



(86) Date de dépôt PCT/PCT Filing Date: 2001/06/21
(87) Date publication PCT/PCT Publication Date: 2001/12/27
(85) Entrée phase nationale/National Entry: 2002/12/19
(86) N° demande PCT/PCT Application No.: US 2001/019942
(87) N° publication PCT/PCT Publication No.: 2001/098354
(30) Priorités/Priorities: 2000/06/21 (60/214,027) US;
2000/08/25 (60/228,045) US; 2000/12/12 (60/255,104) US

(51) Cl.Int.⁷/Int.Cl.⁷ C07K 14/705, A61K 39/395, A61K 38/17,
A01K 67/027, C12N 15/12, C07K 16/28, C12Q 1/68,
G01N 33/577, G01N 33/53

(71) Demandeur/Applicant:
INCYTE GENOMICS, INC., US

(72) Inventeurs/Inventors:
GRIFFIN, JENNIFER A., US;
KALLICK, DEBORAH A., US;
TRIBOULEY, CATHERINE M., US;
YUE, HENRY, US;
NGUYEN, DANNIEL B., US;
...

(74) Agent: SMART & BIGGAR

(54) Titre : RECEPTEURS
(54) Title: HUMAN RECEPTORS

(57) **Abrégé/Abstract:**

The invention provides human receptors (REPTR) and polynucleotides which identify and encode REPTR. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of REPTR.



(72) **Inventeurs(suite)/Inventors(continued)**: TANG, Y. TOM, US; LAL, PREETI, US; POLICKY, JENNIFER L., US; AZIMZAI, YALDA, US; LU, DYUNG AINA M., US; GRAUL, RICHARD, US; YAO, MONIQUE G., US; BURFORD, NEIL, US; HAFALIA, APRIL J. A., US; BAUGHN, MARIAH R., US; BANDMAN, OLGA, US; PATTERSON, CHANDRA, US; YANG, JUNMING, US; XU, YUMING, US; GANDHI, AMEENA R., US; WARREN, BRIDGET A., US; DING, LI, US; SANJANWALA, MADHU S., US; DUGGAN, BRENDAN M., US; LU, YAN, US

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/098354 A3

(51) International Patent Classification⁷: **C07K 14/705**,
16/28, C12N 15/12, C12Q 1/68, A01K 67/027, A61K
38/17, 39/395, G01N 33/53, 33/577

(21) International Application Number: PCT/US01/19942

(22) International Filing Date: 21 June 2001 (21.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/214,027 21 June 2000 (21.06.2000) US
60/228,045 25 August 2000 (25.08.2000) US
60/255,104 12 December 2000 (12.12.2000) US

(71) Applicant (for all designated States except US): **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **KALLICK, Deborah, A.** [US/US]; 900 Olive Street, Menlo Park, CA 94025 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **NGUYEN, Danniell, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **GRAUL, Richard** [US/US]; 682-29th Avenue, San Francisco, CA 94121 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago

Road, San Leandro, CA 94577 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **YANG, Junming** [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014 (US). **DING, Li** [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US). **SANJAN-WALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). **LU, Yan** [CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, 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, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
17 October 2002

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: HUMAN RECEPTORS

(57) Abstract: The invention provides human receptors (REPTR) and polynucleotides which identify and encode REPTR. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of REPTR.



WO 01/098354 A3

RECEPTORS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of receptors and to the use of these sequences in the diagnosis, treatment, and prevention of autoimmune/inflammatory, reproductive, gastrointestinal, developmental, endocrine, neurological, and cell proliferative disorders including cancer, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of receptors.

BACKGROUND OF THE INVENTION

The term receptor describes proteins that specifically recognize other molecules. Most receptors are cell surface proteins which bind extracellular ligands and produce cellular responses in the areas of growth, differentiation, endocytosis, and immune response. Other receptors facilitate the selective transport of proteins out of the endoplasmic reticulum and localize enzymes to particular locations in the cell.

Cell surface receptors are typically integral plasma membrane proteins. These receptors recognize hormones such as catecholamines; peptide hormones, e.g., glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, parathyroid hormone, and vasopressin; growth and differentiation factors, e.g., epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, platelet-derived growth factor, nerve growth factor, colony-stimulating factors, and erythropoietin; small peptide factors such as thyrotropin-releasing hormone; galanin, somatostatin, and tachykinins; cytokines, e.g., chemokines, interleukins, interferons, and tumor necrosis factor; small peptide factors such as bombesin, oxytocin, endothelin, angiotensin II, vasoactive intestinal peptide, and bradykinin; neurotransmitters such as neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, e.g., enkephalins, endorphins and dynorphins; galanin, somatostatin, and tachykinins; and circulatory system-borne signaling molecules, e.g., angiotensin, complement, calcitonin, endothelins, and formyl-methionyl peptides. They also recognize cell adhesion molecules in the extracellular matrix, or molecules on the surface of other cells. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptides. Other cell surface receptors bind ligands to be internalized by the cell. This receptor-mediated endocytosis functions in the uptake of low density lipoproteins (LDL), transferrin, glucose- or mannose-terminal glycoproteins, galactose-terminal glycoproteins, immunoglobulins, phosphovitellogenins, fibrin, proteinase-inhibitor complexes, plasminogen activators, and thrombospondin (Lodish, H. et al. (1995) Molecular Cell Biology,

Scientific American Books, New York NY, p. 723; Mikhailenko, I. et al. (1997) J. Biol. Chem. 272:6784-6791).

Transmembrane proteins (TM) are characterized by extracellular, transmembrane, and intracellular domains. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an α -helical conformation. TM proteins are classified as bitopic (Types I and II) proteins, which span the membrane once, and polytopic (Types III and IV) (Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96) proteins, which contain multiple membrane-spanning segments. TM proteins that act as cell-surface receptor proteins involved in signal transduction include growth and differentiation factor receptors, and receptor-interacting proteins such as Drosophila pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins) and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins function as vesicle and organelle-forming molecules, such as calveolins; or cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins.

Many membrane proteins (MPs) contain amino acid sequence motifs that serve to localize proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in targeted cancer treatment of tumor vasculature (Arap, W. et al. (1998) Science, 279:377-380). Membrane proteins may also contain amino acid sequence motifs that serve to interact with extracellular or intracellular molecules, such as carbohydrate recognition domains.

Chemical modification of amino acid residue side chains alters the manner in which MPs interact with other molecules, such as membrane phospholipids. Examples of such chemical modifications include the formation of covalent bonds with glycosaminoglycans, oligosaccharides, phospholipids, acetyl and palmitoyl moieties, ADP-ribose, phosphate, and sulphate groups.

RNA encoding membrane proteins may have alternative splice sites which give rise to proteins encoded by the same gene but with different messenger RNA and amino acid sequences. Splice variant membrane proteins may interact with other ligand and protein isoforms.

Receptors bound to growth factors trigger intracellular signal transduction pathways which activate various downstream effectors that regulate gene expression, cell division, cell differentiation, cell motility, and other cellular processes. Many growth factor receptors, including receptors for epidermal growth factor, platelet-derived growth factor, fibroblast growth factor, and the growth modulator α -thrombin, contain intrinsic protein kinase activities. These signaling proteins contain a common domain referred to as a Src homology (SH) domain. SH2 domains and SH3 domains are

found in phospholipase C- γ , PI-3-K p85 regulatory subunit, Ras-GTPase activating protein, and pp60^{c-src} (Lowenstein, E.J. et al. (1992) Cell 70:431-442). The cytokine family of receptors share a different common binding domain and include transmembrane receptors for growth hormone (GH), interleukins, erythropoietin, and prolactin. Other receptors and second messenger-binding proteins have intrinsic serine/threonine protein kinase activity. These include activin/TGF- β /BMP-superfamily receptors, calcium- and diacylglycerol-activated/phospholipid-dependant protein kinase (PK-C), and RNA-dependant protein kinase (PK-R). In addition, other serine/threonine protein kinases, including nematode Twitchin, have fibronectin-like, immunoglobulin C2-like domains.

G-protein coupled receptors (GPCRs) are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which span the plasma membrane and form a bundle of antiparallel alpha (α) helices. These proteins range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) Eur. J. Biochem. 196:1-10; Coughlin, S.R. (1994) Curr. Opin. Cell Biol. 6:191-197). The amino-terminus of the GPCR is extracellular, of variable length and often glycosylated; the carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops of the GPCR alternate with intracellular loops and link the transmembrane domains. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. The activated receptor, in turn, interacts with an intracellular heterotrimeric guanine nucleotide binding (G) protein complex which mediates further intracellular signaling activities, generally the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate, or interactions with ion channel proteins (Baldwin, J.M. (1994) Curr. Opin. Cell Biol. 6:180-190).

The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. Cysteine disulfide bridges connect the second and third extracellular loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily (ExpASY PROSITE document PS00237; and Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp 2-6).

GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators, as well as those for acetylcholine, adenosine, epinephrine and norepinephrine, bombesin, bradykinin, chemokines, dopamine, endothelin, γ -aminobutyric acid (GABA), follicle-stimulating hormone (FSH), glutamate, gonadotropin-releasing hormone (GnRH), hepatocyte growth factor, histamine, leukotrienes, melanocortins, neuropeptide Y, opioid peptides, opsins, prostanoids, serotonin, somatostatin, tachykinins, thrombin, thyrotropin-releasing hormone (TRH), vasoactive intestinal polypeptide family, vasopressin and oxytocin, and orphan receptors. Neuropeptide Y (NPY) is a 36 amino acid amidated peptide which produces a pronounced feeding

response in a variety of species. The actions of NPY are believed to be mediated by a family of receptor subtypes named Y1 - Y6. The Y1 and Y5 receptor subtypes are intimately involved in NPY-induced feeding (Doods, H.N. (2000) Expert Opin. Investig. Drugs 9:1327-1346).

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, supra). For instance, retinitis pigmentosa may arise from mutations in the rhodopsin gene. Rhodopsin is the retinal photoreceptor which is located within the discs of the eye rod cell. Parma, J. et al. (1993, Nature 365:649-651) report that somatic activating mutations in the thyrotropin receptor cause hyperfunctioning thyroid adenomas and suggest that certain GPCRs susceptible to constitutive activation may behave as protooncogenes. Other mutations and changes in transcriptional activation of GPCR-encoding genes have been associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding disorders.

The frizzled cell surface receptor was originally identified in Drosophila melanogaster, where it is important for proper bristle and hair polarity on the wing, leg, thorax, abdomen, and eye of the developing insect. (Wang, Y. et al. (1996) J. Biol. Chem. 271:4468-4476.) Frizzled proteins act as putative Wnt receptors. Distinct intracellular pathways may be activated as a result of Wnt/Frizzled interactions. The canonical pathway involves activation of the cytoplasmic protein Dsh via both beta-catenin-dependent and independent mechanisms (Boutros, M. et al. (2000) Science 288:1825-1828), while a second involves the activation of protein kinase C (Medina A. and Steinbeisser, H. (2000) Dev. Dyn. 218:671-680). The secreted signaling molecules encoded by Wnt genes bind to frizzled receptors and stabilize cytosolic beta-catenin, which induces resistance to apoptosis. Two frizzled-related proteins can act as Wnt antagonists, and are associated with human overload-induced heart failure (Schumann, H. et al. (2000) 45:720-728). The frizzled gene encodes a 587 amino acid protein which contains an N-terminal signal sequence and seven putative transmembrane regions. The N-terminus is cysteine-rich and is probably extracellular while the C-terminus is probably cytosolic. Multiple frizzled gene homologs have been found in rat, mouse, and human. The frizzled receptors are not homologous to other seven-transmembrane-region receptors.

Cell adhesion molecules

Families of cell adhesion molecules include the cadherins, integrins, and lectins. Cadherins comprise a family of calcium-dependent glycoproteins that function in mediating cell-cell adhesion in virtually all solid tissues of multicellular organisms. These proteins share multiple repeats of a cadherin-specific motif, and the repeats form the folding units of the cadherin extracellular domain. Cadherin molecules cooperate to form focal contacts, or adhesion plaques, between adjacent epithelial cells. Cadherins preferentially bind one another on cells in contact, acting as both receptor and ligand. The cadherin family includes the classical cadherins and protocadherins.

Classical cadherins include the E-cadherin, N-cadherin, and P-cadherin subfamilies. E-cadherin is present on many types of epithelial cells and is especially important for embryonic development. N-cadherin is present on nerve, muscle, and lens cells and is also critical for embryonic development. P-cadherin is present on cells of the placenta and epidermis. Recent studies report that

5 protocadherins are involved in a variety of cell-cell interactions (Suzuki, S.T. (1996) *J. Cell Sci.* 109:2609-2611). The intracellular anchorage of cadherins is regulated by their dynamic association with catenins, a family of cytoplasmic signal transduction proteins associated with the actin cytoskeleton. The anchorage of cadherins to the actin cytoskeleton appears to be regulated by protein tyrosine phosphorylation, and the cadherins are the target of phosphorylation-induced junctional

10 disassembly (Aberle, H. et al. (1996) *J. Cell. Biochem.* 61:514-523).

Nuclear receptors bind small molecules such as hormones or second messengers, leading to increased receptor-binding affinity to specific chromosomal DNA elements. In addition the affinity for other nuclear proteins may also be altered. Such binding and protein-protein interactions may regulate and modulate gene expression. Examples of such receptors include the steroid hormone

15 receptors family, the retinoic acid receptors family, and the thyroid hormone receptors family.

Ligand-gated receptor ion channels include extracellular (ELG) and intracellular (ILG) channels. ELGs rapidly transduce neurotransmitter-binding events into electrical signals, such as fast synaptic neurotransmission. ELGs include channels directly gated by neurotransmitters such as acetylcholine, L-glutamate, glycine, ATP, serotonin, GABA, and histamine. ELG genes encode

20 proteins having strong structural and functional similarities. ILGs are activated by many intracellular second messengers. ILGs are encoded by distinct and unrelated gene families and include receptors for cAMP, cGMP, calcium ions, ATP, and metabolites of arachidonic acid.

Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are

25 trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer domain, an α -helical coiled-coil domain, and a triple helical collagenous domain. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides,

30 phospholipids, and asbestos (Matsumoto, A. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:9133-9137; Elomaa, O. et al. (1995) *Cell* 80:603-609). Scavenger receptors have been implicated in the development of atherosclerosis and other macrophage-associated functions. The bovine type I and type II scavenger receptors are multidomain transmembrane proteins that differ only by the presence in the type I receptor of an additional, extracellular cysteine-rich C-terminal domain. The type

35 I-specific scavenger receptor cysteine-rich (SRCR) (one, three, or four per polypeptide chain) is

found in diverse secreted and cell-surface proteins including CD5, complement factor I, Ly-1, and speract receptor (Freeman, M. et al. (1990) Proc. Natl. Acad. Sci. U S A 87:8810-8814).

T cell receptors (TCRs) stimulate T cell antigen recognition and the transmission of signals that both induce death in infected cells and stimulate proliferation of other immune cells. A T cell
 5 recognizes an antigen when it is presented to the TCR as a peptide complexed with a major histocompatibility molecule (MHC) on the surface of an antigen presenting cell. The TCR on most T cells consists of immunoglobulin-like integral membrane glycoproteins containing two polypeptide subunits, α and β , of similar molecular weight. Both TCR subunits have an extracellular domain containing both variable and constant regions, a transmembrane domain that traverses the membrane
 10 once, and a short intracellular domain (Saito, H. et al. (1984) Nature 309:757-762). The genes for the TCR subunits are constructed through somatic rearrangement of different gene segments. Interaction of antigen in the proper MHC context with the TCR initiates signaling cascades that induce the proliferation, maturation, and function of cellular components of the immune system (Weiss, A. (1991) Annu. Rev. Genet. 25: 487-510). Rearrangements in TCR genes and alterations in TCR
 15 expression have been noted in lymphomas, leukemias, autoimmune disorders, and immunodeficiency disorders (Aisenberg, A.C. et al. (1985) N. Engl. J. Med. 313:529-533; Weiss, supra).

Selectins, or LEC-CAMs, comprise a specialized lectin subfamily involved primarily in inflammation and leukocyte adhesion (reviewed in Lasky, L. A. (1991) J. Cell. Biochem. 45:139-146). Selectins mediate the recruitment of leukocytes from the circulation to sites of acute
 20 inflammation and are expressed on the surface of vascular endothelial cells in response to cytokine signaling. Selectins bind to specific ligands on the leukocyte cell membrane and enable the leukocyte to adhere to and migrate along the endothelial surface. Binding of selectin to its ligand leads to polarized rearrangement of the actin cytoskeleton and stimulates signal transduction within the leukocyte (Brenner, B. et al. (1997) Biochem. Biophys. Res. Commun. 231:802-807; Hidari, K. I. et
 25 al. (1997) J. Biol. Chem. 272:28750-28756). Members of the selectin family possess three characteristic motifs: a lectin or carbohydrate recognition domain; an epidermal growth factor-like domain; and a variable number of short consensus repeats (scr or "sushi" repeats). Sushi domains, also known as complement control protein (CCP) modules, or short consensus repeats (SCR), occur in a wide variety of complement and adhesion proteins (Norman, D.G. et al. (1991) J. Mol. Biol.
 30 219:717-725).

Leucine rich repeats (LRR) are short motifs found in numerous proteins from a wide range of species. LRR motifs are of variable length, most commonly 20-29 amino acids and multiple repeats are typically present in tandem. LRR is important for protein/protein interactions and cell adhesion, and LRR proteins are involved in cell/cell interactions, morphogenesis, and development (Kobe, B.
 35 and Deisenhofer, J. (1995) Curr. Opin. Struct. Biol. 5:409-416). The human ISLR (immunoglobulin

superfamily containing leucine-rich repeat) protein contains a C2-type immunoglobulin domain as well as LRR. The ISLR gene is linked to the critical region for Bardet-Biedl syndrome, a developmental disorder of which the most common feature is retinal dystrophy (Nagasawa, A. et al. (1999) *Genomics* 61:37-43).

5 The discovery of new receptors and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of autoimmune/inflammatory, reproductive, gastrointestinal, developmental, endocrine, neurological, and cell proliferative disorders including cancer, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of receptors.

10

SUMMARY OF THE INVENTION

The invention features purified polypeptides, receptors, referred to collectively as "REPTR" and individually as "REPTR-1," "REPTR-2," "REPTR-3," "REPTR-4," "REPTR-5," "REPTR-6," "REPTR-7," "REPTR-8," "REPTR-9," "REPTR-10," "REPTR-11," and "REPTR-12." In one
15 aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of
20 SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-12.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the
25 group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.
30 In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-12. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:13-24.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group
35 consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting

of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group

consisting of SEQ ID NO:13-24, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional REPTR, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino

acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional REPTR, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional REPTR, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the

activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:13-24, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting

of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

5

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

25

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

35

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now
5 described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

10 “REPTR” refers to the amino acid sequences of substantially purified REPTR obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which intensifies or mimics the biological activity of REPTR. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other
15 compound or composition which modulates the activity of REPTR either by directly interacting with REPTR or by acting on components of the biological pathway in which REPTR participates.

An “allelic variant” is an alternative form of the gene encoding REPTR. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or
20 many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding REPTR include those sequences with deletions,
25 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as REPTR or a polypeptide with at least one functional characteristic of REPTR. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding REPTR, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding
30 REPTR. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent REPTR. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of REPTR is retained. For example,
35 negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged

amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine.

Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

5 The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to a sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

10 “Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

 The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of REPTR. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small
15 molecules, or any other compound or composition which modulates the activity of REPTR either by directly interacting with REPTR or by acting on components of the biological pathway in which REPTR participates.

 The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant.

20 Antibodies that bind REPTR polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin,
25 thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

 The term “antigenic determinant” refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies
30 which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

 The term “antisense” refers to any composition capable of base-pairing with the “sense” (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA;
35 RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as

phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once
5 introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical
10 functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic REPTR, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid
15 sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution.
20 Compositions comprising polynucleotide sequences encoding REPTR or fragments of REPTR may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

25 "Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison
30 WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows
35 amino acids which may be substituted for an original amino acid in a protein and which are regarded

as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
5	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
10	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
15	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
20	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation,
 25 (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide.
 30 Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

35 A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a
 40 diseased and a normal sample.

A "fragment" is a unique portion of REPTR or the polynucleotide encoding REPTR which is

identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:13-24 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:13-24, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:13-24 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:13-24 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:13-24 and the region of SEQ ID NO:13-24 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-12 is encoded by a fragment of SEQ ID NO:13-24. A fragment of SEQ ID NO:1-12 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-12. For example, a fragment of SEQ ID NO:1-12 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-12. The precise length of a fragment of SEQ ID NO:1-12 and the region of SEQ ID NO:1-12 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length

supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes
5 in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases “percent identity” and “% identity,” as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some
10 alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e
15 sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the “percent similarity” between aligned polypeptide sequence pairs.

20 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

25 *Open Gap: 11 and Extension Gap: 1 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

30 Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment
35 length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be

used to describe a length over which percentage identity may be measured.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

5 The term “humanized antibody” refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

 “Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific
10 hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive
15 conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

20 Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and
25 conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

 High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS,
30 for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular
35 circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions

will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of REPTR which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of REPTR which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of REPTR. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of REPTR.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where

necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition.

5 PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an REPTR may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will
10 vary by cell type depending on the enzymatic milieu of REPTR.

"Probe" refers to nucleic acid sequences encoding REPTR, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

15 "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous
20 nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

25 Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs
30 can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to
35 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to

5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

“Reporter molecules” are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

5 An “RNA equivalent,” in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

10 The term “sample” is used in its broadest sense. A sample suspected of containing REPTR, nucleic acids encoding REPTR, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

15 The terms “specific binding” and “specifically binding” refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope “A,” the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

20 The term “substantially purified” refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

25 A “substitution” refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

 “Substrate” refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

30 A “transcript image” refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

 “Transformation” describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid
35 sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based

on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term “transformed cells” includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A “transgenic organism,” as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass “single nucleotide polymorphisms” (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a

propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-
 5 1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

10 THE INVENTION

The invention is based on the discovery of new human receptors (REPTR), the polynucleotides encoding REPTR, and the use of these compositions for the diagnosis, treatment, or prevention of autoimmune/inflammatory, reproductive, gastrointestinal, developmental, endocrine, neurological, and cell proliferative disorders including cancer.

15 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is
 20 denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte
 25 polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

30 Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the
 35 MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group,

Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are receptors. For example, SEQ ID NO:1 is 68% identical from residue C221 to residue C842 to rat transmembrane receptor UNC5H1 (GenBank ID g2055392) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0 (rounded down from a very small value by the BLAST program), which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a ZU5 domain (a domain present in ZO1 and Unc5-like netrin receptors) as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from MOTIFS and BLAST_PRODOME analyses provide further corroborative evidence that SEQ ID NO:1 is an Unc5-like netrin receptor. SEQ ID NO:8 is 40% identical from residue Q263 to residue G973 to *Drosophila melanogaster* adherin (GenBank ID g4887715) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:8 also contains a cadherin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:8 is a cell surface receptor. SEQ ID NO:12 is 40% identical from residue M1 to residue P304 to human complement receptor 1 (GenBank ID g451303) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $4.8e-107$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains Sushi (complement) repeat domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:12 is a complement receptor. SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-12 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide

consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:13-24 or that distinguish between SEQ ID NO:13-24 and
5 related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective
10 full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 3974950F6 is the identification number of an Incyte cDNA sequence, and ADRETUT06 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from
15 pooled cDNA libraries (e.g., 55106555H1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g2229606) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g5926688_010.edit is the identification number of a Genscan-predicted coding sequence, with
20 g5926688 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL023814_00001 represents a "stitched" sequence in which 023814 is the identification number of
25 the cluster of sequences to which the algorithm was applied, and 00001 is the number of the prediction generated by the algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. For example, FL6977010_g8176711_000001_g5832711 is the identification number of a "stretched" sequence, with 6977010 being the Incyte project identification
30 number, g8176711 being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, and g5832711 being the GenBank identification number of the nearest GenBank protein homolog. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not
35 shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors
5 which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses REPTR variants. A preferred REPTR variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the REPTR amino acid sequence, and which contains at least one functional or structural characteristic of REPTR.

10 The invention also encompasses polynucleotides which encode REPTR. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:13-24, which encodes REPTR. The polynucleotide sequences of SEQ ID NO:13-24, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the
15 sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding REPTR. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding REPTR. A particular aspect of the invention encompasses a variant of a
20 polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:13-24 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:13-24. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of REPTR.

25 It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding REPTR, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These
30 combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring REPTR, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode REPTR and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring REPTR under appropriately
35 selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding

REPTR or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding REPTR and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode REPTR and REPTR derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding REPTR or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:13-24 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding REPTR may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed,

restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.)

Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising
5 a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme
10 digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed
15 using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been
20 size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze
25 the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire
30 process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode REPTR may be cloned in recombinant DNA molecules that direct expression of
35 REPTR, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent

degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express REPTR.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter REPTR-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of REPTR, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding REPTR may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, REPTR itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of REPTR, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or

a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.)

The composition of the synthetic peptides may be confirmed by amino acid analysis or by
5 sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active REPTR, the nucleotide sequences encoding REPTR or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers,
10 constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding REPTR. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding REPTR. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding REPTR and its initiation codon and
15 upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression
20 may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding REPTR and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques,
25 and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences
30 encoding REPTR. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or
35 animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster

(1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and
 5 Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al.
 10 (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding REPTR. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding REPTR can be achieved using a
 15 multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding REPTR into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of
 20 nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of REPTR are needed, e.g. for the production of antibodies, vectors which direct high level expression of REPTR may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of REPTR. A number of vectors
 25 containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994)
 30 Bio/Technology 12:181-184.)

Plant systems may also be used for expression of REPTR. Transcription of sequences encoding REPTR may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock
 35 promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al.

(1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.)

These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

5 In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding REPTR may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses REPTR in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc.
10 Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are
15 constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of REPTR in cell lines is preferred. For example, sequences encoding REPTR can be transformed
20 into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which
25 successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *apr⁻* cells, respectively. (See, e.g., Wigler, M. et
30 al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981)
35 J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which

alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to
5 quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding REPTR is inserted within a marker gene sequence, transformed cells containing
10 sequences encoding REPTR can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding REPTR under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding REPTR and that express
15 REPTR may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of REPTR using either
20 specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on REPTR is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See,
25 e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and
30 may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding REPTR include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding REPTR, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available,
35 and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase

such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding REPTR may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode REPTR may be designed to contain signal sequences which direct secretion of REPTR through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding REPTR may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric REPTR protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of REPTR activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the REPTR encoding sequence and the heterologous protein sequence, so that REPTR may be cleaved away from the heterologous moiety following purification.

Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

5 In a further embodiment of the invention, synthesis of radiolabeled REPTR may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

10 REPTR of the present invention or fragments thereof may be used to screen for compounds that specifically bind to REPTR. At least one and up to a plurality of test compounds may be screened for specific binding to REPTR. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of REPTR, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a 15 natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which REPTR binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express REPTR, either as a secreted 20 protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing REPTR or cell membrane fractions which contain REPTR are then contacted with a test compound and binding, stimulation, or inhibition of activity of either REPTR or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is 25 detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with REPTR, either in solution or affixed to a solid support, and detecting the binding of REPTR to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical 30 libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

REPTR of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of REPTR. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for 35 REPTR activity, wherein REPTR is combined with at least one test compound, and the activity of

REPTR in the presence of a test compound is compared with the activity of REPTR in the absence of the test compound. A change in the activity of REPTR in the presence of the test compound is indicative of a compound that modulates the activity of REPTR. Alternatively, a test compound is combined with an in vitro or cell-free system comprising REPTR under conditions suitable for
5 REPTR activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of REPTR may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding REPTR or their mammalian homologs may be “knocked out” in an animal model system using homologous recombination in embryonic stem
10 (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo;
15 Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic
20 Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding REPTR may also be manipulated in vitro in ES cells derived from
25 human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding REPTR can also be used to create “knockin” humanized animals
30 (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding REPTR is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease.
35 Alternatively, a mammal inbred to overexpress REPTR, e.g., by secreting REPTR in its milk, may

also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists
 5 between regions of REPTR and receptors. In addition, the expression of REPTR is closely associated with brain tumor tissue, hippocampal tissue, a liver tumor cell line, nasal polyp tissue, and spleen tissue. Therefore, REPTR appears to play a role in autoimmune/inflammatory, reproductive, gastrointestinal, developmental, endocrine, neurological, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased REPTR expression or activity, it is
 10 desirable to decrease the expression or activity of REPTR. In the treatment of disorders associated with decreased REPTR expression or activity, it is desirable to increase the expression or activity of REPTR.

Therefore, in one embodiment, REPTR or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or
 15 activity of REPTR. Examples of such disorders include, but are not limited to, an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis,
 20 cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
 25 polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and hematopoietic cancer including lymphoma, leukemia, and myeloma; a reproductive
 30 disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis; cancer of the breast, fibrocystic breast disease, galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology,

cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism, azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and anejaculation, haemangioblastomas, cystsphaeochromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, α_1 -antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; an endocrine disorder such as a disorder of the hypothalamus and/or pituitary resulting from lesions such as a primary brain tumor, adenoma, infarction associated with pregnancy, hypophysectomy, aneurysm, vascular malformation, thrombosis, infection, immunological disorder, and complication due to head trauma, a disorder associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, a disorder associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, a disorder associated with hypothyroidism

including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, a disorder associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, a disorder associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), a pancreatic disorder such as Type I or Type II diabetes mellitus and associated complications, a disorder associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, a disorder associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbation of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, a hypergonadal disorder associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorder, a peripheral nervous system disorder, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathy, myasthenia gravis, periodic paralysis, a mental disorder including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including

adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

5 In another embodiment, a vector capable of expressing REPTR or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of REPTR including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified REPTR in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent
10 a disorder associated with decreased expression or activity of REPTR including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of REPTR may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of REPTR including, but not limited to, those listed above.

15 In a further embodiment, an antagonist of REPTR may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of REPTR. Examples of such disorders include, but are not limited to, those autoimmune/inflammatory, reproductive, gastrointestinal, developmental, endocrine, neurological, and cell proliferative disorders including cancer, described above. In one aspect, an antibody which specifically binds REPTR may be used
20 directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express REPTR.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding REPTR may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of REPTR including, but not limited to, those described above.

25 In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the
30 various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of REPTR may be produced using methods which are generally known in the art. In particular, purified REPTR may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind REPTR. Antibodies to REPTR may
35 also be generated using methods that are well known in the art. Such antibodies may include, but are

not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans,
5 and others may be immunized by injection with REPTR or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in
10 humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to REPTR have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches
15 of REPTR amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to REPTR may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma
20 technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate
25 antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce REPTR-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be
30 generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA
35 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for REPTR may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between REPTR and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering REPTR epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for REPTR. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of REPTR-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple REPTR epitopes, represents the average affinity, or avidity, of the antibodies for REPTR. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular REPTR epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the REPTR-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of REPTR, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of REPTR-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding REPTR, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding REPTR. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding REPTR. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding REPTR may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides

brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in REPTR expression or regulation causes disease, the expression of REPTR from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

5 In a further embodiment of the invention, diseases or disorders caused by deficiencies in REPTR are treated by constructing mammalian expression vectors encoding REPTR and introducing these vectors by mechanical means into REPTR-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene
10 transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of REPTR include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA),
15 PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). REPTR may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA*
20 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous
25 gene encoding REPTR from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental
30 parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to REPTR expression are treated by constructing a retrovirus vector consisting of (i) the
35 polynucleotide encoding REPTR under the control of an independent promoter or the retrovirus long

terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding REPTR to cells which have one or more genetic abnormalities with respect to the expression of REPTR. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding REPTR to target cells which have one or more genetic abnormalities with respect to the expression of REPTR. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing REPTR to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.

169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding REPTR to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for REPTR into the alphavirus genome in place of the capsid-coding region results in the production of a large number of REPTR-coding RNAs and the synthesis of high levels of REPTR in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of REPTR into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of

polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding REPTR.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding REPTR. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding REPTR.

Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased REPTR expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding REPTR may be therapeutically useful, and in the treatment of disorders associated with decreased REPTR expression or activity, a compound which specifically promotes expression of the polynucleotide encoding REPTR may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding REPTR is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding REPTR are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding REPTR. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S.

Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient.

5 Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and
10 monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of
15 Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of REPTR, antibodies to REPTR, and mimetics, agonists, antagonists, or inhibitors of REPTR.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary,
20 intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of
25 fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

30 Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising REPTR or fragments thereof. For example, liposome preparations
35 containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of

the macromolecule. Alternatively, REPTR or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

5 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

10 A therapeutically effective dose refers to that amount of active ingredient, for example REPTR or fragments thereof, antibodies of REPTR, and agonists, antagonists or inhibitors of REPTR, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the
15 dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity.
20 The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the
25 severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about $0.1 \mu\text{g}$ to $100,000 \mu\text{g}$, up to a total dose of
30 about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind REPTR may be used for the diagnosis of disorders characterized by expression of REPTR, or in assays to monitor patients being treated with REPTR or agonists, antagonists, or inhibitors of REPTR. Antibodies useful for
5 diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for REPTR include methods which utilize the antibody and a label to detect REPTR in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in
10 the art and may be used.

A variety of protocols for measuring REPTR, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of REPTR expression. Normal or standard values for REPTR expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to
15 REPTR under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of REPTR expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding REPTR may be used
20 for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of REPTR may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of REPTR, and to monitor regulation of REPTR levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide
25 sequences, including genomic sequences, encoding REPTR or closely related molecules may be used to identify nucleic acid sequences which encode REPTR. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the
30 probe identifies only naturally occurring sequences encoding REPTR, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the REPTR encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:13-24 or from
35 genomic sequences including promoters, enhancers, and introns of the REPTR gene.

Means for producing specific hybridization probes for DNAs encoding REPTR include the cloning of polynucleotide sequences encoding REPTR or REPTR derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding REPTR may be used for the diagnosis of disorders associated with expression of REPTR. Examples of such disorders include, but are not limited to, an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and hematopoietic cancer including lymphoma, leukemia, and myeloma; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis; cancer of the breast, fibrocystic breast disease, galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism, azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and anejaculation, haemangioblastomas, cystsphaeochromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma,

anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative

5 colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, α_1 -antitrypsin deficiency, Reye's syndrome, primary

10 sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy,

15 epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly,

20 craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; an endocrine disorder such as a disorder of the hypothalamus and/or pituitary resulting from lesions such as a primary brain tumor, adenoma, infarction associated with pregnancy, hypophysectomy, aneurysm, vascular malformation, thrombosis, infection, immunological disorder, and complication due to head trauma, a disorder associated with hypopituitarism including hypogonadism, Sheehan syndrome,

25 diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, a disorder associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, a disorder associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis

30 associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, a disorder associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, a disorder associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), a pancreatic disorder such as Type I or Type II diabetes mellitus and associated complications, a

35 disorder associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal

cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, a disorder associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbation of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, a hypergonadal disorder associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorder, a peripheral nervous system disorder, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathy, myasthenia gravis, periodic paralysis, a mental disorder including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding REPTR may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered REPTR

expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding REPTR may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding REPTR may be labeled by standard methods and added to a fluid or tissue
5 sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding REPTR in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate
10 the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of REPTR, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a
15 sequence, or a fragment thereof, encoding REPTR, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard
20 values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several
25 days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals
30 to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding REPTR may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide
35 encoding REPTR, or a fragment of a polynucleotide complementary to the polynucleotide encoding

REPTR, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding REPTR may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding REPTR are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of REPTR include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the

activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her

5 pharmacogenomic profile.

In another embodiment, REPTR, fragments of REPTR, or antibodies specific for REPTR may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to
10 generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by
15 hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

20 Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of
25 pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test
30 compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes
35 are used to normalize the rest of the expression data. The normalization procedure is useful for

comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of
5 Environmental Health Sciences, released February 29, 2000, available at
<http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the
10 treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

15 Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under
20 given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson,
25 supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are
30 compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be
35 obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for REPTR to quantify the levels of REPTR expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-

2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding REPTR may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding REPTR on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, REPTR, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes
5 between REPTR and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with REPTR, or fragments thereof,
10 and washed. Bound REPTR is then detected by methods well known in the art. Purified REPTR can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing
15 antibodies capable of binding REPTR specifically compete with a test compound for binding REPTR. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with REPTR.

In additional embodiments, the nucleotide sequences which encode REPTR may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on
20 properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific
25 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/214,027, U.S. Ser. No. 60/228,045, and U.S. Ser. No. 60/255,104, are hereby expressly incorporated by reference.

30 EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a
35 suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of

phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in

384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

5 Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cyclor or the PTC-200 thermal cyclor (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared
10 using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI
15 protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing
20 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family
25 databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or
30 Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention
35 may begin at any of the methionine residues of the full length translated polypeptide. Full length

polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:13-24. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative receptors were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode receptors, the encoded polypeptides were analyzed by querying against PFAM models for receptors. Potential receptors were also identified by homology to Incyte cDNA sequences that had been annotated as receptors. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis

was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs

(HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of REPTR Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:13-24 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:13-24 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Génethon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Génethon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:19 was mapped to chromosome 8 within the interval from 60.0 to 64.6 centiMorgans.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel (1995) *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is

much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding REPTR are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed.

VIII. Extension of REPTR Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham

Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

5 The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA
10 recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

15 In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:13-24 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base
20 pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a
25 SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

30 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

35 X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated

using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

5 Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia
10 Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water,
15 and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic
20 apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in
25 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered
30 with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X

SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The

software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the REPTR-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring REPTR. Although use of
5 oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of REPTR. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a
10 complementary oligonucleotide is designed to prevent ribosomal binding to the REPTR-encoding transcript.

XII. Expression of REPTR

Expression and purification of REPTR is achieved using bacterial or virus-based expression systems. For expression of REPTR in bacteria, cDNA is subcloned into an appropriate vector
15 containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express REPTR upon induction with isopropyl beta-D-
20 thiogalactopyranoside (IPTG). Expression of REPTR in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding REPTR by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong
25 polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

30 In most expression systems, REPTR is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham
35 Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from

REPTR at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified REPTR obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

XIII. Functional Assays

REPTR function is assessed by expressing the sequences encoding REPTR at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of REPTR on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding REPTR and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding REPTR and other genes of interest can be analyzed by

northern analysis or microarray techniques.

XIV. Production of REPTR Specific Antibodies

REPTR substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to
5 immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the REPTR amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well
10 described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the
15 oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-REPTR activity by, for example, binding the peptide or REPTR to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring REPTR Using Specific Antibodies

20 Naturally occurring or recombinant REPTR is substantially purified by immunoaffinity chromatography using antibodies specific for REPTR. An immunoaffinity column is constructed by covalently coupling anti-REPTR antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

25 Media containing REPTR are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of REPTR (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/REPTR binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and REPTR is collected.

30 XVI. Identification of Molecules Which Interact with REPTR

REPTR, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled REPTR, washed, and any wells with labeled REPTR complex are assayed. Data obtained using different
35 concentrations of REPTR are used to calculate values for the number, affinity, and association of

REPTR with the candidate molecules.

Alternatively, molecules interacting with REPTR are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

5 REPTR may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of REPTR Activity

10 REPTR activity is measured by combining a purified epitope-tagged sample with a selected radiolabeled REPTR ligand. Ligands for SEQ ID NO:1 include acetylated low density lipoprotein (Ashkenas, J. et al. (1993) *J. Lipid Res.* 34:983-1000). Ligands for SEQ ID NO:11 include OX (Wright, G. J. (2000) *Immunity* 13:233-242). Ligands for SEQ ID NO:3 include complement proteins C3 and C5 (Tausk, F. and Gigli, I. (1990) *J. Invest. Dermatol.* 94:141S-145S). REPTR/ligand
15 complexes are recovered by immunoprecipitation with a commercial antibody against the epitope. REPTR activity is proportional to the amount of ligand bound.

Alternatively, REPTR activity is measured by phosphorylation of a protein substrate using γ -labeled [^{32}P]-ATP and quantitation of the incorporated radioactivity using a radioisotope counter. REPTR is incubated with the protein substrate, [^{32}P]-ATP, and an appropriate kinase buffer. The
20 [^{32}P] incorporated into the product is separated from free [^{32}P]-ATP by electrophoresis and the incorporated [^{32}P] is counted. The amount of [^{32}P] recovered is proportional to the activity of REPTR in the assay. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

In the alternative, REPTR activity is measured by the increase in cell proliferation resulting
25 from transformation of a mammalian cell line such as COS7, HeLa or CHO with an eukaryotic expression vector encoding REPTR. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression of REPTR. Phase microscopy is then used to compare the mitotic index of transformed
30 versus control cells. An increase in the mitotic index indicates REPTR activity.

An assay for REPTR activity measures the expression of REPTR on the cell surface. cDNA encoding REPTR is subcloned into an appropriate mammalian expression vector suitable for high levels of cDNA expression. The resulting construct is transfected into a nonhuman cell line such as NIH3T3. Cell surface proteins are labeled with biotin using methods known in the art. Precipitations
35 are performed using streptavidin-coated beads; precipitated and total cellular protein samples are then

analyzed using SDS-PAGE and blotting techniques. The ratio of biotin-labeled precipitant to the total amount of REPTR expressed in the cell is proportional to the amount of REPTR expressed on the cell surface.

In a further alternative, an assay for REPTR activity is based upon the ability of GPCR
5 family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) J. Biol. Chem. 273:4990-4996). A plasmid encoding full length REPTR is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the
10 attached cells are gently washed with PBS. The cells are then incubated in culture medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand compared to those without ligand are proportional to the amount of REPTR present in the transfected
15 cells.

An alternative assay for REPTR activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the amount of newly synthesized DNA in Swiss mouse 3T3 cells expressing REPTR. An appropriate mammalian expression vector containing cDNA encoding REPTR is added to quiescent 3T3 cultured cells using transfection
20 methods well known in the art. The transfected cells are incubated in the presence of [³H]thymidine and varying amounts of REPTR ligand. Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a tritium radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold REPTR ligand concentration range is indicative of
25 receptor activity. One unit of activity per milliliter is defined as the concentration of REPTR producing a 50% response level, where 100% represents maximal incorporation of [³H]thymidine into acid-precipitable DNA (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY; p. 73).

Alternatively, an assay for REPTR activity measures the effect of REPTR expression on the
30 regulation of cell growth. To demonstrate that increased levels of REPTR expression correlates with decreased cell motility and increased cell proliferation, expression vectors encoding REPTR are electroporated into highly motile cell lines, such as U-937 (ATCC CRL 1593), HEL 92.1.7 (ATCC TIB 180) and MAC10, and the motility of the electroporated and control cells are compared. Methods for the design and construction of an expression vector capable of expressing REPTR in the
35 desired mammalian cell line(s) chosen are well known to the art. Assays for examining the motility

of cells in culture are known to the art (cf Miyake, M. et al. (1991) J. Exp. Med. 174:1347-1354 and Ikeyama, S. et al. (1993) J. Exp. Med. 177:1231-1237). Increasing the level of REPTR in highly motile cell lines by transfection with an REPTR expression vector inhibits or reduces the motility of these cell lines, and the amount of this inhibition is proportional to the activity of REPTR in the
5 assay.

Alternatively, an assay for cadherin activity measures the expression of REPTR on the cell surface. cDNA encoding REPTR is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using REPTR-specific antibodies, and immunoprecipitated samples are analyzed using
10 SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of REPTR expressed on the cell surface.

Alternatively, an assay for REPTR activity measures the amount of cell aggregation induced by overexpression of REPTR. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding REPTR contained within a suitable mammalian expression vector under control of a
15 strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of REPTR activity.

Various modifications and variations of the described methods and systems of the invention
20 will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the
25 scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
6052371	1	6052371CD1	13	6052371CB1
2642942	2	2642942CD1	14	2642942CB1
3798924	3	3798924CD1	15	3798924CB1
4586653	4	4586653CD1	16	4586653CB1
5951460	5	5951460CD1	17	5951460CB1
1534444	6	1534444CD1	18	1534444CB1
6777669	7	6777669CD1	19	6777669CB1
1897612	8	1897612CD1	20	1897612CB1
6977010	9	6977010CD1	21	6977010CB1
926992	10	926992CD1	22	926992CB1
1002055	11	1002055CD1	23	1002055CB1
3998749	12	3998749CD1	24	3998749CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	6052371CD1	g2055392	0	transmembrane receptor UNC5H1 [Rattus norvegicus] Leonardo, E. D. et al. (1997) Nature 386:833-838.
2	2642942CD1	g439296	3.80E-82	[Homo sapiens] garp Ollendorff, V. et al. (1994) Cell Growth Differ. 5(2):213-219.
3	3798924CD1	g6683905	9.50E-106	Dispatched [Drosophila melanogaster] Burke, R. et al. (1999) Cell 99(7):803-815.
4	4586653CD1	g57734	1.20E-137	potential ligand-binding protein [Rattus rattus] Dear, T. N. et al. (1991) EMBO J. 10(10):2813-2819.
5	5951460CD1	g1387996	1.30E-87	lens intrinsic membrane protein 19 [Rattus norvegicus] Church, R. L. and Wang, J. H. (1993) Curr. Eye Res. 12(12):1057-1065.
6	1534444CD1	g1151260	0	Transmembrane receptor [Mus musculus] Wang, Y. et al. (1996) J. Biol. Chem. 271:4468-4476
7	6777669CD1	g3800736	1.20E-21	Seven-pass transmembrane receptor precursor [Mus musculus] Hadjantonakis, A.K. et al. (1997) Genomics 45:97-104
8	1897612CD1	g4887715	0	Adherin [Drosophila melanogaster] Clark, H.F. et al. (1995) Genes Dev. 9:1530-1542
9	6977010CD1	g5832711	0	Flamingo 1 [Mus musculus] Usui, T. et al. (1999) Cell 98:585-595
10	926992CD1	g293746	2.70E-65	[Mus musculus] macrophage scavenger receptor type I Ashkenas, J. et al. (1993) J. Lipid Res. 34:983-1000
11	1002055CD1	g9796480	9.60E-93	[Rattus norvegicus] OX2 receptor precursor Wright, G.J. (2000) Immunity 13:233-242
12	3998749CD1	g451303	4.80E-107	[Homo sapiens] complement receptor 1 Vik, D.P. and Wong, W.W. (1993) J. Immunol. 151:6214- 6224

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	6052371CD1	842	S137 S232 S298 S352 S356 S389 S417 S532 S539 S543 S655 S671 S700 S765 S838 T134 T281 T604 T684 T836 Y219	N107 N218 N287 N441 N682 N725 N816	TRANSMEMBRANE RECEPTOR UNC5: PD011882:W544-C842 signal_peptide:M1-A25 signal_cleavage:M1-A25 transmem_domain:Y306-V326 ZU5 domain:T439-G542 Receptor_Cytokines_2: G243-S249, S246-S252	BLAST_PRODOM HMMER SPSCAN HMMER HMMER_PFAM MOTIFS
2	2642942CD1	692	S158 S175 S295 S317 S323 S403 S447 S454 S463 S517 S569 S624 T244 T379 T429 T488 T612 T673	N155 N21 N232 N292 N309 N312 N408 N427 N500 N622 N74	Leucine-rich repeat signature: PR00019A:L378-L391 PR00019B:F535-L548 signal_peptide:M1-R20 transmem_domain:L653-T673 Leucine Rich Repeat (LRR): E251-S272, K273-S294, D329-P352, S353-G376, A377-G402, S403-R426, N427-S447, S463-S486, N537-L558, A559-L582, L82-G105, H106-P132, G133-S157, S158-E181, R182-A205, E206-R227 Leucine Zipper: L48-L69, L492-L513	BLIMPS_PRINTS HMMER HMMER HMMER_PFAM
3	3798924CD1	1124	S394 S557 S56 S569 S579 S761 S763 S785 S872 S980 T184 T19 T38 T455 T519 T58 T803 T808 T809 T843 T85 Y792 T1005 S1023 S1044 T1110	N159 N182 N226 N280 N436 N517 N76 N1078	transmem_domain:L100-Y121, F152-G170, M599-F621, L708-F727	HMMER

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
4	4586653CD1	419	S152 S165 S351 S365 S404 T144 T316 T329	N142 N288 N400	LIGAND BINDING PROTEIN RYA3: DM05385 S17448 1-473:V4-L347 LIGAND BINDING PROTEIN RYA3: PD177882:F86-F261 signal_peptide:M1-P20 transmem_domain:L213-L235 signal_cleavage:M1-A18 BY SIMILARITY TRANSMEM: DM02609 P20274 1-172:M1-R173 LENS FIBER INTRINSIC MEMBRANE: PD152448:M1-R173 PMP-22/EMP/MP20 family: BL01221A:M1-W28 BL01221B:A38-C51 BL01221C:A59-I103 BL01221D:F136-R162 transmem_domain:M1-L19, F67-A85, M104-T123, I141-C159 PMP-22/EMP/MP20/Claudin family: PMP22_Claudin:M1-Y157 Signal_cleavage: M1-G24 Signal peptide: M1-A25 Transmembrane domain: L283-M303, V398-L417, V490-F508, F586-W605 Frizzled/Smoothed family membrane region Frizzled:P267-A623 Frizzled domain Fz: C35-M149 FRIZZLED PROTEIN SIGNATURE PR00489: W280-D302, Y308-R330, V398-F422, F441-G464, L486-F508, L529-C550, V585-W605	BLAST_DOMO BLAST_PRODUM HMMER HMMER SPSCAN BLAST_DOMO BLAST_PRODUM BLIMPS_BLOCKS HMMER HMMER_PFAM SPSCAN HMMER HMMER HMMER_PFAM HMMER_PFAM BLIMPS_PRINTS
5	5951460CD1	173	S170 T171	N62		
6	1534444CD1	694	S109 S243 S270 S29 S450 S460 S606 S612 T523 T61 T674	N152 N475 N49		

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6					FRIZZLED, FZ-1 DM03929: A45054 54-474:C35-G194, C396-G474, P208-S341, A188-P230, P345-G366, G196-Q233 P18537 1-415:S11-R165, G374-G474, V252-C340 DM05386: A45054 476-641:S477-G626 TRANSMEMBRANE PROTEIN FRIZZLED HOMOLOG PD003058:E389-G626, PD001435:C35-M149, PD003033:H182-S341	BLAST_DOMO
7	6777669CD1	1331	S101 S224 S246 S33 S331 S36 S393 S449 S62 S691 S790 S794 S969 S990 T591 T653 T858 S1034 S1218 S1072 S1109 S1226 T1244 S1272 S1283	N155 N200 N268 N329 N429 N595 N652 N683 N730 N77 N787 N94	Rgd cell attachment sequence R355-D357 Signal peptide: M1-G26 HMM score 23.98 Signal_cleavage:M1-G26 score 11.4 Transmembrane domain: P768-L786, P876-Y903, S917-L937 Leucine Rich Repeat LR:S597-P623, G78-S101, L102-G125, E126-P149, R150-P173 7 Transmembrane receptor (secretin family) 7tm_2: L762-V1069 Leucine rich repeat C-terminal domain LRRCT: E183-E233 Molluscan rhodopsin C-terminal tail PR00239E: P569-P580 EMR1 7tm receptor DM05221 A57172 465-886:P701-L937, F1020-G1099 TRANSMEMBRANE GPROTEIN COUPLED RECEPTOR PD000752: A771-L937	MOTIFS HMMER SPSCAN HMMER HMMER_PFAM HMMER_PFAM HMMER_PFAM BLIMPS_PRINTS BLAST_DOMO BLAST_PRODOR

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
8	1897612CD1	3217	S68 T231 S267 S286 S343 S345 S369 T397 T450 S540 T555 T557 S605 T612 T688 S752 S1021 T1073 S1235 T1283 S1386 S1505 S1557 T1593 S1694 T1763 S1826 S1849 T1868 T1930 T1995 S2027 T2079 T2218 S2318 T2320 S2336 T2426 T2510 T2529 S2546 S2561 S2633 S2663 S2682 T2723 T2757 T2809 T2998 S3205 S3 T90 T117 T177 T450 T594 S727 T791 T858 T947 T1033 S1045 T1073 S1257	N204 N243 N356 N538 N1192 N1637 N1915 N2280 N2347 N2488 N2680 N2711 N2781	Cadherin motif: V118-P128, V230-P240, L414-P424, V520-P530, V627-P637, V825-P835, V1142-P1152, V1658-P1668, V1762-P1772, I1867-P1877, L2078-P2088, V2184-P2194, V2283-P2293, V2389-P2399, V2509-P2519, I2613-P2623 Signal peptide: M1-G29 Signal_cleavage:M1-G29 Transmembrane domain: H7-W28, L2855-L2877 Cadherin: cadherin.prf I501-F551, T2490-L2540, I1744-L1793, V1535-L1587, V1123-V1173, V396-F445, L212-V261, D2368-L2420, T1845-T1897 Cadherin domain: Y537-S630, L2630-T2723, Y644-V735, D842-Q932, F2737-T2842, Y749-D828, Y948-L1039, T1053-L1145, L1165-L1255, L1280-E1372, F1469-A1559, Y1573-E1661, L1675-L1765, Y1779-R1870, P1893-Q1978, S1992-Q2081, Y2095-I2187, W2200-E2286, Y2300-Q2392, Y2406-L2512, Y2526-Q2616, L34-A121, A135-L233, Y247-D324, G325-T417, R432-Q523 Cadherin extracellular repeat BL00232B: V2504-G2551 Cadherin signature PR:002058 S1135-P1152	MOTIFS HMMER SPSCAN HMMER PROFILES CAN HMMER_PFAM BLIMPS_BLOCKS BLIMPS_PRINTS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
8			T1426 S1500 S1541 T1589 T1624 T1655 T1728 T1868 T1904 T1995 S2008 T2044 S2056 T2169 T2320 S2525 S1328 T1382 T2577 S2699 T2809 S2844 S2950 S3153		CADHERIN REPEAT DM00030 P33450 187-298: T164-D270, Y2439-D2549, L384-D455 ADHERIN CELL ADHESION GLYCOPROTEIN TRANSMEMBRANE CALCIUM BINDING REPEAT PD138796: L1055-L1165, A2976-E3108, F2737-E2959, P2309-D2375, L436-L522, S540-D601, D2522-D2600, F1469-L1526, R246-A381, L1165-L1217, G2211-E2259, L845-D915, S1676-D1748, V1787-G1828, A1907-L1962, A2000-K2031, P653-V698, P1582-D1644, R2089-I2148	BLAST_DOMO BLAST_PRODOM
9	6977010CD1	2936	S114 S148 S163 S203 S298 S336 S389 S401 S425 S461 S634 S736 S824 S839 T131 T190 T243 T244 T261 T311 T315 T423 T467 T472 T514 T600 T612 T617 T687 T692 T70 T708 T738 T770 T787 T800 T841 T876 T904 T911 T945 T947 Y307 S1849 T1160 Y1311 S2469 T1288 S1079 T1869 S1432 T1903 S1781 T2038 S1793 T2738 S2054 T2797 S2250 T2817 S2295	N487 N558 N702 N1037 N1077 N1183 N1213 N1828 N1502 N1901 N1566 N2033 N1742 N2052 N2332 N2354 N2434	Cadherin motif I278-P288, L388-P398, V494-P504, V599- P609, I804-P814, V910-P920, V1012-P1022 EGF motif C1275-C1286, C1313-C1324, C1599-C1610, C1818-C1829, C1856-C1867, C1944-C1955 Signal cleavage: M1-G32 Signal peptide: M1-G32 Transmembrane domain: V2406-I2423, P2584-L2601 Cadherins extracellular repeated domain signature cadherin.prf: A785-V835, F576-V630, A894-V941, V260-V309, T367-V419, T472-V525 7 transmembrane receptor (Secretin family) 7tm_2:I2393-V2636 Cadherin domain cadherin: Y187-T281, Y295-E391, Y405-L497, F511- L602, Y616-T704, Y718-N807, Y821-L913, F927-L1015 EGF-like (extracellular) domain EGF: C1293-C1324 C1333-C1366, C1579-C1610, C1798-C1829, C1833-C1867	MOTIFS MOTIFS SPSCAN HMMER HMMER PROFILESAN HMMER_PFAM HMMER_PFAM HMMER_PFAM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9			T2866 S2375 S2534 T2926 S2534 S2740 T1036 S2745 T1131 S2760 T1369 S2762 T1369 S2762 T1380 S2889 T1845 S2907 T2099 S2932 T2655 S1327 T2689 S1330 S1514 S1687 S2058 S2278 S2921 S2344 S2084 S2868 S2645 S2692		Laminin G (extracellular) domain: F1396-Y1460, C1505-D1558, C1579-C1610, F1645-H1702, V1745-G1774 Latrophilin/CL-1-like GPS domain (exocytosis GPCR) GPS: T2324-R2377 Cadherins extracellular repeat BL00232B: P905-G952 G-protein coupled receptor BL00649: A2403-L2448, C2459-L2484, G2506-F2530, C2619-C2644 Type II EGF-like signature PR00010C: G1309-Y1319 Type III EGF-like signature PR00011: N1801-C1829, C1937-C1955 CADHERIN SIGNATURE PR00205:Q779-P794, S797-P814, I838-F852 CALCIUM-BINDING PRECURSOR PD00919: C1579-C1590, V1326-N1340, V1337-C1366, S1810-P1860, Y1816-C1844, D222-Q263 7tm receptor EMR1 DM05221 I37225 347-738:R2323-C2644 CADHERIN REPEAT DM00030 P08641 189-298: R850-E951, A539-D639 SEVENPASS TRANSMEMBRANE RECEPTOR PRECURSOR PD183649:L2560-E2935 PD155621:M1614-L1794 TRANSMEMBRANE CELL ADHESION CALCIUM BINDING REPEAT PRECURSOR PD017898: L1024-E1325 TRANSMEMBRANE GPROTEIN COUPLED PD000752: I2394-K2641, P1848-E1906, W1594-I1635, C1985-E2053	HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLIMPS_BLOCKS BLIMPS_PRINTS BLIMPS_PRINTS BLIMPS_PRINTS BLIMPS_PRODOM BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10	926992CD1	437	S28 S29 S46 S327 T201 T358 T387 T394 Y385	N44 N76 N135 N173 N196 N242 N339	Signal cleavage: M1-S24 Signal peptide: M1-R25 Transmembrane domain: L9-S29 Collagen triple helix repeat (20 copies): G257-K316 Scavenger receptor cysteine-rich domain: V338-N435 Speract (scavenger) receptor repeat proteins domain proteins BL00420: G251-E279, N339-G393, C424-C434 C1q domain (complement system activation) protein: BL01113 G266-D292 Speract receptor repeated domain signature speract_receptor.prf: G320-M399 Speract receptor signature PR00258: I335-Y351, R354-D365, G369-R379, D400-C414, N423-N435 ANTIGEN PRECURSOR SIGNAL M130 TRANSMEMBRANE GLYCOPROTEIN REPEAT VARIANT CYTOPLASMIC PROTEIN PD000767: V338-N435 PRECURSOR SIGNAL COLLAGEN ALPHA 3IX CHAIN EXTRACELLULAR MATRIX CONNECTIVE TISSUE PD028299: K247-R322 COLLAGEN ALPHA PRECURSOR CHAIN REPEAT SIGNAL CONNECTIVE TISSUE EXTRACELLULAR MATRIX PD000007: A246-D321 SIMILAR TO CUTICULAR COLLAGEN PD067228: G248-D325 SPERACT RECEPTOR DOMAIN DM04833 P21758 1-345: L127-G314 P30204 1-349: Q101-G324 DM00148 P21758 347-452: I335-C434 P21757 345-450: I335-C434	SPSCAN HMMER HMMER HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLIMPS_BLOCKS PROFILESAN BLIMPS_PRINTS BLAST_PRODUM BLAST_PRODUM BLAST_PRODUM BLAST_PRODUM BLAST_PRODUM

91

[illegible]

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
13	6052371CB1	3580	1-168, 2259-2867, 3543-3580	3974950F6 (ADRETUT06)	3302	3562
				5015170F8 (BRAXNOT03)	326	885
				92229606	3039	3580
				70484939V1	2678	3190
				6052371J1 (BRABDIR03)	651	1213
				4019505F7 (BRAXNOT01)	1	341
				6989724H1 (BRAIFER05)	1248	1888
				70485324V1	2141	2791
				70485074V1	2882	3432
				4783281H1 (BRATNOT03)	291	552
				5960193H1 (BRATNOT05)	1602	2156
				70482468V1	2139	2656
				7745508J1 (ADRETUE04)	872	1614
14	2642942CB1	2429	1-31, 936- 960, 2334- 2361, 2407- 2429	70844253V1	1219	1885
				71189744V1	1124	1686
				71189259V1	584	1193
				71191574V1	461	956
				7170082H1 (MCLRNOC01)	1	544
				4099778T6 (BRAITUT26)	1795	2429
				71412118V1	2625	3322
				71413432V1	2460	3187
				4368556F6 (THYMNOT11)	1	587
				7216513H1 (LUNGFEC01)	3330	3933
15	3798924CB1	3934	1-882, 2703-3367, 1624-1670	71412435V1	1184	1866
				5028029H1 (COLCDIT01)	3782	3934
				3798924T6 (SPLNNOT12)	3274	3904
				3798924F6 (SPLNNOT12)	1347	1907
				70796677V1	1952	2565
				6479006H1 (PROSTMC01)	700	1345
				71412140V1	1778	2502
				4368556T6 (THYMNOT11)	259	935
				70466381V1	744	1255
				70477595V1	1074	1633
16	4586653CB1	1633	1-164, 309- 1633	70467133V1	558	1188
				8010625H1 (NOSEDIC02)	1	628

Table 4 (cont.)

17	5951460CB1	879	1-54	5289452F8 (LIVRTUS02)	67	682
				5951460F6 (LIVRTUN04)	1	458
				95113551	471	879
18	1534444CB1	2085	1-108, 1785-2085, 473-1149	70682068V1	2074	2085
				FL023814_00001	1	2085
19	6777669CB1	5497	1-4052, 4577-4653	70691012V1	4720	5321
				8096141H1 (EYERNOA01)	2303	2988
				8020818J1 (BMARTXE01)	791	1336
				7612480J1 (KIDCTME01)	1774	2326
				7016962H1 (KIDNNOC01)	4181	4798
				7356264H1 (HEARNON03)	4111	4674
				6810945H1 (SKIRNOR01)	3680	4186
				70686433V1	4876	5411
				7643339J1 (SEMTDE01)	546	1297
				7735364J1 (BRAITUE01)	2405	3065
				7635088H1 (SINTDIE01)	3115	3666
				7663664H1 (UTRSTME01)	1309	1914
				7724763J1 (THYRDIE01)	3435	4115
				70688492V1	5087	5497
				GNN.g5926688_010.edit	1	2279
				6777669H1 (OVARDIR01)	181	720
20	1897612CB1	10123	1-4592, 5177-5497	7171221H1 (BRSTTMC01)	2986	3260
				6776039R8 (OVARDIR01)	6498	7257
				1349048F1 (LATRTUT02)	9356	9826
				4426155H1 (BRAPDIT01)	8859	9126
				7070373H1 (BRAUTDR02)	1253	1599
				70159017V1	7417	8055
				7035625H1 (SINTFER03)	5691	6303
				6782025H1 (OVARDIR01)	3204	3861
				3604927H1 (LUNGNOT30)	8729	9066
				7404288H1 (UTREDME05)	2379	2807
				70157736V1	7180	7707
				7440469H1 (ADRETUE02)	545	1086
				8067830J1 (BRAIFEE05)	4641	5194
				71763526V1	5263	5790

Table 4 (cont.)

				2349726F6 (COLSUCT01)	9107	9798
				6456564H1 (COLNDIC01)	7926	8533
				6777080J1 (OVARDIR01)	8073	8850
				1897612F6 (BLADTUT06)	5824	6431
				1456075R1 (COLNFET02)	9602	10108
				5512895F6 (BRADDIR01)	1462	2013
				8068573J1 (BRAIFEE05)	2697	3454
				2255632R6 (OVARTUT01)	9912	10123
				6984134F8 (BRAIFER05)	4979	5723
				7724578J1 (THYRDIE01)	3516	4051
				4756468F6 (BRAHNOT01)	970	1459
				7647279J1 (UTRSTUE01)	1	664
				GNN.g8570385_000017_00	1	7042
				2.edit		
				70986678V1	1647	2240
				7261410H1 (UTRETM01)	6391	7047
21	6977010CB1	9321	1986-5341, 1-1324, 5876-6986, 8509-9321, 7357-7388	7069926H1 (BRAUTDR02)	8729	9321
				7013707H1 (KIDNNOC01)	8412	9058
				6950239H1 (BRAITDR02)	1	681
				FL6977010_g8176711_000	223	9030
				001_g5832711		
22	926992CB1	3900	1232-2651, 3284-3900, 1-162, 3179-3236	926992R1 (BRAINOT04)	2185	2789
				7256460H2 (SKIRTD01)	1	474
				3084755H1 (HEAONOT03)	3188	3342
				1960144R6 (BRSTNOT04)	1539	2052
				72150249D1	2540	3197
				4241654H1 (SYNWDIT01)	1769	2106
				1720922F6 (BLADNOT06)	1014	1591
				1995327R6 (BRSTTUT03)	3241	3900
				7751654H1 (HEAONOE01)	654	1341
				7722451H2 (THYRDIE01)	215	779
				1599092F6 (BLADNOT03)	1965	2551
				8176242H1 (FETANON01)	464	1026
23	1002055CB1	2076	64-598, 804-839, 1024-1753	71573380V1	1373	1971
				71231319V1	465	1114
				71573050V1	190	937

Table 4 (cont.)

				2810401F6 (BRSTNOT17)	1	411
				71570657V1	1492	2076
				702459T6 (SYNORAT03)	1063	1699
24	3998749CB1	3991	1-424, 855- 1155, 2485- 2612	60207650U1	3347	3991
				8243689H1 (BONEUNR01)	1	650
				7982690H1 (UTRSTMC01)	1360	2056
				8243689J1 (BONEUNR01)	690	1320
				7989604H1 (UTRCDIC01)	1955	2677
				623369R6 (PGANNOT01)	2732	3342
				55106555H1	1002	1816
				55142628J1	466	1212
				7006315H1 (COLNFEC01)	2891	3502
				6482765H1 (MIXDUNB01)	2197	2767

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
13	6052371CB1	BRABDIR03
14	2642942CB1	HEAONOT04
15	3798924CB1	SPLNNOT12
16	4586653CB1	NOSEDIC02
17	5951460CB1	LIVRTUS02
18	1534444CB1	SPLNNOT04
19	6777669CB1	THP1AZT01
20	1897612CB1	OVARDIR01
21	6977010CB1	BRAHTDR04
22	926992CB1	BRAITUT22
23	1002055CB1	SYNORAT03
24	3998749CB1	PLACNOB01

Table 6

Library	Vector	Library Description
BRABDIR03	pINCY	This random primed library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day for 40 years).
BRAHTDR04	PCDNA2.1	This random primed library was constructed using RNA isolated from archaecortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAITUT22	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the right frontal/parietal lobe of a 76-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a meningioma. Family history included senile dementia.
HEAONOT04	pINCY	Library was constructed using RNA isolated from aortic tissue removed from a 12-year-old Caucasian female, who died from a closed head injury.
LIVRTUS02	pINCY	This subtracted C3A liver tumor cell line tissue library was constructed using 6.4 million clones from a 3-methylcholthrene-treated hepatocyte library and was subjected to two rounds of subtraction hybridization with 1.72 million clones from an untreated C3A hepatocyte library. The starting library for subtraction was constructed using RNA isolated from a treated C3A hepatocyte cell line which is a derivative of Hep G2, a cell line derived from a hepatoblastoma removed from a 15-year-old Caucasian male. The cells were treated with 3-methylcholanthrene (MCA), 5 mM for 48 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated C3A hepatocyte cells from the same cell line. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6(1996):791.
NOSEDIC02	PSPORT1	This large size fractionated library was constructed using RNA isolated from nasal polyp tissue.
OVARDIR01	PCDNA2.1	This random primed library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, vaginal suspension and fixation, and incidental appendectomy. Pathology indicated stromal hyperthecosis of the

Table 6 (cont.)

Library	Vector	Library Description
		right and left ovaries. Pathology for the matched tumor tissue indicated a dermoid cyst (benign cystic teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified. The cervix showed squamous metaplasia. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, and primary tuberculous complex.
PLACNOB01	PBLUESCRIPT	Library was constructed using RNA isolated from placenta.
SPLNNOT04	pINCY	Library was constructed using RNA isolated from the spleen tissue of a 2-year-old Hispanic male, who died from cerebral anoxia. Past medical history and serologies were negative.
SPLNNOT12	pINCY	Library was constructed using RNA isolated from spleen tissue removed from a 65-year-old female. Pathology indicated the spleen was negative for metastasis. Pathology for the associated tumor tissue indicated well-differentiated neuroendocrine carcinoma (islet cell tumor), nuclear grade 1, forming a dominant mass in the distal pancreas.
SYNORAT03	PSPORT1	Library was constructed using RNA isolated from the wrist synovial membrane tissue of a 56-year-old female with rheumatoid arthritis.
THP1AZT01	pINCY	Library was constructed using RNA isolated from THP-1 promonocyte cells treated for three days with 0.8 micromolar 5-aza-2'-deoxycytidine. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (Int. J. Cancer (1980) 26:171).

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified “HIGH” value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

<110> INCYTE GENOMICS, INC.
 GRIFFIN, Jennifer A.
 KALLICK, Deborah A.
 TRIBOULEY, Catherine M.
 YUE, Henry
 NGUYEN, Danniell B.
 TANG, Y. Tom
 LAL, Preeti
 POLICKY, Jennifer L.
 AZIMZAI, Yalda
 LU, Dyung Aina M.
 GRAUL, Richard
 YAO, Monique G.
 BURFORD, Neil
 HAFALIA, April J. A.
 BAUGHN, Mariah R.
 BANDMAN, Olga
 PATTERSON, Chandra
 YANG, Junming
 XU, Yuming
 GANDHI, Ameena R.
 WARREN, Bridget A.
 DING, Li
 SANJANWALA, Madhu S.
 DUGGAN, Brendan M.
 LU, Yan

<120> RECEPTORS

<130> PF-0793 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/214,027; 60/228,045; 60/255,104

<151> 2000-06-21; 2000-08-25; 2000-12-12

<160> 24

<170> PERL Program

<210> 1

<211> 842

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6052371CD1

<400> 1

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

Arg	Ser	Thr	Asp	Gly	Ser	Ser	Gly	Leu	Pro	Thr	Met	Glu	Val	Arg
				95					100					105
Ile	Asn	Val	Ser	Arg	Gln	Gln	Val	Glu	Lys	Val	Phe	Gly	Leu	Glu
				110					115					120
Glu	Tyr	Trp	Cys	Gln	Cys	Val	Ala	Trp	Ser	Ser	Ser	Gly	Thr	Thr
				125					130					135
Lys	Ser	Gln	Lys	Ala	Tyr	Ile	Arg	Ile	Ala	Tyr	Leu	Arg	Lys	Asn
				140					145					150
Phe	Glu	Gln	Glu	Pro	Leu	Ala	Lys	Glu	Val	Ser	Leu	Glu	Gln	Gly
				155					160					165
Ile	Val	Leu	Pro	Cys	Arg	Pro	Pro	Glu	Gly	Ile	Pro	Pro	Ala	Glu
				170					175					180
Val	Glu	Trp	Leu	Arg	Asn	Glu	Asp	Leu	Val	Asp	Pro	Ser	Leu	Asp
				185					190					195
Pro	Asn	Val	Tyr	Ile	Thr	Arg	Glu	His	Ser	Leu	Val	Val	Arg	Gln
				200					205					210
Ala	Arg	Leu	Ala	Asp	Thr	Ala	Asn	Tyr	Thr	Cys	Val	Ala	Lys	Asn
				215					220					225
Ile	Val	Ala	Arg	Arg	Arg	Ser	Ala	Ser	Ala	Ala	Val	Ile	Val	Tyr
				230					235					240
Val	Asp	Gly	Ser	Trp	Ser	Pro	Trp	Ser	Lys	Trp	Ser	Ala	Cys	Gly
				245					250					255
Leu	Asp	Cys	Thr	His	Trp	Arg	Ser	Arg	Glu	Cys	Ser	Asp	Pro	Ala
				260					265					270
Pro	Arg	Asn	Gly	Gly	Glu	Glu	Cys	Gln	Gly	Thr	Asp	Leu	Asp	Thr
				275					280					285
Arg	Asn	Cys	Thr	Ser	Asp	Leu	Cys	Val	His	Thr	Ala	Ser	Gly	Pro
				290					295					300
Glu	Asp	Val	Ala	Leu	Tyr	Val	Gly	Leu	Ile	Ala	Val	Ala	Val	Cys
				305					310					315
Leu	Val	Leu	Leu	Leu	Leu	Val	Leu	Ile	Leu	Val	Tyr	Cys	Arg	Lys
				320					325					330
Lys	Glu	Gly	Leu	Asp	Ser	Asp	Val	Ala	Asp	Ser	Ser	Ile	Leu	Thr
				335					340					345
Ser	Gly	Phe	Gln	Pro	Val	Ser	Ile	Lys	Pro	Ser	Lys	Ala	Asp	Asn
				350					355					360
Pro	His	Leu	Leu	Thr	Ile	Gln	Pro	Asp	Leu	Ser	Thr	Thr	Thr	Thr
				365					370					375
Thr	Tyr	Gln	Gly	Ser	Leu	Cys	Pro	Arg	Gln	Asp	Gly	Pro	Ser	Pro
				380					385					390
Lys	Phe	Gln	Leu	Thr	Asn	Gly	His	Leu	Leu	Ser	Pro	Leu	Gly	Gly
				395					400					405
Gly	Arg	His	Thr	Leu	His	His	Ser	Ser	Pro	Thr	Ser	Glu	Ala	Glu
				410					415					420
Glu	Phe	Val	Ser	Arg	Leu	Ser	Thr	Gln	Asn	Tyr	Phe	Arg	Ser	Leu
				425					430					435
Pro	Arg	Gly	Thr	Ser	Asn	Met	Thr	Tyr	Gly	Thr	Phe	Asn	Phe	Leu
				440					445					450
Gly	Gly	Arg	Leu	Met	Ile	Pro	Asn	Thr	Gly	Ile	Ser	Leu	Leu	Ile
				455					460					465
Pro	Pro	Asp	Ala	Ile	Pro	Arg	Gly	Lys	Ile	Tyr	Glu	Ile	Tyr	Leu
				470					475					480
Thr	Leu	His	Lys	Pro	Glu	Asp	Val	Arg	Leu	Pro	Leu	Ala	Gly	Cys
				485					490					495
Gln	Thr	Leu	Leu	Ser	Pro	Ile	Val	Ser	Cys	Gly	Pro	Pro	Gly	Val
				500					505					510
Leu	Leu	Thr	Arg	Pro	Val	Ile	Leu	Ala	Met	Asp	His	Cys	Gly	Glu
				515					520					525
Pro	Ser	Pro	Asp	Ser	Trp	Ser	Leu	Arg	Leu	Lys	Lys	Gln	Ser	Cys
				530					535					540
Glu	Gly	Ser	Trp	Glu	Asp	Val	Leu	His	Leu	Gly	Glu	Glu	Ala	Pro
				545					550					555
Ser	His	Leu	Tyr	Tyr	Cys	Gln	Leu	Glu	Ala	Ser	Ala	Cys	Tyr	Val

WO 01/98354

PCT/US01/19942

	560		565		570
Phe Thr Glu Gln	Leu Gly Arg Phe Ala	Leu Val Gly Glu Ala	Leu		
	575		580		585
Ser Val Ala Ala	Ala Lys Arg Leu Lys	Leu Leu Leu Phe Ala	Pro		
	590		595		600
Val Ala Cys Thr	Ser Leu Glu Tyr Asn	Ile Arg Val Tyr Cys	Leu		
	605		610		615
His Asp Thr His	Asp Ala Leu Lys Glu	Val Val Gln Leu Glu	Lys		
	620		625		630
Gln Leu Gly Gly	Gln Leu Ile Gln Glu	Pro Arg Val Leu His	Phe		
	635		640		645
Lys Asp Ser Tyr	His Asn Leu Arg Leu	Ser Ile His Asp Val	Pro		
	650		655		660
Ser Ser Leu Trp	Lys Ser Lys Leu Leu	Val Ser Tyr Gln Glu	Ile		
	665		670		675
Pro Phe Tyr His	Ile Trp Asn Gly Thr	Gln Arg Tyr Leu His	Cys		
	680		685		690
Thr Phe Thr Leu	Glu Arg Val Ser Pro	Ser Thr Ser Asp Leu	Ala		
	695		700		705
Cys Lys Leu Trp	Val Trp Gln Val Glu	Gly Asp Gly Gln Ser	Phe		
	710		715		720
Ser Ile Asn Phe	Asn Ile Thr Lys Asp	Thr Arg Phe Ala Glu	Leu		
	725		730		735
Leu Ala Leu Glu	Ser Glu Ala Gly Val	Pro Ala Leu Val Gly	Pro		
	740		745		750
Ser Ala Phe Lys	Ile Pro Phe Leu Ile	Arg Gln Lys Ile Ile	Ser		
	755		760		765
Ser Leu Asp Pro	Pro Cys Arg Arg Gly	Ala Asp Trp Arg Thr	Leu		
	770		775		780
Ala Gln Lys Leu	His Leu Asp Ser His	Leu Ser Phe Phe Ala	Ser		
	785		790		795
Lys Pro Ser Pro	Thr Ala Met Ile Leu	Asn Leu Trp Glu Ala	Arg		
	800		805		810
His Phe Pro Asn	Gly Asn Leu Ser Gln	Leu Ala Ala Ala Val	Ala		
	815		820		825
Gly Leu Gly Gln	Pro Asp Ala Gly Leu	Phe Thr Val Ser Glu	Ala		
	830		835		840
Glu Cys					

<210> 2

<211> 692

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2642942CD1

<400> 2

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

				95					100				105	
His	Leu	Arg	Ser	Leu	Val	Leu	Gly	Asp	Asn	Cys	Leu	Ser	Glu	Asn
				110					115					120
Tyr	Glu	Glu	Thr	Ala	Ala	Ala	Leu	His	Ala	Leu	Pro	Gly	Leu	Arg
				125					130					135
Arg	Leu	Asp	Leu	Ser	Gly	Asn	Ala	Leu	Thr	Glu	Asp	Met	Ala	Ala
				140					145					150
Leu	Met	Leu	Gln	Asn	Leu	Ser	Ser	Leu	Arg	Ser	Val	Ser	Leu	Ala
				155					160					165
Gly	Asn	Thr	Ile	Met	Arg	Leu	Asp	Asp	Ser	Val	Phe	Glu	Gly	Leu
				170					175					180
Glu	Arg	Leu	Arg	Glu	Leu	Asp	Leu	Gln	Arg	Asn	Tyr	Ile	Phe	Glu
				185					190					195
Ile	Glu	Gly	Gly	Ala	Phe	Asp	Gly	Leu	Ala	Glu	Leu	Arg	His	Leu
				200					205					210
Asn	Leu	Ala	Phe	Asn	Asn	Leu	Pro	Cys	Ile	Val	Asp	Phe	Gly	Leu
				215					220					225
Thr	Arg	Leu	Arg	Val	Leu	Asn	Val	Ser	Tyr	Asn	Val	Leu	Glu	Trp
				230					235					240
Phe	Leu	Ala	Thr	Gly	Gly	Glu	Ala	Ala	Phe	Glu	Leu	Glu	Thr	Leu
				245					250					255
Asp	Leu	Ser	His	Asn	Gln	Leu	Leu	Phe	Phe	Pro	Leu	Leu	Pro	Gln
				260					265					270
Tyr	Ser	Lys	Leu	Arg	Thr	Leu	Leu	Leu	Arg	Asp	Asn	Asn	Met	Gly
				275					280					285
Phe	Tyr	Arg	Asp	Leu	Tyr	Asn	Thr	Ser	Ser	Pro	Arg	Glu	Met	Val
				290					295					300
Ala	Gln	Phe	Leu	Leu	Val	Asp	Gly	Asn	Val	Thr	Asn	Ile	Thr	Thr
				305					310					315
Val	Ser	Leu	Trp	Glu	Glu	Phe	Ser	Ser	Ser	Asp	Leu	Ala	Asp	Leu
				320					325					330
Arg	Phe	Leu	Asp	Met	Ser	Gln	Asn	Gln	Phe	Gln	Tyr	Leu	Pro	Asp
				335					340					345
Gly	Phe	Leu	Arg	Lys	Met	Pro	Ser	Leu	Ser	His	Leu	Asn	Leu	His
				350					355					360
Gln	Asn	Cys	Leu	Met	Thr	Leu	His	Ile	Arg	Glu	His	Glu	Pro	Pro
				365					370					375
Gly	Ala	Leu	Thr	Glu	Leu	Asp	Leu	Ser	His	Asn	Gln	Leu	Ser	Glu
				380					385					390
Leu	His	Leu	Ala	Pro	Gly	Leu	Ala	Ser	Cys	Leu	Gly	Ser	Leu	Arg
				395					400					405
Leu	Phe	Asn	Leu	Ser	Ser	Asn	Gln	Leu	Leu	Gly	Val	Pro	Pro	Gly
				410					415					420
Leu	Phe	Ala	Asn	Ala	Arg	Asn	Ile	Thr	Thr	Leu	Asp	Met	Ser	His
				425					430					435
Asn	Gln	Ile	Ser	Leu	Cys	Pro	Leu	Pro	Ala	Ala	Ser	Asp	Arg	Val
				440					445					450
Gly	Pro	Pro	Ser	Cys	Val	Asp	Phe	Arg	Asn	Met	Ala	Ser	Leu	Arg
				455					460					465
Ser	Leu	Ser	Leu	Glu	Gly	Cys	Gly	Leu	Gly	Ala	Leu	Pro	Asp	Cys
				470					475					480
Pro	Phe	Gln	Gly	Thr	Ser	Leu	Thr	Tyr	Leu	Asp	Leu	Ser	Ser	Asn
				485					490					495
Trp	Gly	Val	Leu	Asn	Gly	Ser	Leu	Ala	Pro	Leu	Gln	Asp	Val	Ala
				500					505					510
Pro	Met	Leu	Gln	Val	Leu	Ser	Leu	Arg	Asn	Met	Gly	Leu	His	Ser
				515					520					525
Ser	Phe	Met	Ala	Leu	Asp	Phe	Ser	Gly	Phe	Gly	Asn	Leu	Arg	Asp
				530					535					540
Leu	Asp	Leu	Ser	Gly	Asn	Cys	Leu	Thr	Thr	Phe	Pro	Arg	Phe	Gly
				545					550					555
Gly	Ser	Leu	Ala	Leu	Glu	Thr	Leu	Asp	Leu	Arg	Arg	Asn	Ser	Leu
				560					565					570

WO 01/98354

PCT/US01/19942

Thr	Ala	Leu	Pro	Gln	Lys	Ala	Val	Ser	Glu	Gln	Leu	Ser	Arg	Gly
				575					580					585
Leu	Arg	Thr	Ile	Tyr	Leu	Ser	Gln	Asn	Pro	Tyr	Asp	Cys	Cys	Gly
				590					595					600
Val	Asp	Gly	Trp	Gly	Ala	Leu	Gln	His	Gly	Gln	Thr	Val	Ala	Asp
				605					610					615
Trp	Ala	Met	Val	Thr	Cys	Asn	Leu	Ser	Ser	Lys	Ile	Ile	Arg	Val
				620					625					630
Thr	Glu	Leu	Pro	Gly	Gly	Val	Pro	Arg	Asp	Cys	Lys	Trp	Glu	Arg
				635					640					645
Leu	Asp	Leu	Gly	Leu	Leu	Tyr	Leu	Val	Leu	Ile	Leu	Pro	Ser	Cys
				650					655					660
Leu	Thr	Leu	Leu	Val	Ala	Cys	Thr	Val	Ile	Val	Leu	Thr	Phe	Lys
				665					670					675
Lys	Pro	Leu	Leu	Gln	Val	Ile	Lys	Ser	Arg	Cys	His	Trp	Ser	Ser
				680					685					690
Val	Tyr													

<210> 3

<211> 1124

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3798924CD1

<400> 3

Met	Ala	Ala	Arg	Arg	Lys	Asp	Gln	Leu	Lys	Cys	Thr	Asn	Val	Pro
1				5					10					15
Arg	Lys	Cys	Thr	Lys	Tyr	Asn	Ala	Val	Tyr	Gln	Ile	Leu	His	Tyr
				20					25					30
Leu	Val	Asp	Lys	Asp	Phe	Met	Thr	Pro	Lys	Thr	Ala	Asp	Tyr	Ala
				35					40					45
Thr	Pro	Ala	Leu	Lys	Tyr	Ser	Met	Leu	Phe	Ser	Pro	Thr	Glu	Lys
				50					55					60
Gly	Glu	Ser	Met	Met	Asn	Ile	Tyr	Leu	Asp	Asn	Phe	Glu	Asn	Trp
				65					70					75
Asn	Ser	Ser	Asp	Gly	Val	Thr	Thr	Ile	Thr	Gly	Ile	Glu	Phe	Gly
				80					85					90
Ile	Lys	His	Ser	Leu	Phe	Gln	Asp	Tyr	Leu	Leu	Met	Asp	Thr	Val
				95					100					105
Tyr	Pro	Ala	Ile	Ala	Ile	Val	Ile	Val	Leu	Leu	Val	Met	Cys	Val
				110					115					120
Tyr	Thr	Lys	Ser	Met	Phe	Ile	Thr	Leu	Met	Thr	Met	Phe	Ala	Ile
				125					130					135
Ile	Ser	Ser	Leu	Ile	Val	Ser	Tyr	Phe	Leu	Tyr	Arg	Val	Val	Phe
				140					145					150
His	Phe	Glu	Phe	Phe	Pro	Phe	Met	Asn	Leu	Thr	Ala	Leu	Ile	Ile
				155					160					165
Leu	Val	Gly	Ile	Gly	Ala	Asp	Asp	Ala	Phe	Val	Leu	Cys	Asp	Val
				170					175					180
Trp	Asn	Tyr	Thr	Lys	Phe	Asp	Lys	Pro	His	Ala	Glu	Thr	Ser	Glu
				185					190					195
Thr	Val	Ser	Ile	Thr	Leu	Gln	His	Ala	Ala	Leu	Ser	Met	Phe	Val
				200					205					210
Thr	Ser	Phe	Thr	Thr	Ala	Ala	Ala	Phe	Tyr	Ala	Asn	Tyr	Val	Ser
				215					220					225
Asn	Ile	Thr	Ala	Ile	Arg	Cys	Phe	Gly	Val	Tyr	Ala	Gly	Thr	Ala
				230					235					240
Ile	Leu	Val	Asn	Tyr	Val	Leu	Met	Val	Thr	Trp	Leu	Pro	Ala	Val
				245					250					255

Val	Val	Leu	His	Glu	Arg	Tyr	Leu	Leu	Asn	Ile	Phe	Thr	Cys	Phe
				260					265					270
Lys	Lys	Pro	Gln	Gln	Gln	Ile	Tyr	Asp	Asn	Lys	Ser	Cys	Trp	Thr
				275					280					285
Val	Ala	Cys	Gln	Lys	Cys	His	Lys	Val	Leu	Phe	Ala	Ile	Ser	Glu
				290					295					300
Ala	Ser	Arg	Ile	Phe	Phe	Glu	Lys	Val	Leu	Pro	Cys	Ile	Val	Ile
				305					310					315
Lys	Phe	Arg	Tyr	Leu	Trp	Leu	Phe	Trp	Phe	Leu	Ala	Leu	Thr	Val
				320					325					330
Gly	Gly	Ala	Tyr	Ile	Val	Cys	Ile	Asn	Pro	Lys	Met	Lys	Leu	Pro
				335					340					345
Ser	Leu	Glu	Leu	Ser	Glu	Phe	Gln	Val	Phe	Arg	Ser	Ser	His	Pro
				350					355					360
Phe	Glu	Arg	Tyr	Asp	Ala	Glu	Tyr	Lys	Lys	Leu	Phe	Met	Phe	Glu
				365					370					375
Arg	Val	His	His	Gly	Glu	Glu	Leu	His	Met	Pro	Ile	Thr	Val	Ile
				380					385					390
Trp	Gly	Val	Ser	Pro	Glu	Asp	Asn	Gly	Asn	Pro	Leu	Asn	Pro	Lys
				395					400					405
Ser	Lys	Gly	Lys	Leu	Thr	Leu	Asp	Ser	Ser	Phe	Asn	Ile	Ala	Ser
				410					415					420
Pro	Ala	Ser	Gln	Ala	Trp	Ile	Leu	His	Phe	Cys	Gln	Lys	Leu	Arg
				425					430					435
Asn	Gln	Thr	Phe	Phe	Tyr	Gln	Thr	Asp	Glu	Gln	Asp	Phe	Thr	Ser
				440					445					450
Cys	Phe	Ile	Glu	Thr	Phe	Lys	Gln	Trp	Met	Glu	Asn	Gln	Asp	Cys
				455					460					465
Asp	Glu	Pro	Ala	Leu	Tyr	Pro	Cys	Cys	Ser	His	Trp	Ser	Phe	Pro
				470					475					480
Tyr	Lys	Gln	Glu	Ile	Phe	Glu	Leu	Cys	Ile	Lys	Arg	Ala	Ile	Met
				485					490					495
Glu	Leu	Glu	Arg	Ser	Thr	Gly	Tyr	His	Leu	Asp	Ser	Lys	Thr	Pro
				500					505					510
Gly	Pro	Arg	Phe	Asp	Ile	Asn	Asp	Thr	Ile	Arg	Ala	Val	Val	Leu
				515					520					525
Glu	Phe	Gln	Ser	Thr	Tyr	Leu	Phe	Thr	Leu	Ala	Tyr	Glu	Lys	Met
				530					535					540
His	Gln	Phe	Tyr	Lys	Glu	Val	Asp	Ser	Trp	Ile	Ser	Ser	Glu	Leu
				545					550					555
Ser	Ser	Ala	Pro	Glu	Gly	Leu	Ser	Asn	Gly	Trp	Phe	Val	Ser	Asn
				560					565					570
Leu	Glu	Phe	Tyr	Asp	Leu	Gln	Asp	Ser	Leu	Ser	Asp	Gly	Thr	Leu
				575					580					585
Ile	Ala	Met	Gly	Leu	Ser	Val	Ala	Val	Ala	Phe	Ser	Val	Met	Leu
				590					595					600
Leu	Thr	Thr	Trp	Asn	Ile	Ile	Ile	Ser	Leu	Tyr	Ala	Ile	Ile	Ser
				605					610					615
Ile	Ala	Gly	Thr	Ile	Phe	Val	Thr	Val	Gly	Ser	Leu	Val	Leu	Leu
				620					625					630
Gly	Trp	Glu	Leu	Asn	Val	Leu	Glu	Ser	Val	Thr	Ile	Ser	Val	Ala
				635					640					645
Val	Gly	Leu	Ser	Val	Asp	Phe	Ala	Val	His	Tyr	Gly	Val	Ala	Tyr
				650					655					660
Arg	Leu	Ala	Pro	Asp	Pro	Asp	Arg	Glu	Gly	Lys	Val	Ile	Phe	Ser
				665					670					675
Leu	Ser	Arg	Val	Gly	Ser	Ala	Met	Ala	Met	Ala	Ala	Leu	Thr	Thr
				680					685					690
Phe	Val	Ala	Gly	Ala	Met	Met	Met	Pro	Ser	Thr	Val	Leu	Ala	Tyr
				695					700					705
Thr	Gln	Leu	Gly	Thr	Phe	Met	Met	Leu	Ile	Met	Cys	Ile	Ser	Trp
				710					715					720
Ala	Phe	Ala	Thr	Phe	Phe	Phe	Gln	Cys	Met	Cys	Arg	Cys	Leu	Gly

WO 01/98354

PCT/US01/19942

Pro	Gln	Gly	Thr	725	Cys	Gly	Gln	Ile	Pro	730	Leu	Pro	Lys	Lys	Leu	Gln	735
				740						745							750
Cys	Ser	Ala	Phe	755	Ser	His	Ala	Leu	Ser	760	Thr	Ser	Pro	Ser	Asp	Lys	765
Gly	Gln	Ser	Lys	770	Thr	His	Thr	Ile	Asn	775	Ala	Tyr	His	Leu	Asp	Pro	780
Arg	Gly	Pro	Lys	785	Ser	Glu	Leu	Glu	His	790	Glu	Phe	Tyr	Glu	Leu	Glu	795
Pro	Leu	Ala	Ser	800	His	Ser	Cys	Thr	Ala	805	Pro	Glu	Lys	Thr	Thr	Tyr	810
Glu	Glu	Thr	His	815	Ile	Cys	Ser	Glu	Phe	820	Phe	Asn	Ser	Gln	Ala	Lys	825
Asn	Leu	Gly	Met	830	Pro	Val	His	Ala	Ala	835	Tyr	Asn	Ser	Glu	Leu	Ser	840
Lys	Ser	Thr	Glu	845	Ser	Asp	Thr	Gly	Ser	850	Ala	Leu	Leu	Gln	Pro	Pro	855
Leu	Glu	Gln	His	860	Thr	Val	Cys	His	Phe	865	Phe	Ser	Leu	Asn	Gln	Arg	870
Cys	Ser	Cys	Pro	875	Asp	Ala	Tyr	Lys	His	880	Leu	Asn	Tyr	Gly	Pro	His	885
Ser	Cys	Gln	Gln	890	Met	Gly	Asp	Cys	Leu	895	Cys	His	Gln	Cys	Ser	Pro	900
Thr	Thr	Ser	Ser	905	Phe	Val	Gln	Ile	Gln	910	Asn	Gly	Val	Ala	Pro	Leu	915
Lys	Ala	Thr	His	920	Gln	Ala	Val	Glu	Gly	925	Phe	Val	His	Pro	Ile	Thr	930
His	Ile	His	His	935	Cys	Pro	Cys	Leu	Gln	940	Gly	Arg	Val	Lys	Pro	Ala	945
Gly	Met	Gln	Asn	950	Ser	Leu	Pro	Arg	Asn	955	Phe	Phe	Leu	His	Pro	Val	960
Gln	His	Ile	Gln	965	Ala	Gln	Glu	Lys	Ile	970	Gly	Lys	Thr	Asn	Val	His	975
Ser	Leu	Gln	Arg	980	Ser	Ile	Glu	Glu	His	985	Leu	Pro	Lys	Met	Ala	Glu	990
Pro	Ser	Ser	Phe	995	Val	Cys	Arg	Ser	Thr	1000	Gly	Ser	Leu	Leu	Lys	Thr	1005
Cys	Cys	Asp	Pro	1010	Glu	Asn	Lys	Gln	Arg	1015	Glu	Leu	Cys	Lys	Asn	Arg	1020
Asp	Val	Ser	Asn	1025	Leu	Glu	Ser	Ser	Gly	1030	Gly	Thr	Glu	Asn	Lys	Ala	1035
Gly	Gly	Lys	Val	1040	Glu	Leu	Ser	Leu	Ser	1045	Gln	Thr	Asp	Ala	Ser	Val	1050
Asn	Ser	Glu	His	1055	Phe	Asn	Gln	Asn	Glu	1060	Pro	Lys	Val	Leu	Phe	Asn	1065
His	Leu	Met	Gly	1070	Glu	Ala	Gly	Cys	Arg	1075	Ser	Cys	Pro	Asn	Asn	Ser	1080
Gln	Ser	Cys	Gly	1085	Arg	Ile	Val	Arg	Val	1090	Lys	Cys	Asn	Ser	Val	Asp	1095
Cys	Gln	Met	Pro	1100	Asn	Met	Glu	Ala	Asn	1105	Val	Pro	Ala	Val	Leu	Thr	1110
His	Ser	Glu	Leu	1115	Ser	Gly	Glu	Ser	Leu	1120	Leu	Ile	Lys	Thr	Leu		

<210> 4

<211> 419

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4586653CD1

WO 01/98354

PCT/US01/19942

<400> 4

Met	Gln	Pro	Val	Met	Leu	Ala	Leu	Trp	Ser	Leu	Leu	Leu	Leu	Trp
1				5					10					15
Gly	Leu	Ala	Thr	Pro	Cys	Gln	Glu	Leu	Leu	Glu	Thr	Val	Gly	Thr
				20					25					30
Leu	Ala	Arg	Ile	Asp	Lys	Asp	Glu	Leu	Gly	Lys	Ala	Ile	Gln	Asn
				35					40					45
Ser	Leu	Val	Gly	Glu	Pro	Ile	Leu	Gln	Asn	Val	Leu	Gly	Ser	Val
				50					55					60
Thr	Ala	Val	Asn	Arg	Gly	Leu	Leu	Gly	Ser	Gly	Gly	Leu	Leu	Gly
				65					70					75
Gly	Gly	Gly	Leu	Leu	Gly	His	Gly	Gly	Val	Phe	Gly	Val	Val	Glu
				80					85					90
Glu	Leu	Ser	Gly	Leu	Lys	Ile	Glu	Glu	Leu	Thr	Leu	Pro	Lys	Val
				95					100					105
Leu	Leu	Lys	Leu	Leu	Pro	Gly	Phe	Gly	Val	Gln	Leu	Ser	Leu	His
				110					115					120
Thr	Lys	Val	Gly	Met	His	Cys	Ser	Gly	Pro	Leu	Gly	Gly	Leu	Leu
				125					130					135
Gln	Leu	Ala	Ala	Glu	Val	Asn	Val	Thr	Ser	Arg	Val	Ala	Leu	Ala
				140					145					150
Val	Ser	Ser	Arg	Gly	Thr	Pro	Ile	Leu	Ile	Leu	Lys	Arg	Cys	Ser
				155					160					165
Thr	Leu	Leu	Gly	His	Ile	Ser	Leu	Phe	Ser	Gly	Leu	Leu	Pro	Thr
				170					175					180
Pro	Leu	Phe	Gly	Val	Val	Glu	Gln	Met	Leu	Phe	Lys	Val	Leu	Pro
				185					190					195
Gly	Leu	Leu	Cys	Pro	Val	Val	Asp	Ser	Val	Leu	Gly	Val	Val	Asn
				200					205					210
Glu	Leu	Leu	Gly	Ala	Val	Leu	Gly	Leu	Val	Ser	Leu	Gly	Ala	Leu
				215					220					225
Gly	Ser	Val	Glu	Phe	Ser	Leu	Ala	Thr	Leu	Pro	Leu	Ile	Ser	Asn
				230					235					240
Gln	Tyr	Ile	Glu	Leu	Asp	Ile	Asn	Pro	Ile	Val	Lys	Ser	Val	Ala
				245					250					255
Gly	Asp	Ile	Ile	Asp	Phe	Pro	Lys	Ser	Arg	Ala	Pro	Ala	Lys	Val
				260					265					270
Pro	Pro	Lys	Lys	Asp	His	Thr	Ser	Gln	Val	Met	Val	Pro	Leu	Tyr
				275					280					285
Leu	Phe	Asn	Thr	Thr	Phe	Gly	Leu	Leu	Gln	Thr	Asn	Gly	Ala	Leu
				290					295					300
Asp	Met	Asp	Ile	Thr	Pro	Glu	Leu	Val	Pro	Ser	Asp	Val	Pro	Leu
				305					310					315
Thr	Thr	Thr	Asp	Leu	Ala	Ala	Leu	Leu	Pro	Glu	Val	Met	Thr	Val
				320					325					330
Arg	Ala	Gln	Leu	Ala	Pro	Ser	Ala	Thr	Lys	Leu	His	Ile	Ser	Leu
				335					340					345
Ser	Leu	Glu	Arg	Leu	Ser	Val	Lys	Val	Ala	Ser	Ser	Phe	Thr	His
				350					355					360
Ala	Phe	Asp	Gly	Ser	Arg	Leu	Glu	Glu	Trp	Leu	Ser	His	Val	Val
				365					370					375
Gly	Ala	Val	Tyr	Ala	Pro	Lys	Leu	Asn	Val	Ala	Leu	Asp	Val	Gly
				380					385					390
Ile	Pro	Leu	Pro	Lys	Val	Leu	Asn	Ile	Asn	Phe	Ser	Asn	Ser	Val
				395					400					405
Leu	Glu	Ile	Val	Glu	Asn	Ala	Val	Val	Leu	Thr	Val	Ala	Ser	
				410					415					

<210> 5

<211> 173

<212> PRT

<213> Homo sapiens

WO 01/98354

PCT/US01/19942

<220>
 <221> misc_feature
 <223> Incyte ID No: 5951460CD1

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

<210> 6
 <211> 694
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1534444CD1

<400> 6
 Met Glu Trp Gly Tyr Leu Leu Glu Val Thr Ser Leu Leu Ala Ala
 1 5 10 15
 Leu Ala Leu Leu Gln Arg Ser Ser Gly Ala Ala Ala Ala Ser Ala
 20 25 30
 Lys Glu Leu Ala Cys Gln Glu Ile Thr Val Pro Leu Cys Lys Gly
 35 40 45
 Ile Gly Tyr Asn Tyr Thr Tyr Met Pro Asn Gln Phe Asn His Asp
 50 55 60
 Thr Gln Asp Glu Ala Gly Leu Glu Val His Gln Phe Trp Pro Leu
 65 70 75
 Val Glu Ile Gln Cys Ser Pro Asp Leu Lys Phe Phe Leu Cys Ser
 80 85 90
 Met Tyr Thr Pro Ile Cys Leu Glu Asp Tyr Lys Lys Pro Leu Pro
 95 100 105
 Pro Cys Arg Ser Val Cys Glu Arg Ala Lys Ala Gly Cys Ala Pro
 110 115 120
 Leu Met Arg Gln Tyr Gly Phe Ala Trp Pro Asp Arg Met Arg Cys
 125 130 135
 Asp Arg Leu Pro Glu Gln Gly Asn Pro Asp Thr Leu Cys Met Asp
 140 145 150
 Tyr Asn Arg Thr Asp Leu Thr Thr Ala Ala Pro Ser Pro Pro Arg
 155 160 165
 Arg Leu Pro Pro Pro Pro Pro Gly Glu Gln Pro Pro Ser Gly Ser

				170					175					180
Gly	His	Gly	Arg	Pro	Pro	Gly	Ala	Arg	Pro	Pro	His	Arg	Gly	Gly
				185					190					195
Gly	Arg	Gly	Gly	Gly	Gly	Gly	Asp	Ala	Ala	Ala	Pro	Pro	Ala	Arg
				200					205					210
Gly	Gly	Gly	Gly	Gly	Gly	Lys	Ala	Arg	Pro	Pro	Gly	Gly	Gly	Ala
				215					220					225
Ala	Pro	Cys	Glu	Pro	Gly	Cys	Gln	Cys	Arg	Ala	Pro	Met	Val	Ser
				230					235					240
Val	Ser	Ser	Glu	Arg	His	Pro	Leu	Tyr	Asn	Arg	Val	Lys	Thr	Gly
				245					250					255
Gln	Ile	Ala	Asn	Cys	Ala	Leu	Pro	Cys	His	Asn	Pro	Phe	Phe	Ser
				260					265					270
Gln	Asp	Glu	Arg	Ala	Phe	Thr	Val	Phe	Trp	Ile	Gly	Leu	Trp	Ser
				275					280					285
Val	Leu	Cys	Phe	Val	Ser	Thr	Phe	Ala	Thr	Val	Ser	Thr	Phe	Leu
				290					295					300
Ile	Asp	Met	Glu	Arg	Phe	Lys	Tyr	Pro	Glu	Arg	Pro	Ile	Ile	Phe
				305					310					315
Leu	Ser	Ala	Cys	Tyr	Leu	Phe	Val	Ser	Val	Gly	Tyr	Leu	Val	Arg
				320					325					330
Leu	Val	Ala	Gly	His	Glu	Lys	Val	Ala	Cys	Ser	Gly	Gly	Ala	Pro
				335					340					345
Gly	Ala	Gly	Gly	Ala	Gly	Gly	Ala	Gly	Gly	Ala	Ala	Ala	Gly	Ala
				350					355					360
Gly	Ala	Ala	Gly	Ala	Gly	Ala	Gly	Gly	Pro	Gly	Gly	Arg	Gly	Glu
				365					370					375
Tyr	Glu	Glu	Leu	Gly	Ala	Val	Glu	Gln	His	Val	Arg	Tyr	Glu	Thr
				380					385					390
Thr	Gly	Pro	Ala	Leu	Cys	Thr	Val	Val	Phe	Leu	Leu	Val	Tyr	Phe
				395					400					405
Phe	Gly	Met	Ala	Ser	Ser	Ile	Trp	Trp	Val	Ile	Leu	Ser	Leu	Thr
				410					415					420
Trp	Phe	Leu	Ala	Ala	Gly	Met	Lys	Trp	Gly	Asn	Glu	Ala	Ile	Ala
				425					430					435
Gly	Tyr	Ser	Gln	Tyr	Phe	His	Leu	Ala	Ala	Trp	Leu	Val	Pro	Ser
				440					445					450
Val	Lys	Ser	Ile	Ala	Val	Leu	Ala	Leu	Ser	Ser	Val	Asp	Gly	Asp
				455					460					465
Pro	Val	Ala	Gly	Ile	Cys	Tyr	Val	Gly	Asn	Gln	Ser	Leu	Asp	Asn
				470					475					480
Leu	Arg	Gly	Phe	Val	Leu	Ala	Pro	Leu	Val	Ile	Tyr	Leu	Phe	Ile
				485					490					495
Gly	Thr	Met	Phe	Leu	Leu	Ala	Gly	Phe	Val	Ser	Leu	Phe	Arg	Ile
				500					505					510
Arg	Ser	Val	Ile	Lys	Gln	Gln	Asp	Gly	Pro	Thr	Lys	Thr	His	Lys
				515					520					525
Leu	Glu	Lys	Leu	Met	Ile	Arg	Leu	Gly	Leu	Phe	Thr	Val	Leu	Tyr
				530					535					540
Thr	Val	Pro	Ala	Ala	Val	Val	Val	Ala	Cys	Leu	Phe	Tyr	Glu	Gln
				545					550					555
His	Asn	Arg	Pro	Arg	Trp	Glu	Ala	Thr	His	Asn	Cys	Pro	Cys	Leu
				560					565					570
Arg	Asp	Leu	Gln	Pro	Asp	Gln	Ala	Arg	Arg	Pro	Asp	Tyr	Ala	Val
				575					580					585
Phe	Met	Leu	Lys	Tyr	Phe	Met	Cys	Leu	Val	Val	Gly	Ile	Thr	Ser
				590					595					600
Gly	Val	Trp	Val	Trp	Ser	Gly	Lys	Thr	Leu	Glu	Ser	Trp	Arg	Ser
				605					610					615
Leu	Cys	Thr	Arg	Cys	Cys	Trp	Ala	Ser	Lys	Gly	Ala	Ala	Val	Gly
				620					625					630
Gly	Gly	Ala	Gly	Ala	Thr	Ala	Ala	Gly	Gly	Gly	Gly	Gly	Pro	Gly
				635					640					645

WO 01/98354

PCT/US01/19942

Gly	Gly	Gly	Gly	Gly	Gly	Pro	Gly	Gly	Gly	Gly	Gly	Gly	Pro	Gly	Gly
				650					655						660
Gly	Gly	Gly	Ser	Leu	Tyr	Ser	Asp	Val	Ser	Thr	Gly	Leu	Thr	Trp	
				665					670						675
Arg	Ser	Gly	Thr	Ala	Ser	Ser	Val	Ser	Tyr	Pro	Lys	Gln	Met	Pro	
				680					685						690
Leu	Ser	Gln	Val												

<210> 7

<211> 1331

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6777669CD1

<400> 7

Met	Arg	Gly	Ala	Pro	Ala	Arg	Leu	Leu	Leu	Pro	Leu	Leu	Pro	Trp	
1				5					10					15	
Leu	Leu	Leu	Leu	Leu	Ala	Pro	Glu	Ala	Arg	Gly	Ala	Pro	Gly	Cys	
				20					25					30	
Pro	Leu	Ser	Ile	Arg	Ser	Cys	Lys	Cys	Ser	Gly	Glu	Arg	Pro	Lys	
				35					40					45	
Gly	Leu	Ser	Gly	Gly	Val	Pro	Gly	Pro	Ala	Arg	Arg	Arg	Val	Val	
				50					55					60	
Cys	Ser	Gly	Gly	Asp	Leu	Pro	Glu	Pro	Pro	Glu	Pro	Gly	Leu	Leu	
				65					70					75	
Pro	Asn	Gly	Thr	Val	Thr	Leu	Leu	Leu	Ser	Asn	Asn	Lys	Ile	Thr	
				80					85					90	
Gly	Leu	Arg	Asn	Gly	Ser	Phe	Leu	Gly	Leu	Ser	Leu	Leu	Glu	Lys	
				95					100					105	
Leu	Asp	Leu	Arg	Asn	Asn	Ile	Ile	Ser	Thr	Val	Gln	Pro	Gly	Ala	
				110					115					120	
Phe	Leu	Gly	Leu	Gly	Glu	Leu	Lys	Arg	Leu	Asp	Leu	Ser	Asn	Asn	
				125					130					135	
Arg	Ile	Gly	Cys	Leu	Thr	Ser	Glu	Thr	Phe	Gln	Gly	Leu	Pro	Arg	
				140					145					150	
Leu	Leu	Arg	Leu	Asn	Ile	Ser	Gly	Asn	Ile	Phe	Ser	Ser	Leu	Gln	
				155					160					165	
Pro	Gly	Val	Phe	Asp	Glu	Leu	Pro	Ala	Leu	Lys	Val	Val	Asp	Leu	
				170					175					180	
Gly	Thr	Glu	Phe	Leu	Thr	Cys	Asp	Cys	His	Leu	Arg	Trp	Leu	Leu	
				185					190					195	
Pro	Trp	Ala	Gln	Asn	Arg	Ser	Leu	Gln	Leu	Ser	Glu	His	Thr	Leu	
				200					205					210	
Cys	Ala	Tyr	Pro	Ser	Ala	Leu	His	Ala	Gln	Ala	Leu	Gly	Ser	Leu	
				215					220					225	
Gln	Glu	Ala	Gln	Leu	Cys	Cys	Glu	Gly	Ala	Leu	Glu	Leu	His	Thr	
				230					235					240	
His	His	Leu	Ile	Pro	Ser	Leu	Arg	Gln	Val	Val	Phe	Gln	Gly	Asp	
				245					250					255	
Arg	Leu	Pro	Phe	Gln	Cys	Ser	Ala	Ser	Tyr	Leu	Gly	Asn	Asp	Thr	
				260					265					270	
Arg	Ile	Arg	Trp	Tyr	His	Asn	Arg	Ala	Pro	Val	Glu	Gly	Asp	Glu	
				275					280					285	
Gln	Ala	Gly	Ile	Leu	Leu	Ala	Glu	Ser	Leu	Ile	His	Asp	Cys	Thr	
				290					295					300	
Phe	Ile	Thr	Ser	Glu	Leu	Thr	Leu	Ser	His	Ile	Gly	Val	Trp	Ala	
				305					310					315	
Ser	Gly	Glu	Trp	Glu	Cys	Thr	Val	Ser	Met	Ala	Gln	Gly	Asn	Ala	
				320					325					330	

WO 01/98354

PCT/US01/19942

Ser	Lys	Lys	Val	Glu	Ile	Val	Val	Leu	Glu	Thr	Ser	Ala	Ser	Tyr
				335					340					345
Cys	Pro	Ala	Glu	Arg	Val	Ala	Asn	Asn	Arg	Gly	Asp	Phe	Arg	Trp
				350					355					360
Pro	Arg	Thr	Leu	Ala	Gly	Ile	Thr	Ala	Tyr	Gln	Ser	Cys	Leu	Gln
				365					370					375
Tyr	Pro	Phe	Thr	Ser	Val	Pro	Leu	Gly	Gly	Gly	Ala	Pro	Gly	Thr
				380					385					390
Arg	Ala	Ser	Arg	Arg	Cys	Asp	Arg	Ala	Gly	Arg	Trp	Glu	Pro	Gly
				395					400					405
Asp	Tyr	Ser	His	Cys	Leu	Tyr	Thr	Asn	Asp	Ile	Thr	Arg	Val	Leu
				410					415					420
Tyr	Thr	Phe	Val	Leu	Met	Pro	Ile	Asn	Ala	Ser	Asn	Ala	Leu	Thr
				425					430					435
Leu	Ala	His	Gln	Leu	Arg	Val	Tyr	Thr	Ala	Glu	Ala	Ala	Ser	Phe
				440					445					450
Ser	Asp	Met	Met	Asp	Val	Val	Tyr	Val	Ala	Gln	Met	Ile	Gln	Lys
				455					460					465
Phe	Leu	Gly	Tyr	Val	Asp	Gln	Ile	Lys	Glu	Leu	Val	Glu	Val	Met
				470					475					480
Val	Asp	Met	Ala	Ser	Asn	Leu	Met	Leu	Val	Asp	Glu	His	Leu	Leu
				485					490					495
Trp	Leu	Ala	Gln	Arg	Glu	Asp	Lys	Ala	Cys	Ser	Arg	Ile	Val	Gly
				500					505					510
Ala	Leu	Glu	Arg	Ile	Gly	Gly	Ala	Ala	Leu	Ser	Pro	His	Ala	Gln
				515					520					525
His	Ile	Ser	Val	Asn	Ala	Arg	Asn	Val	Ala	Leu	Glu	Ala	Tyr	Leu
				530					535					540
Ile	Lys	Pro	His	Ser	Tyr	Val	Gly	Leu	Thr	Cys	Thr	Ala	Phe	Gln
				545					550					555
Arg	Arg	Glu	Gly	Gly	Val	Pro	Gly	Thr	Arg	Pro	Gly	Ser	Pro	Gly
				560					565					570
Gln	Asn	Pro	Pro	Pro	Glu	Pro	Glu	Pro	Pro	Ala	Asp	Gln	Gln	Leu
				575					580					585
Arg	Phe	Arg	Cys	Thr	Thr	Gly	Arg	Pro	Asn	Val	Ser	Leu	Ser	Ser
				590					595					600
Phe	His	Ile	Lys	Asn	Ser	Val	Ala	Leu	Ala	Ser	Ile	Gln	Leu	Pro
				605					610					615
Pro	Ser	Leu	Phe	Ser	Ser	Leu	Pro	Ala	Ala	Leu	Ala	Pro	Pro	Val
				620					625					630
Pro	Pro	Asp	Cys	Thr	Leu	Gln	Leu	Leu	Val	Phe	Arg	Asn	Gly	Arg
				635					640					645
Leu	Phe	His	Ser	His	Ser	Asn	Thr	Ser	Arg	Pro	Gly	Ala	Ala	Gly
				650					655					660
Pro	Gly	Lys	Arg	Arg	Gly	Val	Ala	Thr	Pro	Val	Ile	Phe	Ala	Gly
				665					670					675
Thr	Ser	Gly	Cys	Gly	Val	Gly	Asn	Leu	Thr	Glu	Pro	Val	Ala	Val
				680					685					690
Ser	Leu	Arg	His	Trp	Ala	Glu	Gly	Ala	Glu	Pro	Val	Ala	Ala	Trp
				695					700					705
Trp	Ser	Gln	Glu	Gly	Pro	Gly	Glu	Ala	Gly	Gly	Trp	Thr	Ser	Glu
				710					715					720
Gly	Cys	Gln	Leu	Arg	Ser	Ser	Gln	Pro	Asn	Val	Ser	Ala	Leu	His
				725					730					735
Cys	Gln	His	Leu	Gly	Asn	Val	Ala	Val	Leu	Met	Glu	Leu	Ser	Ala
				740					745					750
Phe	Pro	Arg	Glu	Val	Gly	Gly	Ala	Gly	Ala	Gly	Leu	His	Pro	Val
				755					760					765
Val	Tyr	Pro	Cys	Thr	Ala	Leu	Leu	Leu	Leu	Cys	Leu	Phe	Ala	Thr
				770					775					780
Ile	Ile	Thr	Tyr	Ile	Leu	Asn	His	Ser	Ser	Ile	Arg	Val	Ser	Arg
				785					790					795
Lys	Gly	Trp	His	Met	Leu	Leu	Asn	Leu	Cys	Phe	His	Ile	Ala	Met

Thr	Ser	Ala	Val	800	Phe	Ala	Gly	Gly	Ile	805	Thr	Leu	Thr	Asn	Tyr	Gln	810
				815						820							825
Met	Val	Cys	Gln	830	Ala	Val	Gly	Ile	Thr	835	Leu	His	Tyr	Ser	Ser	Leu	840
Ser	Thr	Leu	Leu	845	Trp	Met	Gly	Val	Lys	850	Ala	Arg	Val	Leu	His	Lys	855
Glu	Leu	Thr	Trp	860	Arg	Ala	Pro	Pro	Pro	865	Gln	Glu	Gly	Asp	Pro	Ala	870
Leu	Pro	Thr	Pro	875	Ser	Pro	Met	Leu	Arg	880	Phe	Tyr	Leu	Ile	Ala	Gly	885
Gly	Ile	Pro	Leu	890	Ile	Ile	Cys	Gly	Ile	895	Thr	Ala	Ala	Val	Asn	Ile	900
His	Asn	Tyr	Arg	905	Asp	His	Ser	Pro	Tyr	910	Cys	Trp	Leu	Val	Trp	Arg	915
Pro	Ser	Leu	Gly	920	Ala	Phe	Tyr	Ile	Pro	925	Val	Ala	Leu	Ile	Leu	Leu	930
Ile	Thr	Trp	Ile	935	Tyr	Phe	Leu	Cys	Ala	940	Gly	Leu	Arg	Leu	Arg	Gly	945
Pro	Leu	Ala	Gln	950	Asn	Pro	Lys	Ala	Gly	955	Asn	Ser	Arg	Ala	Ser	Leu	960
Glu	Ala	Gly	Glu	965	Glu	Leu	Arg	Gly	Ser	970	Thr	Arg	Leu	Arg	Gly	Ser	975
Gly	Pro	Leu	Leu	980	Ser	Asp	Ser	Gly	Ser	985	Leu	Leu	Ala	Thr	Gly	Ser	990
Ala	Arg	Val	Gly	995	Thr	Pro	Gly	Pro	Pro	1000	Glu	Asp	Gly	Asp	Ser	Leu	1005
Tyr	Ser	Pro	Gly	1010	Val	Gln	Leu	Gly	Ala	1015	Leu	Val	Thr	Thr	His	Phe	1020
Leu	Tyr	Leu	Ala	1025	Met	Trp	Ala	Cys	Gly	1030	Ala	Leu	Ala	Val	Ser	Gln	1035
Arg	Trp	Leu	Pro	1040	Arg	Val	Val	Cys	Ser	1045	Cys	Leu	Tyr	Gly	Val	Ala	1050
Ala	Ser	Ala	Leu	1055	Gly	Leu	Phe	Val	Phe	1060	Thr	His	His	Cys	Ala	Arg	1065
Arg	Arg	Asp	Val	1070	Arg	Ala	Ser	Trp	Arg	1075	Ala	Cys	Cys	Pro	Pro	Ala	1080
Ser	Pro	Ala	Ala	1085	Pro	His	Ala	Pro	Pro	1090	Arg	Ala	Leu	Pro	Ala	Ala	1095
Ala	Glu	Asp	Gly	1100	Ser	Pro	Val	Phe	Gly	1105	Glu	Gly	Pro	Pro	Ser	Leu	1110
Lys	Ser	Ser	Pro	1115	Ser	Gly	Ser	Ser	Gly	1120	His	Pro	Leu	Ala	Leu	Gly	1125
Pro	Cys	Lys	Leu	1130	Thr	Asn	Leu	Gln	Leu	1135	Ala	Gln	Ser	Gln	Val	Cys	1140
Glu	Ala	Gly	Ala	1145	Ala	Gly	Gly	Glu	Gly	1150	Glu	Pro	Glu	Pro	Ala		1155
Gly	Thr	Arg	Gly	1160	Asn	Leu	Ala	His	Arg	1165	His	Pro	Asn	Asn	Val	His	1170
His	Gly	Arg	Arg	1175	Ala	His	Lys	Ser	Arg	1180	Ala	Lys	Gly	His	Arg	Ala	1185
Gly	Glu	Ala	Cys	1190	Gly	Lys	Asn	Arg	Leu	1195	Lys	Ala	Leu	Arg	Gly	Gly	1200
Ala	Ala	Gly	Ala	1205	Leu	Glu	Leu	Leu	Ser	1210	Ser	Glu	Ser	Gly	Ser	Leu	1215
His	Asn	Ser	Pro	1220	Thr	Asp	Ser	Tyr	Leu	1225	Gly	Ser	Ser	Arg	Asn	Ser	1230
Pro	Gly	Ala	Gly	1235	Leu	Gln	Leu	Glu	Gly	1240	Glu	Pro	Met	Leu	Thr	Pro	1245
Ser	Glu	Gly	Ser	1250	Asp	Thr	Ser	Ala	Ala	1255	Pro	Leu	Ser	Glu	Ala	Gly	1260
Arg	Ala	Gly	Gln	1265	Arg	Arg	Ser	Ala	Ser	1270	Arg	Asp	Ser	Leu	Lys	Gly	1275

WO 01/98354

PCT/US01/19942

Gly	Gly	Ala	Leu	Glu	Lys	Glu	Ser	His	Arg	Arg	Ser	Tyr	Pro	Leu
				1280					1285					1290
Asn	Ala	Ala	Ser	Leu	Asn	Gly	Ala	Pro	Lys	Gly	Gly	Lys	Tyr	Asp
				1295					1300					1305
Asp	Val	Thr	Leu	Met	Gly	Ala	Glu	Val	Ala	Ser	Gly	Gly	Cys	Met
				1310					1315					1320
Lys	Thr	Gly	Leu	Trp	Lys	Ser	Glu	Thr	Thr	Val				
				1325					1330					

<210> 8

<211> 3217

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1897612CD1

<400> 8

Met	Lys	Ser	Pro	Arg	Pro	His	Leu	Leu	Leu	Pro	Leu	Leu	Leu	Leu
1				5					10					15
Leu	Leu	Leu	Leu	Leu	Gly	Ala	Gly	Val	Pro	Gly	Ala	Trp	Gly	Gln
				20					25					30
Ala	Gly	Ser	Leu	Asp	Leu	Gln	Ile	Asp	Glu	Glu	Gln	Pro	Ala	Gly
				35					40					45
Thr	Leu	Ile	Gly	Asp	Ile	Ser	Ala	Gly	Leu	Pro	Ala	Gly	Thr	Ala
				50					55					60
Ala	Pro	Leu	Met	Tyr	Phe	Ile	Ser	Ala	Gln	Glu	Gly	Ser	Gly	Val
				65					70					75
Gly	Thr	Asp	Leu	Ala	Ile	Asp	Glu	His	Ser	Gly	Val	Val	Arg	Thr
				80					85					90
Ala	Arg	Val	Leu	Asp	Arg	Glu	Gln	Arg	Asp	Arg	Tyr	Arg	Phe	Thr
				95					100					105
Ala	Val	Thr	Pro	Asp	Gly	Ala	Thr	Val	Glu	Val	Thr	Val	Arg	Val
				110					115					120
Ala	Asp	Ile	Asn	Asp	His	Ala	Pro	Ala	Phe	Pro	Gln	Ala	Arg	Ala
				125					130					135
Ala	Leu	Gln	Val	Pro	Glu	His	Thr	Ala	Phe	Gly	Thr	Arg	Tyr	Pro
				140					145					150
Leu	Glu	Pro	Ala	Arg	Asp	Ala	Asp	Ala	Gly	Arg	Leu	Gly	Thr	Gln
				155					160					165
Gly	Tyr	Ala	Leu	Ser	Gly	Asp	Gly	Ala	Gly	Glu	Thr	Phe	Arg	Leu
				170					175					180
Glu	Thr	Arg	Pro	Gly	Pro	Asp	Gly	Thr	Pro	Val	Pro	Glu	Leu	Val
				185					190					195
Val	Thr	Gly	Glu	Leu	Asp	Arg	Glu	Asn	Arg	Ser	His	Tyr	Met	Leu
				200					205					210
Gln	Leu	Glu	Ala	Tyr	Asp	Gly	Gly	Ser	Pro	Pro	Arg	Arg	Ala	Gln
				215					220					225
Ala	Leu	Leu	Asp	Val	Thr	Leu	Leu	Asp	Ile	Asn	Asp	His	Ala	Pro
				230					235					240
Ala	Phe	Asn	Gln	Ser	Arg	Tyr	His	Ala	Val	Val	Ser	Glu	Ser	Leu
				245					250					255
Ala	Pro	Gly	Ser	Pro	Val	Leu	Gln	Val	Phe	Ala	Ser	Asp	Ala	Asp
				260					265					270
Ala	Gly	Val	Asn	Gly	Ala	Val	Thr	Tyr	Glu	Ile	Asn	Arg	Arg	Gln
				275					280					285
Ser	Glu	Gly	Asp	Gly	Pro	Phe	Ser	Ile	Asp	Ala	His	Thr	Gly	Leu
				290					295					300
Leu	Gln	Leu	Glu	Arg	Pro	Leu	Asp	Phe	Glu	Gln	Arg	Arg	Val	His
				305					310					315
Glu	Leu	Val	Val	Gln	Ala	Arg	Asp	Asp	Gly	Ser	Pro	Gln	Val	Ser
				320					325					330

Glu	Ala	Ala	Pro	Pro	Gly	Gln	Leu	Val	Ala	Arg	Ile	Ser	Val	Ser
				335					340					345
Asp	Pro	Asp	Asp	Gly	Asp	Phe	Ala	His	Val	Asn	Val	Ser	Leu	Glu
				350					355					360
Gly	Gly	Glu	Gly	His	Phe	Ala	Leu	Ser	Thr	Gln	Asp	Ser	Val	Ile
				365					370					375
Tyr	Leu	Val	Cys	Val	Ala	Arg	Arg	Leu	Asp	Arg	Glu	Glu	Arg	Asp
				380					385					390
Ala	Tyr	Asn	Leu	Arg	Val	Thr	Ala	Thr	Asp	Ser	Gly	Ser	Pro	Pro
				395					400					405
Leu	Arg	Ala	Glu	Ala	Ala	Phe	Val	Leu	His	Val	Thr	Asp	Val	Asn
				410					415					420
Asp	Asn	Ala	Pro	Ala	Phe	Asp	Arg	Gln	Leu	Tyr	Arg	Pro	Glu	Pro
				425					430					435
Leu	Pro	Glu	Val	Ala	Leu	Pro	Gly	Ser	Phe	Val	Val	Arg	Val	Thr
				440					445					450
Ala	Arg	Asp	Pro	Asp	Gln	Gly	Thr	Asn	Gly	Gln	Val	Thr	Tyr	Ser
				455					460					465
Leu	Ala	Pro	Gly	Ala	His	Thr	His	Trp	Phe	Ser	Ile	Asp	Pro	Thr
				470					475					480
Ser	Gly	Ile	Ile	Thr	Thr	Ala	Ala	Ser	Leu	Asp	Tyr	Glu	Leu	Glu
				485					490					495
Pro	Gln	Pro	Gln	Leu	Ile	Val	Val	Ala	Thr	Asp	Gly	Gly	Leu	Pro
				500					505					510
Pro	Leu	Ala	Ser	Ser	Ala	Thr	Val	Ser	Val	Ala	Leu	Gln	Asp	Val
				515					520					525
Asn	Asp	Asn	Glu	Pro	Gln	Phe	Gln	Arg	Thr	Phe	Tyr	Asn	Ala	Ser
				530					535					540
Leu	Pro	Glu	Gly	Thr	Gln	Pro	Gly	Thr	Cys	Phe	Leu	Gln	Val	Thr
				545					550					555
Ala	Thr	Asp	Ala	Asp	Ser	Gly	Pro	Phe	Gly	Leu	Leu	Ser	Tyr	Ser
				560					565					570
Leu	Gly	Ala	Gly	Leu	Gly	Ser	Ser	Gly	Ser	Pro	Pro	Phe	Arg	Ile
				575					580					585
Asp	Ala	His	Ser	Gly	Asp	Val	Cys	Thr	Thr	Arg	Thr	Leu	Asp	Arg
				590					595					600
Asp	Gln	Gly	Pro	Ser	Ser	Phe	Asp	Phe	Thr	Val	Thr	Ala	Val	Asp
				605					610					615
Gly	Gly	Gly	Leu	Lys	Ser	Met	Val	Tyr	Val	Lys	Val	Phe	Leu	Ser
				620					625					630
Asp	Glu	Asn	Asp	Asn	Pro	Pro	Gln	Phe	Tyr	Pro	Arg	Glu	Tyr	Ala
				635					640					645
Ala	Ser	Ile	Ser	Ala	Gln	Ser	Pro	Pro	Gly	Thr	Ala	Val	Leu	Arg
				650					655					660
Leu	Arg	Ala	His	Asp	Pro	Asp	Gln	Gly	Ser	His	Gly	Arg	Leu	Ser
				665					670					675
Tyr	His	Ile	Leu	Ala	Gly	Asn	Ser	Pro	Pro	Leu	Phe	Thr	Leu	Asp
				680					685					690
Glu	Gln	Ser	Gly	Leu	Leu	Thr	Val	Ala	Trp	Pro	Leu	Ala	Arg	Arg
				695					700					705
Ala	Asn	Ser	Val	Val	Gln	Leu	Glu	Ile	Gly	Ala	Glu	Asp	Gly	Gly
				710					715					720
Gly	Leu	Gln	Ala	Glu	Pro	Ser	Ala	Arg	Val	Asp	Ile	Ser	Ile	Val
				725					730					735
Pro	Gly	Thr	Pro	Thr	Pro	Pro	Ile	Phe	Glu	Gln	Leu	Gln	Tyr	Val
				740					745					750
Phe	Ser	Val	Pro	Glu	Asp	Val	Ala	Pro	Gly	Thr	Ser	Val	Gly	Ile
				755					760					765
Val	Gln	Ala	His	Asn	Pro	Pro	Gly	Gly	Asp	Pro	Arg	Gly	Leu	Phe
				770					775					780
Ser	Leu	Asp	Ala	Val	Ser	Gly	Leu	Leu	Gln	Thr	Leu	Arg	Pro	Leu
				785					790					795
Asp	Arg	Glu	Leu	Leu	Gly	Pro	Val	Leu	Glu	Leu	Glu	Val	Arg	Ala

Gly Ser Gly Val	800	805	810
Pro Pro Ala Phe Ala Val Ala Arg Val Arg Val	815	820	825
Leu Leu Asp Asp Val Asn Asp Asn Ser Pro Ala Phe Pro Ala Pro	830	835	840
Glu Asp Thr Val Leu Leu Pro Pro Asn Thr Ala Pro Gly Thr Pro	845	850	855
Ile Tyr Thr Leu Arg Ala Leu Asp Pro Asp Ser Gly Val Asn Ser	860	865	870
Arg Val Thr Phe Thr Leu Leu Ala Gly Gly Gly Gly Ala Phe Thr	875	880	885
Val Asp Pro Thr Thr Gly His Val Arg Leu Met Arg Pro Leu Gly	890	895	900
Pro Ser Gly Gly Pro Ala His Glu Leu Glu Leu Glu Ala Arg Asp	905	910	915
Gly Gly Ser Pro Pro Arg Thr Ser His Phe Arg Leu Arg Val Val	920	925	930
Val Gln Asp Val Gly Thr Arg Gly Leu Ala Pro Arg Phe Asn Ser	935	940	945
Pro Thr Tyr Arg Val Asp Leu Pro Ser Gly Thr Thr Ala Gly Thr	950	955	960
Gln Val Leu Gln Val Gln Ala Gln Ala Pro Asp Gly Gly Pro Ile	965	970	975
Thr Tyr His Leu Ala Ala Glu Gly Ala Ser Ser Pro Phe Gly Leu	980	985	990
Glu Pro Gln Ser Gly Trp Leu Trp Val Arg Ala Ala Leu Asp Arg	995	1000	1005
Glu Ala Gln Glu Leu Tyr Ile Leu Lys Val Met Ala Val Ser Gly	1010	1015	1020
Ser Lys Ala Glu Leu Gly Gln Gln Thr Gly Thr Ala Thr Val Arg	1025	1030	1035
Val Ser Ile Leu Asn Gln Asn Glu His Ser Pro Arg Leu Ser Glu	1040	1045	1050
Asp Pro Thr Phe Leu Ala Val Ala Glu Asn Gln Pro Pro Gly Thr	1055	1060	1065
Ser Val Gly Arg Val Phe Ala Thr Asp Arg Asp Ser Gly Pro Asn	1070	1075	1080
Gly Arg Leu Thr Tyr Ser Leu Gln Gln Leu Ser Glu Asp Ser Lys	1085	1090	1095
Ala Phe Arg Ile His Pro Gln Thr Gly Glu Val Thr Thr Leu Gln	1100	1105	1110
Thr Leu Asp Arg Glu Gln Gln Ser Ser Tyr Gln Leu Leu Val Gln	1115	1120	1125
Val Gln Asp Gly Gly Ser Pro Pro Arg Ser Thr Thr Gly Thr Val	1130	1135	1140
His Val Ala Val Leu Asp Leu Asn Asp Asn Ser Pro Thr Phe Leu	1145	1150	1155
Gln Ala Ser Gly Ala Ala Gly Gly Gly Leu Pro Ile Gln Val Pro	1160	1165	1170
Asp Arg Val Pro Pro Gly Thr Leu Val Thr Thr Leu Gln Ala Lys	1175	1180	1185
Asp Pro Asp Glu Gly Glu Asn Gly Thr Ile Leu Tyr Thr Leu Thr	1190	1195	1200
Gly Pro Gly Ser Glu Leu Phe Ser Leu His Pro His Ser Gly Glu	1205	1210	1215
Leu Leu Thr Ala Ala Pro Leu Ile Arg Ala Glu Arg Pro His Tyr	1220	1225	1230
Val Leu Thr Leu Ser Ala His Asp Gln Gly Ser Pro Pro Arg Ser	1235	1240	1245
Ala Ser Leu Gln Leu Leu Val Gln Val Leu Pro Ser Ala Arg Leu	1250	1255	1260
Ala Glu Pro Pro Pro Asp Leu Ala Glu Arg Asp Pro Ala Ala Pro	1265	1270	1275

Val	Pro	Val	Val	Leu	Thr	Val	Thr	Ala	Ala	Glu	Gly	Leu	Arg	Pro
				1280					1285					1290
Gly	Ser	Leu	Leu	Gly	Ser	Val	Ala	Ala	Pro	Glu	Pro	Ala	Gly	Val
				1295					1300					1305
Gly	Ala	Leu	Thr	Tyr	Thr	Leu	Val	Gly	Gly	Ala	Asp	Pro	Glu	Gly
				1310					1315					1320
Thr	Phe	Ala	Leu	Asp	Ala	Ala	Ser	Gly	Arg	Leu	Tyr	Leu	Ala	Arg
				1325					1330					1335
Pro	Leu	Asp	Phe	Glu	Ala	Gly	Pro	Pro	Trp	Arg	Ala	Leu	Thr	Val
				1340					1345					1350
Gln	Val	Gln	Asp	Glu	Asn	Glu	His	Ala	Pro	Ala	Phe	Ala	Arg	Asp
				1355					1360					1365
Pro	Leu	Gly	Ala	Leu	Pro	Glu	Asn	Pro	Glu	Pro	Gly	Ala	Ala	Leu
				1370					1375					1380
Tyr	Thr	Phe	Arg	Ala	Ser	Asp	Ala	Asp	Gly	Pro	Gly	Pro	Asn	Ser
				1385					1390					1395
Asp	Val	Arg	Tyr	Arg	Leu	Leu	Arg	Gln	Glu	Pro	Pro	Val	Pro	Gly
				1400					1405					1410
Phe	Ala	Trp	Thr	Arg	Ala	Pro	Gly	Arg	Gln	Leu	Arg	Ala	Ala	Trp
				1415					1420					1425
Thr	Glu	Arg	Pro	Leu	Pro	Arg	Cys	Cys	Cys	Trp	Trp	Lys	Pro	Pro
				1430					1435					1440
Thr	Gly	Pro	Pro	Thr	Pro	Ala	Ala	Val	Val	Gln	Arg	Ala	Phe	Gln
				1445					1450					1455
Arg	Ile	Tyr	Val	Thr	Asp	Ala	Asn	Glu	Asn	Ala	Pro	Val	Phe	Ala
				1460					1465					1470
Ser	Pro	Cys	Thr	Gln	Asp	Gln	Pro	Pro	Gly	Pro	Ala	Ala	Gly	Thr
				1475					1480					1485
Leu	Leu	Ala	Arg	Asp	Pro	His	Leu	Gly	Glu	Ala	Ala	Arg	Val	Ser
				1490					1495					1500
Tyr	Arg	Leu	Ala	Ser	Gly	Gly	Asp	Gly	His	Phe	Arg	Leu	His	Ser
				1505					1510					1515
Ser	Thr	Gly	Ala	Leu	Ser	Val	Val	Arg	Pro	Leu	Asp	Arg	Glu	Gln
				1520					1525					1530
Arg	Ala	Glu	His	Val	Leu	Thr	Val	Val	Ala	Ser	Asp	Arg	Ala	Pro
				1535					1540					1545
Arg	Pro	Arg	Ser	Ala	Thr	Gln	Val	Leu	Thr	Val	Ser	Val	Ala	Asp
				1550					1555					1560
Val	Asn	Asp	Glu	Ala	Pro	Thr	Phe	Gln	Gln	Gln	Glu	Tyr	Ser	Val
				1565					1570					1575
Leu	Leu	Leu	Glu	Asn	Asn	Pro	Pro	Gly	Thr	Ser	Leu	Leu	Thr	Leu
				1580					1585					1590
Arg	Ala	Thr	Asp	Pro	Asp	Val	Gly	Ala	Asn	Gly	Gln	Val	Thr	Tyr
				1595					1600					1605
Gly	Gly	Val	Ser	Ser	Glu	Ser	Phe	Ser	Leu	Asp	Pro	Asp	Thr	Gly
				1610					1615					1620
Val	Leu	Thr	Thr	Leu	Arg	Ala	Leu	Asp	Arg	Glu	Glu	Gln	Glu	Glu
				1625					1630					1635
Ile	Asn	Leu	Thr	Val	Tyr	Ala	Gln	Asp	Arg	Gly	Ser	Pro	Pro	Gln
				1640					1645					1650
Leu	Thr	His	Val	Thr	Val	Arg	Val	Ala	Val	Glu	Asp	Glu	Asn	Asp
				1655					1660					1665
His	Ala	Pro	Thr	Phe	Gly	Ser	Ala	His	Leu	Ser	Leu	Glu	Val	Pro
				1670					1675					1680
Glu	Gly	Gln	Asp	Pro	Gln	Thr	Leu	Thr	Met	Leu	Arg	Ala	Ser	Asp
				1685					1690					1695
Pro	Asp	Val	Gly	Ala	Asn	Gly	Gln	Leu	Gln	Tyr	Arg	Ile	Leu	Asp
				1700					1705					1710
Gly	Asp	Pro	Ser	Gly	Ala	Phe	Val	Leu	Asp	Leu	Ala	Ser	Gly	Glu
				1715					1720					1725
Phe	Gly	Thr	Met	Arg	Pro	Leu	Asp	Arg	Glu	Val	Glu	Pro	Ala	Phe
				1730					1735					1740
Gln	Leu	Arg	Ile	Glu	Ala	Arg	Asp	Gly	Gly	Gln	Pro	Ala	Leu	Ser

1745	1750	1755
Ala Thr Leu Leu Leu Thr Val Thr Val Leu Asp Ala Asn Asp His		
1760	1765	1770
Ala Pro Ala Phe Pro Val Pro Ala Tyr Ser Val Glu Val Pro Glu		
1775	1780	1785
Asp Val Pro Ala Gly Thr Leu Leu Leu Gln Leu Gln Ala His Asp		
1790	1795	1800
Pro Asp Ala Gly Ala Asn Gly His Val Thr Tyr Tyr Leu Gly Ala		
1805	1810	1815
Gly Thr Ala Gly Ala Phe Leu Leu Glu Pro Ser Ser Gly Glu Leu		
1820	1825	1830
Arg Thr Ala Ala Ala Leu Asp Arg Glu Gln Cys Pro Ser Tyr Thr		
1835	1840	1845
Phe Ser Val Ser Ala Val Asp Gly Ala Ala Ala Gly Pro Leu Ser		
1850	1855	1860
Thr Thr Val Ser Val Thr Ile Thr Val Arg Asp Val Asn Asp His		
1865	1870	1875
Ala Pro Thr Phe Pro Thr Ser Pro Leu Arg Leu Arg Leu Pro Arg		
1880	1885	1890
Pro Gly Pro Ser Phe Ser Thr Pro Thr Leu Ala Leu Ala Thr Leu		
1895	1900	1905
Arg Ala Glu Asp Arg Asp Ala Gly Ala Asn Ala Ser Ile Leu Tyr		
1910	1915	1920
Arg Leu Ala Gly Thr Pro Pro Pro Gly Thr Thr Val Asp Ser Tyr		
1925	1930	1935
Thr Gly Glu Ile Arg Val Ala Arg Ser Pro Val Ala Leu Gly Pro		
1940	1945	1950
Arg Asp Arg Val Leu Phe Ile Val Ala Thr Asp Leu Gly Arg Pro		
1955	1960	1965
Ala Arg Ser Ala Thr Gly Val Ile Ile Val Gly Leu Gln Gly Glu		
1970	1975	1980
Ala Glu Arg Gly Pro Arg Phe Pro Arg Ala Ser Ser Glu Ala Thr		
1985	1990	1995
Ile Arg Glu Asn Ala Pro Pro Gly Thr Pro Ile Val Ser Pro Arg		
2000	2005	2010
Ala Val His Ala Gly Gly Thr Asn Gly Pro Ile Thr Tyr Ser Ile		
2015	2020	2025
Leu Ser Gly Asn Glu Lys Gly Thr Phe Ser Ile Gln Pro Ser Thr		
2030	2035	2040
Gly Ala Ile Thr Val Arg Ser Ala Glu Gly Leu Asp Phe Glu Val		
2045	2050	2055
Ser Pro Arg Leu Arg Leu Val Leu Gln Ala Glu Ser Gly Gly Ala		
2060	2065	2070
Phe Ala Phe Thr Val Leu Thr Leu Thr Leu Gln Asp Ala Asn Asp		
2075	2080	2085
Asn Ala Pro Arg Phe Leu Arg Pro His Tyr Val Ala Phe Leu Pro		
2090	2095	2100
Glu Ser Arg Pro Leu Glu Gly Pro Leu Leu Gln Val Glu Ala Asp		
2105	2110	2115
Asp Leu Asp Gln Gly Ser Gly Gly Gln Ile Ser Tyr Ser Leu Ala		
2120	2125	2130
Ala Ser Gln Pro Ala Arg Gly Leu Phe His Val Asp Pro Thr Thr		
2135	2140	2145
Gly Thr Ile Thr Thr Thr Ala Ile Leu Asp Arg Glu Ile Trp Ala		
2150	2155	2160
Glu Thr Arg Leu Val Leu Met Ala Thr Asp Arg Gly Ser Pro Ala		
2165	2170	2175
Leu Val Gly Ser Ala Thr Leu Thr Val Met Val Ile Asp Thr Asn		
2180	2185	2190
Asp Asn Arg Pro Thr Ile Pro Gln Pro Trp Glu Leu Arg Val Ser		
2195	2200	2205
Glu Asp Ala Leu Leu Gly Ser Glu Ile Ala Gln Val Thr Gly Asn		
2210	2215	2220

Asp	Val	Asp	Ser	Gly	Pro	Val	Leu	Trp	Tyr	Val	Leu	Ser	Pro	Ser
				2225					2230					2235
Gly	Pro	Gln	Asp	Pro	Phe	Ser	Val	Gly	Arg	Tyr	Gly	Gly	Arg	Val
				2240					2245					2250
Ser	Leu	Thr	Gly	Pro	Leu	Asp	Phe	Glu	Gln	Cys	Asp	Arg	Tyr	Gln
				2255					2260					2265
Leu	Gln	Leu	Leu	Ala	His	Asp	Gly	Pro	His	Glu	Gly	Arg	Ala	Asn
				2270					2275					2280
Leu	Thr	Val	Leu	Val	Glu	Asp	Val	Asn	Asp	Asn	Ala	Pro	Ala	Phe
				2285					2290					2295
Ser	Gln	Ser	Leu	Tyr	Gln	Val	Met	Leu	Leu	Glu	His	Thr	Pro	Pro
				2300					2305					2310
Gly	Ser	Ala	Ile	Leu	Ser	Val	Ser	Ala	Thr	Asp	Arg	Asp	Ser	Gly
				2315					2320					2325
Ala	Asn	Gly	His	Ile	Ser	Tyr	His	Leu	Ala	Ser	Pro	Ala	Asp	Gly
				2330					2335					2340
Phe	Ser	Val	Asp	Pro	Asn	Asn	Gly	Thr	Leu	Phe	Thr	Ile	Val	Gly
				2345					2350					2355
Thr	Val	Ala	Leu	Gly	His	Asp	Gly	Ser	Gly	Ala	Val	Asp	Val	Val
				2360					2365					2370
Leu	Glu	Ala	Arg	Asp	His	Gly	Ala	Pro	Gly	Arg	Ala	Ala	Arg	Ala
				2375					2380					2385
Thr	Val	His	Val	Gln	Leu	Gln	Asp	Gln	Asn	Asp	His	Ala	Pro	Ser
				2390					2395					2400
Phe	Thr	Leu	Ser	His	Tyr	Arg	Val	Ala	Val	Thr	Glu	Asp	Leu	Pro
				2405					2410					2415
Pro	Gly	Ser	Thr	Leu	Leu	Thr	Leu	Glu	Ala	Thr	Asp	Ala	Asp	Gly
				2420					2425					2430
Ser	Arg	Ser	His	Ala	Ala	Val	Asp	Tyr	Ser	Thr	Ile	Ser	Gly	Asn
				2435					2440					2445
Trp	Gly	Arg	Val	Phe	Gln	Leu	Glu	Pro	Arg	Leu	Ala	Glu	Ala	Gly
				2450					2455					2460
Glu	Ser	Ala	Gly	Pro	Gly	Pro	Arg	Ala	Leu	Gly	Cys	Leu	Val	Leu
				2465					2470					2475
Leu	Glu	Pro	Leu	Asp	Phe	Glu	Ser	Leu	Thr	Gln	Tyr	Asn	Leu	Thr
				2480					2485					2490
Val	Ala	Ala	Ala	Asp	Arg	Gly	Gln	Pro	Pro	Gln	Ser	Ser	Val	Val
				2495					2500					2505
Pro	Val	Thr	Val	Thr	Val	Leu	Asp	Val	Asn	Asp	Asn	Pro	Pro	Val
				2510					2515					2520
Phe	Thr	Arg	Ala	Ser	Tyr	Arg	Val	Thr	Val	Pro	Glu	Asp	Thr	Pro
				2525					2530					2535
Val	Gly	Ala	Glu	Leu	Leu	His	Val	Glu	Ala	Ser	Asp	Ala	Asp	Pro
				2540					2545					2550
Gly	Pro	His	Gly	Leu	Val	Arg	Phe	Thr	Val	Ser	Ser	Gly	Asp	Pro
				2555					2560					2565
Ser	Gly	Leu	Phe	Glu	Leu	Asp	Glu	Ser	Ser	Gly	Thr	Leu	Arg	Leu
				2570					2575					2580
Ala	His	Ala	Leu	Asp	Cys	Glu	Thr	Gln	Ala	Arg	His	Gln	Leu	Val
				2585					2590					2595
Val	Gln	Ala	Ala	Asp	Pro	Ala	Gly	Ala	His	Phe	Ala	Leu	Ala	Pro
				2600					2605					2610
Val	Thr	Ile	Glu	Val	Gln	Asp	Val	Asn	Asp	His	Gly	Pro	Ala	Phe
				2615					2620					2625
Pro	Leu	Asn	Leu	Leu	Ser	Thr	Ser	Val	Ala	Glu	Asn	Gln	Pro	Pro
				2630					2635					2640
Gly	Thr	Leu	Val	Thr	Thr	Leu	His	Ala	Ile	Asp	Gly	Asp	Ala	Gly
				2645					2650					2655
Ala	Phe	Gly	Arg	Leu	Arg	Tyr	Ser	Leu	Leu	Glu	Ala	Gly	Pro	Gly
				2660					2665					2670
Pro	Glu	Gly	Arg	Glu	Ala	Phe	Ala	Leu	Asn	Ser	Ser	Thr	Gly	Glu
				2675					2680					2685
Leu	Arg	Ala	Arg	Val	Pro	Phe	Asp	Tyr	Glu	His	Thr	Glu	Ser	Phe

Arg Leu Leu Val Gly	Ala Ala Asp Ala Gly	Asn Leu Ser Ala Ser
2690	2695	2700
Val Thr Val Ser Val	Leu Val Thr Gly Glu	Asp Glu Tyr Asp Pro
2705	2710	2715
Val Phe Leu Ala Pro	Ala Phe His Phe Gln	Val Pro Glu Gly Ala
2720	2725	2730
Arg Arg Gly His Ser	Leu Gly His Val Gln	Ala Thr Asp Glu Asp
2735	2740	2745
Gly Gly Ala Asp Gly	Leu Val Leu Tyr Ser	Leu Ala Thr Ser Ser
2750	2755	2760
Pro Tyr Phe Gly Ile	Asn Gln Thr Thr Gly	Ala Leu Tyr Leu Arg
2765	2770	2775
Val Asp Ser Arg Ala	Pro Gly Ser Gly Thr	Ala Thr Ser Gly Gly
2780	2785	2790
Gly Gly Arg Thr Arg	Arg Glu Ala Pro Arg	Glu Leu Arg Leu Glu
2795	2800	2805
Val Ile Ala Arg Gly	Pro Leu Pro Gly Ser	Arg Ser Ala Thr Val
2810	2815	2820
Pro Val Thr Val Asp	Ile Thr His Thr Ala	Leu Gly Leu Ala Pro
2825	2830	2835
Asp Leu Asn Leu Leu	Leu Val Gly Ala Val	Ala Ala Ser Leu Gly
2840	2845	2850
Val Val Val Val Leu	Ala Leu Ala Ala Leu	Val Leu Gly Leu Val
2855	2860	2865
Arg Ala Arg Ser Arg	Lys Ala Glu Ala Ala	Pro Gly Pro Met Ser
2870	2875	2880
Gln Ala Ala Pro Leu	Ala Ser Asp Ser Leu	Gln Lys Leu Gly Arg
2885	2890	2895
Glu Pro Pro Ser Pro	Pro Pro Ser Glu His	Leu Tyr His Gln Thr
2900	2905	2910
Leu Pro Ser Tyr Gly	Gly Pro Gly Ala Gly	Gly Pro Tyr Pro Arg
2915	2920	2925
Gly Gly Ser Leu Asp	Pro Ser His Ser Ser	Gly Arg Gly Ser Ala
2930	2935	2940
Glu Ala Ala Glu Asp	Asp Glu Ile Arg Met	Ile Asn Glu Phe Pro
2945	2950	2955
Arg Val Ala Ser Val	Ala Ser Ser Leu Ala	Ala Arg Gly Pro Asp
2960	2965	2970
Ser Gly Ile Gln Gln	Asp Ala Asp Gly Leu	Ser Asp Thr Ser Cys
2975	2980	2985
Glu Pro Pro Ala Pro	Asp Thr Trp Tyr Lys	Gly Arg Lys Ala Gly
2990	2995	3000
Leu Leu Leu Pro Gly	Ala Gly Ala Thr Leu	Tyr Arg Glu Glu Gly
3005	3010	3015
Pro Pro Ala Thr Ala	Thr Ala Phe Leu Gly	Gly Cys Gly Leu Ser
3020	3025	3030
Pro Ala Pro Thr Gly	Asp Tyr Gly Phe Pro	Ala Asp Gly Lys Pro
3035	3040	3045
Cys Val Ala Gly Ala	Leu Thr Ala Ile Val	Ala Gly Glu Glu Glu
3050	3055	3060
Leu Arg Gly Ser Tyr	Asn Trp Asp Tyr Leu	Leu Ser Trp Cys Pro
3065	3070	3075
Gln Phe Gln Pro Leu	Ala Ser Val Phe Thr	Glu Ile Ala Arg Leu
3080	3085	3090
Lys Asp Glu Ala Arg	Pro Cys Pro Pro Ala	Pro Arg Ile Asp Pro
3095	3100	3105
Pro Pro Leu Ile Thr	Ala Val Ala His Pro	Gly Ala Lys Ser Val
3110	3115	3120
Pro Pro Lys Pro Ala	Asn Thr Ala Ala Ala	Arg Ala Ile Phe Pro
3125	3130	3135
Pro Ala Ser His Arg	Ser Pro Ile Ser His	Glu Gly Ser Leu Ser
3140	3145	3150
Pro Ala Ser His Arg	Ser Pro Ile Ser His	Glu Gly Ser Leu Ser
3155	3160	3165

WO 01/98354

PCT/US01/19942

Ser Ala Ala Met Ser Pro Ser Phe Ser Pro Ser Leu Ser Pro Leu
 3170 3175 3180
 Ala Ala Arg Ser Pro Val Val Ser Pro Phe Gly Val Ala Gln Gly
 3185 3190 3195
 Pro Ser Ala Ser Ala Leu Ser Ala Glu Ser Gly Leu Glu Pro Pro
 3200 3205 3210
 Asp Asp Thr Glu Leu His Ile
 3215

<210> 9

<211> 2936

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6977010CD1

<400> 9

Met Arg Ser Pro Ala Thr Gly Val Pro Leu Pro Thr Pro Pro Pro
 1 5 10 15
 Pro Pro Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Pro Pro Pro Leu
 20 25 30
 Leu Gly Asp Gln Val Gly Pro Cys Arg Ser Leu Gly Ser Arg Gly
 35 40 45
 Arg Gly Ser Ser Gly Ala Cys Ala Pro Met Gly Trp Leu Cys Pro
 50 55 60
 Ser Ser Ala Ser Asn Leu Trp Leu Tyr Thr Ser Arg Cys Arg Asp
 65 70 75
 Ala Gly Thr Glu Leu Thr Gly His Leu Val Pro His His Asp Gly
 80 85 90
 Leu Arg Val Trp Cys Pro Glu Ser Glu Ala His Ile Pro Leu Pro
 95 100 105
 Pro Ala Pro Glu Gly Cys Pro Trp Ser Cys Arg Leu Leu Gly Ile
 110 115 120
 Gly Gly His Leu Ser Pro Gln Gly Lys Leu Thr Leu Pro Glu Glu
 125 130 135
 His Pro Cys Leu Lys Ala Pro Arg Leu Arg Cys Gln Ser Cys Lys
 140 145 150
 Leu Ala Gln Ala Pro Gly Leu Arg Ala Gly Glu Arg Ser Pro Glu
 155 160 165
 Glu Ser Leu Gly Gly Arg Arg Lys Arg Asn Val Asn Thr Ala Pro
 170 175 180
 Gln Phe Gln Pro Pro Ser Tyr Gln Ala Thr Val Pro Glu Asn Gln
 185 190 195
 Pro Ala Gly Thr Pro Val Ala Ser Leu Arg Ala Ile Asp Pro Asp
 200 205 210
 Glu Gly Glu Ala Gly Arg Leu Glu Tyr Thr Met Asp Ala Leu Phe
 215 220 225
 Asp Ser Arg Ser Asn Gln Phe Phe Ser Leu Asp Pro Val Thr Gly
 230 235 240
 Ala Val Thr Thr Ala Glu Glu Leu Asp Arg Glu Thr Lys Ser Thr
 245 250 255
 His Val Phe Arg Val Thr Ala Gln Asp His Gly Met Pro Arg Arg
 260 265 270
 Ser Ala Leu Ala Thr Leu Thr Ile Leu Val Thr Asp Thr Asn Asp
 275 280 285
 His Asp Pro Val Phe Glu Gln Gln Glu Tyr Lys Glu Ser Leu Arg
 290 295 300
 Glu Asn Leu Glu Val Gly Tyr Glu Val Leu Thr Val Arg Ala Thr
 305 310 315
 Asp Gly Asp Ala Pro Pro Asn Ala Asn Ile Leu Tyr Arg Leu Leu
 320 325 330

WO 01/98354

PCT/US01/19942

Glu	Gly	Ser	Gly	Gly	Ser	Pro	Ser	Glu	Val	Phe	Glu	Ile	Asp	Pro
				335					340					345
Arg	Ser	Gly	Val	Ile	Arg	Thr	Arg	Gly	Pro	Val	Asp	Arg	Glu	Glu
				350					355					360
Val	Glu	Ser	Tyr	Gln	Leu	Thr	Val	Glu	Ala	Ser	Asp	Gln	Gly	Arg
				365					370					375
Asp	Pro	Gly	Pro	Arg	Ser	Thr	Thr	Ala	Ala	Val	Phe	Leu	Ser	Val
				380					385					390
Glu	Asp	Asp	Asn	Asp	Asn	Ala	Pro	Gln	Phe	Ser	Glu	Lys	Arg	Tyr
				395					400					405
Val	Val	Gln	Val	Arg	Glu	Asp	Val	Thr	Pro	Gly	Ala	Pro	Val	Leu
				410					415					420
Arg	Val	Thr	Ala	Ser	Asp	Arg	Asp	Lys	Gly	Ser	Asn	Ala	Val	Val
				425					430					435
His	Tyr	Ser	Ile	Met	Ser	Gly	Asn	Ala	Arg	Gly	Gln	Phe	Tyr	Leu
				440					445					450
Asp	Ala	Gln	Thr	Gly	Ala	Leu	Asp	Val	Val	Ser	Pro	Leu	Asp	Tyr
				455					460					465
Glu	Thr	Thr	Lys	Glu	Tyr	Thr	Leu	Arg	Val	Arg	Ala	Gln	Asp	Gly
				470					475					480
Gly	Arg	Pro	Pro	Leu	Ser	Asn	Val	Ser	Gly	Leu	Val	Thr	Val	Gln
				485					490					495
Val	Leu	Asp	Ile	Asn	Asp	Asn	Ala	Pro	Ile	Phe	Val	Ser	Thr	Pro
				500					505					510
Phe	Gln	Ala	Thr	Val	Leu	Glu	Ser	Val	Pro	Leu	Gly	Tyr	Leu	Val
				515					520					525
Leu	His	Val	Gln	Ala	Ile	Asp	Ala	Asp	Ala	Gly	Asp	Asn	Ala	Arg
				530					535					540
Leu	Glu	Tyr	Arg	Leu	Ala	Gly	Val	Gly	His	Asp	Phe	Pro	Phe	Thr
				545					550					555
Ile	Asn	Asn	Gly	Thr	Gly	Trp	Ile	Ser	Val	Ala	Ala	Glu	Leu	Asp
				560					565					570
Arg	Glu	Glu	Val	Asp	Phe	Tyr	Ser	Phe	Gly	Val	Glu	Ala	Arg	Asp
				575					580					585
His	Gly	Thr	Pro	Ala	Leu	Thr	Ala	Ser	Ala	Ser	Val	Ser	Val	Thr
				590					595					600
Val	Leu	Asp	Val	Asn	Asp	Asn	Asn	Pro	Thr	Phe	Thr	Gln	Pro	Glu
				605					610					615
Tyr	Thr	Val	Arg	Leu	Asn	Glu	Asp	Ala	Ala	Val	Gly	Thr	Ser	Val
				620					625					630
Val	Thr	Val	Ser	Ala	Val	Asp	Arg	Asp	Ala	His	Ser	Val	Ile	Thr
				635					640					645
Tyr	Gln	Ile	Thr	Ser	Gly	Asn	Thr	Arg	Asn	Arg	Phe	Ser	Ile	Thr
				650					655					660
Ser	Gln	Ser	Gly	Gly	Gly	Leu	Val	Ser	Leu	Ala	Leu	Pro	Leu	Asp
				665					670					675
Tyr	Lys	Leu	Glu	Arg	Gln	Tyr	Val	Leu	Ala	Val	Thr	Ala	Ser	Asp
				680					685					690
Gly	Thr	Arg	Gln	Asp	Thr	Ala	Gln	Ile	Val	Val	Asn	Val	Thr	Asp
				695					700					705
Ala	Asn	Thr	His	Arg	Pro	Val	Phe	Gln	Ser	Ser	His	Tyr	Thr	Val
				710					715					720
Asn	Val	Asn	Glu	Asp	Arg	Pro	Ala	Gly	Thr	Thr	Val	Val	Leu	Ile
				725					730					735
Ser	Ala	Thr	Asp	Glu	Asp	Thr	Gly	Glu	Asn	Ala	Arg	Ile	Thr	Tyr
				740					745					750
Phe	Met	Glu	Asp	Ser	Ile	Pro	Gln	Phe	Arg	Ile	Asp	Ala	Asp	Thr
				755					760					765
Gly	Ala	Val	Thr	Thr	Gln	Ala	Glu	Leu	Asp	Tyr	Glu	Asp	Gln	Val
				770					775					780
Ser	Tyr	Thr	Leu	Ala	Ile	Thr	Ala	Arg	Asp	Asn	Gly	Ile	Pro	Gln
				785					790					795
Lys	Ser	Asp	Thr	Thr	Tyr	Leu	Glu	Ile	Leu	Val	Asn	Asp	Val	Asn

Asp	Asn	Ala	Pro	Gln	Phe	Leu	Arg	Asp	Ser	Tyr	Gln	Gly	Ser	Val	800	805	810
															815	820	825
Tyr	Glu	Asp	Val	Pro	Pro	Phe	Thr	Ser	Val	Leu	Gln	Ile	Ser	Ala			
															830	835	840
Thr	Asp	Arg	Asp	Ser	Gly	Leu	Asn	Gly	Arg	Val	Phe	Tyr	Thr	Phe			
															845	850	855
Gln	Gly	Gly	Asp	Asp	Gly	Asp	Gly	Asp	Phe	Ile	Val	Glu	Ser	Thr			
															860	865	870
Ser	Gly	Ile	Val	Arg	Thr	Leu	Arg	Arg	Leu	Asp	Arg	Glu	Asn	Val			
															875	880	885
Ala	Gln	Tyr	Val	Leu	Arg	Ala	Tyr	Ala	Val	Asp	Lys	Gly	Met	Pro			
															890	895	900
Pro	Ala	Arg	Thr	Pro	Met	Glu	Val	Thr	Val	Thr	Val	Leu	Asp	Val			
															905	910	915
Asn	Asp	Asn	Pro	Pro	Val	Phe	Glu	Gln	Asp	Glu	Phe	Asp	Val	Phe			
															920	925	930
Val	Glu	Glu	Asn	Ser	Pro	Ile	Gly	Leu	Ala	Val	Ala	Arg	Val	Thr			
															935	940	945
Ala	Thr	Asp	Pro	Asp	Glu	Gly	Thr	Asn	Ala	Gln	Ile	Met	Tyr	Gln			
															950	955	960
Ile	Val	Glu	Gly	Asn	Ile	Pro	Glu	Val	Phe	Gln	Leu	Asp	Ile	Phe			
															965	970	975
Ser	Gly	Glu	Leu	Thr	Ala	Leu	Val	Asp	Leu	Asp	Tyr	Glu	Asp	Arg			
															980	985	990
Pro	Glu	Tyr	Val	Leu	Val	Ile	Gln	Ala	Thr	Ser	Ala	Pro	Leu	Val			
															995	1000	1005
Ser	Arg	Ala	Thr	Val	His	Val	Arg	Leu	Leu	Asp	Arg	Asn	Asp	Asn			
															1010	1015	1020
Pro	Pro	Val	Leu	Gly	Asn	Phe	Glu	Ile	Leu	Phe	Asn	Asn	Tyr	Val			
															1025	1030	1035
Thr	Asn	Arg	Ser	Ser	Ser	Phe	Pro	Gly	Gly	Ala	Ile	Gly	Arg	Val			
															1040	1045	1050
Pro	Ala	His	Asp	Pro	Asp	Ile	Ser	Asp	Ser	Leu	Thr	Tyr	Ser	Phe			
															1055	1060	1065
Glu	Arg	Gly	Asn	Glu	Leu	Ser	Leu	Val	Leu	Leu	Asn	Ala	Ser	Thr			
															1070	1075	1080
Gly	Glu	Leu	Lys	Leu	Ser	Arg	Ala	Leu	Asp	Asn	Asn	Arg	Pro	Leu			
															1085	1090	1095
Glu	Ala	Ile	Met	Ser	Val	Leu	Val	Ser	Asp	Gly	Val	His	Ser	Val			
															1100	1105	1110
Thr	Ala	Gln	Cys	Ala	Leu	Arg	Val	Thr	Ile	Ile	Thr	Asp	Glu	Met			
															1115	1120	1125
Leu	Thr	His	Ser	Ile	Thr	Leu	Arg	Leu	Glu	Asp	Met	Ser	Pro	Glu			
															1130	1135	1140
Arg	Phe	Leu	Ser	Pro	Leu	Leu	Gly	Leu	Phe	Ile	Gln	Ala	Val	Ala			
															1145	1150	1155
Ala	Thr	Leu	Ala	Thr	Pro	Pro	Asp	His	Val	Val	Val	Phe	Asn	Val			
															1160	1165	1170
Gln	Arg	Asp	Thr	Asp	Ala	Pro	Gly	Gly	His	Ile	Leu	Asn	Val	Ser			
															1175	1180	1185
Leu	Ser	Val	Gly	Gln	Pro	Pro	Gly	Pro	Gly	Gly	Gly	Pro	Pro	Phe			
															1190	1195	1200
Leu	Pro	Ser	Glu	Asp	Leu	Gln	Glu	Arg	Leu	Tyr	Leu	Asn	Arg	Ser			
															1205	1210	1215
Leu	Leu	Thr	Ala	Ile	Ser	Ala	Gln	Arg	Val	Leu	Pro	Phe	Asp	Asp			
															1220	1225	1230
Asn	Ile	Cys	Leu	Arg	Glu	Pro	Cys	Glu	Asn	Tyr	Met	Arg	Cys	Val			
															1235	1240	1245
Ser	Val	Leu	Arg	Phe	Asp	Ser	Ser	Ala	Pro	Phe	Ile	Ala	Ser	Ser			
															1250	1255	1260
Ser	Val	Leu	Phe	Arg	Pro	Ile	His	Pro	Val	Gly	Gly	Leu	Arg	Cys			
															1265	1270	1275

Arg	Cys	Pro	Pro	Gly	Phe	Thr	Gly	Asp	Tyr	Cys	Glu	Thr	Glu	Val
				1280					1285					1290
Asp	Leu	Cys	Tyr	Ser	Arg	Pro	Cys	Gly	Pro	His	Gly	Arg	Cys	Arg
				1295					1300					1305
Ser	Arg	Glu	Gly	Gly	Tyr	Thr	Cys	Leu	Cys	Arg	Asp	Gly	Tyr	Thr
				1310					1315					1320
Gly	Glu	His	Cys	Glu	Val	Ser	Ala	Arg	Ser	Gly	Arg	Cys	Thr	Pro
				1325					1330					1335
Gly	Val	Cys	Lys	Asn	Gly	Gly	Thr	Cys	Val	Asn	Leu	Leu	Val	Gly
				1340					1345					1350
Gly	Phe	Lys	Cys	Asp	Cys	Pro	Ser	Gly	Asp	Phe	Glu	Lys	Pro	Tyr
				1355					1360					1365
Cys	Gln	Val	Thr	Thr	Arg	Ser	Phe	Pro	Ala	His	Ser	Phe	Ile	Thr
				1370					1375					1380
Phe	Arg	Gly	Leu	Arg	Gln	Arg	Phe	His	Phe	Thr	Leu	Ala	Leu	Ser
				1385					1390					1395
Phe	Ala	Thr	Lys	Glu	Arg	Asp	Gly	Leu	Leu	Leu	Tyr	Asn	Gly	Arg
				1400					1405					1410
Phe	Asn	Glu	Lys	His	Asp	Phe	Val	Ala	Leu	Glu	Val	Ile	Gln	Glu
				1415					1420					1425
Gln	Val	Gln	Leu	Thr	Phe	Ser	Ala	Gly	Glu	Ser	Thr	Thr	Thr	Val
				1430					1435					1440
Ser	Pro	Phe	Val	Pro	Gly	Gly	Val	Ser	Asp	Gly	Gln	Trp	His	Thr
				1445					1450					1455
Val	Gln	Leu	Lys	Tyr	Tyr	Asn	Lys	Pro	Leu	Leu	Gly	Gln	Thr	Gly
				1460					1465					1470
Leu	Pro	Gln	Gly	Pro	Ser	Glu	Gln	Lys	Val	Ala	Val	Val	Thr	Val
				1475					1480					1485
Asp	Gly	Cys	Asp	Thr	Gly	Val	Ala	Leu	Arg	Phe	Gly	Ser	Val	Leu
				1490					1495					1500
Gly	Asn	Tyr	Ser	Cys	Ala	Ala	Gln	Gly	Thr	Gln	Gly	Gly	Ser	Lys
				1505					1510					1515
Lys	Ser	Leu	Asp	Leu	Thr	Gly	Pro	Leu	Leu	Leu	Gly	Gly	Val	Pro
				1520					1525					1530
Asp	Leu	Pro	Glu	Ser	Phe	Pro	Val	Arg	Met	Arg	Gln	Phe	Val	Gly
				1535					1540					1545
Cys	Met	Arg	Asn	Leu	Gln	Val	Asp	Ser	Arg	His	Ile	Asp	Met	Ala
				1550					1555					1560
Asp	Phe	Ile	Ala	Asn	Asn	Gly	Thr	Val	Pro	Gly	Cys	Pro	Ala	Lys
				1565					1570					1575
Lys	Asn	Val	Cys	Asp	Ser	Asn	Thr	Cys	His	Asn	Gly	Gly	Thr	Cys
				1580					1585					1590
Val	Asn	Gln	Trp	Asp	Ala	Phe	Ser	Cys	Glu	Cys	Pro	Leu	Gly	Phe
				1595					1600					1605
Gly	Gly	Lys	Ser	Cys	Ala	Gln	Glu	Met	Ala	Asn	Pro	Gln	His	Phe
				1610					1615					1620
Leu	Gly	Ser	Ser	Leu	Val	Ala	Trp	His	Gly	Leu	Ser	Leu	Pro	Ile
				1625					1630					1635
Ser	Gln	Pro	Trp	Tyr	Leu	Ser	Leu	Met	Phe	Arg	Thr	Arg	Gln	Ala
				1640					1645					1650
Asp	Gly	Val	Leu	Leu	Gln	Ala	Ile	Thr	Arg	Gly	Arg	Ser	Thr	Ile
				1655					1660					1665
Thr	Leu	Gln	Leu	Arg	Glu	Gly	His	Val	Met	Leu	Ser	Val	Glu	Gly
				1670					1675					1680
Thr	Gly	Leu	Gln	Ala	Ser	Ser	Leu	Arg	Leu	Glu	Pro	Gly	Arg	Ala
				1685					1690					1695
Asn	Asp	Gly	Asp	Trp	His	His	Ala	Gln	Leu	Ala	Leu	Gly	Ala	Ser
				1700					1705					1710
Gly	Gly	Pro	Gly	His	Ala	Ile	Leu	Ser	Phe	Asp	Tyr	Gly	Gln	Gln
				1715					1720					1725
Arg	Ala	Glu	Gly	Asn	Leu	Gly	Pro	Arg	Leu	His	Gly	Leu	His	Leu
				1730					1735					1740
Ser	Asn	Ile	Thr	Val	Gly	Gly	Ile	Pro	Gly	Pro	Ala	Gly	Gly	Val

				1745					1750					1755
Ala	Arg	Gly	Phe	Arg	Gly	Cys	Leu	Gln	Gly	Val	Arg	Val	Ser	Asp
				1760					1765					1770
Thr	Pro	Glu	Gly	Val	Asn	Ser	Leu	Asp	Pro	Ser	His	Gly	Glu	Ser
				1775					1780					1785
Ile	Asn	Val	Glu	Gln	Gly	Cys	Ser	Leu	Pro	Asp	Pro	Cys	Asp	Ser
				1790					1795					1800
Asn	Pro	Cys	Pro	Ala	Asn	Ser	Tyr	Cys	Ser	Asn	Asp	Trp	Asp	Ser
				1805					1810					1815
Tyr	Ser	Cys	Ser	Cys	Asp	Pro	Gly	Tyr	Tyr	Gly	Asp	Asn	Cys	Thr
				1820					1825					1830
Asn	Val	Cys	Asp	Leu	Asn	Pro	Cys	Glu	His	Gln	Ser	Val	Cys	Thr
				1835					1840					1845
Arg	Lys	Pro	Ser	Ala	Pro	His	Gly	Tyr	Thr	Cys	Glu	Cys	Pro	Pro
				1850					1855					1860
Asn	Tyr	Leu	Gly	Pro	Tyr	Cys	Glu	Thr	Arg	Ile	Asp	Gln	Pro	Cys
				1865					1870					1875
Pro	Arg	Gly	Trp	Trp	Gly	His	Pro	Thr	Cys	Gly	Pro	Cys	Asn	Cys
				1880					1885					1890
Asp	Val	Ser	Lys	Gly	Phe	Asp	Pro	Asp	Cys	Asn	Lys	Thr	Ser	Gly
				1895					1900					1905
Glu	Cys	His	Cys	Lys	Glu	Asn	His	Tyr	Arg	Pro	Pro	Gly	Ser	Pro
				1910					1915					1920
Thr	Cys	Leu	Leu	Cys	Asp	Cys	Tyr	Pro	Thr	Gly	Ser	Leu	Ser	Arg
				1925					1930					1935
Val	Cys	Asp	Pro	Glu	Asp	Gly	Gln	Cys	Pro	Cys	Lys	Pro	Gly	Val
				1940					1945					1950
Ile	Gly	Arg	Gln	Cys	Asp	Arg	Cys	Asp	Asn	Pro	Phe	Ala	Glu	Val
				1955					1960					1965
Thr	Thr	Asn	Gly	Cys	Glu	Gly	Pro	Leu	Phe	Ala	Ser	Tyr	Cys	Pro
				1970					1975					1980
Arg	Pro	Met	Arg	Cys	Trp	Pro	Pro	Ala	Glu	Pro	Leu	Ser	Gln	Ser
				1985					1990					1995
Gln	Gly	Leu	Pro	Val	Cys	Leu	Pro	Glu	Ala	Gly	Pro	Phe	Gly	Phe
				2000					2005					2010
Leu	Pro	Pro	Gly	Thr	Ala	Val	Arg	His	Cys	Asp	Glu	His	Arg	Gly
				2015					2020					2025
Trp	Leu	Pro	Pro	Asn	Leu	Phe	Asn	Cys	Thr	Ser	Ile	Thr	Phe	Ser
				2030					2035					2040
Glu	Leu	Lys	Gly	Phe	Ala	Glu	Arg	Leu	Gln	Arg	Asn	Glu	Ser	Gly
				2045					2050					2055
Leu	Asp	Ser	Gly	Arg	Ser	Gln	Gln	Leu	Ala	Leu	Leu	Leu	Arg	Asn
				2060					2065					2070
Ala	Thr	Gln	His	Thr	Ala	Gly	Tyr	Phe	Gly	Ser	Asp	Val	Lys	Val
				2075					2080					2085
Ala	Tyr	Gln	Leu	Ala	Thr	Arg	Leu	Leu	Ala	His	Glu	Ser	Thr	Gln
				2090					2095					2100
Arg	Gly	Phe	Gly	Leu	Ser	Ala	Thr	Gln	Asp	Val	His	Phe	Thr	Glu
				2105					2110					2115
Asn	Leu	Leu	Arg	Val	Gly	Ser	Ala	Leu	Leu	Asp	Thr	Ala	Asn	Lys
				2120					2125					2130
Arg	His	Trp	Glu	Leu	Ile	Gln	Gln	Thr	Glu	Gly	Gly	Thr	Ala	Trp
				2135					2140					2145
Leu	Leu	Gln	His	Tyr	Glu	Ala	Tyr	Ala	Ser	Ala	Leu	Ala	Gln	Asn
				2150					2155					2160
Met	Arg	His	Thr	Tyr	Leu	Ser	Pro	Phe	Thr	Ile	Val	Thr	Pro	Asn
				2165					2170					2175
Ile	Val	Ile	Ser	Val	Val	Arg	Leu	Asp	Lys	Gly	Asn	Phe	Ala	Gly
				2180					2185					2190
Ala	Lys	Leu	Pro	Arg	Tyr	Glu	Ala	Leu	Arg	Gly	Glu	Gln	Pro	Pro
				2195					2200					2205
Asp	Leu	Glu	Thr	Thr	Val	Ile	Leu	Pro	Glu	Ser	Val	Phe	Arg	Glu
				2210					2215					2220

Thr	Pro	Pro	Val	Val	Arg	Pro	Ala	Gly	Pro	Gly	Glu	Ala	Gln	Glu	2225	2230	2235
Pro	Glu	Glu	Leu	Ala	Arg	Arg	Gln	Arg	Arg	His	Pro	Glu	Leu	Ser	2240	2245	2250
Gln	Gly	Glu	Ala	Val	Ala	Ser	Val	Ile	Ile	Tyr	Arg	Thr	Leu	Ala	2255	2260	2265
Gly	Leu	Leu	Pro	His	Asn	Tyr	Asp	Pro	Asp	Lys	Arg	Ser	Leu	Arg	2270	2275	2280
Val	Pro	Lys	Arg	Pro	Ile	Ile	Asn	Thr	Pro	Val	Val	Ser	Ile	Ser	2285	2290	2295
Val	His	Asp	Asp	Glu	Glu	Leu	Leu	Pro	Arg	Ala	Leu	Asp	Lys	Pro	2300	2305	2310
Val	Thr	Val	Gln	Phe	Arg	Leu	Leu	Glu	Thr	Glu	Glu	Arg	Thr	Lys	2315	2320	2325
Pro	Ile	Cys	Val	Phe	Trp	Asn	His	Ser	Ile	Leu	Val	Ser	Gly	Thr	2330	2335	2340
Gly	Gly	Trp	Ser	Ala	Arg	Gly	Cys	Glu	Val	Val	Phe	Arg	Asn	Glu	2345	2350	2355
Ser	His	Val	Ser	Cys	Gln	Cys	Asn	His	Met	Thr	Ser	Phe	Ala	Val	2360	2365	2370
Leu	Met	Asp	Val	Ser	Arg	Arg	Glu	Val	Gly	Pro	Thr	Gly	Ala	Ala	2375	2380	2385
Ala	Glu	Pro	Trp	Asn	Gly	Glu	Ile	Leu	Pro	Leu	Lys	Thr	Leu	Thr	2390	2395	2400
Tyr	Val	Ala	Leu	Gly	Val	Thr	Leu	Ala	Ala	Leu	Leu	Leu	Thr	Phe	2405	2410	2415
Phe	Phe	Leu	Thr	Leu	Leu	Arg	Ile	Leu	Arg	Ser	Asn	Gln	His	Gly	2420	2425	2430
Ile	Arg	Arg	Asn	Leu	Thr	Ala	Ala	Leu	Gly	Leu	Ala	Gln	Leu	Val	2435	2440	2445
Phe	Leu	Leu	Gly	Ile	Asn	Gln	Ala	Asp	Leu	Pro	Phe	Ala	Cys	Thr	2450	2455	2460
Val	Ile	Ala	Ile	Leu	Leu	His	Phe	Leu	Tyr	Leu	Cys	Thr	Phe	Ser	2465	2470	2475
Trp	Ala	Leu	Leu	Glu	Ala	Leu	His	Leu	Tyr	Arg	Ala	Leu	Thr	Glu	2480	2485	2490
Val	Arg	Asp	Val	Asn	Thr	Gly	Pro	Met	Arg	Phe	Tyr	Tyr	Met	Leu	2495	2500	2505
Gly	Trp	Gly	Val	Pro	Ala	Phe	Ile	Thr	Gly	Leu	Ala	Val	Gly	Leu	2510	2515	2520
Asp	Pro	Glu	Gly	Tyr	Gly	Asn	Pro	Asp	Phe	Cys	Trp	Leu	Ser	Ile	2525	2530	2535
Tyr	Asp	Thr	Leu	Ile	Trp	Ser	Phe	Ala	Gly	Pro	Val	Ala	Phe	Ala	2540	2545	2550
Val	Ser	Met	Ser	Val	Phe	Leu	Tyr	Ile	Leu	Ala	Ala	Arg	Ala	Ser	2555	2560	2565
Cys	Ala	Ala	Gln	Arg	Gln	Gly	Phe	Glu	Lys	Lys	Gly	Pro	Val	Ser	2570	2575	2580
Gly	Leu	Gln	Pro	Ser	Phe	Ala	Val	Leu	Leu	Leu	Leu	Ser	Ala	Thr	2585	2590	2595
Trp	Leu	Leu	Ala	Leu	Leu	Ser	Val	Asn	Ser	Asp	Thr	Leu	Leu	Phe	2600	2605	2610
His	Tyr	Leu	Phe	Ala	Thr	Cys	Asn	Cys	Ile	Gln	Gly	Pro	Phe	Ile	2615	2620	2625
Phe	Leu	Ser	Tyr	Val	Val	Leu	Ser	Lys	Glu	Val	Arg	Lys	Ala	Leu	2630	2635	2640
Lys	Leu	Ala	Cys	Ser	Arg	Lys	Pro	Ser	Pro	Asp	Pro	Ala	Leu	Thr	2645	2650	2655
Thr	Lys	Ser	Thr	Leu	Thr	Ser	Ser	Tyr	Asn	Cys	Pro	Ser	Pro	Tyr	2660	2665	2670
Ala	Asp	Gly	Arg	Leu	Tyr	Gln	Pro	Tyr	Gly	Asp	Ser	Ala	Gly	Ser	2675	2680	2685
Leu	His	Ser	Thr	Ser	Arg	Ser	Gly	Lys	Ser	Gln	Pro	Ser	Tyr	Ile			

WO 01/98354

PCT/US01/19942

2690	2695	2700
Pro Phe Leu Leu Arg Glu Glu Ser Ala Leu Asn Pro Gly Gln Gly		
2705	2710	2715
Pro Pro Gly Leu Gly Asp Pro Gly Ser Leu Phe Leu Glu Gly Gln		
2720	2725	2730
Asp Gln Gln His Asp Pro Asp Thr Asp Ser Asp Ser Asp Leu Ser		
2735	2740	2745
Leu Glu Asp Asp Gln Ser Gly Ser Tyr Ala Ser Thr His Ser Ser		
2750	2755	2760
Asp Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Ala Ala Phe		
2765	2770	2775
Pro Gly Glu Gln Gly Trp Asp Ser Leu Leu Gly Pro Gly Ala Glu		
2780	2785	2790
Arg Leu Pro Leu His Ser Thr Pro Lys Asp Gly Gly Pro Gly Pro		
2795	2800	2805
Gly Lys Ala Pro Trp Pro Gly Asp Phe Gly Thr Thr Ala Lys Glu		
2810	2815	2820
Ser Ser Gly Asn Gly Ala Pro Glu Glu Arg Leu Arg Glu Asn Gly		
2825	2830	2835
Asp Ala Leu Ser Arg Glu Gly Ser Leu Gly Pro Leu Pro Gly Ser		
2840	2845	2850
Ser Ala Gln Pro His Lys Gly Ile Leu Lys Lys Lys Cys Leu Pro		
2855	2860	2865
Thr Ile Ser Glu Lys Ser Ser Leu Leu Arg Leu Pro Leu Glu Gln		
2870	2875	2880
Cys Thr Gly Ser Ser Arg Gly Ser Ser Ala Ser Glu Gly Ser Arg		
2885	2890	2895
Gly Gly Pro Pro Pro Arg Pro Pro Pro Arg Gln Ser Leu Gln Glu		
2900	2905	2910
Gln Leu Asn Gly Val Met Pro Ile Ala Met Ser Ile Lys Ala Gly		
2915	2920	2925
Thr Val Asp Glu Asp Ser Ser Gly Ser Glu Gly		
2930	2935	

<210> 10

<211> 437

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 926992CD1

<400> 10

Met Ser Gln Thr Ala Gly Lys His Leu Leu Val Phe Leu Ile Leu		
1	5	10
Val Gly Ile Phe Ile Leu Ala Val Ser Arg Pro Arg Ser Ser Pro		
	20	25
Asp Asp Leu Lys Ala Leu Thr Arg Asn Val Asn Arg Leu Asn Glu		
	35	40
Ser Phe Arg Asp Leu Gln Leu Arg Leu Leu Gln Ala Pro Leu Gln		
	50	55
Ala Asp Leu Thr Glu Gln Val Trp Lys Val Gln Asp Ala Leu Gln		
	65	70
Asn Gln Ser Asp Ser Leu Leu Ala Leu Ala Gly Ala Val Gln Arg		
	80	85
Leu Glu Gly Ala Leu Trp Gly Leu Gln Ala Gln Ala Val Gln Thr		
	95	100
Glu Gln Ala Val Ala Leu Leu Arg Asp Arg Thr Gly Gln Gln Ser		
	110	115
Asp Thr Ala Gln Leu Glu Leu Tyr Gln Leu Gln Val Glu Ser Asn		
	125	130
Ser Ser Gln Leu Leu Leu Arg Arg His Ala Gly Leu Leu Asp Gly		

WO 01/98354

PCT/US01/19942

Leu	Ala	Arg	Arg	Val	Gly	Ile	Leu	Gly	Glu	Glu	Leu	Ala	Asp	Val	140	145	150
Gly	Gly	Val	Leu	Arg	Gly	Leu	Asn	His	Ser	Leu	Ser	Tyr	Asp	Val	155	160	165
Ala	Leu	His	Arg	Thr	Arg	Leu	Gln	Asp	Leu	Arg	Val	Leu	Val	Ser	170	175	180
Asn	Ala	Ser	Glu	Asp	Thr	Arg	Arg	Leu	Arg	Leu	Ala	His	Val	Gly	185	190	195
Met	Glu	Leu	Gln	Leu	Lys	Gln	Glu	Leu	Ala	Met	Leu	Asn	Ala	Val	200	205	210
Thr	Glu	Asp	Leu	Arg	Leu	Lys	Asp	Trp	Glu	His	Ser	Ile	Ala	Leu	215	220	225
Arg	Asn	Ile	Ser	Leu	Ala	Lys	Gly	Pro	Pro	Gly	Pro	Lys	Gly	Asp	230	235	240
Gln	Gly	Asp	Glu	Gly	Lys	Glu	Gly	Arg	Pro	Gly	Ile	Pro	Gly	Leu	245	250	255
Pro	Gly	Leu	Arg	Gly	Leu	Pro	Gly	Glu	Arg	Gly	Thr	Pro	Gly	Leu	260	265	270
Pro	Gly	Pro	Lys	Gly	Asp	Asp	Gly	Lys	Leu	Gly	Ala	Thr	Gly	Pro	275	280	285
Met	Gly	Met	Arg	Gly	Phe	Lys	Gly	Asp	Arg	Gly	Pro	Lys	Gly	Glu	290	295	300
Lys	Gly	Glu	Lys	Gly	Asp	Arg	Ala	Gly	Asp	Ala	Ser	Gly	Val	Glu	305	310	315
Ala	Pro	Met	Met	Ile	Arg	Leu	Val	Asn	Gly	Ser	Gly	Pro	His	Glu	320	325	330
Gly	Arg	Val	Glu	Val	Tyr	His	Asp	Arg	Arg	Trp	Gly	Thr	Val	Cys	335	340	345
Asp	Asp	Gly	Trp	Asp	Lys	Lys	Asp	Gly	Asp	Val	Val	Cys	Arg	Met	350	355	360
Leu	Gly	Phe	Arg	Gly	Val	Glu	Glu	Val	Tyr	Arg	Thr	Ala	Arg	Phe	365	370	375
Gly	Gln	Gly	Thr	Gly	Arg	Ile	Trp	Met	Asp	Asp	Val	Ala	Cys	Lys	380	385	390
Gly	Thr	Glu	Glu	Thr	Ile	Phe	Arg	Cys	Ser	Phe	Ser	Lys	Trp	Gly	395	400	405
Val	Thr	Asn	Cys	Gly	His	Ala	Glu	Asp	Ala	Ser	Val	Thr	Cys	Asn	410	415	420
Arg	His														425	430	435

<210> 11

<211> 325

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1002055CD1

<400> 11

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

WO 01/98354

PCT/US01/19942

				80					85					90
Glu	Thr	Asn	Glu	Thr	Lys	Glu	Thr	Asn	Cys	Thr	Asp	Glu	Arg	Ile
				95					100					105
Thr	Trp	Val	Ser	Arg	Pro	Asp	Gln	Asn	Ser	Asp	Leu	Gln	Ile	Arg
				110					115					120
Pro	Val	Ala	Ile	Thr	His	Asp	Gly	Tyr	Tyr	Arg	Cys	Ile	Met	Val
				125					130					135
Thr	Pro	Asp	Gly	Asn	Phe	His	Arg	Gly	Tyr	His	Leu	Gln	Val	Leu
				140					145					150
Val	Thr	Pro	Glu	Val	Thr	Leu	Phe	Gln	Asn	Arg	Asn	Arg	Thr	Ala
				155					160					165
Val	Cys	Lys	Ala	Val	Ala	Gly	Lys	Pro	Ala	Ala	Gln	Ile	Ser	Trp
				170					175					180
Ile	Pro	Glu	Gly	Asp	Cys	Ala	Thr	Lys	Gln	Glu	Tyr	Trp	Ser	Asn
				185					190					195
Gly	Thr	Val	Thr	Val	Lys	Ser	Thr	Cys	His	Trp	Glu	Val	His	Asn
				200					205					210
Val	Ser	Thr	Val	Thr	Cys	His	Val	Ser	His	Leu	Thr	Gly	Asn	Lys
				215					220					225
Ser	Leu	Tyr	Ile	Glu	Leu	Leu	Pro	Val	Pro	Gly	Ala	Lys	Lys	Ser
				230					235					240
Ala	Lys	Leu	Tyr	Ile	Pro	Tyr	Ile	Ile	Leu	Thr	Ile	Ile	Ile	Leu
				245					250					255
Thr	Ile	Val	Gly	Phe	Ile	Trp	Leu	Leu	Lys	Val	Asn	Gly	Cys	Arg
				260					265					270
Lys	Tyr	Lys	Leu	Asn	Lys	Thr	Glu	Ser	Thr	Pro	Val	Val	Glu	Glu
				275					280					285
Asp	Glu	Met	Gln	Pro	Tyr	Ala	Ser	Tyr	Thr	Glu	Lys	Asn	Asn	Pro
				290					295					300
Leu	Tyr	Asp	Thr	Thr	Asn	Lys	Val	Lys	Ala	Ser	Gln	Ala	Leu	Gln
				305					310					315
Ser	Glu	Val	Asp	Thr	Asp	Leu	His	Thr	Leu					
				320					325					

<210> 12

<211> 1251

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3998749CD1

<400> 12

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

				140					145					150
Cys	Val	Gly	Gly	Phe	Phe	Leu	Arg	Gly	Asn	Ser	Thr	Thr	Leu	Cys
				155					160					165
Gln	Pro	Asp	Gly	Thr	Trp	Ser	Ser	Pro	Leu	Pro	Glu	Cys	Val	Pro
				170					175					180
Val	Glu	Cys	Pro	Gln	Pro	Glu	Glu	Ile	Pro	Asn	Gly	Ile	Ile	Asp
				185					190					195
Val	Gln	Gly	Leu	Ala	Tyr	Leu	Ser	Thr	Ala	Leu	Tyr	Thr	Cys	Lys
				200					205					210
Pro	Gly	Phe	Glu	Leu	Val	Gly	Asn	Thr	Thr	Thr	Leu	Cys	Gly	Glu
				215					220					225
Asn	Gly	His	Trp	Leu	Gly	Gly	Lys	Pro	Thr	Cys	Lys	Ala	Ile	Glu
				230					235					240
Cys	Leu	Lys	Pro	Lys	Glu	Ile	Leu	Asn	Gly	Lys	Phe	Ser	Tyr	Thr
				245					250					255
Asp	Leu	His	Tyr	Gly	Gln	Thr	Val	Thr	Tyr	Ser	Cys	Asn	Arg	Gly
				260					265					270
Phe	Arg	Leu	Glu	Gly	Pro	Ser	Ala	Leu	Thr	Cys	Leu	Glu	Thr	Gly
				275					280					285
Asp	Trp	Asp	Val	Asp	Ala	Pro	Ser	Cys	Asn	Ala	Ile	His	Cys	Asp
				290					295					300
Ser	Pro	Gln	Pro	Ile	Glu	Asn	Gly	Phe	Val	Glu	Gly	Ala	Asp	Tyr
				305					310					315
Ser	Tyr	Gly	Ala	Ile	Ile	Ile	Tyr	Ser	Cys	Phe	Pro	Gly	Phe	Gln
				320					325					330
Val	Ala	Gly	His	Ala	Met	Gln	Thr	Cys	Glu	Glu	Ser	Gly	Trp	Ser
				335					340					345
Ser	Ser	Ile	Pro	Thr	Cys	Met	Pro	Ile	Asp	Cys	Gly	Leu	Pro	Pro
				350					355					360
His	Ile	Asp	Phe	Gly	Asp	Cys	Thr	Lys	Leu	Lys	Asp	Asp	Gln	Gly
				365					370					375
Tyr	Phe	Glu	Gln	Glu	Asp	Asp	Met	Met	Glu	Val	Pro	Tyr	Val	Thr
				380					385					390
Pro	His	Pro	Pro	Tyr	His	Leu	Gly	Ala	Val	Ala	Lys	Thr	Trp	Glu
				395					400					405
Asn	Thr	Lys	Glu	Ser	Pro	Ala	Thr	His	Ser	Ser	Asn	Phe	Leu	Tyr
				410					415					420
Gly	Thr	Met	Val	Ser	Tyr	Thr	Cys	Asn	Pro	Gly	Tyr	Glu	Leu	Leu
				425					430					435
Gly	Asn	Pro	Val	Leu	Ile	Cys	Gln	Glu	Asp	Gly	Thr	Trp	Asn	Gly
				440					445					450
Ser	Ala	Pro	Ser	Cys	Ile	Ser	Ile	Glu	Cys	Asp	Leu	Pro	Thr	Ala
				455					460					465
Pro	Glu	Asn	Gly	Phe	Leu	Arg	Phe	Thr	Glu	Thr	Ser	Met	Gly	Ser
				470					475					480
Ala	Val	Gln	Tyr	Ser	Cys	Lys	Pro	Gly	His	Ile	Leu	Ala	Gly	Ser
				485					490					495
Asp	Leu	Arg	Leu	Cys	Leu	Glu	Asn	Arg	Lys	Trp	Ser	Gly	Ala	Ser
				500					505					510
Pro	Arg	Cys	Glu	Ala	Ile	Ser	Cys	Lys	Lys	Pro	Asn	Pro	Val	Met
				515					520					525
Asn	Gly	Ser	Ile	Lys	Gly	Ser	Asn	Tyr	Thr	Tyr	Leu	Ser	Thr	Leu
				530					535					540
Tyr	Tyr	Glu	Cys	Asp	Pro	Gly	Tyr	Val	Leu	Asn	Gly	Thr	Glu	Arg
				545					550					555
Arg	Thr	Cys	Gln	Asp	Asp	Lys	Asn	Trp	Asp	Glu	Asp	Glu	Pro	Ile
				560					565					570
Cys	Ile	Pro	Val	Asp	Cys	Ser	Ser	Pro	Pro	Val	Ser	Ala	Asn	Gly
				575					580					585
Gln	Val	Arg	Gly	Asp	Glu	Tyr	Thr	Phe	Gln	Lys	Glu	Ile	Glu	Tyr
				590					595					600
Thr	Cys	Asn	Glu	Gly	Phe	Leu	Leu	Glu	Gly	Ala	Arg	Ser	Arg	Val
				605					610					615

WO 01/98354

PCT/US01/19942

Cys	Leu	Ala	Asn	Gly	Ser	Trp	Ser	Gly	Ala	Thr	Pro	Asp	Cys	Val
				620					625					630
Pro	Val	Arg	Cys	Ala	Thr	Pro	Pro	Gln	Leu	Ala	Asn	Gly	Val	Thr
				635					640					645
Glu	Gly	Leu	Asp	Tyr	Gly	Phe	Met	Lys	Glu	Val	Thr	Phe	His	Cys
				650					655					660
His	Glu	Gly	Tyr	Ile	Leu	His	Gly	Ala	Pro	Lys	Leu	Thr	Cys	Gln
				665					670					675
Ser	Asp	Gly	Asn	Trp	Asp	Ala	Glu	Ile	Pro	Leu	Cys	Lys	Pro	Val
				680					685					690
Asn	Cys	Gly	Pro	Pro	Glu	Asp	Leu	Ala	His	Gly	Phe	Pro	Asn	Gly
				695					700					705
Phe	Ser	Phe	Ile	His	Gly	Gly	His	Ile	Gln	Tyr	Gln	Cys	Phe	Pro
				710					715					720
Gly	Tyr	Lys	Leu	His	Gly	Asn	Ser	Ser	Arg	Arg	Cys	Leu	Ser	Asn
				725					730					735
Gly	Ser	Trp	Ser	Gly	Ser	Ser	Pro	Ser	Cys	Leu	Pro	Cys	Arg	Cys
				740					745					750
Ser	Thr	Pro	Val	Ile	Glu	Tyr	Gly	Thr	Val	Asn	Gly	Thr	Asp	Phe
				755					760					765
Asp	Cys	Gly	Lys	Ala	Ala	Arg	Ile	Gln	Cys	Phe	Lys	Gly	Phe	Lys
				770					775					780
Leu	Leu	Gly	Leu	Ser	Glu	Ile	Thr	Cys	Glu	Ala	Asp	Gly	Gln	Trp
				785					790					795
Ser	Ser	Gly	Phe	Pro	His	Cys	Glu	His	Thr	Ser	Cys	Gly	Ser	Leu
				800					805					810
Pro	Met	Ile	Pro	Asn	Ala	Phe	Ile	Ser	Glu	Thr	Ser	Ser	Trp	Lys
				815					820					825
Glu	Asn	Val	Ile	Thr	Tyr	Ser	Cys	Arg	Ser	Gly	Tyr	Val	Ile	Gln
				830					835					840
Gly	Ser	Ser	Asp	Leu	Ile	Cys	Thr	Glu	Lys	Gly	Val	Trp	Ser	Gln
				845					850					855
Pro	Tyr	Pro	Val	Cys	Glu	Pro	Leu	Ser	Cys	Gly	Ser	Pro	Pro	Ser
				860					865					870
Val	Ala	Asn	Ala	Val	Ala	Thr	Gly	Glu	Ala	Pro	Thr	Tyr	Glu	Ser
				875					880					885
Glu	Val	Lys	Leu	Arg	Cys	Leu	Glu	Gly	Tyr	Thr	Met	Asp	Thr	Asp
				890					895					900
Thr	Asp	Thr	Phe	Thr	Cys	Gln	Lys	Asp	Gly	Arg	Trp	Phe	Pro	Glu
				905					910					915
Arg	Ile	Ser	Cys	Ser	Pro	Lys	Lys	Cys	Pro	Leu	Pro	Glu	Asn	Ile
				920					925					930
Thr	His	Ile	Leu	Val	His	Gly	Asp	Asp	Phe	Ser	Val	Asn	Arg	Gln
				935					940					945
Val	Ser	Val	Ser	Cys	Ala	Glu	Gly	Tyr	Thr	Phe	Glu	Gly	Val	Asn
				950					955					960
Ile	Ser	Val	Cys	Gln	Leu	Asp	Gly	Thr	Trp	Glu	Pro	Pro	Phe	Ser
				965					970					975
Asp	Glu	Ser	Cys	Ser	Pro	Val	Ser	Cys	Gly	Lys	Pro	Glu	Ser	Pro
				980					985					990
Glu	His	Gly	Phe	Val	Val	Gly	Ser	Lys	Tyr	Thr	Phe	Glu	Ser	Thr
				995					1000					1005
Ile	Ile	Tyr	Gln	Cys	Glu	Pro	Gly	Tyr	Glu	Leu	Glu	Gly	Asn	Arg
				1010					1015					1020
Glu	Arg	Val	Cys	Gln	Glu	Asn	Arg	Gln	Trp	Ser	Gly	Gly	Val	Ala
				1025					1030					1035
Ile	Cys	Lys	Glu	Thr	Arg	Cys	Glu	Thr	Pro	Leu	Glu	Phe	Leu	Asn
				1040					1045					1050
Gly	Lys	Ala	Asp	Ile	Glu	Asn	Arg	Thr	Thr	Gly	Pro	Asn	Val	Val
				1055					1060					1065
Tyr	Ser	Cys	Asn	Arg	Gly	Tyr	Ser	Leu	Glu	Gly	Pro	Ser	Glu	Ala
				1070					1075					1080
His	Cys	Thr	Glu	Asn	Gly	Thr	Trp	Ser	His	Pro	Val	Pro	Leu	Cys

WO 01/98354

PCT/US01/19942

1085	1090	1095
Lys Pro Asn Pro Cys Pro Val Pro Phe Val Ile Pro Glu Asn Ala		
1100	1105	1110
Leu Leu Ser Glu Lys Glu Phe Tyr Val Asp Gln Asn Val Ser Ile		
1115	1120	1125
Lys Cys Arg Glu Gly Phe Leu Leu Gln Gly His Gly Ile Ile Thr		
1130	1135	1140
Cys Asn Pro Asp Glu Thr Trp Thr Gln Thr Ser Ala Lys Cys Glu		
1145	1150	1155
Lys Ile Ser Cys Gly Pro Pro Ala His Val Glu Asn Ala Ile Ala		
1160	1165	1170
Arg Gly Val His Tyr Gln Tyr Gly Asp Met Ile Thr Tyr Ser Cys		
1175	1180	1185
Tyr Ser Gly Tyr Met Leu Glu Gly Phe Leu Arg Ser Val Cys Leu		
1190	1195	1200
Glu Asn Gly Thr Trp Thr Ser Pro Pro Ile Cys Arg Ala Val Cys		
1205	1210	1215
Arg Phe Pro Cys Gln Asn Gly Gly Ile Cys Gln Arg Pro Asn Ala		
1220	1225	1230
Cys Ser Cys Gln Arg Ala Gly Trp Gly Ala Ser Val Lys Asn Gln		
1235	1240	1245
Ser Ala Phe Phe Pro Val		
1250		

<210> 13

<211> 3580

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6052371CB1

<400> 13

```

gccatggcgcg tccgggcccg cctgtggcca gcgctcctgg gcatagtcct cgccgcttgg 60
ctccgcggct cgggtgcccc gcagagtgcc accgtggcca acccagtgcc tgggtgccaac 120
ccggacctgc ttccccactt cctggtggag cccgaggatg tgtacatcgt caagaacaag 180
ccagtgtcgc ttgtgtgcaa ggccgtgccc gccacgcaga tcttcttcaa gtgcaacggg 240
gagtgggtgc gccaggtgga ccacgtgatc gagcgcagca cagacgggag cagtgggctg 300
cccaccatgg aggtccgcat taatgtctca aggcagcagg tcgagaaggt gttcgggctg 360
gaggaatact ggtgccagtg cgtggcatgg agctcctcgg gcaccaccaa gagtcagaag 420
gcctacatcc gcatagccta tttgcgcaag aacttcgagc aggagccgct ggccaaggag 480
gtgtccctgg agcagggcat cgtgctgccc tgccgtccac cggagggcat ccctccagcc 540
gaggtggagt ggctccggaa cgaggacctg gtggaccctg ccctggaccc caatgtatac 600
atcacgcggg agcacagcct ggtggtgcga caggcccgcc ttgctgacac ggccaactac 660
acctgcgtgg ccaagaacat cgtggcacgt cgccgcagcg cctccgctgc tgtcatcgtc 720
tacgtggacg gcagctggag cccgtggagc aagtggtcgg cctgtgggct ggactgcacc 780
cactggcgga gccgtgagtg ctctgaccca gcaccccgca acggagggga ggagtgccag 840
ggcactgacc tggacacccg caactgtacc agtgacctct gtgtacacac tgcttctggc 900
cctgaggacg tggccctcta tgtgggcctc atcgccgtgg ccgtctgcct ggtcctgctg 960
ctgcttgtcc tcatcctcgt ttattgccgg aagaaggagg ggctggactc agatgtggct 1020
gactcgtcca ttctcacctc aggcttccag cccgtcagca tcaagcccag caaagcagac 1080
aacccccatc tgctcaccat ccagccggac ctacgacca ccaccaccac ctaccagggc 1140
agtctctgtc cccggcagga tgggcccagc cccaagttcc agctcaccaa tgggcacctg 1200
ctcagccccc tgggtggcgg ccgccacaca ctgcaccaca gctctccac ctctgaggcc 1260
gaggagtctg tctcccgct ctccaccag aactacttcc gctccctgcc ccgaggcacc 1320
agcaacatga cctatgggac cttcaacttc ctcggggggc ggctgatgat ccctaataca 1380
ggaatcagcc tcctcatccc ccagatgcc ataccccgag ggaagatcta tgagatctac 1440
ctcacgctgc acaagccgga agacgtgagg ttgcccctag ctggctgtca gaccctgctg 1500
agtcccatcg ttagctgtgg accccctggc gtcctgctca cccggccagt catcctggct 1560
atggaccact gtggggagcc cagccctgac agctggagcc tgcgcctcaa aaagcagtcg 1620
tgcgagggca gctgggagga tgtgctgcac ctggggcgagg aggcgcctc ccacctctac 1680
tactgccagc tggaggccag tgcttctac gtcttcaccg agcagctggg ccgctttgcc 1740

```

```

ctggtgggag aggccctcag cgtggctgcc gccagcgcc tcaagctgct tctgtttgcg 1800
ccggtggcct gcacctccct cgagtacaac atccgggtct actgcctgca tgacacccac 1860
gatgcactca aggaggtggt gcagctggag aagcagctgg ggggacagct gatccaggag 1920
ccacgggtcc tgcacttcaa ggacagttac cacaacctgc gcctatccat ccacgatgtg 1980
cccagctccc tgtggaagag taagctcctt gtcagctacc aggagatccc cttttatcac 2040
atctggaatg gcacgcagcg gtacttgac tgcaccttca ccctggagcg tgtcagcccc 2100
agcactagtg acctggcctg caagctgtgg gtgtggcagg tggagggcga cgggcagagc 2160
ttcagcatca acttcaacat caccaaggac acaaggtttg ctgagctgct ggctctggag 2220
agtgaagcgg gggctcccagc cctgggtgggc ccagtgctct tcaagatccc cttcctcatt 2280
cggcagaaga taatttccag cctggaccca ccctgtaggc ggggtgccga ctggcggact 2340
ctggcccaga aactccacct ggacagccat ctcagcttct ttgcctcaa gccagcccc 2400
acagccatga tcctcaacct gtgggagggc cggcacttcc ccaacggcaa cctcagccag 2460
ctggctgcag cagtggctgg actgggcccag ccagacgctg gcctcttcac agtgtcggag 2520
gctgagtgtg gagggccggc agggccgaca cctacactct caccagcttt ggcacccacc 2580
aaggacaggc agaagccgga caggggcccct tccccacacc ggggagagct gctcggacag 2640
gccccctccc ggccgaagct gtcccttaat gctggctcct cagaccctgc ccgaactccc 2700
acctctccat ggctgccta gccaggctgg cactgccact cacactcggc cccagggccc 2760
aggagggaca gtgcctggag cctgggcccag gccagccca tctgtgtgtg tgtatgtgcg 2820
tgtgatgcta cctctcctcc cgtccctctc caggggcccc gcatacacac ggccatgcac 2880
gcacacactg ggctggggc agggccccag agctcctgcc tgagctggac cttatgcaaa 2940
catttctgtg cctgctgggt aggggcacgt ctgaggggccc ctgctccaag cctgcaggac 3000
cgagggccac agccggacag ggggtagccc ctggattcag gcacacgacc accacacgag 3060
cacgtgccac gcatgcctcg tgtgctcatc tcacacacac cccctcccg ggtcacgcag 3120
acacccccca accacacaca tctcatgccg tacacctgag gctgctcacg tctcacgccc 3180
agtgttggtg cacatttgcc tctcacatgc tgccctctcc acccaccag ggacacccca 3240
cggctcctcc ctgcccctgc cctcccccca gccttgaggt gccctgcccg gcggggcctg 3300
tgaatatgca atgggagtc caggctgtac agtgggtgag gtgtgtgtgg cgtggcgtgc 3360
ccgtccccag ggctggctgg tgccccacgc ggggcctgtc atgtgaagct cgtgtcctga 3420
ctttgtctta agtgcattca cgcacttact cttggcctta tgtacacagc cttgcccggc 3480
cgccggggca cataggggtt ttatcgggcg tgaatgtaaa taaattatat atatataattg 3540
ctaaaaaaaa aaaaaaaaaa attctgcggc cgcaagctta 3580

```

<210> 14

<211> 2429

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2642942CB1

<400> 14

```

cacagcaagg aggtagccca gcccgcgctt cggctgctct cgaggaggcc ggagtccccg 60
gagacgatgc gcccgcgca gccgcctgcg cctgcgggag ccggctgccc ttgagatgga 120
gttgctgcct ctttggtctt gcctgggttt tcaacttctg accgtgggct ggaggaacag 180
aagcggaaac gccacagcag cctcccaagg agtctgcaag ttggtgggtg gagccgctga 240
ctgccgaggg cagagcctcg cttcggtgcc cagcagcctc ccgccccacg cccggatgct 300
caccctggat gccaaccttc tcaagacctt gtggaatcac tccctccagc cttacctctt 360
cctggagagc ctcagcctgc acagctgcc a cctggagcgc atcagccgcg gcgccttcca 420
ggagcaaggc cacctgcgca gcctggctct gggggacaac tgcctctcag agaactacga 480
agagacggca gccgcccctc acgcccctgc gggcctgcg aggctggact tgtcaggaaa 540
cgccctgacg gaggacatgg cagccctcat gctccagaac ctctcctcgc tgcggctcgt 600
gtccctggcg gggaacacca tcatcgggct ggacgactcc gtcttcgagg gcctggagcg 660
tctccgggag ctggatctgc agaggaacta catcttcgag atcgagggag gcgctttcga 720
cggcctggct gagctgaggc acctcaacct ggccttcaac aacctcccct gcatcgtgga 780
cttcgggctc acgcggtgc gggctcctcaa cgtcagctac aacgtcctgg agtggttcct 840
cgcgaccggg ggagaggctg ccttcgagct ggagacgctg gacctgtctc acaaccagct 900
gctgttcttc ccgctgctgc ccagttacag caagttgcgg acctcctgc tgcgcgacaa 960
caacatgggc ttctaccggg acctgtacaa cacctcgtcg ccgagggaga tgggtggccca 1020
gttcctcctc gtggacggca acgtgaccaa catcaccacc gtcagcctct gggaagaatt 1080
ctcctccagc gacctcgcag atctccgctt cctggacatg agccagaacc agttccagta 1140
cctgccagac ggcttcctga ggaaaatgcc ttccctctcc cacctgaacc tccaccagaa 1200
ttgcctgatg acgcttcaca ttcgggagca cgagccccc ggagcgctca ccgagctgga 1260

```

```

cctgagccac aaccagctgt cggagctgca cctggctccg gggctggcca gctgcctggg 1320
cagcctgcgc ttgttcaacc tgagctccaa ccagctcctg ggcgtccccc ctggcctctt 1380
cgccaatgct aggaacatca ctacacttga catgagccac aatcagatct cactttgtcc 1440
cctgccagct gcctcggacc ggggtgggccc ccctagctgt gtggatttca ggaatatggc 1500
atctttaaagg agcctgtctc tggagggctg tggcctgggg gcattgccag actgcccatt 1560
ccaagggacc tccctgacct acttagacct ctcaagcaac tgggggggttc tgaatgggag 1620
cctcgcccca ctccaggatg ttgcccccat gttacaggtc ctgtctctca ggaacatggg 1680
cctccactcc agcttttatgg cgttggactt ctctggggtt gggaatctca gggacttaga 1740
tctgtcgggg aattgcttga ccaccttccc aagggttggg ggcagcctgg ccctggagac 1800
cctggatctc cgtagaaact cgctcacagc ccttccccag aaggctgtgt ctgagcagct 1860
ctcgagaggt ctgcggacca tctacctcag tcagaatcca tatgactgct gtgggggtga 1920
tggctggggg gccctgcagc atgggcagac ggtggccgac tgggccatgg tcacctgcaa 1980
cctctcctcc aagatcatcc gcgtgacgga gctgcccgga ggtgtgcctc gggactgcaa 2040
gtgggagcgg ctggacctgg gcctgctcta cctcgtgctc atcctcccca gctgcctcac 2100
cctgctgggtg gcctgcactg tcatcgtcct cacttttaag aagcctctgc ttcagggtcat 2160
caagagccgc tgccactggt cctccgttta ctgacctggc tgtgtgccaa gactcgaaat 2220
tcgggtccgca cacaacagga cactttctct gccagctttc aagatgtgat gcagaggcca 2280
agtctgacga attgaagttt caattaaaat ttaatatgtt tccattcctc atcgcccacc 2340
ccacccccgc ccccaccacc tgcccaagtt ctttttccat cattataaat tcatcctcat 2400
tatcttacta catagttatt aaagtactt                2429

```

<210> 15

<211> 3934

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3798924CB1

<400> 15

```

aggaacagca ataggccaga gattgggtcac atggaataat atggtgaaaa atacaggata 60
caaagcaaca ttagcaaatt atccctttaa atatgcagat gaacaagcca aaagccatcg 120
ggatgataga tggtcagatg atcattatga aagagagaaa agagaagttg actggaactt 180
ccacaaggac agctttttct gcgacgttcc aagtgaccga tattccagag tggatatttac 240
ttcatctgga ggggagacat tatggaattt acctgcaatt aaatcaatgt gcaatgtaga 300
taattccagg atcagatctc atccccagtt tggatgatctc tgccagagga ccactgctgc 360
ctcctgctgc cccagctgga cactgggaaa ctacatcgcc atctgaacaa tagatcgtcc 420
tgtcagaaaa tagttgagcg agacgtttct catacctttg aagctgcttc ggacttgtgc 480
caaacactac caaaatggca ctctggggcc agactgctgg gacatggcag ccagaagaaa 540
ggaccagctc aagtgcacca atgtgccacg caaatgtacc aagtacaatg ctgtgtacca 600
gacccctcat tacttggtgg acaaagactt tatgaccca aagacggctg actatgccac 660
gccagcttta aaatacagca tgctcttctc tcccacagag aaaggggaga gcatgatgaa 720
catttacttg gacaactttg aaaactggaa ctcttctgac ggcgtgacta ccatcaccgg 780
gattgagttt ggtatcaaac acagtttgtt tcaggattat cttctaattg atactgtgta 840
tcttgccata gccatcgtga ttgtcctttt agttatgtgt gtctacacca agtccatgtt 900
tatcactctg atgacaatgt ttgcaataat cagttctttg attgtttcct attttctcta 960
tcgtgtagta tttcacttcg aattttttcc ttttatgaac ctactgccc tcattatttt 1020
ggttgggaatt ggagcagatg atgcttttgt cctgtgtgat gtttggaaact acacaaaatt 1080
tgataagcct catgccgaaa cctcagaaac agtaagcatc accttgacgc acgctgccct 1140
ctccatgttc gtcaccagtt ttaccactgc tgctgccttt tatgctaact atgttagcaa 1200
cattacagca atccgatgct ttgggggttta tgcggggaca gctatatattg tgaattacgt 1260
tttgatggtc acatggcttc cagcagttgt tgtgctgcat gagcggtatc ttcttaatat 1320
attcacttgc ttcaaaaagc cccagcagca aatatatgat aacaaaagct gctggacagt 1380
ggcttgccag aagtgccaca aagtactctt tgccatttca gaagcatctc gaattttttt 1440
cgaaaaagta ttgccatgca ttgtcattaa gtttcgctac ctttggctgt tttggttcct 1500
tgctttaact gtaggtgggg cctacattgt atgtataaat ccaaagatga aactgccctc 1560
actggagtta tccgagttcc aggtgttccg gtctgcccac ccttttgagc gttatgatgc 1620
tgaatacaaa aagcttttca tgtttgaacg tgttcaccat ggcgaggagc tccacatgcc 1680
catcacagta atctggggcg tgtcccaga agacaatggc aaccactaa atcccaagag 1740
taaagggag ttgacattag atagcagttt taacatcgcc agccagctt cccaggcctg 1800
gattttgcac ttctgtcaaa aactgagaaa ccaaacattc ttttaccaga ctgatgaaca 1860
ggacttcacc agctgcttca ttgagacatt caaacagtgg atggaaaacc aggactgtga 1920

```

```

tgagcctgcc ctgtacccat gctgcagcca ctggagcttc ccctacaagc aagagatttt 1980
tgaactgtgc atcaagagag ctatcatgga gctggaaagg agtacagggt accattttgga 2040
tagcaaaacc ccagggccga ggtttgatat caatgatact atcagggcag tgggtgttaga 2100
gttccagagt acctacctct tcacactggc ttatgaaaag atgcatcagt ttataaaga 2160
ggtggactcg tggatatcca gtgagctgag ttccggccct gaaggcctca gcaatgggtg 2220
gtttgtcagc aatctggagt tctatgacct ccaggatagc ctctccgatg gcaccctcat 2280
tgccatgggg ctgtcagttg ctgttgcat ttagcgtgatg ctgctgacaa cttggaacat 2340
catcataagc ctttatgcca tcatttcaat tgctggaacg atatttgtca ctggttggtc 2400
tcttgtcctg ctgggctggg agctcaatgt gttggaatct gtcaccattt cgggtgccgt 2460
cggcttgtct gtagactttg ccgtccatta tgggggttgc taccgcttgg ctccagatcc 2520
cgaccgagaa ggcaaagtga tcttctctct gagtcgctg ggctctgcga tggccatggc 2580
tgccctgacc accttcgtgg caggggccat gatgatgcc tccacagttc tagcttacac 2640
ccagctgggc accttcatga tgctcatcat gtgtatcagt tgggctttcg ccaccttctt 2700
tttccagtgc atgtgccggg gccttggacc acagggtacc tgtggtcaga ttcctttacc 2760
taaaaaacta cagtgcagtg ccttttccca tgccttgtct acaagtccca gtgacaagg 2820
acaaagcaaa acacatacca taaatgctta tcathtagat cccagggggc caaatctga 2880
actggagcat gagttttatg aattagaacc tctggcttcc cacagctgca ctgcccctga 2940
gaagaccact tatgaagaga cccacatctg ctctgaattt ttcaacagcc aagcaaagaa 3000
tttagggatg cctgtgcatg cagcttacia cagtgaactc agcaaaagca ctgaaagtga 3060
cactggctct gccttggtac agccccctct tgaacagcat accgtgtgtc acttcttctc 3120
tctgaatcag agatgtagct gccccgatgc ctacaaacac ttgaactatg gccacactc 3180
ttgccagcag atgggggact gcttgtgcca ccagtgtct cctaccacta gcagctttgt 3240
ccagatccaa aacggcgtgg cacctctgaa ggccacacac caagctgtcg agggctttgt 3300
gcaccccatc acgcacatcc accactgtcc ctgcctgcag ggcagagtaa agccagccg 3360
aatgcagaat tctctgccta ggaatttttt cctccacca gtgcagcaca ttcaggccca 3420
agaaaaaatt ggcaagacca atgtacacag tcttcagagg agcatagaag agcatcttcc 3480
aaagatggca gagccatcgt catttgtctg cagaagcact ggatcgttac tcaaaacgtg 3540
ttgcgacccc gagaataaac aaagggaact ctgtaaaaat agagacgtga gcaatctgga 3600
gagcagtgga gggactgaaa acaaggcagg agggaaagtg gagctgagct tgtcacagac 3660
ggatgcaagt gtgaactcag aacatttcaa tcagaatgaa ccaaaagtcc tatttaata 3720
tttaatgggg gaggctggtt gtaggtcttg ccaaataat tcacaaagtt gtggcagaat 3780
tgtgagagtg aagtgcatt ctgtggactg tcaaatgcca aacatggaag ccaatgtgcc 3840
tgctgtatta acacactcgg aactttctgg tgaaagtttg ttaataaaaa cactataata 3900
aatgcagcat tcaattcaga aaaaaaaaaa aaaa 3934

```

<210> 16

<211> 1633

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4586653CB1

<400> 16

```

ggggtggagg aaccgacagg aggcggggag cccccaccta ccccttgtgg agctgcagga 60
gcaagggcat gcagccagtc atgctggccc tgtggtccct gcttctgctc tggggcctgg 120
cgactccatg ccaggagctg ctagagacgg tgggcacgct cgctcggatt gacaaggatg 180
aactcggcaa agccatccag aactcactgg ttggggagcc cattctgcag aatgtgctgg 240
gatcggtcac agctgtgaac cggggcctct tgggctcagg agggctgctt ggaggaggcg 300
gcttgtctgg ccacggaggg gtttttggcg ttgtcgagga gctctctggg ctgaagattg 360
aggagctcac gctgccaaag gtgttgctga agctgctgcc gggatttggg gtgcagctga 420
gcctgcacac caaagtgggc atgcattgct ctggccccct tgggtggcctt ctgcagctgg 480
ctgcggagggt gaacgtgaca tcgcgggtgg cgctggccgt gagctcaagg ggcacacca 540
tccttatcct caagcgtgc agcacgctcc tgggccacat cagcctgttc tcagggtgc 600
tgccacacc actctttggg gtcgtggaac agatgctctt caaggtgctt ccgggactgc 660
tgtgccccgt ggtggacagt gtgctgggtg tggatgaatga gctcctgggg gctgtgctgg 720
gcctggtgtc ccttggggct cttgggtccg tggaattctc tctggccaca ttgcctctca 780
tctccaacca gtacatagaa ctggacatca accctatcgt gaagagtgtg gctgggtgata 840
tcattgactt ccccaagtcc cgtgccccag ccaaggtgcc ccccaagaag gaccacacat 900
cccaggtgat ggtgccactg tacctcttca acaccacgtt tggactcctg cagaccaacg 960
gcgccctcga catggacatc acccctgagc tgggtcccg cgatgtccca ctgacaacta 1020
cagacctggc agctttgctc cctgaggtca tgactgtgcg tgcccagctg gctccctcgg 1080

```

WO 01/98354

PCT/US01/19942

```

ctaccaagct gcacatctcc ctgtccctgg aacggctcag tgtcaagggtg gcctcctcct 1140
ttacccatgc ctttgacgga tcgctgttag aagaatggct cagccatgtg gtcggggcag 1200
tgtatgcacc aaagcttaac gtggccctgg atgttggaat tcccctgcct aaggttctta 1260
atatcaatth ttccaattca gttctggaga tcgtagagaa tgctgttggtg ctgaccgtgg 1320
catcctgagg ctgagacatg gccaccagcc ttccctgttg actactagag accacctgtc 1380
tactctgcct caatttccct cccagtctct agctgatgtt ggtgacagta aaaatgccct 1440
ctggccctga agcactactc caagtttggg gtgggaactg cctggctaata catagaactg 1500
cctcagagag agctctgggg cctcagtagc aaaccctgag ctttcaataa atgactcctg 1560
tatctcggta aaaaaaaaaa aaaggggggg cgcgcgaaaa gggagcccgt ggaccgggga 1620
ataaatcccg gac 1633

```

<210> 17

<211> 879

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5951460CB1

<400> 17

```

gcgtgcttac acagctcgga caaagccagg ttgctttgag caaggctggc gacaagatca 60
ccatgtacag cttcatgggt ggtggcctgt tctgtgcctg ggtggggacc atcctcctgg 120
tggtggccat ggcaacagac cactggatgc agtaccggct gtcagggtcc ttcgcccacc 180
agggcctgtg gcggtactgc ctgggcaaca agtgctacct gcagacagac agcatcgcat 240
actggaatgc caccggggcc ttcatgatcc tgtctgccct atgcgccatc tccggcatca 300
tcatgggcat catggccttc gctcatcagc ctaccttctc ccgcctctcc cggcccttct 360
ctgctggcat catgtttttt tcctcaacc ttttcgtcgt gttggccttg gccatctaca 420
ctggagtcac cgtcagcttc ctgggcccgc gctttgggga ctggcgcttt tcctggtcct 480
acatcctggg ctgggtggca gtgctcatga cgttcttcgc agggattttc tacatgtgcg 540
cctaccgggt gcatgaatgc cggcgctgt ctacaccccg ctgagcccaa atgtgtcccc 600
caacttcac tggaaagttaa agtgaggcca ctgaagagga ggaggagggt ctagaggcct 660
gaaatcctgg ttcctagggg aatgaggggg ctcagttctg gactgtgggt ttgtgggggg 720
aggctgactc ctggtcctag gctggaagga ggaagaatag ggcccatggg agggagctga 780
gaagactcaa gtccccgtct gcctggcagg ttgttagaaa aatggactat ccattagagc 840
aactttctgg ggcctaataa aactgatgtg aaactaaaa 879

```

<210> 18

<211> 2085

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1534444CB1

<400> 18

```

atggagtggg gttacctgtt ggaagtgacc tcgctgctgg ccgccttggc gctgctgcag 60
cgctctagcg gcgctgcggc cgccctcgcc aaggagctgg catgccaaga gatcaccgtg 120
ccgctgtgta agggcatcgg ctacaactac acctacatgc ccaatcagtt caaccacgac 180
acgcaagacg aggcgggcct ggagggtgcac cagttctggc cgctgggtgga gatccagtgc 240
tcgcccgatc tcaagttctt cctgtgcagc atgtacacgc ccatctgcct agaggactac 300
aagaagccgc tgcgcacctg ccgctcgggt tgcgagcgcg ccaaggccgg ctgcgcgccc 360
ctcatgcgcc agtacggctt cgccctggccc gaccgcatgc gctgcgaccg gctgcccag 420
caaggcaacc ctgacacgct gtgcatggac tacaaccgca ccgacctaac caccgccgcg 480
cccagcccgc cgcgcgcctt gccgcgcgcg ccgcccggcg agcagccgcc ttcgggcagc 540
ggccacggcc gccgcgcggg ggccaggccc ccgcaccgcg gcggcgggcag gggcggtggc 600
ggcggggacg cggcggcgcc ccagctcgc ggcgcgggcg gtggcgggaa ggcgcggccc 660
cctggcgggc gcgcggctcc ctgcgagccc ggggtgccagt gccgcgcgcc tatggtgagc 720
gtgtccagcg agcgccaccc gctctacaac cgcgtcaaga caggccagat cgctaactgc 780
gcgctgccct gccacaacc ctttttcagc caggacgagc gcgccttcac cgtcttctgg 840
atcggcctgt ggtcgggtgt ctgcttcgtg tccaccttcg ccaccgtctc caccttctct 900
atcgacatgg agcgcttcaa gtacccggag cggcccatta tcttctctc ggctgctac 960

```

```

ctcttcgtgt cggtgggcta cctagtgcgc ctggtggcgg gccacgagaa ggtggcgtgc 1020
agcgggtggcg cgccggggcgc gggggggcgct ggggggcgcgg gcggcgcggc ggcggggcgcg 1080
ggcgcgggcgg gcgcggggcgc gggcgggcccg ggcggggcgcg gcgagtacga ggagctgggc 1140
gcggtggagc agcacgtgcg ctacgagacc accggccccg cgctgtgcac cgtgggtcttc 1200
ttgctggtct acttcttcgg catggccagc tccatctggt gggatgatctt gtcgctcaca 1260
tggttcctgg cggccggtat gaagtggggc aacgaagcca tcgccggcta ctgcagtac 1320
ttccacctgg ccgctgggct tgtgcccagc gtcaagtcca tcgctgggtgct ggcgctcagc 1380
tcggtggacg gcgaccgggt ggcggggcatc tgctacgtgg gcaaccagag cctggacaac 1440
ctgcgcggct tcgtgctggc gccgctgggtc atctacctct tcatcggcac catgttcctg 1500
ctggccggct tcgtgtccct cttccgcac cgcctcggtca tcaagcaaca ggacggcccc 1560
accaagacgc acaagctgga gaagctgatg atccgcctgg gcctgttcac cgtgctctac 1620
accgtgcccc cgcggtgggt ggtcgctgc ctcttctacg agcagcaca cgcggcgcg 1680
tgaggaggcca cgcacaactg cccgtgcctg cgggacctgc agcccgacca ggcacgcagg 1740
cccgactacg ccgtcttcat gctcaagtac ttcagtgtgc tagtggtggg catcacctcg 1800
ggcgtgtggg tctgggtccg caagacgtg gagtcctggc gctccctgtg caccgctgc 1860
tgctgggcca gcaagggcgc cgcggtgggc ggggggcgcgg gcgccacggc cgcggggggg 1920
ggcgggcggg cggggggcgg cggcgggcgg ggaccggcg gcggcggggg gccggggcgg 1980
ggcggggggt ccctctacag cgacgtcagc actggcctga cgtggcggtc gggcacggcg 2040
agctccgtgt cttatccaaa gcagatgcca ttgtcccagg tctga 2085

```

<210> 19

<211> 5497

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6777669CB1

<400> 19

```

atgcgggggg cgcccgcgcg cctgctgctg ccgctgctgc cgtggctcct gctgctcctg 60
gcgcccagag ctccggggcgc gcccggtgc ccgctatcca tccgcagctg caagtgctcg 120
ggggagcggc ccaaggggct gagcggcggc gtccctggcc cggctcggcg gaggggtggg 180
tgcagcggcg gggacctccc ggagcctccc gagcccgccc ttctgcctaa cggcacctgt 240
accctgctct tgagcaataa caagatcacg gggctccgca atggctcctt cctgggactg 300
tactgctgg agaagctgga cctgaggaac aacatcatca gcacagtga gccgggcgc 360
ttcctgggccc tgggggagct gaagcgttta gatctctcca acaaccggat tggctgtctc 420
acctccgaga ccttccaggg cctccccagg cttctccgac taaacatata tggaaacatc 480
ttctccagtc tgcaacctgg ggtctttgat gagctgccag cccttaaggt tgtggacttg 540
ggcacccagt tcctgacctg tgactgccac ctgcgctggc tgctgccctg ggcccagaat 600
cgctccctgc agctgtcggg acacacgctc tgtgcttacc ccagtgcctt gcatgctcag 660
gccctgggca gcctccagga ggcccagctc tgctgcgagg gggccctgga gctgcacaca 720
caccacctca tcccgctccc acgccaagtg gtgttccagg gggatcggct gcccttccag 780
tgctctgcca gctacctggg caacgacacc cgcctccgct ggtaccacaa ccgagcccc 840
gtggagggtg atgagcaggc gggcatcctc ctggccgaga gcctcatcca cgactgcacc 900
ttcatcacca gtgagctgac gctgtctcac atcggcgtgt gggcctcagg cgagtgggag 960
tgcaccgtgt ccatggccca aggcaacgcc agcaagaagg tggagatcgt ggtgctggag 1020
acctctgct cctactgccc cgccgagcgt gttgccacaa accgcgggga cttcaggtgg 1080
ccccgaactc tggctggcat cacagcctac cagtcctgcc tgcagtatcc cttcacctca 1140
gtgcccctgg gcgggggtgc cccgggcacc cgagcctccc gccggtgtga ccgtgccggc 1200
cgctgggagc caggggacta ctcccactgt ctctacacca acgacatcac caggggtgctg 1260
tacaccttcg tgctgatgcc catcaatgcc tccaatgcgc tgaccctggc tcaccagctg 1320
cgctgtaca cagccgaggc cgctagcttt tcagacatga tggatgtagt ctatgtggct 1380
cagatgatcc agaaattttt gggttatgtc gaccagatca aagagctggg agaggtgatg 1440
gtggacatgg ccagcaacct gatgctgggt gacgagcacc tgctgtggct ggcccagcgc 1500
gaggacaagg cctgcagccg catcgtgggt gccctggagc gcattggggg ggcgcgccctc 1560
agcccccatg cccagcacat ctcagtgaat gcgaggaacg tggcatgtga ggcctacctc 1620
atcaagccgc acagctacgt gggcctgacc tgcacagcct tccagaggag ggaggagg 1680
gtgccgggca cacggccagg aagccctggc cagaaccccc cacctgagcc cgagccccca 1740
gctgaccagc agctccgctt ccgctgcacc accgggaggc ccaatgtttc tctgtcgtcc 1800
ttccacatca agaacagcgt ggccctggcc tccatccagc tgcccccgag tctattctca 1860
tcccttccgg ctgcccctgg tcccccggtg cccccagact gcaccctgca actgctcgtc 1920
ttccgaaatg gccgcctctt ccacagccac agcaacacct cccgcccctg agctgctggg 1980

```

cctggcaaga	ggcgtggcgt	ggccaccccc	gtcatcttctg	caggaaccag	tggctgtggc	2040
gtgggaaacc	tgacagagcc	agtggccggt	tcgctgcggc	actgggctga	gggagccgaa	2100
cctgtggccg	cttgggtggag	ccaggagggg	cccggggagg	ctgggggctg	gacctcggag	2160
ggctgccagc	tccgctccag	ccagcccaat	gtcagcgccc	tgactgccca	gcacttgggc	2220
aatgtggccg	tgctcatgga	gctgagcgcc	tttcccaggg	aggtgggggg	cgccggggca	2280
gggctgcacc	ccgtggtata	cccctgcacg	gccttgctgc	tgctctgcct	cttcgccacc	2340
atcatcacct	acatcctcaa	ccacagctcc	atccgtgtgt	cccggaaagg	ctggcacatg	2400
ctgctgaact	tgtgcttcca	catagccatg	acctctgctg	tctttgcggg	gggcatcaca	2460
ctcaccaact	accagatggt	ctgccaggcg	gtgggcatca	ccctgcacta	ctcctcccta	2520
tccacgctgc	tctggatggg	cgtgaaggcg	cgagtgtctc	ataaggagct	cacctggagg	2580
gcacccctc	cgcaagaagg	ggaccccgct	ctgcctactc	ccagtcctat	gctccgggtc	2640
tatttgatcg	ctggagggat	tccactcatt	atctgtggca	tcacagctgc	agtcaacatc	2700
cacaactacc	gggaccacag	cccctactgc	tggctgggtg	ggcgtccaag	ccttggcgcc	2760
ttctacatcc	ctgtggcttt	gattctgctc	atcacctgga	tctatttctc	gtgcgccggg	2820
ctacgcttac	ggggtcctct	ggcacagaac	ccaaggcgg	gcaacagcag	ggcctccctg	2880
gaggcagggg	aggagctgag	gggttccacc	aggctcaggg	gcagcggccc	cctcctgagt	2940
gactcaggtt	cccttcttgc	tactgggagc	gcgcgagtgg	ggacgcccgg	gcccccgagg	3000
gatggtgaca	gcctctattc	tccgggagtc	cagctagggg	cgctgggtgac	cacgcacttc	3060
ctgtacttgg	ccatgtgggc	ctgcggggct	ctggcagtgt	cccagcgctg	gctgccccgg	3120
gtggtgtgca	gctgcttgta	cgggggtggca	gcctccgccc	tgggcctctt	cgtcttctact	3180
caccactgtg	ccaggcggag	ggacgtgaga	gcctcgtggc	gcgcctgctg	ccccctgcc	3240
tctcccgcgg	ccccccatgc	cccgccccgg	gccttgcccc	ccgcccgcaga	ggacgggttc	3300
ccggtgttcg	gggagggccc	cccctccctc	aagtcctccc	caagcggcag	cagcggccat	3360
ccgctggctc	tgggcccctg	caagctcacc	aacctgcagc	tggcccagag	tcaggtgtgc	3420
gaggcggggg	cggcgggccg	cggggaagga	gagccggagc	cggcgggcac	ccggggaaac	3480
ctcgcccacc	gccaccccaa	caacgtgcac	cacgggcgtc	gggcgcacaa	gagccggggc	3540
aagggaacacc	gcgcggggga	ggcctgcggc	aagaaccggc	tcaaggccct	gcgcgggggg	3600
gcggcggggg	cgctggagct	gctgtccagc	gagagcggtg	gtctgcacaa	cagccccacc	3660
gacagctacc	tgggcagcag	ccgcaacagc	ccgggcgccc	gcctgcagct	ggaaggcgag	3720
cccatgctca	cgccgtccga	gggcagcgac	accagcgccg	cgccgctttc	tgaggcgggc	3780
cgggcaggcc	agcgccgcag	cgccagccgc	gacagtctca	agggcgggcg	cgcgctggag	3840
aaggagagcc	atcgccgctc	gtacccgctc	aacgcccga	gcctaaacgg	cgcccccaag	3900
gggggcaagt	acgacgacgt	caccctgatg	ggcgcgagg	tagccagcgg	cggctgcatg	3960
aagaccggac	tctggaagag	cgaaactacc	gtctaagggt	gggcgggcga	cgcggtagac	4020
gggctggcca	cgcggtctcg	tccccgctc	ctcgggggcc	tccaagggtg	ctccgtagtc	4080
agcaggttgg	aggcagagga	gccgatggct	ggaggaagcc	cacaggcgga	tgttccccac	4140
ttgcctagag	ggcatccctc	tggggtagcg	acagacaatc	ccagaaacac	gcataataca	4200
tttccgtcca	gccccggggc	gtctgactgt	cggtgccctc	ccaggaacgg	ggaaggcctc	4260
cgtctgtgtg	aaagggcaca	gcacatccca	ggtgcaccct	ccccaaagtac	tcccaccccc	4320
cctactgtcc	atgcggcctc	actggggggc	atcagcctca	ccagcaaagc	agagatgaga	4380
gcgtgggaac	tgtgttcttt	cctccctgcc	ctctactgat	ttcagcccag	cccctgccta	4440
gatcctaggt	cccttttctc	cccaggtttg	gctggcacga	gagctagccc	agcacatgaa	4500
gcagggtgatg	ttaagtcaca	agggtgctgct	tttcagatcc	actatgcaag	aggggagggg	4560
ggggccacgt	gaaaggcagc	tctagacatc	aaccagtcct	gggggagggg	agtgggaacc	4620
gggcacaact	aggaacaatg	ccaccattcc	cacaggagtg	gtacttaaac	cagacagcag	4680
ggttcagagg	tggcacaccg	ggacaaagct	gaggccctgc	acctcaacag	ctgactgccca	4740
ggtgcctgtg	ggtgaactga	ggggagtaga	gggagagggc	aggtggaact	ggggcagaat	4800
ctagtcatgc	cctaaagcta	gtcctgtaaa	caatggtgcc	ccagaaagct	gcagggtggtg	4860
tttgggagaag	cagttacttt	tcagttacaa	gacccatctc	cctagtctca	gccttacaac	4920
accacgggac	taaggaagag	cacttccttg	cctccgtaag	gccagaggaa	gaaccatccc	4980
aatcatttga	tctccagctc	cacagtagag	agaaacctac	aaaatgtcaa	accagcttcc	5040
cgactcccag	gagctcaagc	caagcccaga	ggcagtggct	ggggtcctctg	caggtcatga	5100
ggggcctatg	cctttactcc	ttttaaacac	cagcaccctg	cttttcccca	acctaaaacc	5160
aaccaccagc	atttccctcc	cagtcttcac	atcactctgg	cctcatcacc	aaggtgacag	5220
aggacacagg	ggagggggaa	aaccacacac	cactccttgg	aatgggtcct	gttatttatg	5280
cttgctgcac	agacatatta	gaagaaaaaa	aaaagctttg	tattattctt	ccacatatgc	5340
tggctgctgt	ttacacaccc	tgccaatgcc	ttagcactgg	agagcttttt	gcaatatgct	5400
ggggaaaggg	gagggagggg	atgaaagtgc	caaagaaaac	atgtttttta	gaactcgggt	5460
tttatacaat	agaatgtttt	ctagcagaaa	aaaaaaa			5497

<210> 20

<211> 10123

<212> DNA

WO 01/98354

PCT/US01/19942

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1897612CB1

<400> 20

atgaagagcc	ccaggcccca	cctcctgcta	ccattgctgc	tgctgctgct	gctgctgctg	60
ggggctgggg	tgccagggtgc	ctgggggtcag	gctgggagcc	tggaacttgca	gattgatgag	120
gagcagccag	cgggtacact	gattggcgac	atcagtgcgg	ggcttccggc	aggcacggca	180
gctcctctca	tgtacttcat	ctctgcccac	gagggcagcg	gcgtgggcac	agacctggcc	240
attgacgaac	acagtggggt	cgtccgtaca	gcccgtgtct	tggaacctga	gcagcgggac	300
cgctaccgct	tcactgcagt	cactcctgat	ggtgccaccg	tagaagttac	agtgcgagtg	360
gctgacatca	acgacctatgc	tccagccttc	ccacaggctc	gggctgccct	gcaggtaacct	420
gagcatacag	cttttggcac	ccgctaccca	ctggagcctg	ctcgtgatgc	agatgctggg	480
cgtctgggaa	cccagggtcta	tgcgctatct	ggtgatgggg	ctggagagac	cttccggctg	540
gagacacgcc	ccggtccaga	tgggactcca	gtacctgagc	tggttagttac	tggggaactg	600
gaccgagaga	accgctcaca	ctatatgcta	cagctggagg	cctatgatgg	tgggttcaccc	660
ccccggaggg	cccaggccct	gctggacgtg	acactgctgg	acatcaatga	ccatgccccg	720
gctttcaatc	agagccgcta	ccatgctgtg	gtgtctgaga	gcctggcccc	tggcagtcct	780
gtcttgcagg	tgttcgcata	tgatgccgat	gctggtgtca	atggggctgt	gacttacgag	840
atcaaccgga	ggcagagcga	gggtgatgga	cccttctcca	tcgacgcaca	cacggggctg	900
ctgcagttag	agcggccact	ggactttgag	cagcggcggg	tccatgaact	ggtggtgcaa	960
gcacgagatg	atggctcccc	ccaagtgtct	gagggccgcc	cacctggaca	gctcgttgct	1020
cgcactctctg	tgtcagaccc	agatgatggg	gactttgccc	atgtcaatgt	gtccctggaa	1080
ggtggagagg	gccactttgc	cctaagcacc	caagacagcg	tcacttatct	ggtgtgtgtg	1140
gctcggcggc	tggatcgaga	ggagagggat	gcctataact	tgagggttac	agccacagac	1200
tcaggctcac	ctccactgcg	ggctgaggct	gcctttgtgc	tgcacgtcac	tgatgtcaac	1260
gacaatgcac	ctgcctttga	ccgccagctc	taccgacctg	agccccctgc	tgagggttgcg	1320
ctgcctggca	gctttgtagt	gcgggtgact	gctcgggatc	ctgaccaagg	caccaatggt	1380
caggctcactt	atagcctagc	ccctggcgcc	cacacccact	ggttctccat	tgaccccacc	1440
tcaggcatta	tcactacggc	tgcctcactg	gactatgagt	tggaaacctca	gccacagctg	1500
attgtggtgg	ccacagatgg	tggcctgccc	cctctagcct	cctctgccac	agttagcgtg	1560
gccttgcaag	atgtgaatga	taatgagccc	caattccaga	ggactttcta	caatgcctca	1620
ctgcctgagg	gcacccagcc	tggaaacttg	ttcctgcagg	tgacagccac	agacgcggat	1680
agtggcccat	ttggcctcct	ctcctattcc	ttgggtgctg	gacttgggtc	ctccggatct	1740
ccccattcc	gcattgatgc	ccacagcggg	gatgtgtgca	caacccggac	cctggaccgt	1800
gaccaggggc	cctcaagctt	tgacttcaca	gtgacagctg	tggatggggg	aggcctcaag	1860
tccatggtat	atgtgaaggt	gtttctgtca	gacgagaatg	acaaccctcc	tcagttttat	1920
ccacgggagt	atgctgccag	tataagtgcc	cagagtccac	caggcacagc	tgtgctgagg	1980
ttgcgtgccc	atgacctga	ccagggatcc	catgggcgac	tctcctacca	tatcctggct	2040
ggcaacagcc	ccccactttt	taacctggat	gagcaatcag	ggctgttgac	agtagcctgg	2100
cccttggcca	gacggggccaa	ttctgtggtg	cagctggaga	tcggggctga	ggacggagggt	2160
ggcctacagg	cagaacccag	tgcccagagt	gacatcagca	ttgtgcctgg	aacccccaca	2220
ccacccatat	ttgagcaact	acagtatggt	ttttctgtgc	cagaggatgt	ggcaccaggc	2280
accagtgtgg	gcatagtcca	ggcacacaac	ccaccagggt	gggatccccg	aggactcttc	2340
tccttagatg	cggtatcagg	actgttgcaa	acacttcgcc	ctctggaccg	ggagctactg	2400
ggaccagtgt	tggagctgga	ggtgcgagca	ggcagtggag	tgcccccagc	tttcgctgta	2460
gctcgggtgc	gtgtgctgct	ggatgatgtg	aatgacaact	cccctgcctt	tcctgcacct	2520
gaagacacgg	tattgctacc	accaaact	gccccaggga	ctcccatcta	tacactgcgg	2580
gctcttgacc	ccgactcagg	tgtaaacagt	cgagtcacct	ttaccctgct	tgctgggggt	2640
ggtggagcct	tcaccgtgga	ccccaccaca	ggccatgtac	ggcttatgag	gcctctgggg	2700
ccctcaggag	ggccagccca	tgagctggag	ctggaggccc	gggatggggg	ctccccacca	2760
cgcaccagcc	actttcgact	acgggtggtg	gtacaggatg	tgggaacccg	tgggctggct	2820
ccccgattca	acagccctac	ctaccgtgtg	gacctgccct	caggcaccac	tgctggaact	2880
caggtcctgc	aagtgcaggc	ccaagcacca	gatggggggc	ctatcaccta	tcaccttgca	2940
gcagagggag	caagtagccc	ctttggcctg	gagccacaga	gtgggtggct	atgggtgcgg	3000
gcagcactag	accgtgaggc	ccaggaattg	tacatactga	aggtaatggc	agtgtctggg	3060
tccaaagctg	agttggggca	gcagacaggc	acagccaccg	tgagggtcag	catcctcaac	3120
cagaatgaac	acagtccccg	cttgtctgag	gatcccacct	tcctggctgt	ggctgagaac	3180
cagccccag	ggaccagcgt	gggccgagtc	tttgccactg	accgagactc	aggacccaat	3240
ggacgtctga	cctacagcct	gcaacagctg	tctgaagaca	gcaaggcctt	ccgcacccac	3300
ccccagactg	gagaagtgac	cacactccaa	accctggacc	gtgagcagca	gagcagctat	3360

cagctcctgg	tgcaggtgca	ggatggaggg	agcccccccc	gcagcaccac	aggcactgtg	3420
catgttgacg	tgcttgacct	caacgacaac	agccccacgt	tcctgcaggc	ttcaggagct	3480
gctggtgggg	gcctccctat	acaggtacca	gaccgcgtgc	ctccaggaac	actggtgacg	3540
actctgcagg	cgaaggatcc	agatgagggg	gagaatggga	ccatcttgta	cacgctaact	3600
ggctcctggc	cagagctttt	ctctctgcac	cctcactcag	gggagctgct	cactgcagct	3660
ccctgatcc	gagcagagcg	gccccactat	gtgctgacac	tgagtgtca	tgaccaaggc	3720
agccctcctc	gaagtgccag	cctccagctg	ctgggtgcagg	tgcttccctc	agctcgcttg	3780
gccgagccgc	ccccagatct	cgcagagcgg	gaccagcggg	caccagtgcc	tgtcgtgctg	3840
acggtgacag	cagctgaggg	actgcggccc	ggctctctgt	tgggctcggg	ggcagcgcca	3900
gagccccgcg	gtgtgggtgc	actcacctac	acactgggtg	gcggtgccga	tcccagagggc	3960
accttcgcgc	tggatgcggc	ctcagggcgc	ttgtacctgg	cgcggccccct	ggacttcgaa	4020
gctggccccg	cgtggcgcgc	gcttacggta	caagtgcagg	acgagaaatga	gcatgcgccc	4080
gcctttgcgc	gcgacccgct	gggcgcgctg	ccagagaacc	cggagccccg	cgcagcgctg	4140
tacactttcc	gcgcgtcgga	cgcgcagcgg	cccggccccca	atagcgacgt	gcgctaccgc	4200
ctgctgcgcc	aggagccgcc	cgtgcggggc	ttcgccctgga	cgcgcgcacc	ggggcgctcag	4260
ctccgcgcgg	cctggaccga	gagaccactc	ccgcgctgct	gctgctgggtg	gaagccaccg	4320
accggccccg	caacgccagc	cgcgcgtcgt	cagcgcgcgt	ttcagcgcat	atacgtcacg	4380
gatgcgaatg	agaacgcgcc	tgtcttcgcc	tcgccgtgca	cgcaggacca	gccgcctggg	4440
cccgcggctg	gcacgctcct	agccccgcgc	ccgcactctg	gcgaggctgc	acgcgtgtcc	4500
tatcggtctg	catctggcgg	ggacggccac	ttccggctgc	actcaagcac	tggagcgctg	4560
tcctggtgct	ggccgttgga	ccgcgaacaa	cgagctgagc	acgtactgac	agtgggtggc	4620
tcagaccgag	ctccccgccc	gcgctcggcc	acgcaggctc	tgaccgtcag	tgtcgctgac	4680
gtcaacgacg	aggcgcctac	tttccagcag	caggagtaca	gcgtcctctt	gcttgagaac	4740
aaccctcctg	gcacatctct	gctcacctct	cgagcaaccg	accccgacgt	ggggggccaac	4800
gggcaagtga	cttatggagg	cgtctctagc	gaaagctttt	ctctggatcc	tgacactggg	4860
gttctcacga	ctcttcgggg	cctggatcga	gaggaacagg	aggagatcaa	cctgacagtg	4920
tatgcccagg	acaggggctc	acctcctcag	ttaacgcctg	tcactgttcg	agtggctgtg	4980
gaggatgaga	atgaccatgc	accaaccttt	gggagtgcct	atctctctct	ggaggtgcct	5040
gagggccagg	acccccagac	ccttaccatg	cttcgggcct	ctgatccaga	tgtgggagcc	5100
aatgggcagt	tgcagtaccg	catcctagat	ggggacccat	caggagcctt	tgtcctagac	5160
cttgcttctg	gagagtttgg	caccatgcgg	ccactagaca	gagaagtgga	gccagctttc	5220
cagctgagga	tagaggcccc	ggatggaggg	cagccagctc	tcagtgccac	gctgcttttg	5280
acagtgcag	tgctggatgc	caatgaccat	gctccagcct	ttcctgtgct	tgcctactcg	5340
gtggagggtg	cggaggatgt	gcctgcaggg	acctgctgc	tgcagctaca	ggctcatgac	5400
cctgatgctg	gagctaattg	ccatgtgacc	tactacctgg	gcgcgggtac	agcaggagcc	5460
ttcctgctgg	agcccagctc	tggagaactg	cgcacagctg	cagccttgga	cagagaacag	5520
tgtcccagct	acaccttttc	tgtgagtgc	gtggatgggt	cagctgctgg	gcccctaagc	5580
accacagtgt	ctgtcaccat	cacgggtgcg	gatgtcaatg	accatgcacc	caccttcccc	5640
accagtcctc	tgcgcctacg	tctgccccgc	ccaggcccca	gcttcagtac	cccaaccttg	5700
gctctggcca	cactgagagc	tgaagatcgt	gatgctgggt	ccaatgcttc	cattctgtac	5760
cggctggcag	gcacaccacc	tcctggcact	actgtggact	cttacactgg	tgaaatccgc	5820
gtggccccgt	ctcctgtagc	tctaggcccc	cgagatcgtg	tcctcttcat	tgtggccact	5880
gatcttgggc	gtccagctcg	ctctgccact	ggtgtgatca	ttgtttggact	gcagggggaa	5940
gctgagcgtg	gacccccgct	tccccgggct	agcagtgagg	ctacgattcg	tgagaatgcg	6000
ccccagggga	ctcctattgt	ctccccaggg	gccgtccatg	caggaggcac	aaatggaccc	6060
atcacctaca	gcattctcag	tgggaatgag	aaagggacat	tctccatcca	gcctagtaca	6120
ggtgccatca	cagttcgctc	agcagagggg	ctagacttct	aggtgagtcc	acggctgcga	6180
ctggtgctgc	aggcagagag	tggaggagcc	tttgccctca	ctgtgctgac	cctgaccctg	6240
caagatgcc	acgacaatgc	tccccgtttc	ctgcggcccc	attatgtggc	cttccttcct	6300
gagtcgccgg	ccttggaggg	gccccctgct	cagggtggagg	cggatgacct	ggatcaaggc	6360
tctggaggac	agatttccta	cagtctggct	gcacccagc	cggcacgtgg	attgtttccac	6420
gtagacccaa	ccacaggcac	tatcactacc	acagccatcc	tggaccgtga	gatctgggct	6480
gaaacacggt	tgggtgctgat	ggccacagac	agagggagcc	cagccctggg	gggctcagct	6540
accttgacgg	tgatgggtcat	cgacaccaat	gacaatcgcc	ccaccatccc	ccaaccttgg	6600
gagctccgag	tgtcagaaga	tgcgttattg	ggctcagaga	ttgcacaggt	aacaggggaat	6660
gatgtggact	caggacccgt	gctgtgggtat	gtgctaagcc	catctggggc	ccaggatccc	6720
ttcagtgttg	gccgctatgg	aggccgtgtc	tccttcacgg	ggccccctgga	ctttgagcag	6780
tgtgaccgct	accagctgca	gctgctggca	catgatgggc	ctcatgaggg	ccgtgccaac	6840
ctcacagtgc	ttgtggagga	tgtcaatgac	aatgcacctg	ccttctcaca	gagcctctac	6900
caggtaatgc	tgcttgagca	cacacccccca	ggcagtgcga	ttctctccgt	ctctgccact	6960
gatcgggact	caggtgccaa	cggtcacatt	tcctaccacc	tggcttcccc	tgccgatggc	7020
ttcagtgttg	accccaacaa	tgggaccttg	ttcacaatag	tgggaacagt	ggccttgggc	7080
catgacgggt	caggagcagt	ggatgtgggtg	ctggaagcac	gagaccacgg	ggctccaggc	7140

WO 01/98354

PCT/US01/19942

cgggcagcac	gagccacagt	gcacgtgcag	ctgcaggacc	agaacgacca	cgccccgagc	7200
ttcacattgt	cacactaccg	tgtggctgtg	actgaagacc	tgccccctgg	ctccactctg	7260
ctcaccctgg	aggctacaga	tgctgatgga	agccgcagcc	atgccgctgt	ggactacagc	7320
accatcagtg	gcaactgggg	ccgagtcttc	cagctggaac	ccaggctggc	tgaggctggg	7380
gagagtgtctg	gaccaggccc	ccgggcaactg	ggctgcctgg	tgttgcttga	acctctagac	7440
tttgaaagcc	tgacacagta	caatctaaca	gtggctgcag	ctgaccgtgg	gcagccaccc	7500
caaagctcag	tcgtgccagt	cactgtcact	gtactagatg	tcaatgacaa	cccacctgtc	7560
tttacccgag	catcctaccg	tgtgacagta	cctgaggaca	cacctgttgg	agctgagctg	7620
ctgcatgtag	aggcctctga	cgctgaccct	ggccctcatg	gcctcgtgcg	tttactgtc	7680
agctcaggcg	acccatcagg	gctcttttgg	ctggatgaga	gctcaggcac	cttgcgactg	7740
gcccattgccc	tggactgtga	gaccacaggct	cgacatcagc	ttgtagtaca	ggctgctgac	7800
cctgctgggtg	cacactttgc	tttggcacca	gtgacaattg	aggtccagga	tgtgaatgat	7860
catggccccag	ccttcccact	gaacttactc	agcaccagcg	tggccgagaa	tcagcctcca	7920
ggcactctcg	tgaccactct	gcatgcaatc	gacggggatg	ctggggcttt	tgggaggctc	7980
cgttacagcc	tgttgagggc	tggggccagga	cctgaggggc	gtgaggcatt	tgcactgaac	8040
agctcaacag	gggagttgcg	tgcgcgagtg	cccttttgact	atgagcacac	agaaagcttc	8100
cggctgctgg	tgggtgctgc	tgatgctggg	aatctctcag	cctctgtcac	tgtgtcggtg	8160
ctagtgactg	gagaggatga	gtatgaccct	gtatttctgg	caccagcttt	ccacttccaa	8220
gtgcccgaag	gtgcccggcg	tggccacagc	ttgggtcacg	tgcaggccac	agatgaggat	8280
gggggtgccg	atggcctggg	tctgtattcc	cttgccacct	cttcccccta	ttttggtatt	8340
aaccagacta	caggagccct	gtacctgcgg	gtggacagtc	gggcaccagg	cagcggaaca	8400
gccacctctg	ggggtggggg	ccggaccctgg	cgggaagcac	cacgggagct	gaggctggag	8460
gtgatagcac	gggggcctct	gcctgggttc	cggagtgcc	cagtgcctgt	gaccgtggat	8520
atcaccacaca	ccgcactggg	cctggcacct	gacctcaacc	tgctattagt	agggggccgtg	8580
gcagcctcct	tgggagttgt	ggtgggtgctt	gcactggcag	ccctggctct	aggacttgtt	8640
cgggcccgtg	gccgcaaggc	tgaggcagcc	cctggcccaa	tgtcacaggc	agcaccctta	8700
gccagtgact	cactgcagaa	actgggcccgg	gagccacct	gtccaccacc	ctctgagcac	8760
ctctatcacc	agactcttcc	cagctatggg	gggccaggag	ctggaggacc	ctacccccgt	8820
ggtggctcct	tggacccttc	acattcaagt	ggccgaggat	cagcagaggc	tgcagaggat	8880
gatgagatcc	gcatgatcaa	tgagttcccc	cgtgtggcca	gtgtggcctc	ctctctggct	8940
gcccgtggcc	ctgactcagg	catccagcag	gatgcagatg	gtctgagtga	cacatcctgc	9000
gaaccacctg	cccctgacac	ctgggtataag	ggccgaaagg	cagggtctgt	gctgccaggt	9060
gcaggagcca	ctctctacag	agaggagggg	ccccagcca	ctgccacagc	cttcctgggg	9120
ggctgtggcc	tgagccctgc	accactggg	gactatggct	tcccagcaga	tggcaagcca	9180
tgtgtggcag	gtgcgctgac	agccattgtg	gccggcgagg	aggagctccg	tggcagctat	9240
aactgggact	acctgctgag	ctgggtgccct	cagttccaac	cactggccag	tgtcttcaca	9300
gagatcgctc	ggctcaagga	tgaagctcgg	ccatgtcccc	cagctccccg	tatcgacca	9360
ccaccctca	tactgcccgt	ggcccaccca	ggagccaagt	ctgtgcccc	caagccagca	9420
aacacagctg	cagcccgggc	catcttccca	ccagcttctc	accgctcccc	catcagccat	9480
gaaggctccc	tgtcctcagc	tgccatgtcc	cccagcttct	cacctctct	gtctcctctg	9540
gctgctcgct	caccggttgt	ctcaccattt	ggggtggccc	agggtccctc	agcctcagca	9600
ctcagcgag	agtctggcct	ggagccacct	gatgacacgg	agctgcacat	ctagctgtgg	9660
cccaggctgg	gccccgacct	gggatgcgca	cagtgtcccc	aacgcaggcc	ccactctgag	9720
cctgccctgg	gcagcctcgg	actatgactg	gctacgggga	ggccaccacc	aggccccagc	9780
tctccaccct	gaactcccca	gccccctcag	agtactagga	ccacagaagc	cctgttgctc	9840
actgacctgt	gaccagggtcc	aatgtgggga	gaaatatgaa	ggaggtagca	gccctgggtt	9900
ctcctcagtg	agggatccct	gccctgcacc	agcaccctga	gatggagctg	agactttatt	9960
tattgggggt	agggggatgg	aggaggtccc	tccaacatgt	ttggaccag	ctcctttggg	10020
ttccactgac	accctgccc	ctgccccctgc	ccagaaccaa	gtgccatttc	tcactctgga	10080
gccttaataa	actgcaattt	gtatccagaa	aaaaaaaaaa	aaa		10123

<210> 21

<211> 9321

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6977010CB1

<400> 21

ccggggggcg	gtctgggtca	gtgtggcagt	gggagccccg	gctcgtcgga	gggtgcagcg	60
cgggggtccc	ccgagccatc	cagacgcagg	ccccgcgggg	cgcacgggag	gccccggggg	120

actggcgccc	tggcccgggc	atgagggcgc	gcggggccgg	caggagccgg	aggaggagcc	180
gccgccgccg	ttgacccggc	cgccggccgg	gagctgggag	agatgcggag	cccggccacc	240
ggcgtccccc	tcccaacgcc	gccgccgccg	ccgctgctgc	tgctgttgct	gctgctgctg	300
ccgccgccac	tattgggaga	ccaagtgggg	ccctgtcggt	ccttgggggtc	caggggacga	360
ggctcttcgg	gggcctgcgc	ccccatgggc	tggctctgtc	catectcagc	gtcgaacctc	420
tggctctaca	ccagccgctg	cagggatgcg	ggcactgagc	tgactggcca	cctgggtacc	480
caccacgatg	gcctgagggg	ttgggtgtcca	gaatccgagg	cccataattcc	cctaccacca	540
gctcctgaag	gctgcccctg	gagctgtcgc	ctcctgggca	ttggaggcca	ccttttccca	600
cagggcaagc	tcacactgcc	cgaggagcac	ccgtgcttaa	aggctccacg	gctcagatgc	660
cagtcctgca	agctggcaca	ggcccccg	ctcagggcag	gggaaagggtc	accagaagag	720
tccctgggtg	ggcgtcggaa	aaggaatgta	aatacagccc	cccagttcca	gccccccagc	780
taccaggcca	cagtgcggga	gaaccagcca	gcaggcacc	ctgttgcatc	cctgagggcc	840
atcgacccgg	acgaggggtg	ggcaggtcga	ctggagtaca	ccatggatgc	cctctttgat	900
agccgctcca	accagttctt	ctccctggac	ccagtcactg	gtgcagtaac	cacagccgag	960
gagctggatc	gtgagaccaa	gagcaccac	gtcttcaggg	tcacggcgca	ggaccacggc	1020
atgccccgac	gaagtgcctt	ggctacactc	accatcttgg	ttactgacac	caatgaccat	1080
gaccctgtgt	tcgagcagca	ggagtacaag	gagagcctca	gggagaacct	ggagggttggc	1140
tatgaggtgc	tcactgtcag	ggccacggat	ggtgatgccc	ctcccaatgc	caatattctg	1200
taccgcctgc	tggaggggtc	tgggggcagc	ccctctgaag	tctttgagat	cgaccctcgc	1260
tctgggggtga	tccgaaccgg	tggccctgtg	gatcgggaag	agggtggaatc	ctaccagctg	1320
acggtagagg	caagtgacca	gggtcgggac	ccgggtcctc	ggagtaccac	agccgctgtt	1380
ttcctttctg	tggaggatga	caatgataat	gccccccagt	ttagtgagaa	gcgctatgtg	1440
gtccaggtga	gggaggatgt	gactccaggg	gccccagtac	tccgagtcac	agcctcggat	1500
cgagacaagg	ggagcaatgc	cgtgggtgcac	tatagcatca	tgagtggcaa	tgctcgggga	1560
cagttttatc	tggatgcccc	gactggagct	ctggatgtgg	tgagccctct	tgactatgag	1620
acgaccaagg	agtacaccct	acgggtgcga	gcacaggatg	gtggccgtcc	cccactctct	1680
aatgtctctg	gcttgggtgac	agtacaggtc	ctggatatca	acgacaatgc	ccccatcttc	1740
gtcagcacc	ctttccaggc	tactgtcctg	gagagcgtcc	ccttaggcta	cctgggtctc	1800
catgtccagg	ctatcgacgc	tgatgctggt	gacaatgccc	gcctggaata	ccgccttgct	1860
ggggtgggac	atgacttccc	cttcaccatc	aacaatggca	caggctggat	ctctgtggct	1920
gctgaactgg	accgggagga	agttgatttc	tacagccttg	gggtagaagc	tcgagaccat	1980
ggcactccag	cactcactgc	ctcggccagt	gtcagcgtga	ctgtcctgga	tgtcaacgac	2040
aacaatccaa	cctttaccca	accagagtac	acagtgcggc	tcaatgagga	tgcagctgtg	2100
ggcaccagcg	tgggtgacgg	gtcagctgtg	gaccgtgatg	ctcatagtgt	catcacctac	2160
cagatcacca	gtggcaatac	tcgaaaccgc	ttctccatca	ccagccaaag	tggtgggtggg	2220
ctggtatccc	ttgccctgcc	actggactac	aaacttgagc	ggcagtatgt	gttgggtgtt	2280
accgcctccg	atggcactcg	gcaggacacg	gcacagattg	tggtgaatgt	caccgacgcc	2340
aacacccatc	gtcctgtctt	tcagagctcc	cactatacag	tgaatgttaa	tgaggaccgg	2400
ccggcaggca	ccacgggtgg	gctgatcagc	gccacggatg	aggacacagg	tgagaatgcc	2460
cgcatacct	acttcatgga	ggacagcatc	ccccagttcc	gcatacatgc	agacacgggg	2520
gctgtcacca	cccaggctga	gctggactac	gaagaccaag	tgtcttacac	cctggccatt	2580
actgctcggg	acaatggcat	tccccagaag	tccgacacca	cctacctgga	gatcctgggtg	2640
aacgacgtga	atgacaatgc	ccctcagttc	ctgcgagact	cctaccaggg	cagtgtctat	2700
gaggatgtgc	cacccttcac	tagcgtcctg	cagatctcag	ccactgatcg	tgattctgga	2760
cttaatggca	gggtcttcta	caccttccaa	ggaggcgacg	atggagacgg	tgactttatt	2820
gttgagtcca	cgtcaggcat	cgtgcgaacg	ctacggaggc	tggatcgaga	gaacgtggcc	2880
cagtatgtct	tgcgggcata	tgcagtggac	aaggggatgc	ccccagcccc	cacacctatg	2940
gaagtgacag	tcactgtgtt	ggatgtgaat	gacaatcccc	ctgtctttga	gcaggatgag	3000
tttgatgtgt	ttgtggaaga	gaacagcccc	attgggctag	ccgtggcccc	ggtcacagcc	3060
actgaccccc	atgaaggcac	caatgcccag	attatgtacc	agattgtgga	gggcaacatc	3120
cctgaggtct	tccagctgga	catcttctcc	ggggagctga	cagccctggg	agacttagac	3180
tacgaggacc	ggcctgagta	cgtcctgggc	atccaggcca	cgtcagctcc	tctgggtgagc	3240
cgggctacag	tccacgtccg	cctccttgac	cgcaatgaca	accaccagt	gctgggcaac	3300
tttgagatcc	ttttcaacaa	ctatgtcacc	aatcgctcaa	gcagcttccc	tgggggtgcc	3360
attggccgag	tacctgcccc	tgaccctgat	atctcagata	gtctgactta	cagctttgag	3420
cggggaaatg	aactcagcct	ggtcctgtct	aatgcctcca	cgggtgagct	gaagctaagc	3480
cgcgcactgg	acaacaaccg	gcctctggag	gccatcatga	gcgtgctggg	gtcagacggc	3540
gtacacagcg	tgaccgcccc	gtgcgcgctg	cgtgtgacca	tcatacccca	tgagatgctc	3600
accacacagc	tcacgctgcg	cctggaggac	atgtcaccgg	agcgcttcc	gtcaccactg	3660
ctaggcctct	tcataccaggc	gggtggccggc	acgctggcca	cgccaccgga	ccacgtgggtg	3720
gtcttcaacg	tacagcggga	caccgacgcc	cccgggggcc	acatcctcaa	cgtgagcctg	3780
tcggtggggc	agccgccagg	gcccgggggc	gggcgcgccc	tcctgccttc	tgaggacctg	3840
caggagcgcc	tatacctcaa	ccgcagcctg	ctgacggcca	tctcggcaca	gcgcgtgctg	3900

cccttcgacg	acaacatctg	cctgcgggag	ccctgcgaga	actacatgcg	ctgcgtgtcg	3960
gtgctgcgct	tcgactcctc	cgcgccttc	atcgccctct	cctccgtgct	cttccggccc	4020
atccaccccg	tcggagggct	gcgctgccgc	tgcccgcg	gcttcacggg	tgactactgc	4080
gagaccgagg	tggacctctg	ctactcgcg	ccctgtggcc	cccacggg	ctgccgcagc	4140
cgcgagggcg	gctacacctg	cctctgtcgt	gatggctaca	cgggtgagca	ctgtgaggtg	4200
agtgtctgct	caggccgttg	caccccggt	gtctgcaaga	atgggggcac	ctgtgtcaac	4260
ctgctggtgg	gcggtttcaa	gtgcgattgc	ccatctggag	acttcgagaa	gccctactgc	4320
caggtgacca	cgcgcagctt	ccccgcccac	tccttcatca	cctttcgcgg	cctgcgccag	4380
cgtttccact	tcacctggc	cctctcgttt	gccacaaagg	agcgcgacgg	gttgctgttg	4440
tacaatgggc	gtttcaatga	gaagcatgac	tttgtggccc	tcgaggtgat	ccaggagcag	4500
gtccagctca	ccttctctgc	aggggagtca	accaccacgg	tgtcccccatt	cgtgcccgga	4560
ggagtccagt	atggccagt	gcatacgg	cagctgaaat	actacaataa	gccactgttg	4620
ggtcagacag	ggctcccaca	gggcccac	gagcagaagg	tggctgtggt	gaccgtggat	4680
ggctgtgaca	caggagtggc	cttgccgttc	ggatctgtcc	tgggcaacta	ctcctgtgct	4740
gcccagggca	cccaggggtg	cagcaagaag	tctctggatc	tgacggggcc	cctgctacta	4800
ggcgggggtg	ctgacctgcc	cgagagcttc	ccagtccgaa	tgcggcagtt	cgtgggctgc	4860
atgcggaacc	tgcaggtgga	cagccggcac	atagacatgg	ctgacttcat	tgccaacaat	4920
ggcacccgtg	ctggctgccc	tgccaagaag	aacgtgtgtg	acagcaacac	ttgccacaat	4980
gggggcactt	gcgtgaacca	gtgggacgcg	ttcagctgcg	agtgcacct	gggctttggg	5040
ggcaagagct	gcgcccagga	aatggccaat	ccacagcact	tcctgggcag	cagcctggtg	5100
gcctggcatg	gcctctcgct	gcccattctc	caaccctggt	acctcagcct	catgttccgc	5160
acgcgccagg	ccgacggtgt	cctgctgcag	gccatcacca	gggggcgcag	caccatcacc	5220
ctacagctac	gagagggcca	cgtgatgctg	agcgtggagg	gcacagggct	tcaggcctcc	5280
tctctccgtc	tggagccagg	ccggggccaat	gacggtgact	ggcaccatgc	acagctggca	5340
ctgggagcca	gcggggggcc	tggccatgcc	attctgtcct	tcgattatgg	gcagcagaga	5400
gcagagggca	acctggggcc	ccggctgcat	ggtctgcacc	tgagcaacat	aacagtgggc	5460
ggaataacct	ggccagccgg	cgggtgtggc	cgtggctttc	ggggctgttt	gcaggggtgtg	5520
cgggtgagcg	atacgccaga	gggggttaac	agcctggatc	ccagccatgg	ggagagcatc	5580
aacgtggagc	aaggctgtag	cctgcctgac	ccttgtgact	caaaccctg	tcctgctaac	5640
agctattgca	gcaacgactg	ggacagctat	tcctgcagct	gtgatccagg	ttactatggt	5700
gacaactgta	ctaattgtgt	tgacctgaac	ccgtgtgagc	accagtctgt	gtgtaccgc	5760
aagcccagtg	ccccccatgg	ctataacctg	gagtgtcccc	caaattacct	tggggcatac	5820
tgtgagacca	ggattgacca	gccttgtccc	cgtggctggt	ggggacatcc	cacatgtggc	5880
ccatgcaact	gtgatgtcag	caaaggcttt	gaccagact	gcaacaagac	aagcggcgag	5940
tgccactgca	aggagaacca	ctaccggccc	ccaggcagcc	ccacctgcct	cttgtgtgac	6000
tgctacccca	caggctcctt	gtccagagtc	tgtgaccctg	aggatggcca	gtgtccatgc	6060
aagccaggtg	tcatcggg	tcagtgtgac	cgctgtgaca	acccttttgc	tgaggtcacc	6120
accaatggct	gtgaaggggc	cttgtttgct	agttactgtc	cccggcccat	gaggtgctgg	6180
cctccagcag	aacctctcag	ccagtctcag	gggcttcctg	tgtgtctccc	tgaggccggc	6240
ccttttggct	tccttcccc	agggactgct	gtgcgccact	gtgatgagca	caggggggtg	6300
ctccccccaa	acctcttcaa	ctgcacgtcc	atcaccttct	cagaactgaa	gggcttcgct	6360
gagcggctac	agcgggaatga	gtcaggccta	gactcagggc	gctcccagca	gctagccctg	6420
ctcctgcgca	acgccacgca	gcacacagct	ggctacttcg	gcagcgacgt	caagggtggc	6480
taccagctgg	ccacgcggct	gctggcccac	gagagcaccc	agcggggctt	tgggctgtct	6540
gccacacagg	acgtgcactt	cactgagaat	ctgctgcggg	tgggcagcgc	cctcctggac	6600
acagccaaca	agcggcactg	ggagctgac	cagcagacag	aggggtggc	cgcctggctg	6660
ctccagcact	atgaggccta	cgccagtgcc	ctggcccaga	acatgcggca	cacctaccta	6720
agccccttca	ccatcgtcac	gcccacatt	gtcatctccg	tagtgcgctt	ggacaaagg	6780
aactttgctg	gggccaagct	gccccgctac	gaggccctgc	gtggggagca	gccccgggac	6840
cttgagacaa	cagtcatctc	gcctgagtct	gtcttcagag	agacgcccc	cgtggtcagg	6900
cccgcaggcc	ccggagaggc	ccaggagcca	gaggagctgg	cacggcgaca	gcgacggcac	6960
ccggagctga	gccaggggtga	ggctgtggcc	agcgtcatca	tctaccgcac	cctggccggg	7020
ctactgcctc	ataactatga	ccctgacaag	cgcagcttga	gagtccccaa	acgcccgatc	7080
atcaacacac	ccgtgggtgag	catcagcgtc	catgatgatg	aggagcttct	gccccggggc	7140
ctggacaaac	ccgtcacgg	gcagttccgc	ctgctggaga	cagaggagcg	gaccaagccc	7200
atctgtgtct	tctggaacca	ttcaatcctg	gtcagtggca	caggtggctg	gtcggccaga	7260
ggctgtgaag	tcgtcttccg	caatgagagc	cacgtcagct	gccagtgcaa	ccacatgacg	7320
agcttcgctg	tgctcatgga	cgtttctcgg	cgggaggtcg	ggcccacagg	ggcagctgca	7380
gagccgtgga	atggggagat	cctgccactg	aagacactga	catacgtggc	tctaggtgtc	7440
accttggtg	cccttctgct	caccttcttc	ttcctcactc	tcttgcgat	cctgcgctcc	7500
aaccaacacg	gcatccgacg	taacctgaca	gctgccctgg	gcctggctca	gctggctctc	7560
ctcctgggaa	tcaaccaggc	tgacctccct	tttgccctgca	cagtcatctg	catcctgctg	7620
cacttcctgt	acctctgcac	cttttccctgg	gctctgctgg	aggccttgca	cctgtaccgg	7680

```

gcactcactg aggtgcgcgga tgtcaacacc ggcccatgc gcttctacta catgctgggc 7740
tggggcgtgc ctgccttcat cacagggcta gccgtgggccc tggaccccgga gggctacggg 7800
aaccctgact tctgctggct ctccatctat gacacgctca tctggagttt tgctggcccg 7860
gtggcctttg ccgtctcgat gagtgtcttc ctgtacatcc tggcgggccc ggcctcctgt 7920
gctgcccagc ggcaggggctt tgagaagaaa ggtcctgtct cgggcctgca gccctccttc 7980
gccgtcctcc tgctgctgag cgccacgtgg ctgctggcac tgctctctgt caacagcgac 8040
accctcctct tccactacct ctttgctacc tgcaattgca tccagggccc cttcatcttc 8100
ctctcctatg tgggtgcttag caaggaggct cggaaagcac tcaagcttgc ctgcagccgc 8160
aagcccagcc ctgaccctgc tctgaccacc aagtccaccc tgacctcgtc ctacaactgc 8220
cccagcccct acgcagatgg gcggctgtac cagccctacg gagactcggc cggctctctg 8280
cacagcacca gtcgctcggg caagagtcag ccagctaca tccccttctt gctgagggag 8340
gagtcgcgac tgaaccctgg ccaagggccc cctggcctgg gggatccagg cagcctgttc 8400
ctggaaggct aagaccagca gcatgacct gacacggact ccgacagtga cctgtcctta 8460
gaagacgacc agagtggctc ctatgcctct acccactcat cagacagtga ggaggaagaa 8520
gaggaggagg aagaggaggc cgccttcctt ggagagcagg gctgggatat cctgctgggg 8580
cctggagcag agagactgcc cctgcacagt actcccaagg atggggggccc agggcctggc 8640
aaggccccct ggccaggaga ctttggggacc acagcaaaag agagtagtgg caacggggcc 8700
cctgaggagc ggctgcggga gaatggagat gccctgtctc gagaggggtc cctaggcccc 8760
cttccaggct cttctgcccc gcctcacaaa ggcatectta agaagaagtg tctgcccacc 8820
atcagcgaga agagcagcct cctgcggctc cccctggagc aatgcacagg gtcttcccgg 8880
ggctcctccg ctagtgaggg cagccggggc ggccccctc cccgccacc gccccggcag 8940
agcctccagg agcagctgaa cgggggtcatg cccatcgcca tgagcatcaa ggcaggcacg 9000
gtggatgagg actcgtcagg ctccgaagga taggacctcc caggatgctt cccagcctct 9060
cctcagtttc ccatctgctg tgctctggg aggagaggga ctctggggg gcctgcccct 9120
catacgccat caccaaaagg aaaggacaaa gccacacgca gccagggctt cacacccttc 9180
aggctgcacc cgggcaggcc tcagaacggg gaggggcccag ggcaaagggt gtgtctcgtc 9240
ctgcccgcac tgcctctccc aggaactgga aaagccctgt ccggtgaggg ggcagaagga 9300
ctcagcgcca ctggaccccc a 9321

```

<210> 22

<211> 3900

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 926992CB1

<400> 22

```

ggacaaggct ctacagcctc agccagggca ctcagctgtt gcaggggtgtg atggagaaca 60
aagctatgta cctacacacc gtcagcgact gtgacaccag ctccatctgt gaggattcct 120
ttgatggcag gagcctgtcc aagctgaacc tgtgtgagga tgggtccatgt cacaacggc 180
gggcaagcat ctgctgggtct tcctgattct tgtgggcatc ttcattcttag cagtgtccag 240
gccgcgcagc tcccctgacg acctgaaggc cctgactcgc aatgtgaacc ggctgaatga 300
gagcttccgg gacttgcagc tgccggctgct gcaggctccg ctgcaagcgg acctgacgga 360
gcaggtgtgg aaggtgcagg acgcgctgca gaaccagtca gactcgttgc tggcgctggc 420
gggcgcagtg cagcggctgg agggcgcgct atgggggctg caggcgcagg cgggtgcagac 480
cgagcaggcg gtggccctgc tgccgggaccg caggggcccag cagagcgaca cggcgcagct 540
ggagctctac cagctgcagg tggagagcaa cagtagccag ctgctgctga ggcgccacgc 600
gggcctgctg gacgggctgg cgcgcagggt gggcatcctg ggcgaggagc tggccgacgt 660
gggcggcgctg ctgcgcggcc tcaaccacag cctgtcctac gacgtggccc tccaccgcac 720
gcggctgcag gacctgcggg tgctgggtgag caacgccagc gaggacacgc gccgcctgcg 780
cctggcgcaç gtaggcatgg agctgcagct gaagcaggag ctggccatgc tcaacgcggc 840
caccgaggac ctgcgccctca aggactggga gcactccatc gcactgcgga acatctccct 900
cgcgaaaggg ccaccgggac ccaaagggtga tcaggggggat gaaggaaagg aaggcaggcc 960
tggcatccct ggattgcctg gacttcgagg tctgcccggg gagagaggta ccccaggatt 1020
gcccggggccc aagggcgatg atgggaagct gggggccaca ggaccaatgg gcatgcgtgg 1080
gttcaaaggc gaccgaggcc caaaaggaga gaaaggagag aaaggagaca gagctgggga 1140
tgccagtggc gtggaggccc cgatgatgat ccgcctgggtg aatggctcag gtccgcacga 1200
gggcccgcgtg gaagtgtacc acgaccggcg ctggggcacc gtgtgtgacg acggctggga 1260
caagaaggac ggagacgtgg tgtgccgcat gctcggcttc cgcggtgtgg aggaggtgta 1320
ccgcacagct cgattcgggc aaggcactgg gaggatctgg atggatgacg ttgcctgcaa 1380
gggcacagag gaaaccattt tccgctgcag cttctccaaa tgggggggtga caaactgtgg 1440

```

```

acatgccgaa gatgccagcg tgacatgcaa cagacactga aagtgggagc agcccaagtt 1500
cggggtcctg cacagagcac ccttcctgca tccctggggt ggggcacagc tcggggccac 1560
cctgaccatg cctcgaccac accccgtcca gcattctcag tcctcacacc tgcattccag 1620
gaccgtgggg gccgggtcatc atttccctct tgaacatgtg ctccgaagta taactctggg 1680
acctactgcc cgtctctctc ttccaccagg ttctctgcatg aggagccctg atcaactgga 1740
tcaccacttt gccagcctc tgaacaccat gcaccaggcc tcaatatccc agttcccttt 1800
ggccttttag ttacaggtga atgctgagaa tgtgtcagag acaagtgcag cagcagcgat 1860
ggttggtagt atagatcatt tactcttcag acaattccca aacctccatt agtccaagag 1920
tttctacatc ttcttcccca gcaagaggca acgtcaagtg atgaatttcc cccctttact 1980
ctgcctctgc tccccatttg ctagtttgag gaagtgcacat agaggagaag ccagctgtag 2040
gggcaagagg gaaatgcaag tcacctgcag gaatccagct agatttgagg aagggaatga 2100
aactaacatt gaatgactac catggcacgc taaatagtat cttgggtgcc aaattcatgt 2160
atccacttag ctgcattggt ccaggggcatg tcagtctgga tacagcctta cctccaggta 2220
gcacttaact ggtccattca cctagactgc aagtaagaag acaaaatgac tgagaccgtg 2280
tgcccacctg aacttattgt ctttacttgg cctgagctaa aagcttgggt gcaggacctg 2340
tgtaactaga aagttgccta cttcagaacc tccaggggcg gagtgcaagg tcaaacaatga 2400
ctggcttcca ggccgaccat caatgtagga ggagagctga tgtggagggt gacatggggg 2460
ctgcccattg taaacctgag tccagtgtc tggcattggg cagtcacggt taaagccaag 2520
tcatgtgtgt ctcagctgtt tggagggtgat gattttgcat cttccaagcc tcttcagggtg 2580
tgaatctgtg gtcaggaaaa cacaagtcct aatggaaccc ttagggggga aggaaatgaa 2640
gattccctat aacctctggg ggtggggagt aggaataagg ggccttgggc ctccataaat 2700
ctgcaatctg caccctctc ctagagacag ggagatcgtg ttctgctttt tacatgagga 2760
gcagaactgg gccatacacg tgttcaagaa ctagggggagc tacctggtag caagtgagtg 2820
cagaccaccc tcaccttggg ggaatctcaa actcataggc ctcagataca cgatcacctg 2880
tcatatcagg tgagcactgg cctgcttggg gagagacctg ggcccctcca ggtgtaggaa 2940
cagcaacact cctggctgac aactaagcca atatggccct aggtcattct tgcttccaat 3000
atgcttgcca ctctttaa atgtcctaatga tgagaaactc tctttctgac caattgctat 3060
gtttacataa cacgcatgta ctcatgcac ccttgccaga gccatatata gtatgcata 3120
ataaacatag cactttttac tacatagctc agcacattgc aaggtttgca tttaagttaa 3180
aaaaaaaaaa aaaaaaaact aaagggtgaaa gatgccacat tgaacaaact aaattcccaa 3240
cccggttctg gcaaagaatc cagttatccc ttccatgaag acgcacataa ctctcttact 3300
tggtctttcc attagggaca acataagtct tgttttacat caaataaaaa caatgtttaa 3360
aagtgtgtga accttaaaaa tggaagtcta ctagtttaca tacctacttc agaggacatg 3420
gaaatgacca tgggcctgca tttcaggggac caagcaa at taggcctggc ctaaaataca 3480
tcagaccctt tgtaagagag aatttcaata aagcaaaaaa catgtcaaaa aaaaaaata 3540
agctcaaata aaaaaagggt gagaaatggg attatagtga ggggtgtgtt gggagaagta 3600
tggggcgtag tgggtgtgga gtgtaagtgg gaggttggtg agttggggta gggagagtaa 3660
acagaaagag gggcgggcgc ttctaggggg gtttcgcggt ttgtgggtcg cgggtgtgtt 3720
ggggcattcc tgtgagcgcc cctgttgggg gggggtcagc cccaggttt ttgctgtgtc 3780
cgagtggggc ggcgctgttt ttatagaaca ggggtggtga aattgggggg aaagactccc 3840
ccctgtggtg tttttccccc cgtttgttgt gtggcgctgc ttttgggggg ttgaatgctc 3900

```

<210> 23

<211> 2076

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1002055CB1

<400> 23

```

ggatcctgta ctgagaagtt gaccagagag ggtctcacca tgccgacagt tccttctgta 60
cctgtgtgga ggaaaagtac tgagtgaagg gcagaaaaag agaaaacaga aatgctctgc 120
ccttgaggaa ctgctaacct agggctactg ttgattttga ctatcttctt agtggccgct 180
tcaagcagtt tatgtatgga tgaaaaacag attacacaga actactcgaa agtactcgca 240
gaagttaaca cttcatggcc tgtaaagatg gctacaaatg ctgtgctttg ttgccctcct 300
atcgcattaa gaaatttgat cataataaca tgggaaataa tcctgagagg ccagccttcc 360
tgcacaaaag cctacaggaa agaaacaaat gagaccaagg aaaccaactg tactgatgag 420
agaataacct ggggtctccag acctgatcag aattcggacc ttcagattcg tccagtggcc 480
atcactcatg acgggtatta cagatgcata atggtaacac ctgatgggaa tttccatcgt 540
ggatatcacc tccaagtgtt agttacacct gaagtgacct tgtttcaaaa caggaaataga 600
actgcagtat gcaaggcagt tgcagggaag ccagctgcgc agatctcctg gatcccagag 660

```

```

ggcgattgtg ccactaagca agaatactgg agcaatggca cagtgactgt taagagtaca 720
tgccactggg aggtccacaa tgtgtctacc gtgacctgcc acgtctccca tttgactggc 780
aacaagagtc tgtacataga gctacttcct gttccagggtg ccaaaaaatc agcaaaatta 840
tatattccat atatcatcct tactattatt attttgacca tcgtgggatt catttggttg 900
ttgaaagtca atggctgcag aaaatataaa ttgaataaaa cagaatctac tccagttggt 960
gaggaggatg aaatgcagcc ctatgccagc tacacagaga agaacaatcc tctctatgat 1020
actacaaaca aggtgaaggc atctcaggca ttacaaagtg aagttgacac agacctccat 1080
actttataag ttgttggact ctagtaccac gaaacaacaa caaacgagat acattataat 1140
tactgtctga ttttcttaca gttctagaat gaagacttat attgaaatta ggttttccaa 1200
ggttcttaga agacatttta atggattctc attcataccc ttgtataatt ggaatttttg 1260
attcttagct gctaccagct agttctctga agaactgatg ttattacaaa gaaaatacat 1320
gcccatgacc aaatattcaa attgtgcagg acagtaaata atgaaaacca aatttcctca 1380
agaaataact gaagaaggag caagtgtgaa cagtttcttg tgtatccttt cagaatattt 1440
taatgtacat atgacatgtg tatatgccta tggatatatg gtcaatttat gtgtcccctt 1500
acataacat gcacatatct ttgtcaaggc accagtggga acaatacact gcattactgt 1560
tctatacata tgaaaaccta ataataaag tcttagagat cattttatat catgacaagt 1620
agagctacct cattcttttt aatgggtata taaaattcca ttgtatagtt atatcattat 1680
ttaattaaaa acaaccctaa tgatggatat ttagattctt ttaagttttg tttatttctt 1740
ttaagttttg tttgtggtat aaacaatacc acatagaatg tttcttgtgc atatatctct 1800
ttgtttttga gtatatctgt aggataactt tcttgagtgg aattgtcagg tcaaaggggt 1860
tgtgcatttt actattgata tatatgttaa attgtgtcaa atatatatgt caaattccct 1920
ccaacattgt ttaaattgtgc ctttccctaa atttctattt taataactgt actattcctg 1980
cttctacagt tgccactttc tctttttaat caaccagatt aaatatgatg tgagattata 2040
ataagaatta tactatttaa taaaaatgga tttata 2076

```

<210> 24

<211> 3991

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3998749CB1

<400> 24

```

ggtgaaccac ctaagggttga gaatggcttt ctggagcata caactggcag gatctttgag 60
agtgaagtga ggtatcagtg taacccgggc tataagtcag tcggaagtc tgtatttgtc 120
tgccaaggcc aatcgccact ggcacagtga atcccctctg atgtgtgttc ctctcgactg 180
tggaaaacct ccccgatcc agaattggctt catgaaagga gaaaactttg aagtaggggtc 240
caagggttcag tttttctgta atgagggtta tgagcttggt ggtgacagtt cttggacatg 300
tcagaaatct ggcaaatgga ataagaagtc aaatccaaag tgcattgctg ccaagtgcc 360
agagccgccc ctcttggaag accagctagt attaaaggag ttgaccaccg aggtaggagt 420
tgtgacattt tcctgtaaag aagggcattg cctgcaaggc ccctctgtcc tgaaatgctt 480
gccatcccag caatggaatg actctttccc tgtttgtaag attgttcttt gtaccccacc 540
tcccctaatt tcctttggtg tcccattcc ttctctgtct cttcattttg gaagtactgt 600
caagtattct tgtgtaggtg ggtttttcct aagaggaaat tctaccacc tctgccaacc 660
tgatggcacc tggagctctc cactgccaga atgtgttcca gtagaatgtc cccaacctga 720
ggaaatcccc aatggaatca ttgatgtgca aggccttgcc tatctcagca cagctctcta 780
tacctgcaag ccaggctttg aattgggtgg aaatactacc accctttgtg gagaaaatgg 840
tcaactggct ggaggaaaac caacatgtaa agccattgag tgctgaaac ccaaggagat 900
tttgaatggc aaattctctt acacggacct acactatgga cagaccgtta cctactcttg 960
caaccgaggc tttcggctcg aagggtccag tgcttgacc tgtttagaga cagggtgattg 1020
ggatgtagat gcccattctt gcaatgccat ccactgtgat tccccacaac ccattgaaaa 1080
tgggttttga gaagggtgcag attacagcta tggtgccata atcatctaca gttgcttccc 1140
tgggttttcag gtggctggtc atgccatgca gacctgtgaa gagtcaggat ggtcaagttc 1200
catcccaaca tgtatgccaa tagactgtgg cctccctcct catatagatt ttggagactg 1260
tactaaactc aaagatgacc agggatattt tgagcaagaa gacgacatga tggaagtctc 1320
atacgtgact cctcaccctc cttatcattt gggagcagtg gctaaaacct gggaaaatac 1380
aaaggagtct cctgctacac attcatcaaa ctttctgtat ggtaccatgg tttcatacac 1440
ctgtaatcca ggatatgaac ttctggggaa ccctgtgctg atctgccagg aagatggaac 1500
ttggaatggc agtgcaccat cctgcatttc aattgaatgt gacttgctta ctgctcctga 1560
aaatggcttt ttgcgtttta cagagactag catgggaagt gctgtgcagt atagctgtaa 1620
acctggacac attctagcag gctctgactt aaggctttgt ctagagaata gaaagtggag 1680

```

tggtgcctcc	ccacgctgtg	aagccatttc	atgcaaaaag	ccaaatccag	tcatgaatgg	1740
atccatcaaa	ggaagcaact	acacatacct	gagcacgttg	tactatgagt	gtgaccccgg	1800
atatgtgctg	aatggcactg	agaggagaac	atgccaggat	gacaaaaact	gggatgagga	1860
tgagcccatt	tgcattcctg	tggactgcag	ttcaccccca	gtctcagcca	atggccaggt	1920
gagaggagac	gagtacacat	tccaaaaaga	gattgaatac	acttgcaatg	aagggtttct	1980
gcttgaggga	gccaggagtc	gggtttgtct	tgccaatgga	agttggagtg	gagccactcc	2040
cgactgtgtg	cctgtcagat	gtgccacccc	gccacaactg	gccaatgggg	tgacggaagg	2100
cctggactat	ggcttcatga	aggaagtaac	attccactgt	cacgagggct	acatcttgca	2160
cggtgctcca	aaactcacct	gtcagtcaga	tggcaactgg	gatgcagaga	ttcctctctg	2220
taaaccagtc	aactgtggac	ctcctgaaga	tcttgcccat	ggtttcccta	atggtttttc	2280
ctttattcat	gggggccata	tacagtatca	gtgctttcct	ggttataagc	tccatggaaa	2340
ttcatcaaga	aggtgcctct	ccaatggctc	ctggagtggc	agctcacctt	cctgcctgcc	2400
ttgcagatgt	tccacaccag	taattgaata	tggaaactgtc	aatgggacag	attttgactg	2460
tggaaaggca	gcccggattc	agtgttcaa	aggcttcaag	ctcctaggac	tttctgaaat	2520
cacctgtgaa	gccgatggcc	agtggagctc	tgggttcccc	cactgtgaac	acacttcttg	2580
tggttctctt	ccaatgatac	caaatgcgtt	catcagttag	accagctctt	ggaaggaaaa	2640
tgtgataact	tacagctgca	ggtctggata	tgtcatacaa	ggcagttcag	atctgatttg	2700
tacagagaaa	ggggtatgga	gccagcctta	tccagtctgt	gagcccttgt	cctgtgggtc	2760
cccaccgtct	gtcgccaatg	cagtggcaac	tggagaggca	cccacctatg	aaagtgaagt	2820
gaaactcaga	tgtctggaag	gttatacgat	ggatacagat	acagatacat	tcacctgtca	2880
gaaagatggg	cgctggttcc	ctgagagaat	ctcctgcagt	cctaaaaaat	gtcctctccc	2940
ggaaaacata	acacatatac	ttgttcatgg	ggacgatttc	agtgtgaata	ggcaagtttc	3000
tgtgtcatgt	gcagaagggt	atacctttga	gggagttaac	atatcagtat	gtcagcttga	3060
tggaaacctg	gagccaccat	tctccgatga	atcttgcagt	ccagtttctt	gtgggaaacc	3120
tgaaagtcca	gaacatggat	ttgtgggttg	cagtaaatac	acctttgaaa	gcacaattat	3180
ttatcagtgt	gagcctggct	atgaactaga	ggggaacagg	gaacgtgtct	gccaggagaa	3240
cagacagtgg	agtggagggg	tggcaatatg	caaagagacc	aggtgtgaaa	ctccacttga	3300
atthctcaat	gggaaagctg	acattgaaaa	caggacgact	ggacccaacg	tggtatatth	3360
ctgcaacaga	ggctacagtc	ttgaaggggc	atctgaggca	cactgcacag	aaaatggaac	3420
ctggagccac	ccagtccttc	tctgcaaacc	aaatccatgc	cctgttcctt	ttgtgattcc	3480
cgagaatgct	ctgctgtctg	aaaaggagtt	ttatgttgat	cagaatgtgt	ccatcaaattg	3540
tagggaagg	tttctgctgc	agggccacgg	catcattacc	tgcaaccccg	acgagacgtg	3600
gacacagaca	agcgccaaat	gtgaaaaaat	ctcatgtggg	ccaccagctc	acgtagaaaa	3660
tgcaattgct	cgaggcgtag	attatcaata	tggagacatg	atcacctact	catgttacag	3720
tggatacatg	ttggaggggt	tcctgaggag	tgtttggtta	gaaaatggaa	catggacatc	3780
acctcctatt	tgcagagctg	tctgtcgatt	tccatgtcag	aatggggggc	tctgccaacg	3840
cccaaattgct	tgttcctgtc	agagggctgg	atggggcgcc	tctgtgaaga	accaatctgc	3900
attcttccct	gtctgaaagg	aggtecgctg	gtggccccct	accagtgtga	ctgcccgcct	3960
ggctggacgg	ggtctcgctg	tcatacagct	g			3991

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of

5 SEQ ID NO:1-12,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12,

c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and

10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-12.

15

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

20

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:13-24.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

25

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

30

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

5 11. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90REPTR identical to a polynucleotide sequence selected from the group consisting of SEQ ID
10 NO:13-24,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

15 12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

20 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if
25 present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide
30 having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

5 17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

18. A method for treating a disease or condition associated with decreased expression of functional REPTR, comprising administering to a patient in need of such treatment the composition
10 of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 15 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

20 21. A method for treating a disease or condition associated with decreased expression of functional REPTR, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide
25 of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22
30 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional REPTR, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- 5 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 10 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound
- 15 with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target

20 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- 25 c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- 30 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- 35 c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

5

29. A diagnostic test for a condition or disease associated with the expression of REPTR in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

10 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,

15 b) a single chain antibody,

c) a Fab fragment,

d) a F(ab')₂ fragment, or

e) a humanized antibody.

20 31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of REPTR in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

25

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of REPTR in a subject, comprising administering to said subject an effective amount of the composition of claim

30 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from

the group consisting of SEQ ID NO:1-12, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the

group consisting of SEQ ID NO:1-12 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

5 57. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:13.

58. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:14.

10 59. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:15.

15 60. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:16.

61. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:17.

20 62. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:18.

63. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:19.

25 64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:20.

30 65. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:21.

66. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:22.

67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:23.

68. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
5 NO:24.