



(86) Date de dépôt PCT/PCT Filing Date: 2002/01/28

(87) Date publication PCT/PCT Publication Date: 2002/08/15

(85) Entrée phase nationale/National Entry: 2003/07/15

(86) N° demande PCT/PCT Application No.: US 2002/002616

(87) N° publication PCT/PCT Publication No.: 2002/062841

(30) Priorités/Priorities: 2001/02/02 (60/266,195) US;
2001/02/08 (60/267,924) US; 2001/02/09 (60/268,112) US;
2001/02/09 (60/267,816) US; 2001/02/26 (60/271,639) US;
2001/09/07 (60/317,818) US; 2001/12/21 (60/343,553) US

(51) Cl.Int.⁷/Int.Cl.⁷ C07K 14/47, A01K 67/027, A61K 38/17,
C12N 15/12, C12N 1/21, C07K 16/18, C12N 5/10,
C12Q 1/68, C12N 15/00, G01N 33/68

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

(72) Inventeurs/Inventors:
TANG, TOM Y., US;
YUE, HENRY, US;
GANDHI, AMEENA R., US;
YAO, MONIQUE G., US;
WARREN, BRIDGET A., US;

...

(74) Agent: SMART & BIGGAR

(54) Titre : PROTEINES SECRETEES

(54) Title: SECRETED PROTEINS

(57) **Abrégé/Abstract:**

The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. 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 SECP.

(72) **Inventeurs(suite)/Inventors(continued)**: DING, LI, US; DUGGAN, BRENDAN M., US; XU, YUMING, US; YANG, JUNMING, US; THANGAVELU, KAVITHA, US; LAL, PREETI G., US; HONCHELL, CYNTHIA D., US; CHAWLA, NARINDER K., US; LEE, SALLY, US; LEE, ERNESTINE A., US; RICHARDSON, THOMAS W., US; BAUGHN, MARIAH R., US; ELLIOTT, VICKI S., US

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 August 2002 (15.08.2002)

PCT

(10) International Publication Number
WO 02/062841 A3

(51) International Patent Classification⁷: **C07K 14/47**,
C12N 15/12, A01K 67/027, C12N 15/00, 1/21, 5/10,
C07K 16/18, C12Q 1/68, A61K 38/17, G01N 33/68

(21) International Application Number: PCT/US02/02616

(22) International Filing Date: 28 January 2002 (28.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/266,195	2 February 2001 (02.02.2001)	US
60/267,924	8 February 2001 (08.02.2001)	US
60/268,112	9 February 2001 (09.02.2001)	US
60/267,816	9 February 2001 (09.02.2001)	US
60/271,639	26 February 2001 (26.02.2001)	US
60/317,818	7 September 2001 (07.09.2001)	US
60/343,553	21 December 2001 (21.12.2001)	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): **TANG, Tom, Y.** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **GANDHI, Ameena, R.** [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). **DING, Li** [CN/US]; 3353 Alma Street, #146, Palo Alto, CA 94306 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **YANG, Junming** [US/US]; 7125 Bark Lane, San Jose, CA 95129 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **LAL, Preeti, G.** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **HONCHELL, Cynthia, D.** [US/US];

400 Laurel Street #203, San Carlos, CA 94070 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **LEE, Sally** [US/US]; 825 East Evelyn, #425, Sunnyvale, CA 94086 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **RICHARDSON, THOMAS, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (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, EC, 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, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, 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, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:

5 June 2003

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: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. 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 SECP.



WO 02/062841 A3

SECRETED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of secreted proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues. Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, vasomediators, cell surface markers, and antigen recognition molecules. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various

differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A.N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection, maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) *J. Biol. Chem.* 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) *J. Biol. Chem.* 268:5879-5885). Hemomucin is a novel Drosophila surface mucin that may be involved in the induction of antibacterial effector molecules (Theopold, U. et al. (1996) *J. Biol. Chem.* 271:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine, C.T. et al. (1998) *Connect Tissue Res.* 38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma

tumor, a non-mineralized odontogenic tumor (Deutsch, D. et al. (1998) Connect. Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al. (2000) Genet. Res. 76:41-50). Research by Yokoyama et al. (1996) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-related ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue (Yokoyama, M. et al. (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al. (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K), a secreted glycoprotein isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is preceded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths, K. et al. (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich et al. (1994) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich, A. et al. (1994) J. Biol. Chem. 269:18401-18407).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor, has been shown to promote neurite outgrowth in vitro. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of

particular sensory systems (Ohta, K. et al. (1995) *Cell* 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) *Cell* 95:903-916).

Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek, C.A. and Chou, J.Y. (1991) *Endocrinology* 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic and chorionic membranes (Plouzek, C.A. et al. (1993) *Placenta* 14:277-285).

The interaction between a sperm cell and an ovum, which ultimately results in the fusion of two gametes into a single zygote, also involves a number of secreted proteins. The process of fusion can be divided into four steps: (i) sperm with intact acrosomes (apical secretory vesicles, see below) interact with the extracellular matrix (zona pellucida, ZP) of the ovum (Eddy, E. and O'brian, D. (1994) in *The Physiology of Reproduction*, Ed. Knobil, E. and Neill, J. New York, NY Raven Press pp. 29-78), (ii) sperm undergo the acrosome reaction and penetrate the ZP, (iii) sperm bind to the plasma membrane of the ovum, and (iv) sperm fuse with the ovum to form the zygote. Much of the seminal work involved in the dissection of these processes has been performed using mice as a model system for mammalian fertilization. The ZP in mice is composed of three N-glycosylated and O-glycosylated secreted proteins, ZP1 (a 200 kDa dimeric protein), ZP2 (120 kDa), and ZP3 (83 kDa) (Wassarman, P. (1988) *Annu. Rev. Biochem.* 57:415-442). ZP2 and ZP3 polypeptides form filaments consisting of structural repeats. ZP2/ZP3 filaments are crosslinked by ZP1 to form the ZP matrix (Wassarman, P. (1999) *Cell* 96:175-183).

Mice defective for N-linked glycosylation are fertile, suggesting that N-linked glycosylation is not essential for ZP function (Rosiere, T. and Wassarman, P. (1992) *Dev. Biol.* 154:309-317; reviewed in Snell, W. and White, J. (1996) *Cell* 85:629-637; and Wassarman, P. *supra*). However, the targeted disruption of both ZP3 alleles in female mice results in decreased ZP thickness and infertility (Liu, C. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:5431-5436; Rankin, T. et al. (1996) *Development* 122:2903-2910). ZP1 and ZP2 have been less extensively studied than ZP3. Based on results obtained from (i) limited-proteolysis experiments (Rosiere, T. and Wassarman, P. *supra*), (ii) inactivation experiments using antibodies directed against defined regions of the ZP3 glycoprotein (Millar, S. (1989) *Science* 246:935-938), and exon-swapping experiments (Kinloch, R. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:263-267), putative O-linked glycosylation sites encoded by the 7th of 8 exons in the murine ZP3 gene were implicated in acrosome binding. Crude site-directed mutagenesis

experiments have further delimited the essential O-linked glycosylation site to five serine residues (and no threonine residues) between positions 328-343 (inclusive) of the 424-residue ZP3 polypeptide (Kinloch, R. et al. supra). Interestingly, murine and human ZP3 polypeptides are only 28% identical in this critical region, while being 67% identical overall and sharing all recognized domains. These phylogenetic data suggest that ZP3 protein may function primarily as a scaffold for the presentation of the correct carbohydrate moieties rather than participate in biological interactions as a polypeptide molecule. (Rosiere, T. supra; Snell, W. and White, J. supra; and Wassarman, P. supra). The primary amino acid sequence of the ZP3 polypeptide is less important to acrosome binding than the specific glycosylation that occurs in the exon-7 region (Kinloch, R. et al. supra).

Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration; therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) *Int. J. Oncol.* 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide of corticotropin (ACTH), a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor polypeptide of the hormone beta-lipotropin (beta-LPH). Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH; while beta-MSH is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC results in an endocrine disorder characterized by early-

onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al. (1979) Can. J. Biochem. 57:1111-1121; Krude, H. et al. (1998) Nat. Genet. 19:155-157; Online Mendelian Inheritance in Man (OMIM) 176830).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin, which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

The Slit protein, first identified in *Drosophila*, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. ((1998) Brain Res. Mol. Brain Res. 62:175-186) have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EGF-like motifs and leucine-rich repeats, both of which are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., supra). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and it is suggested that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al. (1999) J. Biol. Chem. 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and Fleetwood-Walker, S.M. (1998) *Trends Pharmacol. Sci.* 19:346-348).

Other proteins that contain signal peptides include secreted proteins with enzymatic activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al. (1995) *Dev. Dyn.* 202:388-396; Firestein, G.S. (1992) *Curr. Opin. Rheumatol.* 4:348-354; Ray, J.M. and Stetler-Stevenson, W.G. (1994) *Eur. Respir. J.* 7:2062-2072; and Mignatti, P. and Rifkin, D.B. (1993) *Physiol. Rev.* 73:161-195). Additional examples are the acetyl-CoA synthetases which

activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-26466). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA. Acetyl-CoA synthetases share a region of sequence similarity identified as the AMP-binding domain signature. Acetyl-CoA synthetase has been shown to be associated with hypertension
5 (Toh, H. (1991) Protein Seq. Data Anal. 4:111-117; and Iwai, N. et al. (1994) Hypertension 23:375-380).

A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds
10 in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs). FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and
15 localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336-29341; Schreiber, S.L. (1991) Science 251:283-287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein
20 activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the *gag* protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication.
25 Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204-23214; Hunter, T. (1998) Cell 92:141-143; and Leverson, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family
30 of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al. (2001) Proc. Natl. Acad. Sci. USA
35 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is

Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A. and C.T. Przysiecki (1987) *Int. J. Biochem.* 19:1-7; C. Vermeer (1990) *Biochem. J.* 266:625-636).

The *Drosophila* sp. gene *crossveinless 2* is characterized as having a putative signal or transmembrane sequence, a partial Von Willebrand Factor D domain similar to those domains known to regulate the formation of intramolecular and intermolecular bonds, and five cysteine-rich domains, known to bind BMP-like (bone morphogenetic proteins) ligands. These features suggest that *crossveinless 2* may act extracellularly or in the secretory pathway to directly potentiate ligand signaling and hence, involvement in the BMP-like signaling pathway known to play a role in vein specification (Conley, C.A. et al., (2000) *Development* 127:3947-3959). The dorsal-ventral patterning in both vertebrate and *Drosophila* embryos requires a conserved system of extracellular proteins to generate a positional informational gradient.

Immunoglobulins

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. Most cell surface and soluble molecules that mediate functions such as recognition, adhesion or binding have evolved from a common evolutionary precursor (i.e., these proteins have structural homology). A number of molecules outside the immune system that have similar functions are also derived from this same evolutionary precursor. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. The criteria for a protein to be a member of the Ig superfamily is to have one or more Ig domains, which are regions of 70-110 amino acid residues in length homologous to either Ig variable-like (V) or Ig constant-like (C) domains. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors, neural cell-adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR). These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic

antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of β -sheets. Conserved cysteine residues form an intrachain disulfide-bonded loop, 55-75 amino acid residues in length, which connects the two layers of the β -sheets. Each β -sheet has three or four anti-parallel β -strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the β -strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an additional pair of β -strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead attached to the plasma membrane by a covalent glycosylphosphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic

interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody and pairing of variable regions of light and heavy chains form the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp.142-145.)

Antibodies

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246.)

Antibodies are multimeric members of the Ig superfamily which are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α , δ , ϵ , γ , and μ H-chain types. There are two types of L-chains, κ and λ , either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H-

and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site.

5 (Reviewed in Alberts, B. et al. *supra*, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain the repeated Ig domains of members of the Ig superfamily. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one
10 of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-
15 specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

20 The discovery of new secreted proteins, 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 cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

25

SUMMARY OF THE INVENTION

The invention features purified polypeptides, secreted proteins, referred to collectively as "SECP" and individually as "SECP-1," "SECP-2," "SECP-3," "SECP-4," "SECP-5," "SECP-6," "SECP-7," "SECP-8," "SECP-9," "SECP-10," "SECP-11," "SECP-12," "SECP-13," "SECP-14,"
30 "SECP-15," "SECP-16," "SECP-17," "SECP-18," "SECP-19," "SECP-20," "SECP-21," "SECP-22," "SECP-23," and "SECP-24." In one 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:1-24, 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
35 ID NO:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence

selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-24.

5 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 group consisting of SEQ ID NO:1-24, 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:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence
10 selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-24. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:25-48.

15 Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding 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:1-24, 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:1-24, c) a
20 biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. 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.

25 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:1-24, 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:1-24, c) a
30 biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. 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.

35 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:1-24, 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:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

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:25-48, 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:25-48, 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:25-48, 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:25-48, 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:25-48, 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:25-48, 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:1-24, 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:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, 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:1-24, 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:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. 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 SECP, 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:1-24, 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:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the

group consisting of SEQ ID NO:1-24. 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
5 treating a disease or condition associated with overexpression of functional SECP, 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:1-24, b) a polypeptide comprising a
10 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. The method comprises a) combining the polypeptide with at least
15 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:1-24, b) a polypeptide comprising a
20 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24. The method comprises a) combining the polypeptide with at least
25 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.

30 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 polynucleotide sequence selected from the group consisting of SEQ ID NO:25-48, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target
35 polynucleotide in the presence of varying amounts of the compound and in the absence of the

compound.

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:25-48, 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:25-48, 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:25-48, 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:25-48, 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.

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 scores for the matches between each polypeptide and its homolog(s) are 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.

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.

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 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

“SECP” refers to the amino acid sequences of substantially purified SECP 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 SECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

An “allelic variant” is an alternative form of the gene encoding SECP. 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 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.

- 5 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 SECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding SECP. 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 SECP. 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 SECP is retained. For example, 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.

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.

“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 SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP 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. Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide
5 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, 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
10 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 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 “aptamer” refers to a nucleic acid or oligonucleotide molecule that binds to a
15 specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include
20 deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH₂), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system.
25 Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

The term “intramer” refers to an aptamer which is expressed in vivo. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl Acad. Sci. USA 96:3606-3610).

The term “spiegelmer” refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term “antisense” refers to any composition capable of base-pairing with the “sense”
35 (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA;

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
5 molecules may be produced by any method including chemical synthesis or transcription. Once 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.

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

15 "Complementary" describes the relationship between two single-stranded nucleic acid 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
20 or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP 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
25 sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"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
30 program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison 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
35 the protein is conserved and not significantly changed by such substitutions. The table below shows

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

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 diseased and a normal sample.

“Exon shuffling” refers to the recombination of different coding regions (exons). Since an

exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP 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:25-48 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:25-48, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:25-48 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:25-48 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:25-48 and the region of SEQ ID NO:25-48 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:1-24 is encoded by a fragment of SEQ ID NO:25-48. A fragment of SEQ ID NO:1-24 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-24. For example, a fragment of SEQ ID NO:1-24 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-24. The precise length of a fragment of SEQ ID NO:1-24 and the region of SEQ ID NO:1-24 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 "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.

5 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
10 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
15 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
20 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
25 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

30 *Open Gap: 5 and Extension Gap: 2 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

35 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 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 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 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.

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

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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 length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be
5 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.

The term “humanized antibody” refers to an antibody molecule in which the amino acid
10 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 hybridization is an indication that two nucleic acid sequences share a high degree of complementarity.
15 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 conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill
20 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.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature
25 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 conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al.
30 (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, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC
35 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 $\mu\text{g/ml}$. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions
5 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
10 hybridization complex may be formed in solution (e.g., C_0t or R_0t 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
15 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.

20 An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SECP 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 SECP which is useful in any of the antibody production methods disclosed herein or known in the art.

25 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 SECP. For example, modulation
30 may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SECP.

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
35 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
5 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. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript
10 elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SECP 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 vary by cell type depending on the enzymatic milieu of SECP.

"Probe" refers to nucleic acid sequences encoding SECP, 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.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target
20 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 nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also
25 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.

Methods for preparing and using probes and primers are described in the references, for
30 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 can be derived from a known sequence, for example, by using computer programs intended for that
35 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 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 use 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.

5 “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.

10 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.

The term “sample” is used in its broadest sense. A sample suspected of containing SECP, nucleic acids encoding SECP, or fragments thereof may comprise a bodily fluid; an extract from a
15 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.

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
20 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.

The term “substantially purified” refers to nucleic acid or amino acid sequences that are
25 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.

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

30 “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.

A “transcript image” or “expression profile” refers to the collective pattern of gene
35 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 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 alternate 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-07-1999) 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.

THE INVENTION

The invention is based on the discovery of new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

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 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 SEQ ID NO:1-2, SEQ ID NO:10-13, and SEQ ID NO:17-22, as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show polypeptides of SEQ ID NO:1-2, SEQ ID NO:10-13, and SEQ ID NO:17-22 and their corresponding Incyte polypeptide sequence numbers (Incyte Polypeptide ID). Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the matches between each polypeptide and its homolog. Column 5 shows the annotation of the GenBank database homologs along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of each 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 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, including the locations of signal peptides (as indicated by "Signal Peptide" and/or "signal_cleavage".) 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 secreted proteins. For example, SEQ ID NO:1 is 30% identical, from residue C101 to residue C449, to Arabidopsis thaliana pectinacetylesterase (GenBank ID g6478931) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $1.1e-22$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a signal peptide as determined by searching for statistically significant matches in the HMMER database of conserved protein family domains. (See Table 3.) In an alternative example, SEQ ID NO:2 is 76% identical, from residue M1 to residue D185, to human pancreatitis-associated protein (PAP) (GenBank ID g482909) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $7.0e-76$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains a lectin C-type 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:2 is a pancreatitis-associated protein. Note that the C-type lectin domain is found in PAP. In an alternative example, SEQ ID NO:7-9 contain signal peptides as determined by the HMMER algorithm used for searching for statistically significant matches in the hidden Markov model (HMM)-based database of conserved protein family domains. Data from Spscan provides further corroborative evidence that SEQ ID NO:7-9 are secreted proteins. In an alternative example, SEQ ID NO:13 is 99% identical, from residue M1 to residue F331, to human pregnancy-specific beta-1-glycoprotein (GenBank ID g190647) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $4.1e-181$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:13 also contains immunoglobulin domains 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.) In an alternative

example, SEQ ID NO:19 is 40% identical, from residue S177 to residue P317, to the murine secreted polypeptide, ZSIG37 (GenBank ID g6274477), as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $6.6e-18$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:19

5 also contains C1q domains 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:19 is a secreted polypeptide with C1q domains. In an alternative example, SEQ ID NO:20 is 68% identical, from residue W9 to residue A631, to the rat zona pellucida (ZP) 1 glycoprotein

10 (GenBank ID g2804566), as determined by BLAST analysis, with a probability score of $3.7e-230$. SEQ ID NO:20 also contains a ZP domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. Data from BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:20 is a ZP-related protein. In an alternative example, SEQ ID NO:21 is 42%

15 identical, from residue L11 to residue S320, to a human immunoglobulin superfamily member protein (GenBank ID g7767239), as determined by BLAST analysis, with a probability score of $5.9e-77$. SEQ ID NO:21 also contains immunoglobulin domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. Data from BLIMPS analysis provides further corroborative evidence that SEQ ID

20 NO:21 is a member of the immunoglobulin superfamily. In an alternative example, SEQ ID NO:22 is 96% identical, from residue M1 to residue D225, to murine gliacolin, a C1q-like protein expressed in Glial cells (GenBank ID g10566471), as determined by BLAST analysis, with a probability score of $1.2e-137$. SEQ ID NO:22 also contains a C1q domain, and a collagen repeat motif characteristic of C1q proteins, as determined by searching for statistically significant matches in the hidden Markov

25 model (HMM)-based PFAM database of conserved protein family domains. Data from BLIMPS and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:22 is a C1q-related polypeptide. The algorithms and parameters for the analysis of SEQ ID NO:3-6, SEQ ID NO:10-12, SEQ ID NO:14-18, and SEQ ID NO:23-24, were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-24 are described in Table 7.

30 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. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of

35 each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3')

positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:25-48 or that distinguish between SEQ ID NO:25-48 and related polynucleotide sequences.

5 The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from
10 the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation “ENST”). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation “NM” or “NT”) or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation “NP”). Alternatively, the polynucleotide fragments described in
15 column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an “exon stitching” algorithm. For example, a polynucleotide sequence identified as FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ represents a “stitched” sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N_{1,2,3...}, if present, represent specific
20 exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an “exon-stretching” algorithm. For example, a polynucleotide sequence identified as FLXXXXXX_gAAAAA_gBBBBB_1_N is a “stretched” sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human
25 genomic sequence to which the “exon-stretching” algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the “exon-stretching” algorithm, a RefSeq identifier (denoted by “NM,” “NP,” or “NT”) may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

30 Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
--------	--

GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not 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 which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses SECP variants. A preferred SECP 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 SECP amino acid sequence, and which contains at least one functional or structural characteristic of SECP.

The invention also encompasses polynucleotides which encode SECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:25-48, which encodes SECP. The polynucleotide sequences of SEQ ID NO:25-48, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. 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 SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:25-48 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:25-48. Any one of the polynucleotide variants described above can encode an amino

acid sequence which contains at least one functional or structural characteristic of SECP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding SECP. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding SECP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding SECP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding SECP. For example, a polynucleotide comprising a sequence of SEQ ID NO:36 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:47, and a polynucleotide comprising a sequence of SEQ ID NO:42 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:48. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

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 SECP, 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 combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

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

SECP 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 SECP or any fragment thereof.

5 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:25-48 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
10 "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
15 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
20 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.)

25 The nucleic acid sequences encoding SECP 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.)
30 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 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,
35 M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme

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 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 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 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 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 SECP may be cloned in recombinant DNA molecules that direct expression of SECP, 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 SECP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SECP-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 No.

5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat.

5 Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SECP, 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
10 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
15 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 SECP 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
20 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, SECP 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

25 may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, 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

30 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 sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active SECP, the nucleotide sequences encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains

35 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, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and 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 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 SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, 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 encoding SECP. 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 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 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.

(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 SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT1 plasmid (Life Technologies). Ligation of sequences encoding SECP 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 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 SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP 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 SECP. A number of vectors 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) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of SECP. Transcription of sequences encoding SECP 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 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.)

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 SECP 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 SECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc.

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 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 SECP in cell lines is preferred. For example, sequences encoding SECP can be transformed 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 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 *aprt* cells, respectively. (See, e.g., Wigler, M. et 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) 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 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 SECP is inserted within a marker gene sequence, transformed cells containing

sequences encoding SECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SECP 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.

5 In general, host cells that contain the nucleic acid sequence encoding SECP and that express SECP 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.

10 Immunological methods for detecting and measuring the expression of SECP using either 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 SECP is preferred, but a
15 competitive binding assay may be employed. These and other assays are well known in the art. (See, 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.)

20 A wide variety of labels and conjugation techniques are known by those skilled in the art and 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 SECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding SECP, or any fragments thereof, may be cloned into a vector
25 for the production of an mRNA probe. Such vectors are known in the art, are commercially available, 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
30 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 SECP 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
35 and/or the vector used. As will be understood by those of skill in the art, expression vectors

containing polynucleotides which encode SECP may be designed to contain signal sequences which direct secretion of SECP 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 SECP 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 SECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP 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 SECP encoding sequence and the heterologous protein sequence, so that SECP 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.

In a further embodiment of the invention, synthesis of radiolabeled SECP 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.

SECP of the present invention or fragments thereof may be used to screen for compounds

that specifically bind to SECP. At least one and up to a plurality of test compounds may be screened for specific binding to SECP. 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 SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a 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 SECP 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 SECP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is 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 SECP, either in solution or affixed to a solid support, and detecting the binding of SECP 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 libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

SECP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP 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 SECP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem

(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 No. 5,175,383 and U.S. Patent No. 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; 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 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 SECP may also be manipulated in vitro in ES cells derived from 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 SECP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding SECP 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. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP 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 between regions of SECP and secreted proteins. In addition, examples of tissues expressing SECP can be found in Table 6. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

Therefore, in one embodiment, SECP 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 SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed
5 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, a cancer 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,
10 salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; 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
15 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
20 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; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive
25 heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula,
30 atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; 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
35 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-

- 5 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 disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other
- 10 neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration,
- 15 and familial frontotemporal dementia; and 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-
- 20 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.

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

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

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

In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such
35 disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory,

cardiovascular, neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

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

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
10 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 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.

15 An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may 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
20 fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have advantages in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

25 For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with SECP 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
30 as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in 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 SECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will
35 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 of SECP 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 SECP 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 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 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 SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be 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* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for SECP 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 SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP 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 SECP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions.

5 The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP 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
10 SECP-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 SECP, 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,
15 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
20 of SECP-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 SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene
25 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 SECP. 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 SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

30 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
35

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 SECP 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 SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-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 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 SECP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors

(Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SECP 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 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 H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP 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 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 SECP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SECP 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 No. 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).

5 In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. 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
10 (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 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.

15 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP 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
20 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 No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which
25 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
30 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
35 deliver polynucleotides encoding SECP 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 SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP 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 SECP 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 SECP.

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
5 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 SECP. 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
10 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
15 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.

20 An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. 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
25 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 SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with
30 decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP 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
35 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 SECP is exposed to at least one test compound thus obtained. The sample
5 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 SECP 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 SECP. The amount of hybridization may be quantified, thus forming
10 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
15 (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
20 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.
25 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
30 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
35 Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may

consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

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, 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 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.

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 SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP 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).

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.

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, 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 ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the 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 LD₅₀/ED₅₀ 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. 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 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 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 SECP may be used for the diagnosis of disorders characterized by expression of SECP, or in assays to monitor patients being treated with SECP or agonists, antagonists, or inhibitors of SECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SECP include methods which utilize the antibody and a label to detect SECP 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 the art and may be used.

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP 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 SECP may be used 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 SECP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SECP, and to monitor regulation of SECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SECP or closely related molecules may be used to identify nucleic acid sequences which encode SECP. 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 probe identifies only naturally occurring sequences encoding SECP, 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 SECP 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:25-48 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for DNAs encoding SECP include the cloning of polynucleotide sequences encoding SECP or SECP 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 ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, 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, a cancer 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; 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
 5 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
 10 colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal
 circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a
 cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris,
 myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic
 valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse,
 15 rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic
 endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy,
 myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac
 transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease,
 aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular
 20 tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary
 artery bypass graft surgery; 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
 25 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
 30 hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental
 disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal
 disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases,
 muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders,
 dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia
 35 gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders,

seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and 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. The polynucleotide sequences encoding SECP 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 SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue 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 SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate 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 SECP, 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 sequence, or a fragment thereof, encoding SECP, 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 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 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 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 SECP 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 encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, 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 SECP 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 SECP 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).

SNPs may be used to study the genetic basis of human disease. For example, at least 16

common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic
5 fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase
10 pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations. (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641.)

Methods which may also be used to quantify the expression of SECP include radiolabeling or
15 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
20 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
25 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
30 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 pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein
35 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 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 No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by 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.

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 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 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 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 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 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
5 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.

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
10 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 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
15 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, *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
20 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 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
25 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 obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray,
30 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; Mendozze, 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
35 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
5 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
10 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
15 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
20 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.,
25 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.
30 (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP 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
35 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 SECP 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, SECP, 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 between SECP 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 SECP, or fragments thereof,

and washed. Bound SECP is then detected by methods well known in the art. Purified SECP 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.

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

In additional embodiments, the nucleotide sequences which encode SECP may be used in any 10 molecular biology techniques that have yet to be developed, provided the new techniques rely on 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 embodiments are, 15 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. 60/267,924, U.S. Ser. No. 60/266,195, U.S. Ser. No. 60/268,112, U.S. Ser. No. 60/267,816, U.S. Ser. No. 60/271,639, U.S. Ser. No. 60/317,818, and U.S. Ser. No. 60/343,553, 20 are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database 25 (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a 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 30 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 35 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), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), 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

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 cycler or the PTC-200 thermal

cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared 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).

5 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 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,
10 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 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
15 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; PROTEOME databases with sequences from Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Candida albicans (Incyte Genomics, Palo Alto CA); hidden
20 Markov model (HMM)-based protein family databases such as PFAM; and HMM-based protein domain databases such as SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (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,
25 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 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
30 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 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, the PROTEOME databases, BLOCKS,
35 PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases

such as PFAM; and HMM-based protein domain databases such as SMART. 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:25-48. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins 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 secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. 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
5 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
10 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
15 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
20 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
25 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
30 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
35 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.

5 VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:25-48 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:25-48 were assembled into clusters of contiguous and overlapping sequences using
10 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 Généthon 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.

15 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
20 distances are based on genetic markers mapped by Généthon 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.

25 In this manner, SEQ ID NO:36 was mapped to chromosome 3 within the interval from 142.20 to 148.70 centiMorgans. SEQ ID NO:37 was mapped to chromosome 19 within the interval from 51.00 to 51.70 centiMorgans, and within the interval from 62.00 to 69.90 centiMorgans, and to chromosome 5 within the interval from 141.40 to 142.60 centiMorgans. More than one map location is reported for SEQ ID NO:37, indicating that sequences having different map locations were
30 assembled into a single cluster. This situation occurs, for example, when sequences having strong similarity, but not complete identity, are assembled into a single cluster.

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
35 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:

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

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 SECP 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; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided

by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of SECP Encoding Polynucleotides

5 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
10 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.

15 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
20 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.

25 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
30 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
35 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

5 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.

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

10 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 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

15 primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

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.

20 IX. Identification of Single Nucleotide Polymorphisms in SECP Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:25-48 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was

25 used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated

30 algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

35 Certain SNPs were selected for further characterization by mass spectrometry using the high

throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezuelan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:25-48 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base 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 μ Ci of [γ - ^{32} P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 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).

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.

XI. 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.)

- 5 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.
- 10 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
- 15 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/ μ l oligo-(dT) primer (21mer), 1X
- 20 first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M 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
- 25 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
- 30 then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

- 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
- 35 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 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia 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, 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 U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic 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 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g 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 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 μ l 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
5 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
10 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
15 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
20 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
25 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).

XII. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to
30 detect, decrease, or inhibit expression of naturally occurring SECP. Although use of 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 SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent
35 promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is

designed to prevent ribosomal binding to the SECP-encoding transcript.

XIII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector

5 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 SECP upon induction with isopropyl beta-D-
10 thiogalactopyranoside (IPTG). Expression of SECP 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 SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong
15 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.)

20 In most expression systems, SECP 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
25 Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from SECP 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,
30 supra, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, and XIX, where applicable.

XIV. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression
35 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 SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP 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 SECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of SECP Specific Antibodies

SECP 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 immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the SECP 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 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
5 oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XVI. Purification of Naturally Occurring SECP Using Specific Antibodies

10 Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP 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.

15 Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SECP 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 SECP is collected.

XVII. Identification of Molecules Which Interact with SECP

SECP, 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 SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations
25 of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate molecules.

Alternatively, molecules interacting with SECP 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).

30 SECP 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).

XVIII. Demonstration of SECP Activity

35 An assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA

synthesis in Swiss mouse 3T3 cells (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of [³H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations.

- 5 Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, 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 SECP concentration range is indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal
- 10 incorporation of [³H]thymidine into acid-precipitable DNA .

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped Xenopus myocyte is manipulated into contact with one of the fibroblasts in SECP-free

15 medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) *Neuron* 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The

20 lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting

25 techniques. The concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to the amount of SECP in transit through the secretory pathway.

Alternatively, AMP binding activity is measured by combining SECP with ³²P-labeled AMP. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to remove unbound label. The radioactivity

30 retained in the gel is proportional to SECP activity.

XIX. Demonstration of Immunoglobulin Activity

An assay for SECP activity measures the ability of SECP to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E.S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.)

35 SECP is isotopically labeled using methods known in the art. Various serum concentrations are

added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a
5 characteristic precipitin curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

10 Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP 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 SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled
15 immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Alternatively, an assay for SECP activity measures the amount of cell aggregation induced by overexpression of SECP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding SECP contained within a suitable mammalian expression vector under control of a strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green
20 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 SECP activity.

Various modifications and variations of the described methods and systems of the invention
25 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
30 scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1849820	1	1849820CD1	25	1849820CB1
70610307	2	70610307CD1	26	70610307CB1
8137559	3	8137559CD1	27	8137559CB1
4801255	4	4801255CD1	28	4801255CB1
6160719	5	6160719CD1	29	6160719CB1
3524602	6	3524602CD1	30	3524602CB1
513718	7	513718CD1	31	513718CB1
896308	8	896308CD1	32	896308CB1
2105862	9	2105862CD1	33	2105862CB1
7939381	10	7939381CD1	34	7939381CB1
7487507	11	7487507CD1	35	7487507CB1
1483931	12	1483931CD1	36	1483931CB1
8175223	13	8175223CD1	37	8175223CB1
4173218	14	4173218CD1	38	4173218CB1
5014679	15	5014679CD1	39	5014679CB1
7487510	16	7487510CD1	40	7487510CB1
2682619	17	2682619CD1	41	2682619CB1
4582105	18	4582105CD1	42	4582105CB1
931619	19	931619CD1	43	931619CB1
2155025	20	2155025CD1	44	2155025CB1
7640495	21	7640495CD1	45	7640495CB1
5960119	22	5960119CD1	46	5960119CB1
7500143	23	7500143CD1	47	7500143CB1
7503605	24	7503605CD1	48	7503605CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	1849820CD1	g6478931	1.10E-22	[Arabidopsis thaliana] putative pectinacetyl esterase
2	70610307CD1	g482909	7.00E-76	[Homo sapiens] pancreatitis-associated protein Duseti, N.J. et al. (1994) Genomics 19:108-114
10	7939381CD1	g3986746	9.10E-25	[Bos taurus] tuftelin Bashir, M.M. et al. (1997) Arch. Oral Biol. 42: 489-496
11	7487507CD1	g854326	0	[Mus musculus] semaphorin B Puschel, A.W. et al. (1995) Neuron 14: 941-948
12	1483931CD1	g186396	7.70E-17	[Homo sapiens] mucin Gum, J.R. Jr. et al. (1992) J. Biol. Chem. 267: 21375-21383
13	8175223CD1	g190647	4.10E-181	[Homo sapiens] pregnancy-specific beta-1-glycoprotein Plouzek, C.A. et al. (1991) Biochem. Biophys. Res. Commun. 176: 1532-1538
17	2682619CD1	g5931953	8.90E-26	[Mus musculus] autocrine motility factor receptor Shimizu, K. et al. (1999) FEBS Lett. 456: 295-300
18	4582105CD1	g7294667	2.20E-41	[Drosophila melanogaster] CG14130 gene product (Adams, M.D. et al. (2000) Science 287: 2185-2195)
19	931619CD1	g13274522	0	complement-c1q tumor necrosis factor-related protein [Homo sapiens]
20	2155025CD1	g2804566	3.70E-230	[Rattus norvegicus] zona pellucida 1 glycoprotein Akatsuka, K. et al. (1998) Mol. Reprod. Dev. 51:454-467.
21	7640495CD1	g7767239	5.90E-77	[Homo sapiens] immunoglobulin superfamily member protein
22	5960119CD1	g10566471	1.20E-137	[Mus musculus] Gliacolin Koide, T. et al. (2000) J. Biol. Chem. 275:27957-27963.

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	1849820CD1	496	S117 S138 S148 S164 S192 S193 S195 S417 S421 T25 T57 T100 T142 T327 T374 T451 T457 Y140	N96	signal_cleavage: M1-G19 Signal Peptides: M1-G19, M1-G22 Transmembrane domain: E214-T236 N-terminus is cytosolic PECTIN ACETYLESTERASE PROTEIN PUTATIVE PRECURSOR F21E10.11 SIGNAL F14J9.21 PD009474: W120-K272	SPSCAN HMMER TMAP BLAST_PRODOM
2	70610307CD1	185	S35 S57 S60 S73 S81 S113 S123 T24 T29		signal_cleavage: M1-G26 Signal Peptides: M5-P25, M5-E28, M1-E27, M1-E28 Lectin C-type domain: L66-C181 Transmembrane domain: P4-S22 N-terminus is non-cytosolic C-type lectin domain proteins BL00615: C51-C68, W168-C181 C-type lectin domain signature and profile: D126-K184 PRECURSOR SIGNAL PROTEIN LECTIN REG LITHOSTATHINE REGENERATING INFLAMMATORY RESPONSE ACUTE PD149843: G26-A67 LECTIN PROTEIN PRECURSOR GLYCOPROTEIN SIGNAL RECEPTOR TRANSMEMBRANE REPEAT CELL DOMAIN PD000254: C68-C181 C-TYPE LECTIN DM00035 Q06141 33-172: L33-F183 DM00035 P35230 33-172: L33-F183 DM00035 P23132 33-172:L33-F183 DM00035 S54979 33-171: L33-F183 C-type lectin domain signature: C156-C181	SPSCAN HMMER HMMER_PFAM TMAP BLIMPS_BLOCKS PROFILESAN BLAST_PRODOM BLAST_PRODOM BLAST_DOMO MOTIFS

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
3	8137559CD1	73			signal_cleavage: M1-A58 Signal Peptides: M8-D28, M8-P31, M3-P29, M1-P31 Transmembrane domains: M8-L26 G36-A58 N-terminus is cytosolic Phosphoenolpyruvate carboxykinase (ATP) signature: I14-L69	SPSCAN HMMER TMAP PROFILESCAN
4	4801255CD1	75	S22 S23 T28 T37 T57	N4 N20	signal_cleavage: M1-P18 Signal Peptides: M1-Q19, M1-M24	SPSCAN HMMER
5	6160719CD1	96			signal_cleavage: M1-A22 Signal Peptide: M1-A36	SPSCAN HMMER
6	3524602CD1	87	S2 S12 S60 S83 T45	N65 N81	Signal Peptide: M25-K47	HMMER
7	513718CD1	96	S17 S26		signal_cleavage: M1-S17 Signal Peptide: M1-S22 Transmembrane domain: K57-F80 N-terminus is non-cytosolic Glycosyl hydrolases family 17 signature: L31-I84	SPSCAN HMMER TMAP PROFILESCAN
8	896308CD1	101	S3 S57 S67 S90 T71 T96		signal_cleavage: M1-P33 Signal Peptides: M6-T36, M6-H38	SPSCAN HMMER
9	2105862CD1	93			signal_cleavage: M1-A19 Signal Peptide: M1-Y20 Transmembrane domains: I4-Y20, G65-I81 N-terminus is cytosolic	SPSCAN HMMER TMAP

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	7939381CD1	466	S23 S59 S155 S229 S247 S261 S291 S334 S348 S376 S396 S407 T64 T177 T242 T332 T419 T437 T460		signal_cleavage: M1-G20 COILED COIL CHAIN MYOSIN REPEAT HEAVY ATP BINDING FILAMENT HEPTAD-K PD000002: L195-E425, Q175-K384, L222-E425, Q131-M370, L174-Q385 CAP-GLY DOMAIN DM03881 P35458 1-1052: T182-I426 TRICHOHYALIN DM03839 P37709 632-1103: E38-L387	SPSCAN BLAST_PRODOR
11	7487507CD1	730	S106 S111 S119 S176 S252 S342 S385 S411 S467 S471 S502 S524 S543 S707 S714 S716 T152 T250 T262 T284 T361 T366 T690	N120 N135 N465 N576	Signal Peptide:M1-A31 Sema domain:F64-V447 Transmembrane domains: P4-T29, V433-V461, S648-L676 N terminus is non-cytosolic. SEMAPHORIN B PRECURSOR IMMUNOGLOBULIN FOLD MULTIGENE FAMILY NEUROGENESIS DEVELOPMENTAL PD116663: P512-A730 PD107003: M1-D63 SEMAPHORIN RECEPTOR KINASE SIGNAL TYROSINE HEPATOCYTEPD001844: L67-F245, V217-C379, G369-V447 SEMAPHORIN; FASCICLIN; COLLAPIN; II DM01606 I48745 1-619: M1-V447, L443-L590 DM01606 A49069 1-646: F64-C379, P238-V447, V450-W560 DM01606 I48747 1-646: L24-C379, P238-V447, V450-W560, P548-L581 DM01606 I48744 1-639: A31-V447, V450-T573	BLAST_DOMO BLAST_DOMO BLAST_DOMO HMMER HMMER_PFAM TMAP BLAST_PRODOR BLAST_PRODOR BLAST_DOMO

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
12	1483931CD1	575	S33 S57 S126 S127 S171 S172 S179 S317 S330 S465 T85 T150 T161 T162 T236 T342 T409 T411 T438 T498 T533	N178 N223 N261 N446 N504 N509	signal_cleavage: M1-A28 Signal Peptide: M1-A28 Transmembrane domains: M1-A28, A470-K496 N terminus is non-cytosolic.	SPSCAN HMMER TMAP
13	8175223CD1	335	S96 S226 S302 T113 T146 T190 T198 T292 T308 T317	N61 N104 N111 N199	signal_cleavage: M1-A34 Signal Peptide: M1-A34 Immunoglobulin domains: G254-A303, T63-T121, M162-I219 Transmembrane domain: I13-L30 N terminus is cytosolic. PRECURSOR GLYCOPROTEIN SIGNAL PREGNANCY-SPECIFIC CARCINOEMBRYONIC ANTIGEN IMMUNOGLOBULIN FOLD PREGNANCY MULTIGENE PD000677: N29-T113 PRECURSOR GLYCOPROTEIN SIGNAL PREGNANCY-SPECIFIC IMMUNOGLOBULIN FOLD MULTIGENE FAMILY BETA1 PD002253: P147-E218 PD149666: R114-T146 CARCINOEMBRYONIC ANTIGEN PRECURSOR AMINO-TERMINAL DOMAIN DM00372 A54312 38-148: I38-P149 DM00372 P11462 38-128: I38-D129 DM00372 JC4121 38-128: I38-D129 RGD motif: R127-D129	SPSCAN HMMER HMMER_PFAM TMAP BLAST_PRODOR BLAST_PRODOR BLAST_DOMO MOTIFS

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
					IMMUNOGLOBULIN DM00001 JC4122 28-112: S150-L235	BLAST_DOMO
14	4173218CD1	120	S32 S46 S96 T89 T92 Y93	N36	signal_cleavage: M1-G26 Signal Peptides: M1-A24, M1-G26 Transmembrane domains: P4-R23, L52-V80 N terminus is cytosolic.	SPSCAN HMME TMAP
15	5014679CD1	103	S38 S50 S95		Signal Peptide: M1-P26	HMME
16	7487510CD1	170	S37 S45 S111 T122		signal_cleavage: M1-A36 Signal Peptide: M18-A36 Transmembrane domain: N116-A136 N terminus is non-cytosolic.	SPSCAN HMME TMAP
17	2682619CD1	617	S100 S265 S399 S444 S506 S542 T280 T328 T485 T512	N132 N541	Signal Peptide: M1-A22 Zinc finger, C3HC4 type (RING finger): C291-C329 Transmembrane domains: M1-K26, K38-F66, F99-A120, H138-H158, G167-T187, T215-F243 N terminus is cytosolic. PROTEIN ZINC FINGER F55A11.3 CHROMOSOME XV READING FRAME YOL013C PD040235: M8-D288 PROLINE-RICH PROTEIN DM03894 A39066 1-159: P339-P468, P339-P407, A340-P475	HMME HMME_PFAM TMAP BLAST_PRODOR
18	4582105CD1	221	S41 S124 S170 S199 S204 T42 T76		signal_cleavage: M1-V18 Transmembrane domain: R115-M143 N terminus is cytosolic.	SPSCAN TMAP

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
19	931619CD1	329	T38, S244, S264		Signal cleavage: M1-A16 Signal Peptide: M1-A16 Clq domain: A31-V156, A178-V311 Clq domain proteins: BL01113: I304-P313, A47-A82, R116-R135 Complement C1Q domain signature: PR00007: P194-V220, F68-H87, R116-L137, K302-Y312 PRECURSOR SIGNAL COLLAGEN REPEAT HYDROXYLATION GLYCOPROTEIN CHAIN PLASMA EXTRACELLULAR MATRIX: PD002992: A178-V311, A31-V156 C1Q DOMAIN DM00777 P23206 477-673: S30-V156, P200-P313, G168-G211 Cell attachment sequence: R280-D282 Clq domain signature: F50-F80 F203-F233	SPSCAN HMMER HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM
20	2155025CD1	640	S183, S253, S486, S519, S539, S559, S600, S637, T59, T85, T252, T299, T552, T565	N76, N381, N563, N598	Signal cleavage: M1-G25 Signal cleavage: M1-G30 Zona pellucida-like domain: Q281-G551 Transmembrane domain: G4-R26, S600-Q628; N-terminus is cytosolic ZP domain proteins: BL00682: I452-L476, R506-S519, I366-N381 ZP domain signature (zp_domain.prf): V440-P499 Zona pellucida sperm-binding protein signature: PR00023: F330-A342, P450-A465, Q469-S486, G525-G540 ZONA PELLUCIDA GLYCOPROTEIN PRECURSOR ZP1 SIGNAL: PD023337: W9-H300	SPSCAN HMMER HMMER-PFAM TMAP BLIMPS-BLOCKS PROFILES SCAN BLIMPS-PRINTS 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
					GLYCOPROTEIN PRECURSOR SIGNAL ZONA PROTEIN PELLUCIDA TRANSMEMBRANE RECEPTOR SPERM-BINDING SPERM: PD002550: C282-R554	BLAST-PRODODM
					ZONA PELLUCIDA GLYCOPROTEIN PRECURSOR ZP1 SIGNAL: PD023287: T550-R639	BLAST-PRODODM
					ZP DOMAIN: DM01748 Q07287 154-492: P246-V578	BLAST-DOMO
21	7640495CD1	388	S83, S170, S264, S294, S348, S361, T33, T43, T155, T183, T225, Y102	N31, N67, N262, N286, N375	Signal cleavage: M1-G24 Signal Peptide: M1-G24 Immunoglobulin domain: G238-A293, G138-A201, G37-L106 Transmembrane domain: Q318-W346; N-terminus is not cytoplasmic	SPSCAN HMMER HMMER-PFAM
					GLYCOPROTEIN ANTIGEN PRESENTATION: PD02327: L118-V129, S144-L165	BLIMPS-PRODODM
					POLIOVIRUS RECEPTOR DM02194 A53437 1-399: V39-F373	BLAST-DOMO
22	5960119CD1	255	T27, T160		Signal Peptide: M1-P18 C1q domain: A128-I252 Collagen triple helix repeat (20 copies): G61-T120 C1q domain proteins: BL01113: G70-P96, V142-V177, D211-K230, S245-A254 Complement C1q domain signature: PR00007: Q136-K162, F163-G182, D211-D232, K243-Y253 PRECURSOR SIGNAL COLLAGEN REPEAT HYDROXYLATION GLYCOPROTEIN CHAIN PLASMA EXTRACELLULAR MATRIX: PD002992: E138-I252 C1Q DOMAIN: DM00777 S23297 465-674: I59-I251 C1q domain signature: F145-Y175	HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODODM BLAST-DOMO MOTIFS

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
23	7500143CD1	105	T28 T63	N34 N39	signal_cleavage: M1-R24 Signal Peptide: M1-C21, M1-K23, M1-C20 signal_cleavage: M1-V18	SPSCAN HMMER
24	7503605CD1	190	S27 S93 S139 S168 S173 T45		DDC / GAD / HDC / TyroDC pyridoxal-phosphate attachment site proteins BL00392: Y32-G41	SPSCAN BLIMPS_BLOCKS

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
25/ 1849820CB1/ 1992	1-253, 1-497, 1-659, 1-835, 67-218, 134-961, 167-557, 176-718, 186-451, 193-480, 256-409, 307-448, 332-553, 346-556, 347-565, 356-626, 380-679, 392-590, 403-595, 420-703, 436-588, 438-1043, 457-728, 462-757, 488-730, 488-1029, 571-865, 589-762, 615-799, 637-872, 639-1215, 641-1185, 651-850, 668-1213, 696-1336, 700-1215, 722-930, 745-1215, 747-1194, 770-1047, 780-1286, 783-1360, 817-1326, 859-1133, 862-1302, 953-1206, 1008-1585, 1013-1352, 1015-1570, 1016-1311, 1029-1214, 1051-1305, 1051-1556, 1054-1400, 1059-1352, 1065-1331, 1108-1658, 1199-1524, 1214-1445, 1215-1291, 1219-1485, 1245-1766, 1326-1646, 1358-1593, 1359-1938, 1370-1637, 1379-1648, 1379-1960, 1441-1890, 1453-1992, 1463-1962, 1466-1761, 1466-1956, 1466-1986, 1485-1897, 1494-1890, 1518-1963, 1519-1974, 1530-1772, 1541-1963, 1545-1992, 1547-1797, 1560-1992, 1649-1975, 1678-1962, 1702-1975, 1780-1975
26/ 70610307CB1/806	1-460, 37-195, 37-333, 37-558, 196-460, 334-460, 424-458, 461-558, 490-806
27/ 8137559CB1/1004	1-657, 98-193, 232-268, 334-657, 372-601, 372-858, 372-1004, 451-996, 523-1003
28/ 4801255CB1/1202	1-247, 1-542, 1-570, 1-611, 234-797, 378-850, 386-850, 475-1202
29/ 6160719CB1/ 1416	1-280, 1-492, 1-586, 1-655, 1-663, 1-711, 1-717, 3-466, 3-717, 3-742, 3-753, 3-845, 4-769, 10-646, 117-1014, 274-1025, 277-806, 343-1020, 474-717, 506-1158, 609-1261, 642-1093, 679-1265, 758-1402, 844-1416, 1032-1363, 1120-1343, 1120-1416, 1133-1381, 1141-1352, 1141-1359
30/ 3524602CB1/ 1707	1-886, 12-63, 12-89, 12-108, 12-185, 12-205, 12-211, 12-224, 12-225, 12-244, 12-282, 12-306, 12-342, 12-391, 12-432, 12-556, 12-618, 12-642, 20-151, 20-205, 20-303, 20-451, 24-219, 55-861, 63-865, 75-834, 77-494, 87-688, 105-976, 118-735, 133-844, 135-768, 137-442, 152-699, 155-921, 162-780, 163-1033, 164-720, 187-613, 192-1002, 192-1008, 210-1026, 215-827, 232-699, 248-543, 273-882, 290-890, 297-1000, 307-901, 309-900, 310-873, 313-1033, 325-899, 328-1038, 361-977, 373-636, 375-1104, 384-613, 428-713, 436-1023, 466-1067, 467-1067, 476-1369, 496-1360, 546-1369, 589-870, 592-1070, 597-1369, 599-1214, 609-924, 619-912, 706-1369, 711-1369, 725-1369, 751-1369, 787-1343, 793-1369, 822-1295, 830-1369, 843-1369, 894-1293, 972-1574, 1030-1633, 1035-1707, 1058-1570, 1059-1633, 1093-1369, 1186-1707, 1207-1500, 1225-1707, 1359-1633, 1373-1633, 1390-1656, 1421-1619

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
31/ 513718CB1/ 1436	1-214, 1-234, 1-462, 1-510, 1-538, 1-549, 1-582, 1-598, 1-600, 1-608, 1-609, 1-612, 4-602, 29-569, 29-611, 96-652, 137-843, 161-693, 207-498, 231-777, 238-698, 261-910, 271-631, 284-854, 286-855, 313-1025, 382-915, 442-1080, 546-1199, 580-1199, 593-1202, 593-1206, 621-840, 621-849, 621-870, 621-920, 621-1030, 621-1036, 621-1085, 621-1194, 622-1101, 623-1085, 629-1030, 632-1089, 635-1197, 660-882, 739-1377, 742-1436, 751-1280, 773-1236, 773-1244, 786-1335, 799-1436, 800-1244, 809-1197, 908-1314, 993-1218, 1034-1434, 1136-1436
32/ 896308CB1/ 1984	1-479, 70-724, 112-752, 205-752, 230-942, 373-1022, 374-1049, 384-1021, 390-1110, 437-1119, 437-1124, 504-1182, 512-1048, 525-1208, 527-1121, 544-1138, 579-1259, 598-1145, 660-1328, 660-1344, 701-1263, 704-1355, 721-1391, 741-1416, 780-1092, 793-1480, 803-1422, 829-1479, 846-1452, 868-1420, 869-1611, 883-1509, 888-1513, 963-1513, 1007-1549, 1010-1637, 1012-1581, 1028-1631, 1075-1627, 1121-1447, 1188-1624, 1209-1938, 1253-1934, 1289-1952, 1292-1821, 1300-1984, 1340-1984, 1347-1871, 1549-1984
33/2105862CB1/ 1131	1-465, 1-1131, 260-926, 515-798, 515-983, 515-1104, 515-1131, 572-1131, 843-970, 1072-1131
34/ 7939381CB1/ 2417	1-632, 17-491, 36-676, 46-346, 47-614, 47-675, 47-720, 50-340, 56-707, 67-687, 103-780, 103-793, 123-416, 178-660, 210-786, 215-458, 220-485, 253-505, 285-516, 285-895, 405-991, 406-831, 432-771, 502-695, 516-1226, 642-882, 679-1151, 711-986, 722-1271, 724-995, 735-1005, 735-1258, 754-1290, 757-1026, 793-1472, 838-1492, 849-1391, 851-1488, 857-1171, 857-1242, 857-1247, 863-1379, 893-1256, 913-1363, 922-1291, 924-1123, 924-1151, 926-1464, 928-1226, 940-1118, 977-1256, 977-1258, 989-1200, 994-1181, 1004-1501, 1089-1705, 1098-1368, 1104-1621, 1123-1358, 1154-1725, 1165-1427, 1165-1445, 1174-1654, 1174-1765, 1198-1771, 1211-1723, 1212-1525, 1225-1729, 1251-1869, 1265-1502, 1265-1798, 1275-1563, 1286-1877, 1289-1486, 1289-1849, 1291-1781, 1293-1515, 1300-1875, 1304-1552, 1306-1878, 1306-1945, 1309-1814, 1323-1933, 1333-2000, 1341-1921, 1343-1551, 1354-1960, 1365-1937, 1367-1590, 1373-1567, 1373-1621, 1392-1620, 1392-1624, 1408-1686, 1411-2034, 1418-1902, 1424-1683, 1426-1766, 1426-2068, 1434-1675, 1435-1896, 1436-2030, 1447-1527, 1470-2107, 1514-2033, 1518-2144, 1551-2179, 1563-2125, 1583-2105, 1602-2209, 1602-2258, 1609-2216, 1634-2251, 1639-2211, 1647-2202, 1666-2197, 1666-2259, 1670-2214, 1671-2279, 1671-2289, 1680-2184, 1686-2369, 1692-2134, 1705-2358, 1710-2259, 1739-2383, 1754-2017, 1754-2063,

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
34/ 7939381CB1/ 2417 (continued)	1754-2417, 1756-2388, 1757-2202, 1757-2346, 1758-2129, 1758-2313, 1758-2392, 1761-2158, 1777-2385, 1780-2360, 1783-2408, 1784-2176, 1784-2414, 1799-2358, 1801-2272, 1809-2056, 1809-2391, 1813-2417, 1829-2415, 1829-2417, 1836-2116, 1845-2100, 1845-2364, 1848-2095, 1858-2075, 1859-2417, 1862-2137, 1862-2143, 1863-2386, 1872-2114, 1875-2087, 1878-2411, 1880-2344, 1894-2364, 1899-2143, 1904-2161, 1904-2176, 1904-2197, 1916-2196, 1916-2197, 1916-2248, 1916-2398, 1919-2220, 1922-2417, 1925-2168, 1927-2348, 1932-2288, 1935-2326, 1938-2412, 1943-2398, 1948-2398, 1958-2189, 1966-2213, 1969-2281, 1978-2399, 1989-2149, 1991-2398, 1995-2415, 1997-2398, 1998-2397, 1998-2400, 2002-2398, 2004-2404, 2005-2399, 2016-2397
35/ 7487507CB1/ 2768	1-620, 1-622, 1-644, 1-647, 1-650, 4-650, 11-611, 52-650, 87-642, 112-650, 161-650, 169- 633, 180-457, 180-595, 180-650, 181-650, 184-650, 187-650, 189-460, 189-544, 189-567, 192-650, 198-418, 199-430, 200-1009, 202-507, 209-621, 219-650, 236-650, 247-648, 277- 650, 336-645, 336-662, 337-508, 366-662, 376-650, 395-650, 403-547, 416-650, 427-1134, 442-650, 443-499, 445-650, 472-635, 477-650, 499-562, 500-650, 508-650, 511-879, 511-880, 514-650, 516-662, 525-562, 528-650, 556-661, 561-640, 561-661, 561-662, 569-661, 572-650, 572-661, 576-650, 577-650, 582-661, 590-650, 604-650, 610-650, 617-650, 619-650, 619- 1232, 622-1248, 628-650, 632-1320, 636-1127, 650-1156, 650-1183, 655-1114, 662-765, 662- 767, 662-1200, 667-1287, 670-701, 692-1239, 693-1388, 703-1250, 706-1413, 708-1430, 716- 1470, 722-1261, 745-1263, 745-1336, 751-1154, 753-1360, 753-1363, 755-1144, 759-1413, 763-1258, 767-1247, 774-1360, 777-1438, 778-1316, 787-1461, 790-1305, 790-1464, 790-1470, 791-1468, 792-1385, 793-1434, 796-973, 805-1430, 807-1081, 822-1468, 824-1398, 831-1355, 834-1433, 844-1552, 866-1456, 869-1629, 876-1506, 877-1579, 880-1572, 883-1488, 890-1469, 897-1572, 898-1605, 901-1541, 907-1465, 910-1414, 919-1464, 923-1531, 929-1507, 944-1578, 947-1508, 951-1641, 952-1692, 958-1510, 958-1541, 975-1472, 979-1700, 989-1842, 995-1541, 1002-1541, 1003-1541, 1008-1541, 1015-1541, 1016-1661, 1018-1541, 1024-1541, 1025-1591, 1025-1666, 1025-1757, 1029-1431, 1037-1541, 1048-1527, 1057-1476, 1064-1541, 1088-1541, 1096-1776, 1123-1541, 1126-1541, 1148-1541, 1162-1541, 1164-1837, 1175-1837, 1189-1995, 1217-1963, 1217-2002, 1220-1662, 1220-1849, 1221-1541, 1236-2090, 1244-1831, 1246-1541, 1252-1908, 1262-1959, 1266-1706, 1267-1541, 1271-2178, 1273-1541, 1275-1462, 1275-1535, 1275-1541, 1276-1541, 1277-1541, 1277-1992, 1281-1541, 1283-1541, 1294-1928, 1295-1541,

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
35/ 7487507CB1/ 2768 (continued)	1297-1785, 1301-1541, 1323-1927, 1326-1872, 1333-1541, 1333-2006, 1344-1541, 1349-1948, 1366-1541, 1369-2097, 1372-2258, 1381-2096, 1382-2035, 1389-1947, 1392-1855, 1392-1984, 1394-2147, 1398-1541, 1401-1541, 1410-2020, 1411-1950, 1413-1541, 1416-2066, 1416-2119, 1421-1541, 1427-2060, 1428-2091, 1430-2058, 1450-2055, 1472-2175, 1480-2215, 1525-2141, 1533-2261, 1539-1673, 1540-1563, 1540-1601, 1540-1604, 1540-1609, 1540-1630, 1540-1666, 1540-1670, 1540-1690, 1540-1735, 1540-1741, 1540-1749, 1540-1765, 1540-1785, 1540-1799, 1540-1808, 1540-1875, 1540-2108, 1541-1580, 1541-1627, 1541-1742, 1541-2119, 1542-2112, 1548-2366, 1551-2269, 1555-2171, 1559-2132, 1571-2081, 1586-2128, 1586-2191, 1590-2269, 1591-2181, 1596-2235, 1598-2344, 1601-2194, 1601-2243, 1606-2258, 1612-2375, 1619-2178, 1626-2305, 1631-2162, 1639-2324, 1640-2301, 1641-2361, 1650-2148, 1657-2140, 1658-2291, 1675-2205, 1681-2289, 1681-2324, 1685-2432, 1688-2240, 1689-2339, 1690-2342, 1690-2389, 1696-2354, 1698-2277, 1702-2322, 1712-2366, 1715-2362, 1715-2383, 1718-2383, 1719-2214, 1728-2424, 1729-2220, 1733-2410, 1737-2344, 1741-2402, 1743-2384, 1747-2328, 1750-2372, 1754-2372, 1773-2498, 1775-2259, 1779-2288, 1786-2451, 1789-2509, 1803-2349, 1803-2422, 1803-2461, 1808-2083, 1812-2423, 1819-2260, 1820-2104, 1820-2526, 1821-2110, 1821-2141, 1822-2055, 1822-2473, 1825-2311, 1830-2321, 1832-2500, 1832-2616, 1833-2482, 1835-2549, 1836-2335, 1840-2305, 1841-2535, 1841-2535, 1852-2607, 1855-2279, 1863-2172, 1863-2481, 1865-2589, 1867-2484, 1870-2104, 1870-2630, 1877-2402, 1886-2081, 1886-2130, 1889-2224, 1898-2133, 1898-2151, 1910-2211, 1913-2360, 1914-2521, 1918-2423, 1929-2631, 1931-2353, 1931-2669, 1938-2612, 1942-2518, 1950-2631, 1952-2481, 1952-2617, 1954-2162, 1960-2514, 1963-2315, 1967-2377, 1981-2641, 1986-2560, 1998-2353, 2001-2400, 2003-2155, 2005-2248, 2005-2461, 2005-2576, 2009-2213, 2011-2768, 2018-2749, 2024-2379, 2026-2710, 2030-2573, 2031-2393, 2048-2503

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
36/ 1483931CB1/ 2096	1-259, 1-385, 1-407, 1-438, 1-500, 5-266, 16-624, 36-475, 58-659, 93-678, 124-262, 124-263, 145-374, 149-499, 232-1642, 456-1167, 528-990, 655-949, 782-1232, 782-1314, 782-1355, 783-909, 817-1167, 863-1484, 1067-1229, 1109-1406, 1121-1378, 1121-1382, 1121-1415, 1121-1436, 1121-1596, 1121-1608, 1121-1628, 1121-1633, 1121-1634, 1121-1648, 1121-1765, 1134-1366, 1150-1526, 1186-1444, 1246-1699, 1251-1683, 1270-1981, 1279-1988, 1292-1988, 1298-1847, 1300-1574, 1300-1865, 1309-1635, 1323-1840, 1330-1841, 1373-1863, 1398-2071, 1399-2095, 1405-2096, 1407-1908, 1411-1977, 1432-2066, 1441-2096, 1443-2096, 1447-2096, 1465-2055, 1511-1891, 1519-1860, 1541-1848, 1546-2096, 1554-2096, 1557-2050, 1579-2096, 1589-1845, 1591-1979, 1595-2083, 1613-1859, 1628-2045, 1642-2095, 1642-2096, 1644-2096, 1649-2096, 1657-2083, 1661-2096, 1663-2096, 1675-1938, 1675-2096, 1680-2096, 1686-2019, 1691-2096, 1692-2031, 1693-2096, 1701-1966, 1704-2096, 1709-2086, 1712-2096, 1715-2082, 1723-1977, 1731-2096, 1735-2096, 1738-2096, 1741-2082, 1751-1965, 1762-2096, 1774-2007, 1785-2096, 1787-2096, 1816-2083, 1818-2096
37/ 8175223CB1/ 1562	1-336, 1-382, 1-388, 1-483, 1-583, 1-628, 1-657, 1-674, 1-680, 1-698, 2-603, 9-197, 12-699, 23-216, 38-237, 63-292, 123-698, 342-698, 429-1109, 447-1045, 451-1125, 480-1151, 490-1207, 512-1226, 558-832, 562-687, 563-687, 612-1155, 652-832, 654-687, 661-687, 669-1164, 707-1066, 714-829, 725-832, 726-1419, 745-1394, 758-1039, 758-1058, 758-1062, 758-1068, 758-1078, 758-1083, 758-1171, 760-1083, 761-823, 763-1083, 764-1457, 766-1083, 774-832, 787-1527, 792-832, 794-1083, 794-1171, 796-1009, 800-1160, 801-1083, 801-1465, 802-1082, 802-1083, 803-1139, 804-1169, 804-1171, 806-1171, 813-1527, 818-1008, 820-1098, 821-1171, 823-1030, 833-1083, 835-1083, 839-1083, 841-1074, 843-1171, 848-1083, 853-1083, 873-1171, 879-1083, 883-1171, 883-1286, 884-1168, 891-1074, 892-1562, 894-1438, 896-1171, 899-1083, 902-1083, 902-1171, 903-1171, 914-1171, 916-1083, 936-1062, 940-1083, 941-1171, 946-1083, 946-1171, 952-1171, 957-1083, 957-1396, 969-1171, 971-1171, 989-1519, 1169-1229, 1170-1220, 1170-1240, 1170-1245, 1170-1265, 1170-1280, 1170-1296, 1170-1298, 1170-1302, 1170-1304, 1170-1308, 1170-1311, 1269-1311, 1269-1479

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
38/ 4173218CB1/ 1110	1-584, 1-718, 10-275, 36-260, 36-593, 36-637, 36-815, 37-838, 85-362, 93-335, 93-600, 130-302, 174-490, 293-490, 293-792, 312-454, 374-588, 374-646, 374-1003, 375-594, 382-553, 388-854, 458-486, 509-725, 509-1048, 512-1085, 529-978, 529-1011, 529-1049, 533-1110, 548-1083, 560-813, 560-1030, 563-957, 634-1094, 645-1107, 648-1110, 652-971, 663-885, 674-1097, 683-1101, 829-1052, 829-1095, 829-1096, 829-1097, 863-1085, 934-1102, 941-1102, 951-1097
39/ 5014679CB1/ 1412	1-256, 42-552, 69-709, 165-709, 219-670, 291-518, 294-552, 460-1125, 490-670, 583-1412, 588-1288, 592-1140, 663-1263, 683-1328, 729-1408, 802-1306, 815-1389, 818-1409, 867-1412, 944-1215, 952-1210, 958-1408
40/7487510CB1/ 1060	1-568, 1-641, 377-1053, 402-766, 572-1060, 606-885
41/ 2682619CB1/ 2268	1-536, 40-675, 65-233, 70-729, 333-361, 333-523, 333-540, 333-823, 369-956, 400-876, 427-843, 467-828, 467-834, 645-929, 645-1035, 661-1115, 666-1286, 678-881, 686-1050, 689-1221, 714-957, 733-1115, 787-1077, 794-1182, 844-1107, 844-1115, 848-1115, 863-1568, 864-1115, 867-1091, 887-1483, 889-1115, 902-1091, 919-1166, 999-1110, 1050-1190, 1170-1525, 1192-1911, 1216-1525, 1222-1620, 1248-1482, 1303-1706, 1342-1518, 1347-1626, 1417-1969, 1421-1724, 1434-1761, 1710-2198, 1725-1857, 1802-2268
42/ 4582105CB1/ 821	1-204, 1-319, 1-347, 1-371, 1-377, 1-381, 8-319, 17-350, 21-336, 23-319, 23-346, 23-362, 24-319, 24-380, 32-319, 34-319, 36-319, 36-350, 42-319, 49-319, 54-319, 54-362, 55-380, 57-216, 57-294, 57-317, 57-319, 57-362, 58-319, 59-319, 61-319, 61-349, 61-350, 62-319, 63-291, 63-319, 63-380, 64-319, 67-319, 76-294, 77-278, 77-380, 78-319, 78-339, 78-349, 78-362, 79-319, 79-346, 79-380, 80-319, 80-326, 81-319, 81-346, 81-348, 81-362, 82-319, 82-331, 82-337, 82-354, 84-319, 84-380, 85-319, 86-319, 86-381, 87-319, 87-335, 88-319, 88-336, 96-352, 96-355, 96-381, 97-379, 97-381, 99-362, 100-384, 102-319, 103-319, 104-350, 107-362, 109-319, 112-319, 112-325, 113-380, 113-471, 114-319, 114-741, 115-319, 115-380, 118-319, 119-319, 121-319, 122-319, 126-748, 126-749, 129-319, 129-380, 130-319, 130-362, 131-319, 144-319, 151-381, 154-381, 156-319, 160-294, 247-444, 282-817, 288-535, 295-534, 295-807, 295-817, 297-742, 332-821, 354-821, 358-584, 372-764, 372-787, 372-800, 372-821, 373-655, 373-670, 373-694, 373-699, 373-767, 373-779, 373-785, 373-786, 373-789, 373-792, 373-798, 373-800, 373-801, 373-802, 373-803, 373-811, 373-815, 373-821, 375-767,

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
42/ 4582105CB1/ 821 (continued)	375-787, 375-793, 376-806, 376-813, 376-817, 376-821, 377-768, 377-802, 377-803, 378-764, 382-815, 387-821, 391-676, 391-684, 391-790, 391-796, 391-799, 398-678, 398-768, 398-778, 398-797, 398-801, 398-807, 398-815, 416-815, 418-821, 420-691, 420-763, 427-814, 435-821, 436-821, 445-620, 450-789, 454-710, 462-691, 462-802, 466-764, 471-754, 477-748, 483-726, 487-728, 490-751, 500-817, 501-814, 504-790, 516-821, 531-821, 547-787, 547-821, 560-821, 564-821, 577-779, 581-821, 616-816, 616-817, 649-817, 694-817
43/931619CB1/ 1325	1-338, 1-368, 170-485, 254-808, 254-866, 360-761, 360-1019, 580-1244, 697-1194, 716-1234, 725-961, 729-974, 734-1181, 736-1099, 741-935, 776-959, 798-1320, 810-1325, 825-1235, 853-1296, 860-1303, 932-1308, 962-1209, 1023-1307, 1152-1307
44/2155025CB1/ 2030	1-2030, 1298-1456, 1353-1431, 1353-1974, 1378-1631, 1456-1477, 1490-1974, 1651-1978, 1710-1977, 1746-1974, 1869-1974
45/ 7640495CB1/ 1307	1-629, 50-597, 85-585, 85-672, 85-743, 85-791, 86-526, 86-537, 86-570, 86-711, 87-440, 87-499, 113-379, 113-513, 113-549, 115-542, 134-659, 134-733, 149-694, 157-484, 159-774, 169-407, 169-544, 207-879, 229-594, 263-836, 306-576, 336-977, 379-627, 379-896, 379-917, 420-670, 449-672, 449-971, 457-1007, 504-1166, 518-775, 552-831, 555-968, 581-828, 581- 970, 602-1307, 687-943, 699-971, 700-1116, 700-1165, 700-1176, 748-988, 802-968, 810- 1027, 819-1083, 928-1307, 977-1257, 988-1307, 996-1280, 1072-1307, 1174-1307, 1192-1307, 1285-1307
46/ 5960119CB1/ 768	1-105, 1-591, 469-768, 734-768
47/ 7500143CB1/ 782	1-151, 20-691, 92-369, 92-749, 115-744, 211-525, 211-528, 251-654, 251-675, 254-705, 258- 705, 266-692, 271-705, 272-705, 279-695, 284-547, 284-705, 310-575, 313-705, 318-695, 324-691, 326-705, 333-586, 350-691, 365-782, 396-705, 425-692, 504-689, 504-719, 508-705, 585-691, 620-710, 630-705

Table 4 (cont.)

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
48/ 7503605CB1/ 893	1-724, 83-226, 173-351, 195-442, 203-441, 203-724, 204-649, 206-714, 239-803, 261-729, 265-491, 279-694, 280-562, 280-577, 280-606, 280-674, 280-686, 280-692, 280-693, 280-696, 280-699, 280-705, 280-707, 280-708, 280-709, 280-710, 280-718, 280-722, 281-366, 281-736, 282-674, 282-694, 282-700, 284-675, 284-709, 284-710, 284-737, 285-670, 288-601, 288-670, 288-707, 288-713, 288-731, 288-733, 288-737, 288-746, 289-722, 294-735, 298-583, 298-697, 298-703, 298-706, 299-591, 300-760, 305-585, 305-675, 305-704, 305-708, 305-722, 315-704, 323-722, 325-784, 327-598, 327-670, 334-721, 336-560, 341-436, 342-740, 352-527, 352-652, 352-682, 352-706, 352-721, 352-722, 352-736, 352-741, 352-759, 352-763, 352-782, 352-783, 352-784, 352-787, 352-789, 352-790, 352-794, 354-582, 356-761, 357-696, 360-782, 361-617, 361-715, 369-598, 369-709, 369-733, 373-671, 378-661, 384-654, 390-633, 394-635, 394-784, 396-629, 397-658, 408-742, 411-697, 417-550, 423-776, 438-739, 454-721, 455-694, 467-744, 471-752, 488-748, 506-749, 509-721, 514-722, 522-761, 523-736, 524-753, 532-893, 539-738, 542-735, 545-760, 555-732, 555-734, 576-747, 580-742, 629-721

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
25	1849820CB1	PLACFEC01
27	8137559CB1	PLACFER01
28	4801255CB1	MYEPUNT01
29	6160719CB1	MONOTXN05
30	3524602CB1	BRACDIK08
31	513718CB1	MPHGNOT03
32	896308CB1	TESTTUT02
33	2105862CB1	BRSTTMT01
34	7939381CB1	HEAONOC01
35	7487507CB1	THYRDIE01
36	1483931CB1	BRAINOR03
37	8175223CB1	PLACNOB01
38	4173218CB1	BLADTUT05
39	5014679CB1	COLNUCT03
40	7487510CB1	COLENOR03
41	2682619CB1	UTRCDIE01
42	4582105CB1	LUNGNOT02
43	931619CB1	LUNGNOT09
44	2155025CB1	BRAINOT09
45	7640495CB1	SCORNON02
46	5960119CB1	TESTTUE02
47	7500143CB1	LUNGTUT03
48	7503605CB1	LATRTUT02

Table 6

Library	Vector	Library Description
BLADTUT05	pINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder. Patient history included lung neoplasm and tobacco abuse in remission. Family history included malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
BRACDIK08	PSPORT1	This amplified and normalized library was constructed using RNA isolated from diseased corpus callosum 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).
BRAINOR03	PBK-CMV	This random primed library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from brain tissue removed from a Caucasian male fetus (donor A) who was stillborn with a hypoplastic left heart at 23 weeks' gestation and from brain tissue removed from a Caucasian male fetus (donor B), who died at 23 weeks' gestation from premature birth. Serologies were negative for both donors and family history for donor B included diabetes in the mother.
BRAINOT09	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
BRSTTMT01	pINCY	Library was constructed using RNA isolated from breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated recurrent grade 4, nuclear grade 3, ductal carcinoma. Angiolymphatic space invasion was identified. Left breast needle biopsy indicated grade 4 ductal adenocarcinoma. Paraffin embedded tissue was estrogen positive. Patient history included breast cancer and deficiency anemia. Family history included cervical cancer.
COLENOR03	PCDNA2.1	Library was constructed using RNA isolated from colon epithelium tissue removed from a 13-year-old Caucasian female who died from a motor vehicle accident.

Table 6 (cont.)

Library	Vector	Library Description
COLNUCT03	pINCY	Library was constructed using RNA isolated from diseased colon tissue obtained from a 69-year-old Caucasian male during a partial colon excision with ileostomy. Pathology indicated severely active idiopathic inflammatory bowel disease most consistent with chronic ulcerative colitis. Patient history included benign neoplasm of the colon. Previous surgeries included cholecystectomy, spinal canal exploration, partial glossectomy, radical cystectomy, and bladder operation. Family history included cerebrovascular disease and benign hypertension.
HEAONOC01	PSPORT1	This large size fractionated library was constructed using RNA isolated from the aorta of a 39-year-old Caucasian male, who died from a gunshot wound. Serology was positive for cytomegalovirus (CMV). Patient history included tobacco abuse (one pack of cigarettes per day for 25 years), and occasionally cocaine, marijuana, and alcohol use.
LATRTUT02	pINCY	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
LUNGNOT02	PBLUESCRIPT	Library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male, who died of a subarachnoid hemorrhage.
LUNGNOT09	pINCY	Library was constructed using RNA isolated from the lung tissue of a 23-week-old Caucasian male fetus. The pregnancy was terminated following a diagnosis by ultrasound of infantile polycystic kidney disease.
LUNGFTUT03	PSPORT1	Library was constructed using RNA isolated from lung tumor tissue removed from the left lower lobe of a 69-year-old Caucasian male during segmental lung resection. Pathology indicated residual grade 3 invasive squamous cell carcinoma. Patient history included acute myocardial infarction, prostatic hyperplasia, malignant skin neoplasm, and tobacco use.

Table 6 (cont.)

Library	Vector	Library Description
MONOTXN05	pINCY	This normalized treated monocyte cell tissue library was constructed from 1.03 million independent clones from a monocyte tissue library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
MPHGNOT03	PBLUESCRIPT	Library was constructed using RNA isolated from plastic adherent mononuclear cells isolated from buffy coat units obtained from unrelated male and female donors.
MYEPUNT01	pINCY	Library was constructed RNA isolated from an untreated K-562 cell line, derived from chronic myelogenous leukemia precursor cells obtained from a 53-year-old female.
PLACFEC01	pINCY	This large size-fractionated library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time). Serology was positive for anti-CMV and remaining serologies were negative. Family history included multiple pregnancies and live births, and an abortion in the mother.
PLACFER01	pINCY	The library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time). Serology was positive for anti-CMV. Family history included multiple pregnancies and live births, and an abortion.
PLACNOB01	PBLUESCRIPT	Library was constructed using RNA isolated from placenta.
SCORNON02	PSPORT1	This normalized spinal cord library was constructed from 3.24M independent clones from the a spinal cord tissue library. RNA was isolated from the spinal cord tissue removed from a 71-year-old Caucasian male who died from respiratory arrest. Patient history included myocardial infarction, gangrene, and end stage renal disease. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).

Table 6 (cont.)

Library	Vector	Library Description
TESTTUE02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma forming a largely necrotic mass involving the entire testicle. Rare foci of residual testicle showed intralobular germ cell neoplasia and tumor was identified at the spermatic cord margin. The patient presented with backache. Patient history included tobacco use. Previous surgeries included a needle biopsy of testis. Patient medications included Colace and antacids.
TESTTUT02	pINCY	Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
THYRDIE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from diseased thyroid tissue removed from a 22-year-old Caucasian female during closed thyroid biopsy, partial thyroidectomy, and regional lymph node excision. Pathology indicated adenomatous hyperplasia. The patient presented with malignant neoplasm of the thyroid. Patient history included normal delivery; alcohol abuse, and tobacco abuse. Previous surgeries included myringotomy. Patient medications included an unspecified type of birth control pills. Family history included hyperlipidemia and depressive disorder in the mother; and benign hypertension, congestive heart failure, and chronic leukemia in the grandparent(s).
UTRCDIE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from uterine cervix tissue removed from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the matched tumor tissue indicated intramural uterine leiomyoma. Patient history included hypothyroidism, pelvic floor relaxation, paraplegia, and self catheterization. Previous surgeries included a normal delivery, a laminectomy, and a rhinoplasty. Patient medications included Synthroid. Family history included benign hypertension in the father; and type II diabetes and hyperlipidemia in the mother.

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.	

WO 02/062841

PCT/US02/02616

<110> INCYTE GENOMICS, INC.
 TANG, Y. Tom
 YUE, Henry
 GANDHI, Ameena R.
 YAO, Monique G.
 WARREN, Bridget A.
 DING, Li
 DUGGAN, Brendan M.
 XU, Yuming
 YANG, Junming
 THANGAVELU, Kavitha
 LAL, Preeti G.
 HONCHELL, Cynthia D.
 WALIA, Narinder K.
 LEE, Sally
 LEE, Ernestine A.
 RICHARDSON, Thomas
 BAUGHN, Mariah R.
 ELLIOTT, Vicki S.

<120> SECRETED PROTEINS

<130> PI-0359 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/266,195; 60/267,924; 60/268,112; 60/267,816; 60/271,639
 60/317,818; 60/343,553

<151> 2001-02-02; 2001-02-08; 2001-02-09; 2001-02-09; 2001-02-26;
 2001-09-07; 2001-12-21

<160> 48

<170> PERL Program

<210> 1

<211> 496

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1849820CD1

<400> 1

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

WO 02/062841

PCT/US02/02616

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

Ser

<210> 2
 <211> 185
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 70610307CD1

<400> 2
 Met Leu Pro Pro Met Ala Leu Pro Ser Val Ser Trp Leu Leu Phe

WO 02/062841

PCT/US02/02616

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

<210> 3
 <211> 73
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 8137559CD1

<400> 3
Met Pro Met Lys Lys Ser Trp Met Pro Lys Thr Cys His Ile Phe
1 5 10 15
Leu Leu Leu Ala Ala Phe Phe Gln Asn Arg Leu Thr Asp Pro Phe
20 25 30
Pro Cys Ser Ile Cys Gly Glu Cys Gln Tyr Gly Phe Ser Phe Pro
35 40 45
Ser Phe Phe Phe Leu Ile Ser Ser Ala Pro Val Lys Ala Phe Pro
50 55 60
Leu Ser Thr Asp His Ser Cys Gly Leu Cys Trp Ala Val
65 70

<210> 4
 <211> 75
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4801255CD1

<400> 4
Met Arg Glu Asn Val Ser Leu Leu Ala Ala Leu Val Ile Ala Thr
1 5 10 15
Cys Leu Pro Gln Asn Arg Ser Ser Met Asp Glu Gln Thr Glu Lys
20 25 30
Cys Val Arg Arg Leu Val Thr Glu Thr Asp Gln Gly Arg His Leu
35 40 45

WO 02/062841

PCT/US02/02616

Lys Lys His Leu Ser Cys Asp Leu Leu Ala Gly Thr Leu Lys Ser
 50 55 60
 Cys Tyr Ala Asn Gly Pro Ala Thr Leu Ala Gly Arg Asn Tyr Cys
 65 70 75

<210> 5
 <211> 96
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6160719CD1

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

<210> 6
 <211> 87
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3524602CD1

<400> 6
 Met Ser Met Trp Asp Val Gln Ser Arg Gly Thr Ser Tyr Gln Glu
 1 5 10 15
 Trp Pro Phe Ile Ser Thr Val Pro Thr Met Ala Lys Ser Cys Gln
 20 25 30
 Ser Leu Met Leu Gly Ser His Leu Leu Ser Asn Pro Pro Gln Thr
 35 40 45
 Gln Lys Pro Gly Thr Leu Ile Leu His Gln Pro Glu Val Met Ser
 50 55 60
 Val Asp Asp Met Asn Asp Ser Ile Gly Gln Val Gln Leu Arg Val
 65 70 75
 Arg Gln Ile Leu Ala Asn His Ser Asp Gly Glu Gly
 80 85

<210> 7
 <211> 96
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 513718CD1

WO 02/062841

PCT/US02/02616

<400> 7

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

<210> 8

<211> 101

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 896308CD1

<400> 8

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

<210> 9

<211> 93

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2105862CD1

<400> 9

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

WO 02/062841

PCT/US02/02616

Phe Lys Lys

<210> 10

<211> 466

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7939381CD1

<400> 10

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

Ser	Leu	Lys	Lys	Lys	Leu	Gln	Gln	Lys	Gln	Leu	Leu	Ile	Leu	Gln
				380					385					390
Leu	Leu	Glu	Lys	Ile	Ser	Phe	Leu	Glu	Gly	Glu	Asn	Asn	Glu	Leu
				395					400					405
Gln	Ser	Arg	Leu	Asp	Tyr	Leu	Thr	Glu	Thr	Gln	Ala	Lys	Thr	Glu
				410					415					420
Val	Glu	Thr	Arg	Glu	Ile	Gly	Val	Gly	Cys	Asp	Leu	Leu	Pro	Ser
				425					430					435
Gln	Thr	Gly	Arg	Thr	Arg	Glu	Ile	Val	Met	Pro	Ser	Arg	Asn	Tyr
				440					445					450
Thr	Pro	Tyr	Thr	Arg	Val	Leu	Glu	Leu	Thr	Met	Lys	Lys	Thr	Leu
				455					460					465

Thr

<210> 11
 <211> 730
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7487507CD1

<400> 11

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

Phe	Leu	Lys	Ala	Gln	Leu	Leu	Cys	Thr	Gln	Pro	Gly	Gln	Leu	Pro
				290					295					300
Phe	Asn	Val	Ile	Arg	His	Ala	Val	Leu	Leu	Pro	Ala	Asp	Ser	Pro
				305					310					315
Thr	Ala	Pro	His	Ile	Tyr	Ala	Val	Phe	Thr	Ser	Gln	Trp	Gln	Val
				320					325					330
Gly	Gly	Thr	Arg	Ser	Ser	Ala	Val	Cys	Ala	Phe	Ser	Leu	Leu	Asp
				335					340					345
Ile	Glu	Arg	Val	Phe	Lys	Gly	Lys	Tyr	Lys	Glu	Leu	Asn	Lys	Glu
				350					355					360
Thr	Ser	Arg	Trp	Thr	Thr	Tyr	Arg	Gly	Pro	Glu	Thr	Asn	Pro	Arg
				365					370					375
Pro	Gly	Ser	Cys	Ser	Val	Gly	Pro	Ser	Ser	Asp	Lys	Ala	Leu	Thr
				380					385					390
Phe	Met	Lys	Asp	His	Phe	Leu	Met	Asp	Glu	Gln	Val	Val	Gly	Thr
				395					400					405
Pro	Leu	Leu	Val	Lys	Ser	Gly	Val	Glu	Tyr	Thr	Arg	Leu	Ala	Val
				410					415					420
Glu	Thr	Ala	Gln	Gly	Leu	Asp	Gly	His	Ser	His	Leu	Val	Met	Tyr
				425					430					435
Leu	Gly	Thr	Thr	Thr	Gly	Ser	Leu	His	Lys	Ala	Val	Gly	Ala	Val
				440					445					450
Phe	Val	Gly	Phe	Ser	Gly	Gly	Val	Trp	Arg	Val	Pro	Arg	Ala	Asn
				455					460					465
Cys	Ser	Val	Tyr	Glu	Ser	Cys	Val	Asp	Cys	Val	Leu	Ala	Arg	Asp
				470					475					480
Pro	His	Cys	Ala	Trp	Asp	Pro	Glu	Ser	Arg	Thr	Cys	Cys	Leu	Leu
				485					490					495
Ser	Ala	Pro	Asn	Leu	Asn	Ser	Trp	Lys	Gln	Asp	Met	Glu	Arg	Gly
				500					505					510
Asn	Pro	Glu	Trp	Ala	Cys	Ala	Ser	Gly	Pro	Met	Ser	Arg	Ser	Leu
				515					520					525
Arg	Pro	Gln	Ser	Arg	Pro	Gln	Ile	Ile	Lys	Glu	Val	Leu	Ala	Val
				530					535					540
Pro	Asn	Ser	Ile	Leu	Glu	Leu	Pro	Cys	Pro	His	Leu	Ser	Ala	Leu
				545					550					555
Ala	Ser	Tyr	Tyr	Trp	Ser	His	Gly	Pro	Ala	Ala	Val	Pro	Glu	Ala
				560					565					570
Ser	Ser	Thr	Val	Tyr	Asn	Gly	Ser	Leu	Leu	Leu	Ile	Val	Gln	Asp
				575					580					585
Gly	Val	Gly	Gly	Leu	Tyr	Gln	Cys	Trp	Ala	Thr	Glu	Asn	Gly	Phe
				590					595					600
Ser	Tyr	Pro	Val	Ile	Ser	Tyr	Trp	Val	Asp	Ser	Gln	Asp	Gln	Thr
				605					610					615
Leu	Ala	Leu	Asp	Pro	Glu	Leu	Ala	Gly	Ile	Pro	Arg	Glu	His	Val
				620					625					630
Lys	Val	Pro	Leu	Thr	Arg	Val	Ser	Gly	Gly	Ala	Ala	Leu	Ala	Ala
				635					640					645
Gln	Gln	Ser	Tyr	Trp	Pro	His	Phe	Val	Thr	Val	Thr	Val	Leu	Phe
				650					655					660
Ala	Leu	Val	Leu	Ser	Gly	Ala	Leu	Ile	Ile	Leu	Val	Ala	Ser	Pro
				665					670					675
Leu	Arg	Ala	Leu	Arg	Ala	Arg	Gly	Lys	Val	Gln	Gly	Cys	Glu	Thr
				680					685					690
Leu	Arg	Pro	Gly	Glu	Lys	Ala	Pro	Leu	Ser	Arg	Glu	Gln	His	Leu
				695					700					705
Gln	Ser	Pro	Lys	Glu	Cys	Arg	Thr	Ser	Ala	Ser	Asp	Val	Asp	Ala
				710					715					720
Asp	Asn	Asn	Cys	Leu	Gly	Thr	Glu	Val	Ala					
				725					730					

<210> 12

<211> 575

WO 02/062841

PCT/US02/02616

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1483931CD1

<400> 12

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

WO 02/062841

PCT/US02/02616

	410		415		420
Val Val Ser Thr	Ala Thr Gly Asn Phe	Leu Asn Arg Leu Val	Pro		
	425		430		435
Ala Gly Thr Trp	Lys Pro Gly Thr Ala	Gly Asn Ile Ser His	Val		
	440		445		450
Ala Glu Gly Asp	Lys Pro Gln His Arg	Ala Thr Ile Cys Leu	Ser		
	455		460		465
Lys Met Asp Ile	Ala Trp Val Ile Leu	Ala Ile Ser Val Pro	Ile		
	470		475		480
Ser Ser Cys Ser	Val Leu Leu Thr Val	Cys Cys Met Lys Arg	Lys		
	485		490		495
Lys Lys Thr Ala	Asn Pro Glu Asn Asn	Leu Ser Tyr Trp Asn	Asn		
	500		505		510
Thr Ile Thr Met	Asp Tyr Phe Asn Arg	His Ala Val Glu Leu	Pro		
	515		520		525
Arg Glu Ile Gln	Ser Leu Glu Thr Ser	Glu Asp Gln Leu Ser	Glu		
	530		535		540
Pro Arg Ser Pro	Ala Asn Gly Asp Tyr	Arg Asp Thr Gly Met	Val		
	545		550		555
Leu Val Asn Pro	Phe Cys Gln Glu Thr	Leu Phe Val Gly Asn	Asp		
	560		565		570
Gln Val Ser Glu	Ile				
	575				

<210> 13

<211> 335

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8175223CD1

<400> 13

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

WO 02/062841

PCT/US02/02616

	215		220		225
Ser Arg Ser Asp	Pro Val Thr Leu Asn	Leu Leu His Gly	Pro Asp		
	230		235		240
Leu Pro Arg Ile	Phe Pro Ser Val Thr	Ser Tyr Tyr Ser	Gly Glu		
	245		250		255
Asn Leu Asp Leu	Ser Cys Phe Ala Asn	Ser Asn Pro Pro	Ala Gln		
	260		265		270
Tyr Ser Trp Thr	Ile Asn Gly Lys Phe	Gln Leu Ser Gly	Gln Lys		
	275		280		285
Leu Phe Ile Pro	Gln Ile Thr Pro Lys	His Asn Gly Leu	Tyr Ala		
	290		295		300
Cys Ser Ala Arg	Asn Ser Ala Thr Gly	Glu Glu Ser Ser	Thr Ser		
	305		310		315
Leu Thr Ile Arg	Val Ile Ala Pro Pro	Gly Leu Gly Thr	Phe Ala		
	320		325		330
Phe Asn Asn Pro	Thr				
	335				

<210> 14

<211> 120

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4173218CD1

<400> 14

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

<210> 15

<211> 103

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5014679CD1

<400> 15

Met Phe Arg Phe Phe	Ser Ser Leu Pro Gly	Leu Cys Arg Cys Cys
1	5	10 15
Cys Ser Leu Cys Leu	Glu Met Phe Ile Trp	Pro Glu Ser Lys Ser
	20	25 30
Leu Ser Lys Leu Leu	Leu Phe Ser Met Thr	Glu Leu Gln Leu Phe
	35	40 45
Val Cys Pro Pro Ser	Pro Leu Glu Tyr Met	Val Pro Lys Gly Lys

WO 02/062841

PCT/US02/02616

Pro	His	Ala	Ile	His	Ser	Pro	Pro	Asn	His	Pro	Gln	Gln	Leu	Val
				50					55					60
Gly	Lys	His	Glu	Thr	Leu	Leu	Leu	Asn	Asn	Val	Phe	Pro	Leu	Ala
				65					70					75
Gln	Gly	Leu	Gln	Ser	Gln	Arg	Leu	Thr	Arg	Gly	Lys	Asp		
				80					85					90
				95					100					

<210> 16
 <211> 170
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7487510CD1

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

<210> 17
 <211> 617
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2682619CD1

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

His	Leu	Leu	Glu	Arg	Ser	Trp	Tyr	Ala	Val	Thr	Glu	Thr	Cys	Leu
				80					85					90
Ala	Phe	Thr	Val	Phe	Arg	Asp	Asp	Phe	Ser	Pro	Arg	Phe	Val	Ala
				95					100					105
Leu	Phe	Thr	Leu	Leu	Leu	Phe	Leu	Lys	Cys	Phe	His	Trp	Leu	Ala
				110					115					120
Glu	Asp	Arg	Val	Asp	Phe	Met	Glu	Arg	Ser	Pro	Asn	Ile	Ser	Trp
				125					130					135
Leu	Phe	His	Cys	Arg	Ile	Val	Ser	Leu	Met	Phe	Leu	Leu	Gly	Ile
				140					145					150
Leu	Asp	Phe	Leu	Phe	Val	Ser	His	Ala	Tyr	His	Ser	Ile	Leu	Thr
				155					160					165
Arg	Gly	Ala	Ser	Val	Gln	Leu	Val	Phe	Gly	Phe	Glu	Tyr	Ala	Ile
				170					175					180
Leu	Met	Thr	Met	Val	Leu	Thr	Ile	Phe	Ile	Lys	Tyr	Val	Leu	His
				185					190					195
Ser	Val	Asp	Leu	Gln	Ser	Glu	Asn	Pro	Trp	Asp	Asn	Lys	Ala	Val
				200					205					210
Tyr	Met	Leu	Tyr	Thr	Glu	Leu	Phe	Thr	Gly	Phe	Ile	Lys	Val	Leu
				215					220					225
Leu	Tyr	Met	Ala	Phe	Met	Thr	Ile	Met	Ile	Lys	Val	His	Thr	Phe
				230					235					240
Pro	Leu	Phe	Ala	Ile	Arg	Pro	Met	Tyr	Leu	Ala	Met	Arg	Gln	Phe
				245					250					255
Lys	Lys	Ala	Val	Thr	Asp	Ala	Ile	Met	Ser	Arg	Arg	Ala	Ile	Arg
				260					265					270
Asn	Met	Asn	Thr	Leu	Tyr	Pro	Asp	Ala	Thr	Pro	Glu	Glu	Leu	Gln
				275					280					285
Ala	Met	Asp	Asn	Val	Cys	Ile	Ile	Cys	Arg	Glu	Glu	Met	Val	Thr
				290					295					300
Gly	Ala	Lys	Arg	Leu	Pro	Cys	Asn	His	Ile	Phe	His	Thr	Ser	Cys
				305					310					315
Leu	Arg	Ser	Trp	Phe	Gln	Arg	Gln	Gln	Thr	Cys	Pro	Thr	Cys	Arg
				320					325					330
Met	Asp	Val	Leu	Arg	Ala	Ser	Leu	Pro	Ala	Gln	Ser	Pro	Pro	Pro
				335					340					345
Pro	Glu	Pro	Ala	Asp	Gln	Gly	Pro	Pro	Pro	Ala	Pro	His	Pro	Pro
				350					355					360
Pro	Leu	Leu	Pro	Gln	Pro	Pro	Asn	Phe	Pro	Gln	Gly	Leu	Leu	Pro
				365					370					375
Pro	Phe	Pro	Pro	Gly	Met	Phe	Pro	Leu	Trp	Pro	Pro	Met	Gly	Pro
				380					385					390
Phe	Pro	Pro	Val	Pro	Pro	Pro	Pro	Ser	Ser	Gly	Glu	Ala	Val	Ala
				395					400					405
Pro	Pro	Ser	Thr	Ser	Ala	Ala	Ala	Leu	Ser	Arg	Pro	Ser	Gly	Ala
				410					415					420
Ala	Thr	Thr	Thr	Ala	Ala	Gly	Thr	Ser	Ala	Thr	Ala	Ala	Ser	Ala
				425					430					435
Thr	Ala	Ser	Gly	Pro	Gly	Ser	Gly	Ser	Ala	Pro	Glu	Ala	Gly	Pro
				440					445					450
Ala	Pro	Gly	Phe	Pro	Phe	Pro	Pro	Pro	Trp	Met	Gly	Met	Pro	Leu
				455					460					465
Pro	Pro	Pro	Phe	Ala	Phe	Pro	Pro	Met	Pro	Val	Pro	Pro	Ala	Gly
				470					475					480
Phe	Ala	Gly	Leu	Thr	Pro	Glu	Glu	Leu	Arg	Ala	Leu	Glu	Gly	His
				485					490					495
Glu	Arg	Gln	His	Leu	Glu	Ala	Arg	Leu	Gln	Ser	Leu	Arg	Asn	Ile
				500					505					510
His	Thr	Leu	Leu	Asp	Ala	Ala	Met	Leu	Gln	Ile	Asn	Gln	Tyr	Leu
				515					520					525
Thr	Val	Leu	Ala	Ser	Leu	Gly	Pro	Pro	Arg	Pro	Ala	Thr	Ser	Val
				530					535					540
Asn	Ser	Thr	Glu	Glu	Thr	Ala	Thr	Thr	Val	Val	Ala	Ala	Ala	Ser

WO 02/062841

PCT/US02/02616

	545		550		555
Ser Thr Ser Ile	Pro Ser Ser Glu Ala	Thr Thr Pro Thr	Pro Gly		
	560		565		570
Ala Ser Pro Pro	Ala Pro Glu Met Glu	Arg Pro Pro Ala	Pro Glu		
	575		580		585
Ser Val Gly Thr	Glu Glu Met Pro Glu	Asp Gly Glu Pro	Asp Ala		
	590		595		600
Ala Glu Leu Arg	Arg Arg Arg Leu Gln	Lys Leu Glu Ser	Pro Val		
	605		610		615
Ala His					

<210> 18
 <211> 221
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4582105CD1

<400> 18

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

<210> 19
 <211> 329
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 931619CD1

<400> 19

WO 02/062841

PCT/US02/02616

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

<210> 20

<211> 640

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2155025CD1

<400> 20

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

WO 02/062841

PCT/US02/02616

Arg	Phe	Lys	Val	Val	Asp	Glu	Phe	Gly	Asn	Arg	Phe	Asp	Val	Asn
				65					70					75
Asn	Cys	Ser	Ile	Cys	Tyr	His	Trp	Val	Thr	Ser	Arg	Pro	Gln	Glu
				80					85					90
Pro	Ala	Val	Phe	Ser	Ala	Asp	Tyr	Arg	Gly	Cys	His	Val	Leu	Glu
				95					100					105
Lys	Val	Gly	Asp	Gly	Arg	Phe	His	Leu	Arg	Val	Phe	Met	Glu	Ala
				110					115					120
Val	Leu	Pro	Asn	Gly	Arg	Val	Asp	Val	Ala	Gln	Asp	Ala	Thr	Leu
				125					130					135
Ile	Cys	Pro	Lys	Pro	Asp	Pro	Ser	Arg	Thr	Leu	Asp	Ser	Gln	Leu
				140					145					150
Ala	Pro	Pro	Ala	Met	Phe	Ser	Val	Ser	Ile	Pro	Gln	Thr	Leu	Ser
				155					160					165
Phe	Leu	Pro	Thr	Ser	Gly	His	Thr	Ser	Gln	Gly	Ser	Gly	His	Ala
				170					175					180
Phe	Pro	Ser	Pro	Leu	Asp	Pro	Gly	His	Ser	Ser	Val	His	Pro	Thr
				185					190					195
Pro	Ala	Leu	Pro	Ser	Pro	Gly	Pro	Gly	Pro	Thr	Leu	Ala	Thr	Leu
				200					205					210
Ala	Gln	Pro	His	Trp	Gly	Thr	Leu	Glu	His	Trp	Asp	Val	Asn	Lys
				215					220					225
Arg	Asp	Tyr	Ile	Gly	Thr	His	Leu	Ser	Gln	Glu	Gln	Cys	Gln	Val
				230					235					240
Ala	Ser	Gly	His	Leu	Pro	Cys	Ile	Val	Arg	Arg	Thr	Ser	Lys	Glu
				245					250					255
Ala	Cys	Gln	Gln	Ala	Gly	Cys	Cys	Tyr	Asp	Asn	Thr	Arg	Glu	Val
				260					265					270
Pro	Cys	Tyr	Tyr	Gly	Asn	Thr	Ala	Thr	Val	Gln	Cys	Phe	Arg	Asp
				275					280					285
Gly	Tyr	Phe	Val	Leu	Val	Val	Ser	Gln	Glu	Met	Ala	Leu	Thr	His
				290					295					300
Arg	Ile	Thr	Leu	Ala	Asn	Ile	His	Leu	Ala	Tyr	Ala	Pro	Thr	Ser
				305					310					315
Cys	Ser	Pro	Thr	Gln	His	Thr	Glu	Ala	Phe	Val	Val	Phe	Tyr	Phe
				320					325					330
Pro	Leu	Thr	His	Cys	Gly	Thr	Thr	Met	Gln	Val	Ala	Gly	Asp	Gln
				335					340					345
Leu	Ile	Tyr	Glu	Asn	Trp	Leu	Val	Ser	Gly	Ile	His	Ile	Gln	Lys
				350					355					360
Gly	Pro	Gln	Gly	Ser	Ile	Thr	Arg	Asp	Ser	Thr	Phe	Gln	Leu	His
				365					370					375
Val	Arg	Cys	Val	Phe	Asn	Ala	Ser	Asp	Phe	Leu	Pro	Ile	Gln	Ala
				380					385					390
Ser	Ile	Phe	Pro	Pro	Pro	Ser	Pro	Ala	Pro	Met	Thr	Gln	Pro	Gly
				395					400					405
Pro	Leu	Arg	Leu	Glu	Leu	Arg	Ile	Ala	Lys	Asp	Glu	Thr	Phe	Ser
				410					415					420
Ser	Tyr	Tyr	Gly	Glu	Asp	Asp	Tyr	Pro	Ile	Val	Arg	Leu	Leu	Arg
				425					430					435
Glu	Pro	Val	His	Val	Glu	Val	Arg	Leu	Leu	Gln	Arg	Thr	Asp	Pro
				440					445					450
Asn	Leu	Val	Leu	Leu	Leu	His	Gln	Cys	Trp	Gly	Ala	Pro	Ser	Ala
				455					460					465
Asn	Pro	Phe	Gln	Gln	Pro	Gln	Trp	Pro	Ile	Leu	Ser	Asp	Gly	Cys
				470					475					480
Pro	Phe	Lys	Gly	Asp	Ser	Tyr	Arg	Thr	Gln	Met	Val	Ala	Leu	Asp
				485					490					495
Gly	Ala	Thr	Pro	Phe	Gln	Ser	His	Tyr	Gln	Arg	Phe	Thr	Val	Ala
				500					505					510
Thr	Phe	Ala	Leu	Leu	Asp	Ser	Gly	Ser	Gln	Arg	Ala	Leu	Arg	Gly
				515					520					525
Leu	Val	Tyr	Leu	Phe	Cys	Ser	Thr	Ser	Ala	Cys	His	Thr	Ser	Gly

	530		535		540
Leu Glu Thr Cys Ser Thr Ala Cys Ser Thr Gly Thr Thr Arg Gln					
	545		550		555
Arg Arg Ser Ser Gly His Arg Asn Asp Thr Ala Arg Pro Gln Asp					
	560		565		570
Ile Val Ser Ser Pro Gly Pro Val Gly Phe Glu Asp Ser Tyr Gly					
	575		580		585
Gln Glu Pro Thr Leu Gly Pro Thr Asp Ser Asn Gly Asn Ser Ser					
	590		595		600
Leu Arg Pro Leu Leu Trp Ala Val Leu Leu Leu Pro Ala Val Ala					
	605		610		615
Leu Val Leu Gly Phe Gly Val Phe Val Gly Leu Ser Gln Thr Trp					
	620		625		630
Ala Gln Lys Leu Trp Glu Ser Asn Arg Gln					
	635		640		

<210> 21

<211> 388

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7640495CD1

<400> 21

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

WO 02/062841

PCT/US02/02616

Asn Gly Thr Tyr	275	Asn Lys His Gly His	285
Thr Cys Glu Ala Ser	280	Asn Lys His Gly His	285
Val Leu Val Val Tyr	290	Asp Pro Gly Ala Val	300
Val Leu Val Val Tyr	295	Asp Pro Gly Ala Val	300
Val Leu Val Val Tyr	305	Asp Pro Gly Ala Val	315
Val Leu Val Val Tyr	310	Asp Pro Gly Ala Val	315
Val Leu Val Val Tyr	320	Asp Pro Gly Ala Val	330
Val Leu Val Val Tyr	325	Asp Pro Gly Ala Val	330
Val Leu Val Val Tyr	335	Asp Pro Gly Ala Val	345
Val Leu Val Val Tyr	340	Asp Pro Gly Ala Val	345
Val Leu Val Val Tyr	350	Asp Pro Gly Ala Val	360
Val Leu Val Val Tyr	355	Asp Pro Gly Ala Val	360
Val Leu Val Val Tyr	365	Asp Pro Gly Ala Val	375
Val Leu Val Val Tyr	370	Asp Pro Gly Ala Val	375
Val Leu Val Val Tyr	380	Asp Pro Gly Ala Val	385

<210> 22

<211> 255

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5960119CD1

<400> 22

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

<210> 23

WO 02/062841

PCT/US02/02616

<211> 105
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7500143CD1

<400> 23
 Met Lys Arg Ser Leu Gln Ala Leu Tyr Cys Gln Leu Leu Thr Val
 1 5 10 15
 Leu Leu Thr Val Cys Cys Met Lys Arg Lys Lys Lys Thr Ala Asn
 20 25 30
 Pro Glu Asn Asn Leu Ser Tyr Trp Asn Asn Thr Ile Thr Met Asp
 35 40 45
 Tyr Phe Asn Arg His Ala Val Glu Leu Pro Arg Glu Ile Gln Ser
 50 55 60
 Leu Glu Thr Ser Glu Asp Gln Leu Ser Glu Pro Arg Ser Pro Ala
 65 70 75
 Asn Gly Asp Tyr Arg Asp Thr Gly Met Val Leu Val Asn Pro Phe
 80 85 90
 Cys Gln Glu Thr Leu Phe Val Gly Asn Asp Gln Val Ser Glu Ile
 95 100 105

<210> 24
 <211> 190
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7503605CD1

<400> 24
 Met Ala Gly Thr Gly Leu Leu Ala Leu Arg Thr Leu Pro Gly Pro
 1 5 10 15
 Ser Trp Val Arg Gly Ser Gly Pro Ser Val Leu Ser Arg Arg Tyr
 20 25 30
 Glu Tyr Asp His Trp Asp Ala Ala Ile His Gly Phe Arg Glu Thr
 35 40 45
 Glu Lys Ser Arg Trp Ser Glu Ala Ser Arg Ala Ile Leu Gln Arg
 50 55 60
 Val Gln Ala Ala Ala Phe Gly Pro Gly Gln Thr Leu Leu Ser Ser
 65 70 75
 Val His Val Leu Asp Leu Glu Ala Arg Gly Tyr Ile Lys Pro His
 80 85 90
 Val Asp Ser Ile Lys Phe Cys Gly Ala Thr Ile Ala Gly Leu Ser
 95 100 105
 Leu Leu Ser Pro Ser Val Met Arg Leu Val His Thr Gln Glu Pro
 110 115 120
 Gly Glu Trp Leu Glu Leu Leu Leu Glu Pro Gly Ser Leu Tyr Ile
 125 130 135
 Leu Arg Gly Ser Ala Arg Tyr Asp Phe Ser His Glu Ile Leu Arg
 140 145 150
 Asp Glu Glu Ser Phe Phe Gly Glu Arg Arg Ile Pro Arg Gly Arg
 155 160 165
 Arg Ile Ser Val Ile Cys Arg Ser Leu Pro Glu Gly Met Gly Pro
 170 175 180
 Gly Glu Ser Gly Gln Pro Pro Pro Ala Cys
 185 190

<210> 25

WO 02/062841

PCT/US02/02616

<211> 1992
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1849820CB1

<400> 25
 ggccgcccgc tccccgcagg tggacgcggc catggggccga ggggtgcgcg tgctgctgct 60
 gctgagcctg ctgcactgcg ccggggggcag cgaggggcagg aagacctggc ggcgcccggg 120
 tcagcagccg cctcctcccc cgcggaccga ggcggcgcgc gcggccggac agcccgtgga 180
 gagcttcccc ctggacttca cggccgtgga gggtaacatg gacagcttca tggcgcaagt 240
 caagagcctg gcgcagtccc tgtacccttg ctccgcgcag cagctcaacg aggacctgcg 300
 cctgcacctc ctactcaaca cctcgggtgac ctgcaacgac ggcagccccg ccggctacta 360
 cctgaaggag tccagggggca gccggcggtg gctcctcttc ctggaaggcg gctgggtactg 420
 cttcaaccgc gagaactgcg actccagata cgacaccatg cggcgccctca tgagctcccc 480
 ggactggccg cgcactcgca caggcacagg gatcctgtcc tcacagccgg aggagaacct 540
 ctactgggtg aacgcaaaca tgggtcttcat cccctactgc tccagtgatg tttggagcgg 600
 ggcttcatcc aagtctgaga agaacgagta cgccttcatg ggcgcccctca tcatccagga 660
 ggtggtgcgc gagcttcttg gcagaggggt gagcggggcc aaggtgctgc tgctggccgg 720
 gagcagcgcg gggggcaccc ggggtgctcct gaatgtggac cgtgtggctg agcagctgga 780
 gaagctgggc taccagcca tccaggtgcg aggcctggct gactccggct ggttcctgga 840
 caacaagcag tatcgccaca cagactgcgt cgacacgatc acgtgcgcgc ccacggaggc 900
 catccgcccgt ggcatcaggt actggaacgg ggtggtcccc gagcgctgcc gacgccagtt 960
 ccaggagggc gaggagtgga actgcttctt tggctacaag gtctaccoga ccctgcgctg 1020
 ccctgtgttc gtggtgcagt ggctgtttga cgaggcacag ctgacgggtg acaacgtgca 1080
 cctgacgggg cagccgggtgc aggagggcct gcggctgtac atccagaacc tcggccgcga 1140
 gctgcgccac aactcaagg acgtgccggc cagctttgcc cccgcctgcc tctcccatga 1200
 gatcatcatc cggagccact ggacggatgt ccagggtgaag gggacgtcgc tgccccgagc 1260
 actgcactgc tgggacagga gcctccatga cagccacaag gccagcaaga cccccctcaa 1320
 gggctgcccc gtccacctgg tggacagctg cccctggccc cactgcaacc cctcatgccc 1380
 caccgtccga gaccagttca cggggcaaga gatgaacgtg gccagttcc tcatgcacat 1440
 gggcttcgac atgcagacgg tggcccagcc gcagggactg gagcccagtg agctgctggg 1500
 gatgctgagc aacggaagct aggcagactg tctggaggag gagccggcac tgagggggcc 1560
 agacacccgc tgccccagtg ccacctcacc cccaccagc aggcctccc gtctcttcgg 1620
 gacagggccc cagccgtccc cctgtcttg gtctgcccac tgccctcctg ccccggttt 1680
 ccctgcccct ctcccacagc ccagccagag acaagggacc tgctgtcatc cccatctgtg 1740
 gcctgggggt ccttcctgac aacgaggggg tagccagaag agaagcactg gattcctcag 1800
 tccaccagct cagacagcac ccaccggccc caccatcaa gcccttttat attattttat 1860
 aaagtgactt ttttattact ttaatttttt aaaaaaagga aaataagaat atatgatgaa 1920
 tgatatgttt ttgtaacttt ttaaaaatga ttttaaagag acaaaaaaga acctcacaaa 1980
 aaaaaaaaaa aa 1992

<210> 26
 <211> 806
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 70610307CB1

<400> 26
 atgctgcctc ccatggccct gccagtggtg tcctggctac tcttcatttt actctctccc 60
 ttttcttcca ccccagggtga agaaacccag aaggaaactgc cctctccacg gatcagctgt 120
 cccaaaggct ccaaggccta tggctcccc tgctatgcct tgtttttgtc accaaaatcc 180
 tggatggatg cagatctggc ttgccagaag cggccctctg gaaaactggt gtctgtgctc 240
 agtggggctg agggatcctt cgtgtcctcc ctggtgagga gcattagtaa cagctattca 300
 tacatctgga ttgggtcca tgacccaca cagggtcttg agcctgatgg agatggatgg 360
 gagtggagta gcactgatgt gatgaattac tttgcatggg agaaaaatcc ctccaccatc 420
 ttaaacctg gccactgtgg gagcctgtca agaagcacag ccccatgcct tttatatctt 480
 gtctccctag gatcttctgaa gtggaaagat tataactgtg atgcaaagtt accctatgtc 540

WO 02/062841

PCT/US02/02616

tgcaagttca	aggactaggg	ccagtgggaa	atcatcagcc	tcaacttcgc	gtgcagctca	600
tcatggacat	gagacccgtg	tgaagactca	ccctgtggag	agaatattct	cccccaactg	660
ccctacctga	ctaccctgtc	atgatccctc	ctctttttcc	cttttcttca	ccctcatttc	720
aggcttttct	tccgtcttcc	aagtcttgag	atctcagaga	attattatta	aaatgttact	780
ttatacttta	aaaaaaaaaa	aaaagg				806

<210> 27

<211> 1004

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8137559CB1

<400> 27

caatcggccg	aggcagtatc	tgtagctggg	agtgtggtac	ttgagtaata	acaaaccctg	60
catcaggagt	ggggcatcaa	gagggacaga	gactgggggtg	tagcacactc	agacaccttc	120
ctgcttctcc	ttgctgctct	ataacaagtg	tgccttgagg	agacaaaatg	gactcatggg	180
ctggtccttg	agcccagaga	gctgagtacc	cacaggagga	gggtgggaag	gaagagctca	240
ttctctcttt	gtatgtgtat	gtgtgtgtgt	gtttgtgtgt	gcatgcagga	gagagagaaa	300
gaggggaggg	ggagagagag	agagagaaaa	aaaaaacctg	ctgcctgagg	gagctggggc	360
tcagtgagcg	aaaacactct	gctttttcaa	ctctacatgg	ttacagctaa	acccagcaag	420
tcctaggtca	tgacctacag	aaatgcccat	gaagaaaagt	tggatgcca	agacttgcca	480
catatttctg	ttacttgcag	ctttctttca	gaacagactg	actgaccctt	ttccatgttc	540
catctgcggg	gaatgccagt	atggcttctc	atttccttct	ttcttcttcc	tcatcagttc	600
agccccagta	aaagcatttc	ccctttccac	agaccactcc	tgtggactct	gctgggctgt	660
gtaatgtaat	gtgccttctt	ctgggctccc	tggactctgc	agtctctgtg	cctgttaatt	720
tctctgtttc	cactggttga	cattttgttc	ctcatggcca	tgggactctg	tctaattctt	780
ccttatcttc	catcctccac	ctccagggtg	ggaacagagc	aggagttag	gaactatttg	840
gtgagtgagt	gagtgataat	agataattag	atgctcatat	atagataatt	aggtgctcat	900
atatcctcta	ttgccttact	tttttctctc	tctgcccttt	cacctcattg	ctttgtcctc	960
cttttgcttt	ccagggttgc	ttcgtgaagg	ggggcggtgc	ggcg		1004

<210> 28

<211> 1202

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4801255CB1

<400> 28

aggctggaga	tttcatcaag	ctactcacag	cagcatgccca	tttaaaattc	atgaattatt	60
tattttctgga	atttttccgtt	taatatTTTT	ggaccacaat	tgacctcaga	ttttactttc	120
tactcttgtc	agtcatatct	gcaaggctga	cagttaccag	gggctgtggc	caggcatgag	180
ggagaatgtg	tcattgctgg	ctgcgttggt	catagcaacg	tgtctgcctc	agaacagatc	240
cagcatggat	gagcaaacag	aaaaatgtgt	aagaaggctt	gtgactgaaa	cagaccaagg	300
cagacacttg	aagaagcatc	tttcttgtga	tctcctggct	ggtaccctga	agtcctgtta	360
tgccaatgga	ccagcaaccc	tcgcaggtag	gaattactgc	tagaaagact	atcctctctg	420
caggcaaatt	tctggaagga	tactatTTTA	tggggaaaaa	tgtgaacaac	tgtaggtaat	480
ttcacagtct	gtagggcttg	tctcctatTT	gtaatttctt	acatgtttac	tcgtgggtga	540
aatgtggggg	tggctaagga	acaagcaggc	ccgctgtggt	tcagtggctt	tcagagaagg	600
tggcgggaaa	tagtctcttc	tcagcagcag	ctgaatccca	gtgttagcag	acggagcctt	660
gtttcttgaa	agtacattcc	cagggattcc	cttgaagaat	gtgagagatc	agctggaatg	720
aaactggatg	tttttagggtc	ctgagggaaa	tgtaaatcca	gcagcttatt	ctcaagatgt	780
ggctccagaa	ccagcagcac	tgattcttct	aaagagttga	gagtgcattt	tagctgtgat	840
atctttgtac	tgtaagaaa	tcctcctcaa	caagacaaaa	cttgaccaag	agcattcagt	900
tgtcttccag	aaggctccct	tgaggactgg	gactcttcct	ggatgcaccc	aagtaatcta	960
atgaaacaag	gactcgtgat	ggcagcccct	cacctgacaa	catggatttt	ccccagagat	1020
catttgtaat	gtttttcctg	aatcattaaa	ttgtgtttcc	ctattagcaa	tgttatgtat	1080
ttctatatcg	catatcataa	gtgaccgcca	tgtagaacat	aacctccatt	tatagcatgg	1140

WO 02/062841

PCT/US02/02616

acacacagcc catcacgtat catgaccttg catacgccta gataacaaag atcctcacca 1200
cc 1202

<210> 29
<211> 1416
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6160719CB1

<400> 29
ctttttttttt gtaaaaatgt tttacttttta attttttatgg atatatagag atgttttaa 60
tactagccca aagtcacagt tggtaaaatg ttgagctgga atttgaattg agcaagcagt 120
ctacccctta gcagaatctg ctttttttta atgaactaag cagcctggcc agagtccatg 180
ccctcagcag ctacgttggtg ttgtccttaa ggaggagcat gtttttagaaa atcagttata 240
aaatcagggtg attgatgaag tgcagggtta agttaagaga agagtgaact gccttcagaa 300
tccagtgtg ttaccactca agtggagatc tcagataaat cacttattgg agcttctgta 360
caatcatctg taaaaccatt acttcccact tggagagatt tttgaggatt aaatgagata 420
atgcatgaaa gccctctagc ttgggcatca gtacacctga gttcccttcc cttgctctgc 480
acagcctgct catcactact gatggggaac tctgtcctct gtagggcccc tgcagacatg 540
ggccttgccct ggatgctgct gctgtcggag cctaggagag ttgtgcctgg catcgcagca 600
cagggtactca cagctctcag aaggagactc ctgtctggga ccctgccctc attcccacgt 660
aggaaaaatc ctttacctga gcatctcctg gccttcattg ttaggttgta gactacaatg 720
aatgatattc tgtgtttaat tacattatgc acaacactct acagagtggg tggttttgaa 780
tcccaaccac taatttacga agtggagcgg ctctgctggc tctgtgaagt atgtgttggtg 840
gagccagagg tgatgctggt ggatgtgggt ggtgatttac gggagagcag cataagcaga 900
ggaaggcaca gagacctggg ttcaaattccc actgccagtg ctatctgacg tgagacttcg 960
gacaagttaa ttaaccttaa ggcttagtgt ccttgcatgt aaaatacaaa taatgctgac 1020
ctcattgatt ccttggtgagg agcccatggg ataattgtgg gtaccatgca tcagtatcat 1080
ttcccttttcc cttgataatg attatatagg acagctagtg tcataggaag ctaatggtta 1140
ggaattcaga ctacacacaag aatagatttt gttttgcagt ggaaacttag tgataacttt 1200
ggactaagac aatgtgaaga cattttggca gacaatcata gagtagttat attcttgtga 1260
ggtctgcaga agtccctagt ggggagagca gacagagcct tggatgctca gacctagttt 1320
ggctgtgaaa ctgaagcttt agtcctttta agtctgtgtt tctgaccctg gaggaaaaaa 1380
gttatacata ccttgaaaga ttttccttga aaaaaa 1416

<210> 30
<211> 1707
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3524602CB1

<400> 30
cagggcgaag taaaatatat ttgaaataat cgttcaggtc agtttttggtt tgtaatggaa 60
atttctcttc caatgtaaat aaagcatttg atatgaaaat tacaacatgc aaagagctgc 120
tgagtacatt ttaggcatct caagtggatt ttaatgtaac attttaatag cttttacatt 180
tctttttaca aaacacagct atttcattag ctattcagat ttccttacca cttccatttg 240
aaccctatga aaatatatga aaaaatatat cccccacccc caaccagcaa ttgtataaga 300
atagcaatta ttttaaaaag aacttgcata gtattctgaa tttatcatgt tgttagaaag 360
gctgcctatg gatttttttg agattttcag aaaatgctgc aaactggaca ctccagttct 420
tccagatatt ggtagccgcg gcgcagaggg aatgctgtgt ggttaggagt gattacacca 480
ccagagactt caggactatg aagaggaaaa agtagtggtt ttaaaaaata aggagatctt 540
actaaaagg atctgggaca gaggcattgaa ggaagtgaac aggattcaac agctgagaac 600
catcagccag aacataaagg atagggaga aaatagatga taacaggctt atggccatgt 660
ggaagtgaac acacactcac cagtgttaat ttacatccaa gactaactga aattatgctg 720
aaaacattat tcaactggctg gtgcatactc gtgacaaaat gacagaacac taagagttct 780
tccagttaaa tttattcatt ttttaaaaat tctgagctat gttattgtta acagataaaa 840
taccacttgt tttataaaac tacactgata gtcctcagtt caggttttcc tggaagctga 900

WO 02/062841

PCT/US02/02616

```

ccttatgatg atgatgtgag tacaagtagt ttatctggga ggcaaaggga accccagtaa 960
gaaaatgtag ccaaaaaatt aagggttaca aagccagcta ccactgtgac ggacagatgc 1020
ttgatctggt gaagaagctc taggcaaagt gtaagtatac atgtaagagt tggcacaccc 1080
caggcaggag cggggaccat atttacagac tctgtgtgtc actgattgaa agctgcttcc 1140
caggtcttaa ctctctggca ctctctggctg ccacatactg gggcagggtt ggctttttgac 1200
aagagccctc agatatgtca atgtgggatg tacagtctcg gggaacaagt taccaagagt 1260
ggcctttttat ttccactgtt cctaccatgg caaaatcatg tcagtcattg atgttgggtt 1320
cccacttact ttccaaccct ccccaaacc aaagcctgg gacactgata ctccaccagc 1380
cagaggatcat gtcagttgat gacatgaatg acagtatagg gcagggtgcag ctcagagtga 1440
gacaaattct tgccaatcat tctgacgggg aaggatgact ctggaactag caagagaagc 1500
actgaaggat gcctttcctg accctcacat ccaacaaggg cctccaaatc ctcgccttcc 1560
tgatccattt gttgttgttg ttgttgttga gatagagcct tgcttgggtca cccaggctgg 1620
agtgcagtgg tgcaatctcg gctcactgca acctccactt ctcaggttca aacgattctc 1680
ctgcctaggc ctcccagagta gctggga 1707

```

<210> 31
 <211> 1436
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 513718CB1

```

<400> 31
gaaaaaggcc ggccaccgag cggctgtggg tgaatgagtt aaccaaagag caaatgtctg 60
agttggatca tgaaacagca gatcaagtgg gctcagtgtc tgtaaacatc tacactgccc 120
atgaggacaa gatctgtatg ctccatctac acccagaggc tgaaggaatg cctgtggtgt 180
agacaactca cttttctaga gaaatggaga agggattaga catgttaacc ctgaagaaaa 240
ggaaacttgg ggtagggtag gctcagaaag tcttcagagg ttgggaacgt ggcccatgaa 300
agagcaggga gactggaaag ggaccctgcc tgtggaggca gctctgggcc ccacttgaa 360
acacacgggt gaaccacttg acccacctta cagtgaaga agctggcacc tgtggaccac 420
tggagaaacc tcagtagcca tctggtgttc ccaatgact gtgttgctca caaccctcaa 480
gagtcctttg ttgcttttgg aattgtaggg ctggcaagat gtcttcaccc actgctagat 540
tcatggctga ggcccctata atgaaggaca gatgaacaag agaagaacat acaaacgcat 600
agaatataag ttttacatga cacagaagcc tttgagggga aatgaagaac cagagaaaca 660
gggaaacctg tgtatttgct atgctcgatt ggatgaagag tggacagtca tggagaggca 720
tgaccgagca aagcgggtga taacctggag ggaatgtcac cacatctgtt ggctcagatt 780
ctcctctgtg tccttgtgtc ttcagaaaaa ggatgttcct ttcctctgag tatgcgggca 840
tactcacac cagggtccaa tgtcctactt caggggaaag tcaggaaatc cttcctggga 900
ctgatgaccg gcctcagaga gaagggtaga acaagggaag gtgagagtgg cctttctgct 960
tttgctgttt tctcaaattg caattttcta ttttcagggg caatatgtcc tgaaccccat 1020
caggatgaag ccccagctcc atgaagcccc tcttgaactg gtccatccta ctgttccagt 1080
ctcagctccc tctgctccag gccttgaact tcacattcca gcctggggac tcctggcagt 1140
tcctaaata tgcaaagagc tcagctttct ttcctgggtc tggctcctgg cacttggccc 1200
tctctgtctc tcccttcctt cttcccatct ctcttccttt ctggctcatg tccctcgcac 1260
atgctgtaaa ctcttcctgg aacatctccc ttcacactgt gaatctcttg ccagctcact 1320
tcttttctcc ctttaagaat ccacttctca aaagaaaatg gcaactctgt gaggtgatag 1380
aatgctgat tcgcttgatc attgtggcgt ccacaatgta tatataaaac aagttg 1436

```

<210> 32
 <211> 1984
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 896308CB1

```

<400> 32
cgccccctttt tttttttttt ctttttcagag ggtctcgctc tgtcaccctg gcaggagcac 60
agtgggtgcaa tcatagctca ctgtagactc aaatttgggc tcaagagatc ctctgtctc 120
agcctcccga gtagctagta gctgagacca caggcgggag ccactgggcc tgtctcttat 180

```

WO 02/062841

PCT/US02/02616

```

atcttttatgt cagctctttt ctgatggttt ccttagtctc tacatatattt tactctttaa 240
cctcaggcat gcacataccc ctctgagggtg cttcccagtc tgtagcctcc tccaatttca 300
catgccccga attaaactca ctattctcct cgggaaaact tcttgctact ctattccaca 360
tttcaacttca tggcccacca cctatccagc ctctgaaggc ggcgggttggg gagacactta 420
ggactgttca caccaacaga agtggccaga cactctgtct cctaattatt ttaaacaatgc 480
attctactat ccgttcccaa aaaatatgta gcactcgtgc aagtttgtgt ctacattatt 540
agaactaatc agcctctacc tccctcgagg cagaggctca tctgtcattg accccgggac 600
acgtctgctc caggcatgct gactctggcc ttgacgttcc ccaagcacat ggtatccttt 660
gacaaggcca catctgcacc ctcagcaacc ccagcagcct cctctgctgg caagctcctg 720
caggcgcatt ccaaacagtt caaatgtcac cgcctctcct ccagttctcc ctcaacaaca 780
gttagtggtg ctcttttact gtggatacac agtgcttctc acaggtctcc attcaaataca 840
agccaaggaa tgttttttaa cacctgttgt gcaaccagta cactgtgctg ggcactggaa 900
aatatgaaaa taactaagag ctgaactcta tccttatgga tttcacctgc tagaaggaga 960
cagacatgta aggaaataat gagcatacag cactataaat gcaaagcagc agagaggcca 1020
ggacagctga ggtggaatct gaagagccaa gagttttcat tggaacgatt tgcttagaca 1080
cccccttatt ccctcaccct ccaaccagta agctccttaa gagggcatac ttgcttaata 1140
ctgtttgtag tctcagaatc tagtacactg cccaacacag agcagcgact ctaagcactt 1200
cttgaatgaa tgggcaagcc tgcactgtga catattaagc ctactcatga gtcattccac 1260
aaagcaatct tctccataat tactgccaaag taagaattgt gtgaaacgcc tggaaaagat 1320
caagcattta caagtgatga tgtccgtgag gatgaaaaga ggacatgtgc ctgcaggcca 1380
cctgcatggc cctgcttggt tggctcctgac ctccctcgcc catccatgtc ctccagcccc 1440
caccacacac agtgctgtgt tcaactgttc tcacgcaccc tacaccctgc ctcacctcgg 1500
ggtctctgca cggctgggtcc ctgtgcctgg aaagtccctc cgcctgacac ccaaatgcct 1560
ccctcctccc ttcctttcag gtgtttgccc aaatgtcgcc ctctcagtga ggcctttctt 1620
gaccacccgt ttaaaaatct aatctccttc actgctttat ttttatttta ttttattttt 1680
tgcttctttt ttttgagata gggctcttgct ctgtcaccca ggctggagtg cagtgggtgcg 1740
acctcagcta gctgcaacct ccacctcca ggttcaagag attctcctgc ctccagggaag 1800
gggaggttgt ggtgagccac gattgcgtca attgcactcc agcctgggca acaagagtga 1860
aactccatct cataaataaa taaataaatt aattaataaa tcaaacaaac aaaactcatg 1920
gccatactaa tagtattgag attaataatg tgttgggtta cagcatcact ggctgggtctg 1980
aatg

```

<210> 33

<211> 1131

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2105862CB1

<400> 33

```

gaatcaaaat ggagttacta atgttaagga aaccagggga agctgtgtac agagccaggg 60
aaggctgtga agagagtgtt ctcacacttg tatgcctgat aatgaaagag actacaaaga 120
ccatatacctt acacaaaggc catcaccacc ttacacaaaa tagtacttct gcaaggacat 180
ctgcccagca actgcctgtt cagcctccaa ctgggtgtcac ccttggtatt gatctttgta 240
gctaaggata tttattttcaa aactattata taacccttgt ttctccttta aaaacctttg 300
tcttccttta cctccctgaa tatgtatggc atagtttact atgttatgca tattcctatt 360
gcaatgctct gttctgaaat aaacatcttt tctttgtgat cctctctctg ttatgtagg 420
tgacagtaaa aattcaaagt ggttttagtga tttgaacttt gctttatata aacatttcct 480
ttttatatgc tttgtgaaac tagaagggtg aataattaat ggcatactta gttaaaaaaa 540
tttatatgat atgcacatac tgcatttggt cctaataatg atctcaacat ttcatttaca 600
attagcttat tccacagtat taagaaaaca taggtttctt cctattttat ataatcagc 660
ctttaaaata aaacaaacca gcttttgcaa aattatttac aaagatacct ggccttgta 720
tttgtcattt gaaaataatt atgggacctg ctttttaa at tttactgaggg ggatctcttt 780
ctgctgtaaa attttacttc tttctgaagt aaaattgtat tttaaaaaat gatataatga 840
aaggccaggt gcagtgggtc atgcctgtaa atcccagtac tttggaggct gaggcaggta 900
gatcacctga ggtgtcagga gttcgagatc agcctgggtc acatggcaaa acccgtctc 960
tactaaaaat acaaaaatta gccgggcctg gtggcacatg cctgtaatcc cagctacttg 1020
ggagggtgag gcacgagaat cacttgactt gggaggcaga ggttgacagt agccgagatc 1080
ctgccactgc actccccagc ctgggcgaca gagtgagact ccaaaaaaaa a 1131

```

<210> 34

WO 02/062841

PCT/US02/02616

<211> 2417

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7939381CB1

<400> 34

```
gggccttcgc ggtgcagctg aggctgcaag tagccggcgc cgtcccgcgt cgccccgcgc 60
cagggcgggc cccgcacgct taccctgccc gggaggaacg ccggcgctcca gcccgctacc 120
ggccgcccgt gcgggatgct gcgctccacg tccacgggtca ccctgctctc gggcggcgcc 180
gccaggacgc ccggggcgcc cagcaggagg gcaaatgttt gcagactacg gctgaccgta 240
cctcctgaga gtccagttcc tgagcaatgt gaaaagaaga ttgagagaaa agagcagctt 300
cttgacctga gcaatggaga acctaccagg aaacttcctc aggggtgttg ttatgggtgtg 360
gtgcgaagat cagatcaaaa tcagcagaaa gaaatgggtg tgtatgggtg gtccaccagt 420
cagctgaaag aagagatgaa ctacatcaaa gatgtgagag ccactttgga aaaggtgaga 480
aagcgaatgt atggagacta tgatgagatg agacagaaga ttcgacagct caccaggaa 540
ctatcagttt cccatgctca gcaggagtat ctggagaatc acatccaaac ccagtcgtct 600
gccctggatc gttttaatgc catgaactca gccttggcgt cagattccat tggcctgcag 660
aaaaccctcg tggatgtgac tttggaaaac agcaacatta aggatcaa at cagaaatctg 720
cagcagacgt atgaagcatc catggacaag ctgagggaaa agcagaggca gttggaggta 780
gcgcaagtgt aaaaccagct gctaaaaatg aaggtggaat cgtcccaaga agccaatgct 840
gaggtgatgc gagagatgac caagaagctg tacagccagt atgaggagaa gctgcaggaa 900
gaacagagga agcacagtgc tgagaaggag gctcttttgg aagaaaccaa tagttttctg 960
aaagcgattg aagaagccaa taaaaagatg caagcagcag agatcagcct agaggagaaa 1020
gaccagagga tcggggagct ggacaggctg attgagcgca tggaaaagga acgtcatcaa 1080
ctgcaacttc aactcctaga acatgaaaca gaaatgtctg gggagttaac tgattctgac 1140
aaggaaaggt atcagcagtt ggaggaggca tcagccagcc tccgtgagcg gatcagacac 1200
ctagatgaca tgggtgcattg ccagcagaag aaagtcaagc agatggtcga ggagattgaa 1260
tcattaaaga aaaagttgca acagaaacag ctcttaatac tgcagctttt agaaaagata 1320
tctttcttag aaggagagaa taatgaacta caaagcaggt tggactat tt aacagaaacc 1380
caggccaaga ccgaagtgga aaccagagag ataggagtgg gctgtgatct tctaccagc 1440
caaacaggca ggactcgtga aattgtgatg ctttctagga actacacccc atacacaaga 1500
gtcctggagt taaccatgaa gaaaactctg acttaggcac tcagaggcat acacttttta 1560
cagatggaca aaagctctgg aaccctgtgg cttcaaatec tttgggaagg gtgactgttg 1620
tttcccctac acacagtgtg agccggaatg ggaatcgctg aggtctgat ccacttctaa 1680
gacaggaagg aaagtgaagg cagagtgagc aggtgaagaga gggatataca aggtcacatt 1740
tcagacaccc actcggcata ccctgccgta ctgcatcatc atttgttttc tttgtagaca 1800
ctgaaatcct atcaggagga ttccttcaca atgtatttta tttgctagac tttggttggg 1860
agggaaaagg acattaattt gaagtttcat gttattcatg ccaggattgt ttgatagagc 1920
atgaaggttt tgtttaccca taaaagtatt agaggcagcg tttctctgat acagagaggc 1980
ctgtccacaa gaagcatggg caccagcca aacttgaacc tggaaggagg ggttcccggc 2040
ctgcagggtg tctttcctct tgggtcccaag catctgtgca gggtcgtggg agccacactg 2100
agagacttgt gtgggcccaga caagcttcat tctgatgcgc tagtcccttg gtttaatttg 2160
tgcttatgct tttcattgga ccagctgaaa tcaactgtatt tattcaactt gtgatttttt 2220
tttctttctc actttaactt aaagagaatt ttatatgtct tggaaattta ataatttagt 2280
gttctcagta tcaattgggtg tttttgttaa acgaatgaat catctgttca tgcagtctct 2340
actttgatat tataacctat gtcacatgtg tttaataaat accatatatt ttgttctaaa 2400
aaaaaaaaa aaaaaaa 2417
```

<210> 35

<211> 2768

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7487507CB1

<400> 35

```
ttggcatgat gggcacctgg agggccgcac tcccgttcca gccaggctga gccttctgtc 60
ccctgcctct ggggcctggg aacccccctt cttctttctc ctgaatggca cccccgcct 120
```

```

agaatccaga caccgagttt cccactgtgg ctgggttcaag ggtatgtgag agctccctgg 180
tgacagtctg tggctgagca tggccctccc agccctgggc ctggacccct ggagcctcct 240
gggccttttc ctcttccaac tgcttcagct gctgctgccg acgacgaccg cgggggggagg 300
cgggcagggg cccatgcccc gggtcagata ctatgcaggg gatgaacgta gggcacttag 360
cttcttccac cagaagggcc tccaggattt tgacactctg ctcttgagtg gtgatggaaa 420
tactctctac gtggggggctc gagaagccat tctggccttg gatataccagg atccaggggt 480
ccccaggcta aagaacatga taccgtggcc agccagtgc agaaaaaaga gtgaatgtgc 540
ctttaagaag aagagcaatg agacacagtg tttcaacttc atccgtgtcc tggttttcta 600
caatgtcacc catctctaca cctgcggcac ctctgccttc agccctgctt gtaccttcat 660
tgaacttcaa gattcctacc tgttgcccat ctcgaggac aagggtcatgg agggaaaagg 720
ccaaagcccc tttgaccccc ctcacaagca tacggctgtc ttgggtggatg ggatgctcta 780
ttctgggtact atgaacaact tcctggggcag tgagcccatc ctgatgcgca cactgggatc 840
ccagcctgtc ctcaagaccg acaacttcct ccgctggctg catcatgacg cctcctttgt 900
ggcagccatc ccttcgaccc aggtcgtcta ctctctcttc gaggagacag ccagcgagtt 960
tgacttcttt gagaggctcc acacatcgcg ggtggctaga gtctgcaaga atgacgtggg 1020
cggcgaaaag ctgctgcaga agaagtggac caccttcttg aaggcccagc tgctctgcac 1080
ccagccgggg cagctgccct tcaacgtcat ccgccacgcg gtccctgctcc ccgccgattc 1140
tcccacagct cccacatct acgcagtctt cacctcccag tggcaggttg gcgggaccag 1200
gagctctgcg gtttgtgcct tctctctctt ggacattgaa cgtgtcttta aggggaaata 1260
caaagagttg aacaaagaaa cttcacgctg gactacttat aggggccctg agaccaacc 1320
ccggccaggc agttgctcag tgggcccctc ctctgataag gccctgacct tcatgaagga 1380
ccatttcttg atggatgagc aagtgggtggg gacgcccctg ctggtgaaat ctggcgtgga 1440
gtatacacgg ctctgcagtg agacagccca gggccttgat gggcacagcc atcttgctcat 1500
gtacctggga accaccacag ggtcgtcca caaggctgtg ggtgcagtgt ttgtaggctt 1560
ctcaggaggt gtctggaggg tgccccgagc caactgtagt gtctatgaga gctgtgtgga 1620
ctgtgtcctt gcccgggacc cccactgtgc ctgggaccct gagtcccgaa cctgttgct 1680
cctgtctgcc cccaacctga actcctggaa gcaggacatg gagcggggga acccagagt 1740
ggcatgtgcc agtggcccca tgagcaggag ccttcggcct cagagccgcc cgcaaatcat 1800
taaagaagtc ctggctgtcc ccaactccat cctggagctc ccctgcccc acctgtcagc 1860
cttggcctct tattattgga gtcattggcc agcagcagtc ccagaagcct ctccactgt 1920
ctacaatggc tccctcttgc tgatagtgc ggatggagtt ggggtctct accagtgtg 1980
ggcaactgag aatggctttt cataccctgt gatctcctac tgggtggaca gccaggacca 2040
gaccctggcc ctggatcctg aactggcagg catcccccg gagcatgtga aggtcccgtt 2100
gaccagggtc agtgggtggg ccgccctggc tgcccagcag tcctactggc cccactttgt 2160
cactgtcact gtcctctttg ccttagtgct ttcaggagcc ctcatcatcc tcgtggcctc 2220
cccattgaga gcactccggg ctcggggcaa ggttcagggc tgtgagaccc tgcgccctgg 2280
ggagaaggcc ccgttaagca gagagcaaca cctccagtct cccaaggaaat gcaggacctc 2340
tgccagtgat gtggacgctg acaacaactg cctaggcact gaggtagctt aaactctagg 2400
cacaggccgg ggctgcggtg caggcaccgg gccatgctgg ctgggcggcc caagcacagc 2460
cctgactagg atgacagcag caaaaagac cacctttctc ccctgagagg agcttctgct 2520
actctgcac actgatgaca ctacagcagg tgatgcacag cagtctgcct cccctatggg 2580
actcccttct accaagcaca tgagctctct aacagggtgg gggctacccc cagacctgct 2640
cctacactga tattgaagaa cctggagagg atccttcagt tctggccatt ccagggacct 2700
tccagaaaca cagtgtttca agagacccta aaaaacctgc ctgtcccaga cactcttgtg 2760
cagcactt 2768

```

<210> 36

<211> 2096

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1483931CB1

<400> 36

```

gagggagccg gagcgcttct cccgagttgg tgatagattg gtggtcatcc aacatgcaga 60
aatgaatgag cagtgaagag cagcagagcc gatgggtcat gaggatgtaa gtgcgtttga 120
aggcttccac accctctact ccaggacaga atcatgaata aactggagga taagcaggac 180
cagatgatac catgaagaga agtttacagg ccctctattg ccaactgtta agtttctgct 240
tgatcttggc actgaccgaa gcgctggcat ttgccatcca ggaaccatct cccagggaat 300
ctcttcaggt cctcccttca ggcactcccc cgggaacctt ggtgacagca cccacagct 360
ctaccagaca tacttctgtg gtgatgctga cccccaatcc cgatggaccc ccctcacagg 420

```

```

ctgcagctcc catggcaaca ccgacacccc gtgcagaggg gcaccctcct acgcacacca 480
tctccacccat cgctgcgaca gtaaccgccc cccattctga aagctccctg tccacagggc 540
ccgctccagc agccatggca accacatcct ccaagccaga gggccgccc ctagggcagg 600
ctgcccccac catcctgctg acaaagccac cggggggccac cagccgcccc accacagcgc 660
ccccccgcac taccacacgc agggccccca ggcccccagg ctcttcccga aaaggggctg 720
gtaattcatc acgcccctgtc ccgcctgcac ctgggtggcca ctccaggagt aaagaaggac 780
agcgaggacg aaatccaagc tccacacctc tggggcagaa gcggcccctg gggaaaatct 840
ttcagatcta caagggcaac ttcacagggg ctgtggaacc agagccctct accctcacc 900
ccaggacccc actctggggc tactcctctt caccacagcc ccagacagt gctgcgacca 960
cagtgccccag caatacctca tggggaccca ccaccacctc cctggggcct gcaaaggaca 1020
agccaggcct tcgcagagca gcccaggggg gtggttctac cttcaccagc caaggaggga 1080
caccagatgc cacagcagcc tcaggtgccc ctgtcagtcc acaagctgcc ccagtgcctt 1140
ctcagcgcgc ccaccacggg gaccacaggg atggccccag ccatagtgac tcttggttta 1200
ctgttaccac tggcaccagc agacctctgt ctaccagctc tgggggtctt acggctgcca 1260
cggggcccac ccagctgcc ttcgatacca gtgtctcagc cccttcccag gggattcctc 1320
agggagcatc cacaacccca caagctccaa cccatccctc caggggtctc gaaagcacta 1380
tttctggagc caaggaggag actgtggcca ccctcaccat gaccgaccgg gtgcccagtc 1440
ctctctccac agtgggtatc acagccacag gcaatttctt caaccgcctg gtccccgccc 1500
ggacctggaa gcctgggaca gcagggaaca tctcccatgt ggccgagggg gacaaaccgc 1560
agcacagagc caccatctgc ctgagcaaga tggatatcgc ctgggtgatc ctggccatca 1620
gcgtgcccac ctctcctgc tctgtcctgc tgacggtgtg ctgcatgaag aggaagaaga 1680
agaccgcaa cccggagaa aacctgagct actggaacaa caccatcacc atggactact 1740
tcaacaggca tgctgtggag ctgcccaggg agatccagtc ccttgaaacc tctgaggacc 1800
agctctcaga gccccgctcc ccagccaatg gcgactatag agacactggg atggtccttg 1860
ttaaccctt ctgtcaagaa acactgtttg tgggaaacga tcaagtatct gagatctaac 1920
tacagcaggc atcactttgc cattccgtat ttttcgtctc taaattataa atatacaaat 1980
atatatatta taaatataac ctttgtgtaa ccctgactta atgagaaaca ttttcagctt 2040
tttttcctat gaattgtcaa catctttttt acaagtgtgg tttaaaaaaa aaaaaa 2096

```

<210> 37

<211> 1562

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8175223CB1

<400> 37

```

gggcaggctc ggaagagtct cagcgcagaa ggaggaagga cagcacagct gacagccgtg 60
ctctggaagc ttctggatcc taggctcatc tccacagagg agaacatgca cgcagcagag 120
atcatggggc ccctctcagc ccctccctgc acagagcaca tcaaatggaa ggggctcctg 180
ctcacagcat tactttttaa cttctggaac ttgcctacca ctgcccaggt catgattgaa 240
gccagccac ccaaagtgtc cgaggggaag gatgttcttc tacttgtcca caatttgccc 300
cagaatctta ctggctacat ctggtacaaa ggacaaatca gggacctcta ccattacatt 360
acatcatatg tagtagacgg tcaaataatt atatatggac cggcatacag tggacgagaa 420
acagtatatt ccaatgcac cctgctgatc cagaatgtca cccgggagga cgcaggatcc 480
tacaccttac acatcataaa gcgagggtgat gggactagag gagtaactgg atatttcacc 540
ttcaccttat acctggagac tcccaagccc tccatctcca gcagcaactt aaaccccagg 600
gaggccatgg agactgtgat cttaacctgt aatcctgaga ctccggacgc aagctacctg 660
tgggtggatga atggtcagag cctccctatg actcatagga tgcagctgtc tgaaaccaac 720
aggaccctct ttctatattg tgtcacaaag tatactgcag gaccctatga atgtgaaata 780
tggaaactcag ggagtgccag ccgcagtgc cagtcaccc tgaatctcct ccatgggtcca 840
gacctccca gaattttccc ttcagtcacc tcttactatt caggagagaa cctcgacttg 900
tcctgcttgc caaactctaa cccaccagca cagtattctt ggacaattaa tgggaagttt 960
cagctatcag gacaaaagct ctttatccct cagattactc caaagcataa tgggctctat 1020
gcttgctctg ctcgtaactc agccactggc gaggaagct ccacatcctt gacaatcaga 1080
gtcattgtct ctccaggatt aggaactttt gctttcaata atccaacgta gcagccgtga 1140
tgtcattttt gtattttcagg aagactggca ggagatttat ggaaaagact atgaaaagga 1200
ctcttgaata caagtctctg ataacttcaa gatcatacca ctggactaag aactttcaaa 1260
attttgatga acaggctgat accttcagta aattcaagac aaagaagaaa agaactccat 1320
ttcattggac taaataacaa aaggataatg ttttcataat tttttatttg aaaatgtgct 1380
gattttttga atgtttttat ctcagattt atgaattttt ttcttcagca attggtaaag 1440

```

WO 02/062841

PCT/US02/02616

```
tataacttttg taaacaaaaa ttgaaacatt tgctttttgct ctctgagtgc cccagaatgg 1500
gaatctattc atgaatatcc atatgtttat ggtactaaag ttatttgcac aagtttaaaa 1560
aa 1562
```

<210> 38
 <211> 1110
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4173218CB1

```
<400> 38
gggtgggggc ggacccacag aacgaccgac ggaccgaggg ttcgagggag ggacacggac 60
caggaacctg agctaggtca aagacgcccc ggccaggtgc cccgtcgcag gtgcccctgg 120
ccggagatgc ggtaggaggg gcgagcgcga gaagccccct cctcggcgct gccaaacccgc 180
caccagccc atggcgaacc ccgggctggg gctgcttctg gcgctgggccc tgccgttcct 240
gctggccccg tggggccgag cctgggggca aatacagacc acttctgcaa atgagaatag 300
cactgttttg ccttcaccca ccagctccag ctccgatggc aacctgcgtc cggaagccat 360
cactgctatc atcgtggtct tctccctctt ggctgccttg ctccctggctg tggggctggc 420
actgttggtg cggaagcttc gggagaagcg gcagacggag ggcacctacc ggcccagtag 480
cgaggagcag gtgggtgccc gcgtgccacc gacccccaac ctcaagttgc cgccggaaga 540
gcggctcatc tgaacgctgg ggcctgctgc agccaccaac actgccagg actgcgggtt 600
gctggcttgt acaccgcagc tgccaccgag acaccagcct ctgatggctc aggaggactt 660
gtggggagag gctgggggca cccatgtggt gggctctgtg cagcatgttg cctctgcttg 720
gctgtgcctg cagctcaggg tgctggggct cgggaccac cccctgctt gcggaaccaa 780
cttttctctg tgtgtccagc agggcccccaca acccctctc ctttctttca gttctcccat 840
gcagccgagg cccgggcccc tcaggactcc aaggagacgg tgcagggctg cctgcccac 900
taggtcccc ctccctgcac tgtctccctt cattgctgtg tgaccttggg gaaaggcagt 960
gccctctctg ggcagtcaga tccaccaggt gcttaatagc agggagaag gtacttcaaa 1020
gactctgccc ctgaggtcaa gagaggatgg ggctattcac ttttatatat ttatataaaa 1080
ttagtagtga gatgtaaaaa aaaaaaaaaa 1110
```

<210> 39
 <211> 1412
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5014679CB1

```
<400> 39
ggatttgtgt agattgtatg catatacaac atcattttat tccagagact tgagcatctg 60
cgaattttgga tatccaaagg aagtcctaga accagttccc gacaaatacc aagagatgac 120
tttacttttag acctaagatc ttgctagaca gcctttgaga ggacgaagga taactagagc 180
agaagcaa at cgagcgaag tagactacaa caatcaagtc aaaaaacata aagctcctat 240
caacacctca agattttaga aaagcctctt aacgggcaat tttcttttaa ataaagactg 300
ggtcccacta tgtagccag gctgggtcaca aactcctgga ctcaagtgat cctcctgcct 360
tggcctccta aagtgcctggg attacaggca tgagccactg tgcccggcct gggcagtcga 420
ttttcaggaa gccacttaat gttcagattc ttttcttctc tcccaggact ttgcagatgc 480
tgctgttccc tctgcctgga gatgttcatc tggcctgaat ccaaaagtct gtccaagtta 540
cttctctttt ccatgacaga gttacaacta tttgtctgtc ccccttcccc attagagtat 600
atgggtcccca aaggaaagcc ccatgccatt cactcaccac ctaatcatcc ccaacaacta 660
gtgggcaaac atgaaacact tttgttgaat aatgtgtttc ctttagccca gggattacaa 720
agtcaaaggc ttacaagggg gaaggactag caggtgtgct cggaggggac tctagccaac 780
tgagagagtaa ctgggtgagaa gacaatgaca tgctcttagg gtgtctgcta ctcagatcct 840
ctcagtttaag tcccgtgcac aaagggtggg caggtctact gttgccagat cttcagatca 900
ttcaaaagtt acccgaattc tggattttca tgcaaatctc ccaattttta aaaatgttgg 960
catcaaattc agaatgtttt aaccagccac actgtgtggg ccaggga aaa caggcccaa 1020
ggccaggtga agctgcagtt ttcaactgct gccttggcct tcccagagc cccttaagtc 1080
ctaagaaggc tcagtcaaca cactggcact cgtttcatcc cactggttga gaaagcaagt 1140
```

WO 02/062841

PCT/US02/02616

cttagcaaag	taatggggga	gtcttaccaa	aactgcaaag	gaaaaaaata	tgataccatg	1200
aaattagaaa	ttcactgcat	ttgaaatgaa	aatgacttat	ttctgtactt	ctataactca	1260
caaatttagc	agacatcatt	tagacatfff	tgtagtgcaa	cagattttaa	tctggaattt	1320
catggtttgt	aaatcactat	gtgatactag	ctataacccat	gagcaatatt	acgggcagta	1380
aatgttgatt	tctatttcgt	ttcttgaaaa	aa			1412

<210> 40

<211> 1060

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7487510CB1

<400> 40

agccccctaac	cgcgagtgat	ccgccagcct	cggcctcccc	aggtgccggg	attgcagacg	60
gagtctcgtt	cactcagtgc	tcaatgggtg	ccaggctgga	gtgcagtggc	gtgatctcgg	120
ctcgctacaa	cctccacctc	ccagccgcct	gccttgccct	cccaaagagc	cgagattgca	180
gcctctgccc	ggccgcccacc	ccgtctggga	agtgaggagc	gtctctgcct	ggccgcccac	240
cgtctgggat	gtgaggagcg	tctctgccc	gctgcccagt	ctgggaagt	aggagcgcct	300
cctccccggc	gccatcccgt	ctaggaagt	aggagcgtct	ctgcccggcc	gcccacgtc	360
tgagatgtgg	ggagcgccac	tgccccggcc	ccccgtccgg	gaggtgcctc	ggcttccgca	420
tctgtcgtat	gacccgtgat	ctctgggaag	ccacacagct	caaggtcttg	gggcacgtca	480
tgagggtctc	ggaagcgtca	cttaccctgt	ccctgtcggc	atcatcatcg	tcagcatcgt	540
ttaagaatca	agccctgttt	tcttcttctg	accactgggt	ggctccgcag	aattgggtct	600
gtgattatcg	cgctctcaaa	ggcggccttg	gggtttgggt	gaacagtatg	ataatgctgg	660
tttgtcgtag	gtcaaaaaca	gcaaattatc	tgcaatgtca	tgtgggttcta	cctaattgctt	720
gcggtgtccc	tgccctgggc	tgtttccctt	cggcttcctc	tcagcgaatc	acgaacacat	780
tccacggact	cacctccttg	gaagcctttt	ggattctctg	cgcagcccaa	gctgcccggg	840
atctgggagg	ccaggctgag	tctatggccc	cggagcccgc	ccggacttgc	cactggagac	900
ctggggccaa	gggcccaccc	gagctgggaa	gagagggcta	gaaagagagc	attagaatcg	960
aggggctggg	tgccggaggct	cacgcctgtc	atcccagcac	tttgggagcc	gagggagatg	1020
gatcacctga	ggtaggaat	tcaagagcag	cctggccaac			1060

<210> 41

<211> 2268

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2682619CB1

<400> 41

ggaagttccg	gttcgcagg	ggtggggagt	gttggttaacc	ggagggggcag	ccgcagtcgc	60
gcggatttag	cgggctcgcg	gcgctgggtt	cctggtctcc	gggcccaggca	atgttccgca	120
cggcagtgat	gatggcggcc	agcctggcgc	tgaccggggc	tgtgggtggc	cacgcctact	180
acctcaaaca	ccagttctac	cccactgtgg	tgtacctgac	caagtccagc	cccagcatgg	240
cagtcctgta	catccaggcc	tttgtccttg	tcttccctct	gggcaagggt	atgggcaagg	300
tgttcttttg	gcaactgagg	gcagcagaga	tgagacacct	tctggaacgt	tcctgggtacg	360
ccgtcacaga	gacttgtctg	gccttcaccg	tttttcggga	tgacttcagc	ccccgctttg	420
ttgcactctt	cactcttctt	ctcttcctca	aatgtttcca	ctggctggct	gaggaccgtg	480
tggactttat	ggaacgcagc	cccaacatct	cctggctctt	tcactgccgc	attgtctctc	540
ttatgttcct	cctgggcatc	ctggacttcc	tcttcgtcag	ccacgcctat	cacagcatcc	600
tgaccctgtg	ggcctctgtg	cagctgggtg	ttggctttga	gtatgccatc	ctgatgacga	660
tggtgctcac	catcttcctc	aagtatgtgc	tgactccgt	ggacctccag	agtgagaacc	720
cctgggacaa	caaggctgtg	tacatgctct	acacagagct	gtttacaggc	ttcatcaagg	780
ttctgctgta	catggccttc	atgaccatca	tgatcaagg	gcacaccttc	ccactctttg	840
ccatccggcc	catgtacctg	gccatgagac	agttcaagaa	agctgtgaca	gatgccatca	900
tgtctcgcgc	agccatccgc	aacatgaaca	ccctgtatcc	agatgccacc	ccagaggagc	960
tccaggcaat	ggacaatgtc	tgcacatct	gccgagaaga	gatgggtgact	ggtgccaaga	1020
gactgccctg	caaccacatt	ttccatacca	gctgcctgcg	ctcctgggtc	cagcggcagc	1080

WO 02/062841

PCT/US02/02616

```

agacctgccc cacctgccgt atggatgtcc ttcgtgcac gctgccagcg cagtcaccac 1140
caccgccgga gcctgcggat caggggccac cccctgcccc ccacccccca ccactcttgc 1200
ctcagccccc caacttcccc cagggcctcc tgcctccttt tcctccaggc atgttccac 1260
tgtggccccc catgggcccc tttccacctg tcccgccctc cccagctca ggagaggctg 1320
tggctcctcc atccaccagt gcagcagccc tttctcggcc cagtggagca gctacaacca 1380
cagctgctgg caccagtgtt actgctgctt ctgccacagc atctggccca ggctctggct 1440
ctgccccaga ggctggccct gccctgggtt tcccttccc tcctccctgg atgggtatgc 1500
ccctgcctcc accctttgcc ttccccccaa tgctgtgcc ccctgcgggc tttgctgggc 1560
tgaccccaga ggagctacga gctctggagg gccatgagcg gcagcacctg gaggcccgcc 1620
tgacagagct gcgtaacatc cacacactgc tggacgccgc catgctgcag atcaaccagt 1680
acctcacctg gctggcctcc ttggggcccc cccggcctgc cacttcagtc aactccactg 1740
aggagactgc cactacagtt gttgctgctg cctcctccac cagcatccct agctcagagg 1800
ccacgacccc aacccagga gcctccccac cagcccctga aatggaaagg cctccagctc 1860
ctgagtcagt gggcacagag gagatgcctg aggatggaga gcccgatgca gcagagctcc 1920
gccggcgccg cctgcagaag ctggagtctc ctgttgccca ctgacactgc cccagcccag 1980
ccccagcctc tgctcttttg agcagccctc gctggaacat gtcctgccac caagtgccag 2040
ctccctctct gtctgcacca gggagtagta ccccagctc tgagaaagag gcggcatccc 2100
ctaggccaag tggaaagagg ctgggggttc catttgactc cagtcccagg cagccatggg 2160
gatctcgggt cagttccagc cttcctctcc aactcttcag ccctgtgttc tgctggggcc 2220
atgaaggcag aaggtttagc ctctgaagag cctcttcttc cccaacc 2268

```

<210> 42

<211> 821

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4582105CB1

<400> 42

```

gaggaagtc aatcgctgcc gcaggtaccg ccaatggctt ttggcggggg cgttcccca 60
ccctgccctc tctcatgacc ccgctccggg attatggccg ggactgggct gctggcgctg 120
cggacgctgc cagggcccag ctgggtgcga ggctcggggc ctcccgctgt gagccgctg 180
caggacgcgg ccgtggtgcg gcctggcttc ctgagcacgg cagaggagga gacgctgagc 240
cgagaactgg agcccagact gcgcgcgcgc cgctacgaat acgatcactg ggacgcggcc 300
atccacggct tccgagagac agagaagtcg cgctggtcag aagccagccg ggccatcctg 360
cagcgcgtgc agggcgccgc ctttggcccc ggccagacc tgcctctctc cgtgcacgtg 420
ctggacctgg aagcccgccg ctacatcaag cccacgctgg acagcatcaa gttctgcggg 480
gccaccatcg ccggcctgtc tctcctgtct cccagcgcta tgccgctggg gcacaccag 540
gagccggggg agtggctgga actcttgctg gagccgggct ccctctacat ccttaggggc 600
tcagcccggt atgacttctc ccatgagatc cttcgggatg aagagtcctt ctttggggaa 660
cgccggattc cccggggccg gcgcctctcc gtgatctgcc gctccctccc tgagggcag 720
gggccagggg agtctggaca gccgccccca gcctgctgac cccagcttt ctacagacac 780
cagatttgtg aataaagttg gggaaatggac agcctaaaaa a 821

```

<210> 43

<211> 1325

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 931619CB1

<400> 43

```

ccgggaaaag cagccggagc ccccgccgcc cctgccgcag cgcggggcgg cagcgcgcag 60
cccggcaccc gcagcctgca gcctgcagcc cgcagcccgc agcccgagc cagatcgccg 120
gctcagaccg aacccgactc gaccgccgcc cccagccagg cgccatgctg ccgcttctgc 180
tgggcctgct gggcccagcg gcctgctggg ccctggggcc gacccccggc ccgggatact 240
ctgagctgcg ctccggcctt tcggcgccac gcaccacccc cctggagggc acgtcggaga 300
tggcggtgac cttcgacaag gtgtacgtga acatcggggg cgacttcgat gtggccaccg 360
gccagtttcg ctgccgcgtg cccggcgccct acttcttctc cttcacggct ggcaaggccc 420

```

WO 02/062841

PCT/US02/02616

```

cgcacaagag cctgtcgggtg atgctgggtgc gaaaccgcga cgaggtgcag gcgctggcct 480
tcgacgagca gcggcggcca ggcgcgcggc gcgcagccag ccagagcgcc atgctgcagc 540
tcgactacgg cgacacagtg tggctgcggc tgcttggcgc cccgcagtac gcgctaggcg 600
cgcccggcgc caccttcagc ggctacctag tctacgccga cgccgacgct gacgcgcctg 660
cgcgcgggcc gcccgcgccc cccgagccgc gctcggcctt ctcggcggcg cgcacgcgca 720
gcttgggtggg ctcggaacgt ggccccgggc cgcggcacca accactcgcc ttcgacaccg 780
agttcgtcaa cattggcggc gacttcgacg cggcggccgg cgtgttccgc tgccgtctgc 840
ccggcgccta cttcttctcc ttcacgctgg gcaagctgcc gcgtaagacg ctgtcggtta 900
agctgatgaa gaaccgcgac gaggtgcagg ccatgattta cgacgacggc gcgtcgcggc 960
gccgcgagat gcagagccag agcgtgatgc tggccctgcg gcgcggcgac gccgtctggc 1020
tgctcagcca cgaccacgac ggctacggcg cctacagcaa ccacggcaag tacatcacct 1080
tctccggctt cctgggtgtac cccgacctcg cccccgcgc cccgcggggc ctcggggcct 1140
cggagctact gtgagccccg ggccagagaa gagcccggga gggccagggg cgtgcatgcc 1200
aggccggggc cgaggctcga aagtcccgcg cgagcgccac ggcctccggg cgcgcctgga 1260
ctctgccaat aaagcggaaa gcgggcacgc gcagcgcccg gcagcccagg aaaaaaaaaa 1320
aaaaa                                           1325

```

<210> 44

<211> 2030

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2155025CB1

<400> 44

```

gggtgtgtct gtggcgtctc atggcaggag gctcagccac gacctggggg taccctgtgg 60
ccctgctact gctggtcgcc accctggggc tgggtagggt gctccagccc gacctgggcc 120
tcccaggcct ccggcacagc tacgactgtg ggatcaaggg aatgcagctg ctgggtgttc 180
ccaggccagg ccagactctc cgcttcaagg tgggtgatga atttgggaac cgatttgatg 240
tcaacaactg ctccatctgc taccactggg tcacctccag gccgcaggag cctgcagtct 300
tctcggccga ttacagaggc tgccacgtgc tggagaaggt aggggatggg cgtttccacc 360
tgagggtgtt catggaggct gtgctgcca atggctcgtg ggatgtggca caagacgcta 420
ctctgatctg tcccaaacct gacctctccc ggactctgga ctcccagctg gcaccacccg 480
ccatgttctc tgtctcaatc ccacaaacct tttccttctt cccacactct ggccatacct 540
cccaaggctc tggccatgcc tttcccagcc cactggacct agggcacagc tctgtccacc 600
caaccctgc tttaccatcc cctggacctg gacctacct cgccaccctg gctcaacccc 660
actggggcac cttggaacac tgggatgtga acaaacgaga ttacataggt acccacctga 720
gccaggagca gtgccagggt gcctcagggc acctccctg catcgtgaga agaacttcaa 780
aagaagcctg tcagcaggct ggctgctgct atgacaacac cagagagggt ccctgttact 840
atggcaacac agctactgtc cagtgcctca gagatggcta cttcgtcctc gtgggtgtcc 900
aagaaatggc cttgacacac aggatcacac tggccaacat ccacctggcc tatgccccca 960
ccagctgctc cccaacacag cacacggaag ctttcgtggg cttctacttc cctctcacc 1020
actgtggaac cacaatgcag gtggctggcg accagctcat ctatgagaac tggctgggtg 1080
ctggcatcca catccaaaag gggccacagg gttccatcac gcgggacagc accttccagc 1140
ttcatgtgcg ctgtgtcttc aacgccagtg acttcttgc cattcaggca tccattttcc 1200
caccctcatc gctgtctcct atgacctcag cgggcccct gcggcttgag ctgcggattg 1260
ccaaagacga gaccttcagc tcgtactatg gggaggatga ctatcccatc gtgaggctgc 1320
tccgagaacc agtccatgtg gaggtccggc ttctgcagag gacagacccc aacctgggtc 1380
tgctgctgca ccagtgtgg ggcgtccca gtgccaacct cttccagcag cccagtggtc 1440
ccatcctgtc agacggatgc cctttcaagg gcgacagcta cagaacccaa atggtagcct 1500
tggacggggc cacaccttct cagtcgcact accagcgatt cactgttgct accttcgccc 1560
tcctggactc aggtcccag agagccctca gaggactggg ttacttgttc tgcagcacct 1620
ctgcctgcc aacctcagg ctggagactt gctccactgc atgtagcact ggcactacaa 1680
gacagcgacg atcctcaggt caccgtaatg aactgccag gcccaggac atcgtgagct 1740
ctccggggcc agtgggcttt gaggattctt atgggcagga gccacactt gggcccacag 1800
actccaatgg gaactccagc ctgagacctc tcctttgggc ggtccttttg ctgccagctg 1860
ttgccctgg ccttgggttt ggtgtctttg tgggcctgag ccagacctgg gcccagaagc 1920
tctgggaaag caacagacag tgaatgggccc caataaacia tcatttcaaa cctactgaaa 1980
ccagggtgtg agaagttatt tgtgacgact agaacagaac tatttcttat 2030

```

<210> 45

WO 02/062841

PCT/US02/02616

<211> 1307

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7640495CB1

<400> 45

```

cgggcccggga gcgagcggcg gcggcggcg gggcggcacc atgggcccggg cccggcgctt 60
ccagtggccg ctgctgctgc tgtgggcccgc cgcggcgggg ccaggggagc gacaggaagt 120
acagacagag aacgtgacag tggctgaggg tgggggtggct gagatcacct gccgtctgca 180
ccagtatgat ggggtccatag ttgtcatcca gaaccagcc cggcagaccc tcttcttcaa 240
tggcaccctgt gccttgaagg atgagcggtt ccagcttgag gagttctccc cacgcccgggt 300
gcggatccgg ctctcagatg cccgcctgga ggacgagggg ggctatctct gccagctcta 360
cacagaagac acccaccacc agattgccac gctcacggta ctagtggccc cagagaatcc 420
tgtgggtggag gtccgggagc aggcggtaga gggcggcgag gtggagctca gctgcctcgt 480
tccgcgggtcc cgtccggctg ccaccctgcg ctggtaaccg gaccgcaagg agctgaaagg 540
agtgagcagc agccaggaaa atggcaaggc ctggagcgtg gcaagcacag tacggtttctg 600
tgtggaccgt aaggacgacg gtggatatcat catctgtgag gcgcagaacc aggcgctgcc 660
ctccggacac agcaagcaga cgcagtacgt gctggatgtg cagtactccc ccacggcccg 720
gattcatgcc tcccaagctg tgggtgaggga gggagacacg ctgggtgttg cgtgtgctgt 780
cacggggaac cccaggccaa accagatccg ctggaaccgc gggaatgagt ctttgccgga 840
gagggcggag gccgtgggag agacgctcac gctgccgggt ctggatatcc cggataacgg 900
cacctacact tgcgaggcgt ccaataagca cggccatgcg agggcgctct acgtacttgt 960
ggtctacgac cctgggtgcg tggtagaggc tcagacgtcg gttccctatg ccattgtggg 1020
cggcatcctg gcgctgctgg tgtttctgat catatgtgtg ctagtgggca tgggtctggtg 1080
ctcgggtacg cagaagggtt cctatctgac ccacgaagcc agtggccttg atgaacaggg 1140
agaagcaaga gaagccttcc tcaatggcag cgacggacac aagaggaaag aggaattctt 1200
catctgacct tatccccacc ccaggcctag gcctgggcct gggctggggg cccccccact 1260
gccagctgca agggaaccag caaagacatt taccagagtc tgggatg 1307

```

<210> 46

<211> 768

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5960119CB1

<400> 46

```

atgctgctgg tgctgggtgg gctcatcccc gtgctgggtga gctcgggccc cccggaaggc 60
cactatgaga tgctgggcac ctgccgcagt gtgtgcgacc cctacggggg caccaaggcg 120
cccagcaccg ctgccacgcc cgaccgcggc ctcatgcagt cctgcccac cttcatccag 180
ggccccaagg gcgaggccgg caggcccggg aaggcgggtc cgcgcggggc ccccgagag 240
cccggggcac ccggcccat ggggcccccg ggcgagaagg gcgagccggg ccgccaaggc 300
ctgccggggc cgcccggggc gcccggcctg aacgcggccg gggccatcag cgccgccacc 360
tacagcacgg tgcccaagat cgccttctac gccggcctca agcggcagca tgaaggctac 420
gaggtgctca agttcgacga cgtggtcacc aacctcgga accactacga cccaccacc 480
ggcaagtcca cctgctccat cccgggcata tacttcttca cctaccacgt cctgatgcgc 540
ggaggggacg gcaccagcat gtgggctgat ctctgcaaaa acaaccaggc gcgtgctagt 600
gcaattgcc aagatgctga tcagaattac gactatgcca gtaacagtgt ggttcttcat 660
ttggagccgg gagatgaagt ctatatcaaa ttagatggcg ggaaagccca tggaggaaac 720
aacaacaaat acagcacgtt ttctggattt attatttatg ctgactga 768

```

<210> 47

<211> 782

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7500143CB1

<400> 47

```

gctttctcct gagccgttgg agggagccgg agcgcttctc ccgagttggg gatagattgg 60
tgggtcatcca acatgcagaa atgaatgagc agtgaaaagc agcagagccg atgggtcatg 120
aggatgtaag tgcgtttgaa ggcttccaca ccctctactc caggacagaa tcatgaataa 180
actggaggat aagcaggacc agatgatacc atgaagagaa gtttacaggc cctctattgc 240
caactgttaa ctgtcctgct gacgggtgtgc tgcataaaga ggaagaagaa gaccgccaac 300
ccggagaaca acctgagcta ctggaacaac accatcacca tggactactt caacaggcat 360
gctgtggagc tgcccaggga gatccagtcc cttgaaacct ctgaggacca gctctcagag 420
ccccgctccc cagccaatgg cgactataga gacactggga tggtccttgt taacccttc 480
tgtcaagaaa cactgtttgt gggaaacgat caagtatctg agatctaact acagcaggca 540
tcactttgcc attccgtatt ttctgtctct aaattataaa tatacaaata tatatattat 600
aaatataacc tttgtgtaac cctgacttaa tgagaaacat tttcagcttt ttttcctatg 660
aattgtcaac atctttttta caagtgtggg ttaaaaaaaaa aaaaacttta cagaatgatc 720
tgtggcttta taaaataaag gtattttctaa gcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 780
tt
782

```

<210> 48

<211> 893

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7503605CB1

<400> 48

```

gaggggaagtc aatcgctgcc gcaggtaccg ccaatggctt ttggcggggg cgttcccca 60
ccctgccctc tctcatgacc ccgctccggg attatggccg ggactgggct gctggcgctg 120
cggacgctgc cagggcccag ctgggtgcga ggctcgggcc cttccgtgct gagccgcgc 180
tacgaatacg atcactggga cgcggccatc cacggcttcc gagagacaga gaagtcgcgc 240
tggtcagaag ccagccgggc catcctgcag cgcgtgcagg cggccgcctt tggccccggc 300
cagaccctgc tctcctccgt gcacgtgctg gacctggaag cccgcggcta catcaagccc 360
cacgtggaca gcatcaagtt ctgcggggcc accatcgccg gcctgtctct cctgtctccc 420
agcgttatgc ggctgggtgca caccaggag ccgggggagt ggctggaact cttgctggag 480
ccgggctccc tctacatcct taggggctca gcccgttatg acttctccca tgagatcctt 540
cgggatgaag agtccttctt tggggaacgc cggattcccc ggggccggcg catctccgtg 600
atctgccgct ccctccctga gggcatgggg ccaggggagt ctggacagcc gccccagcc 660
tgctgacccc cagctttcta cagacaccag atttgtgaat aaagtggggg aatggacagc 720
ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 780
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaagg gggggggcgc ccccaaaagg 840
gagggccgcc ccccgcccgg ggaatttttt cgcccggccg ggcccccgcg gga 893

```

What is claimed is:

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

- 5 a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24,
- 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:1-20,
- 10 c) a polypeptide comprising a naturally occurring amino acid sequence at least 97% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:21-22,
- d) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:23-24,
- 15 e) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, and
- f) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

20 2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

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

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

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:25-48.

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

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

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

9. A method of 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. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

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

12. 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:25-48,
- 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:25-44 and SEQ ID NO:47-48,
- c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 97% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-46,
- d) a polynucleotide complementary to a polynucleotide of a),
- e) a polynucleotide complementary to a polynucleotide of b),
- f) a polynucleotide complementary to a polynucleotide of c), and
- g) an RNA equivalent of a)-f).

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

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 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.

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

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 10 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.

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

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

20 19. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 17.

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

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

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

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

35

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide 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.

5

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

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

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

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- 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.

15

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

- 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 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.

25

30

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

35

- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

- 5 29. A method of assessing toxicity of a test compound, the 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
10 at least 20 contiguous nucleotides of a polynucleotide of claim 12 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 12 or fragment thereof,
 - c) quantifying the amount of hybridization complex, and
 - d) comparing the amount of hybridization complex in the treated biological sample with
15 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.

30. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:

- 20 a) combining the biological sample with an antibody of claim 11, under conditions
suitable for the antibody to bind the polypeptide and form an antibody:polypeptide
complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the
presence of the polypeptide in the biological sample.

25

31. The antibody of claim 11, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- 30 d) a F(ab')₂ fragment, or
- e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

- 35 33. A method of diagnosing a condition or disease associated with the expression of SECP in

a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

5

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

10 36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- 15 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 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24.

20 37. A polyclonal antibody produced by a method of claim 36.

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

25 39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-24, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- 30 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 comprising an amino acid sequence selected from the group consisting of
- 35 SEQ ID NO:1-24.

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

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

5 42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

10

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24 in a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 15 b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-24 from a sample, the method comprising:

- a) incubating the antibody of claim 11 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 comprising an amino acid sequence selected from the group consisting of SEQ ID
25 NO:1-24.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

30 47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a
35 hybridization complex, and

- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first
5 oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
10 completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
15 completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a
20 nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

25

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at
30 another distinct physical location on the substrate.

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

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

35

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

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

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

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

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

10

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

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

15

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

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

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

20

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

25

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

30

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

35

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

5 78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

10 80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:25.

81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:26.

15 82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:27.

83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:28.

20 84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.

25 85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:30.

86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.

30 87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:32.

88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:33.

35

89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:34.

5 90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:35.

91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:36.

10 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:37.

15 93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:38.

94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:39.

20 95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:40.

96. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:41.

25 97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:42.

98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:43.

30 99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:44.

35 100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:45.

101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.

102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:47.

103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:48.