzyme reverse transcriptase, then inserted into “vectors” (other DNA molecules
25 that can continue to replicate after addition of foreign DNA). Exemplary vectors for libraries include bacteriophage (also known as “phage”), viruses that infect bacteria, for example, lambda phage. The library can then be probed for the specific cDNA (and thus mRNA) of interest.

30 As used herein, “expression” refers to the process by which polynucleotides are transcribed into mRNA and/or the process by which the transcribed mRNA is subsequently being translated into peptides, polypeptides, or proteins. If the polynucleotide is derived from genomic DNA, expression may

include splicing of the mRNA in a eukaryotic cell. "Differentially expressed" as applied to a gene, refers to the differential production of the mRNA transcribed and/or translated from the gene or the protein product encoded by the gene. A differentially expressed gene may be overexpressed or underexpressed as compared to the expression level of a normal or control cell. In one aspect, it refers to a differential that is 2.5 times, preferably 5 times, or preferably 10 times higher or lower than the expression level detected in a control sample. The term "differentially expressed" also refers to nucleotide sequences in a cell or tissue which are expressed where silent in a control cell or not expressed where expressed in a control cell.

As used herein, "solid phase support" or "solid support", used interchangeably, is not limited to a specific type of support. Rather a large number of supports are available and are known to one of ordinary skill in the art. Solid phase supports include silica gels, resins, derivatized plastic films, glass beads, cotton, plastic beads, alumina gels, microarrays and chips. As used herein, "solid support" also includes synthetic antigen-presenting matrices, cells, and liposomes. A suitable solid phase support may be selected on the basis of desired end use and suitability for various protocols. For example, for peptide synthesis, solid phase support may refer to resins such as polystyrene (*e.g.*, PAM-resin obtained from Bachem Inc., Peninsula Laboratories, *etc.*), POLYHIPE® resin (obtained from Aminotech, Canada), polyamide resin (obtained from Peninsula Laboratories), polystyrene resin grafted with polyethylene glycol (TentaGel®, Rapp Polymere, Tubingen, Germany) or polydimethylacrylamide resin (obtained from Milligen/Bioscience, California).

A polynucleotide also can be attached to a solid support for use in high throughput screening assays. PCT WO 97/10365, for example, discloses the construction of high density oligonucleotide chips. See also, U.S. Patent Nos. 5,405,783; 5,412,087; and 5,445,934. Using this method, the probes are synthesized on a derivatized glass surface also known as chip arrays. Photoprotected nucleoside phosphoramidites are coupled to the glass surface, selectively deprotected by photolysis through a photolithographic mask, and

reacted with a second protected nucleoside phosphoramidite. The coupling/deprotection process is repeated until the desired probe is complete.

“Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogsteen binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

Hybridization reactions can be performed under conditions of different “stringency”. In general, a low stringency hybridization reaction is carried out at about 40 °C in 10 x SSC or a solution of equivalent ionic strength/temperature. A moderate stringency hybridization is typically performed at about 50 °C in 6 x SSC, and a high stringency hybridization reaction is generally performed at about 60 °C in 1 x SSC.

When hybridization occurs in an antiparallel configuration between two single-stranded polynucleotides, the reaction is called “annealing” and those polynucleotides are described as “complementary”. A double-stranded polynucleotide can be “complementary” or “homologous” to another polynucleotide, if hybridization can occur between one of the strands of the first polynucleotide and the second. “Complementarity” or “homology” (the degree that one polynucleotide is complementary with another) is quantifiable in terms of the proportion of bases in opposing strands that are expected to form hydrogen bonding with each other, according to generally accepted base-pairing rules.

A polynucleotide or polynucleotide region (or a polypeptide or polypeptide region) has a certain percentage (for example, 80%, 85%, 90% or 95%) of “sequence identity” to another sequence means that, when aligned, that percentage of bases (or amino acids) are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those

described in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F.M. Ausubel *et al.*, eds., 1987) Supplement 30, section 7.7.18, Table 7.7.1.

Preferably, default parameters are used for alignment. A preferred alignment program is BLAST, using default parameters. In particular, preferred programs are BLASTN and BLASTP, using the following default parameters: Genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + SwissProtein + SPupdate + PIR. Details of these programs can be found at the following Internet address: <http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST>.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in structure or function. Metaplasia is a form of controlled cell growth in which one type of fully differentiated cell substitutes for another type of differentiated cell. Metaplasia can occur in epithelial or connective tissue cells. Atypical metaplasia involves a somewhat disorderly metaplastic epithelium.

As used herein, the terms "pathological cells", "neoplasia", "tumor", "tumor cells", "cancer" and "cancer cells", (used interchangeably) refer to cells which exhibit relatively autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation (*i.e.*, de-regulated cell division). Pathological cells can be malignant or benign. A metastatic cell or tissue means that the cell can invade and destroy neighboring body structures.

"Suppressing" tumor growth indicates a growth state that is curtailed when compared to growth without contact with educated, antigen-specific immune effector cells described herein. Tumor cell growth can be assessed by any means known in the art, including, but not limited to, measuring tumor size, determining whether tumor cells are proliferating using a ³H-thymidine incorporation assay, or counting tumor cells. "Suppressing" tumor cell growth means any or all of the following states: slowing, delaying, and stopping tumor growth, as well as tumor shrinkage.

As used herein a second polynucleotide "corresponds to" another (a first) polynucleotide if it is related to the first polynucleotide by any of the following relationships:

- 5 1) The second polynucleotide comprises the first polynucleotide and the second polynucleotide encodes a gene product.
- 10 2) The second polynucleotide is 5' or 3' to the first polynucleotide in cDNA, RNA, genomic DNA, or fragment of any of these polynucleotides. For example, a second polynucleotide may be a fragment of a gene that includes the first and second polynucleotides. The first and second polynucleotides are related in that they are components of the gene coding for a gene product, such as a protein or antibody. However, it is not necessary that the second polynucleotide comprises or overlaps with the first polynucleotide to be encompassed within the definition of "corresponding to" as used herein. For example, the first polynucleotide
15 may be a fragment of a 3' untranslated region of the second polynucleotide, for example a promoter sequence. The first and second polynucleotide may be fragment of a gene coding for a gene product. The second polynucleotide may be an exon of the gene while the first polynucleotide may be an intron of the gene.
- 20 3) The second polynucleotide is the complement of the first polynucleotide.

The "genotype" of a cell refers to the genetic makeup of the cell and/or its gene expression profile. Modulation of the genotype of a cell can be achieved by introducing additional DNA or RNA either as episomes or as an integral part of
25 the chromosomal DNA of the recipient cell. The genotype can also be modulated by altering the expression level, e.g. mRNA abundance, of a particular gene using agents that regulate gene expression.

A "database" denotes a set of stored data which represent a collection of sequences including nucleotide and peptide sequences, which in turn represent a
30 collection of biological reference materials.

The term "sequence motif" refers to a pattern present in a group of molecules. For instance, in one embodiment, the present invention provides for identification of a sequence motif among peptides. In this embodiment, a typical pattern may be identified by characteristic amino acid residues, such as hydrophobic, hydrophilic, basic, acidic, and the like.

The terms "major histocompatibility complex" or "MHC" refers to a complex of genes encoding cell-surface molecules that are required for antigen presentation to T cells and for rapid graft rejection. In humans, the MHC complex is also known as the HLA complex. The proteins encoded by the MHC complex are known as "MHC molecules" and are classified into Class I and Class II MHC molecules. Class I MHC molecules include membrane heterodimeric proteins made up of an α chain encoded in the MHC associated noncovalently with β 2-microglobulin. Class I MHC molecules are expressed by nearly all nucleated cells and have been shown to function in antigen presentation to CD8+ T cells. Class I molecules include HLA-A, -B, and -C in humans. Class I molecules generally bind peptides 8-10 amino acids in length. Class II MHC molecules also include membrane heterodimeric proteins consisting of noncovalently associated α and β chains. Class II MHC are known to participate in antigen presentation to CD4+ T cells and, in humans, include HLA-DP, -DQ, and DR. Class II molecules generally bind peptides 12-20 amino acid residues in length. The term "MHC restriction" refers to a characteristic of T cells that permits them to recognize antigen only after it is processed and the resulting antigenic peptides are displayed in association with either a self Class I or Class II MHC molecule. Methods of identifying and comparing MHC are known in the art and are described in Allen et al. (1994) Human Imm. 40:25-32; Santamaria et al. (1993) Human Imm. 37:39-50 and Hurley et al. (1997) Tissue Antigens 50:401-415.

The term "antigen-presenting matrix," as used herein, intends a molecule or molecules which can present antigen in such a way that the antigen can be bound by a T-cell antigen receptor on the surface of a T cell. An antigen-presenting matrix can be on the surface of an antigen-presenting cell (APC), on a vesicle preparation of an APC, or can be in the form of a synthetic matrix on a

solid support such as a bead or a plate. An example of a synthetic antigen-presenting matrix is purified MHC Class I molecules complexed to β 2-microglobulin, or purified MHC Class II molecules, or functional portions thereof, attached to a solid support.

5 The term "antigen presenting cell," as used herein, intends any cell which presents on its surface an antigen in association with a major histocompatibility complex molecule, or portion thereof, or, alternatively, one or more non-classical MHC molecules, or a portion thereof. Examples of suitable APCs are discussed in detail below and include, but are not limited to, whole cells such as
10 macrophages, dendritic cells, B cells, hybrid APCs, and foster antigen presenting cells. Methods of making hybrid APCs have been described. See, for example, International Patent Application No. WO 98/46785; and WO 95/16775.

 Dendritic cells (DCs) are potent antigen-presenting cells. It has been shown that DCs provide all the signals required for T cell activation and
15 proliferation. These signals can be categorized into two types. The first type, which gives specificity to the immune response, is mediated through interaction between the T-cell receptor/CD3 ("TCR/CD3") complex and an antigenic peptide presented by a major histocompatibility complex ("MHC") Class I or II protein on the surface of APCs. This interaction is necessary, but not sufficient, for T
20 cell activation to occur. In fact, without the second type of signals, the first type of signals can result in T cell anergy. The second type of signals, called co-stimulatory signals, is neither antigen-specific nor MHC-restricted, and can lead to a full proliferation response of T cells and induction of T cell effector functions in the presence of the first type of signals. As used herein, "dendritic
25 cell" is to include, but not be limited to a pulsed dendritic cell, a foster cell or a dendritic cell hybrid.

 "Co-stimulatory molecules" are involved in the interaction between receptor-ligand pairs expressed on the surface of antigen presenting cells and T cells. Research accumulated over the past several years has demonstrated
30 convincingly that resting T cells require at least two signals for induction of cytokine gene expression and proliferation (Schwartz (1990) Science **248**:1349-1356; Jenkins (1992) Immunol. Today **13**:69-73). One signal, the one that

confers specificity, can be produced by interaction of the TCR/CD3 complex with an appropriate MHC/peptide complex. The second signal is not antigen specific and is termed the "co-stimulatory" signal. This signal was originally defined as an activity provided by bone-marrow-derived accessory cells such as macrophages and dendritic cells, the so called "professional" APCs. Several molecules have been shown to enhance co-stimulatory activity. These are heat stable antigen (HSA) (Liu et al. (1992) *J. Exp. Med.* **175**:437-445), chondroitin sulfate-modified MHC invariant chain (Ii-CS) (Naujokas et al. (1993) *Cell* **74**:257-268), intracellular adhesion molecule 1 (ICAM-1) (Van Seventer (1990) *J. Immunol.* **144**:4579-4586), B7-1, and B7-2/B70 (Schwartz (1992) *Cell* **71**:1065-1068). These molecules each appear to assist co-stimulation by interacting with their cognate ligands on the T cells. Co-stimulatory molecules mediate co-stimulatory signal(s) which are necessary, under normal physiological conditions, to achieve full activation of naïve T cells. One exemplary receptor-ligand pair is the B7 co-stimulatory molecule on the surface of APCs and its counter-receptor CD28 or CTLA-4 on T cells (Freeman et al. (1993) *Science* **262**:909-911; Young et al. (1992) *J. Clin. Invest.* **90**: 229; Nabavi et al. (1992) *Nature* **360**:266-268). Other important co-stimulatory molecules are CD40, CD54, CD80, CD86. The term "co-stimulatory molecule" encompasses any single molecule or combination of molecules which, when acting together with a peptide/MHC complex bound by a TCR on the surface of a T cell, provides a co-stimulatory effect which achieves activation of the T cell that binds the peptide. The term thus encompasses B7, or other co-stimulatory molecule(s) on an antigen-presenting matrix such as an APC, fragments thereof (alone, complexed with another molecule(s), or as part of a fusion protein) which, together with peptide/MHC complex, binds to a cognate ligand and results in activation of the T cell when the TCR on the surface of the T cell specifically binds the peptide. Co-stimulatory molecules are commercially available from a variety of sources, including, for example, Beckman Coulter. It is intended, although not always explicitly stated, that molecules having similar biological activity as wild-type or purified co-stimulatory molecules (e.g., recombinantly produced or muteins thereof) are intended to be used within the spirit and scope of the invention.

As used herein, "solid phase support" or "solid support," used interchangeably, is not limited to a specific type of support. Rather a large number of supports are available and are known to one of ordinary skill in the art. Solid phase supports include silica gels, resins, derivatized plastic films, glass beads, cotton, plastic beads, alumina gels. As used herein, "solid support" also includes synthetic antigen-presenting matrices, cells, and liposomes. A suitable solid phase support may be selected on the basis of desired end use and suitability for various protocols. For example, for peptide synthesis, solid phase support may refer to resins such as polystyrene (e.g., PAM-resin obtained from Bachem Inc., Peninsula Laboratories, etc.), POLYHIPE® resin (obtained from Aminotech, Canada), polyamide resin (obtained from Peninsula Laboratories), polystyrene resin grafted with polyethylene glycol (TentaGel®, Rapp Polymere, Tubingen, Germany) or polydimethylacrylamide resin (obtained from Milligen/Bioscience, California).

The term "immunomodulatory agent," as used herein, is a molecule, a macromolecular complex, or a cell that modulates an immune response. The term "modulate an immune response" includes inducing (increasing, eliciting) an immune response; and reducing (suppressing) an immune response. An immunomodulatory method (or protocol) is one that modulates an immune response in a subject.

The term "immune effector cells" refers to cells capable of binding an antigen or which mediate an immune response. These cells include, but are not limited to, T cells, B cells, monocytes, macrophages, NK cells and cytotoxic T lymphocytes (CTLs), for example CTL lines, CTL clones, and CTLs from tumor, inflammatory, or other infiltrates. Certain diseased tissue expresses specific antigens and CTLs specific for these antigens have been identified. For example, approximately 80% of melanomas express the antigen known as gp100.

The term "immune effector molecule," as used herein, refers to molecules capable of antigen-specific binding, and includes antibodies, T cell antigen receptors, and MHC Class I and Class II molecules.

A "naïve" immune effector cell is an immune effector cell that has never been exposed to an antigen.

As used herein, the term “educated, antigen-specific immune effector cell,” is an immune effector cell as defined above, which has encountered antigen and which is specific for that antigen. An educated, antigen-specific immune effector cell may be activated upon binding antigen. “Activated” implies that the cell is no longer in G₀ phase, and begins to produce cytokines characteristic of the cell type. For example, activated CD4⁺ T cells secrete IL-2 and have a higher number of high affinity IL-2 receptors on their cell surfaces relative to resting CD4⁺ T cells.

As used herein, the term “a disease or condition related to a population of CD4⁺ or CD8⁺ T cells” is one which can be related to a population of CD4⁺ or CD8⁺ T cells, such that these cells are primarily responsible for the pathogenesis of the disease; it is also one in which the presence of CD4⁺ or CD8⁺ T cells is an indicia of a disease state; it is also one in which the presence of a population CD4⁺ or CD8⁺ T cells is not the primary cause of the disease, but which plays a key role in the pathogenesis of the disease; it is also one in which a population of CD4⁺ or CD8⁺ T cells mediates an undesired rejection of a foreign antigen. Examples of a condition related to a population of CD4⁺ or CD8⁺ T cells include, but are not limited to, autoimmune disorders, graft rejection, immunoregulatory disorders, and anaphylactic disorders.

As used herein, the terms “pathological cells,” “neoplasia,” “tumor,” “tumor cells,” “cancer” and “cancer cells,” (used interchangeably) refer to cells which exhibit relatively autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation (i.e., de-regulated cell division). Pathological cells can be malignant or benign.

“Suppressing” tumor growth indicates a growth state that is curtailed when compared to growth without contact with educated, antigen-specific immune effector cells described herein. Tumor cell growth can be assessed by any means known in the art, including, but not limited to, measuring tumor size, determining whether tumor cells are proliferating using a ³H-thymidine incorporation assay, or counting tumor cells. “Suppressing” tumor cell growth

means any or all of the following states: slowing, delaying, and stopping tumor growth, as well as tumor shrinkage.

5 The term “culturing” refers to the *in vitro* propagation of cells or organisms on or in media of various kinds. It is understood that the descendants of a cell grown in culture may not be completely identical (morphologically, genetically, or phenotypically) to the parent cell. By “expanded” is meant any proliferation or division of cells.

10 As used herein, the term “cytokine” refers to any one of the numerous factors that exert a variety of effects on cells, for example, inducing growth or proliferation. Non-limiting examples of cytokines which may be used alone or in combination in the practice of the present invention include, interleukin-2 (IL-2), stem cell factor (SCF), interleukin-3 (IL-3), interleukin-6 (IL-6), interleukin-12 (IL-12), G-CSF, granulocyte macrophage-colony stimulating factor (GM-CSF), interleukin-1 alpha (IL-1 α), interleukin-11 (IL-11), MIP-1 α , leukemia inhibitory factor (LIF), c-kit ligand, thrombopoietin (TPO) and flt3 ligand. The present invention also includes culture conditions in which one or more cytokine is specifically excluded from the medium. Cytokines are commercially available from several vendors such as, for example, Genzyme (Framingham, MA), Genentech (South San Francisco, CA), Amgen (Thousand Oaks, CA), R&D Systems and Immunex (Seattle, WA). It is intended, although not always explicitly stated, that molecules having similar biological activity as wild-type or purified cytokines (e.g., recombinantly produced or muteins thereof) are intended to be used within the spirit and scope of the invention.

25 A “composition” is intended to mean a combination of active agent and another compound or composition, inert (for example, a detectable agent or label) or active, such as an adjuvant.

A “pharmaceutical composition” is intended to include the combination of an active agent with a carrier, inert or active, making the composition suitable for diagnostic or therapeutic use *in vitro*, *in vivo* or *ex vivo*.

30 As used herein, the term “pharmaceutically acceptable carrier” encompasses any of the standard pharmaceutical carriers, such as a phosphate

buffered saline solution, water, and emulsions, such as an oil/water or water/oil emulsion, and various types of wetting agents. The compositions also can include stabilizers and preservatives. For examples of carriers, stabilizers and adjuvants, see Martin, REMINGTON'S PHARM. SCI., 15th Ed. (Mack Publ. Co., Easton (1975)).

An "effective amount" is an amount sufficient to effect beneficial or desired results such as prevention or treatment. An effective amount can be administered in one or more administrations, applications or dosages.

A "subject," "individual" or "patient" is used interchangeably herein, which refers to a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are not limited to, murines, simians, humans, farm animals, sport animals, and pets.

A "control" is an alternative subject or sample used in an experiment for comparison purpose. A control can be "positive" or "negative". For example, where the purpose of the experiment is to determine a correlation of an altered expression level of a gene with a particular type of cancer, it is generally preferable to use a positive control (a subject or a sample from a subject, carrying such alteration and exhibiting syndromes characteristic of that disease), and a negative control (a subject or a sample from a subject lacking the altered expression and clinical syndrome of that disease).

DIAGNOSTIC METHODS

As noted above, this invention provides various methods for aiding in the diagnosis of the state of a cell or tissue that is characterized by differential expression of a protein identified in Table 1. The methods are particularly useful for aiding in the diagnosis of disorders relating to the immune response, e.g., cancers, autoimmune diseases and viral infections.

When referring to cancerous disorders, the pathological state of a cell or tissue generally is determined by noting whether the growth of the cell is not governed by the usual limitation of normal growth. For the purposes of this invention, the term also is to include genotypic changes that occur prior to detection of this growth in the form of a tumor and are causative of these

phenotypic changes. The phenotypic changes associated with the pathological state of a cell (a set of *in vitro* characteristics associated with a tumorigenic ability *in vivo*) include a more rounded cell morphology, looser substratum attachment, loss of contact inhibition, loss of anchorage dependence, release of proteases such as plasminogen activator, increased sugar transport, decreased serum requirement, expression of fetal antigens and the like. (See, Luria *et al.* (1978) GENERAL VIROLOGY, 3^d edition, 436-446 (John Wiley & Sons, New York)).

Accordingly, one embodiment is a method of diagnosing the condition of a cell or tissue by screening for the presence of a differentially expressed gene isolated from a sample. In one aspect, the gene is expressed more in a pathological cell or tissue (e.g., autoimmune disorders), and is selected from the proteins identified in Table 1. Detection can be by any appropriate method, including for example, detecting the quantity of mRNA transcribed from the gene, or the quantity of cDNA produced from the reverse transcription of the mRNA transcribed from the gene, or the quantity of the polypeptide or protein encoded by the gene. Probes for each of these methods are provided in Table 1. These methods can be performed on a sample by sample basis or modified for high throughput analysis. Additionally, databases containing quantitative full or partial transcripts or protein sequences isolated from a cell sample can be searched and analyzed for the presence and amount of transcript or expressed gene product. In one aspect, the database contains at least one of the sequences shown in Table 1.

For the purpose of illustration only, gene expression is determined by noting the amount (if any, e.g., altered) expression of the gene in the test system at the level of an mRNA transcribed from at least one gene identified in Table 1. In a separate embodiment, augmentation of the level of the polypeptide or protein encoded by the gene is indicative of the presence of the pathological condition of the cell or tissue. For example, an enhanced or increase in the expression level of a gene identified in Table 1 may indicate infection or alternatively, an over-stimulated immune response that occurs, for example, in certain autoimmune disorders. In yet a further embodiment, a decrease in the level of polypeptide or

protein encoded by the gene is indicative of the pathological condition, e.g., cancer. The method can be used for aiding in the diagnosis of cancer by detecting a genotype that is correlated with a phenotype characteristic of primary tumor cells. Thus, by detecting this genotype prior to tumor growth, one can
5 predict a predisposition to cancer and/or provide early diagnosis and treatment.

Cell or tissue samples used for this invention encompass body fluid, solid tissue samples, tissue cultures or cells derived there from and the progeny thereof, and sections or smears prepared from any of these sources, or any other samples that may contain a cell having a gene described herein. In one
10 embodiment, the sample comprises cells prepared from a subject's tissue.

In assaying for an alteration in mRNA level, nucleic acid contained in the aforementioned samples is first extracted according to standard methods in the art. For instance, mRNA can be isolated using various lytic enzymes or chemical solutions according to the procedures set forth in Sambrook *et al.* (1989) *supra*,
15 or extracted by nucleic-acid-binding resins following the accompanying instructions provided by manufactures. The mRNA of a proto-oncogene of interest contained in the extracted nucleic acid sample is then detected by hybridization (e.g., Northern blot analysis) and/or amplification procedures according to methods widely known in the art or based on the methods
20 exemplified herein.

Nucleic acid molecules having at least 10 nucleotides and exhibiting sequence complementarity or homology to at least one gene identified in Table 1 find utility as hybridization probes. It is known in the art that a "perfectly matched" probe is not needed for a specific hybridization. Minor changes in
25 probe sequence achieved by substitution, deletion or insertion of a small number of bases do not affect the hybridization specificity. In general, as much as 20% base-pair mismatch (when optimally aligned) can be tolerated. Preferably, a probe useful for detecting mRNA is at least about 80% identical to the homologous region of comparable size contained in the genes or polynucleotides
30 identified in Table 1 identified sequences, which have the Locus Link numbers identified in Table 1. In one aspect, the probe is 85% identical to the corresponding gene sequence after alignment of the homologous region, or

alternatively, it exhibits 90% identity. Additional probes can be derived from sequences for the genes identified by the Locus Link Nos. provided in Table 1, or to a homologous region of comparable size contained in the previously identified sequences, which have the Locus Link Nos. identified in Table 1. These probes
5 can be used in radioassays (*e.g.*, Southern and Northern blot analysis) to detect, prognose, diagnose or monitor various pathological states resulting from differential expression of a gene of interest. The total size of fragment, as well as the size of the complementary stretches, will depend on the intended use or application of the particular nucleic acid segment. Smaller fragments derived
10 from the known sequences will generally find use in hybridization embodiments, wherein the length of the complementary region may be varied, such as between about 10 and about 100 nucleotides, or even full length according to the complementary sequences one wishes to detect.

In one aspect, nucleotide probes having complementary sequences over
15 stretches greater than about 10 nucleotides in length are used, so as to increase stability and selectivity of the hybrid, and thereby improving the specificity of particular hybrid molecules obtained. Alternatively, one can design nucleic acid molecules having gene-complementary stretches of more than about 25 or
alternatively more than about 50 nucleotides in length, or even longer where
20 desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCRTM technology with two priming oligonucleotides as described in U.S. Pat. No. 4,603,102 or by introducing
selected sequences into recombinant vectors for recombinant production. In one
25 aspect, a probe is about 50 to about 75, nucleotides or alternatively, about 50 to about 100 nucleotides in length.

In certain embodiments, it will be advantageous to employ nucleic acid sequences as described herein in combination with an appropriate means, such as a label, for detecting hybridization and therefore complementary sequences. A
30 wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. One can employ a fluorescent label or

an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

Hybridization reactions can be performed under conditions of different "stringency". Relevant conditions include temperature, ionic strength, time of incubation, the presence of additional solutes in the reaction mixture such as formamide, and the washing procedure. Higher stringency conditions are those conditions, such as higher temperature and lower sodium ion concentration, which require higher minimum complementarity between hybridizing elements for a stable hybridization complex to form. Conditions that increase the stringency of a hybridization reaction are widely known and published in the art. See, for example, Sambrook *et al.* (1989) *supra*.

The nucleotide probes of the present invention can also be used as primers and detection of genes or gene transcripts that are differentially expressed in certain body tissues. Additionally, a primer useful for detecting the aforementioned differentially expressed mRNA is at least about 80% identical to the homologous region of comparable size contained in the previously identified sequences, which have the Locus Link Nos. numbers identified in Table 1. For the purpose of this invention, amplification means any method employing a primer-dependent polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA-polymerases such as T7 DNA polymerase, Klenow fragment of *E.coli* DNA polymerase, and reverse transcriptase.

A known amplification method is PCR, MacPherson *et al.*, PCR: A PRACTICAL APPROACH, (IRL Press at Oxford University Press (1991)). However, PCR conditions used for each application reaction are empirically determined. A number of parameters influence the success of a reaction. Among them are annealing temperature and time, extension time, Mg^{2+} ATP concentration, pH, and the relative concentration of primers, templates, and deoxyribonucleotides.

After amplification, the resulting DNA fragments can be detected by agarose gel electrophoresis followed by visualization with ethidium bromide staining and ultraviolet illumination. A specific amplification of differentially expressed genes of interest can be verified by demonstrating that the amplified DNA fragment has the predicted size, exhibits the predicated restriction digestion pattern, and/or hybridizes to the correct cloned DNA sequence.

The probes also can be attached to a solid support for use in high throughput screening assays using methods known in the art. PCT WO 97/10365 and U.S. Patent Nos. 5,405,783; 5,412,087 and 5,445,934; for example, disclose the construction of high density oligonucleotide chips which can contain one or more of the sequences disclosed herein. Using the methods disclosed in U.S. Patent Nos. 5,405,783; 5,412,087 and 5,445,934; the probes of this invention are synthesized on a derivatized glass surface. Photoprotected nucleoside phosphoramidites are coupled to the glass surface, selectively deprotected by photolysis through a photolithographic mask, and reacted with a second protected nucleoside phosphoramidite. The coupling/deprotection process is repeated until the desired probe is complete.

The expression level of a gene can also be determined through exposure of a nucleic acid sample to a probe-modified chip. Extracted nucleic acid is labeled, for example, with a fluorescent tag, preferably during an amplification step. Hybridization of the labeled sample is performed at an appropriate stringency level. The degree of probe-nucleic acid hybridization is quantitatively measured using a detection device, such as a confocal microscope. See U.S. Pat Nos. 5,578,832 and 5,631,734. The obtained measurement is directly correlated with gene expression level.

The probes and high density oligonucleotide probe arrays also provide an effective means of monitoring expression of the genes identified in Table 1. They are also useful to screen for compositions that upregulate or downregulate the expression of the genes identified in Table 1.

In another embodiment, the methods of this invention are used to monitor expression of the genes identified in Table 1 which specifically hybridize to the

probes of this invention in response to defined stimuli, such as an exposure of a cell or subject to a drug.

In one embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample nucleic acids. The labels may be incorporated by any of a number of means known to those of skill in the art. However, in one aspect, the label is simultaneously incorporated during the amplification step in the preparation of the sample nucleic acid. Thus, for example, polymerase chain reaction (PCR) with labeled primers or labeled nucleotides will provide a labeled amplification product. In a separate embodiment, transcription amplification, as described above, using a labeled nucleotide (*e.g.*, fluorescein-labeled UTP and/or CTP) incorporates a label in to the transcribed nucleic acids.

Alternatively, a label may be added directly to the original nucleic acid sample (*e.g.*, mRNA, polyA, mRNA, cDNA, *etc.*) or to the amplification product after the amplification is completed. Means of attaching labels to nucleic acids are known to those of skill in the art and include, for example nick translation or end-labeling (*e.g.*, with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (*e.g.*, a fluorophore).

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (*e.g.*, Dynabeads™), fluorescent dyes (*e.g.*, fluorescein, Texas red, rhodamine, green fluorescent protein, and the like), radiolabels (*e.g.*, ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P) enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads. Patents teaching the use of such labels include U.S. Patents Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

Means of detecting such labels are known to those of skill in the art. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label.

As described in more detail in WO 97/10365, the label may be added to the target (sample) nucleic acid(s) prior to, or after the hybridization. These are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected. For a detailed review of methods of labeling nucleic acids and detecting labeled hybridized nucleic acids see LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY, Vol. 24: Hybridization with Nucleic Acid Probes, P. Tijssen, ed. Elsevier, N.Y. (1993).

The nucleic acid sample also may be modified prior to hybridization to the high density probe array in order to reduce sample complexity thereby decreasing background signal and improving sensitivity of the measurement using methods known in the art, *e.g.*, the methods disclosed in WO 97/10365.

Results from the chip assay are typically analyzed using a computer software program. *See*, for example, EP 0717 113 A2 and WO 95/20681. The hybridization data is read into the program, which calculates the expression level of the targeted gene(s) *i.e.*, the genes identified in Table 1. This figure is compared against existing data sets of gene expression levels for diseased and healthy individuals. A correlation between the obtained data and that of a set of diseased individuals indicates the onset of a disease in the subject patient.

Also within the scope of this application is a data base useful for the detection of pathological tissue comprising one or more of the sequences (or parts thereof) of the genes listed Table 1.

5 These polynucleotide sequences are stored in a digital storage medium such that a data processing system for standardized representation of the genes that correlate to a pathological condition. The data processing system is useful to analyze gene expression between two cells by first selecting a cell suspected of being of a pathological phenotype or genotype and then isolating polynucleotides from the cell or tissue. The isolated polynucleotides are then sequenced. The
10 sequences from the sample are compared with the sequence(s) present in the database using homology search techniques described above. In one aspect, greater than 90% is selected, or alternatively greater than 95% is selected, or alternatively greater than or equal to 97% sequence identity is selected, between the test sequence and at least one sequence identified in Table 1 or its
15 complement, is a positive indication that the polynucleotide has been isolated from a cell or tissue as defined above.

Alternatively, one can compare a sample against a database. Briefly, multiple RNAs are isolated from cell or tissue samples using methods known in the art and described for example, in Sambrook *et al.* (1989) *supra*. Optionally,
20 the gene transcripts can be converted to cDNA. A sampling of the gene transcripts are subjected to sequence-specific analysis and quantified. These gene transcript sequence abundances are compared against reference database sequence abundances including normal data sets for diseased and healthy patients. The patient has the disease(s) with which the patient's data set most
25 closely correlates which includes the expression of the transcripts identified herein.

Differential expression of the genes of interest can also be determined by examining the protein product. A variety of techniques are available in the art for protein analysis. They include but are not limited to radioimmunoassays, ELISA
30 (enzyme linked immunoradiometric assays), "sandwich" immunoassays, immunoradiometric assays, *in situ* immunoassays (using *e.g.*, colloidal gold, enzyme or radioisotope labels), western blot analysis, immunoprecipitation

assays, immunofluorescent assays, and PAGE-SDS. One means to determine protein level involves (a) providing a biological sample containing polypeptides; and (b) measuring the amount of any immunospecific binding that occurs between an antibody reactive to the expression product of a gene of interest and a component in the sample, in which the amount of immunospecific binding indicates the level of the expressed proteins.

Antibodies that specifically recognize and bind to the protein products of these genes are required for these immunoassays. These may be purchased from commercial vendors or generated and screened using methods known in the art. See Harlow and Lane (1988) *supra*. and Sambrook *et al.* (1989) *supra*. Alternatively, polyclonal or monoclonal antibodies that specifically recognize and bind the protein product of a gene of interest can be made and isolated using known methods.

In diagnosing pathologies characterized by a differential expression of genes, one typically conducts a comparative analysis of the subject and appropriate controls. Preferably, a diagnostic test includes a control sample derived from a subject (hereinafter "positive control"), that exhibits the predicted change in expression of a gene of interest, *e.g.*, at a level of at least 2.5 fold and clinical characteristics of the pathology of interest. Alternatively, a diagnosis also includes a control sample derived from a subject (hereinafter "negative control"), that lacks the clinical characteristics of the pathological state and whose expression level of the gene at question is within a normal range. A positive correlation between the subject and the positive control with respect to the identified alterations indicates the presence of or a predisposition to said disease. A lack of correlation between the subject and the negative control confirms the diagnosis.

There are various methods available in the art for quantifying mRNA or protein level from a cell sample and indeed, any method that can quantify these levels is encompassed by this invention. For example, determination of the mRNA level of the aforementioned genes may involve, in one aspect, measuring the amount of mRNA in a sample isolated from the cell by hybridization or quantitative amplification using at least one oligonucleotide probe that is

complementary to the mRNA. Determination of the aforementioned gene products requires measuring the amount of immunospecific binding that occurs between an antibody reactive to the gene product of a gene identified in Table 1. To detect and quantify the immunospecific binding, or signals generated during hybridization or amplification procedures, digital image analysis systems including but not limited to those that detect radioactivity of the probes or chemiluminescence can be employed.

SCREENING ASSAYS

The present invention also provides a screen for identifying leads, drugs, therapeutic biologics, and methods for reversing the pathological condition of the subject by modifying the immune response. In one aspect, the screen identifies lead compounds or biological agents which are useful for the treatment of malignancy, hyperplasia metaplasia, autoimmune diseases and viral infections characterized by differential expression of a gene identified in Table 1.

Thus, to practice the method *in vitro*, suitable cell cultures or tissue cultures are first provided. The cell can be a cultured cell or a genetically modified cell which differentially expresses a gene associated with a pathological cell *e.g.*, at least one gene identified in Table 1. Alternatively, the cells can be from a tissue biopsy. The cells are cultured under conditions (temperature, growth or culture medium and gas (CO₂)) and for an appropriate amount of time to attain exponential proliferation without density dependent constraints. It also is desirable to maintain an additional separate cell culture; one which does not receive the agent being tested as a control.

As is apparent to one of skill in the art, the method can be modified for high throughput analysis and suitable cells may be cultured in microtiter plates and several agents may be assayed at the same time by noting genotypic changes, phenotypic changes and/or cell death.

When the agent is a composition other than a DNA or RNA nucleic acid molecule, the suitable conditions comprise directly added to the cell culture or added to culture medium for addition. As is apparent to those skilled in the art, an "effective" amount must be added which can be empirically determined.

The screen involves contacting the agent with a test cell characterized by differential expression of a gene of interest and then assaying the cell for the level of said gene expression. In some aspects, it may be necessary to determine the level of gene expression prior to the assay. This provides a base line to compare expression after administration of the agent to the cell culture. In another embodiment, the test cell is a cultured cell from an established cell line that differentially expresses a gene of interest. An agent is a possible therapeutic agent if gene expression is returned (reduced or increased) to a level that is present in a cell in a normal or non-pathological state, or the cell selectively dies, or exhibits reduced rate of growth.

In yet another aspect, the test cell or tissue sample is isolated from the subject to be treated and one or more potential agents are screened to determine the optimal therapeutic and/or course of treatment for that individual patient.

For the purposes of this invention, an "agent" is intended to include, but not be limited to a biological or chemical compound such as a simple or complex organic or inorganic molecule, a peptide, a protein or an oligonucleotide. A vast array of compounds can be synthesized, for example oligomers, such as oligopeptides and oligonucleotides, and synthetic organic compounds based on various core structures, and these are also included in the term "agent". In addition, various natural sources can provide compounds for screening, such as plant or animal extracts, and the like. It should be understood, although not always explicitly stated that the agent is used alone or in combination with another agent, having the same or different biological activity as the agents identified by the inventive screen. The agents and methods also are intended to be combined with other therapies.

As noted above, cells having differential expression of a gene of interest that results in the pathological state are suitably treated by this method. These cells can be identified by any method known in the art that allows for the identification of differential expression of the gene.

When the agent is a nucleic acid, it can be added to the cell cultures by methods known in the art, which includes, but is not limited to calcium phosphate precipitation, microinjection or electroporation. Alternatively or additionally, the

nucleic acid can be incorporated into an expression or insertion vector for incorporation into the cells. Vectors that contain both a promoter and a cloning site into which a polynucleotide can be operatively linked are known in the art and briefly described *infra*.

5 Polynucleotides are inserted into vector genomes using methods known in the art. For example, insert and vector DNA can be contacted, under suitable conditions, with a restriction enzyme to create complementary ends on each molecule that can pair with each other and be joined together with a ligase. Alternatively, synthetic nucleic acid linkers can be ligated to the termini of
10 restricted polynucleotide. These synthetic linkers contain nucleic acid sequences that correspond to a particular restriction site in the vector DNA. Additionally, an oligonucleotide containing a termination codon and an appropriate restriction site can be ligated for insertion into a vector containing, for example, some or all of the following: a selectable marker gene, such as the neomycin gene for
15 selection of stable or transient transfectants in mammalian cells; enhancer/promoter sequences from the immediate early gene of human CMV for high levels of transcription; transcription termination and RNA processing signals from SV40 for mRNA stability; SV40 polyoma origins of replication and ColE1 for proper episomal replication; versatile multiple cloning sites; and T7 and SP6
20 RNA promoters for *in vitro* transcription of sense and antisense RNA. Other means are known and available in the art.

One can determine if the object of the method, *i.e.*, reversal of the pathological state of the cell, using methods known in the art, some of which are described herein.

25 When the subject is an animal such as a rat or mouse, the method provides a convenient animal model system which can be used prior to clinical testing of the therapeutic agent or alternatively, for lead optimization. In this system, a candidate agent is a potential drug if gene expression is returned to a normal level or if symptoms associated or correlated to the presence of cells containing
30 differential expression of a gene of interest are ameliorated, each as compared to untreated, animal having the pathological cells. It also can be useful to have a

separate negative control group of cells or animals which are healthy and not treated, which provides a basis for comparison.

THERAPEUTIC METHODS

5 The polypeptides and polynucleotides of the invention are useful to modulate immune responses by administration to a subject, alone or in combination with vaccines, immune effector cells, antigen presenting cells, immunogens, and other known biological modifiers. For the purpose of illustration only, suitable biological modifiers include, but are not limited to, cytokines such as IL-2, IL-4, IL-10, TNF- α , IL-12, IFN- γ ; non-specific modifiers
10 such as phytohemagglutinin (PHA), phorbol esters such as phorbol myristate acetate (PMA), concanavalin-A, and ionomycin; antibodies specific for cell surface markers, such as anti-CD2, anti-CD3, anti-IL-2 receptor, anti-CD28; chemokines, including, for example, lymphotactin. The biological modifiers can be native factors obtained from natural sources, factors produced by recombinant
15 DNA technology, chemically synthesized polypeptides or other molecules, or any derivative having the functional activity of the native factor. If more than one biological modifier is used, the exposure can be simultaneous or sequential.

In one embodiment, the composition of this invention is administered in an amount effective to modulate the immune response. In a further preferred
20 embodiment, an agent of the invention is administered in an amount effective to treat cancer, autoimmune disease or viral infections. The compositions of the invention can also be used to prevent progression from a pre-pathological or non-malignant state into a pathological or a malignant state.

Various delivery systems are known and can be used to administer a
25 composition or other therapeutic agent of the invention, *e.g.*, encapsulation in liposomes, microparticles, microcapsules, expression by recombinant cells, receptor-mediated endocytosis (*See, e.g.*, Wu and Wu, (1987), J. Biol. Chem. **262**:4429-4432), construction of a therapeutic nucleic acid as part of a retroviral or other vector, *etc.* Methods of delivery include but are not limited to intra-
30 arterial, intra-muscular, intravenous, intranasal, and oral routes. In a specific embodiment, it may be desirable to administer the pharmaceutical compositions

of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, by injection, or by means of a catheter.

5 The compositions can be administered to subjects or individuals susceptible to or at risk of developing a disease or pathological condition. When the composition is administered to a subject such as a mouse, a rat or a human patient, the composition can be added to a pharmaceutically acceptable carrier and systemically or topically administered to the subject. To determine patients that can be beneficially treated, a tumor sample is removed from the patient and
10 the cells are assayed for the differential expression of the gene. Therapeutic amounts can be empirically determined and will vary with the pathology being treated, the subject being treated and the efficacy and toxicity of the composition. When delivered to an animal, the method is useful to further confirm efficacy of the composition. As an example of an animal model, groups of nude mice
15 (Balb/c NCR nu/nu female, Simonsen, Gilroy, CA) are each subcutaneously inoculated with about 10^5 to about 10^9 hyperproliferative, cancer or target cells as defined herein. When the tumor is established, the composition is administered, for example, by subcutaneous injection around the tumor. Tumor measurements to determine reduction of tumor size are made in two dimensions using venier
20 calipers twice a week. Other animal models may also be employed as appropriate.

Administration *in vivo* can be effected in one dose, continuously or intermittently throughout the course of treatment. Methods of determining the most effective means and dosage of administration are known to those of skill in
25 the art and will vary with the composition used for therapy, the purpose of the therapy, the target cell being treated, and the subject being treated. Single or multiple administrations can be carried out with the dose level and pattern being selected by the treating physician. Suitable dosage formulations and methods of administering the compositions can be found below.

30 The compositions of the present invention can be used in the manufacture of medicaments and for the treatment of humans and other animals by

administration in accordance with conventional procedures, such as an active ingredient in pharmaceutical compositions.

5 The pharmaceutical compositions can be administered orally, intranasally, parenterally or by inhalation therapy, and may take the form of tablets, lozenges, granules, capsules, pills, ampoules, suppositories or aerosol form. They may also take the form of suspensions, solutions and emulsions of the active ingredient in aqueous or nonaqueous diluents, syrups, granulates or powders. In addition to an agent of the present invention, the pharmaceutical compositions can also contain other pharmaceutically active compounds or a plurality of compounds of the invention.

10 More particularly, an agent of the present invention also referred to herein as the active ingredient, may be administered for therapy by any suitable route including oral, rectal, nasal, topical (including transdermal, aerosol, buccal and sublingual), vaginal, parental (including subcutaneous, intramuscular, intravenous and intradermal) and pulmonary. It will also be appreciated that the preferred route will vary with the condition and age of the recipient, and the disease being treated.

15 Ideally, the composition should be administered to achieve peak concentrations of the active compound at sites of disease. This may be achieved, for example, by the intravenous injection of the composition, optionally in saline, or orally administered, for example, as a tablet, capsule or syrup containing the active ingredient. Desirable blood levels of the composition may be maintained by a continuous infusion to provide a therapeutic amount of the active ingredient within disease tissue. The use of operative combinations is contemplated to provide therapeutic combinations requiring a lower total dosage of each component than may be required when each individual therapeutic compound or drug is used alone, thereby reducing adverse effects.

20 While it is possible for the composition to be administered alone, it is preferable to present it as a pharmaceutical formulation comprising at least one active ingredient, as defined above, together with one or more pharmaceutically acceptable carriers therefor and optionally other therapeutic compositions. Each

carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient.

Formulations include those suitable for oral, rectal, nasal, topical (including transdermal, buccal and sublingual), vaginal, parenteral (including subcutaneous, intramuscular, intravenous and intradermal) and pulmonary administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods known in the art of pharmacy. Such methods include the step of bringing into association the active ingredient with the carrier which constitutes one or more accessory ingredients. In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers or finely divided solid carriers or both, and then if necessary shaping the product.

Formulations of the present invention suitable for oral administration may be presented as discrete units such as capsules, cachets or tablets, each containing a predetermined amount of the active ingredient; as a powder or granules; as a solution or suspension in an aqueous or non-aqueous liquid; or as an oil-in-water liquid emulsion or a water-in-oil liquid emulsion. The active ingredient may also be presented a bolus, electuary or paste.

A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets may be prepared by compressing in a suitable machine the active ingredient in a free-flowing form such as a powder or granules, optionally mixed with a binder (*e.g.*, povidone, gelatin, hydroxypropylmethyl cellulose), lubricant, inert diluent, preservative, disintegrant (*e.g.*, sodium starch glycolate, cross-linked povidone, cross-linked sodium carboxymethyl cellulose) surface-active or dispersing agent. Molded tablets may be made by molding in a suitable machine a mixture of the powdered compound moistened with an inert liquid diluent. The tablets may optionally be coated or scored and may be formulated so as to provide slow or controlled release of the active ingredient therein using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile. Tablets may optionally be provided with an enteric coating, to provide release in parts of the gut other than the stomach.

Formulations suitable for topical administration in the mouth include lozenges comprising the active ingredient in a flavored basis, usually sucrose and acacia or tragacanth; pastilles comprising the active ingredient in an inert basis such as gelatin and glycerin, or sucrose and acacia; and mouthwashes comprising the active ingredient in a suitable liquid carrier.

Pharmaceutical compositions for topical administration according to the present invention may be formulated as an ointment, cream, suspension, lotion, powder, solution, past, gel, spray, aerosol or oil. Alternatively, a formulation may comprise a patch or a dressing such as a bandage or adhesive plaster impregnated with active ingredients and optionally one or more excipients or diluents.

If desired, the aqueous phase of the cream base may include, for example, at least about 30% w/w of a polyhydric alcohol, *i.e.*, an alcohol having two or more hydroxyl groups such as propylene glycol, butane-1,3-diol, mannitol, sorbitol, glycerol and polyethylene glycol and mixtures thereof. The topical formulations may desirably include a compound which enhances absorption or penetration of the composition through the skin or other affected areas. Examples of such dermal penetration enhancers include dimethylsulfoxide and related analogues.

The oily phase of the emulsions of this invention may be constituted from known ingredients in a known manner. While this phase may comprise merely an emulsifier (otherwise known as an emulgent), it desirably comprises a mixture of at least one emulsifier with a fat or an oil or with both a fat and an oil. Preferably, a hydrophilic emulsifier is included together with a lipophilic emulsifier which acts as a stabilizer. It is also preferred to include both an oil and a fat. Together, the emulsifier(s) with or without stabilizer(s) make up the so-called emulsifying wax, and the wax together with the oil and/or fat make up the so-called emulsifying ointment base which forms the oily dispersed phase of the cream formulations.

Emulgents and emulsion stabilizers suitable for use in the formulation of the present invention include Tween 60, Span 80, cetostearyl alcohol, myristyl alcohol, glyceryl monostearate and sodium lauryl sulphate.

The choice of suitable oils or fats for the formulation is based on achieving the desired cosmetic properties, since the solubility of the active compound in most oils likely to be used in pharmaceutical emulsion formulations is very low. Thus the cream should preferably be a non-greasy, non-staining and washable product with suitable consistency to avoid leakage from tubes or other containers. Straight or branched chain, mono- or dibasic alkyl esters such as diisoadipate, isocetyl stearate, propylene glycol diester of coconut fatty acids, isopropyl myristate, decyl oleate, isopropyl palmitate, butyl stearate, 2-ethylhexyl palmitate or a blend of branched chain esters known as Crodamol CAP may be used, the last three being preferred esters. These may be used alone or in combination depending on the properties required. Alternatively, high melting point lipids such as white soft paraffin and/or liquid paraffin or other mineral oils can be used.

Formulations suitable for topical administration to the eye also include eye drops wherein the active ingredient is dissolved or suspended in a suitable carrier, especially an aqueous solvent for the composition.

Formulations for rectal administration may be presented as a suppository with a suitable base comprising, for example, cocoa butter or a salicylate.

Formulations suitable for vaginal administration may be presented as pessaries, tampons, creams, gels, pastes, foams or spray formulations containing in addition to the composition, such carriers as are known in the art to be appropriate.

Formulations suitable for nasal administration, wherein the carrier is a solid, include a coarse powder having a particle size, for example, in the range of about 20 to about 500 microns which is administered in the manner in which snuff is taken, *i.e.*, by rapid inhalation through the nasal passage from a container of the powder held close up to the nose. Suitable formulations wherein the carrier is a liquid for administration as, for example, nasal spray, nasal drops, or by aerosol administration by nebulizer, include aqueous or oily solutions of the composition.

Formulations suitable for parenteral administration include aqueous and non-aqueous isotonic sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents, and liposomes or other microparticulate systems which are designed to target the compound to blood components or one or more organs. The formulations may be presented in unit-dose or multi-dose sealed containers, for example, ampoules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets of the kind previously described.

Preferred unit dosage formulations are those containing a daily dose or unit, daily subdose, as herein above-recited, or an appropriate fraction thereof, of an composition.

It should be understood that in addition to the ingredients particularly mentioned above, the formulations of this invention may include other agents conventional in the art having regard to the type of formulation in question, for example, those suitable for oral administration may include such further agents as sweeteners, thickeners and flavoring agents. It also is intended that the agents, compositions and methods of this invention be combined with other suitable compositions and therapies.

TRANSGENIC ANIMALS

In another aspect, the genes of Table 1 can be used to generate transgenic animal models. In recent years, geneticists have succeeded in creating transgenic animals, for example mice, by manipulating the genes of developing embryos and introducing foreign genes into these embryos. Once these genes have integrated into the genome of the recipient embryo, the resulting embryos or adult animals can be analyzed to determine the function of the gene. The mutant animals are produced to understand the function of known genes *in vivo* and to create animal models of human diseases. (See, e.g., Chisaka *et al.* (1992) 355:516-520; Joyner

et al. (1992) in POSTIMPLANTATION DEVELOPMENT IN THE MOUSE (Chadwick and Marsh, eds., John Wiley & Sons, United Kingdom) pp:277-297; Dorin *et al.* (1992) Nature **359**:211-215).

U.S. Patent Nos. 5,464,764 and 5,487,992 describe one type of transgenic animal in which the gene of interest is deleted or mutated sufficiently to disrupt its function. (*See, also* U.S. Patent Nos. 5,631,153 and 5,627,059). These “knock-out” animals, made by taking advantage of the phenomena of homologous recombination, can be used to study the function of a particular gene sequence *in vivo*.

ANTIBODIES

Also provided by this invention is an antibody capable of specifically forming a complex with the expression product of a gene of interest. The term “antibody” includes polyclonal antibodies and monoclonal antibodies. The antibodies include, but are not limited to mouse, rat, and rabbit or human antibodies. The antibodies are useful to identify and purify gene expression products as well as APCs expressing the polypeptides.

Laboratory methods for producing polyclonal antibodies and monoclonal antibodies, as well as deducing their corresponding nucleic acid sequences, are known in the art, *see* Harlow and Lane (1988) *supra* and Sambrook *et al.* (1989) *supra*. The monoclonal antibodies of this invention can be biologically produced by introducing protein or a fragment thereof into an animal, *e.g.*, a mouse or a rabbit. The antibody producing cells in the animal are isolated and fused with myeloma cells or hetero-myeloma cells to produce hybrid cells or hybridomas. Accordingly, the hybridoma cells producing the monoclonal antibodies of this invention also are provided.

Thus, using the protein or fragment thereof, and known methods, one of skill in the art can produce and screen the hybridoma cells and antibodies of this invention for antibodies having the ability to bind the proteins or polypeptides.

If a monoclonal antibody being tested binds with the protein or polypeptide, then the antibody being tested and the antibodies provided by the hybridomas of this invention are equivalent. It also is possible to determine

without undue experimentation, whether an antibody has the same specificity as the monoclonal antibody of this invention by determining whether the antibody being tested prevents a monoclonal antibody of this invention from binding the protein or polypeptide with which the monoclonal antibody is normally reactive.

5 If the antibody being tested competes with the monoclonal antibody of the invention as shown by a decrease in binding by the monoclonal antibody of this invention, then it is likely that the two antibodies bind to the same or a closely related epitope. Alternatively, one can pre-incubate the monoclonal antibody of this invention with a protein with which it is normally reactive, and determine if

10 the monoclonal antibody being tested is inhibited in its ability to bind the antigen. If the monoclonal antibody being tested is inhibited then, in all likelihood, it has the same, or a closely related, epitopic specificity as the monoclonal antibody of this invention.

The term "antibody" also is intended to include antibodies of all isotypes.

15 Particular isotypes of a monoclonal antibody can be prepared either directly by selecting from the initial fusion, or prepared secondarily, from a parental hybridoma secreting a monoclonal antibody of different isotype by using the sib selection technique to isolate class switch variants using the procedure described in Steplewski *et al.* (1985) Proc. Natl. Acad. Sci. USA **82**:8653 or Spira *et al.*

20 (1984) J. Immunol. Meth. **74**:307.

This invention also provides biological active fragments of the polyclonal and monoclonal antibodies described above. These "antibody fragments" retain some ability to selectively bind with its antigen or immunogen. Such antibody fragments can include, but are not limited to:

- 25
- (1) Fab,
 - (2) Fab',
 - (3) F(ab')₂,
 - (4) Fv, and
 - (5) SCA

A specific example of “a biologically active antibody fragment” is a CDR region of the antibody. Methods of making these fragments are known in the art, see for example, Harlow and Lane (1988) *supra*.

5 The antibodies of this invention also can be modified to create chimeric antibodies and humanized antibodies (Oi *et al.* (1986) *BioTechniques* 4(3):214). Chimeric antibodies are those in which the various domains of the antibodies’ heavy and light chains are coded for by DNA from more than one species.

10 The isolation of other hybridomas secreting monoclonal antibodies with the specificity of the monoclonal antibodies of the invention can also be accomplished by one of ordinary skill in the art by producing anti-idiotypic antibodies (Herlyn *et al.* (1986) *Science* 232:100). An anti-idiotypic antibody is an antibody which recognizes unique determinants present on the monoclonal antibody produced by the hybridoma of interest.

15 Idiotypic identity between monoclonal antibodies of two hybridomas demonstrates that the two monoclonal antibodies are the same with respect to their recognition of the same epitopic determinant. Thus, by using antibodies to the epitopic determinants on a monoclonal antibody it is possible to identify other hybridomas expressing monoclonal antibodies of the same epitopic specificity.

20 It is also possible to use the anti-idiotypic technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the mirror image of the epitope bound by the first monoclonal antibody. Thus, in this instance, the anti-idiotypic monoclonal antibody could be used for immunization for production of these
25 antibodies.

As used in this invention, the term “epitope” is meant to include any determinant having specific affinity for the monoclonal antibodies of the invention. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have
30 specific three dimensional structural characteristics, as well as specific charge characteristics.

The antibodies of this invention can be linked to a detectable agent or label. There are many different labels and methods of labeling known to those of ordinary skill in the art.

5 The coupling of antibodies to low molecular weight haptens can increase the sensitivity of the assay. The haptens can then be specifically detected by means of a second reaction. For example, it is common to use haptens such as biotin, which reacts avidin, or dinitrophenol, pyridoxal, and fluorescein, which can react with specific anti-hapten antibodies. *See*, Harlow and Lane (1988) *supra*.

10 The monoclonal antibodies of the invention also can be bound to many different carriers. Thus, this invention also provides compositions containing the antibodies and another substance, active or inert. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and
15 magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such, using routine experimentation.

20 Compositions containing the antibodies, fragments thereof or cell lines which produce the antibodies, are encompassed by this invention. When these compositions are to be used pharmaceutically, they are combined with a pharmaceutically acceptable carrier.

25 Specific combination(s) of cytokines have been used successfully to amplify (or partially substitute) for the activation/conversion achieved with calcium ionophore: these cytokines include but are not limited to purified or recombinant human ("rh") rhGM-CSF, rhIL-2, and rhIL-4. Each cytokine when given alone is inadequate for optimal upregulation.

MATERIALS AND METHODS

Phenotypic Characterization of Invention Polypeptides

These methods are provided herein to determine the optimal invention polypeptide for treatment of a disease or for personalized therapy.

5 A polynucleotide, *e.g.*, cDNA sequence, encoding an invention polypeptide is cloned into a plasmid expression vector, and placed under the control of a strong promoter such as CMV. The cDNA sequences are inserted in frame and cloned 5' of a short sequence encoding a convenient epitope tag. The encoded secreted factor possesses the epitope tag appended to its carboxy
10 terminus. Plasmids encoding tagged candidate factors are transfected into cell lines such as HEK293 or COS cells. Conditioned medium derived from transfected cells is harvested and epitope tagged proteins are immunoprecipitated by using anti-epitope tag antibodies. Immunoprecipitated proteins are detected following polyacrylamide gel electrophoresis and the molecular mass of the
15 candidate factor derived from the conditioned medium is determined and compared to the mass predicted on the basis of DNA sequence data.

The *in vitro* biological activity of the secreted polypeptide is then confirmed. Transfected cells producing secreted factors detected above are separated from target cells by a permeable membrane in a transwell tissue culture
20 plate and the effect of the secreted factor on the behavior of target cells is determined. Chemotactic factors such as chemokines will stimulate the migration of the target cells. Alternatively, conditioned media from plasmid-transfected cells can be added directly to target cells. Growth factors will stimulate the proliferation of target cells. Cytokines would be predicted to stimulate the
25 proliferation or activation of immune effector cells.

In Vitro Assays

Migration Assay

Cell lines transfected with plasmids encoding candidate secreted factors are placed into the lower chamber of a transwell chamber. *See, e.g.*, Fan & Malik
30 (2003) Nature Medicine 9:315. Target cells, potentially responsive to the effects of the secreted factor, are added to the upper chamber of the transwell plate

where they can be separated from the transfected cells by a permeable membrane. The target cells can be labeled (chemical, fluorescent, radioactive) to facilitate quantitation and detection. Suitable target cells include, but are not limited to monocytes, T cells, endothelial cells, dendritic cells, macrophages or other cell types potentially capable of responding to external soluble factors. The labeled cells are monitored as a function of time to determine the proportion that migrate through the permeable membrane and enter the lower chamber in response to chemotactic proteinaceous factors produced by the transfected cells. In instances where the concentration of the secreted factor within the conditioned medium is not high enough to detect a chemotactic activity, conditioned media can be collected, and concentrated prior to addition to the lower chamber using the method described in Cao *et al.* (2000) J. Immunology **165**:2588. To confirm the observed effect is due to the secreted factor, anti-epitope tag antibodies can be used to deplete conditioned medium of the secreted factor. Vigne *et al.* (2001) J. Biol. Chem. **276**:17101.

To confirm that the agent is to alter or modulate soluble factors capable of inducing transendothelial migration of T cells or other cells, the insert of a transwell plate can be coated with endothelial cells to produce a confluent monolayer. The cell-coated permeable insert would be placed between target cells (T cells or other) in the upper chamber and gene transfected cells or conditioned medium from gene transfected cells in the lower chamber. As above, the number of target cells that are able to traverse the endothelial cell-coated membrane is determined.

Proliferation

Target cells, potentially responsive to the effects of the secreted factor, are added to a well of a tissue culture plate. Conditioned medium from gene transfected cells are added to the well along with ^3H -thymidine. Target cells could include, but are not limited to monocytes, T cells, endothelial cells, dendritic cells, macrophages or other cell types potentially capable of responding to external soluble factors. Following several days, the level of ^3H -thymidine incorporated into target cells is determined by scintillation counting. To confirm

the observed effect is due to the candidate secreted factor, anti-epitope tag antibodies can be used to deplete conditioned medium of the secreted factor. The assay can also be conducted using transwell plates by adding target cells to the upper chamber and plasmid-transfected, irradiated cells to the lower chamber. Incorporation of ^3H -thymidine into live cells in the upper chamber can be determined while no radioactivity would be incorporated into irradiated cells in the lower chamber.

Activation

Target cells, potentially responsive to the effects of the secreted factor, are added to a well of a micro culture plate. Conditioned medium from gene transfected cells is then added to the well. Target cells include, but are not limited to monocytes, T cells, endothelial cells, dendritic cells, macrophages or other cell types potentially capable of responding to external soluble factors. Following several hours, labeled antibodies specific for activation markers displayed on the target cells or produced by the target cells, are added to the well and the intensity of the antibody signal could be quantitated. Soluble factors such as cytokines, capable of inducing the upregulation of activation markers, can be detected. To confirm the observed effect is due to the secreted factor, anti-epitope tag antibodies can be used to deplete conditioned medium of the secreted factor. The assay can also be conducted using transwell plates by adding target cells to the upper chamber and plasmid transfected cells to the lower chamber.

In Vivo Assays

High pressure tail vein plasmid injection could be conducted to force expression of the soluble factor in the liver of recipient rodents. Anti-epitope tag antibodies can be used to detect the presence of the secreted tagged protein in the blood of injected animals by immunoprecipitation followed by gel electrophoresis. At various time points following plasmid injection, blood samples are drawn and cellular and chemical analyses is conducted. Growth factors might be expected to cause the proliferation of specific circulating blood cell types that could be enumerated by conducting FACS using cell type specific antibodies. Activating factors might be expected to promote the display of

activation markers on specific cell types that could be detected by FACS using marker specific antibodies. Finally liver biopsies can be collected from from plasmid injected animals and cellular infiltrates are detected and identified by conducting histology and immunohistochemistry. Chemotactic factors are
5 expected to cause the recruitment and concentration of specific circulating cell types within the liver. Alternatively, subcutaneous or intradermal injection of plasmids encoding candidate chemotactic factors or of cells producing candidate soluble factors can be performed. Fan & Malik (2003) Nature Medicine 9:315. Skin biopsies or injection site lavages can be harvested from injected animals
10 and cellular infiltrates could be detected and identified by conducting histology and immunohistochemistry.

Therapeutic Methods

Enhancement of Antigens By Antigen-Presenting Matrices

The compositions of this invention can be used in conjunction with
15 antigen-presenting matrices. For use in immunomodulatory methods and diagnostic methods of the invention, an antigen-presenting matrix can presents an antigenic epitope bound to an MHC molecule. Any known method can be used to achieve presentation of an epitope by an antigen-presenting matrix. The following are non-limiting examples of methods which can be used.

20 An antigenic peptide can be delivered to antigen-presenting cells as polypeptide or peptide or in the form of cDNA encoding the protein/peptide.

Another method to deliver an antigenic epitope to an APC is by pulsing. Pulsing can be accomplished *in vitro/ex vivo* by exposing APCs to the antigenic polypeptide(s) or peptide(s) of this invention. The polypeptide(s) or peptide(s)
25 are added to APCs at a concentration of 1-10 μ m for approximately 3 hours. Pulsed APCs can subsequently be administered to the host via an intravenous, subcutaneous, intranasal, intramuscular or intraperitoneal route of delivery.

Antigenic epitopes can also be delivered *in vivo*, for example, as part of a polypeptide or complexed with another macromolecule, with or without adjuvant

via the intravenous, subcutaneous, intranasal, intramuscular or intraperitoneal route of delivery.

Various other techniques can be used, including the following. Paglia *et al.* (1996) J. Exp. Med. **183**:317-322 has shown that APC incubated with whole protein *in vitro* are recognized by MHC Class I-restricted CTLs, and that immunization of animals with these APCs led to the development of antigen-specific CTLs *in vivo*. In addition, several different techniques have been described which lead to the expression of antigen in the cytosol of APCs, such as DCs. These include (1) the introduction into the APCs of RNA isolated from tumor cells, (2) infection of APCs with recombinant vectors to induce endogenous expression of antigen, and (3) introduction of tumor antigen into the DC cytosol using liposomes. See Boczkowski *et al.* (1996) J. Exp. Med. **184**:465-472; Rouse *et al.* (1994) J. Virol. **68**:5685-5689; and Nair *et al.* (1992) J. Exp. Med. **175**:609-612).

Another method which can be used is termed "painting." It has been demonstrated that glycosyl-phosphatidylinositol (GPI)-modified proteins possess the ability to reincorporate themselves back into cell membranes after purification. Hirose *et al.* (1995) Methods Enzymol. **250**:582-614; Medof *et al.*, (1984) J. Exp. Med. **160**:1558-1578; Medof (1996) FASEB J. **10**:574-586; and Huang *et al.* (1994) Immunity **1**:607-613 have exploited this property in order to create APCs of specific composition for the presentation of antigen to CTLs. They devised expression vectors for β 2-microglobulin and the HLA-A2.1 allele. The proteins were expressed in Schneider S2 *Drosophila melanogaster* cells, known to support GPI-modification. After purification, the proteins could be incubated together with a purified antigenic peptide which resulted in a trimolecular complex capable of efficiently inserting itself into the membranes of autologous cells. In essence, these protein mixtures were used to "paint" the APC surface, conferring the ability to stimulate a CTL clone that was specific for the antigenic peptide. Cell coating was shown to occur rapidly and to be protein concentration dependent. This method of generating APCs bypasses the need for gene transfer into the APC and permits control of antigenic peptide densities at the cell surfaces.

Stimulation of Immune Effector Cells

The present invention makes use of the above-described antigen-presenting matrices, including APCs, to stimulate production of an enriched population of antigen-specific immune effector cells. The APCs prepared as described above are mixed with naïve immune effector cells. Preferably, the cells are cultured in the presence of a suitable polypeptide of this invention. The culture conditions are such that the antigen-specific immune effector cells expand (i.e. proliferate) at a much higher rate than the APCs. Multiple infusions of APCs and optional cytokines can be performed to further expand the population of antigen-specific cells.

In one embodiment, the immune effector cells are T cells. In a separate embodiment, the immune effector cells can be genetically modified by transduction with a transgene coding for a polypeptide of this invention. Methods for introducing transgenes *in vitro*, *ex vivo* and *in vivo* are known in the art. See Sambrook, *et al.* (1989) *supra*.

An effector cell population suitable for use in the methods of the present invention can be autogeneic or allogeneic, preferably autogeneic. When effector cells are allogeneic, preferably the cells are depleted of alloreactive cells before use. This can be accomplished by any known means, including, for example, by mixing the allogeneic effector cells and a recipient cell population and incubating them for a suitable time, then depleting CD69⁺ cells, or inactivating alloreactive cells, or inducing anergy in the alloreactive cell population.

Hybrid immune effector cells can also be used. Immune effector cell hybrids are known in the art and have been described in various publications. See, for example, International Patent Application Nos. WO 98/46785; and WO 95/16775.

The effector cell population can comprise unseparated cells, i.e., a mixed population, for example, a PBMC population, whole blood, and the like. The effector cell population can be manipulated by positive selection based on expression of cell surface markers, negative selection based on expression of cell surface markers, stimulation with one or more antigens *in vitro* or *in vivo*,

treatment with one or more polynucleotides and/or polypeptides of the invention, alone or in combination with biological modifiers *in vitro* or *in vivo*, subtractive stimulation with one or more antigens or biological modifiers, or a combination of any or all of these.

5 Effector cells can be obtained from a variety of sources, including but not limited to, PBMC, whole blood or fractions thereof containing mixed populations, spleen cells, bone marrow cells, tumor infiltrating lymphocytes, cells obtained by leukapheresis, biopsy tissue, lymph nodes, e.g., lymph nodes draining from a tumor. Suitable donors include an immunized donor, a non-
10 immunized (naïve) donor, treated or untreated donors. A “treated” donor is one that has been exposed to one or more biological modifiers. An “untreated” donor has not been exposed to one or more biological modifiers.

 Methods of extracting and culturing effector cells are known in the art. For example, effector cells can be obtained by leukapheresis, mechanical
15 apheresis using a continuous flow cell separator. For example, lymphocytes and monocytes can be isolated from the buffy coat by any known method, including, but not limited to, separation over Ficoll-Hypaque™ gradient, separation over a Percoll gradient, or elutriation. The concentration of Ficoll-Hypaque™ can be adjusted to obtain the desired population, for example, a population enriched in T
20 cells. Other methods based on affinity are known and can be used. These include, for example, fluorescence-activated cell sorting (FACS), cell adhesion, magnetic bead separation, and the like. Affinity-based methods may utilize antibodies, or portions thereof, which are specific for cell-surface markers and which are available from a variety of commercial sources, including, the
25 American Type Culture Collection (Manassas, MD). Affinity-based methods can alternatively utilize ligands or ligand analogs, of cell surface receptors.

 The effector cell population can be subjected to one or more separation protocols based on the expression of cell surface markers. For example, the cells can be subjected to positive selection on the basis of expression of one or more
30 cell surface polypeptides, including, but not limited to, “cluster of differentiation” cell surface markers such as CD2, CD3, CD4, CD8, TCR, CD45, CD45RO, CD45RA, CD11b, CD26, CD27, CD28, CD29, CD30, CD31, CD40L; other

markers associated with lymphocyte activation, such as the lymphocyte activation gene 3 product (LAG3), signaling lymphocyte activation molecule (SLAM), T1/ST2; chemokine receptors such as CCR3, CCR4, CXCR3, CCR5; homing receptors such as CD62L, CD44, CLA, CD146, a4b7, aEb7; activation
5 markers such as CD25, CD69 and OX40; and lipoglycans presented by CD1. The effector cell population can be subjected to negative selection for depletion of non-T cells and/or particular T cell subsets. Negative selection can be performed on the basis of cell surface expression of a variety of molecules, including, but not limited to, B cell markers such as CD19, and CD20; monocyte
10 marker CD14; the NK cell marker CD56.

Various methods are known to evaluate T cell activation. CTL activation can be detected by any known method, including but not limited to, tritiated thymidine incorporation (indicative of DNA synthesis), and examination of the population for growth or proliferation, e.g., by identification of colonies.
15 Alternatively, the tetrazolium salt MTT (3-(4,5-dimethyl-thazol-2-yl)-2,5-diphenyl tetrazolium bromide) may be added. Mossman (1983) J. Immunol. Methods **65**:55-63; Niks and Otto (1990) J. Immunol. Methods **130**:140-151. Succinate dehydrogenase, found in mitochondria of viable cells, converts the MTT to formazan blue. Thus, concentrated blue color would indicate
20 metabolically active cells. In yet another embodiment, incorporation of radiolabel, e.g., tritiated thymidine, may be assayed to indicate proliferation of cells. Similarly, protein synthesis may be shown by incorporation of ³⁵S-methionine. In still another embodiment, cytotoxicity and cell killing assays, such as the classical chromium release assay, may be employed to evaluate
25 epitope-specific CTL activation. To detect activation of CD4⁺ T cells, any of a variety of methods can be used, including, but not limited to, measuring cytokine production; and proliferation, for example, by tritiated thymidine incorporation

Release of ⁵¹Cr from labeled target cells is a standard assay which can be used to assess the number of peptide-specific CTLs in a biological sample.
30 Tumor cells, or APCs of the invention, are radiolabeled as targets with about 200 μ Ci of Na₂ ⁵¹CrO₄ for 60 minutes at 37° C, followed by washing. T cells and target cells ($\sim 1 \times 10^4$ /well) are then combined at various effector-to-target ratios

in 96-well, U-bottom plates. The plates are centrifuged at $100 \times g$ for 5 minutes to initiate cell contact, and are incubated for 4-16 hours at 37°C with 5% CO_2 . Release of ^{51}Cr is determined in the supernatant, and compared with targets incubated in the absence of T cells (negative control) or with 0.1% TRITONTM X-100 (positive control). *See, e.g.,* Mishell and Shiigi, eds. SELECTED METHODS IN CELLULAR IMMUNOLOGY (1980) W.H. Freeman and Co.

Polynucleotides of the invention can be administered in a gene delivery vehicle or by inserting into a host cell which in turn recombinantly transcribes, translates and processed the encoded polypeptide. Isolated host cells containing a polynucleotide of the invention in a pharmaceutically acceptable carrier can be combined with appropriate and effective amount of an adjuvant or other biological modifier. In some embodiments, the host cell is an APC, such as a dendritic cell. The agents provided herein as effective for their intended purpose can be administered to subjects having a disease to be treated with an immunomodulatory method of the invention or to individuals susceptible to or at risk of developing such a disease. When the agent is administered to a subject such as a mouse, a rat or a human patient, the agent can be added to a pharmaceutically acceptable carrier and systemically or topically administered to the subject. Therapeutic amounts can be empirically determined and will vary with the pathology or condition being treated, the subject being treated and the efficacy and toxicity of the therapy.

The amount of a peptide or immune effector cell of the invention will vary depending, in part, on its intended effect, and is ultimately at the discretion of the medical or veterinary practitioner. The factors to be considered include the condition being treated, the route of administration, and nature of the formulation, the mammal's body weight, surface area, age, and general condition and the particular peptide to be administered. A suitable effective dose of peptides of the invention generally lies in the range of from about 0.0001 $\mu\text{mol/kg}$ to about 1000 $\mu\text{mol/kg}$ bodyweight. The total dose may be given as a single dose or multiple doses, e.g., two to six times per day. For example, for a 75 kg mammal (e.g., a human) the dose range would be about 2.25 $\mu\text{mol/kg/day}$ and a typical dose could be about 100 μmol of peptide. If discrete multiple doses

are indicated treatment might typically be 25 μmol of a peptide of the invention given up to 4 times per day. In an alternative administrative regimen, peptides of the invention may be given on alternate days or even once or twice a week. A suitable effective dose of an immune effector cell of the invention generally lies in the range of from about 10^2 to about 10^9 cells per administration. Cells can be administered once, followed by monitoring of the clinical response, such as diminution of disease symptoms or tumor mass. Administration may be repeated on a monthly basis, for example, or as appropriate. Those skilled in the art will appreciate that an appropriate administrative regimen would be at the discretion of the physician or veterinary practitioner.

Administration *in vivo* can be effected in one dose, continuously or intermittently throughout the course of treatment. Methods of determining the most effective means and dosage of administration are known to those of skill in the art and will vary with the composition used for therapy, the purpose of the therapy, the target cell being treated, and the subject being treated. Single or multiple administrations can be carried out with the dose level and pattern being selected by the treating physician. Suitable dosage formulations and methods of administering the agents can be found below.

The agents and compositions of the present invention can be used in the manufacture of medicaments and for the treatment of humans and other animals by administration in accordance with conventional procedures, such as an active ingredient in pharmaceutical compositions.

Enhancement of Vaccines for Cancer Treatment and Prevention

The compositions of this invention can be used in conjunction with vaccines for cancer treatment. Cancer cells contain many new antigens potentially recognizable by the immune system. Given the speed with which epitopes can be identified, custom anticancer vaccines can be generated for affected individuals by isolating TILs from patients with solid tumors, determining their MHC restriction, and assaying these CTLs against the appropriate library for reactive epitopes. These vaccines will be both treatments for affected individuals as well as preventive therapy against recurrence (or

establishment of the disease in patients which present with a familial genetic predisposition to it). Inoculation of individuals who have never had the cancer is expected to be quite successful as preventive therapy, even though a tumor antigen-specific CTL response has not yet been elicited, because in most cases
5 high affinity peptides seem to be immunogenic suggesting that holes in the functional T cell repertoire, if they exist, may be relatively rare. Sette *et al.* (1994) *J. Immunol.*, **153**:5586-5592. In mice, vaccination with appropriate epitopes not only eliminates established tumors but also protects against tumor re-establishment after inoculation with otherwise lethal doses of tumor cells.
10 Bystryn *et al.* (1993) *supra*.

Recent advances in vaccine adjuvants provide effective means of administering peptides so that they impact maximally on the immune system. Del-Giudice (1994) *Experientia* **50**:1061-1066. These peptide vaccines will be of great value in treating metastatic tumors that are generally unresponsive to
15 conventional therapies. Tumors arising from the homozygous deletion of recessive oncogenes are less susceptible to elimination by a humoral (antibody) response and would thus be treated more effectively by eliciting a cellular, CTL response.

Enhancement of Vaccines for Diseases Caused by Pathogenic Organisms

20 The compositions of this invention are also useful in methods to induce (or increase, or enhance) an immune response to a pathogenic organism. These include pathogenic viruses, bacteria, and protozoans.

Viral infections are ideal candidates for immunotherapy. Immunological responses to viral pathogens are sometimes ineffective as in the case of the
25 lentiviruses such as HIV which causes AIDS. The high rates of spontaneous mutation make these viruses elusive to the immune system. However, a saturating profile of CTL epitopes presented on infected cells will identify shared antigens among different serotypes in essential genes that are largely intolerant to mutation which would allow the design of more effective vaccines.

Adoptive Immunotherapy Methods

The expanded populations of antigen-specific immune effector cells and APCs of the present invention find use in adoptive immunotherapy regimes and as vaccines.

5 Adoptive immunotherapy methods involve, in one aspect, administering to a subject a substantially pure population of educated, antigen-specific immune effector cells made by culturing naïve immune effector cells with APCs as described above. In some embodiments, the APCs are dendritic cells.

10 In one embodiment, the adoptive immunotherapy methods described herein are autologous. In this case, the APCs are made using parental cells isolated from a single subject. The expanded population also employs T cells isolated from that subject. Finally, the expanded population of antigen-specific cells is administered to the same patient.

15 In a further embodiment, APCs or immune effector cells are administered with an effective amount of a stimulatory cytokine, such as IL-2 or a co-stimulatory molecule.

Methods of inducing T cell anergy

20 Antigenic peptides are useful in methods to induce T cell unresponsiveness, or anergy. Disorders which can be treated using these methods include autoimmune disorders, allergies, and allograft rejection.

25 Autoimmune disorders are diseases in which the body's immune system responds against self tissues. They include most forms of arthritis, ulcerative colitis, and multiple sclerosis. Antigenic peptides corresponding to endogenous elements that are recognized as foreign can be used in the development of treatments using gene therapy or other approaches. For example, CTL epitopes, which can act as "suicide substrates" for CTLs that mediate autoimmunity, can be designed as described above. That is to say, peptides which have a high affinity for the MHC allele but fail to activate the TCR could effectively mask the cellular immune response against cells presenting the antigen in question. In support of this approach, it is believed that the long latency period of the HIV

30

virus is due to an antiviral immune response and a mechanism by which the virus finally evades the immune system is by generating epitopes that occupy the MHC molecules but do not stimulate a TCR lytic response, inducing specific T cell anergy. Klenerman et al. (1995) Eur. J. Immunol. **25**:1927-1931.

5 *In vitro* stimulation of T cells through the complex of T cell-antigen receptor and CD3 alone in the absence of other signals, induces T cell anergy or paralysis. T cell activation as measured by interleukin-2 production and proliferation *in vitro* requires both antigenic and co-stimulatory signals engendered by cell to cell interactions among antigen-specific T cells and antigen
10 presenting cells. Various interactions of these CD2 proteins on the T-cell surface with CD58 (LFA-3) proteins and antigen-presenting cells, those of CD11a/CD18 (LFA-1) proteins with CD54 (ICAM-1) proteins and those of CD5 proteins with CD72 proteins can impart such a co-stimulatory signal *in vitro*. Cytokines derived from antigen-presenting cells (e.g., interleukin-1 and interleukin-6) can
15 also provide co-stimulatory signals that result in T-cell activation *in vitro*. The delivery of both antigenic and co-stimulatory signals leads to stable transcription of the interleukin-2 gene and other pivotal T cell-activation genes. The foregoing co-stimulatory signals depend on protein kinase C and calcium. Potent antigen presenting cells express CD80 (B7 and BB1) and other related surface proteins
20 and many T cells express B7 binding proteins, namely CD28 and CTLA-4 proteins. Binding of CD80 by CD28 and CDLA-4 stimulates a T cell co-stimulatory pathway that is independent of protein kinase C and calcium leading to vigorous T cell proliferation. The stimulation of B cells also depends on the interaction between the specific antigen and the cell-surface immunoglobulin.
25 T cell derived cytokines (e.g., interleukins-1 and -4), physical contact between T cells and B cells through specific pairs of receptors and co-receptors, or both, provide the signal or signals essential for B cell stimulation.

Conventional routes of administration are used. A T-cell stimulating or anergy producing amount (or therapeutically effective amount as described
30 above) of an immunotherapeutic antigen-superantigen polymer according to the invention is contacted with the target cells. By "T-cell anergy effective amount" is intended an amount which is effective in producing a statistically significant

inhibition of a cellular activity mediated by a TCR. This may be assessed *in vitro* using T-cell activation tests. Typically, T-cell anergy or activation is assayed by tritiated thymidine incorporation in response to specific antigen.

5 One way in which T cell anergy can be induced is to present to a T cell an antigen-presenting matrix which presents an antigenic peptide of the invention in an MHC Class I or Class II molecule, but which lack co-stimulatory molecules necessary to activate the T cell. For example, a cell other than a normal antigen presenting cell (APC), which has been transfected with MHC antigen to which a selected T cell clone is restricted, can be used. Resting T cells are provided with
10 an appropriate peptide recognized by the resting T cells in the context of the MHC transfected into a cellular host other than an APC. The MHC is expressed as a result of introduction into a mammalian cell other than an antigen presenting cell of genes constitutively expressing the α and β chains of the MHC Class II, or an MHC Class I molecule together with invariant chain. Importantly, these cells
15 do not provide other proteins, either cell surface proteins or secreted proteins, associated with antigen presenting cells, which together with the MHC and peptide result in co-stimulatory signals.

To determine whether anergy has been induced, the T cells to be tested can be cultured together with an antigen presenting matrix which presents an
20 antigenic peptide in an MHC Class I or Class II molecule together with co-stimulatory molecules necessary to activate the T cell. The cultures are incubated for about 48 hours, then pulsed with tritiated thymidine and incorporation measured about 18 hours later. The absence of incorporation above control levels, where the T-cells are presented with antigen presenting cells which do not
25 stimulate the T cells, either due to using an MHC to which the T cells are not restricted or using a peptide to which the T cells are not sensitive, is indicative of an absence of activation. One may use other conventional assays to determine the extent of activation, such as assaying for IL-2, -3, or -4, cell surface proteins associated with activation, e.g., CD71 or other convenient techniques. Another
30 method is to determine the expression of a protein which is expressed on quiescent T cells, but not on anergic T cells. U.S. Patent No. 5,747,299.

Enhancement of Adoptive Immunotherapy and Vaccines

The compositions of this invention are useful in conjunction with expanded populations of antigen-specific immune effector cells which are further useful in adoptive immunotherapy regimes and as vaccines.

5 Adoptive immunotherapy methods involve, in one aspect, administering to a subject a substantially pure population of educated, antigen-specific immune effector cells made by culturing naïve immune effector cells with APCs as described above. Preferably, the APCs are dendritic cells.

10 In one embodiment, the adoptive immunotherapy methods described herein are autologous. In this case, the APCs are made using parental cells isolated from a single subject. The expanded population also employs T cells isolated from that subject. Finally, the expanded population of antigen-specific cells is administered to the same patient.

15 It is to be understood that while the invention has been described in conjunction with the above embodiments, that the foregoing description and examples are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

CLAIMS

1. An isolated polynucleotide encoding a peptide selected from the group identified in Table 1 and the complement of said polynucleotide.
2. The isolated polynucleotide of claim 1 and a carrier.
- 5 3. The isolated polynucleotide of claim 1 and a pharmaceutically acceptable carrier.
4. A host cell comprising a polynucleotide of claim 1.
5. The host cell of claim 4, wherein the host cell is an antigen presenting cell and the peptide is presented in the context of an MHC molecule.
- 10 6. The host cell of claim 5, wherein the antigen presenting cell is a dendritic cell.
7. A composition comprising the host cell of any one of claims 4 to 6, and a carrier.
8. A composition comprising the host cell of claim 4 and a pharmaceutically acceptable carrier.
- 15 9. An isolated peptide selected from the group identified in Table 1.
10. A host cell comprising the peptide of claim 9.
11. A composition comprising the peptide of claim 9 and a carrier.
12. A composition comprising the host cell of claim 10 and a carrier.
- 20 13. An antibody that specifically recognizes and binds the peptide of claim 9.
14. A composition comprising the antibody of claim 13 and a carrier.

15. An immune effector cell raised in the presence and at the expense of a host cell of claim 5.

16. A method for eliciting a cytolytic response in a subject comprising administering to said subject an effective amount of a composition of any one of
5 claims 3, 8, 11, 12 or 14.

SEQUENCE LISTING

<110> ROBERTS, BRUCE

<120> BLOOD FACTOR DOMAINS

<130> GZ 2123.40

<140>

<141>

<160> 64

<170> PatentIn Ver. 2.1

<210> 1

<211> 10

<212> DNA

<213> Homo sapiens

<400> 1

gtggtatggc

10

<210> 2

<211> 206

<212> PRT

<213> Homo sapiens

<400> 2

Met Arg Trp Ala Arg Leu Ser Ala Leu Leu Cys Gln Arg Arg Val Val
1 5 10 15Trp Leu Ser Tyr Gly Val Ala Ser Ser Tyr Val Leu Ala Asp Ala Ile
20 25 30Asp Lys Gly Lys Lys Ala Gly Glu Val Pro Ser Pro Glu Ala Gly Arg
35 40 45Ser Ala Arg Val Thr Val Ala Val Val Asp Thr Phe Val Trp Gln Ala
50 55 60Leu Ala Ser Val Ala Ile Pro Gly Phe Thr Ile Asn Arg Val Cys Ala
65 70 75 80Ala Ser Leu Tyr Val Leu Gly Thr Ala Thr Arg Trp Pro Leu Ala Val
85 90 95Arg Lys Trp Thr Thr Thr Ala Leu Gly Leu Leu Thr Ile Pro Ile Ile
100 105 110Ile His Pro Ile Asp Arg Ser Val Asp Phe Leu Leu Asp Ser Ser Leu
115 120 125Arg Lys Ser Thr Gln Gln Trp Gly Ser Pro Ala Pro Pro Asp His Thr
130 135 140Leu Val Pro Gly Leu Cys Ile Gly Leu Leu Leu His Val Asn Leu Leu
145 150 155 160Leu Leu Pro Gly Asn Val Asp Thr Trp Leu Pro Gly Val Gln Arg Pro
165 170 175

Trp His Leu Gly Gly Phe Glu Leu Asp Arg Ser Leu Glu Thr Lys Ala

180 185 190
 Ser Arg Ser Ser Gly Cys Arg Glu Ser Gln Lys Gly Arg Thr
 195 200 205

<210> 3
 <211> 1294
 <212> DNA
 <213> Homo sapiens

<400> 3
 acgtgacgaa aaccagtctg taaaacccga gcctggggaga ggggcttcgg tgcgcggggg 60
 gaatttgacg acgctccctg ctggcgaggaga ttctctgacc tgtccttcgg cgcgggactt 120
 tcggcggggtc ccggccggca gacccaagtg ccggcgggcg atgactgcag tggagccagt 180
 accggctgta gtggccgggg cccgtggcgg gagagtcatt tcagagccgc agccgcgggg 240
 cgaggagcgc agatcgtcta ccgggacacg tgggtgcatg acctgggcta tgcaatgagg 300
 tgggcgaggc ttccgctctt cttgtgccag cggcggggtg tgtggctgag ctatggcgtg 360
 gccagctcct acgtgctggc ggatgccatt gacaaaggca agaaggctgg agaggtgccc 420
 agccctgaag caggccgcag cgccagggtg actgtggctg tgggtggacac ctttgtatgg 480
 caggctctag cctctgtggc cattccgggc ttaccatca accgcgtgtg tgctgcctct 540
 ctctatgtcc tgggcaactg caccgcgtgg cccctggctg tccgcaagtg gaccaccacc 600
 gcgcttgggc tggtagcatc ccccatcatt atccacccca ttgacaggctc ggtggatttc 660
 ctctggactc ccagcctgcg caagtctacc caacagtggg gaagcccagc tcctcctgat 720
 catactctgg tacctggcct gtgcatcggc ctctgcttc atgtcaacct cctactcctg 780
 ccagggaatg tggacacctg gctccctggg gtccaaagac cctggcacct ggggtgggtt 840
 gagctggaca gaagcttaga gacaaaggct tcaagaagca gtggctgcag ggagtcacag 900
 aagggcagga cctgaacgct gtctgcttcc ctggaatcca agatgctgag tggagtgga 960
 ccctgggtgg gcccggccct gtctttttca ggaaaattac atcctcccat ggaggatgag 1020
 agactgaggc tcaggagagg caagggaatg gcccagatc acttggcaag ctgggcaccc 1080
 aggaccccca ggtgcttgac agagtcaccc catggtggta tggctgaaca aggagcggca 1140
 gacaactcag ggagaaactc aggagtgcag taccagggac acctcaggac agattctctg 1200
 gccaggccct tccctgacct aataaatcct gaagaggtaa aaaaaaaaaa aaaaaaaaaa 1260
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1294

<210> 4
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 4
 ctaataaact 10

<210> 5
 <211> 283
 <212> PRT
 <213> Homo sapiens

<400> 5
 Met Thr Ser Gly Pro Gly Gly Pro Ala Ala Ala Gly Gly Arg Lys
 1 5 10 15
 Glu Asn His Gln Trp Tyr Val Cys Asn Arg Glu Lys Leu Cys Glu Ser
 20 25 30
 Leu Gln Ala Val Phe Val Gln Ser Tyr Leu Asp Gln Gly Thr Gln Ile
 35 40 45
 Phe Leu Asn Asn Ser Ile Glu Lys Ser Gly Trp Leu Phe Ile Gln Leu
 50 55 60
 Tyr His Ser Phe Val Ser Ser Val Phe Ser Leu Phe Met Ser Arg Thr
 65 70 75 80

```
<210> 6
<211> 3163
<212> DNA
<213> Homo sapiens
```

<400> 6							
cccttggccct	gaaggggggct	ggatgggcaa	ggcgggccg	atggctcgag	ctcggg	cgggt	60
ggcggcggtg	gccggaggcg	gcggtgcctc	ctcctcctcg	ccccggcgcc	ggcgggtgatc		120
cgagcgagcg	gccgcggccc	ccgatgagac	tgctggcg	ctggctgtgc	ctgagcctgg		180
cgtccgtgtg	gctggcgcg	aggatgtgga	cgttgcggag	cccgctcacc	cgctccctgt		240
acgtgaacat	gactagcggc	ccgggtgggc	cggcgggcggc	cgcggggcggc	aggaaggaga		300
accaccagt	gtatgtgtgc	aacagagaga	aattatgcga	atcactccag	gctgtctttg		360
ttcagagtta	ccttgatcaa	ggaacacaga	tcttcttaaa	caacagcatt	gagaaatcgg		420
gctggctatt	tatccaatta	tatcattctt	ttgtgtcatc	tgtttttagc	ctgtttatgt		480
ctagaacatc	tatcaatggg	ttgtctaggaa	gaggctcaat	gtttgtgttt	tcaccagatc		540
agtttcagag	actgcttaaa	attaatccag	actggaaaac	ccacagactt	cttgatttag		600
gtgctggaga	tggagaagtc	acaaaaatca	tgagccctca	ttttgaagaa	atctatgcc		660
ctgagctttc	tgaaaactatg	atatggcagc	ttcagaaaaa	gaaatacaga	gtccttggt		720
taaatgaatg	gcagaatacg	gggttcag	atgatgtcat	cagctgcctg	aacttgctgg		780
accgctgtga	tcagcccctg	actttgttaa	aagatatcag	aagtgtcttg	gagccaacta		840
gaggcaggg	catccttgcc	cttgtcctcc	cctttcatcc	ctatgtggaa	aacgtaggtg		900
gcaagtggga	gaaaccatga	gaaatttttg	aaatcaaagg	acagacactgg	gaagaacaag		960
tgaatagtct	gcctgaagtt	ttcagaaaag	ctggttttgt	tatcgaagct	ttcaccagac		1020
taccatacct	gtgtgaaggc	gacatgtata	atgactacta	cgttctggat	gacgctgtct		1080

```

ttgtttctcaa accagtataa acacgtggag gtcgaagtct tcagagtccg caccctccgg 1140
gatgtgccct tggaagaggg tctgtgttca caattacgtg aaggaggagac ccttggggac 1200
cgccatttcta aatatcatgt aggaatttaa aaagccaaaa tactaattat ttctttgtag 1260
tgtgtaaagg aatgttttta aaagacaaaa acccaactct ttgtggattt ttatcaactc 1320
tttactcaga gccactctcc aatgcaggtc aactccaat tatgatggaa gatatttttt 1380
atacttaatt gcagtaggga ctcatccca gacaaagcaa tagtcacgac ttcatggaac 1440
caatcaatgg attgtttttt gaagactggc aataaagctg tccattcaat tccaaatact 1500
ggttttaagg ttagccact gatattcttt catgtttaga aattctttct gttattattc 1560
aagaaaaatgt ttttaatcat gctaataaac ttttttgagg atgactttgg catcatgttt 1620
gaattcatat aaagctcccc tagcattttt tattggtttg gcttcaggag tacccaaata 1680
gtagcattat gagaatgacg cagacaattt gaataggggg gaagggaaggc ttcagacttg 1740
ggggaagggg agattattgc aaattgcagt gaacactgag tcagtaaaaa aaaaaaacag 1800
aaacaaaaac ccagcctcat tcagatacag aacttcaggg accctcccc ccaccccccc 1860
agttaatgct gctgtgaaaa atgcaaaaaa acctgggtct gactttgtga tcactcatgt 1920
cccatacact gaactttgtt tttttctgga caatctcagg ttctcagaat tgaaacattc 1980
agttttgtct actgacaaaa tgcaactaaa aatgttttaa ttcaacttct tactctacac 2040
ttatatacct ctctcccaga gcttggtctt tccctgaaat ccttaaagga gtttattatt 2100
tgccaagaat atggtttgat gaagaccgtg tcagactcat gcattctgga atccgtgagt 2160
gcctccaacc ataaaccag aaatgctgcc atagaggaa ccattagcaa gatcaataac 2220
catcgtggat attccaaaga gcagagcatt taaccagaa tctgaaatct cagtaataac 2280
agaaatgtca atatgagatt gtggttatta acaatatgtt atcctcacta ctttatttct 2340
tttttcatcc aggtgtttca ttaagtattc tcatttatat gctcatagaa gtcatcagat 2400
caacacattc actccaacac ctgggtataa aggcgaatta aaacctgaag aattacttta 2460
ttattcatta agaactgcat tcccacatcc tcacaccaat atcttccact gttactgtgc 2520
agtcataata agagttgggt tatgtgattt atgcaaaatt gttattggtg tgttctaata 2580
atttttttcc aggaattcaa gtaagtgaag tacaataact tagcaccata gattgtaaaa 2640
atagaaagag gaaatgagaa aacttaatgt cgttcttgta cagttgagtt ttgcttcttt 2700
ataaactgtc tgttaaaatc aggggttgct gttagactgt caagcagaag tgatggatgt 2760
tcaaggaggc aggttgcccc aaagtgtttt ctaactttca agtagagaac tttatgctct 2820
aataagaaaa atacaactca ttcttaacca ccccccaaa acagagaaaa cacatacatt 2880
tatgcatttt aaagataaaa agtagtttct caaaaaaatc tgaagttctc tgaaaaccag 2940
cttattaaaa gtgctgtttt gtggtgaaaa ttgaaaccaa atccactcac atttcctatc 3000
gggtcttgaa attcttggtg gtgtgtcacc attccactag atggcagtggt tgcttactga 3060
gtgtcaatgg cttttttccc cgtaaattgga atcatgtttt tctcccccaa agtacaataa 3120
agctgccttg tctgcaccaa aaacaaaaaa aaaaaaaaaa aaa 3163

```

<210> 7
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 7
 gctccctacg

10

<210> 8
 <211> 229
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Thr Pro Gln Leu Leu Leu Ala Leu Val Leu Trp Ala Ser Cys Pro
 1 5 10 15
 Pro Cys Ser Gly Arg Lys Gly Pro Pro Ala Ala Leu Thr Leu Pro Arg
 20 25 30
 Val Gln Cys Arg Ala Ser Arg Tyr Pro Ile Ala Val Asp Cys Ser Trp
 35 40 45
 Thr Leu Pro Pro Ala Pro Asn Ser Thr Ser Pro Val Ser Phe Ile Ala
 50 55 60
 Thr Tyr Arg Leu Gly Met Ala Ala Arg Gly His Ser Trp Pro Cys Leu

```
<210> 9
<211> 1161
<212> DNA
<213> Homo sapiens
```

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

<400> 10
tgtattttga

10

<210> 11
<211> 227
<212> PRT
<213> Homo sapiens

<400> 11
Met Arg Val Ala Gly Ala Ala Lys Leu Val Val Ala Val Ala Val Phe
1 5 10 15
Leu Leu Thr Phe Tyr Val Ile Ser Gln Val Phe Glu Ile Lys Met Asp
20 25 30
Ala Ser Leu Gly Asn Leu Phe Ala Arg Ser Ala Leu Asp Thr Ala Ala
35 40 45
Arg Ser Thr Lys Pro Pro Arg Tyr Lys Cys Gly Ile Ser Lys Ala Cys
50 55 60
Pro Glu Lys His Phe Ala Phe Lys Met Ala Ser Gly Ala Ala Asn Val
65 70 75 80
Val Gly Pro Lys Ile Cys Leu Glu Asp Asn Val Leu Met Ser Gly Val
85 90 95
Lys Asn Asn Val Gly Arg Gly Ile Asn Val Ala Leu Ala Asn Gly Lys
100 105 110
Thr Gly Glu Val Leu Asp Thr Lys Tyr Phe Asp Met Trp Gly Gly Asp
115 120 125
Val Ala Pro Phe Ile Glu Phe Leu Lys Ala Ile Gln Asp Gly Thr Ile
130 135 140
Val Leu Met Gly Thr Tyr Asp Asp Gly Ala Thr Lys Leu Asn Asp Glu
145 150 155 160
Ala Arg Arg Leu Ile Ala Asp Leu Gly Ser Thr Ser Ile Thr Asn Leu
165 170 175
Gly Phe Arg Asp Asn Trp Val Phe Cys Gly Gly Lys Gly Ile Lys Thr
180 185 190
Lys Ser Pro Phe Glu Gln His Ile Lys Asn Asn Lys Asp Thr Asn Lys
195 200 205
Tyr Glu Gly Trp Pro Glu Val Val Glu Met Glu Gly Cys Ile Pro Gln
210 215 220
Lys Gln Asp
225

<210> 12
<211> 2475
<212> DNA
<213> Homo sapiens

<400> 12
ccgagaggagt ccgagaggaa gcggaggcgc gagctggagg cggcggtcc cgtcggcctc 60
cggcaggact gagcgctggg aggccggaag gcgggcgcgc acggcggaga ggcgggcggg 120

```

aggccggagc atattaatga aaagtgccat aaactgaaaa accaaacatg agggtagcag 180
gtgctgcaaa gttggtggta gctgtggcag tgtttttact gacattttat gttattttctc 240
aagtatttga aataaaaaatg gatgcaagtt taggaaatct atttgcaaga tcagcatttg 300
acacagctgc acgttctaca aagcctccca gatataagtg tgggatctca aaagcttgcc 360
ctgagaagca ttttgctttt aaaatggcaa gtggagcagc caacgtggtg ggacccaaaa 420
tctgcctgga agataatgtt ttaatgagtg gtgttaagaa taatgttgga agagggatca 480
atgttgccct ggcaaatgga aaaacaggag aagtattaga cactaaatat tttgacatgt 540
ggggaggaga tgtggcacca tttattgagt ttctgaaggc catacaagat ggaacaatag 600
ttttaatggg aacatacgat gatggagcaa ccaaactcaa tgatgaggca ggcgggtca 660
ttgctgattt ggggagcaca tctattacta atcttggttt tagagacaac tgggtcttct 720
gtggtgggaa gggcattaag acaaaaagcc cttttgaaca gcacataaag aacaataagg 780
atacaaacaa atatgaagga tggcctgaag ttgtagaaat ggaaggatgc atccccaga 840
agcaagacta atggaaatgt ggagagaatt gaagaaagcg cactttcact cttaatggga 900
gagctataaa tggcagagct atgtgtaaat attttaagag catgcagcca tcttggtgtg 960
tgcattgagta ttgtctcttt tgatatcagg attattttatt gctaacgtaa atagatagca 1020
ttgtaaataa tcatcacaat gatcaaatca ctgaaccatg tctccgcaca ttccctaaa 1080
agtacaatgt ttagactgct atggtaatac atatttttaa ttctaaaagc atacacaatg 1140
tgtaactgaa tggtttgtga aaaatatatt gatatatata ctagtgtgta tgaaaatatc 1200
atggaataat agggatttta ggggtggata tttatgtttc tttatgtttc tataatgttg 1260
gttgtagtga cattatcttt taaattaaaa agagatttgg ctagtgtgtg gtgtaatgtt 1320
actttacagt ccgactctcc tgatgtacct cttttcatga tctttttctt tccttcccaa 1380
gaaactgagg aatgtttaat atgaaaacat acatcggata tgtgaaaagc acaacaaaat 1440
tcttaatgta cacagtaaaa aagtaaatat ataaatgtag atggcattta ggaccacagc 1500
ttgctggatt tgtgttagct atgggaataa cttgatittg tataagctat ttagagttag 1560
gtggagggtg gcagcttcac agaactggag aaccaggcca agtccccctc ccaacctaata 1620
taggtcattc aggcagccta agtcagtata ttttagagcaa tactagcata cgtttttctt 1680
aattgttatc agcattgacc aagtggtttg gaaggaggca tgctttaata tcacaataat 1740
tttgatttgt aaaccaagaa attaatcctg tgtttatcta acttcataat agcaattatt 1800
gcccgaagct atagtggcat atttcaaaaa gttcttatta ctgggcggac tgataacatt 1860
taaaaaataa ttgtgtttga ccccaaatga ctttataccc aattctacat aaaaatatag 1920
aagatctatc tttttttgtt accttcagat gttcactaaa taactcagtt ttttagcaga 1980
agttttcagg gcattaaata tatgttgtgt atgaagtatc tcaaactgga acataaattt 2040
agtgatcaaa ctgccattca cagtgttaagg cagcacttaa atttcgaacc taaagttag 2100
atgcattgta taaaaaaacc taaaagcagt atctgttatt tagctgtaaa ccaagttaga 2160
agctattcgg ataatttctt aaatattgat gaactttgga gtactgtttc ttccttcaaa 2220
ctgaatgtaa ttaattcatg aataaatgca ctttataatg ttaaacaatc tttgtatact 2280
tttgggattt ttggtgctta tatgctaaat cacattcagc atgtgtattt tgacatttaa 2340
aatacttccc tcaattctgt aaattaaaag aatagttatt ttacagttcc agggattgtg 2400
aaataaatgt tgcagttttt taaaataatg aaaataaata ctcttggttt tgctttgtga 2460
aaaaaaaaaa aaaaa 2475

```

<210> 13
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 13
 gtccccccac

10

<210> 14
 <211> 85
 <212> PRT
 <213> Homo sapiens

<400> 14
 Met Ile Tyr Val Ser Thr Gly Ala Trp Gly Trp Ala Gly Ser Cys Phe
 1 5 10 15
 Trp Gln Trp Ser Phe Cys Pro Pro Ala Cys Val Gly Cys Ile Glu Glu
 20 25 30
 Glu Ala Gln Gly Ser Ser Glu Gln Glu Leu His Tyr Ala Ser Leu Gln
 35 40 45

Arg Leu Pro Val Pro Ser Ser Glu Gly Pro Asp Leu Arg Gly Arg Asp
 50 55 60

Lys Arg Gly Thr Lys Glu Asp Pro Arg Ala Asp Tyr Ala Cys Ile Ala
 65 70 75 80

Glu Asn Lys Pro Thr
 85

<210> 15
 <211> 6853
 <212> DNA
 <213> Homo sapiens

<400> 15

aagcttctga	agggtttaga	tgtaaaatgt	actgatattt	gtaattttta	aagatatttta	60
aaagtgagat	ggattgataa	atatgtgata	aagcaaatata	aaaatgttaa	tagagccagg	120
tgcatgtggc	cactcctgta	attccagcac	tttggaaggc	taagggtggaa	agattgcttg	180
agaccaggag	ttcaaaatca	gcctgggcaa	catagtaaga	ccccatttct	acaaagcctc	240
atgtggtagc	tggtgtctgt	agtcctagct	actcagaagg	ctaagggtggg	aggacctctg	300
agcccaggaa	ttcaattctg	cagttagcta	tgatttcacc	actgcacttc	aggctgagtg	360
acagagttag	accccatctc	aaaaacaaaa	caaaacaaaa	aatgttaata	gtagcatcta	420
ggtggtaaga	atatgttcac	tgtacaatta	ttctcatttc	accctatggt	tgcaactttt	480
aataataaaa	tgtaaaaaaa	acaaaacaaa	caaacaaaaa	accctgaata	ttattcagca	540
tggggaacat	ggaggatggg	gagaaggggtg	ggggaggaag	tagaagggttc	ttgaatttgg	600
aaggggaaac	gcaaattaat	atggaccacac	ccaggcacca	catctcctcc	tcaccccttg	660
ccttacaggc	gctccccagt	cttcaccctc	ctcaaggagt	gggtgtgcaa	tcctccagca	720
cccattctct	tctccatcac	agtgccacta	agaagccttc	acccagggtct	ctccagagag	780
cctcaggccg	ctgcccttac	ttagtctgtg	gttcaatgcc	agaatgctgc	ctcctacagg	840
aagtccacct	gtattgcccc	cacctccttt	cctgtcacca	acttgtcacc	aactttctgt	900
ccttgatcta	tccacagggc	tcattgttag	ctagtatggc	tgcccttaac	tctcatgttt	960
gttaatcaga	cagccaagca	gcctgctgca	tagagctgca	gaacaccaa	tatgccagag	1020
ctcccagctc	gaactggagc	agggtagatg	tgtccacaga	catatgccaa	gatcaagagg	1080
tctcaacaga	tgcatgttaa	gaggtaattag	agaagagtta	atcaagggaag	acacctgaag	1140
gtggtgggtg	tttgctgact	agtggcagga	tcagtgaat	gactggagct	gaggcagatt	1200
atggccctag	ctacaggccc	agaagtttga	aaagaaagat	gttgtaaccc	taaccctgga	1260
gccgaacttc	ctctcctaac	aatgctgggg	aggaacccag	gctggggggag	aagttaaagc	1320
cagaggaggg	gcaggaatgt	ctgagggtggc	aacacttctc	ttcagccaga	cagcactggc	1380
cagtttgtag	tctgtccatc	ctgcaggcca	caagctctgg	gtaagctggg	aatgggcagg	1440
gaccttggtg	gaaggatggt	cacaccccac	agtgggggtga	agctaagatg	aggggagggg	1500
gagtatgggt	ttgagtttcc	ctgggccgtc	gaggaatcct	ctgagtctct	gctccccaaa	1560
gaaattaaag	acaattcatt	tctgtgccca	cggcccttat	ggctccacct	gcacttctgc	1620
tccccacccc	ccagaattcc	tcttaaacc	agaaggtccc	agtttccaga	ccctagtcag	1680
tatatctggc	tctgggggtga	agagaacggc	cccctcttca	ccctcaaaca	ggaaccagtg	1740
gttggagggg	tgaggagtgcc	ttatggggcc	ccagatactc	ctccatgccc	ccaagggggc	1800
cacttcagcc	ctagcagcat	ctgcctgtgg	gaagcagctc	tccacaccag	gtgtgtctgat	1860
ccccacactc	ccgcgctgct	ctgcggctca	gggagcagcc	cacctgctgg	ccttccccctc	1920
atcacccctc	cttcttcccc	ccagtgccca	caccacacca	ggcccagggt	cttccccctc	1980
catcatcccc	ttaccagcac	ctagaacctg	ccagggtgta	aaagtcacct	ccaaaccagc	2040
tggctagccc	agggcagagg	aaagggctgg	gctctggagt	tgggcagagc	tgcccttaaa	2100
ccccagctcc	acctttctgg	gatgggtgac	ctagttaaagt	ccaggcttga	atctcggttc	2160
tttacttggt	caacgggcac	catgataccc	tatgttctgg	ggattagcag	tgagggaatgg	2220
aaagtgccca	gctcgggttg	gcacataagg	gaggctcccc	agcctgggaa	cgattataac	2280
agagggcccc	tactttcaca	gatgaggaac	ttgaggcaag	tcaccagccc	ctgatcattt	2340
cgcctaaaag	agcaaggact	agagttcctg	acctccaggc	cagtccctga	tccctgacct	2400
aatgttatcg	cggaatgatg	gtaagttaa	tgtctcttgc	atctgcatag	agagagtcct	2460
gggagcttag	gaagtgatgg	ggaacagtga	tgtatgcagc	tcatgactag	gtggacaggc	2520
ctctggggac	agctgggtaca	ggagggaaa	ggacctcacg	ggagggccag	aaacctggta	2580
agaggtgagg	tattaagggtc	tgggatggag	aagctctgag	ggtatatattt	tctgcctcta	2640
aaactgttgg	agaggggaatc	tgagaaagct	gcaaccaacc	aggaggctgg	ggtagcgtgg	2700
agaaggaatg	ggcttcctaa	ccttgagccc	tcttccctga	agatatatgt	atctacgggg	2760
gcctggggct	gggcgggctc	ctgcttctgg	cagtggctct	tctgtccgcc	tgccctgtgt	2820
ggctgcatcg	aagaggttag	cgctgcactc	cctccctccc	cctgcagcag	tgccccctgt	2880

gccccaccc	ccacacgctt	tcccactgct	ttcccagaac	actgcctggc	cctggagcca	2940
ctgggaagcc	aacaggggag	tccacgcctg	ctgggtgggg	gagcccggga	ggcgggagaa	3000
gcacaaaggg	tgggctgtgt	tgagcttctt	cttttcttcc	agtaaagagg	ctggagagga	3060
gctgggtgag	tctgggggaca	gggaaggggg	agggcaagag	agatcctgag	tgggtgagtg	3120
gggagaagca	tggtctgagcg	ctgagaggag	gggtggggac	gggagacaag	gagagagaaa	3180
gtaggagcat	gagagaggca	gagaaaatcg	aggcaaaaga	gaaagagaaa	atgagacaga	3240
aaccaagaga	aaaagtgaga	cagaggatag	gagagacagg	gagaaaatga	gagtgagaga	3300
gacacaaga	gaagagcaat	gaaagagaga	gagagagaga	ggctccagaa	ccaggcacag	3360
tggctcacgt	ctgtcattcc	actatcgcaa	ggctgaggca	ggaagatagc	ttgagctcag	3420
gggttgaaga	caatcctgga	caacatagtg	gactctgtct	ccaaagaaaa	aagagagaga	3480
gagagagaga	gagagagagg	gagagagaga	gagagagagg	gagagaagta	agaaaggctg	3540
gaggtgggag	cagaactcac	agggaaaggat	ctgacgcata	gcctcccata	agcaccttct	3600
gtcctggctc	caggccagg	gctcctcaga	gcaggaactc	cactatgcat	ctctgcagag	3660
gctgccagtg	cccagcagtg	agggacctga	cctcaggggc	agagacaaga	gaggcaccaa	3720
ggaggatcca	agagctgact	atgcctgcat	tgctgagaac	aaacccacct	gagcacccca	3780
gacaccttcc	tcaaccagg	cgggtggaca	gggtccccct	gtggtccagc	cagtaaaaaac	3840
catggtcccc	ccacttctgt	gtctcagtc	tctcagtc	tctcagcct	ccgttcaaat	3900
tgatcatcat	caaaacttat	gtggcttttt	gacctttgaa	tagggaattt	tttaaatttt	3960
ttaaaaatta	aaataaaaaa	aacacatggc	tcaccttcc	acctactctg	gggtcaata	4020
gtaatttgtt	gggtgaatga	cagtgttcag	ggacccaagc	ttcccataca	gccagaagag	4080
ggtatgtgtg	ggcctggcag	gaaagggcag	ttgccaagga	ggagtcata	ctgatccttc	4140
ccatttctca	ggacaatcag	gctcagcctc	ctgggactgg	gggaagcaga	tgtgctgagc	4200
tcccacatgg	tggtgggagg	ggcgctggga	ccacagccgg	cagctgcctt	cttgaccctt	4260
tccaggtcag	acctggtgga	agggaaagt	cagagttggg	ggaatccgga	gagagtagat	4320
ttggcatctg	gagaatggag	aagaaaaac	ttgagactca	tgaggagtta	gtggtggggc	4380
agatttattg	gggtcttttg	aagaggacta	gggacatctg	ggctctggaa	tcactcctcg	4440
gggcccctct	gaggagtggc	agtgtgttcc	catgtgacag	tgccctgggtc	agagagagga	4500
caggagctgc	tcagtgttgc	agtcccagg	ctctcctctt	cctgggtctct	gtcctccctc	4560
ctcccactct	cttactgccc	ctcccactcc	gtccactatt	gcccctgggt	ccattactca	4620
catttgcctt	ggtaatagac	ggtgctgccc	acggccacag	agagaaagct	gacagcatag	4680
aatccagccc	gaaggaggag	gactgtacca	gcccctagct	gaggatgttc	tgcatggggc	4740
aatggagacg	ggggttgggg	aagaagtga	cacaggctca	gggagggag	gggcctcaga	4800
ggagcatccc	tgccctccaa	ggacattgcc	tcttggggcc	tccagccagg	aggagacacc	4860
acctcccagc	atctcacctt	tctccaccac	cagccgagtc	ccattccctg	tcccgcacc	4920
aaggcccagc	acctccactc	tgcacacgta	gatgttggcg	tcattggcctc	gcacctccg	4980
gatgtgcagc	tcagcctggt	ggtcatggag	gaaacgggaa	gaagcaagt	gggcccaggc	5040
gcccctgaac	tctggggttc	cattcctcac	ctccttccct	ggaaccacct	catctcgga	5100
ccacgtgacg	gagccaatgg	ccagtctccc	ttggctggca	ttgaaggagc	agggcaggaa	5160
ggcagaggat	ccttcagggg	tacgaatctc	agggggctgg	gacaccagga	gagcacagga	5220
atcctggggg	gaaaagggag	accagagaa	acacctcccc	agttattcca	aagagaaaaa	5280
acaacagagc	ttggagtaga	acatcccagg	tttctccagg	catagggtgc	atgggaatag	5340
atactttggg	gcctcattaa	acccttccct	cttaaccaat	ctgatttctt	aacattgctt	5400
attaaatcat	ttttcggctg	ggtgcagtcg	cacgcctgta	atcccagcac	tttgggaggc	5460
cgaggtgggc	ggatcaccag	gtcaggagat	cagagccatc	ctggccaaca	tggtgaaacc	5520
ccgtctctac	taaaaaaata	caaaaattag	ccgggcatgg	tggtgtgcac	ctgtaatccc	5580
agctactcgg	gaggtctagg	caggagaatc	gcttgaaccc	gggaggcaga	gggtgcctg	5640
agccaagatt	gcgccattgc	actccagcct	gggcgacaaa	gcaagactcc	atctcaaaaa	5700
ataaaaaata	aaaatcattt	ttcaaatctt	tcctatacca	actctcactc	tcacctctctg	5760
ccatcattct	ccagccagtt	cagtagtaac	ttgtctagct	gaaatgtaaa	ccatcatggt	5820
gaaattaagc	tcattaatga	attgcagctg	cctagttaac	taatatcact	cattatatta	5880
tgcaggtatt	attttagtac	aaatggcatt	gtacagtaag	ccatccttcc	tcttttctt	5940
ttttcttttt	tttagatggg	gtcttgctct	gttgccagg	ctggaatgca	gtggtgcaat	6000
cttggtcac	tgcaactcc	gtcccctggg	ttcaagcgat	cctgggtgct	cagcctccca	6060
agtagctggg	actacaggca	cccaccacca	cgactggcta	atttttgat	tttcagtcga	6120
gacagggttt	ccaccatctg	gtctcaaa	cctgacctca	agtgatccac	ccacctcgga	6180
ccaggctggc	tcaaaactct	gatctcaagt	gatccacctg	cctcggcctc	ccaaagtgc	6240
accagccac	tcttggtttt	cgttaaagaa	agtaaactaat	taaaatctcca	gggtgaagacg	6300
tggccttaat	tggttgagat	tcctatttaa	ccgtccatg	ttgatgaatt	aaaccaaata	6360
ttaaaatccc	tgattaaatt	atctacttag	ggaaatttac	aagtcattct	atttcagtg	6420
ttctcaaa	tgagtgtgta	tggaaattac	ctggagcatc	tgctagaaca	gattcctggg	6480
cctaccccc	gagtttttga	ctcagtaggt	ctggagtggg	gcctaagaat	ttgttctag	6540
ttcccagaaa	tccacatttt	gagaactcct	gcatctagtt	aataatatgc	ctgatagtta	6600
aggctctctca	gttcatttaa	aacagtttcg	ccgggtgca	gttcgcacgc	ctataatccc	6660
aacacttttg	gaggccaagg	cgagtggatc	acctgaggtc	aggagtgtga	gaccagcctg	6720
gccaacatgg	tgaacctcg	tctctactaa	aaatacacia	gttagccagc	agtaatggca	6780

tgcacctgta atcctagcta cttgggaggc tgagacagga gaatcatttt tacccaggag 6840
gtggaggctg cag 6853

<210> 16
<211> 10
<212> DNA
<213> Homo sapiens

<400> 16
accttgaagt

10

<210> 17
<211> 277
<212> PRT
<213> Homo sapiens

<400> 17
Met Ile Ile Leu Ile Tyr Leu Phe Leu Leu Leu Trp Glu Asp Thr Gln
1 5 10 15
Gly Trp Gly Phe Lys Asp Gly Ile Phe His Asn Ser Ile Trp Leu Glu
20 25 30
Arg Ala Ala Gly Val Tyr His Arg Glu Ala Arg Ser Gly Lys Tyr Lys
35 40 45
Leu Thr Tyr Ala Glu Ala Lys Ala Val Cys Glu Phe Glu Gly Gly His
50 55 60
Leu Ala Thr Tyr Lys Gln Leu Glu Ala Ala Arg Lys Ile Gly Phe His
65 70 75 80
Val Cys Ala Ala Gly Trp Met Ala Lys Gly Arg Val Gly Tyr Pro Ile
85 90 95
Val Lys Pro Gly Pro Asn Cys Gly Phe Gly Lys Thr Gly Ile Ile Asp
100 105 110
Tyr Gly Ile Arg Leu Asn Arg Ser Glu Arg Trp Asp Ala Tyr Cys Tyr
115 120 125
Asn Pro His Ala Lys Glu Cys Gly Gly Val Phe Thr Asp Pro Lys Gln
130 135 140
Ile Phe Lys Ser Pro Gly Phe Pro Asn Glu Tyr Glu Asp Asn Gln Ile
145 150 155 160
Cys Tyr Trp His Ile Arg Leu Lys Tyr Gly Gln Arg Ile His Leu Ser
165 170 175
Phe Leu Asp Phe Asp Leu Glu Asp Asp Pro Gly Cys Leu Ala Asp Tyr
180 185 190
Val Glu Ile Tyr Asp Ser Tyr Asp Asp Val His Gly Phe Val Gly Arg
195 200 205
Tyr Cys Gly Asp Glu Leu Pro Asp Asp Ile Ile Ser Thr Gly Asn Val
210 215 220
Met Thr Leu Lys Phe Leu Ser Asp Ala Ser Val Thr Ala Gly Gly Phe
225 230 235 240
Gln Ile Lys Tyr Val Ala Met Asp Pro Val Ser Lys Ser Ser Gln Gly

245 250 255
 Lys Asn Thr Ser Thr Thr Ser Thr Gly Asn Lys Asn Phe Leu Ala Gly
 260 265 270
 Arg Phe Ser His Leu
 275

<210> 18
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 18
 ggaccgaggg

10

<210> 19
 <211> 194
 <212> PRT
 <213> Homo sapiens

<400> 19
 Met Arg Leu Leu Pro Cys Ala Ala Gly Thr Ser Ala Phe Leu Ser Ser
 1 5 10 15
 Gln Leu Thr Ile Gly Ala Ser Ser Ser Ser Arg Pro Ser Thr Pro Ala
 20 25 30
 Trp Phe Ser Val Leu Glu Tyr Ile Gln Pro Phe Arg Ser Cys Ser Phe
 35 40 45
 Leu Leu Ser Ser Ser Pro Pro Gly Ala Leu Pro Cys Ser Leu Cys Thr
 50 55 60
 Ser Asp Ser Ile Leu Pro Val Ser Asp Asp Phe Thr Ser Ser Lys Arg
 65 70 75 80
 Pro Leu Gly Pro Ala Phe Ser Lys Ala Phe Leu Glu Asn Ala Gly Gln
 85 90 95
 Asp Leu Pro Phe Asn Phe Val Phe Ser Val Cys Ser Val Ser Ser Ala
 100 105 110
 Gly Cys Leu Ser Ser Arg Ile Thr Glu Asp Met Ser Val Leu Phe Pro
 115 120 125
 Leu Met Pro Gly Ala Arg Leu Gly His Val Cys Ser Ala Met Ala Asn
 130 135 140
 Val Arg Arg Gln Gly Trp Ser His Ser Gly Leu Leu Ala Leu Trp Pro
 145 150 155 160
 Ser Gly Arg Ser Lys Gly Arg Ser Cys Gln Lys Glu Thr Cys Leu Val
 165 170 175
 Thr Gly Arg Ser Gln Gly Phe Gln Asn Pro Ile Gln Val Ile Pro Ser
 180 185 190
 Trp Trp

<210> 20
 <211> 1841

<212> DNA

<213> Homo sapiens

<400> 20

```

tgttttgttt ctttatttgg acttaacggc cactcatttc tgttggcgcc atgcgactgt 60
tgccgtgtgc agcagggacc tctgccttcc tctcctccca gctcacaata ggtgcctctt 120
cctcctcccg gccatccacg ccggcctggt tttcggttct ggaatacatc caaccctttc 180
gttcgtgttc ttttctcctg tcattcttct cgcttggagc tcttccatgt tctctttgca 240
caagcgactc tatccttcca gtttcagatg atttcacctc ctcaaagagg cctttgggac 300
cagccttttc taaagctttc cttgagaatg ctggtcaaga tttacctttt aattttgttt 360
tttctgtttg ttctgtctcc tctgccggct gcttgagctc caggatcaca gaggacatga 420
gtgttctgtt cccactgatg cctggtgcca gactgggaca cgtgtgctca gccatggcta 480
atgtgcgag gcagggatgg tcacacagtg gcctcttggc gctctggccg agtgggagga 540
gcaaggggag gagctgtcag aaggagacct gcctagtcac aggcaggtca cagggcttcc 600
agaatcccat tcaggtgatc ccctcctggt ggtaaacagc acacctcaga agtgggtgtt 660
tgtcatagt gttcttatgg ctcccacagc ctcatctctg ttttgtccat ttgggattga 720
ggagtgtttg agtccagcgt caccttcaga gccagtgcac tccttatcaa tagtgagggc 780
cgctcgtttt gattgttccg tttcgcttga taaagctgag tgagaattat attggccaac 840
agaaacgtct ctctccatgt gaacacagcc actgggtcac cagggacgtg gcttacagtt 900
ccagcaggct cctgtcacac tgggtgcctt gggcagtaca gaactgtgct gctgggtggtg 960
acgtccgggg ctgtgcagtg tcgcttccct ttgcatatgt tgctattggt gtcatttcct 1020
ctctgagatc agtggaagtc tcaggtcagt cccgtggccc tgtattgctt agtgggccat 1080
gtgttttaac ttccaaaagg catattttta atagttattg gagtgttgat gtgaagagct 1140
cctttgataa tgtattttct ttgtagaaat ttaaaagcct gttgattcaa gaactttcta 1200
aagtatttcc ggaagacatg gctaagtatc gaagcatccg gggggaggat caccgcctt 1260
cttaaccagc tcaccctccc tgtgtgaaga tcccctggga ctgcgatgcg gcgtgaggct 1320
gggactgcga gtgctgacgc caccttccct ctgaggtggg actgggccct ggacacaccc 1380
ctcagcccct ctgtcctcat tgtttggcct catgggaccg aggggctgga ggagaggcgg 1440
agctgtgccc cagctgttcc agcagcttgt ctggcgtaaa ctggctttca gagtgtgac 1500
ccctcatcac tgtggggatc attctctctg agggcagatg aggcgcagga aaatagctt 1560
ggaaatgtta aatatgatgg gtaaatataa agttttacaa cattctacct aatatttttc 1620
ttttaacata ctctttctgt tctattgtat tatggtgtcc gaaagctaaa taacgactag 1680
gaaaaatttt tttaaaaaaa gaaaaatcag tttaatgtgg gaagtactta agtgggtatta 1740
tattttacat tttcaagtat agtgcataaa gaatgtttta aatgtaactg ttttcatgga 1800
tttcaattag acatgcctat aataaactaa gtatgtggct t 1841

```

<210> 21

<211> 10

<212> DNA

<213> Homo sapiens

<400> 21

gccggcttga

10

<210> 22

<211> 439

<212> PRT

<213> Homo sapiens

<400> 22

```

Met Thr Ser Ile Leu Thr Val Leu Ile Cys Leu Gly Leu Ser Leu Asp
  1           5           10           15

```

```

Pro Arg Thr His Val Gln Ala Gly Pro Leu Pro Lys Pro Thr Leu Trp
      20           25           30

```

```

Ala Glu Pro Gly Ser Val Ile Thr Gln Gly Ser Pro Val Thr Leu Arg
      35           40           45

```

```

Cys Gln Gly Ser Leu Glu Thr Gln Glu Tyr His Leu Tyr Arg Glu Lys
      50           55           60

```

```

Lys Thr Ala Leu Trp Ile Thr Arg Ile Pro Gln Glu Leu Val Lys Lys

```

65					70					75				80
Gly	Gln	Phe	Pro	Ile ₈₅	Leu	Ser	Ile	Thr	Trp ₉₀	Glu	His	Ala	Gly	Arg ₉₅ Tyr
Cys	Cys	Ile	Tyr ₁₀₀	Gly	Ser	His	Thr	Ala ₁₀₅	Gly	Leu	Ser	Glu	Ser ₁₁₀	Ser Asp
Pro	Leu	Glu ₁₁₅	Leu	Val	Val	Thr	Gly ₁₂₀	Ala	Tyr	Ser	Lys	Pro ₁₂₅	Thr	Leu Ser
Ala	Leu ₁₃₀	Pro	Ser	Pro	Val	Val ₁₃₅	Thr	Ser	Gly	Gly	Asn ₁₄₀	Val	Thr	Ile Gln
Cys ₁₄₅	Asp	Ser	Gln	Val	Ala ₁₅₀	Phe	Asp	Gly	Phe	Ile ₁₅₅	Leu	Cys	Lys	Glu Gly ₁₆₀
Glu	Asp	Glu	His	Pro ₁₆₅	Gln	Cys	Leu	Asn ₁₇₀	Ser	His	Ser	His	Ala	Arg ₁₇₅ Gly
Ser	Ser	Arg	Ala ₁₈₀	Ile	Phe	Ser	Val	Gly ₁₈₅	Pro	Val	Ser	Pro	Ser	Arg Arg
Trp	Ser	Tyr ₁₉₅	Arg	Cys	Tyr	Gly	Tyr ₂₀₀	Asp	Ser	Arg	Ala	Pro ₂₀₅	Tyr	Val Trp
Ser	Leu ₂₁₀	Pro	Ser	Asp	Leu	Leu ₂₁₅	Gly	Leu	Leu	Val	Pro ₂₂₀	Gly	Val	Ser Lys
Lys ₂₂₅	Pro	Ser	Leu	Ser	Val ₂₃₀	Gln	Pro	Gly	Pro	Val ₂₃₅	Val	Ala	Pro	Gly Glu ₂₄₀
Lys	Leu	Thr	Phe	Gln ₂₄₅	Cys	Gly	Ser	Asp	Ala ₂₅₀	Gly	Tyr	Asp	Arg	Phe Val ₂₅₅
Leu	Tyr	Lys	Glu ₂₆₀	Trp	Gly	Arg	Asp	Phe ₂₆₅	Leu	Gln	Arg	Pro	Gly ₂₇₀	Arg Gln
Pro	Gln	Ala ₂₇₅	Gly	Leu	Ser	Gln	Ala ₂₈₀	Asn	Phe	Thr	Leu	Gly ₂₈₅	Pro	Val Ser
Arg	Ser ₂₉₀	Tyr	Gly	Gly	Gln	Tyr ₂₉₅	Thr	Cys	Ser	Gly	Ala ₃₀₀	Tyr	Asn	Leu Ser
Ser ₃₀₅	Glu	Trp	Ser	Ala	Pro ₃₁₀	Ser	Asp	Pro	Leu	Asp ₃₁₅	Ile	Leu	Ile	Thr Gly ₃₂₀
Gln	Ile	Arg	Ala	Arg ₃₂₅	Pro	Phe	Leu	Ser	Val ₃₃₀	Arg	Pro	Gly	Pro	Thr Val ₃₃₅
Ala	Ser	Gly	Glu ₃₄₀	Asn	Val	Thr	Leu	Leu ₃₄₅	Cys	Gln	Ser	Gln	Gly ₃₅₀	Gly Met
His	Thr	Phe ₃₅₅	Leu	Leu	Thr	Lys	Glu ₃₆₀	Gly	Ala	Ala	Asp	Ser ₃₆₅	Pro	Leu Arg
Leu	Lys ₃₇₀	Ser	Lys	Arg	Gln	Ser ₃₇₅	His	Lys	Tyr	Gln	Ala ₃₈₀	Glu	Phe	Pro Met
Ser ₃₈₅	Pro	Val	Thr	Ser	Ala ₃₉₀	His	Ala	Gly	Thr	Tyr ₃₉₅	Arg	Cys	Tyr	Gly Ser ₄₀₀
Leu	Ser	Ser	Asn	Pro ₄₀₅	Tyr	Leu	Leu	Thr	His ₄₁₀	Pro	Ser	Asp	Pro	Leu Glu ₄₁₅

Leu Val Val Ser Gly Ala Ala Glu Thr Leu Ser Pro Pro Gln Asn Lys
 420 425 430

Ser Asp Ser Lys Ala Gly Glu
 435

<210> 23
 <211> 1571
 <212> DNA
 <213> Homo sapiens

<400> 23
 ctctgtcctg ccagcaccga gggctcatcc atccacagag cagtgcagtg ggaggagacg 60
 ccatgacctc catcctcacg gtcctgatct gtctcgggct gagcctggac cccaggaccc 120
 acgtgcaggc agggccccct cccaagccca ccctctgggc tgagccaggc tctgtgatca 180
 cccaagggag tcctgtgacc ctcaagggtc agggggagcct ggagacgcag gaggaccatc 240
 tatatagaga aaagaaaaca gcaactctgga ttacacggat cccacaggag ctgtgaaga 300
 agggccagtt ccccatccta tccatcacct gggaacatgc agggcgggtat tgctgtatct 360
 atggcagcca cactgcaggc ctctcagaga gcagtacccc cctggagctg gtggtgacag 420
 gagcctacag caaaccacc ctctcagctc tgcccagccc tgtggtgacc tcaggagggg 480
 atgtgaccat ccagtgtgac tcacagggtg catttgatgg cttcattctg tgtaagggaag 540
 gagaagatga acaccacaa tgcctgaact cccattccca tgcccgtggg tcatcccggg 600
 ccattctctc cgtgggcccc gtgagcccaa gtcgcagggtg gtcgtacagg tgctatgggt 660
 atgactcgcg cgctccctat gtgtggtctc taccagtgga tctcctgggg ctcctgggtc 720
 caggtgtttc taagaagcca tcactctcag tgcagccggg tcctgtcgtg gcccctgggg 780
 agaagctgac cttccagtgt ggctctgatg ccggctacga cagatttgtt ctgtacaagg 840
 agtggggacg tgacttcctc cagcgccctg gccggcagcc ccaggctggg ctctcccagg 900
 ccaacttcac cctgggcccc gtgagccgct cctacggggg ccagtacaca tgctccgggt 960
 catacaacct ctctccgag tggtcggccc ccagcgaccc cctggacatc ctgatcacag 1020
 gacagatccg tgccagacc ttctctccg tgcggccggg cccacagtg gcctcaggag 1080
 agaacgtgac cctgctgtgt cagtcacagg gagggatgca cactttcctt ttgaccaagg 1140
 agggggcagc tgattccccg ctgctgtctaa aatcaaagcg ccaatctcat aagaccagg 1200
 ctgaattccc catgagtcct gtgacctcgg cccacgcggg gacctacagg tgctacggct 1260
 cactcagctc caaccctac ctgctgactc accccagtga cccctggag ctctggtgct 1320
 caggagcagc tgagaccctc agcccaccac aaaacaagtc cgactccaag gctggtgagt 1380
 gaggagatgc ttgccgtgat gacgctgggc acagagggtc aggtcctgtc aagaggagct 1440
 ggggtgtcctg ggtggacatt tgaagaatta tattcattcc aacttgaaga attattcaac 1500
 acctttaaca atgtatatgt gaagtacttt attctttcat attttaaaaa taaaagataa 1560
 ttatccatga a 1571

<210> 24
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 24
 ttcactgtaa 10

<210> 25
 <211> 298
 <212> PRT
 <213> Homo sapiens

<400> 25
 Met Lys Thr Leu Gln Ser Thr Leu Leu Leu Leu Leu Val Pro Leu
 1 5 10 15
 Ile Lys Pro Ala Pro Pro Thr Gln Gln Asp Ser Arg Ile Ile Tyr Asp
 20 25 30
 Tyr Gly Thr Asp Asn Phe Glu Glu Ser Ile Phe Ser Gln Asp Tyr Glu
 35 40 45

Asp Lys Tyr Leu Asp Gly Lys Asn Ile Lys Glu Lys Glu Thr Val Ile
 50 55 60
 Ile Pro Asn Glu Lys Ser Leu Gln Leu Gln Lys Asp Glu Ala Ile Thr
 65 70 75 80
 Pro Leu Pro Pro Lys Lys Glu Asn Asp Glu Met Pro Thr Cys Leu Leu
 85 90 95
 Cys Val Cys Leu Ser Gly Ser Val Tyr Cys Glu Glu Val Asp Ile Asp
 100 105 110
 Ala Val Pro Pro Leu Pro Lys Glu Ser Ala Tyr Leu Tyr Ala Arg Phe
 115 120 125
 Asn Lys Ile Lys Lys Leu Thr Ala Lys Asp Phe Ala Asp Ile Pro Asn
 130 135 140
 Leu Arg Arg Leu Asp Phe Thr Gly Asn Leu Ile Glu Asp Ile Glu Asp
 145 150 155 160
 Gly Thr Phe Ser Lys Leu Ser Leu Leu Glu Glu Leu Ser Leu Ala Glu
 165 170 175
 Asn Gln Leu Leu Lys Leu Pro Val Leu Pro Pro Lys Leu Thr Leu Phe
 180 185 190
 Asn Ala Lys Tyr Asn Lys Ile Lys Ser Arg Gly Ile Lys Ala Asn Ala
 195 200 205
 Phe Lys Lys Leu Asn Asn Leu Thr Phe Leu Tyr Leu Asp His Asn Ala
 210 215 220
 Leu Glu Ser Val Pro Leu Asn Leu Pro Glu Ser Leu Arg Val Ile His
 225 230 235 240
 Leu Gln Phe Asn Asn Ile Ala Ser Ile Thr Asp Asp Thr Phe Cys Lys
 245 250 255
 Ala Asn Asp Thr Ser Tyr Ile Arg Asp Arg Ile Glu Glu Ile Arg Leu
 260 265 270
 Glu Gly Asn Pro Ile Val Leu Gly Lys His Pro Asn Ser Phe Ile Cys
 275 280 285
 Leu Lys Arg Leu Pro Ile Gly Ser Tyr Phe
 290 295

<210> 26
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 26
 aattctgtaa

10

<210> 27
 <211> 305
 <212> PRT
 <213> Homo sapiens
 <400> 27

Met₁ Ala Ala Gly Leu₅ Ala Arg Leu Leu Leu₁₀ Leu Leu Gly Leu Ser₁₅ Ala
 Gly Gly Pro Ala₂₀ Pro Ala Gly Ala Ala₂₅ Lys Met Lys Val Val₃₀ Glu Glu
 Pro Asn Ala₃₅ Phe Gly Val Asn Asn₄₀ Pro Phe Leu Pro Gln Ala Ser Arg
 Leu Gln Ala Lys Arg Asp Pro₅₅ Ser Pro Val Ser Gly₆₀ Pro Val His Leu
 Phe Arg Leu Ser Gly Lys₇₀ Cys Phe Ser Leu Val₇₅ Glu Ser Thr Tyr Lys₈₀
 Tyr Glu Phe Cys Pro₈₅ Phe His Asn Val Thr₉₀ Gln His Glu Gln Thr₉₅ Phe
 Arg Trp Asn Ala₁₀₀ Tyr Ser Gly Ile Leu₁₀₅ Gly Ile Trp His Glu Trp Glu
 Ile Ala Asn₁₁₅ Asn Thr Phe Thr Gly₁₂₀ Met Trp Met Arg Asp Gly Asp Ala
 Cys Arg Ser Arg Ser Arg Gln₁₃₅ Ser Lys Val Glu Leu Ala Cys Gly Lys
 Ser Asn Arg Leu Ala His Val Ser Glu Pro Ser₁₅₅ Thr Cys Val Tyr Ala₁₆₀
 Leu Thr Phe Glu Thr₁₆₅ Pro Leu Val Cys His₁₇₀ Pro His Ala Leu Leu Val
 Tyr Pro Thr Leu₁₈₀ Pro Glu Ala Leu Gln₁₈₅ Arg Gln Trp Asp Gln Val Glu
 Gln Asp Leu₁₉₅ Ala Asp Glu Leu Ile Thr Pro Gln Gly His₂₀₅ Glu Lys Leu
 Leu Arg Thr Leu Phe Glu Asp Ala Gly Tyr Leu Lys₂₂₀ Thr Pro Glu Glu
 Asn Glu Pro Thr Gln Leu₂₃₀ Glu Gly Gly Pro Asp₂₃₅ Ser Leu Gly Phe Glu₂₄₀
 Thr Leu Glu Asn Cys₂₄₅ Arg Lys Ala His Lys₂₅₀ Glu Leu Ser Lys Glu Ile
 Lys Arg Leu Lys₂₆₀ Gly Leu Leu Thr Gln₂₆₅ His Gly Ile Pro Tyr Thr Arg
 Pro Thr Glu Thr Ser Asn Leu Glu His Leu Gly His Glu Thr Pro Arg
 Ala Lys Ser Pro Glu Gln Leu₂₉₅ Arg Gly Asp Pro Gly₃₀₀ Leu Arg Gly Ser
 Leu
 305

<210> 28

<211> 37476

<212> DNA

<213> Homo sapiens

<400> 28

gatcaccatg	gcaaccggcc	tgtgtttgtt	ccaacagggtg	ggcagaacct	ggtgcaagac	60
cgcagaggca	ggggcgctac	tcccagaccc	ggtctccaga	cccaccagcc	cctcaggaga	120
acactgctac	agagtggggc	tgagggccca	ggccccctcc	cctgcttccc	ccaccgctc	180
cctgtcccag	ccagtccttg	catttcctgc	aagctccctg	ccctccccc	ccggcctggc	240
tgtgctgggt	caggatccgg	cagagcctgg	gccgagcgtc	ccgaaggaga	gcccacccag	300
gggtccttgc	cccagtcctg	cacagacagg	tgtggggcgg	gcatggattt	gagtgaattg	360
ggatccagaa	gccaaagacag	ttgttggcag	acgggcccc	tccctttgaa	gagctgtgtt	420
ctggaagcct	gtctggcctt	gggtgccca	ggtgcccttc	gggtgaaaac	tccaaagtct	480
gaccaccccc	cacccttgtt	ccccaaagg	ctgcgagtg	ccagctgtat	tgcaccaagg	540
gcaggtgcga	ccggggggcc	gctgcgcctt	cggcaagtct	gagagcctca	ggcccccgac	600
aggaccatgc	ttagcacgga	cttctcaaa	tgtggtccct	ggactggcgg	cagcagactg	660
gagctgggaa	gtcgaggcgc	tggctggcgt	ggggaccaca	cctgtgatct	cagcgcttta	720
ggaggccggg	gaagggggat	cccttgagcc	tgggagtttg	agaccagcct	gggcaacata	780
gggagacccc	catctccaca	aaagaaacca	aaaaacatta	gccggcgggt	ggtcccagct	840
actggggagg	ctgaagcggg	aggaccactt	gagccggaga	gttcgaggtt	gcagcgagtg	900
gagatgggtg	cactgcgctc	caacctgggt	gagagctatt	tttgagagct	ttttcgagac	960
tctgtctgaa	aaaaaaagg	aaaaaacgca	gatcgcgccc	ccaccagac	ttcgggacca	1020
gcgggttggg	gacaggaccc	atcaatgcct	taacaaaccc	tgcaggagac	cgcgggcccg	1080
ctgggggttcc	aggacagcgc	agcggaggaa	gagcggggag	ccaccggga	ggggctggcg	1140
gggtccagtg	cagacccag	gggtccggg	ggctctggcc	ggctgtgacc	ccacaagcgc	1200
caactggggc	ctcaccagag	cccaggggac	gatcgcccca	cagaccgtcc	gcgtgggcgt	1260
ggggcgctcc	gagtggatcc	ctgcaggacg	catccccg	ccgagaagag	agcggcgggg	1320
cccgcgaggt	gtcccttagt	caggcccttc	cagaaaggct	tgcaccgctg	ctggtcttcg	1380
tgaccgtggc	ttccctgggg	tccgcggagg	agaccccgcc	gaccaccccc	ttctagaccc	1440
cacgcgccac	ggggagggcg	cgcgcgcca	ggcggggtcc	cgaggcagcc	aagcccgtc	1500
cccgtcccgc	agccacctgt	gggttgactc	acagccccgc	atcccggggg	agggggctcc	1560
ggcccggctc	aaggcacgga	gacccggcc	gggtccccta	aggcgccgtg	gacgcggggg	1620
gcggggcgag	cgcggtcagg	ccccgcccc	gtcagggccc	gccccgcccc	cagtcagacc	1680
cgcgccccgc	agcggcgctc	gcccagcagt	tgctcagtcg	tccggcgcg	cggtgggag	1740
cggtagctgt	gcctcggcgt	ccccgagctg	gtgccgagcg	cagcgcccca	ggagcccccc	1800
catccccgcg	ccggcccacg	gtaagtgcg	cgggtgcgg	ccagtgtgt	tgggcccgcg	1860
tctgggctcg	gccccgcgcc	ccaaccggga	cggaccgacg	gccgcccgc	ccttggtccg	1920
ggtccccccg	gaggcacggg	ggtggcctcc	gagggcgggg	gcgggtggaa	ggtctcgagc	1980
tccctcggc	ctgaggcgcg	ctgaggccgc	gcattgcttg	aggggcgggg	tctcggggcg	2040
atttccctc	cgaacgcgca	ggagaggggtg	cggcggtggg	gtccgcactg	cggagcggac	2100
ggggagggta	tccgggcgct	ggcgacggcc	tccccgcac	ttgagccgcc	aagcgaggct	2160
tgcacccccg	cgctccctc	cccgggggag	cagacgggcc	cggggcctcc	cagcaccctt	2220
cccgggcggg	ctgcaccccc	cagcgtcctc	gtcccttccc	tcctcccga	ggtctctgta	2280
aggcccttta	gtcccccgca	aatccctgcg	agtgcgaagc	ggccctcct	tcagccccac	2340
tcggcaaagg	gctgtgggct	cttgccctgg	ccggggtctc	cgggggagga	gtttaaagt	2400
acacactaga	cccgcctcca	gactcacccc	ggaccgccc	ccccctccga	ggggagagac	2460
ccccatagaa	ctgggggtgt	ggctcctgcg	ggcgcccg	gccccgggaa	agggggtg	2520
gcgaggcggg	gcagggagg	ggtcgggtg	cagagggagg	agggagcgct	gcgcagacac	2580
gggcgagggg	aggaagaagg	cgctcgggc	tccggtctcc	tgcgctcctg	ctgggcgaca	2640
gcggctggac	cttggcagcc	gtctgaggt	tccgggagcc	caaggccacc	ggcagcgctt	2700
gttagaatgc	agattcccg	ggttcgactg	gaatgagacc	ccggggagca	gcgtttgcag	2760
cgggcccgc	aggtgacctg	cacctgggca	ccgcatcgg	cttcgcagg	gccatctgga	2820
ggagtccgtg	agaccaggga	gagattcctg	cgggaaaccc	ccagggcctg	gtgaagggga	2880
gtcggggggc	tgagggagga	agagcaggga	agctggagag	aagaccaaga	gggcagcgct	2940
gagagccggc	ctgtggggca	gatgtcctgc	cacttcccgc	cccaggcaca	caccagggg	3000
accatcagct	aacccctgc	cccaaagtgt	gtaagaagg	cgcccccaac	acgttgaagg	3060
cgtttgttcc	tagacagcct	gcactgggga	ggaaggctga	cccagcctct	gctccgctgc	3120
ccccagactc	agctgcctca	ctgcccccaa	ccctgagccc	ccaaatggcc	tggcaggaga	3180
cagcttggct	taagagtccc	aggttcacat	ctgggctctg	ggcaagtcgc	ccccaccctg	3240
ggctgtgaaa	tgagatgtgc	catcctcctt	gcaaaatgtc	attccccatt	cagccaggcg	3300
ctgggggaca	gccagggagc	aagggtggat	gggcagggtg	ggctcagccc	cagtgcaca	3360
aggcaggagc	tccttggcgg	tcctccctct	gctcactgga	cagatgggag	tgtggttctg	3420
gtcctggcat	caggcatgag	cagagtatgg	gggggttcat	cctgcctcat	gctactgtcc	3480
cagggtgaga	gactgggaag	gagttacatc	acgtctgggc	ccacgggaaa	gacggggggg	3540
cccctgcacc	tggtcactc	ctgccccgta	agtgtgaga	gccaccaggg	cagccagtgc	3600
ttcctgtact	ggtcacagac	agatgcata	gccaccatt	cacagcagct	ccggcctctg	3660
tggcagggtg	ctcccagggc	ccacgcagtt	ggctgcagcc	aggctatcag	ctgcccagga	3720
agcctgggg	ttccagggga	accaaagg	tttctgcgca	cccaggatga	gccaagtctt	3780

ctgtacacct	gtgccaaagac	gccattattg	cacccatcct	agcaacgggc	agtgggaggg	3840
gcaggagctg	cagttgggggt	cagggggggct	ggcctagctg	gttgccccag	gagcttagca	3900
ctggggacac	caaagagacc	ttgactgccg	tccccagccc	acttgctggc	tctgggacct	3960
tgggctgggt	cccagccctg	cgctccgtt	tgccatctg	agatgcaggg	agcaaggggc	4020
acccctctcg	ggccgctctg	gggatgagga	ggtgaatgtg	gagaggaaaa	tgaggggagg	4080
gtaccgagcc	ccctttttct	gcctggaaca	tcttgttatc	acaaagacac	cccctcccca	4140
accccgagct	tcaggccact	tggctcctgc	gtcacatttg	tctccaggac	tgcgagcccc	4200
catcgccccct	gtcacacgca	catgccgcgg	cggttgccct	ggcaacagct	ccagcagcgg	4260
ctgctccact	ggcttcgggg	gacagcaggg	ggagcagcct	cttccaccca	gcaaggaggg	4320
ggaagggggc	ttgtgactcc	caggccctg	aggatttcac	gactggctgt	gggggtgaca	4380
gcttcagaa	gcactgtacc	cagcctacca	agcccccca	tcttttgggg	gggaggctca	4440
caggggaagg	caactgatag	tttgaaggag	gtgagtcccc	cccgaagca	ttgcctctca	4500
caggggtggga	gacgcaaacc	ctccttctct	tcctcagctc	taggggtctg	tcccgtaaga	4560
gcctgcaccc	agggctcggg	ggctggcagg	gaccttgaa	tcaagacgag	atatccggga	4620
atggtgtctg	agctgggtgtg	ggtatagctt	ggggctgggtg	accctgaacc	cccacagaca	4680
cagcaggggc	ccgggagctg	gctcgacaaa	gtccgttttg	aaccaccac	tgccatgggtg	4740
tgctcaggcc	cccacgctgc	ccactgagct	tatgagtcta	gtgccactgt	ggggagctgg	4800
cctggcattgc	agggccaact	gtccgtgaga	gcctgagtgg	gcggtagaac	ctgctggctct	4860
tcgtttccct	tcctgtcccc	tgagaatgac	tatgcccctt	gcgagggcca	gctgtccag	4920
cccgaggca	ccctccgcag	gtggatgagg	gagcactggc	aagggaagg	accagacag	4980
caaggccagg	accccggggg	gctgggggag	cgcggtgtcc	tagaaatcag	aggccggcaa	5040
ggctgggaca	ggaacccaag	agtgggtacc	tctctgtgg	tggggcagag	gtggcctggc	5100
tctgggacag	tttccatgga	gcggggagga	gacctccctg	gggctgctgg	ggggcagagg	5160
cagtgcagag	gtggagggtc	tgacactgca	ccatgcaga	ggaagagggt	ctcaggaggc	5220
tacggtgtct	tagccatcat	cctaccctcg	ttgaaggggg	tgtgcagggc	aggaccaccc	5280
ggccttccct	cctgacacag	gtgggctggg	cggtgctggc	tcatgcctgt	aatcccagca	5340
ctttgggagg	ccgaggctgg	tggatcacct	gaggtcagga	gttcaagacc	agcctggcca	5400
acatagcgaa	accccgctct	caactgaaagt	acaaaaaaat	cagctgggct	tgggtggcggg	5460
cgctgttaat	cccagttact	tgggaggctg	aggcaggaga	atcgcttgaa	cctggaaggc	5520
ggagggttgc	gtgagcggag	actgcgccac	tgcaactcac	ttgaacctgg	agggtggaggc	5580
tgcaagtgc	tgagatcatg	ccactgctct	ccagcctgga	ctccatctca	acaaaaacaaa	5640
acaaaaagtg	ggtggggcct	tccagggggc	ttccgggtct	gcctgggccc	aggctccggg	5700
ccttggcaga	gggcagctct	gtcccagcgg	tgacatctgg	gctgctgctg	ctgctgctgc	5760
tccaagcagg	aggctgacag	ggttggtcag	ttgtcatggc	aacatgaggg	gaactgacgc	5820
tggatggact	gggagggaac	tgtagcccc	acaccggca	ttttccagcg	gggcccggct	5880
cgcttcattgt	tctcgtttgt	accaactatg	gggagcaggc	cctgctttgt	ggggacactg	5940
ccccccacct	tctactgtcc	acctcaggtt	tgctgaccag	ctcctggggg	gaggaggggg	6000
cattgacccc	cgctgtgctt	ggtgagcgag	gacccccagg	cagtccccct	cagggcctcc	6060
tggccacatg	ggcatcctcc	acccaggatg	tcagggttct	ggcgagtttt	tacctcctac	6120
ataggagagc	tgggtacctg	gcacccgtcc	tggccccaga	gtcggcacgg	gggctggggg	6180
agctgagtgc	tggtggccag	atcttggtcc	tcttgatgca	gccccaggaa	acaggcgccct	6240
cctgctgtgc	tgtcgcgggg	aggggtgtgg	ggggcgagtg	tggcaaccgc	ctccctcgcc	6300
agtgcagccc	cataagtccc	tgatggaccc	ccagctctgg	gtgggcacca	gccccgcctt	6360
tgacagcatt	tgctacctcc	tccctgaagc	cgccctgcag	ggggcgcagg	gacttagccg	6420
gtcaccacagc	aggtgtgggg	acaggagccc	tcaccgcctg	accagggtgc	cagcgccatc	6480
caggctgtgg	gtgagctccc	aggcctctga	cccgcctctt	cccgcctcac	ccccccacct	6540
ccccgcctgc	tgtggtgcac	ggagggttga	gtggacgacc	tgaggctgcg	ggctgtgctc	6600
tctgctgtag	gtcaccgcgc	atgtcgacct	tgctggacat	taagagcagc	gtgctcaggc	6660
aggtgcagggt	gtgcccgtcc	ttccgcgcga	ggactgagca	ggacccaggg	agtgccagcg	6720
ccgacccgca	ggagcctgcc	acgggggcct	ggtgggtgcc	gagggggcca	gccccacacg	6780
cccagaaggc	catttcccgc	cagacttcac	acatggccgc	ccctgcccc	gcacctctgg	6840
gcgggaagga	gggtctgcag	ttagctgtgc	catagctgtg	atgtcacatg	gccccctctg	6900
cccctggagt	cgagggtccc	aggcaggggc	cagcgtgctg	agagctgtga	gctgaccccg	6960
ctctttctcc	aggaaacccg	gggatggcgt	ggagttcttt	gcccacatgc	gcctcatgct	7020
gaagaagggg	gaaggcagac	agggcttgcc	gtgcctcgag	gtaagggtgc	cacccccagg	7080
gcccgatacc	acagcccagc	atggggctcc	ccctttgctc	ctccccgctc	cctcctgccc	7140
tcgctccccac	ctcctggggc	accactgttc	ttccctgctg	caggcctggg	gctctccctg	7200
ccgtcagagc	cctcaccctg	ggccccacac	acagggtccc	ctgcgcagtg	gctcgccagc	7260
acccccggag	cctgtggatc	ccagcctcgg	cctgagagcc	ctggccccag	aggaggtaaa	7320
ggtgggggtc	ggaaccaggg	gcagtcgtct	gcagcggctg	ctccctcagc	cgtttagagg	7380
cacctactgt	gtgcacaggg	tctggtgatg	gcagagagcc	aggggtgagtg	ggtgggtgag	7440
tgaggagcgg	aggggtcctc	ggtgcaaaga	ggcagaggtg	gggcttgggc	tcaggcagac	7500
aggaggtgct	ggtggcctgg	cgaggaggcc	gacgcggcac	tgtggccgcc	cttcccccac	7560
ctgcagggtg	agatgtctta	cgaggaggcc	ctgtacacgg	tgctttaccg	cgcggttacc	7620
atgggcccctg	accagggtgga	cgacgaggag	gccctgctca	gctatctcca	gcaggctcagc	7680

ccaccctgac	cccgaccag	accctgacag	cttccccacc	tcaaccctt	ccccaccac	7740
tgacaccatc	atcacccaaa	cagtatgcag	aatctccccg	atccgtgcac	agcacacaga	7800
gacggctgca	ggtgacacaa	cacacacgca	cacaggctgc	aggtgcacac	agacacacgc	7860
acacaggctg	caggtgcaca	cagacacaca	cacaggctgc	aggtgcacac	agacacacgc	7920
acacaggctg	caggtgcaca	cagacacacg	cacacaggct	gcaggtgcac	acagacacgc	7980
acacaggctg	caggtgcaca	cagacacgca	cacaggctgc	aggtgcacac	agacgcacgc	8040
acacacacag	gctgcagggtg	catacagatg	cacacgcaca	caggctgcag	gtgcacacag	8100
acacacgcac	acaggctgta	ggtgcacaca	gacacgcaca	cacacaggct	gcaggtgcac	8160
acagacacag	gcacacaggc	tgacaggtgca	cacagatgca	caggcacaca	ggctgcagggt	8220
gcacacagac	acacgcacac	aggctgcagg	tgacacacaga	cacacgcaca	caggctgcag	8280
gtgcacacag	acacacacac	aggttgcagg	tgacacacaga	cgcatacgca	cacaggctgc	8340
aggtgcacac	agacacacac	acaggctgca	ggtgcgcaca	gatgcccagg	cacacaggct	8400
gcaggtgcac	acagacacac	gcacacaggc	tgacaggtgca	cacagacacg	cacacagggt	8460
gcaggtgcac	acagatgcac	gcacataggc	tgacaggtgta	cacagacatg	cacacagggt	8520
gcaggtgcac	acagatgcac	acgcacacag	gctgcagggtg	tacacagaca	cacgcacaca	8580
ggctgcagggt	gcacacagac	acgcacacac	acacacaggc	tgacaggtgca	cacagacaca	8640
cacaggctgc	atgtctacac	agacacacac	acatacatgc	acataggctg	acagacgcac	8700
ctcgaaggac	gctccccggt	tccacctgca	tgtgctctga	actccccctgt	cctccccgtgg	8760
ccccattccc	cacgcatagca	cttgctggc	ctcatcaca	cccaccatct	ctgggcatgc	8820
ccaggccccct	tgccacaaac	cctctgcccgc	tgccctccag	gcttccccca	accctgcacc	8880
cgtgaggact	gaggttgggt	gtctgtgacg	ggctgagccc	gaggtcccag	cacagtggcc	8940
agcctcgctt	caggggttgc	ctaggccctg	ccaaccagc	caccctccac	aggtgtttgg	9000
caccagcctt	gaggagcaca	ctgaggccat	cgagcgagtg	aggaaggcca	aggtgaggcc	9060
gccactgcct	gggcaggcac	tgaccagcac	gtgcctgcgt	ggcgggggca	cggggcaagg	9120
ccctgtcacc	tccatgctaa	gtagggcatac	tggggggcag	ctcagcctca	ccaggccccc	9180
cacgccccctc	tgtccacagg	ccccacgta	tgccctgaaa	gtctctgtca	tgctgtccaa	9240
gaaccttctg	gccaaaggacc	ccaacggtga	gtggggaccc	agcagacctc	ggctgcgcgg	9300
aggcctggcg	gccatggagg	gccagagggc	gtggggccag	ggctgagacg	cctgccgtgc	9360
ccaggcttca	gcgaccata	ctgcatgctg	ggcatcctgc	ctgcctcgga	cgccacgcgg	9420
gagccccgctg	cacagaagga	gcagcgcttc	ggcttccgca	agggcagcaa	gcgcggtgga	9480
cccctgcctg	ccaagtgcac	ccaggctacc	gaggtgaaga	gcagcaccct	gaaccccgctc	9540
tggaaaggagc	acttctctctt	gtgaggccct	cgcccgctctg	ggtgcgggag	gggggctctg	9600
cctgggggtcc	agagctgggc	tgtgcctgga	gggtgggtggc	tcgcagcagc	tcccggctctc	9660
tttggggccg	tgtgcccact	ttccagagga	gaaaactgag	gcccagagat	gggcgtttgtg	9720
ccaccggggc	cacacggctg	gcaacacctg	gcgggatcag	gatttgacc	tgaacagctg	9780
gtctggagag	cgggtgcttg	tggcctggct	cctgctgctt	ctggactccc	tgaccccgcc	9840
ctgggcagag	cctgcgcggg	gaccctcatg	ggcatctctg	ttctccttgt	agcgagattg	9900
aggatgtgag	cacggaccag	ctgcacctgg	acatctggta	caggccctg	cgccatgcag	9960
ggcggcgggg	atgcacacct	tttgcgggg	agcaggacgg	actcagggcc	cggctgcggg	10020
ctccagacaa	gccaggctct	cccctgtccc	caccagggat	catgacgacg	atgtatccct	10080
ggtagaagcg	gtgagaagc	tgaatgaagt	cactggcctg	aagggcagtg	gggatacgtg	10140
ctgctgggac	ctttctaccc	atcacccgtg	ccctcagctt	ggtcatgggg	cccggcgggc	10200
agtggctccc	caggagccat	cagcgtgatg	ctcaccacct	gtggtggcca	ggtacttcaa	10260
acagatcgtc	aagtcagccc	gcgcaaacgg	gacagcagga	cccaccgagg	accacaccga	10320
tgacttctctg	gggtgcctca	acatacctgt	ccgggtgagt	gggtgtgggg	tggggctggt	10380
gacacttaga	gaagggcctg	gggagtgtct	cagccttcaa	ggggcagggc	actgcctgca	10440
gaagctcctg	cttttgggccc	tccactgagt	gcgcttgctc	ccctcattcc	ccggagaagg	10500
gaagcccagg	ccatgtggctc	tctccagggt	ctcacagagt	cagtgtggaa	ggggaggggg	10560
aggagtctgg	tgggcctgac	ccccatgcta	ccccaggag	gtgcctgtgg	ctggcgctga	10620
ccgctggttc	aagctggagc	cacgctccag	tgccctcgct	gtgcagggac	actgccacct	10680
ggttctcaag	ctgatcacta	cgcagggtgg	gaaagtgggc	gtccccgtcc	tccacccccg	10740
tccttggggc	gggggcagag	ctgggtgact	gggtggctct	gcgttgacga	gggatacgtg	10800
catgagccag	cgcgggcgat	ccggcttcc	gtcccacctg	ctgctgctca	gccatctgct	10860
gcggttggag	cactcagcag	aggaggtagt	gggtgcgcct	aggattggaa	cagcccgcca	10920
gggggcctga	ggaggggatg	tggggctgtc	ccgcctccac	ccacgacaga	cctcctcccc	10980
cagcccaact	ccagagctg	gcgaggagag	ctcagcacac	cagccgccac	catcctctgc	11040
ctgcacggag	cccagagcaa	cctgtcaccc	ttgcagctgg	ccgtgctgtg	agtgggtgga	11100
gctacgagtg	ggcggggaat	gtgggcgggc	gtgagcagtg	ggggcatcgg	ggggcatcgg	11160
gtgagtggtg	gtcgtggctg	ggaggggtgg	gcagggaaag	gggcgggtgct	gagtaggtgg	11220
ggctgcgctg	attgggcggg	aagggaaggg	ggcgggtgcta	cgggtagggtg	aggcagcgaa	11280
aggggcgggtg	ccatgagtag	gtacggcagc	gctaattgga	gggcagggaa	aggggcagtg	11340
ctatgagtag	gcgcggcagtg	gctgattggg	tggggcagag	aaaggggtag	tgctgtaggc	11400
ggtgtctagtg	gggcatggca	ggggcggggt	tcataccctt	tgacccatggg	ccgggccccca	11460
caggcacttg	agggtcagca	gccgccacca	tcaaacctgc	acgctggact	acagctacct	11520
gctggggctg	ctggaggaca	tgcaggcaca	ctgggaagag	gctccttcac	tgccccagga	11580

gcaggtgggt	gcagccggga	ccttcttgcc	agccatggcg	aagggagtgc	tggggactgg	11640
ggtcgctcag	tgccggtcgg	cgagggggcag	agtcctgccc	gcgtgggggt	caaagacaaa	11700
tggacaaact	gtatgacgtg	ggcgagagcc	agcgctctgc	agaaaggaga	cgggagactc	11760
gtgggagtc	cctcggggg	cgccaggac	ttaggacagg	gctggggaaa	gccgccaagg	11820
tggcgcccat	ctgagcgaca	ggagggagcg	gccctggcag	gacggacgtg	ggaactgcag	11880
gggcacaggc	cctacgtgag	ctgcgtgggt	ggaaaccgag	gctgggacag	gccgtctggg	11940
acaggccgac	tgggctgtg	gaggggtgctg	ctcagagcag	agctgggctg	accatgcccg	12000
gggaagggtg	ttgcgggcca	aggcaggag	gatgtttctc	ctcatcagtg	gccggtcggg	12060
gctgctggca	ctgaagggcc	ctgtccccac	aggaggagag	cctggctgat	agcctttccg	12120
ccttctctga	gttcgggctg	cagctgctgc	gccagctccg	agactacttc	cctgccacca	12180
acagcaccgc	tgtccaccgc	ctggagctgc	tgctgaagtg	ggtgcagcgc	cgctgtcag	12240
cgtgggtggg	ggcggctggc	agggcaggcc	ccacgtcagc	gtgctgcccc	caccaggtg	12300
tctgggcaag	ctgcagctct	tccaaccctc	ccttgagatc	tgcccccttcg	agtcggagct	12360
gaacatggac	attgctgcgg	ccctgaagg	ggtttccaaa	gcccagtgga	ggcggctgc	12420
taagccctcc	cttccccctc	cctgccttc	cgccccctgc	acctctgcac	cacggctctc	12480
tgccacctc	ttcctcttc	ctgcaatcca	ccatgctgtc	ttggtggtgt	ctgttgaga	12540
gaggcaaccg	tgagtgttac	gacaggatcc	tgaatgacaa	gagtccccga	gagcaggtgc	12600
agttgtgggg	gaccctggcc	atgaggggta	cgggcagccg	agaggtgggt	acgagctagg	12660
cttccttgca	gccaggacca	cagcgctgc	ctgggctggg	tgtgctggct	gacgctgtc	12720
atgataccct	tcagtgtctg	tacagtgtgt	accgcagcct	cttccacagg	tgggctggc	12780
ccctcagctg	tgtcctgggg	ctggggtcgg	agccaacagg	gcctgcaggt	ggggcgagg	12840
tggatgaggt	gtgtcccttg	ctctgcagca	tcctcaatgt	ggacgtcttc	accctgacct	12900
tccggcagct	ggagcgtctg	gtgaggagg	tccctgaccc	cgggtgcctg	ccaggcatgg	12960
ggaagtgggc	agatgccctt	ggctaggacg	gtcctgggat	tccttgctgc	tggaggcctg	13020
atcctgctgt	cccggacagg	tggctgagg	ggcgtgggtg	ctgacggagg	agctgagccc	13080
caagatgacc	ctggagggtg	cctcggggct	ctttgagctc	tacctgacct	tggctgacct	13140
ccagcgcttc	tgggatagca	tccctggccg	gtgggtgccc	cgccccctatc	tcttgacagc	13200
agactttggg	accagggccc	caggacctgt	gagggggacg	gtcctggagc	tgtgagcctg	13260
gaacgtgctg	gttaagggtg	ctgagtcagg	gccgagccct	cacaaccctc	ccaccacagg	13320
gacagccgct	ctctggccct	ggctggcatc	cacgccccct	tcctgcctgc	tgtgaagctc	13380
tggttccaag	tgctgaggga	ccaggccaag	tggaggcttc	agggagccgt	ggacatggac	13440
acgggtgacag	ctgccctggc	ctgaggacac	tggggctggg	ccaggccccc	agcctcccc	13500
gcctccccag	caaccccagc	ctccccagca	acccccagcc	tcctccgcct	ccccagcctc	13560
cccctcaacc	ccagcctccc	cggcctcctc	cgcaacccca	gcctccccag	caaccccagc	13620
ctccccagca	acccccagct	cctcagcaac	cccagcctcc	cccgaacccc	cagcctccctc	13680
agcaacccca	gcctccccca	caaccccagc	ctccccggcc	tcctccgcaa	ccccagcctc	13740
ccccacaacc	cccgccctcc	ccacctcccc	agcctcccct	gcctccccag	caaacccagc	13800
ctccccctgc	tccctagctg	gagcccgtgg	acgcctcctc	caggcacagc	agctccgcag	13860
ccactgctgg	tctctgcctc	agccacatcc	aggagtgtg	ggtgctgctg	gcgtggcctg	13920
accctgccc	ggctcagggg	ctgggcaccc	agcttgcca	ggtgtgtggg	tgggcccctg	13980
gggtgagggg	aacgggtggg	agaggagatg	ggcgagggga	gggctccatg	gctccccacc	14040
ggcatccct	cctcaggacg	tgtgtgaggc	caccctcttc	tatacggagc	tgcttcggaa	14100
gaaggtggac	actcagccag	gggcggccgg	tgaagcagtg	agcgaggcgg	tgagtgacca	14160
gctggggagg	ggagccggca	ggaggtgggg	ggccatcacc	aggccccgga	cccatcgctg	14220
cctggccaca	cctcctccag	ctctgcgtgg	tcctcaacaa	tgtggagctc	gtgcgcaagg	14280
ctgctgggga	ggccttgaag	ggcctggcat	ggccagaggg	ggccacgggg	cccagggggg	14340
tgtcccccg	ccctctgctc	agctgcacac	aggccctgga	cgatgatctg	caacgggagg	14400
cccacacggt	gacagcgac	ctgacctcta	agggtgggtg	ggcctggaga	ccaaggcggtg	14460
gaggcagctc	ctggtgcccc	ctgcccgtgc	tgagcactgc	tcctgcccctc	agatggtggg	14520
cgacatccgc	aagtatgtac	agcacatcag	tctctgcctc	gactccatcc	agaacgatga	14580
ggtgagtgcc	ggggcgaggg	caccgtggag	actgtgtgta	ctgggggtag	gtcagaggtg	14640
cggggtaagc	ctggcctgac	gccccctgcc	cgtaggctgt	ggccccgctc	atgaagtacc	14700
tggatgagaa	gctggccctg	ctgaacgcct	cgctgggtgaa	ggggaacctg	agcaggtgcg	14760
ggcgggtggg	gtgggatggg	ctgggctggc	ccgtggctac	tgaggccgcc	ctgcaggggtg	14820
ccggaggccc	tgtgggagct	actcctccag	gccattctgc	agggtggggcc	agcggggcag	14880
gtggggctgg	cccgtggtca	ctgatgctgc	cctgcagggt	gctggaggcc	ctgtgggagc	14940
tactcctcca	ggccattctg	caggcgctgg	gtgcaaaccg	tgacgtctct	gctgatttct	15000
acagccgctt	ccatttcacg	ctggaggtag	agctctgtga	aggagtcctc	cccggccggc	15060
ccgcctcag	ggctgctctg	atccctccct	ttgtccccac	tggcctgcct	ggtctcctgt	15120
ggctggagag	cagccactcc	agggctgtcc	acagcgcttc	ctcccgttaa	tgccgaggtg	15180
tggcctctgg	tgatggcctt	tccccacgct	tcctgggagc	ttccgatgcc	cgccaggggtg	15240
tgagaaaggg	ctctgccctt	ggctgctgag	ccccagtg	tgctggggcc	cctttgtgtc	15300
gggggtctct	acctgtcccc	agcacaagcc	aggctccctg	acacctctgc	cttctttccc	15360
tgccccaggc	cctggtcagt	ttttccacg	cagagggtca	gggtttgccc	ctggagagcc	15420
tgagggatgg	aagctacaag	gtgaggtggg	accctctgtg	gggcgggggt	gtggatgagg	15480

ggactgagtt	caaactgcag	ccccacacgg	gctgtgtcct	tgagcatgag	gtggcccccac	15540
gggcagtcag	gtctccaagt	gccagcgctg	accaccagca	cccagggggcc	cttcctgatg	15600
cgcttcccc	tgacagagct	gaaggaggag	ctgcggtg	acaaatgttc	caccgcgag	15660
tgcatcgagc	agttctacct	ggacaagctc	aaacaggtag	ggaggcgcca	gggacagggt	15720
gctgcctccg	aggctcccag	agcccagctg	cgctcaccgc	cctttccgca	gaggaccctg	15780
gagcagaacc	ggtttgagcg	cctgagcgctc	cgttgccatt	acgaggcggc	tgagcagcgg	15840
ctggccgtgg	agggtgtgca	cgccgcggac	ctgctcccc	tggaagccaa	cggtgagttg	15900
cagcggggag	gggtcggtg	gtggtgggat	gggggcaggg	gggctcacga	ctgtgctcct	15960
gcaggcttaa	gtgacccctt	tgtgatcggtg	gagctggg	caccgcatct	ctttccactg	16020
gtccgcagcc	agaggaccca	ggtgaagacc	cggacgctgc	accctgtata	cgacgaactc	16080
ttctacttgt	gagtgtccta	agccccagcc	ccagccccag	gctccagggt	gccggagcga	16140
gactccccgac	tgccctctgt	ccgcagttcc	gtgcctgccc	aggcgtgccc	ccgccgcgcg	16200
gcctgtgtgt	tgttcaccgt	catggaccac	gactggctgt	ccaccaacga	cttcgctggg	16260
gaggcgggcc	tcggcctagg	tgccgtcact	gggtgctg	ggccccagg	gggccccggg	16320
gcaagggctg	ggcagcctgt	caccctgcac	ctgtgcccgc	ccagagccca	gggtgagtga	16380
gcatctgggt	ggaggcaggg	gcagcggg	tgaccccg	ccccacacac	ctgcccgcgc	16440
tgcttgacgt	gagatctgcg	ctgaggaggc	tggaaggccg	caccagcaag	gaggcgag	16500
agttcgtgaa	gaaactcaag	gagctggaga	agtgcattga	ggcggacccc	tgagtccatc	16560
agctgccagc	cccggccctg	gccccacccc	gagttccct	gaagcatcct	ccagctcact	16620
gtggccagct	ttgtgcaacc	agggccccc	gcgcccctc	tgtgtgtgta	cgtgtgtgtc	16680
gtggctggcc	ccgcggcgcc	taccgccc	gccgtgtctg	tctggtgtgt	gctgtgaacc	16740
cctgcaccca	acccacatc	tggtgtggca	acttggcagg	acttggccag	cagctgcccc	16800
ggacacagtg	caggccagag	cggtgtgac	cacctgggtg	gcctccctgc	ccgcttcctt	16860
gggctcccc	gccctgggtg	ggcgtgtggc	agctggtctc	cagggaactca	gtgagtggtc	16920
gtgctctctg	cacaacgggc	aatgtgcaga	cgattttttg	gtaatcacag	ctggggagtg	16980
aaaaggggtg	cactggcacc	actgggtgga	tggtccagag	cctccaccca	cagaggggat	17040
gcaaagggca	ggtgagtcaa	gaaccgcata	ggtctccagt	ccccacgggg	ctcccaggcc	17100
ggggaaaagg	tcccctgagg	tactctgag	gccagggacg	tcacccaagg	ctggtggtca	17160
gtgtgaagg	ctccgtgcca	actggtcagc	tgtccttcac	gcacatatcc	gtggccacct	17220
gagacctgtc	ccacgaccct	tccaggcaga	cccagagagt	cgcccccaacc	cttccccagg	17280
cccagtgtga	aaaacagact	cacaaggggc	ttcttggcct	gcagcttcat	ttgcgagagc	17340
gccgaggcag	gacacagagc	acagctgtgc	tggaagtgtg	gggagaaccc	ggacagctca	17400
gtcctgccag	cagccgcaaa	gagccgaggc	tgccaggccc	atztatgtcc	ctcatgtctc	17460
tagattttct	cgtcacccag	cctcaaaaat	atatgtgtct	gcaaccctca	gtctctctgc	17520
cgtttcttga	ttcctttcca	aacctcagcc	ggggccctgg	cctcagcccc	ccgtgccgtc	17580
gtctgtcctt	cttcacagca	gctgctgtg	gctcctccgc	gctcggcgcc	aggccctggg	17640
tcctcagacg	catcactgtc	atcagtgtcc	gagggcagcc	tgggagagga	gggggagcac	17700
gttcccaaat	gacgtcctcc	ctggctagct	ggagtcaacg	tcatccacaa	aaagccccc	17760
gccgaggaca	gattcctctt	cctccctgtg	ctctccacac	ccaaacctct	ccagtgggac	17820
atgcactgca	ggttttctag	tcagaagacc	atcgggggtg	tgttttacaac	accaccccca	17880
ggcagcagca	agggctcagg	ccagaggtag	aagagagccc	tggttctgag	tgatcatgaa	17940
attaagccca	aggacctacc	gggtgtgtgg	cacaggcctg	ttgggggtttc	caaactctct	18000
ccctgaatcc	acatcgaagg	gatctgagcc	gagagaggaa	agtggcctct	aagtgaagctc	18060
agagcaggac	caggcctccc	accccccagg	ctttggaggg	ccttgggttc	caccccgcga	18120
aggctcaccg	atctcctcct	cctgggggct	ctccttgga	ttggccaaga	actcctgtct	18180
cgctgtcagc	acctcctccg	ggctgcccga	ggccgcgtac	ttgtccagga	gctggccggtt	18240
caggtccaag	aagcccctgc	cccatttaaa	cttccgcagc	aaaatgacag	caaggctctgg	18300
aaagcctgac	ggtgtgagaa	acaggaaacc	caaagaagtc	gacggtcccc	tcaagtgcct	18360
cctgatcacc	ctgctgccgg	atggcggcgc	tctcccaagt	aaccgcaggg	tccccagcaa	18420
acatcagctc	acagtccctga	gaaccctgtc	tgcccccagc	ctgcacctcc	cttgctacag	18480
tccacacccc	tgcaaccgac	caagtcccc	tcctttgtctg	cgaaacaaac	attccccaca	18540
agactgaccc	agggcctcatc	agggcctcct	tagatcagca	attcagtaag	gaggacgtgg	18600
aggataagga	ggattctctg	ggcagtgggtg	ccctgataat	gaatggcacc	ccacacaccg	18660
gccctgatag	aacagcgatc	ccattctgcc	tctactgcaa	ctccacaggc	cgaaaacctg	18720
aagtgcaggt	tggggctgga	accgtggtct	gtctgaactg	ttcccctgtc	cgctcccttat	18780
gctcctcctg	gggtctgtcg	gcaggaaaca	tgatgtctgg	agcatagcag	gaacagcagg	18840
ccgccgggtc	acactgctgt	ggggccctg	gactcaccta	cgatgcagaa	ggtggcgagca	18900
aacgcttcca	cgcaggaaaag	tctgtagggc	cggccatagt	tcacgggggtt	ggcggccacc	18960
aggtagggca	acaggcgcaa	gtggctccct	cgcactctcc	caaacgggtg	ctcgtccagc	19020
ctggcccagg	agcagtcgat	gacggcgacc	ccagactgcg	ccaccagctg	tctgccaggg	19080
acagcatggg	ataagttcag	gagccagcac	tacccaaggt	ggagcatgcc	ccggagaatg	19140
gcctagctaa	gggattcagg	gacatcattc	cgccatcccc	aaacaccctc	aaacggagct	19200
cacgagcaag	cttttgaggc	agctgcttgg	ggcacttgga	gcatgtgttt	attcactcaa	19260
caaacgttta	ttaccccaaa	acctccttta	tcccaggctc	tgcgctacgc	aacggaggca	19320
aactagacag	acgttccctg	cccacgtggg	tgtggatgtg	ggtgcgacat	acaagtgcc	19380

ctaaaccatc	cctgaaggga	accacgcgct	ggaaatgaag	acgacctcgg	gcaggcctcc	19440
cgcgctacc	tgtctgcggg	ggacgcgtac	tgtctgcccc	cggggctcag	caccagaccg	19500
ccgaatctgt	ggcccaggcg	caggcagcgc	accagcccca	ggcgggccaag	cttgcgggcc	19560
gtgcagcgcc	gggggtcgca	gtggcccaac	tcccacatgg	ccagcgtgca	gggcagcgcc	19620
gccgggccc	ggccgccctc	gccgtcagcc	gcccctggct	ccacggaagc	tgcacgagag	19680
agagaagggc	actcggcctc	agcgtgggac	ccccgggcag	gccctcaaac	cctgaccgcg	19740
tctccccatc	acgcctcgct	caccctgcag	cgcggcgcgc	acctcctcgg	cgaaggcctc	19800
cagggagcgc	gtcgggaggt	gccgagggcg	gcccgccttc	gcccccgcc	cgcgcgctgc	19860
cctcctgcgg	cccatggcgc	ggacctgggg	tgccggggac	tccccacccc	acggccgcgc	19920
ccctcggcct	cccaatgggc	tgtgcggctg	ccagcgcttg	cgccggacgc	cacttcgggt	19980
ccccgtggtc	acgtgaccgt	cacttcacgt	gaccgcgcgc	cgcccgctgc	ggcgcgatgg	20040
cggcggggct	ggcgcggtc	ctgttgctcc	tccgggtctc	ggccggcggt	gagtggcccc	20100
gccgtccgcg	tccccaggcc	ccgcgcgcgc	tctgcacccc	ggcctccccg	ctcaccgtct	20160
cgctccccgt	agggcccgcg	ccggcaggtg	cagcgaagat	gaagggtggtg	gaggagccca	20220
acgcgttttg	gtgagcagcc	tcgcgggctg	gcggctcgag	cgggggacgg	cccgggcccc	20280
ttccccgctg	accttgccgc	ttcccgtagg	gtgaacaacc	cgttcttgcc	tcaggccagt	20340
cgctccagg	ccaagaggga	tccttcaccc	gtgtctgggtg	agtgaggagc	gctgttggtc	20400
ggcgcggggg	tggccgcggg	gagcagcaac	agtggaggat	ggatgagtga	ccctgagagt	20460
tacggccccg	agtctaaagc	attggtttgt	cagccttcta	aaattaatgt	ggaagctacga	20520
gagaatatac	cagaacgttc	ttgtagcgaa	ttaaattatc	tccagactta	gtccttcctt	20580
tcctttccga	gaggtaactg	ctgttaatag	tctggtgact	ttgcagacgc	gtttctgtgc	20640
ttttacgtgc	gtgtctatgc	aactggaaaa	cgtatagttt	ttgtaatcta	atcgcttttt	20700
taaaaaatca	tgctgcgcga	gtatttggct	gatttgtcgt	tgtcactcgc	tgtatctctc	20760
gggacgcttc	atggtgatga	tataggtctt	cgctctttct	cttgggtgctc	agcacgcgat	20820
aggatggcac	tgcactcccc	gattgtattt	aggttcttgc	catttttttg	gccattagaa	20880
atagttctgt	ggagaacatg	ctgtgtaggt	gctttgggtt	gccctcaggt	gagcgggttc	20940
ccagtgcagg	attccagaag	tgggaattgg	gggtcatgag	tatatatatc	atttaaaccg	21000
tgaatgagta	ttaaatataa	tagacaaaaa	tccaagaatt	tttttttttt	tttttttgag	21060
acaagagtct	cgctctgtcg	cccaggctgg	agtgcagtgg	cacgatctcg	gctcactgca	21120
acctccgcct	cccaggttca	agcgatccct	atgcctcagc	ctcccagagta	gctgggatca	21180
caggcatgca	tcaccatcac	gcccagcaaa	tttttgtatt	tttagtagag	ataggggtta	21240
tttcaccgta	ttgaccaggc	gggtctcgaa	ctcctgacct	caagtgatcc	gcccactctc	21300
gcctccaaaa	gtgctgggat	tataggcgtg	agccaccgcg	cccagcttta	gtaattttat	21360
tatcgaatac	ctttatgtgt	tagggctctg	gctgggggtg	cagaaacgaa	cacagatata	21420
ccaatgcttg	tcctcatgga	gctgacagtc	taattggggag	ggatggagga	acagcagctg	21480
agccttagga	atacatagaa	atagcagtaa	tgcccaatga	tgaaaaatgc	catgtagatg	21540
tgtgaagcag	aggagggagt	gtgtgtatgt	gtgttagagt	attgtttggg	ggatcgcttt	21600
ttttattttt	tagagacggg	gtcttgctct	ctcactcggg	ctggagtgcg	gtagtgcagt	21660
catagctcac	tgcactcttg	aactcctggg	ctgaagcaat	ccaccaacct	cagcttcctg	21720
agtagctggg	actacagggt	cacgccacca	tgcccggcaa	atgttaagtg	ttcttgtaga	21780
gatggaagtc	tcactatggt	gcccaggctg	gcccacaact	cccagcgtca	agcaatcttc	21840
ctgtcttggt	ggggattgct	ttaaataggg	tgctgaggaa	aagtctaaga	agtgcatttt	21900
cagccaaaag	ctaaaggaga	aaaggagta	aagctgtggc	taattggggg	aagggatcca	21960
ggaaagggag	tggccaaagg	ccttgagcag	gacagccagg	gccagagggg	attcagctgg	22020
ggtagccaca	gatcagagat	aaaggggagc	agatggtgcc	gcaccgtgtt	ggtagggttg	22080
gtggactttg	actccaaaca	catcttttta	agttttttgt	tgtgtggctg	ggcgtagtgg	22140
cgcacctggg	gttaggctca	cttcctgcgc	gctcagaagc	catgggatgg	tggcttccca	22200
gggcctgaag	tgctggagtc	gggccaccac	ctatcccaag	aagctgggga	ggaggcaggc	22260
cactgtgtaa	cagcgcagtt	cccagaccga	ccccttcag	taggagagag	gcagccttgg	22320
ctaattgggac	cacgtctttt	tggaggtgag	agtcctggaa	agaaaaacaga	atatgtctgt	22380
cgtggtgggt	cgtgcatggt	gctcgggagc	ctgaggtgtg	acgatcgctt	gaggtcagga	22440
gttcgagacc	agcctggcca	acatggtgaa	accggtcttc	tactaaaaaa	aaccacaaaa	22500
acaaaagtta	gccgggcata	gtggtatcct	cctgtaatcc	cagctgcttg	ggaggccgag	22560
gctggagaat	tgcttgaatc	caggaggcgg	aggctgcggt	gagctgagat	catgccccctg	22620
cactccagcc	tgggtgacag	agcaagactt	cgctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	22680
cccataaaaa	tttagaaaag	aatacctctt	ctagttttcc	cacccccatc	cccacatcag	22740
ggccagtcag	gttcacatgt	gcatttggtt	cattatgtgt	ctttagtgtt	ttaatttgga	22800
acagttctct	tcccactttt	ggatgtaagt	ggcacatttg	tttcccagtt	tcctgatctt	22860
tttgttttgt	tgatggtgtt	tttctccaga	ccattttttc	ctgcgcctgt	ggggtcaggc	22920
gtggatagtc	ccccgcgtct	gtagggctgg	gcgtggatac	tccccgcgt	ctgtagggcc	22980
gggcgtggat	tgtctccac	gtctgtgggg	tggggccggg	cgtggaaagt	ccccgcgtc	23040
tgtgggggtg	ggcctggata	ctccccgcga	tctgcggggt	ggggcggtga	tcgtctccc	23100
catgtgcggg	gtcgggtgtg	gatggtctgc	cgcgctcgcg	gggtcgggtg	tggtgtggtc	23160
cccgcgtctg	cggggtcggg	tgtggatggg	ctccgccttc	tgccgggtcg	ggtgtggatg	23220
gtctcccgcg	tctgtggggg	cgggtgtgga	tcgtctcccc	cgctgtggg	gtcgggtgtg	23280

gatacctctcc	tgctgtctgtg	gggtcgggtg	tggatattct	acttgttttt	tctgtctttg	23340
atgtagcgct	ccttcaacat	caacgtgtga	ggcggattgt	gtgcagtggg	taggagagtg	23400
tggctgtgtg	aatgagctga	gcagaaagat	aatcaggctg	agaaactcag	gggcgtggtg	23460
attgtggcct	agggctcaga	aggtccccgc	agaagagtgt	ctgccaagga	gaaggaaggg	23520
gtgggcttgc	cagggccgag	gctggtgggg	ctgggctggg	cagtgggcag	ctccatggcc	23580
tccaccagat	gagtcgggag	cctcgtccct	tgacagcagg	aggagcagag	gttggtaggt	23640
ctgtgatattg	gagacagtgt	cccgtctgtg	catggagagg	gaatggggca	cggtcagtgt	23700
ggtgttcctg	ccagcgccca	gcagagtcca	gcatagccca	gcacactgtc	caggaggag	23760
ctctgcctcg	gtggcggtgg	gagggagcgg	ccagctgggc	aggactctgg	ggccttggca	23820
gggtccttga	tctggggagg	gggcggggga	ggagtagagt	ggccccggga	tcctggctag	23880
agaagccatg	gcaagtgcgg	ctcccagcat	gtgagttgca	ggggtcaggc	tgggagcttg	23940
ctttcggctg	gggccgggct	gggtctgcca	ggcagtcctt	ggtgcttggg	gagggctacg	24000
gtgggcacag	gactcaccag	ggggtgcgtg	aggccggagg	gtgaggggtg	ggtccccggt	24060
gaagtctctg	gtggcggtgtg	ggggttgtgg	aggcttggag	ggtactggga	ccaggtgtgg	24120
ggccgggcca	gggagcagcg	tgggagggag	gaccacctt	gagaagggat	tgtaccaag	24180
agacgcagcc	ctagggggag	tgtgtccac	cgaagcctga	ccggggcaga	ggggaaacgg	24240
agtccaggct	cagccaagct	gatgggcctt	gagctgtggc	ctcgggtgca	gggacagctg	24300
gcagtaaggc	cacctggggc	cgcgggagcc	aagggcaccg	gcgagtgtcc	tggccatgca	24360
ggcgtgaac	acggagcctg	caggggacgt	gcgggtgaag	caggagcaga	ggagaaggcct	24420
cccgcgagg	gtgggtggcat	ttgggggtgg	tggatatttg	gggtagagac	ccccgtcag	24480
gtatggctgt	gtctggccca	agaggcgtgg	aggggccagc	ccccacctgt	ctcaggagga	24540
caggctgtctg	aaacctgagg	gccaggcggg	ggggaggggc	cgtggccggg	gttctgggtg	24600
gggccagcag	cagaggacct	ggggtagctg	gcactgccag	ttgggggctt	gtggtacct	24660
cccctgcggg	aggctgcacc	tcagccagta	ggagtccact	gaagctgtgg	acagaccaca	24720
ggcagaggcc	tcctcccagc	tggggagcca	ggccaggccc	cgctgccctc	acacagctct	24780
gtttagaccg	atggccctgc	aggtgcagag	ctaagcccgg	cgcccacgcc	cctcctgggg	24840
tgtctccaag	cagtgtctgc	ctggggcaag	cctgggaacg	tgagttccca	ggacacagga	24900
gcttgaccag	ccaccctccc	agcgggagtg	tgagagcggg	aatctgcggg	cgggggtctg	24960
cgggcgggag	tgtgccccg	agagtgtgcg	ggcgggggtc	tgccggcggg	agtgtgcggg	25020
tggcggtctg	cgggcaggag	tctgagtggc	agatgggtga	gggacacatc	agggtctctg	25080
ggatggctcc	aaaggcttcc	ctatgtgcct	gttgggcagg	gccctagggc	cagactggag	25140
ggaccaagcc	caagcaccac	aggctgtggg	acagcagggt	tgacagggct	gaggtagagc	25200
ccccagcggg	caggctggga	ggacagtggg	tgagcctggg	agtagggggc	accagatgg	25260
tatttgacaa	ggaaggacca	ggacctttct	atatccacag	cccctgtccc	ctgagcaggg	25320
cagcctggcc	atcttggctc	ccatctgtct	tggaggggga	cagccccagg	ggtgagtggg	25380
atgtcaggga	acagaaggcc	aggccgggcc	agcaccagag	aggctctcag	ggcattagcc	25440
tctgcctccc	gggggcgggg	gccaggctga	gggtggcccc	agtgtggcat	ctggctcctg	25500
ggctggccag	attcccaggg	aacagtcacc	cagcacctcg	tgtgtatttc	gactttatta	25560
ttattattat	tattattatt	agagccgagt	gtcgtgtgtg	cgcccagggt	ggagtgtctg	25620
gtcggccagg	ctggagtgtg	gtggcgcgat	ctcggctcag	tgccagctcc	gcctcccggg	25680
ttcacgccat	ttctctgcct	cagcctcctg	agtagctggg	actacaggcg	cccgcaccga	25740
tgcccgctca	attttttgtg	tttttaacag	agacgggggt	tcaccgtgtt	agccaggacg	25800
gtctcgatct	cctgacctcg	tgatctgccc	gcctcggcct	cccacagtgc	tgggatgaca	25860
gtcatgagcc	accgcgcctg	gcctcgactt	tattattaat	acggtaaaac	tacagagagt	25920
ttgaaacatc	agtctggggc	gtcctcctct	aaccagtgtg	cttgctgtctg	cagcgaggc	25980
ccttcagggg	acggggatgg	gggccagtc	cctccctgcc	tccacggccc	aagccccagc	26040
cttcccaca	gtcctggagg	atggcccgtg	gctgacagtc	tcgggtggcg	ctgtgttaag	26100
gtggcgctcc	cagccaggct	ccgactcccc	agcttaaaaga	acagggtggc	tgctgtggct	26160
tgactcacga	gttgccactg	gccttgatgt	gtacacgccc	ttgctctagc	actgcctggg	26220
gctggcctgg	ctggagggtg	gggacgctgc	acgccttccc	tcttccccca	ggcgctgggtg	26280
acactgcctt	gcaaccgtca	gcaccatcat	tccacaact	tgctctgtcc	ccttcaaacc	26340
ctgactcccc	gtgccccgc	tccccagcc	ccggcaccga	ccattctgct	gtctctgaac	26400
ttggctattc	atagacgtca	cataggtgac	tcataactgt	gtgcgacctt	ttgtgtctga	26460
ccaactcctc	ttggcacagt	gtccccaagg	tccccaggc	tggaggctct	cagtgtctcc	26520
tttttatatg	gctgaggact	gttccactgc	aaggacagac	cacgttgttg	tttaccgctc	26580
atctgtccat	ggacactgga	gtggcttcca	ccctctgcct	gggaataggc	tgctgtgaac	26640
ttgggtgtaa	atatttgagt	ctctgcttcc	aatttctgag	agtagctttg	ctggatctta	26700
ttagaattcc	atgttttagct	ttttgaggaa	ccaccaact	gtttgcatag	tggctgcacc	26760
gtttccatt	ccagccagca	ttgcacgagg	ggtccagttt	ctccacattt	tcaccaacac	26820
tgttttttgc	taacagccat	cctagtgggt	gtgaagcggg	accttagggg	tgttttaatt	26880
tgcatttccc	taatgatcag	tcagtgtggg	tcttttcagt	tgcttatggg	gtttgtcctt	26940
tttttttttt	gagatggagt	ctcactctgt	ggccagact	ggagtacagt	ggcgccatct	27000
tggctcactg	caagctctgc	ctcccaggct	caggcgattc	tcctgcctca	gcctctcaag	27060
tagctgggat	tacaggcgtg	agccaccaca	cctggctcat	ttttatttag	tagagacggg	27120
gtttcaccat	gttggccagg	ctggtctcaa	actcctgacc	tcagtatccg	cctgcctcgg	27180

cctcccgaag	tgctgggatt	cccggcttct	ctgcccattt	ttaacttgag	tttattgttg	27240
ggctgtagtt	ctgtgttctg	gatgtgagtc	tcttacatac	gtgggttgcc	agtgtttatt	27300
tcttcccatt	ctgtgggttg	tcttttctct	cccttgatgg	gtttgttttt	ttttccatat	27360
agagtctaac	tctgtagccc	aggctggagt	gcagtagcac	gatcttggtc	cacggcaacc	27420
tctccacctc	ccaggctcaa	ggaatttctc	tgcctcagcc	tcccagtagt	ctgggactgc	27480
aggcatccgc	catcacgcct	ggctaatttt	tttgtacttt	attagagatc	aggtttcacc	27540
atgtttggcca	ggatggtctc	gatctcctga	cttcgtgatc	cgtctgcctc	ggcctcccaa	27600
agcactggga	ttacaggcgt	gagaccctgc	gcccggcctc	tatgccagtc	tttatgccag	27660
tatcacactg	ttttgatcat	aacagctttg	aagtaatttt	tggaaatcagt	atgtacgagt	27720
ctctcaactc	tgtctttaag	atltgttcga	tgatttgagg	tcctctgaga	ctccctatga	27780
acttttaggat	gtgtttttct	atltctgcaa	aaactgtcat	tgggggtttg	atgggggttg	27840
attgaatctg	cagatcactc	tgggtgggtat	tgtcatctta	aaaatattga	atccatcagg	27900
ctgggcgctg	gggctcacgc	ctgtaatccc	agtactttag	gaggctgagg	cagggtgcat	27960
acctgagggt	aggagtttga	gaccatcctg	gccaacatgg	agaaaccccg	tctctacaaa	28020
aatacaaaaa	ttagccagcc	gtgatggcag	gcaccggtaa	tcctagctac	tggggagact	28080
gaggcgggaa	aatcggttga	acccagtggg	ggagatttga	gtgagccaag	atcgtgccac	28140
tgcaatccag	cctggcgaca	atgctagact	ccgtctcaaa	aaaaaaaaag	tggcgtatgg	28200
atgtctcatg	tcagctttat	gtagtgggtg	gcactgtttg	tgctgtgata	attgtgggtg	28260
cttggcacat	tgaccctttt	atcaatgtat	aacatctttg	tctcttaaaa	tagtttcagc	28320
cacttcagct	cttttttggc	tgtttatatg	aatacttttc	cataccttca	ctttcaacct	28380
gtttatgtct	ttaaacttga	agttagtctt	tcaggccggg	catagtggct	cacacctgta	28440
atcccagcac	ttgggaggcc	aagggtggacg	gatcacttca	ggtccacagt	tgcagaccag	28500
cctggccaac	acagtgaaac	cccgcctctg	ctgaagatac	aaaaattagc	tgggcgtggg	28560
ggcgggcacg	tgggatccca	gctactcggg	aggctgaggc	aggggaatca	cttgaaccca	28620
ggaggcgagg	gttgacagtga	gccgagatcg	cgccactgca	tttcaccttg	gggtgacagag	28680
tgagactgtc	tccaaaagaa	aaaaaagaat	ctaattgcctg	atgagctgag	gtggaacagt	28740
ttcatcccca	aaccacccat	ccccaccccc	ggctggtaga	aaaactgcct	tccatgaaac	28800
cagtccctgg	tgccaaaaag	attggggacc	actggtttaa	gtcctgtagc	tttacagacc	28860
atagctagaa	aggcaactgg	tattaattca	ccctgcacga	ggacctccgt	ctgcctcccg	28920
tgagctgctg	tctgctcact	tccccgggtg	gcacaccggc	ctgcatgtaa	ccaactcctg	28980
aagcttttat	ctgggaatgt	ctcttttttt	gggggggtgg	gaagacaggg	tcttgctctg	29040
tcacgcaggc	tggagtgcag	tggtagatc	tcagctcact	gcactctcca	cctcctgggt	29100
tcaagagatt	ctcctgcctc	agcctcctga	gtagctggga	ttacaggtgc	gcgccactac	29160
actcagctca	ttttttctgt	gtgctttttg	tgtagacacg	gggttctcac	agtgttgccc	29220
aggctggtgt	catactcctg	gcctcaagca	atcttcccg	cttggcctcc	caaagtgcct	29280
ggattacagg	catgagccac	gattagcaag	cttaactcta	atltttgaag	ggctatlttt	29340
agaattctcg	gttttgtcag	tttcttccat	tgaatggtac	ctgttttctt	gtttctttga	29400
acgtcttgtg	ctttttgttg	aaaactggtc	cttggccggg	cgcggtggct	caaacctgta	29460
atcccagcac	tttgggaggc	cgaggtgggt	ggatcacagag	gtcaggagat	cgagaccatc	29520
ctggctaacy	cggtgaaacc	ccgtctctac	taaaaatata	aaaaattagc	cgggcatggt	29580
ggcgggcgcc	tgtagtccca	gctacttggg	aggctgaggc	aggagaatgg	cgtagaacatg	29640
ggaggcgagg	cttgaagtga	gccaagatca	gtccactgca	ctccagcctg	gggtgacagag	29700
tgagactcca	tctcaaaaaa	aaagaaaact	ggctctttga	aaacagactc	tgccagtctt	29760
tgcagacagg	ttctgtgctt	ggaccctggg	gatcagtgtg	aggctctctc	caggaccctg	29820
gcactctctc	cgactctcgg	gcaagtgtct	cagcctgggt	gagtccacgt	gagtgcaggg	29880
tgggtgcgag	ggtgggctgg	ggcgagcctt	gcggaccccc	ctcatgccat	ctgtgtcccc	29940
aggtacaagt	atgagttctg	cccgttccac	aacgtgaccc	agcacgagca	gaccttccgc	30000
tggaaacgct	acagtgggat	cctcggtgta	gtggggccgg	ggcagggatc	ccaaagcagc	30060
agcgcagctc	cccacccggc	ctcacgtgcc	gtgcccgtgt	ctccccagca	tctggcacga	30120
gtgggagatc	gccaaacaac	ccttcacggg	catgtggatg	agggacgggt	acgcctgccc	30180
ttcccggagc	cggcagagca	agggtggggc	ctcagcggga	gcccgggaag	aggggcccc	30240
gtctccccaa	cccacttacc	cacgttccct	ccccaggtgg	agctggcggt	tggaaaaagc	30300
aaaccggctg	cccattgtgt	cgagccgagc	acctgcgtct	acgcgctgac	gttcgagacc	30360
cccctcgtct	gccaccccc	cgcttggcta	ggtaggggtg	cgggacgcag	ttgagcccag	30420
tggggctcag	cgcgacgcga	gccctgctgg	aggccctgta	gtgctggggg	ccaggggttg	30480
gacatggggg	gcagctgagc	ctggcttctc	ttgggtcctc	agtgtaccca	accctgccag	30540
aggccctgca	gcggcagtgg	gaccaggtag	agcaggacct	ggccgatgag	ctgatcacc	30600
cccaggtaa	ggtgcgctcg	gggtggcccc	ttgtgggccc	ggctgggagc	tgggtgctgc	30660
ccctgcatcc	tccaccttca	gggcatgag	aagttgctga	ggacactttt	tgaggatgct	30720
ggctacttaa	agaccccaga	agaaaatgaa	cccacccagc	tggagggagg	tcctgacagc	30780
ttgggggttt	agaccctgga	aaactgcagg	aaggtaccgt	attggggggg	ggtgggtggc	30840
cgcagtagcc	ctccagcacc	tgggctttcc	cttgaactct	ttttgtggtt	ggttaggtct	30900
ttaaagaactg	tcaaaggaga	tcaaaaggct	gaaagggttt	ctcaccagc	acggcatccc	30960
ctacacagag	cccacaggtg	agtcacctgt	ggggagaggg	ccaggctcac	catcacactc	31020
gccacctgtg	ggtccaggtg	aggactggcc	acctggtgtt	ttggcagaaa	cttccaactt	31080

ggagcacttg	ggccacgaga	cgcccagagc	caagtctcca	gagcagctgc	ggggtgaccc	31140
aggactgcgt	gggagtttgt	gaccttgttg	tgggagagca	gaggtggacg	cgcccagagc	31200
ccctacagag	aagctggctg	gtaggacccg	cagggaccag	ctgaccaggc	ttgtgctcag	31260
agaagcagac	aaaacaaaga	ttcaagggtt	taattaattc	ccatactgat	aaaaataact	31320
ccatgaattc	tgtaaaccat	tgcataaatg	ctatagtgtg	aaaaaattta	aacaagtgtt	31380
aactttaaac	agttcgctac	aagtaaataa	ttataaatac	taccttctgg	gttaagaaaa	31440
ttccattcaa	ataacattct	catgtaaata	ctcaaacata	cagattgagg	ctccagaaaa	31500
ttaatccact	tgaggcgctc	acgcggaaca	aggtctgctg	accacagtta	cacacgtcgt	31560
gacaccactg	tatcacggcg	aatgtcgaac	actagagtta	cagacgacag	gcaacaagaa	31620
catgcagagc	cggccgcccag	ccagctccta	cgccccgtgc	ccgccatggc	cacagggggg	31680
agctgctggg	cgcgctcca	cctcagcctc	cacggggcac	cttccagcca	ctgctgtgcc	31740
tcagccacct	aggagccatc	cctgaggcca	tggccacca	gacaggtgag	ggaggcccca	31800
gggcacacag	ccccaggatc	ttgcccccat	ttccaaccct	ggcagcgctg	ccagccattc	31860
tttggctctt	ccactgtcta	cctgtcctca	gagcaggtac	tgatctaggg	gagcaggtgc	31920
cacctccatc	tcccagctgt	ccccccacc	ctggggctcc	ccagctctac	caccccgga	31980
gctgagccct	ggagacaccc	tgagggcg	gcgccacctc	cttcccagaa	agagcagagc	32040
agctggcaga	ctagagctca	gctgctcctg	accactgaca	accgggagat	gtctcggcag	32100
acaccactta	tcccagaaaa	gaccagaggc	tccaaatggg	aaccaagtgc	ataaatacaa	32160
ataaaaaacag	agtttacaca	tcgatgacag	tagcacaaaa	tatacgtaat	ttctagatac	32220
tgagcttaata	aatcactgta	gttaaaaaa	aaatagattc	actgaaacca	agtcgatagt	32280
taccttggag	tagtggacta	gctgcagaat	gtcacggacc	cacttctctt	aaaacaacag	32340
aatgttgcta	tcagttgtc	ttacgttaaa	acagttctta	aaccacctac	tatatataat	32400
acattctaata	ttgggtcactg	atgataaaga	tttttaccta	gacgttttgc	acttaaaaaa	32460
tgctattaaa	agtctttggc	aaagccacgg	tcggggagaa	acacgtcggg	gccgttccc	32520
actcgtgtgc	cctcaccacc	cacgtggccc	gcgagagacc	aggtcgccgt	aaaccacaga	32580
tgctgttccc	tgctcaactc	ctagggacag	tcccagccct	gagcagaagg	gtcaggtcat	32640
cgcggggatg	cgttctcagt	cacgtttctg	ctgcgccga	aaccacactg	tcgctccagg	32700
ggcagacggg	acaggggacc	acagcagagc	cgggacctgg	ggggggctgc	ttccccctg	32760
tgctccaggc	cactgagagg	ctgtggccgg	gtcctctcct	gtgacctcac	tggccacatg	32820
gatgagctcc	gatgataact	ggcagcgcg	gtcaggtgac	tgccacatac	acccggcacc	32880
ccggacccta	gcgcacacac	gccccgagca	tctcccacac	gagcaggacc	agggtggtca	32940
cggtgtctgg	ggaggcaacc	actctcccag	ttcactgcct	gacccctgcc	aggacgcca	33000
gcgacatctc	caaggagcta	caaagcacca	ttgcctagga	cagctctgag	taatcagaaa	33060
acagcttttt	tttttttttt	ttgagccaga	gtctcgtctt	gtcaccggg	ctggagtgcg	33120
gtggcgcgat	ctcggtcac	tgcaacctct	gcttcccagg	ttcaagcgat	tctcctgctt	33180
cagcttccc	agtagctggg	actacaggtg	cccgctctca	gccaggcca	atttttgtat	33240
tttactaag	acggggttgc	accatgttgg	ccaggctggg	cttgaactcc	tgaccttagg	33300
tgatccaccc	gcctcccaa	gtgctgggat	cacaggcatg	agccaccgca	cccagccaga	33360
agacagctgc	ttaaaaaagt	aattctaaaa	gctctgacgt	cggatcccc	ctgctgcgtt	33420
ttttgttagt	atcaaaactgt	ctttcagagg	gttaagggag	aaaacaaaac	ccatgatgtc	33480
ccttacttg	ctctccgagg	cgcaggcgca	cagaaacagc	ggccacggac	cacagaaatg	33540
caggacggag	ctctcctgct	cccacgccac	gaggctggaa	catcagcccc	agtgcctggg	33600
gcacacacag	gctcaggaaa	actagctcct	tccatcaagt	taaaaacatc	tcaatccata	33660
atcactctgc	ttataaatac	tatttgcact	taattggaaa	atgtaaaaat	cctaggtctt	33720
actactagat	aaacactcat	ttcataagta	ctagtttcag	aatatgaaat	ctaataatatt	33780
tagcatacta	tgaattgaca	ataaactgat	atgaaatatt	taagttcaca	gtacacatat	33840
aaagcttcat	attatcgcat	aacagaaaag	agcttagagc	agcttagagc	aattttataac	33900
ctctaactaa	attcctcggg	ttacaaagtg	ctttgaaaga	ccccagctcc	cagtggccta	33960
gtcacagtag	tgctcaggacc	acaaactgct	gtgatcacag	cgccgctggg	atcccggagg	34020
caccaggccc	ctcagggaca	ggcaggcggg	ttctgtgggt	ttgcatttaa	ggtttttgag	34080
gaaaatacct	tgaaaccgtc	ggtaggacta	gataggtgac	aacgtgtgac	aggaaaggct	34140
gtcaggccaa	gcgcaggcgg	ggctcccaga	ctatagataa	cgtgtaacag	gaagggctcc	34200
cagcctcggg	cctcacgtcg	gtggcaccag	aagcgagtga	cgacatgtcc	gtggtcagga	34260
ggagcgctgg	agccagggtg	cccagcagga	gggtgctgtc	cccgtgagg	tcaccactgc	34320
aggggctggc	ccttgagta	ggggcactca	gggtgcgggtg	ccgcacacgg	ctcacagagg	34380
atgtgtgtgct	gacagggccg	caggacagca	ccgtggggcc	gtccccggca	ggccacacac	34440
tgcttggcgc	ggagctggaa	gatcacctgc	agggccagaa	caatgacggg	ctcaggaggc	34500
ccctgcccgc	gctgggcccag	ctggggagcc	ggagggcctt	ccgggagga	ctcagcatct	34560
caggccgccc	ggccaccagc	tacagagaca	gggtcaggga	gacgccccag	ccagggccac	34620
ctcaggggaga	aggcagctgc	gggtcagggg	ggtagctctt	gtcttaagag	ccctgggcag	34680
aggacagcct	cctggggcca	ccacagacac	ggcagtgggc	agacaggcca	ggccccctgag	34740
ggcagctggg	tgcccaggat	gtgggctacg	caggaccctc	tgctctatag	cagcgagtcc	34800
aagagcccac	ccttagcagg	cgaggtgagg	aggggacagc	cgaagtggag	gggtcatgtg	34860
tgagggtcac	agctgtgaac	ggggagggat	ggggccacgg	ggagcagtgt	ggcattcgct	34920
gacaagaggg	aggggtgaaa	gggcaagggc	cagcctgaca	ggaaaggcgg	ctcccaggga	34980

```

gacaggcaag gaggccacag ggggctgtgc tgggggtgggg caccgtctcc tcctccctag 35040
gcccggagca gaccctgggg caggggagga aggggacacc gcaccagcca gggccctccc 35100
cagacaggga cagcagccct gcaggcagggc tggccctcca ccctgcccag agcaggactc 35160
acgccgtcca ccgcctccag gtccaggcgc agctgactct gcagcgagtg cagcttcggc 35220
aggggaatgg tgccgatgtc cccacagccc cgcagccccg gcagtgtgga ggctacgccc 35280
aggccctcca gtcctcctg cagctgcttc acctgtgcct ccacctcctc cttcttctgc 35340
agcgccagct gccggtcgct atcgccaca cgggcacgct ccttggcctc ctgcgctct 35400
cgctgccagg caatgcagac ctgaaacca gggcccgctc cagcaccccc cacctacat 35460
gtgccacctg cagaccctct tgcctgtgtc accagcgatc ctcccctccc ctcccagca 35520
tcagctgtgg cccctcacc tgagaccctt gcggccctcc ctcccctcct cctcccctcc 35580
ctcccctccc cgtctcacc cccacacgct cacttgagtc acctgcggcc ctcccctcct 35640
cccctcccat ctcaccccca cagctcacc tgagtcctcc gcggccctcc ctcccctcct 35700
cctgtctggc cccccacac actcacctgc ttcacctgct gccaggactc ctcccactgc 35760
cggatcttcc tcttggcctc gtccagctgc cgctgaccc gggccagctc agctccgttt 35820
ggacttgca tcaagagga tggggggccg gcactcagga tgggggaggg gctgggggag 35880
aagctgccgg aaacaaagt ccagatgtc ccggggacac cgttcaaacc tgagtgtgaa 35940
acggctcgatg acggcccagc cctgctgtgc tcgcggcctg gtgggattgg tgggtgctat 36000
gccccaggcc ccaccgctac gtgggcaccc tctggggctc acctgccctg gcagcagggg 36060
tgacgccagc cctgcccctt gcccactcc tggacccac caggctcagg aagaggcgct 36120
gacagtga cctgccccgg ccggtctccc caccagatct acgccttctt ggccacctga 36180
ggagtctaaa cagagaggaa ttaaattcaa gtctctggga ccacaccatg aggtgggagg 36240
ggcagtggtt ctcagactcc actcccagga accctcaca cactgagact ctcaaagacc 36300
ccacagagct gtgtccatgt ggtgactgcc aatgtctatg cagctttaga gactatgact 36360
gaaatagttt aaatgttact taaaaaaac aggcggggcg cggtggtcga cgctgtaat 36420
cccagcactt tgggaggccg aggcgggtgg atcatgaggt caggagatcg agaccatcct 36480
ggctaacaag gtgaaacccc gtctctacta aaaatacaaa aaattagcca ggcgtggtgg 36540
cgggcgccgt tagtcccagc tactcgggag gctgaggcag gagaatggcg tgaacccggg 36600
aagcggagct tgcagtgagc cgagattgcg ccactgcagt ccgcagtccg gcctgggcca 36660
cagaatgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa aaaaaacagg cctgtcatcc 36720
cagcactttg ggaggtcgag gccagcagat cacttgaggt caggagtttg agaccagtat 36780
ggccaacatg gtgaaactcc atctctacta aaaatacaaa aattagcctg gcgtggtagc 36840
aggtacctgt gatcccagct actcgggagg ctgagacatg acaatcattt taacccggaa 36900
ggcagaggct gcagtgagcc aagattgctg gagcctggac gacagagtga aactttgttt 36960
taaaaaata aaaactttta aaattttgta tttttttgta gactcaggag tctcactatg 37020
gtgcctagcc tggctttgaa cccctggcct caagtgaaca cccacctcga cctcccaaag 37080
tgctgggatt acaggcttga gccactatgt tggcccaaag tattagtttt tttttttttt 37140
tttttttttt ttttgaaatg gagtctagct ctatcgccag gctggattgg agtgcagtgg 37200
cacaatctca gctcactgcg acctccgact cccgggttca agcgattctc ctgccttagc 37260
ctcccaagta gctgggatta caggcatgtg ccaccacgcc cgctaattt ttgtattttt 37320
agtagagacg gggtttcacc atgttgcca ggatggctc gatatcctga cgtcgtgatc 37380
cgcccacctc ggcgcccaca gtgctgggat gacaggcgtg agccaccacg ccgggccttt 37440
ttctcttttt tttttttgag atggagtttc actctt 37476

```

<210> 29
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 29
 gtctacaatt

10

<210> 30
 <211> 293
 <212> PRT
 <213> Homo sapiens

<400> 30
 Met Thr Gly Leu Tyr Glu Leu Val Trp Arg Val Leu His Ala Leu Leu
 1 5 10 15
 Cys Leu His Arg Thr Leu Thr Ser Trp Leu Arg Val Arg Phe Gly Thr
 20 25 30

Trp Asn Trp₃₅ Ile Trp Arg Arg Cys₄₀ Cys Arg Ala Ala Ser₄₅ Ala Ala Val
 Leu Ala₅₀ Pro Leu Gly Phe Thr₅₅ Leu Arg Lys Pro Pro₆₀ Ala Val Gly Arg
 Asn₆₅ Arg Arg His His₇₀ Arg His Pro Arg Gly₇₅ Ser Cys Leu Ala Ala₈₀
 Ala His His Arg Met₈₅ Arg Trp Arg Ala Asp₉₀ Gly Arg Ser Leu Glu₉₅ Lys
 Leu Pro Val His₁₀₀ Met Gly Leu Val Ile₁₀₅ Thr Glu Val Glu₁₁₀ Gln Glu Pro
 Ser Phe Ser₁₁₅ Asp Ile Ala Ser Leu₁₂₀ Val Val Trp Cys Met₁₂₅ Ala Val Gly
 Ile Ser₁₃₀ Tyr Ile Ser Val Tyr₁₃₅ Asp His Gln Gly Ile₁₄₀ Phe Lys Arg Asn
 Asn₁₄₅ Ser Arg Leu Met Asp₁₅₀ Glu Ile Leu Lys Gln₁₅₅ Gln Gln Glu Leu Leu₁₆₀
 Gly Leu Asp Cys₁₆₅ Ser Lys Tyr Ser Pro Glu₁₇₀ Phe Ala Asn Ser Asn₁₇₅ Asp
 Lys Asp Asp Gln₁₈₀ Val Leu Asn Cys His₁₈₅ Leu Ala Val Lys₁₉₀ Val Leu Ser
 Pro Glu Asp₁₉₅ Gly Lys Ala Asp Ile₂₀₀ Val Arg Ala Ala Gln₂₀₅ Asp Phe Cys
 Gln Leu₂₁₀ Val Ala Gln Lys Gln₂₁₅ Lys Arg Pro Thr Asp₂₂₀ Leu Asp Val Asp
 Thr₂₂₅ Leu Ala Ser Leu₂₃₀ Leu Ser Ser Asn Gly Cys₂₃₅ Pro Asp Pro Asp Leu₂₄₀
 Val Leu Lys Phe Gly₂₄₅ Pro Val Asp Ser Thr₂₅₀ Leu Gly Phe Leu₂₅₅ Pro Trp
 His Ile Arg Leu₂₆₀ Thr Glu Ile Val Ser₂₆₅ Leu Pro Ser His Leu₂₇₀ Asn Ile
 Ser Tyr Glu₂₇₅ Asp Phe Phe Ser Ala₂₈₀ Leu Arg Gln Tyr Ala₂₈₅ Ala Cys Glu
 Gln Arg Leu Gly Lys
 290

<210> 31
 <211> 2645
 <212> DNA
 <213> Homo sapiens

<400> 31
 ggcacgaggt ggcggcgggg gcgaggtgag gtgttggcag tggaaagggg ttcgggctcg 60
 gggggcgggg ggacgcggag cgatggcccg cgccggccgc aggggcggat aaaaagccgt 120
 cgcgctgctg gagtgggctg gaggagagg ggggtgtccga gggccacaag agtatgacgg 180
 ggctgtacga gctggtgtgg cgggtgctgc acgcgctgct ctgtctgcac cgcacgctca 240
 cctcctggct ccgcgttcgg ttcggcacct ggaactggat ctggcggcgc tgctgccg 300
 ccgcctctgc cgcggtccta gcgcgctcgc gcttcacgct ccgcaagccc ccggcagtcg 360
 gcaggaaccg ccgtcaccac cggcaccgc gcgggggggc gtgcctggca gccgcacacc 420

```

accggatgcg ctggcgcgcg gacgggtcgtt ccttggagaa gctgcctgtg catatggggc 480
tgggtgatcac cgaggtggag caggaaccca gcttctcggg catcgcgagc ctctgtggtg 540
gggtgatggc cgtgggcatc tcctacatta gcgtctacga ccaccaaggt attttcaaaa 600
gaaataattc cagattgatg gatgaaattt taaaacaaca gcaagaactt ctgggcctag 660
attgttcaaa atactacca gaatttgcaa atagtaatga caaagatgat caagttttaa 720
attgccattt ggcagtgaag gtgctgtctc cggaagatgg aaaagcagat attgtaagag 780
ctgctcagga cttttgccag ttagtagccc agaagcaaaa gagaccaca gatttggatg 840
tagatacggt agccagttaa cttagttaa atggtgtcc tgatcctgat ttagtattga 900
agttcgggtc tgtggacagc acattaggct ttcttccctg gcacatcaga ttgactgaga 960
ttgtctcttt gccttccac ctaaacaatca gttatgagga ctttttctct gcccttcgtc 1020
aatatgcagc ctgtgaacag cgtctgggaa agtagtgggc attggttgca taatttgatt 1080
tgaggcttgt ggaggaaagg aaccaagtga ctctgatgtt tacaaagcac ctatgaaacc 1140
ctgtacacac ctagtccata atcctcataa tttatcaaca aacacaaaaa agtgccttac 1200
ttgagagtga gtgtgtgtgt gtgctgtgtc acgtgcacac atgtgcacgt ttgtatgtat 1260
ggaaataaac ttataaatgg ggacgtattg gagaaggaaa tacatagacc tacaactttg 1320
agcaaatagc agtgatgttt taggaactga aatgtcacac ttaaagtcct cagcccagct 1380
acttccctat ttttgtgggg agaagagggc ctgattagaa ctgttctggt tgtgtttggc 1440
gggaggggaa taatttttgt tcagtccttc ttagtgacca aactttaatt tttagaata 1500
atatattgac ttactgaact gaagcattct gagttgaaag gagctccaga ggagtggagt 1560
tctgtgttgc tcacatgtta aaatcttgct caccttcaga gcagagggaa tacctatctt 1620
cagatatccg tccattttca tctcttaatt gtagtcaaaa gtatgacttg agagtgttgc 1680
tctgggtattc tgggttctga agtctgggtat tctgggtattc tgggttcaaa agtatgactt 1740
gagagtgttg ctctgggtatt ctgagagttg ctctgtattc tgggttctga agattatttg 1800
aaaaataaact cctactacat tgaaatgcag acttaaaaat ttaaaccattg gattaggcag 1860
tcaaaaaaac caagcaagca taaaagggtca ataagttgta atcttgatag taaagggtgga 1920
aaacttatta taaatggaaa gaaagtttta tttccttttt tgtttgatgg gcagtatgcc 1980
atattataacc caaagtctt ttaaaaaata tttccatcaa ccatttttat ttaaaataaa 2040
catttgaggg aagttaccaa ggcagctttt ttctcaaaa gtaacctgtt cctctttgga 2100
atagcacatt ttaggggcat ggtaataacc tgagattttt actcagtaaa tcctgatggt 2160
tactgtgtgt aaaaatctt taagtaggat tgaaggcctc tgtgggggaa taaaatatta 2220
ccaaagtcta taaaaataaa ttttacatgt tctcttttat gacagagagc agcactggtt 2280
ctgttatttt taaaatgaat aattgatttc ttgatagggt ttttaatttt cttccctcac 2340
tgctgattct tagatagaaa ccattcttta tatttgatag actgctttca gaaaaccctt 2400
atcaacaagt gtacaatact tatctaaaac tatacattta gaatggagca gtttaatact 2460
agatctcaga agttttgaaa aatagcaag aagactggat ttggaaagca tgggtctaca 2520
ttggttggtt aattctgaag ctatgaagaa taaatgtttc aactttggat tatgaaacc 2580
catttatgat tttttaaata cacttgaat aaaaatgatt aaactaaaaa aaaaaaaaaa 2640
aaaaa 2645

```

<210> 32
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 32
 gggctcgggg

10

<210> 33
 <211> 323
 <212> PRT
 <213> Homo sapiens

<400> 33
 Met Gln Ala Ala Trp Leu Leu Gly Ala Leu Val Val Pro Gln Leu Leu
 1 5 10 15
 Gly Phe Gly His Gly Ala Arg Gly Ala Glu Arg Glu Trp Glu Gly Gly
 20 25 30
 Trp Gly Gly Ala Gln Glu Glu Arg Glu Arg Glu Ala Leu Met Leu
 35 40 45
 Lys His Leu Gln Glu Ala Leu Gly Leu Pro Ala Gly Arg Gly Asp Glu

50 55 60
 Asn Pro Ala Gly Thr Val Glu Gly Lys Glu Asp Trp Glu Met Glu Glu
 65 70 75 80
 Asp Gln Gly Glu Glu Glu Glu Glu Ala Thr Pro Thr Pro Ser Ser
 85 90 95
 Gly Pro Ser Pro Ser Pro Thr Pro Glu Asp Ile Val Thr Tyr Ile Leu
 100 105 110
 Gly Arg Leu Ala Gly Leu Asp Ala Gly Leu His Gln Leu His Val Arg
 115 120 125
 Leu His Ala Leu Asp Thr Arg Val Val Glu Leu Thr Gln Gly Leu Arg
 130 135 140
 Gln Leu Arg Asn Ala Ala Gly Asp Thr Arg Asp Ala Val Gln Ala Leu
 145 150 155 160
 Gln Glu Ala Gln Gly Arg Ala Glu Arg Glu His Gly Arg Leu Glu Gly
 165 170 175
 Cys Leu Lys Gly Leu Arg Leu Gly His Lys Cys Phe Leu Leu Ser Arg
 180 185 190
 Asp Phe Glu Ala Gln Ala Ala Ala Gln Ala Arg Cys Thr Ala Arg Gly
 195 200 205
 Gly Ser Leu Ala Gln Pro Ala Asp Arg Gln Gln Met Glu Ala Leu Thr
 210 215 220
 Arg Tyr Leu Arg Ala Ala Leu Ala Pro Tyr Asn Trp Pro Val Trp Leu
 225 230 235 240
 Gly Val His Asp Arg Arg Ala Glu Gly Leu Tyr Leu Phe Glu Asn Gly
 245 250 255
 Gln Arg Val Ser Phe Phe Ala Trp His Arg Ser Pro Arg Pro Glu Leu
 260 265 270
 Gly Ala Gln Pro Ser Ala Ser Pro His Pro Leu Ser Pro Asp Gln Pro
 275 280 285
 Asn Gly Gly Thr Leu Glu Asn Cys Val Ala Gln Ala Ser Asp Asp Gly
 290 295 300
 Ser Trp Trp Asp His Asp Cys Gln Arg Arg Leu Tyr Tyr Val Cys Glu
 305 310 315 320
 Phe Pro Phe

<210> 34
 <211> 1407
 <212> DNA
 <213> Homo sapiens

<400> 34
 cgaccaacgg accggacaga gacgaggaga ggaacaggaa gagagaagct gggagaatcg 60
 ggaacctggg ggctagtgc ctgcacacag ggcaggggca ctcggcagtt cccagaggcc 120
 accctccca cccagacat ccagacatct ggaactttgg gtgccaagag tccagcttaa 180
 tgcaggcagc ctggcttttg ggggcttttg tggccccca gctcttgggc tttggccatg 240
 gggctcgggg agcagagagg gagtgggagg gaggctgggg aggtgcccag gaggaggagc 300
 gggagagggg ggccctgatg ctgaagcatc tgcaggaagc cctaggactg cctgctggga 360

```

gggggggatga gaatcctgcc ggaactgttg agggaaaaga ggactgggag atggaggagg 420
accaggggga ggaagaggag gaggaagcaa cgccaacccc atcctccggc cccagcccct 480
ctcccacccc tgaggacatc gtcacttaca tcctgggccg cctggccggc ctggacgcag 540
gcctgcacca gctgcacgtc cgtctgcacg cgttggacac ccgctgggtc gagctgacct 600
aggggctgcg gcagctgcgg aacgcggcag gcgacacccg cgatgccgtg caagccctgc 660
aggaggcgca gggtcgcgcc gagcgcgagc acggccgctt ggaggggtgc ctgaaggggc 720
tgcgcctggg ccacaagtgc ttctgtctct cgcgcgactt cgaagctcag gcggcggcgc 780
aggcgcggtg cacggcgcgg ggcgggagcc tggcgagcc ggagaccgc cagcagatgg 840
aggcgctcac tcggtacctg cgcgcggcgc tcgctcccta caactggccc gtgtggctgg 900
gcgtgcacga tcggcgcgcc gagggcctct acctcttcga aaacggccag egcgtgtcct 960
tcttcgcctg gcatcgctca ccccgccccg agctcggcgc ccagcccagc gcctcgccgc 1020
atccgctcag cccggaccag cccaacggtg gcacgctcga gaactgcgtg gcgcaggcct 1080
ctgacgacgg ctcttggtgg gaccacgact gccagcggcg tctctactac gtctgcgagt 1140
tccccctcta gcggggccgg taccgccct ccttgcccat cccaccaccc ggcctttccc 1200
tgcgccgtgc ccaccctcct ccggaatcgc ccttcccttc ctggccacga atggcagcgt 1260
cctccccgac cccagctctg ggcgttctg ggagggctct tgcggtgccg gcactcctcc 1320
ttgttagtgt ctttccttga aggggcgggc accaggctag gtccggtgcc aataaatcct 1380
tgtggaatct gaaaaaaaaa aaaaaaa 1407

```

<210> 35
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 35
 tgcagggtggc

10

<210> 36
 <211> 204
 <212> PRT
 <213> Homo sapiens

<400> 36
 Met Gly Leu Gly Ala Val Pro Gly Arg Gln Ala Ala Phe Arg Glu Gly
 1 5 10 15
 Leu Glu Gln Ala Val Arg Tyr Ala Lys Ala Leu Gly Cys Pro Arg Ile
 20 25 30
 His Leu Met Ala Gly Arg Val Pro Gln Gly Ala Asp Arg Ile Ala Val
 35 40 45
 Lys Ala Glu Met Glu Ala Val Phe Leu Glu Asn Leu Arg His Ala Ala
 50 55 60
 Gly Val Leu Ala Gln Glu Asp Leu Val Gly Leu Leu Glu Pro Ile Asn
 65 70 75 80
 Thr Arg Ile Thr Asp Pro Gln Tyr Phe Leu Asp Thr Pro Gln Gln Ala
 85 90 95
 Ala Ala Ile Leu Gln Lys Val Gly Arg Pro Asn Leu Gln Leu Gln Met
 100 105 110
 Asp Ile Phe His Trp Gln Ile Met Asp Gly Asn Leu Thr Gly Asn Ile
 115 120 125
 Arg Glu Phe Leu Pro Ile Val Gly His Val Gln Val Ala Gln Val Pro
 130 135 140
 Gly Arg Gly Glu Pro Ser Ser Pro Gly Glu Leu Asn Phe Pro Tyr Leu
 145 150 155 160

<400> 39
Met₁ Leu Val Asp Gly₅ Pro Ser Glu Arg Pro₁₀ Ala Leu Cys Phe Leu₁₅ Leu
Leu Ala Val Ala₂₀ Met Ser Phe Phe Gly₂₅ Ser Ala Leu Ser Ile₃₀ Asp Glu
Thr Arg Ala₃₅ His Leu Leu Leu Lys₄₀ Glu Lys Met Met Arg₄₅ Leu Gly Gly
Arg Leu₅₀ Val Leu Asn Thr Lys₅₅ Glu Glu Leu Ala Asn₆₀ Glu Arg Leu Met
Thr₆₅ Leu Lys Ile Ala₇₀ Glu Met Lys Glu Ala Met₇₅ Arg Thr Leu Ile Phe₈₀

Pro Pro Ser Met His₈₅ Phe Phe Gln Ala Lys₉₀ His Leu Ile Glu Arg₉₅ Ser
 Gln Val Phe Asn₁₀₀ Ile Leu Arg Met₁₀₅ Pro Lys Gly Ala₁₁₀ Ala Leu His
 Leu His Asp₁₁₅ Ile Gly Ile Val Thr₁₂₀ Met Asp Trp Leu Val₁₂₅ Arg Asn Val
 Thr Tyr Arg Pro His Cys His₁₃₅ Ile Cys Phe Thr Pro₁₄₀ Arg Gly Ile Met
 Gln Phe Arg Phe Ala His₁₅₀ Pro Thr Pro Arg Pro₁₅₅ Ser Glu Lys Cys Ser₁₆₀
 Lys Trp Ile Leu Leu₁₆₅ Glu Asp Tyr Arg Lys₁₇₀ Arg Val Gln Asn Val₁₇₅ Thr
 Glu Phe Asp Asp₁₈₀ Ser Leu Leu Arg Asn₁₈₅ Phe Thr Leu Val Thr₁₉₀ Gln His
 Pro Glu Val₁₉₅ Ile Tyr Thr Asn Gln Asn Val Val Trp Ser₂₀₅ Lys Phe Glu
 Thr Ile Phe Phe Thr Ile Ser₂₁₅ Gly Leu Ile His Tyr₂₂₀ Ala Pro Val Phe
 Arg Asp Tyr Val Phe Arg₂₃₀ Ser Met Gln Glu Phe₂₃₅ Tyr Glu Asp Asn Val₂₄₀
 Leu Tyr Met Glu Ile₂₄₅ Arg Ala Arg Leu Leu₂₅₀ Pro Val Tyr Glu Leu Ser₂₅₅
 Gly Glu His His₂₆₀ Asp Glu Glu Trp Ser₂₆₅ Val Lys Thr Tyr Gln Glu Val₂₇₀
 Ala Gln Lys₂₇₅ Phe Val Glu Thr His₂₈₀ Pro Glu Phe Ile Gly₂₈₅ Ile Lys Ile
 Ile Tyr Ser Asp His Arg Ser₂₉₅ Lys Asp Val Ala Val₃₀₀ Ile Ala Glu Ser
 Ile Arg Met Ala Met Gly₃₁₀ Leu Arg Ile Lys Phe₃₁₅ Pro Thr Val Val Ala₃₂₀
 Gly Phe Asp Leu Val₃₂₅ Gly His Glu Asp Thr₃₃₀ Gly His Ser Leu His Asp₃₃₅
 Tyr Lys Glu Ala Leu Met Ile Pro Ala₃₄₅ Lys Asp Gly Val Lys₃₅₀ Leu Pro
 Tyr Phe Phe His Ala Gly Glu Thr₃₆₀ Asp Trp Gln Gly Thr₃₆₅ Ser Ile Asp
 Arg Asn Ile Leu Asp Ala Leu₃₇₅ Met Leu Asn Thr Thr₃₈₀ Arg Ile Gly His
 Gly Phe Ala Leu Ser Lys₃₉₀ His Pro Ala Val Arg₃₉₅ Thr Tyr Ser Trp Lys₄₀₀
 Lys Asp Ile Pro Ile₄₀₅ Glu Val Cys Pro Ile₄₁₀ Ser Asn Gln Val Leu Lys₄₁₅
 Leu Val Ser Asp₄₂₀ Leu Arg Asn His Pro Val Ala Thr Leu Met Ala Thr₄₃₀

Gly His Pro Met Val Ile Ser Ser Asp Asp Pro Ala Met Phe Gly Ala
 435 440 445
 Lys Gly Leu Ser Tyr Asp Phe Tyr Glu Val Phe Met Gly Ile Gly Gly
 450 455 460
 Met Lys Ala Asp Leu Arg Thr Leu Lys Gln Leu Ala Met Asn Ser Ile
 465 470 475
 Lys Tyr Ser Thr Leu Leu Glu Ser Glu Lys Asn Thr Phe Met Glu Ile
 485 490 495
 Trp Lys Lys Arg Trp Asp Lys Phe Ile Ala Asp Val Ala Thr Lys
 500 505 510

<210> 40
 <211> 270
 <212> PRT
 <213> Homo sapiens

<400> 40
 Met Asp Ser Leu Glu Trp Asn Trp Ala Leu Val Tyr Glu Leu Ser Gly
 1 5 10 15
 Glu His His Asp Glu Glu Trp Ser Val Lys Thr Tyr Gln Glu Val Ala
 20 25 30
 Gln Lys Phe Val Glu Thr His Pro Glu Phe Ile Gly Ile Lys Ile Ile
 35 40 45
 Tyr Ser Asp His Arg Ser Lys Asp Val Ala Val Ile Ala Glu Ser Ile
 50 55 60
 Arg Met Ala Met Gly Leu Arg Ile Lys Phe Pro Thr Val Val Ala Gly
 65 70 75 80
 Phe Asp Leu Val Gly His Glu Asp Thr Gly His Ser Leu His Asp Tyr
 85 90 95
 Lys Glu Ala Leu Met Ile Pro Ala Lys Asp Gly Val Lys Leu Pro Tyr
 100 105 110
 Phe Phe His Ala Gly Glu Thr Asp Trp Gln Gly Thr Ser Ile Asp Arg
 115 120 125
 Asn Ile Leu Asp Ala Leu Met Leu Asn Thr Thr Arg Ile Gly His Gly
 130 135 140
 Phe Ala Leu Ser Lys His Pro Ala Val Arg Thr Tyr Ser Trp Lys Lys
 145 150 155 160
 Asp Ile Pro Ile Glu Val Cys Pro Ile Ser Asn Gln Val Leu Lys Leu
 165 170 175
 Val Ser Asp Leu Arg Asn His Pro Val Ala Thr Leu Met Ala Thr Gly
 180 185 190
 His Pro Met Val Ile Ser Ser Asp Asp Pro Ala Met Phe Gly Ala Lys
 195 200 205
 Gly Leu Ser Tyr Asp Phe Tyr Glu Val Phe Met Gly Ile Gly Gly Met
 210 215 220

Lys Ala Asp Leu Arg Thr Leu Lys Gln Leu Ala Met Asn Ser Ile Lys
225 230 235 240

Tyr Ser Thr Leu Leu Glu Ser Glu Lys Asn Thr Phe Met Glu Ile Trp
245 250 255

Lys Lys Arg Trp Asp Lys Phe Ile Ala Asp Val Ala Thr Lys
260 265 270

<210> 41
<211> 3071
<212> DNA
<213> Homo sapiens

<400> 41
gtcctaataag ccacgcccac ctcaaatggc gatcccagag agcatctggc ttcaattttc 60
ctttttccgg aggggagatg gattccctgg aatggaactg ggctctgggtg tatgagctca 120
gtggagagca ccatgacgaa gagtggctcag tgaagactta ccaggaagta gctcagaagt 180
ttgtggaaac tcaccctgag ttatttggaa tcaaaatcat ttattcggat cacagatcca 240
aagatgtggc tgtcatcgca gaatccatcc gaatggccat ggggctccga atcaagttcc 300
ccacgggtggg ggcagggttt gacctgggtg ggcattgagga cactggccac tccttgctgtg 360
actacaagga agctctgatg atccccgcca aggatggcgt taagctgcct tacttcttcc 420
acgccggagg aacagactgg cagggtactt ccatagacag gaacattctg gatgctctga 480
tgctgaacac taccagaatc ggccatggat ttgctttgag caaacacccc gcagtcagga 540
cttactcctg gaaaaaggac atccccatag aagtctgtcc catctctaac cagggtgctga 600
aactgggtgtc tgacttgagg aaccaccctg tagccactct gatggccact gggcacccca 660
tggtgatcag ctctgatgac ccagctatgt ttggtgccaa aggcttgctc tatgatttct 720
atgaggtctt catgggcatt ggggggatga aggctgacct gaggaccctc aaacagctgg 780
ccatgaactc tatcaagtac agtaccctgt tggagagtga gaaaaatact ttcattggaaa 840
tctggaagaa gagatgggat aagttcatag cagatgtggc tacaagtgga ggagaagcta 900
gccagccctc tacaagctgt cctcttgac acgctgtcac ttctctcac tctgtcttga 960
atcagctcca tgtgcccatt aaatcaatgg cctctgtatg gagcgacct gtgagaagca 1020
cttggtctggc tgagcaaatt catcctctgg aaataattctc tctcagccac agtgacattg 1080
accctcttgg ttttctctg tctctggcca ttctctccag tttccctatt tcagagtctt 1140
ctcctctctc tgatctctgt gctgtttcct caggactcag tcctgggctc tcttctattc 1200
tggtctcttt atttttttat ttttgtattt ttctcagatg gagttttgct cttgttgccc 1260
aggctggagt acaatggtgc gatctcagct cagtgcaccc tccgccaccc gggttcaggc 1320
aattctcttg catcagcctc gcgagtagtt ggaattatag gcatgtgcca ccacaccag 1380
ctgatttttg catttttagt agagacaggt tttcaccatg ttggcgaggc tggatccaa 1440
ctcttgacct caggtagtcc actcgccctc ttgtctccaa agtgctggaa ttacaggcat 1500
tagccaccat gcctggccta ttctgggtctc tttaactctc tctcttttat ttctcttctc 1560
tctctgtaca ctttctctgg gtggctctcat ccattccttt gctttttcat accattttatt 1620
tgttaatgat tcccacattt atttatgcac ttggagagct cacaggaatc tcagaaactg 1680
atgaggtaca attctgaacc ctctcagctct tcccttttaa cctttctttt tctctacttt 1740
aatttttcta aagatgtgtc tgctatgttg ccaggtctgg tctccaactc aagtgatcct 1800
cctgcccagc tctcccgaag tgctgggatt ctgacatga gccaccacac tcagcccttt 1860
aaaccttttc ctggcctttc ccatagctgg tgaaggacac ctccatccat tccacgcagt 1920
tgctcaaagc agaaattttc agtgcaagtc ttgatgctgc gccgtccccc actccctaca 1980
tcagaacgca tccctcatct ggactccagc ggtggcttct tgatgctgcg cgggtccccc 2040
ctccctacat cagaatgcat cccgcatcca cactccagcg gctgagctct acctgcacgc 2100
tgttgccaag tccaagctac catactcctg catctcctat gacaacagcc tcctcactga 2160
tctccctttt cttccctttg cctcctccag ctcatttttc acagtgtaga atgacatttt 2220
gtttgtttgt tttgttttgt ttgagatgga gtctcgctct gttgccaggg gtggagtga 2280
gcgggtgcgat ctgggtcac tgcaacctcc acctcccggg ttcaagcgga ttctcgtgcc 2340
tcagcctcct gagtagctgg gattacaggc atgcaccacc atgcccggat aatttttga 2400
tttttagtag agatggggtt tcactatgtt ggccaggctg gtctcgaaca cctgacctcg 2460
tgggtccacc gcctcggcct cccaaagcac tgggattaca ggcgtgagcc acccggcctg 2520
gcctagaatg acttttaaaa gatcaaatta aatcagggtc ctcttttgct tacaacgcag 2580
tgctgttaga ggtacacccc atgtctccac agggcataca gcatccgatt taatctggat 2640
ccattccggc gccttctctc cccagtcacc cagaggggcc caaccccggc ggcctttct 2700
tcctcaaagc tcctcggctc tataccgtgc ctgggtcttt tctctttctc tctgacctga 2760
agattccttc tttccctttt tgctctgccc actcctggtt acccttcaag tttcaagttc 2820
atgtcactgt ctgagagagg ttttctgtgt ctgcacctgt ttctctcagg aagccttgct 2880
cttttccatc atgcctctaa tcacagctta taatcggata tttatttctg tgtctacagt 2940

cttgccctgc cagactgtat gcccacatgtg ggcaggcgct catgattggt tctgattggt 3000
 tcacgcatgc tgctaaccga gagcctgggc ccaaagctag ttagtactca ataaacaatg 3060
 cattgaatga g 3071

<210> 42
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 42
 agcctgcaga

10

<210> 43
 <211> 173
 <212> PRT
 <213> Homo sapiens

<400> 43
 Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Gly Ala Ser Leu Trp
 1 5 10 15
 Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu Ala Val
 20 25 30
 Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val
 35 40 45
 His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys Met
 50 55 60
 Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met Ser
 65 70 75 80
 Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp Arg
 85 90 95
 Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu Val
 100 105 110
 Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala Phe
 115 120 125
 Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu Val Thr
 130 135 140
 Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser
 145 150 155 160
 Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu
 165 170

<210> 44
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 44
 aaaagaaatg

10

<210> 45
 <211> 142

<212> PRT
 <213> Homo sapiens

<400> 45
 Met Leu Asp Ala Ala Cys Pro Gln Thr Ser Asp Cys Lys Phe Phe Ser
 1 5 10 15
 Phe Trp Thr Leu Gly Leu Thr Pro Val Ile Cys Gln Glu Leu Ser Gly
 20 25 30
 Leu Gln Gln Gln Thr Glu Gly Cys Ala Val Gly Phe Pro Thr Phe Glu
 35 40 45
 Val Leu Gly Leu Arg Leu Ala Ser Trp Leu Phe Ser Leu Gln Met Ala
 50 55 60
 Tyr Cys Gly Thr Ser Pro Cys Asp His Leu Ala Arg Ala Ser Trp Gly
 65 70 75 80
 Asn Gly Ala Pro Gly Val Val Val Gln Thr Ser Ala Ser Leu Gln Pro
 85 90 95
 Val Cys Gly Gly Ala Arg Cys Ser Arg Ser Leu Arg Thr Ala Gly Arg
 100 105 110
 Pro Ala His Pro Trp Lys Arg Glu Leu Ala Ala Trp Ser Thr Pro Pro
 115 120 125
 Arg Arg Ala Arg Thr Ala Gly Thr Pro Met Phe Leu Pro Leu
 130 135 140

<210> 46
 <211> 1972
 <212> DNA
 <213> Homo sapiens

<400> 46
 aatagcagga gtggcctgtg cctcaacaac atgcttttgc tgcagataat cagccagagc 60
 ctgtttctct actcctcgct aagaatgctt tgtttcccgt aaggaaagct tttagctaata 120
 ctataatcta tagaaactat gcttatcatt ggctgtcagt aaatatgtgg gtcagactct 180
 gtttgtggct ttcagctctg aaggctgtca gctccctgat cccactctgc actctatttc 240
 tgtgtctgtc tttaattctt ctagtgctgc tgggttaggg tcttcacgac cgaactgggtc 300
 tcggcagtggt aagaacatgg aaacactagg ctgacttagt ctcccagcct acatctttct 360
 cccatgctgg acgctgcctg ccctcaaaca tcagactgca agttcttcag cttttggact 420
 cttggactta caccagtgat ttgccaggag ctctcaggcc ttcagcaaca gactgaaggc 480
 tgtgtctgtcg gcttccctac ttttgagggt ttgggactca gactggcttc ctggctcttc 540
 agcttgcaga tggcctattg tgggacttca ccttgtgatc atctcgcccg tgcttcgtgg 600
 gggaatggag cccctggagt ggttgtgcag accagtgcaa gcctacaacc cgtgtgcgga 660
 ggcgctcggg gcagcaggag cctcagaacg gcggggcgcc ctgcccaccc ctggaagaga 720
 gagctggctg cctggagtac tccaccccg caggccagga ctgcgggcac acctatgttc 780
 ctgccittat aactacctct gcattcaaca aggagagaac acgacaagct acgtctccac 840
 actggtctac acacacagag gatgctggat actgtatgga gttaaagaca gactccttga 900
 ctctcactg tgctctggaa aaccggccct tgactagatg gatgcagtat ctccgagagg 960
 gatacacggg gtgtgtggat tgtcagcctc cagctatgaa ctctgtgagc cttcgttgtt 1020
 ctggagatgg cctggactcc gatggaaatc agactctcca ttggcaagca attggtaatc 1080
 ctcggtgtca aggaacttgg aaaaaagttc ggcgagtaga ccagtgttct tgtccagctg 1140
 ttcacagttt tatttttata tagatggtga tataaatatt tccaaatgca tttgtaaaaca 1200
 ttctaataat tctcaagtca tgttcaatgt ttctaaacc ttcaattttg gccaaagtcc 1260
 ccaaacacat cattgccaca ctctgaagta gagaaagaaa atttaggggc cagtctcaa 1320
 ggaacacagg tcctttatatt ttattttaac taagtgtgaag acccactcaa aaagctcttg 1380
 tggttttatg ttcttgacct ttcaactgga gtctctcat tcagcagggt gcccgtaga 1440
 cacagaatac atgtctgttt gctaaagtaa aattactgta actcagtcgg attattgggtg 1500
 acggaagtgt catttaaggg gatctatgtt ttgaacctg cagtctattt tataatcttt 1560
 taaaagtttc tcactttaaa acatgtatac aggtatgcac actctacaca tatatttcca 1620

```

tataacttaat attccataaa ttctgaaata atttcaagca aattatcaca aataatTTTT 1680
ccacagggca aattatttaa aatttttagta agcattctgt aatgaaaaac caagttatac 1740
taaaaacatt ttttgaagaa gaaatTTTT ttttgcttaa gaagtgaag gatcaaatgc 1800
tcattgtagg tcaatgggag ttcttttaaa tgttataatt ttcatgaaaa gaaatgttga 1860
caccagtga tgaacaatt gctttcattc tgaaaattct accaccattt gcatctaaga 1920
ttatttccaa ggcttaaagc ctgaagctga ataaaaataat ctttcagagt cc 1972

```

<210> 47
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 47
 atgacctgaa 10

<210> 48
 <211> 332
 <212> PRT
 <213> Homo sapiens

<400> 48
 Met Leu Leu Leu Ala Ala Ala Phe Leu Val Ala Phe Val Leu Leu Leu
 1 5 10 15
 Tyr Met Val Ser Pro Leu Ile Ser Pro Lys Pro Leu Ala Leu Pro Gly
 20 25 30
 Ala His Val Val Val Thr Gly Gly Ser Ser Gly Ile Gly Lys Cys Ile
 35 40 45
 Ala Ile Glu Cys Tyr Lys Gln Gly Ala Phe Ile Thr Leu Val Ala Arg
 50 55 60
 Asn Glu Asp Lys Leu Leu Gln Ala Lys Lys Glu Ile Glu Met His Ser
 65 70 75 80
 Ile Asn Asp Lys Gln Val Val Leu Cys Ile Ser Val Asp Val Ser Gln
 85 90 95
 Asp Tyr Asn Gln Val Glu Asn Val Ile Lys Gln Ala Gln Glu Lys Leu
 100 105 110
 Gly Pro Val Asp Met Leu Val Asn Cys Ala Gly Met Ala Val Ser Gly
 115 120 125
 Lys Phe Glu Asp Leu Glu Val Ser Thr Phe Glu Arg Leu Met Ser Ile
 130 135 140
 Asn Tyr Leu Gly Ser Val Tyr Pro Ser Arg Ala Val Ile Thr Thr Met
 145 150 155 160
 Lys Glu Arg Arg Val Gly Arg Ile Val Phe Val Ser Ser Gln Ala Gly
 165 170 175
 Gln Leu Gly Leu Phe Gly Phe Thr Ala Tyr Ser Ala Ser Lys Phe Ala
 180 185 190
 Ile Arg Gly Leu Ala Glu Ala Leu Gln Met Glu Val Lys Pro Tyr Asn
 195 200 205
 Val Tyr Ile Thr Val Ala Tyr Pro Pro Asp Thr Asp Thr Pro Gly Phe
 210 215 220

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

<210> 49
 <211> 2272
 <212> DNA
 <213> Homo sapiens

<400> 49
 aggcgccccg cgcgcgcgcg tgattctcgc ctgcgcgcag cccagccctg cgcgccctgc 60
 ccggcgcccc ccgccccggc gctccggggc cctggccccg cggagcgatg ctgctgctgg 120
 ctgcccgcctt cctcgtggcc ttcgtgctgc tgctgtacat ggtgtctccg ctcatcagcc 180
 ccaagccccct cgccctgccc ggggcgcgat tgggtggttac aggagggtcc agtggcatcg 240
 ggaagtgcag tgctatcgag tgctataaac aaggagcttt tataactctg gttgcacgaa 300
 atgaggataa gctgctgcag gcaaagaaag aaattgaaat gcactctatt aatgacaaac 360
 aggtggtgct ttgcatatca gttgatgtat ctcaagacta taaccaagta gagaatgtca 420
 taaaacaagc acaggagaaa ctgggtccag tggacatgct ggtaaattgt gcaggaatgg 480
 cagtgtcagg aaaatttgaa gatcttgaag ttagtacctt tgaaaggtta atgagcatca 540
 attacctggg cagcgtgtac cccagccggg ccgtgatcac caccatgaag gagcgccggg 600
 tgggcaggat cgtgtttgtg tcctcccagg caggacagtt gggattattc ggtttcacag 660
 cctactctgc atccaagttt gccataaggg gatggcaga agctttgcag atggagggtga 720
 agccatataa tgtctacatc acagtgtgctt acccaccaga cacagacaca cctggctttg 780
 ccgaagaaaa cagaacaaag cctttggaga ctgcacttat ttcagagacc acatctgtgt 840
 gcaaaccaga acaggtggcc aaacaaattg ttaaagatgc catacaagga aatttcaaca 900
 gttcccttgg ctcatatggg tacatgctct cggccctgac ctgtgggatg gctccagtaa 960
 ctctctattac tgaggggctc cagcaggtgg tcaccatggg ccttttccgc actattgtct 1020
 tgttttacct tggaagtgtt gacagcatag ttcgtcgtcg catgatgcag agagaaaaat 1080
 ctgaaaatgc agacaaaact gcctaattctt cttacccctt ggaagaagac tgtttccaaa 1140
 taatttgaac agcttgctgc taaatgggac ccaatttttg gcctatagac acttatgtat 1200
 tgttttcgaa tacgtcagat tggaccagtg ctcttcagga atgtggctgc aagcaagggg 1260
 ctagaagttc acctcctgac agtattatta atactatgca aatatggaat aggagaccat 1320
 ttgattttct aggctttgtg gtagagaggt gaaggtatga gaattaatag cgtgtgaaca 1380
 aagtaaagaa caggattcca gaatgatcat taaatttgtt tctattttatt cttttttgcc 1440
 cccctagaga ttaagtccag aaatgtactt tctggcacat aaagaaatct tgaggacttt 1500
 gtttaaacct tccataaaaa aacaattttc ggtttctcgg gttctctctc tctgtctctc 1560
 tgtctctctg tctctctgtc tctctgtctc tctctctctc tctctttctt tctttgtgta 1620
 ttttattcaa gatgagttgg acccattgcc agtgagtctg aatgtcactg acagccctgt 1680
 gttgtgctca ggactcactc tgctgctggg ggaaactcat ggcttctctc tctctttgat 1740
 cccataaagc tacgaggggg acgggagagg gcagtgcaat gggaagtaaa gagatatttt 1800
 ccagtaggaa aagcaatgct ttcttgctct tagactcaaa tgcttaggga acgtttcatt 1860
 tctcattcat ggggaaaggc agcctcctta aatgttttct gaagagcggg aaaatctaga 1920
 agcttaagaa tttaacagttc cttcaataac catgatgacc tgaagttcac ctatcccat 1980
 ttgacatcta cttgtttttc ccatctcttc ctttccaatt ttgcttatac tgcgtgaata 2040
 tttttgtaaa aaaaaaaaaa aaggaaaaaa aagaccagct aaaattttcg acttgacttt 2100
 ttaacttaac tcatgaatta attaaagcaa atgaaaaaat taaaaagtgt gactttttct 2160
 cggagcatat atgtagcttt taggaaaggc tgatgatggg ataaagtgtg ctcatthaaga 2220

aaaaaagaca aggctgattt tgaagagagt tgcttttgaa ataaaatgat ca 2272

<210> 50
<211> 10
<212> DNA
<213> Homo sapiens

<400> 50
tttaatacat 10

<210> 51
<211> 160
<212> PRT
<213> Homo sapiens

<400> 51
Met Gln Arg Val Ser Gly Leu Leu Ser Trp Thr Leu Ser Arg Val Leu
1 5 10 15
Trp Leu Ser Gly Leu Ser Glu Pro Gly Ala Ala Arg Gln Pro Arg Ile
20 25 30
Met Glu Glu Lys Ala Leu Glu Val Tyr Asp Leu Ile Arg Thr Ile Arg
35 40 45
Asp Pro Glu Lys Pro Asn Thr Leu Glu Glu Leu Glu Val Val Ser Glu
50 55 60
Ser Cys Val Glu Val Gln Glu Ile Asn Glu Glu Glu Tyr Leu Val Ile
65 70 75 80
Ile Arg Phe Thr Pro Thr Val Pro His Cys Ser Leu Ala Thr Leu Ile
85 90 95
Gly Leu Cys Leu Arg Val Lys Leu Gln Arg Cys Leu Pro Phe Lys His
100 105 110
Lys Leu Glu Ile Tyr Ile Ser Glu Gly Thr His Ser Thr Glu Glu Asp
115 120 125
Ile Asn Lys Gln Ile Asn Asp Lys Glu Arg Val Ala Ala Ala Met Glu
130 135 140
Asn Pro Asn Leu Arg Glu Ile Val Glu Gln Cys Val Leu Glu Pro Asp
145 150 155 160

<210> 52
<211> 1019
<212> DNA
<213> Homo sapiens

<400> 52
ggcacgaggg tctggccccg ccttttctct gctctcctga accttttaggc ttgtctcggc 60
ccatttgaag accaggaagt tgatcaatcc cgaggctgct gagagacggg ggcgcgattg 120
ggacagtcgc cagggatggc tgagcgtgaa gatgcagcgg gtgtccgggc tgctctcctg 180
gacgctgagc agagtcctgt ggctctccgg cctctctgag ccgggagctg cccggcagcc 240
ccgcatcatg gaagagaaag cgctagaggt ttatgatttg attagaacta tccgggaccc 300
agaaaagccc aatacttttag aagaactgga agtgggtctcg gaaagttgtg tggagttca 360
ggagataaat gaagaagaat atctggttat tatcaggttc acgccaacag tacctcattg 420
ctctttggcg actcttattg ggctgtgctt aagagtaaaa cticagcgat gtttaccatt 480
taaacataag ttggaaatct acatttctga aggaaccac tcaacagaag aagacatcaa 540
taagcagata aatgacaaag agcgagtggc agctgcaatg gaaaacccca acttacggga 600

```

aattgtggaa cagtgtgtcc ttgaacctga ctgatatgctg ttttaagagc cactggcctg 660
taattgtttg atatatttgt ttaaactctt tgtataatgt cagagactca tgtttaatac 720
ataggtgatt tgtacctcag agcatttttt aaaggattct ttccaagcga gatttaatta 780
taaggtagta cctaatttgt tcaatgtata acattctcag gatttgtaac acttaaataa 840
tcagacagaa taatattttc tagttattat gtgtaagatg agttgctatt tttctgatgc 900
tcattctgat acaactattt ttcgtgtcaa atatctactg tgcccaaata tactcaattt 960
aaatcattac tctgtaaaaa aaataagcag atgattctta taaaaaaaaa aaaaaaaaaa 1019

```

<210> 53
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 53
 tgtaaacaat

10

<210> 54
 <211> 375
 <212> PRT
 <213> Homo sapiens

<400> 54
 Met Lys Val Trp Leu Leu Leu Gly Leu Leu Leu Val His Glu Ala Leu
 1 5 10 15
 Glu Asp Val Thr Gly Gln His Leu Pro Lys Asn Lys Arg Pro Lys Glu
 20 25 30
 Pro Gly Glu Asn Arg Ile Lys Pro Thr Asn Lys Lys Val Lys Pro Lys
 35 40 45
 Ile Pro Lys Met Lys Asp Arg Asp Ser Ala Asn Ser Ala Pro Lys Thr
 50 55 60
 Gln Ser Ile Met Met Gln Val Leu Asp Lys Gly Arg Phe Gln Lys Pro
 65 70 75 80
 Ala Ala Thr Leu Ser Leu Leu Ala Gly Gln Thr Val Glu Leu Arg Cys
 85 90 95
 Lys Gly Ser Arg Ile Gly Trp Ser Tyr Pro Ala Tyr Leu Asp Thr Phe
 100 105 110
 Lys Asp Ser Arg Leu Ser Val Lys Gln Asn Glu Arg Tyr Gly Gln Leu
 115 120 125
 Thr Leu Val Asn Ser Thr Ser Ala Asp Thr Gly Glu Phe Ser Cys Trp
 130 135 140
 Val Gln Leu Cys Ser Gly Tyr Ile Cys Arg Lys Asp Glu Ala Lys Thr
 145 150 155 160
 Gly Ser Thr Tyr Ile Phe Phe Thr Glu Lys Gly Glu Leu Phe Val Pro
 165 170 175
 Ser Pro Ser Tyr Phe Asp Val Val Tyr Leu Asn Pro Asp Arg Gln Ala
 180 185 190
 Val Val Pro Cys Arg Val Thr Val Leu Ser Ala Lys Val Thr Leu His
 195 200 205
 Arg Glu Phe Pro Ala Lys Glu Ile Pro Ala Asn Gly Thr Asp Ile Val
 210 215 220

Tyr Asp Met Lys Arg Gly Phe Val Tyr Leu Gln Pro His Ser Glu His
 225 230 235 240
 Gln Gly Val Val Tyr Cys Arg Ala Glu Ala Gly Gly Arg Ser Gln Ile
 245 250 255
 Ser Val Lys Tyr Gln Leu Leu Tyr Val Ala Val Pro Ser Gly Pro Pro
 260 265 270
 Ser Thr Thr Ile Leu Ala Ser Ser Asn Lys Val Lys Ser Gly Asp Asp
 275 280 285
 Ile Ser Val Leu Cys Thr Val Leu Gly Glu Pro Asp Val Glu Val Glu
 290 295 300
 Phe Thr Trp Ile Phe Pro Gly Gln Lys Asp Glu Arg Pro Val Thr Ile
 305 310 315 320
 Gln Asp Thr Trp Arg Leu Ile His Arg Gly Leu Gly His Thr Thr Arg
 325 330 335
 Ile Ser Gln Ser Val Ile Thr Val Glu Asp Phe Glu Thr Ile Asp Ala
 340 345 350
 Gly Tyr Tyr Ile Cys Thr Ala Gln Asn Leu Gln Gly Gln Thr Thr Val
 355 360 365
 Ala Thr Thr Val Glu Phe Ser
 370 375

<210> 55
 <211> 1502
 <212> DNA
 <213> Homo sapiens

<400> 55
 cctgcgtccc cgccccgcgc agccgcgcgc ctccctgcgt ccgaggtccg aggttcccga 60
 gatgaaggct tggctctctc ttggctctct gctgggtgcac gaagcgctgg aggatgttac 120
 tggccaacac cttcccaaga acaagcgtcc aaaagaacca ggagagaata gaatcaaac 180
 taccaacaag aaggtgaagc ccaaaattcc taaaatgaag gacagggact cagccaattc 240
 agcaccaaag acgcagtcta tcatgatgca agtgctggat aaaggctcgt tccagaaacc 300
 cgccgctacc ctgagtctgc tggcggggca aactgtagag cttcgatgta aagggtagtag 360
 aattgggtgg agctaccctg cgtatctgga caccittaaag gattctcggc tcagcgtcaa 420
 gcagaatgag cgctacggcc agttgactct ggtcaactcc acctcggcag acacaggtga 480
 attcagctgc tgggtgcagc tctgcagcgg ctacatctgc aggaaggacg aggccaaaac 540
 gggctccacc tacatctttt ttacagagaa aggagaactc tttgtacctt ctcccagcta 600
 cttcgatgtt gtctacttga acccgagacag acaggctgtg gttccttgtc ggggtgaccgt 660
 gctgtcggcc aaagtcacgc tccacaggga attcccagcc aaggagatcc cagccaatgg 720
 aacggacatt gtttatgaca tgaagcgggg ctttgtgtat ctgcaacctc attccgagca 780
 ccagggtgtg gtttactgca gggcggaggc cgggggcaga tctcagatct ccgtcaagta 840
 ccagctgctc tacgtggcgg ttcccagtgg cctccctca acaaccatct tggcttcttc 900
 aaacaaagtg aaaagtgggg acgacatcag tgtgctctgc actgtcctgg gggagcccga 960
 tgtggagggt gagttcacct ggatcttccc agggcagaag gatgaaaggc ctgtgacgat 1020
 ccaagacact tggaggttga tccacagagg actgggacac accacgagaa tctcccagag 1080
 tgtcattaca gtggaagact tgcagacgat tgatgcagga tattacattt gcactgctca 1140
 gaatcttcaa ggacagacca cagtagctac cactgttgag ttttcctgac ttggaaaagg 1200
 aaatgtaatg aacttatgga aagcccattt gtgtacacag tcagctttgg ggttcctttt 1260
 attagtgtt tgccagaggc tgatgtcaag caccacaccc caaccccagc gtctcgtgag 1320
 tccgaccag acatccaaac taaaaggaag tcattccagtc tattcacaga agtggttaact 1380
 tttctaacag aaagcatgat tttgattgct tacctacata cgtgttccta gtttttatac 1440
 atgtgtaaac aattttatat aatcaatcat ttctattaaa tgagcacggt tttgtaaaaa 1500
 at 1502

<210> 56
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 56
 ggaccgaggg

10

<210> 57
 <211> 194
 <212> PRT
 <213> Homo sapiens

<400> 57
 Met Arg Leu Leu Pro Cys Ala Ala Gly Thr Ser Ala Phe Leu Ser Ser
 1 5 10 15
 Gln Leu Thr Ile Gly Ala Ser Ser Ser Ser Arg Pro Ser Thr Pro Ala
 20 25 30
 Trp Phe Ser Val Leu Glu Tyr Ile Gln Pro Phe Arg Ser Cys Ser Phe
 35 40 45
 Leu Leu Ser Ser Ser Pro Pro Gly Ala Leu Pro Cys Ser Leu Cys Thr
 50 55 60
 Ser Asp Ser Ile Leu Pro Val Ser Asp Asp Phe Thr Ser Ser Lys Arg
 65 70 75 80
 Pro Leu Gly Pro Ala Phe Ser Lys Ala Phe Leu Glu Asn Ala Gly Gln
 85 90 95
 Asp Leu Pro Phe Asn Phe Val Phe Ser Val Cys Ser Val Ser Ser Ala
 100 105 110
 Gly Cys Leu Ser Ser Arg Ile Thr Glu Asp Met Ser Val Leu Phe Pro
 115 120 125
 Leu Met Pro Gly Ala Arg Leu Gly His Val Cys Ser Ala Met Ala Asn
 130 135 140
 Val Arg Arg Gln Gly Trp Ser His Ser Gly Leu Leu Ala Leu Trp Pro
 145 150 155 160
 Ser Gly Arg Ser Lys Gly Arg Ser Cys Gln Lys Glu Thr Cys Leu Val
 165 170 175
 Thr Gly Arg Ser Gln Gly Phe Gln Asn Pro Ile Gln Val Ile Pro Ser
 180 185 190
 Trp Trp

<210> 58
 <211> 1841
 <212> DNA
 <213> Homo sapiens

<400> 58
 tgttttgttt ctttatttgg acttaacggc cactcatttc tgttggcgcc atgcgactgt 60
 tgccgtgtgc agcagggacc tctgccttcc tctcctccca gctcacaata ggtgcctctt 120
 cctcctcccg gccatccacg ccggcctggg tttcggttct ggaatacatc caaccctttc 180
 gttcgtgttc ttttctcctg tcatcttctc cgcctggagc tcttccatgt tctctttgca 240

```

caagcgactc taticcttcca gtttcagatg atttcacctc ctcaaagagg cctttggggac 300
cagccttttc taaagctttc cttgagaatg ctgggcaaga tttacctttt aattttgttt 360
tttctgtttg ttctgtctcc tctgccggct gcttgagctc caggatcaca gaggacatga 420
gtgttctgtt cccactgatg cctgggtgcca gactgggaca cgtgtgctca gccatggcta 480
atgtgcgag gcagggatgg tcacacagtg gcctcttggc gctctggccg agtgggagga 540
gcaaggggag gagctgtcag aaggagacct gcctagtcac aggcagggtca cagggcttcc 600
agaatcccat tcagggtgatc ccctcctggg ggtaaacagc acacctcaga agtgggtgtt 660
tgtcatagtt gttcttatgg ctcccacagc ctcattctctg tttgttccat ttgggattga 720
ggagttgttg agtccagcgt caccttcaga gccagtgcac tccttatcaa tagtgagggc 780
cgctcgtttt gattgttccg tttcgcttga taaagctgag tgagaattat attggccaac 840
agaaacgtct ctctccatgt gaacacagcc actgggtcac cagggacgtg gcttacagtt 900
ccagcaggct cctgtcacac tgggtgccttt gggcagtaca gaactgtgct gctgggtggg 960
acgtccgggg ctgtgcagtg tcgcttcctt ttgcatatgt tgctattggg gtcatttcct 1020
ctctgagatc agtggaaagtc tcagggtcagt cccgtggccc tgtattgctt agtgggcat 1080
gtgttttaaac ttccaaaagg catattttaa atagtattg gagtggtgat gtgaagagct 1140
cctttgataa tgtattttct ttgtagaaat ttaaaagcct gttgattcaa gaactttcta 1200
aagtatttcc ggaagacatg gctaagtatc gaagcatccg gggggaggat caccgcctt 1260
cttaaccagc tcaccctccc tgtgtgaaga tcccctggga ctgcgatgcg gcgtgaggct 1320
gggactgcga gtgctgacgc caccttcctg ctgagggtgg actgggccct ggacacacc 1380
ctcagccctt ctgtcctcat tgtttggcct catgggaccg aggggctgga ggagaggcgg 1440
agctgtgccc cagctgttcc agcagcttgt ctggcgtcaa ctggcttca gagtgctgac 1500
ccctcatcac tgtggggatc attctctctg agggcagatg aggcgcagga aaatagtctt 1560
ggaaatgtta aatatgatgg gtaaattaaa agttttacaa cattctacct aatatttttc 1620
ttttaacata ctctttctgt tctattgtat tatgggtgcc gaaagctaaa taacgactag 1680
gaaaaatttt tttaaaaaaa gaaaaatcag tttaatgtgg gaagtactta atgggtatta 1740
tattttacat tttcaagtat agtgcataaa gaatgtttta aatgtaactg ttttcatgga 1800
tttcaattag acatgcctat aataaactaa gtatgtggct t 1841

```

<210> 59
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 59
 tccaaccag 10

<210> 60
 <211> 355
 <212> PRT
 <213> Homo sapiens

<400> 60
 Met Gly Asn Leu Phe Met Leu Trp Ala Ala Leu Gly Ile Cys Cys Ala
 1 5 10 15
 Ala Phe Ser Ala Ser Ala Trp Ser Val Asn Asn Phe Leu Ile Thr Gly
 20 25 30
 Pro Lys Ala Tyr Leu Thr Tyr Thr Thr Ser Val Ala Leu Gly Ala Gln
 35 40 45
 Ser Gly Ile Glu Glu Cys Lys Phe Gln Phe Ala Trp Glu Arg Trp Asn
 50 55 60
 Cys Pro Glu Asn Ala Leu Gln Leu Ser Thr His Asn Arg Leu Arg Ser
 65 70 75 80
 Ala Thr Arg Glu Thr Ser Phe Ile His Ala Ile Ser Ser Ala Gly Val
 85 90 95
 Met Tyr Ile Ile Thr Lys Asn Cys Ser Met Gly Asp Phe Glu Asn Cys
 100 105 110

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

<210> 61

<211> 1597

<212> PRT

<213> Homo sapiens

<400> 61

Cys Ala Gly Ala Ala Thr Thr Thr Thr Cys Thr Cys Ala Cys Ala Thr
 1 5 10 15

Ala Ala Ala Thr Ala Cys Thr Gly Ala Gly Gly Ala Ala Gly Ala Cys
 20 25 30

Cys Cys Thr Gly Cys Cys Cys Thr Cys Thr Cys Cys Thr Cys Ala Cys
 35 40 45

Thr Cys Cys Thr Cys Thr Gly Gly Ala Cys Thr Thr Gly Gly Cys Cys

50					55					60						
Cys	Thr	Gly	Ala	Gly	Cys	Thr	Gly	Gly	Ala	Cys	Cys	Thr	Gly	Gly	Thr	
65					70					75						80
Cys	Cys	Ala	Cys	Thr	Gly	Gly	Gly	Gly	Thr	Ala	Gly	Gly	Cys	Ala	Gly	
				85					90					95		
Gly	Gly	Cys	Gly	Ala	Thr	Gly	Gly	Gly	Gly	Ala	Ala	Cys	Cys	Thr	Gly	
			100					105					110			
Thr	Thr	Thr	Ala	Thr	Gly	Cys	Thr	Cys	Thr	Gly	Gly	Gly	Cys	Ala	Gly	
		115					120						125			
Cys	Thr	Cys	Thr	Gly	Gly	Gly	Cys	Ala	Thr	Ala	Thr	Gly	Cys	Thr	Gly	
	130						135					140				
Thr	Gly	Cys	Thr	Gly	Cys	Ala	Thr	Thr	Cys	Ala	Gly	Thr	Gly	Cys	Cys	
145					150					155					160	
Thr	Cys	Thr	Gly	Cys	Cys	Thr	Gly	Gly	Thr	Cys	Ala	Gly	Thr	Gly	Ala	
				165					170					175		
Ala	Cys	Ala	Ala	Thr	Thr	Thr	Cys	Cys	Thr	Gly	Ala	Thr	Ala	Ala	Cys	
			180					185						190		
Ala	Gly	Gly	Thr	Cys	Cys	Cys	Cys	Ala	Ala	Gly	Gly	Cys	Cys	Thr	Ala	
		195						200					205			
Cys	Thr	Gly	Ala	Cys	Cys	Thr	Ala	Cys	Ala	Cys	Gly	Ala	Cys	Thr	Ala	
	210						215					220				
Gly	Thr	Gly	Thr	Gly	Gly	Cys	Cys	Thr	Thr	Gly	Gly	Gly	Thr	Gly	Cys	
225						230					235				240	
Cys	Cys	Ala	Gly	Ala	Gly	Thr	Gly	Gly	Cys	Ala	Thr	Cys	Gly	Ala	Gly	
				245					250					255		
Gly	Ala	Gly	Thr	Gly	Cys	Ala	Ala	Gly	Thr	Thr	Cys	Cys	Ala	Gly	Thr	
			260					265					270			
Thr	Thr	Gly	Cys	Thr	Thr	Gly	Gly	Gly	Ala	Ala	Cys	Gly	Cys	Thr	Gly	
		275					280						285			
Gly	Ala	Ala	Cys	Thr	Gly	Cys	Cys	Cys	Thr	Gly	Ala	Ala	Ala	Ala	Thr	
	290					295					300					
Gly	Cys	Thr	Cys	Thr	Thr	Cys	Ala	Gly	Cys	Thr	Cys	Thr	Cys	Cys	Ala	
305						310					315				320	
Cys	Cys	Cys	Ala	Cys	Ala	Ala	Cys	Ala	Gly	Gly	Cys	Thr	Gly	Ala	Gly	
				325					330					335		
Ala	Ala	Gly	Thr	Gly	Cys	Thr	Ala	Cys	Cys	Ala	Gly	Ala	Gly	Ala	Gly	
			340					345					350			
Ala	Cys	Thr	Thr	Cys	Cys	Thr	Thr	Cys	Ala	Thr	Ala	Cys	Ala	Thr	Gly	
		355					360						365			
Cys	Thr	Ala	Thr	Cys	Ala	Gly	Cys	Thr	Cys	Thr	Gly	Cys	Thr	Gly	Gly	
	370					375					380					
Ala	Gly	Thr	Cys	Ala	Thr	Gly	Thr	Ala	Cys	Ala	Thr	Cys	Ala	Thr	Cys	
385						390					395				400	

Ala Cys Cys Ala Ala Gly Ala Ala Cys Thr Gly Thr Ala Gly Cys Ala
 405 410 415
 Thr Gly Gly Gly Thr Gly Ala Cys Thr Thr Cys Gly Ala Ala Ala Ala
 420 425 430
 Cys Thr Gly Thr Gly Gly Cys Thr Gly Thr Gly Ala Thr Gly Gly Gly
 435 440 445
 Thr Cys Ala Ala Ala Cys Ala Ala Thr Gly Gly Ala Ala Ala Ala Ala
 450 455 460
 Cys Ala Gly Gly Ala Gly Gly Cys Cys Ala Thr Gly Gly Cys Thr Gly
 465 470 475 480
 Gly Ala Thr Cys Thr Gly Gly Gly Gly Ala Gly Gly Cys Thr Gly Cys
 485 490 495
 Ala Gly Cys Gly Ala Cys Ala Ala Thr Gly Thr Gly Gly Ala Ala Thr
 500 505 510
 Thr Thr Gly Gly Gly Gly Ala Ala Ala Gly Gly Ala Thr Cys Thr Cys
 515 520 525
 Cys Ala Ala Ala Cys Thr Cys Thr Thr Thr Gly Thr Gly Gly Ala Cys
 530 535 540
 Ala Gly Thr Thr Thr Gly Gly Ala Gly Ala Ala Gly Gly Gly Gly Ala
 545 550 555 560
 Ala Gly Gly Ala Thr Gly Cys Cys Ala Gly Ala Gly Cys Cys Cys Thr
 565 570 575
 Gly Ala Thr Gly Ala Ala Thr Cys Thr Thr Cys Ala Cys Ala Ala Cys
 580 585 590
 Ala Ala Cys Ala Gly Gly Gly Cys Cys Gly Gly Cys Ala Gly Ala Cys
 595 600 605
 Thr Gly Gly Cys Ala Gly Thr Gly Ala Gly Ala Gly Cys Cys Ala Cys
 610 615 620
 Cys Ala Thr Gly Ala Ala Ala Ala Gly Gly Ala Cys Ala Thr Gly Cys
 625 630 635 640
 Ala Ala Ala Thr Gly Thr Cys Ala Thr Gly Gly Cys Ala Thr Cys Thr
 645 650 655
 Cys Thr Gly Gly Gly Ala Gly Cys Thr Gly Cys Ala Gly Cys Ala Thr
 660 665 670
 Ala Cys Ala Gly Ala Cys Ala Thr Gly Cys Thr Gly Gly Cys Thr Gly
 675 680 685
 Cys Ala Gly Cys Thr Gly Gly Cys Thr Gly Ala Ala Thr Thr Cys Cys
 690 695 700
 Gly Gly Gly Ala Gly Ala Thr Gly Gly Gly Ala Gly Ala Cys Thr Ala
 705 710 715 720
 Cys Cys Thr Ala Ala Ala Gly Gly Cys Cys Ala Ala Gly Thr Ala Thr
 725 730 735
 Gly Ala Cys Cys Ala Gly Gly Cys Gly Cys Thr Gly Ala Ala Ala Ala
 740 745 750

Thr Thr Gly Ala Ala Ala Thr Gly Gly Ala Thr Ala Ala Gly Cys Gly
 755 760 765
 Gly Cys Ala Gly Cys Thr Gly Ala Gly Ala Gly Cys Thr Gly Gly Gly
 770 775 780
 Ala Ala Cys Ala Gly Cys Gly Cys Cys Gly Ala Gly Gly Gly Cys Cys
 785 790 795 800
 Ala Cys Thr Gly Gly Gly Thr Gly Cys Cys Cys Gly Cys Thr Gly Ala
 805 810 815
 Gly Gly Cys Cys Thr Thr Cys Cys Thr Thr Cys Cys Thr Ala Gly Cys
 820 825 830
 Gly Cys Ala Gly Ala Gly Gly Cys Gly Gly Ala Ala Cys Thr Gly Ala
 835 840 845
 Thr Cys Thr Thr Thr Thr Thr Ala Gly Ala Gly Gly Ala Ala Thr Cys
 850 855 860
 Ala Cys Cys Ala Gly Ala Thr Thr Ala Cys Thr Gly Thr Ala Cys Cys
 865 870 875 880
 Thr Gly Cys Ala Ala Thr Thr Cys Cys Ala Gly Cys Cys Thr Gly Gly
 885 890 895
 Gly Cys Ala Thr Cys Thr Ala Thr Gly Gly Cys Ala Cys Ala Gly Ala
 900 905 910
 Gly Gly Gly Thr Cys Gly Thr Gly Ala Gly Thr Gly Cys Cys Thr Ala
 915 920 925
 Cys Ala Gly Ala Ala Cys Ala Gly Cys Cys Ala Cys Ala Ala Cys Ala
 930 935 940
 Cys Ala Thr Cys Cys Ala Gly Gly Thr Gly Gly Gly Ala Gly Cys Gly
 945 950 955 960
 Ala Cys Gly Thr Ala Gly Cys Thr Gly Thr Gly Gly Gly Cys Gly Cys
 965 970 975
 Cys Thr Gly Thr Gly Cys Ala Cys Thr Gly Ala Gly Thr Gly Thr Gly
 980 985 990
 Gly Gly Cys Thr Gly Cys Ala Gly Gly Thr Gly Gly Ala Ala Gly Ala
 995 1000 1005
 Gly Ala Gly Gly Ala Ala Ala Ala Cys Thr Gly Ala Gly Gly Thr Cys
 1010 1015 1020
 Ala Thr Ala Ala Gly Cys Ala Gly Cys Thr Gly Thr Ala Ala Cys Thr
 1025 1030 1035 1040
 Gly Cys Ala Ala Ala Thr Thr Cys Cys Ala Gly Thr Gly Gly Thr Gly
 1045 1050 1055
 Cys Thr Gly Thr Ala Cys Gly Gly Thr Cys Ala Ala Gly Thr Gly Thr
 1060 1065 1070
 Gly Ala Cys Cys Ala Gly Thr Gly Thr Ala Gly Gly Cys Ala Thr Gly
 1075 1080 1085
 Thr Gly Gly Thr Gly Ala Gly Cys Ala Ala Gly Thr Ala Thr Thr Ala

1090	1095	1100
Cys Thr Gly Cys Gly Cys Ala Cys Gly Cys Thr Cys Cys Cys Cys Ala		
1105	1110	1115
Gly Gly Cys Ala Gly Thr Gly Cys Cys Cys Ala Gly Thr Cys Cys Cys		
	1125	1130
Thr Gly Gly Gly Gly Ala Gly Ala Gly Thr Thr Thr Gly Gly Thr Thr		
	1140	1145
Thr Gly Gly Gly Gly Thr Cys Thr Ala Thr Ala Thr Cys Thr Ala Gly		
	1155	1160
Ala Gly Gly Gly Ala Cys Cys Thr Thr Cys Ala Ala Ala Gly Thr Ala		
	1170	1175
Thr Thr Thr Gly Thr Thr Cys Cys Thr Thr Thr Ala Ala Ala Thr Thr		
	1185	1190
Thr Cys Ala Gly Ala Cys Cys Ala Thr Gly Thr Cys Cys Ala Ala Cys		
	1205	1210
Cys Cys Ala Gly Cys Thr Gly Thr Gly Cys Thr Gly Cys Thr Gly Gly		
	1220	1225
Gly Ala Ala Thr Cys Ala Gly Gly Ala Gly Ala Ala Thr Ala Gly Ala		
	1235	1240
Ala Gly Cys Ala Ala Ala Ala Ala Cys Gly Ala Ala Ala Gly Ala		
	1250	1255
Gly Thr Thr Cys Thr Gly Thr Thr Cys Ala Gly Ala Cys Thr Thr Cys		
	1265	1270
Thr Gly Ala Ala Gly Ala Gly Cys Ala Gly Cys Cys Thr Gly Thr Gly		
	1285	1290
Gly Cys Thr Ala Cys Ala Ala Ala Thr Cys Thr Ala Thr Gly Cys Thr		
	1300	1305
Gly Ala Thr Ala Ala Ala Thr Gly Ala Gly Ala Thr Thr Gly Ala Gly		
	1315	1320
Ala Ala Cys Thr Cys Ala Ala Cys Thr Gly Thr Ala Thr Thr Thr Thr		
	1330	1335
Gly Cys Cys Ala Thr Ala Ala Ala Thr Gly Cys Thr Thr Cys Thr Ala		
	1345	1350
Ala Gly Ala Thr Ala Thr Ala Thr Cys Cys Ala Gly Cys Thr Gly Gly		
	1365	1370
Gly Ala Cys Thr Thr Cys Thr Ala Thr Thr Ala Cys Thr Cys Cys Cys		
	1380	1385
Thr Thr Thr Gly Gly Ala Ala Ala Cys Cys Thr Thr Ala Ala Gly Ala		
	1395	1400
Thr Cys Ala Ala Ala Ala Ala Gly Gly Gly Ala Ala Thr Ala Ala Gly		
	1410	1415
Ala Ala Ala Cys Cys Cys Thr Thr Cys Thr Thr Cys Thr Gly Thr Ala		
	1425	1430
		1435
		1440

Thr Cys Cys Cys Ala Ala Thr Ala Ala Thr Cys Cys Ala Cys Cys Ala
 1445 1450 1455
 Gly Gly Ala Thr Ala Ala Ala Gly Gly Ala Gly Ala Ala Ala Cys Thr
 1460 1465 1470
 Ala Gly Ala Ala Ala Thr Ala Thr Gly Cys Ala Ala Cys Thr Cys Cys
 1475 1480 1485
 Cys Thr Thr Gly Ala Thr Thr Thr Cys Ala Gly Thr Gly Thr Thr Thr
 1490 1495 1500
 Gly Gly Cys Ala Gly Gly Thr Ala Ala Cys Ala Ala Ala Ala Ala Ala
 1505 1510 1515 1520
 Thr Thr Gly Ala Gly Ala Cys Cys Cys Ala Gly Ala Cys Ala Cys Thr
 1525 1530 1535
 Gly Gly Thr Cys Ala Ala Cys Ala Gly Gly Ala Ala Ala Ala Cys Ala
 1540 1545 1550
 Ala Thr Ala Cys Ala Gly Ala Cys Thr Cys Cys Cys Ala Gly Ala Ala
 1555 1560 1565
 Thr Thr Ala Gly Ala Ala Ala Gly Thr Gly Thr Thr Ala Thr Thr Thr
 1570 1575 1580
 Thr Ala Ala Thr Gly Cys Ala Ala Cys Cys Thr Ala Gly
 1585 1590 1595

<210> 62
 <211> 10
 <212> DNA
 <213> Homo sapiens

<400> 62
 actcgaatat

10

<210> 63
 <211> 248
 <212> PRT
 <213> Homo sapiens

<400> 63
 Met Gln Thr Cys Pro Leu Ala Phe Pro Gly His Val Ser Gln Ala Leu
 1 5 10 15
 Gly Thr Leu Leu Phe Leu Ala Ala Ser Leu Ser Ala Gln Asn Glu Gly
 20 25 30
 Trp Asp Ser Pro Ile Cys Thr Glu Gly Val Val Ser Val Ser Trp Gly
 35 40 45
 Glu Asn Thr Val Met Ser Cys Asn Ile Ser Asn Ala Phe Ser His Val
 50 55 60
 Asn Ile Lys Leu Arg Ala His Gly Gln Glu Ser Ala Ile Phe Asn Glu
 65 70 75 80
 Val Ala Pro Gly Tyr Phe Ser Arg Asp Gly Trp Gln Leu Gln Val Gln
 85 90 95
 Gly Gly Val Ala Gln Leu Val Ile Lys Gly Ala Arg Asp Ser His Ala

100	105	110
Gly Leu Tyr Met Trp His Leu Val	Gly His Gln Arg Asn Asn Arg Gln	
115	120	125
Val Thr Leu Glu Val Ser Gly Ala Glu Pro Gln Ser Ala Pro Asp Thr		
130	135	140
Gly Phe Trp Pro Val Pro Ala Val Val Thr Ala Val Phe Ile Leu Leu		
145	150	155
Val Ala Leu Val Met Phe Ala Trp Tyr Arg Cys Arg Cys Ser Gln Gln		
165	170	175
Arg Arg Glu Lys Lys Phe Phe Leu Leu Glu Pro Gln Met Lys Val Ala		
180	185	190
Ala Leu Arg Ala Gly Ala Gln Gln Gly Leu Ser Arg Ala Ser Ala Glu		
195	200	205
Leu Trp Thr Pro Asp Ser Glu Pro Thr Pro Arg Pro Leu Ala Leu Val		
210	215	220
Phe Lys Pro Ser Pro Leu Gly Ala Leu Glu Leu Leu Ser Pro Gln Pro		
225	230	235
240		
Leu Phe Pro Tyr Ala Ala Asp Pro		
245		

<210> 64
 <211> 2000
 <212> DNA
 <213> Homo sapiens

<400> 64
 attttcctgg ggctccgggg cgcgagagaag ctgcatccca gaggagcgcg tccaggagcg 60
 gacccgggag tgtttcaaga gccagtgaaca aggaccaggg gcccaagtcc caccagccat 120
 gcagacctgc cccctggcat tccctggcca cgtttcccag gcccttggga cctctctgtt 180
 tttggctgac tccttgagtg ctcagaatga aggctgggac agcccatct gcacagagg 240
 ggtagtctct gtgtcttggg gcgagaacac cgtcatgtcc tgcaacatct ccaacgcctt 300
 ctcccatgtc aacatcaagc tgcgtgcccc cgggcaggag agcgccatct tcaatgaggt 360
 ggctccaggc tacttctccc gggacggctg gcagctccag gttcaggag gcgtggcaca 420
 gctggtgatc aaaggcgccc gggactccca tgctgggctg tacatgtggc acctcgtggg 480
 acaccagaga aataacagac aagtcacgct ggaggtttca ggtgcagaac cccagtcgac 540
 ccctgacact gggttctggc ctgtgcccag cgtgtcttca tctcttgggt 600
 cgctctggtc atgttcgcct ggtacagggt cgctgttcc cagcaacgcc gggagaagaa 660
 gttcttcttc ctagaacccc agatgaaggt cgagccctc agagcgggag cccagcaggg 720
 cctgagcaga gcctccgctg aactgtggac cccagactcc gagccacccc caaggccgct 780
 ggcactgggt ttcaaaccct caccacttgg agccctggag ctgtgttccc cccaaccctt 840
 gtttccatat gccgcagacc catagccgcc tgcaaggcag agaggacaca ggagagccag 900
 ccctgagatg cgaccttggg tggcggggcc gtgggtcttc gtcccacccc gagggcacag 960
 acaccggctt gcttggcagg ctgggcctct gtgtcaccca ctctgggtg cgtgcagacc 1020
 cttcccctcc acccccagc tcttccaagc tctgttccct cagtttccaa aatggaacca 1080
 cctcacctcc gcagacccc acttaccagg acgcatgccc ctccctctgc cctcatcaaa 1140
 cccacagacc cggactccct ttctgccacc ccaggctggg ccggccccag gtgtgggggtc 1200
 cgctctctcc actcccagg ctccgcgccc aagtgagggg gcccctgccg gacccctcaga 1260
 cacactggag ttcagggtg ggggggcctt ggcacatacc tgtcccttgg ctatgagcag 1320
 gctttggggg cccttccgag gcagccccgg gggccgaggt agggcttggg ggcttagagg 1380
 ctgggatggc tcctggcccc accgccaggg ggcaagcgca ggccgggctg ggaggcggcg 1440
 gcggcggtc gggctggggg gtcaggtgga cgctgcctcc ggggctgggt gcgcattcct 1500
 cagtccctcg gccaccggg ggtcgctccc tcgtgcccac cgcacctgcc gagcctcttt 1560
 ggacccagat ctgttcattg tttgtcttc gtcactgcgg cgggggcctt tgatgtcttc 1620
 atctgtatgg ggtggaaaaa tcaccgggaa tcccccttca gttctttgaa aaagtcccat 1680
 gactcgaata tctgaaatga agaaaacaaa ccgactcaca aacctccaag tagctccaaa 1740

tgcaatTTTT	aaaatggaaa	acaaaaatCT	gaaagaaacg	TCCTTtagtg	ctttaagccc	1800
caaaacgtcc	ctaaggcgTC	ctcgagatga	agacgggggg	gagccccag	ccaggTggag	1860
accccgCagg	acgcggcgGC	gcccggTgac	CgaggCctcg	cacagccggc	cgccctgagg	1920
gtcgggCcg	agccagggtc	caagaggggc	gcgtttgtgt	ctcgggttaa	aataaggTtc	1980
cgTccgcgtg	ctgggtcaga					2000