

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/26950 A2

(51) International Patent Classification⁷: **C12N 9/00**

(21) International Application Number: PCT/US01/30424

(22) International Filing Date:
28 September 2001 (28.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/236,523 29 September 2000 (29.09.2000) US
60/238,481 6 October 2000 (06.10.2000) US
60/244,025 27 October 2000 (27.10.2000) US
60/246,001 3 November 2000 (03.11.2000) US
60/247,931 9 November 2000 (09.11.2000) US
60/249,639 16 November 2000 (16.11.2000) US
60/252,819 21 November 2000 (21.11.2000) US

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LAL, Preeti, G.** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue, #1, Redwood City, CA 94025 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94023 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee, #5, San Francisco, CA 94107 (US). **BAUGHN, Mariah, R.** [US/US];

14244 Santiago Road, San Leandro, CA 94577 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **LU, Yan** [CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33690 Mello Way, Fremont, CA 94555 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **DING, Li** [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US).

(74) Agents: **HAMLET-COX, Diana Incyte Genomics, Inc.** et al.; 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, ZW.

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

Published:

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

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

(54) Title: TRANSFERASES

(57) Abstract: The invention provides human transferases (TRNFR) and polynucleotides which identify and encode TRNFR. 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 TRNFR.



WO 02/26950 A2

TRANSFERASES

TECHNICAL FIELD

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

BACKGROUND OF THE INVENTION

Transferases are enzymes that catalyze the transfer of molecular groups. The reaction may involve an oxidation, reduction, or cleavage of covalent bonds, and is often specific to a substrate or to particular sites on a type of substrate. Transferases participate in reactions essential to such functions as synthesis and degradation of cell components, regulation of cell functions including cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Transferases are involved in key steps in disease processes involving these functions. Transferases are frequently classified according to the type of group transferred. For example, methyl transferases transfer one-carbon methyl groups, amino transferases transfer nitrogenous amino groups, and similarly denominated enzymes transfer aldehyde or ketone, acyl, glycosyl, alkyl or aryl, isoprenyl, saccharyl, phosphorous-containing, sulfur-containing, or selenium-containing groups, as well as small enzymatic groups such as Coenzyme A.

Acyl transferases include peroxisomal carnitine octanoyl transferase, which is involved in the fatty acid beta-oxidation pathway, and mitochondrial carnitine palmitoyl transferases, involved in fatty acid metabolism and transport. Choline O-acetyl transferase catalyzes the biosynthesis of the neurotransmitter acetylcholine. N-acyltransferase enzymes catalyze the transfer of an amino acid conjugate to an activated carboxylic group. Endogenous compounds and xenobiotics are activated by acyl-CoA synthetases in the cytosol, microsomes, and mitochondria. The acyl-CoA intermediates are then conjugated with an amino acid (typically glycine, glutamine, or taurine, but also ornithine, arginine, histidine, serine, aspartic acid, and several dipeptides) by N-acyltransferases in the cytosol or mitochondria to form a metabolite with an amide bond. One well-characterized enzyme of this class is the bile acid-CoA:amino acid N-acyltransferase (BAT) responsible for generating the bile acid conjugates which serve as detergents in the gastrointestinal tract (Falany, C. N. et al. (1994) J. Biol. Chem. 269:19375-9; Johnson, M. R. et al. (1991) J. Biol. Chem. 266:10227-33). BAT is also useful as a predictive indicator for prognosis of hepatocellular carcinoma patients after partial hepatectomy (Furutani, M. et al. (1996) Hepatology 24:1441-5).

N-acetyltransferases are cytosolic enzymes which utilize the cofactor acetyl-coenzyme A (acetyl-CoA) to transfer the acetyl group to aromatic amines and hydrazine containing compounds. In humans, there are two highly similar N-acetyltransferase enzymes, NAT1 and NAT2; mice appear to have a third form of the enzyme, NAT3. The human forms of N-acetyltransferase have independent regulation (NAT1 is widely-expressed, whereas NAT2 is in liver and gut only) and overlapping substrate preferences. Both enzymes appear to accept most substrates to some extent, but NAT1 does prefer some substrates (para-aminobenzoic acid, para-aminosalicylic acid, sulfamethoxazole, and sulfanilamide), while NAT2 prefers others (isoniazid, hydralazine, procainamide, dapsone, aminogluthethimide, and sulfamethazine). A recently isolated human gene, tubedown-1, is homologous to the yeast NAT-1 N-acetyltransferases and encodes a protein associated with acetyltransferase activity. The expression patterns of tubedown-1 suggest that it may be involved in regulating vascular and hematopoietic development (Gendron, R.L. et al. (2000) Dev. Dyn. 218:300-315).

Lysophosphatidic acid acyltransferase (LPAAT) catalyzes the acylation of lysophosphatidic acid (LPA) to phosphatidic acid. LPA is the simplest glycerophospholipid, consisting of a glycerol molecule, a phosphate group, and a mono-saturated fatty acyl chain. LPAAT adds a second fatty acyl chain to LPA, producing phosphatidic acid (PA). PA is the precursor molecule for diacylglycerols, which are necessary for the production of phospholipids, and for triacylglycerols, which are essential biological fuel molecules. In addition to being a crucial precursor molecule in biosynthetic reactions, LPA has recently been added to the list of intercellular lipid messenger molecules. LPA interacts with G protein-coupled receptors, coupling to various independent effector pathways including inhibition of adenylate cyclase, stimulation of phospholipase C, activation of MAP kinases, and activation of the small GTP-binding proteins Ras and Rho. (Moolenaar, W.H. (1995) J. Biol. Chem 28-:12949-12952.) The physiological effects of LPA have not been fully characterized yet, but they include promoting growth and invasion of tumor cells. PA, the product of LPAAT, is a key messenger in a common signaling pathway activated by proinflammatory mediators such as interleukin-1 β , tumor necrosis factor α , platelet activating factor, and lipid A. (Bursten, S.L. et al. (1992) Am. J. Physiol. 262:C328-C338; Bursten S.L. et al. (1991) J. Biol. Chem. 255:20732-20743; Kester, M. (1993) J. Cell Physiol. 156:317-325.) Thus, LPAAT activity may mediate inflammatory responses to various proinflammatory agents.

Aminotransferases comprise a family of pyridoxal 5'-phosphate (PLP) -dependent enzymes that catalyze transformations of amino acids. Amino transferases play key roles in protein synthesis and degradation, and they contribute to other processes as well. For example, GABA aminotransferase (GABA-T) catalyzes the degradation of GABA, the major inhibitory amino acid neurotransmitter. The activity of GABA-T is correlated to neuropsychiatric disorders such as

alcoholism, epilepsy, and Alzheimer's disease (Sherif, F.M. and Ahmed, S.S. (1995) Clin. Biochem. 28:145-154). Other members of the family include pyruvate aminotransferase, branched-chain amino acid aminotransferase, tyrosine aminotransferase, aromatic aminotransferase, alanine:glyoxylate aminotransferase (AGT), and kynurenine aminotransferase (Vacca, R.A. et al. (1997) J. Biol. Chem. 272:21932-21937). Kynurenine aminotransferase catalyzes the irreversible transamination of the L-tryptophan metabolite L-kynurenine to form kynurenic acid. The enzyme may also catalyzes the reversible transamination reaction between L-2-aminoadipate and 2-oxoglutarate to produce 2-oxoadipate and L-glutamate. Kynurenic acid is a putative modulator of glutamatergic neurotransmission, thus a deficiency in kynurenine aminotransferase may be associated with pleiotropic effects (Buchli, R. et al. (1995) J. Biol. Chem. 270:29330-29335). Defects in AGT are the cause of primary hyperoxaluria type I (PH), a potentially lethal autosomal recessive disorder characterized by an increased urinary excretion of calcium oxalate, leading to recurrent urolithiasis, nephrocalcinosis, and accumulation of insoluble oxalate throughout the body (Cochat, P. et al. (1999) Eur. J. Pediatr. 158 Suppl 2:S75-S80).

Glycosyl transferases include the mammalian UDP-glucouronosyl transferases, a family of membrane-bound microsomal enzymes catalyzing the transfer of glucouronic acid to lipophilic substrates in reactions that play important roles in detoxification and excretion of drugs, carcinogens, and other foreign substances. Another mammalian glycosyl transferase, mammalian UDP-galactose-ceramide galactosyl transferase, catalyzes the transfer of galactose to ceramide in the synthesis of galactocerebrosides in myelin membranes of the nervous system. Galactosyltransferases are a subset of glycosyltransferases that transfer galactose (Gal) to the terminal N-acetylglucosamine (GlcNAc) oligosaccharide chains that are part of glycoproteins or glycolipids that are free in solution (Kolbinger, F. et al. (1998) J. Biol. Chem. 273:433-440; Amado, M. et al. (1999) Biochim. Biophys. Acta 1473:35-53). β 1,3-galactosyltransferases form Type I carbohydrate chains with Gal (β 1-3)GlcNAc linkages. Known human and mouse β 1,3-galactosyltransferases appear to have a short cytosolic domain, a single transmembrane domain, and a catalytic domain with eight conserved regions. (Kolbinger, F. supra and Hennet, T. et al. (1998) J. Biol. Chem. 273:58-65). A variant of a sequence found within mouse UDP-galactose: β -N-acetylglucosamine β 1,3-galactosyltransferase-I region 8 is also found in bacterial galactosyltransferases, suggesting that this sequence defines a galactosyltransferase sequence motif (Hennet, T. supra). The human LARGE gene and its mouse ortholog both encode a predicted N-acetylglucosaminyltransferase protein that is much longer than other members of its family and contains putative coiled-coil domains. Mutations in this gene are associated with meningioma, and suggest that the mutant protein may be involved in altering the composition of gangliosides in tumor cells (Peyrard, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:598-603).

The sialyltransferases are required for the biosynthesis of gangliosides, glycopospholipids containing sialic acids in the carbohydrate moiety. Gangliosides play a number of critical roles in cellular processes including cell-cell interaction, cell adhesion, mediation of invasion of vectors, and protein targeting. The sialyltransferase ST6GalNAc V shows brain-specific expression and is
5 involved in the synthesis of GD1 α , a ganglioside important for communication between neuronal cells and their supportive cells in brain tissues. ST6GalNAc V also contains glutamine repeats which may be associated with neurodegenerative diseases (Okajima, T. et al. (1999) *J. Biol. Chem.* 274:30557-30562).

Methyl transferases are involved in a variety of pharmacologically important processes.

10 Nicotinamide N-methyl transferase catalyzes the N-methylation of nicotinamides and other pyridines, an important step in the cellular handling of drugs and other foreign compounds. Phenylethanolamine N-methyl transferase catalyzes the conversion of noradrenalin to adrenalin. 6-O-methylguanine-DNA methyl transferase reverses DNA methylation, an important step in carcinogenesis. Uroporphyrin-III C-methyl transferase, which catalyzes the transfer of two methyl groups from S-adenosyl-L-
15 methionine to uroporphyrinogen III, is the first specific enzyme in the biosynthesis of cobalamin, a dietary enzyme whose uptake is deficient in pernicious anemia. Protein-arginine methyl transferases catalyze the posttranslational methylation of arginine residues in proteins, resulting in the mono- and dimethylation of arginine on the guanidino group. Substrates include histones, myelin basic protein, and heterogeneous nuclear ribonucleoproteins involved in mRNA processing, splicing, and transport.
20 Protein-arginine methyl transferase interacts with proteins upregulated by mitogens, with proteins involved in chronic lymphocytic leukemia, and with interferon, suggesting an important role for methylation in cytokine receptor signaling (Lin, W.-J. et al. (1996) *J. Biol. Chem.* 271:15034-15044; Abramovich, C. et al. (1997) *EMBO J.* 16:260-266; and Scott, H.S. et al. (1998) *Genomics* 48:330-340).

25 Phospho transferases catalyze the transfer of high-energy phosphate groups and are important in energy-requiring and -releasing reactions. The metabolic enzyme creatine kinase catalyzes the reversible phosphate transfer between creatine/creatine phosphate and ATP/ADP. Glycocyamine kinase catalyzes phosphate transfer from ATP to guanidoacetate, and arginine kinase catalyzes phosphate transfer from ATP to arginine. A cysteine-containing active site is conserved in this family
30 (PROSITE: PDOC00103).

Prenyl transferases are heterodimers, consisting of an alpha and a beta subunit, that catalyze the transfer of an isoprenyl group. A particularly important member of this group is the Ras farnesyltransferase (FTase) enzyme, which transfers a farnesyl moiety from cytosolic farnesylpyrophosphate to a cysteine residue at the carboxyl terminus of the Ras oncogene protein.

This modification is required to anchor Ras to the cell membrane so that it can perform its role in signal transduction. FTase inhibitors have been shown to be effective in blocking Ras function, and demonstrate antitumor activity in vitro and in vivo (Buolamwini, J.K. (1999) Curr. Opin. Chem. Biol. 3:500-509). FTase shares structural similarity with geranylgeranyl transferase, or Rab GG

5 transferase. This enzyme prenylates Rab proteins, allowing them to perform their roles in regulating vesicle transport (Seabra, M.C. (1996) J. Biol. Chem. 271:14398-14404).

Saccharyl transferases are glycosylating enzymes involved in a variety of metabolic processes. Oligosaccharyl transferase-48, for example, is a receptor for advanced glycation endproducts.

Accumulation of these endproducts is observed in vascular complications of diabetes, macrovascular
10 disease, renal insufficiency, and Alzheimer's disease (Thornalley, P.J. (1998) Cell Mol. Biol. (Noisy-Le-Grand) 44:1013-1023).

Coenzyme A (CoA) transferase catalyzes the transfer of CoA between two carboxylic acids. Succinyl CoA:3-oxoacid CoA transferase, for example, transfers CoA from succinyl-CoA to a recipient such as acetoacetate. Acetoacetate is essential to the metabolism of ketone bodies, which
15 accumulate in tissues affected by metabolic disorders such as diabetes (PROSITE: PDOC00980).

NAD:arginine mono-ADP-ribosyltransferases catalyse the transfer of ADP-ribose from NAD to the guanido group of arginine on a target protein. Substrates for these enzymes have been identified in myotubes and activated lymphocytes, and include alpha integrin subunits. These proteins contain characteristic domains involved in NAD binding and ADP-ribose transfer, including a highly
20 acidic region near the carboxy terminus which is required for enzymatic activity (Moss, J. et al. (1999) Mol. Cell. Biochem. 193:109-113).

Phosphoribosyltransferases catalyze the synthesis of beta-n-5'-monophosphates from phosphoribosylpyrophosphate and an amine. These enzymes are involved in the biosynthesis of purine and pyrimidine nucleotides, and in the purine and pyrimidine salvage pathways. For example, the
25 enzyme hypoxanthine-guanine phosphoribosyltransferase (HGPRT) is a purine salvage enzyme that catalyzes the conversion of hypoxanthine and guanine to their respective mononucleotides. HGPRT is ubiquitous, is known as a 'housekeeping' gene, and is frequently used as an internal control for reverse transcriptase polymerase chain reactions. There is a serine-tyrosine dipeptide that is conserved among all members of the HGPRT family and is essential for the phosphoribosylation of purine bases
30 (Jardim, A. and Ullman, B. (1997) J. Biol. Chem. 272:8967-8973). A partial deficiency of HGPRT can lead to overproduction of uric acid, causing a severe form of gout. An absence of HGPRT causes Lesch-Nyhan syndrome, characterized by hyperuricaemia, mental retardation, choreoathetosis, and compulsive self-mutilation (Sculley, D.G. et al. (1992) Hum. Genet. 90:195-207). Many parasitic organisms are unable to synthesize purines de novo and must rely on the enzymes in salvage pathways

for the synthesis of purine nucleotides; thus these enzymes are potential targets for the treatment of parasitic infections (Craig, S. P., and Eakin, A. R. (2000) *J. Biol. Chem.* 275:20231-20234).

Transglutaminase (Tgases) transferases are Ca^{2+} dependent enzymes capable of forming isopeptide bonds by catalyzing the transfer of the γ -carboxy group from protein-bound glutamine to the ϵ -amino group of protein-bound lysine residues or other primary amines. TGases are the enzymes responsible for the cross-linking of cornified envelope (CE), the highly insoluble protein structure on the surface of the corneocytes, into a chemically and mechanically resistant protein polymer. Seven known human Tgases have been identified. Individual transglutaminase gene products are specialized in the cross-linking of specific proteins or tissue structures, such as factor XIIIa which stabilizes the fibrin clot in hemostasis, prostrate transglutaminase which functions in semen coagulation, and tissue transglutaminase which is involved in GTP-binding in receptor signaling. Four (TGases 1, 2, 3, and X) are expressed in terminally differentiating epithelia such as the epidermis. TGases are critical for the proper cross-linking of the CE as seen in the pathology of patients suffering from one form of the skin diseases referred to as congenital ichthyosis which has been linked to mutations in the keratinocyte transglutaminase (TG_K) gene (Nemes, Z. et al., (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:8402-8407; Aeschlimann, D. et al., (1998) *J. Biol. Chem.* 273:3452-3460).

The discovery of new transferases, 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, developmental, neurological, and autoimmune/inflammatory disorders, and parasitic infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transferases.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, transferases, referred to collectively as "TRNFR" and individually as "TRNFR-1," "TRNFR-2," "TRNFR-3," "TRNFR-4," "TRNFR-5," "TRNFR-6," "TRNFR-7," "TRNFR-8," "TRNFR-9," "TRNFR-10," "TRNFR-11," "TRNFR-12," "TRNFR-13," "TRNFR-14," "TRNFR-15," "TRNFR-16," "TRNFR-17," "TRNFR-18," "TRNFR-19," and "TRNFR-20." 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:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the

invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-20.

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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence
5 at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-20. In
10 another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:21-40.

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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least
15 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the
20 invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least
25 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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
30 polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

5 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:21-40, 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:21-40, 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
10 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
15 ID NO:21-40, 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:21-40, 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
20 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.

25 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:21-40, 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:21-40, c) a
30 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:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TRNFR, 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:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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 TRNFR, 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:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a

pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TRNFR, 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:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, 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, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20

contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:21-40, 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:21-40, 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:21-40, 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 score for the match between each polypeptide and its GenBank homolog is also shown.

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

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

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

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

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

DESCRIPTION OF THE INVENTION

5 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.

10 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.

15 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
20 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

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

An “allelic variant” is an alternative form of the gene encoding TRNFR. 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. 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 TRNFR include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TRNFR or a polypeptide with at least one functional characteristic of TRNFR. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TRNFR, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TRNFR.

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 TRNFR. 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 TRNFR 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 TRNFR. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRNFR either by directly interacting with TRNFR or by acting on components of the biological pathway in which TRNFR 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 TRNFR polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired.

- 5 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 makes contact with a particular antibody. When a protein or a fragment of a protein is used to
10 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 deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The
20 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. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a
25 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-
30 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” (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 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.

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 TRNFR, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

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

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

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

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the

protein is conserved and not significantly changed by such substitutions. The table below shows 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

25 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.

30 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
35 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.

40 "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 TRNFR or the polynucleotide encoding TRNFR 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:21-40 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:21-40, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:21-40 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:21-40 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:21-40 and the region of SEQ ID NO:21-40 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:1-20 is encoded by a fragment of SEQ ID NO:21-40. A fragment of SEQ ID NO:1-20 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-20. For example, a fragment of SEQ ID NO:1-20 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-20. The precise length of a fragment of SEQ ID NO:1-20 and the region of SEQ ID NO:1-20 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

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

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

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example,
5 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
10 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.

15 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
20 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
25 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
30 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 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 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. 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 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 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. (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.

5 Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, 10 such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid 15 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect 25 cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRNFR 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 TRNFR which is useful in any of the antibody production methods disclosed herein or known in the 30 art.

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

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

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

The phrases “nucleic acid” and “nucleic acid sequence” refer to a nucleotide, oligonucleotide, 5 polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

“Operably linked” refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably 10 linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

“Peptide nucleic acid” (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of 15 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 elongation, and may be pegylated to extend their lifespan in the cell.

“Post-translational modification” of an TRNFR may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the 20 art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TRNFR.

“Probe” refers to nucleic acid sequences encoding TRNFR, 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. 25 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 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).

30 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 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 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
5 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 purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

10 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
15 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
20 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
25 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
30 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.

5 Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

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

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

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

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 TRNFR, nucleic acids encoding TRNFR, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

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

30 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with

which they are naturally associated.

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

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

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

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid 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 transferases (TRNFR), the polynucleotides encoding TRNFR, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, developmental, neurological, and autoimmune/inflammatory disorders, and parasitic infections.

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

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transferases.

For example, SEQ ID NO:1 is 98% identical from amino acids 48 to 314 to human mono-ADP-ribosyltransferase (GenBank ID g1495421) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.9×10^{-142} , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a NAD:arginine ADP-ribosyltransferase 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 analyses provide further corroborative evidence that SEQ ID NO:1 is an ADP-ribosyltransferase.

For example, SEQ ID NO:6 is 92% identical to mouse glycerol-3-phosphate acyltransferase (GenBank ID g193367) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:6 also contains an acyltransferase 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.)

For example, SEQ ID NO:10 is 42% identical to human beta-1,3-galactosyltransferase

(GenBank ID g7799921) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $1.5e-64$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:10 also contains a galactosyltransferase 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.)

For example, SEQ ID NO:11 is 90% identical to mouse GalNAc alpha-2, 6-sialyltransferase V (GenBank ID g6691443) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $1.1e-167$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:11 also contains a sialyltransferase family 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.)

Pi-0286 For example, SEQ ID NO:12 is 29% identical to Aquifex aeolicus rRNA methylase (GenBank ID g2984156) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains a SpoU rRNA methylase family 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 analysis provide further corroborative evidence that SEQ ID NO:12 is an rRNA methylase.

For example, SEQ ID NO:15 is 70% identical to human serine palmitoyltransferase (GenBank ID g2564249) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $2.4e-208$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:15 also contains an aminotransferases class II 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:15 is an aminotransferase.

For example, SEQ ID NO:20 is 49% identical to human transglutaminase X (GenBank ID g6690087) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $5.6e-175$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:20 also contains transglutaminase family 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 PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:20 is a transglutaminase transferase.

SEQ ID NO:2-5, SEQ ID NO:7-9, SEQ ID NO:13-14, and SEQ ID NO:16-19 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-20 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:21-40 or that distinguish between SEQ ID NO:21-40 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 168639H1 is the identification number of an Incyte cDNA sequence, and LIVRNOT01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 70853185V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g681754) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 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 identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, 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 exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an “exon-stretching” algorithm. For example,

- 5 FLXXXXXX_gAAAAA_gBBBBB_1_N is the identification number of a “stretched” sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human 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
10 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).

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
15 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
25 column 5 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
30 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 TRNFR variants. A preferred TRNFR 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 TRNFR amino acid sequence, and which contains at least one functional or structural characteristic of TRNFR.

The invention also encompasses polynucleotides which encode TRNFR. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:21-40, which encodes TRNFR. The polynucleotide sequences of SEQ ID NO:21-40, 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 TRNFR. 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 TRNFR. 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:21-40 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:21-40. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TRNFR.

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 TRNFR, 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 TRNFR, and all such variations are to be considered as being specifically disclosed.

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

5 The invention also encompasses production of DNA sequences which encode TRNFR and TRNFR 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 TRNFR or any fragment thereof.

10 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:21-40 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

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

 The nucleic acid sequences encoding TRNFR may be extended utilizing a partial nucleotide
30 sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown

sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme 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 TRNFR may be cloned in recombinant DNA molecules that direct expression of TRNFR, 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 TRNFR.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TRNFR-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

5 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.

10 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TRNFR, 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
15 selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively,
20 fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding TRNFR 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
25 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, TRNFR itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved
30 using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of TRNFR, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid

chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, *supra*, pp. 28-53.)

In order to express a biologically active TRNFR, the nucleotide sequences encoding TRNFR or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TRNFR. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TRNFR. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TRNFR 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 TRNFR 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 TRNFR. 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 TRNFR. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding TRNFR can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TRNFR 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 TRNFR are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRNFR 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 TRNFR. 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 TRNFR. Transcription of sequences encoding TRNFR 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.)

5 In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TRNFR 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 TRNFR 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
15 constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of TRNFR in cell lines is preferred. For example, sequences encoding TRNFR can be transformed into
20 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
25 introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or
30 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 TRNFR is inserted within a marker gene sequence, transformed cells containing sequences encoding TRNFR can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRNFR under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

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

Immunological methods for detecting and measuring the expression of TRNFR 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 TRNFR is preferred, but a 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.)

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 TRNFR include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding TRNFR, or any fragments thereof, may be cloned into a vector

for the production of an mRNA probe. Such vectors are known in the art, are commercially available, 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
5 (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

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

In addition, a host cell strain may be chosen for its ability to modulate expression of the
15 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
20 (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 TRNFR may be ligated to a heterologous sequence resulting in translation of a
25 fusion protein in any of the aforementioned host systems. For example, a chimeric TRNFR protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TRNFR 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
30 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 TRNFR encoding sequence and the heterologous protein sequence, so that TRNFR 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 TRNFR 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.

TRNFR of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TRNFR. At least one and up to a plurality of test compounds may be screened for specific binding to TRNFR. 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 TRNFR, 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 TRNFR 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 TRNFR, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing TRNFR or cell membrane fractions which contain TRNFR are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRNFR 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 TRNFR, either in solution or affixed to a solid support, and detecting the binding of TRNFR 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.

TRNFR of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of TRNFR. Such compounds may include agonists, antagonists, or partial or

inverse agonists. In one embodiment, an assay is performed under conditions permissive for TRNFR activity, wherein TRNFR is combined with at least one test compound, and the activity of TRNFR in the presence of a test compound is compared with the activity of TRNFR in the absence of the test compound. A change in the activity of TRNFR in the presence of the test compound is indicative of a compound that modulates the activity of TRNFR. Alternatively, a test compound is combined with an in vitro or cell-free system comprising TRNFR under conditions suitable for TRNFR activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TRNFR 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 TRNFR 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 TRNFR 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 TRNFR 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 TRNFR 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 TRNFR, e.g., by secreting TRNFR 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

5 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TRNFR and transferases. In addition, the expression of TRNFR is closely associated with adrenal, adrenal tumor, aortic, bone, brain, brain meningioma, breast, gastrointestinal, kidney, lung, ovarian, placental, pancreatic, pancreatic tumor, prostate, prostate tumor, reproductive, and thyroid
10 tissues, and with T-cells. Therefore, TRNFR appears to play a role in cell proliferative, developmental, neurological, and autoimmune/inflammatory disorders, and parasitic infections. In the treatment of disorders associated with increased TRNFR expression or activity, it is desirable to decrease the expression or activity of TRNFR. In the treatment of disorders associated with decreased TRNFR expression or activity, it is desirable to increase the expression or activity of TRNFR.

Therefore, in one embodiment, TRNFR or a fragment or derivative thereof may be
15 administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRNFR. 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,
20 lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy,
25 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
30 glaucoma, cataract, and sensorineural hearing loss; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral

meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental 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 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, familial frontotemporal dementia, and Lesch-Nyhan syndrome; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and infections by parasites classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, and cestodes (tapeworm).

In another embodiment, a vector capable of expressing TRNFR 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 TRNFR including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified TRNFR 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 TRNFR including, but not limited to, those provided above.

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

10 In a further embodiment, an antagonist of TRNFR may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRNFR. Examples of such disorders include, but are not limited to, those cell proliferative, developmental, neurological, and autoimmune/inflammatory disorders, and parasitic infections described above. In one aspect, an antibody which specifically binds TRNFR 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 TRNFR.

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

20 In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the 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.

25 An antagonist of TRNFR may be produced using methods which are generally known in the art. In particular, purified TRNFR may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TRNFR. Antibodies to TRNFR 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 fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit
30 dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TRNFR or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels

such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in 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 TRNFR have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TRNFR 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 TRNFR 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 TRNFR-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 TRNFR 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 TRNFR and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TRNFR 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 TRNFR. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of TRNFR-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TRNFR epitopes, represents the average affinity, or avidity, of the antibodies for TRNFR. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRNFR 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 TRNFR-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 TRNFR, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRNFR-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 TRNFR, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TRNFR. 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 TRNFR. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense
5 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
10 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.*
15 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding TRNFR 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
20 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), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal,
25 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),
30 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 TRNFR expression or regulation causes disease, the expression of TRNFR 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 TRNFR are treated by constructing mammalian expression vectors encoding TRNFR and introducing these vectors by mechanical means into TRNFR-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 TRNFR 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). TRNFR 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 Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TRNFR 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 TRNFR expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TRNFR 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).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TRNFR to cells which have one or more genetic abnormalities with respect to the expression of TRNFR. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent 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 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRNFR to target cells which have one or more genetic abnormalities with respect to the expression of TRNFR. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRNFR to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent 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 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.

5 (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

10 In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TRNFR 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
15 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 TRNFR into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRNFR-coding RNAs and the synthesis of high levels of TRNFR in vector transduced cells. While
20 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 TRNFR into a variety of cell types. The specific transduction of a subset of
25 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
30 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.

5 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 TRNFR.

10 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
15 oligonucleotides using ribonuclease protection assays.

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

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

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TRNFR. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides,

transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased TRNFR expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TRNFR may be therapeutically useful, and in the treatment of disorders associated with decreased TRNFR expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TRNFR may be therapeutically useful.

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

5 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 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.

10 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 Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of TRNFR, antibodies to TRNFR, and mimetics, agonists, antagonists, or inhibitors of TRNFR.

The compositions utilized in this invention may be administered by any number of routes
15 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
20 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
25 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
30 macromolecules comprising TRNFR or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TRNFR 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 TRNFR or fragments thereof, antibodies of TRNFR, and agonists, antagonists or inhibitors of TRNFR, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. 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 μg to 100,000 μg , 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 TRNFR may be used for the diagnosis of disorders characterized by expression of TRNFR, or in assays to monitor patients being

treated with TRNFR or agonists, antagonists, or inhibitors of TRNFR. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRNFR include methods which utilize the antibody and a label to detect TRNFR 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 TRNFR, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TRNFR expression. Normal or standard values for TRNFR expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TRNFR under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TRNFR 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 TRNFR 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 TRNFR may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TRNFR, and to monitor regulation of TRNFR levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TRNFR or closely related molecules may be used to identify nucleic acid sequences which encode TRNFR. 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 TRNFR, 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 TRNFR 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:21-40 or from genomic sequences including promoters, enhancers, and introns of the TRNFR gene.

Means for producing specific hybridization probes for DNAs encoding TRNFR include the cloning of polynucleotide sequences encoding TRNFR or TRNFR derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may

be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

- 5 Polynucleotide sequences encoding TRNFR may be used for the diagnosis of disorders associated with expression of TRNFR. 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
- 10 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and
- 15 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,
- 20 craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases,
- 25 bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental
- 30 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 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, familial frontotemporal dementia, and Lesch-Nyhan syndrome; an autoimmune/inflammatory disorder such as acquired

- 5 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,
- 10 erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis,
- 15 thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and infections by parasites classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as
- 20 ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, and cestodes (tapeworm). The polynucleotide sequences encoding TRNFR 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 TRNFR expression. Such qualitative or quantitative methods are well known in the art.

- 25 In a particular aspect, the nucleotide sequences encoding TRNFR may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TRNFR 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
- 30 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 TRNFR 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 TRNFR, 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 TRNFR, 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 TRNFR 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 TRNFR, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRNFR, 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 TRNFR 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 TRNFR 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 fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

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

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

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

A particular embodiment relates to the use of the polynucleotides of the present invention to

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 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
10 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
15 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)
20 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
25 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
30 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 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 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 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 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 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 TRNFR to quantify the levels of TRNFR expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoz, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each

array element.

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

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

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

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding TRNFR may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during

chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) *Proc. Natl. Acad. Sci. USA* 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TRNFR 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, TRNFR, 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 TRNFR 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 TRNFR, or fragments thereof, and washed. Bound TRNFR is then detected by methods well known in the art. Purified TRNFR can also be coated directly onto plates for use in the aforementioned drug screening techniques.

Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

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

In additional embodiments, the nucleotide sequences which encode TRNFR may be used in any 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 preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No.60/236,523, U.S. Ser. No.60/238,481, U.S. Ser. No.60/244,025, U.S. Ser. No.60/246,001, U.S. Ser. No.60/247,931, U.S. Ser. No.60/249,639, and U.S. Ser. No.60/252,819, 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 (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles

(QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), 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). 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, 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 Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden

Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as
5 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
10 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
15 NO:21-40. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative transferases were initially identified by running the Genscan gene identification
20 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
25 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 transferases, the encoded polypeptides were analyzed by querying against PFAM models for transferases. Potential transferases were also identified by homology to Incyte cDNA
30 sequences that had been annotated as transferases. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-

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

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

“Stitched” Sequences

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

“Stretched” Sequences

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

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

VI. Chromosomal Mapping of TRNFR Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:21-40 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:21-40 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and 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.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by 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.

VII. Analysis of Polynucleotide Expression

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

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer

search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

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

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 TRNFR 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 TRNFR.

VIII. Extension of TRNFR Encoding Polynucleotides

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

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

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

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

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

clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:21-40 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 -³²P] 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 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), supra.).

Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

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

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X 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 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
5 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 uses
10 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
15 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.
20 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.
25 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
30 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
5 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
10 resolution of 20 micrometers.

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

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

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

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated

to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

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

XII. Expression of TRNFR

Expression and purification of TRNFR is achieved using bacterial or virus-based expression systems. For expression of TRNFR in bacteria, cDNA is subcloned into an appropriate vector 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 TRNFR upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRNFR 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 TRNFR by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong 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.)

In most expression systems, TRNFR 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 Pharmacia

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

XIII. Functional Assays

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

The influence of TRNFR on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TRNFR 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 TRNFR and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of TRNFR Specific Antibodies

5 TRNFR 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 rabbits and to produce antibodies using standard protocols.

Alternatively, the TRNFR amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is
10 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-
15 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 oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TRNFR activity by, for example, binding the peptide or TRNFR to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat
20 anti-rabbit IgG.

XV. Purification of Naturally Occurring TRNFR Using Specific Antibodies

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

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

XVI. Identification of Molecules Which Interact with TRNFR

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

5 Alternatively, molecules interacting with TRNFR 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).

TRNFR 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
10 between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of TRNFR Activity

TRNFR transferase activity is measured through assays such as a methyl transferase assay in which the transfer of radiolabeled methyl groups between a donor substrate and an acceptor substrate
15 is measured (Bokar, J.A. et al. (1994) *J. Biol. Chem.* 269:17697-17704). Reaction mixtures (50 μ l final volume) contain 15 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 10 mM dithiothreitol, 3% polyvinylalcohol, 1.5 μ Ci [*methyl*- ^3H]AdoMet (0.375 μ M AdoMet) (DuPont-NEN), 0.6 μ g HEM, and acceptor substrate (0.4 μ g [^{35}S]RNA or 6-mercaptopurine (6-MP) to 1 mM final concentration). Reaction mixtures are incubated at 30°C for 30 minutes, then 65°C for 5 minutes. The products are
20 separated by chromatography or electrophoresis and the level of methyl transferase activity is determined by quantification of *methyl*- ^3H recovery.

Lysophosphatidic acid acyltransferase activity of TRNFR is measured by incubating samples containing TRNFR with 1 mM of the thiol reagent 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB), 50 mM LPA, and 50 mM acyl-CoA in 100 mM Tris-HCl, pH 7.4. The reaction is initiated by addition of acyl-
25 CoA, and allowed to reach equilibrium. Transfer of the acyl group from acyl-CoA to LPA releases free CoA, which reacts with DTNB. The product of the reaction between DTNB and free CoA absorbs at 413 nm. The change in absorbance at 413 nm is measured using a spectrophotometer, and is proportional to the lysophosphatidic acid acyltransferase activity of TRNFR in the sample.

N-acyltransferase activity of TRNFR is measured using radiolabeled amino acid substrates
30 and measuring radiolabel incorporation into conjugated products. TRNFR is incubated in a reaction buffer containing an unlabeled acyl-CoA compound and radiolabeled amino acid, and the radiolabeled acyl-conjugates are separated from the unreacted amino acid by extraction into n-butanol or other appropriate organic solvent. For example, Johnson, M. R. et al. (1990; *J. Biol. Chem.* 266:10227-10233) measured bile acid-CoA:amino acid N-acyltransferase activity by incubating the enzyme with

cholyl-CoA and ^3H -glycine or ^3H -taurine, separating the tritiated cholate conjugate by extraction into n-butanol, and measuring the radioactivity in the extracted product by scintillation. Alternatively, N-acetyltransferase activity is measured using the spectrophotometric determination of reduced CoA (CoASH) described below.

5 N-acetyltransferase activity of TRNFR is measured using the transfer of radiolabel from [^{14}C]acetyl-CoA to a substrate molecule (for example, see Deguchi, T. (1975) *J. Neurochem.* 24:1083-5). Alternatively, a newer spectrophotometric assay based on DTNB reaction with CoASH may be used. Free thiol-containing CoASH is formed during N-acetyltransferase catalyzed transfer of an acetyl group to a substrate. CoASH is detected using the absorbance of DTNB conjugate at 412
10 nm (De Angelis, J. et al. (1997) *J. Biol. Chem.* 273:3045-3050). TRNFR activity is proportional to the rate of radioactivity incorporation into substrate, or the rate of absorbance increase in the spectrophotometric assay.

Aminotransferase activity of TRNFR is measured by determining the activity of purified TRNFR or crude samples containing TRNFR toward various amino and oxo acid substrates under
15 single turnover conditions by monitoring the changes in the UV/VIS absorption spectrum of the enzyme-bound cofactor, PLP. The reactions are performed at 25 °C in 50 mM 4-methylmorpholine (pH 7.5) containing 9 μM purified TRNFR or TRNFR containing samples and substrate to be tested (amino and oxo acid substrates). The half-reaction from amino acid to oxo acid is followed by measuring the decrease in absorbance at 360 nm and the increase in absorbance at 330 nm due to the
20 conversion of enzyme-bound PLP to PMP. The specificity and relative activity of TRNFR is determined by the activity of the enzyme preparation against specific substrates (Vacca, R.A. et al. (1997) *J. Biol. Chem.* 272:21932-21937).

Galactosyltransferase activity of TRNFR is determined by measuring the transfer of galactose from UDP-galactose to a GlcNAc-terminated oligosaccharide chain in a radioactive assay.
25 (Kolbinger, F. et al. (1998) *J. Biol. Chem.* 273:58-65.) The TRNFR sample is incubated with 14 μl of assay stock solution (180 mM sodium cacodylate, pH 6.5, 1 mg/ml bovine serum albumin, 0.26 mM UDP-galactose, 2 μl of UDP- ^3H galactose), 1 μl of MnCl_2 (500 mM), and 2.5 μl of GlcNAc $\beta\text{O}-(\text{CH}_2)_8-\text{CO}_2\text{Me}$ (37 mg/ml in dimethyl sulfoxide) for 60 minutes at 37°C. The reaction is quenched by the addition of 1 ml of water and loaded on a C18 Sep-Pak cartridge (Waters), and the column is
30 washed twice with 5 ml of water to remove unreacted UDP- ^3H galactose. The ^3H galactosylated GlcNAc $\beta\text{O}-(\text{CH}_2)_8-\text{CO}_2\text{Me}$ remains bound to the column during the water washes and is eluted with 5 ml of methanol. Radioactivity in the eluted material is measured by liquid scintillation counting and is proportional to galactosyltransferase activity of TRNFR in the starting sample.

Phosphoribosyltransferase activity of TRNFR is measured as the transfer of a phosphoribosyl

group from phosphoribosylpyrophosphate (PRPP) to a purine or pyrimidine base.

Assay mixture (20 ml) containing 50 mM Tris acetate, pH 9.0, 20 mM 2-mercaptoethanol, 12.5 mM MgCl_2 , and 0.1 mM labeled substrate, for example, $[^{14}\text{C}]$ uracil, is mixed with 20 ml of TRNFR diluted in 0.1 M Tris acetate, pH 9.7, and 1 mg/ml bovine serum albumin. Reactions are preheated for 1 min at 37°C, initiated with 10 ml of 6 mM PRPP, and incubated for 5 min at 37°C. The reaction is stopped by heating at 100°C for 1 min. The product $[^{14}\text{C}]$ UMP is separated from $[^{14}\text{C}]$ uracil on DEAE-cellulose paper (Turner, R.J. et al. (1998) J. Biol. Chem. 273:5932-5938). The amount of $[^{14}\text{C}]$ UMP produced is proportional to the phosphoribosyltransferase activity of TRNFR.

ADP-ribosyltransferase activity of TRNFR is measured as the transfer of radiolabel from adenine-NAD to agmatine (Weng, B. et al. (1999) J. Biol. Chem. 274:31797-31803). Purified TRNFR is incubated at 30°C for 1 hr in a total volume of 300 ml containing 50 mM potassium phosphate (pH. 7.5), 20 mM agmatine, and 0.1 mM [adenine-U- ^{14}C]NAD (0.05 mCi). Samples (100 ml) are applied to Dowex columns and $[^{14}\text{C}]$ ADP-ribosylagmatine eluted with 5 ml of water for liquid scintillation counting. The amount of radioactivity recovered is proportional to ADP-ribosyltransferase activity of TRNFR.

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

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
168639	1	168639CD1	21	168639CB1
2792817	2	2792817CD1	22	2792817CB1
3090127	3	3090127CD1	23	3090127CB1
7480989	4	7480989CD1	24	7480989CB1
2280673	5	2280673CD1	25	2280673CB1
1517230	6	1517230CD1	26	1517230CB1
5665262	7	5665262CD1	27	5665262CB1
2119916	8	2119916CD1	28	2119916CB1
8186259	9	8186259CD1	29	8186259CB1
70250400	10	70250400CD1	30	70250400CB1
2778782	11	2778782CD1	31	2778782CB1
2715885	12	2715885CD1	32	2715885CB1
1742628	13	1742628CD1	33	1742628CB1
2124971	14	2124971CD1	34	2124971CB1
2258250	15	2258250CD1	35	2258250CB1
2626035	16	2626035CD1	36	2626035CB1
4831382	17	4831382CD1	37	4831382CB1
2122183	18	2122183CD1	38	2122183CB1
7484338	19	7484338CD1	39	7484338CB1
8326588	20	8326588CD1	40	8326588CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank NO:	ID	Probability Score	GenBank Homolog
1	168639	g1495421		2.90E-142	[Homo sapiens] mono-ADP-ribosyltransferase Koch-Nolte, F. (1997) Two novel human members of an emerging mammalian gene family related to mono-ADP-ribosylating bacterial toxins. Genomics 39, 370-376
2	2792817	g171529		1.90E-44	[Saccharomyces cerevisiae] uracil phosphoribosyltransferase (FUR1) Kern, L. et al. (1990) The FUR1 gene of <i>Saccharomyces cerevisiae</i> : Cloning, structure and expression of wild-type and mutant alleles. Gene 88, 149-157
3	3090127	g2317725		6.20E-124	[Mus musculus] putative lysophosphatidic acid acyltransferase
4	7480989	g7799921		9.80E-55	[Homo sapiens] beta-1,3-galactosyltransferase
5	2280673	g8497318		4.60E-302	[Mus musculus] acetyltransferase Tubedown-1 Gendron, R.L. et al. (2000) Tubedown-1, A novel acetyltransferase associated with blood vessel development. Dev. Dyn. 218, 300-315
6	1517230	g193367		0	[Mus musculus] glycerol-3-phosphate acyltransferase Shin, D.H. et al. (1991) Transcriptional regulation of p90 with sequence homology to <i>Escherichia</i> <i>coli</i> glycerol-3-phosphate acyltransferase. J. Biol. Chem. 266, 23834-23839
7	5665262	g193367		2.60E-97	[Mus musculus] glycerol-3-phosphate acyltransferase; Shin, <i>supra</i>
8	2119916	g6862930		2.60E-24	[Arabidopsis thaliana] putative glycosyl transferase
9	8186259	g6706726		1.90E-19	[Schizosaccharomyces pombe] putative mannosyltransferase
10	70250400	g7799921		1.50E-64	[Homo sapiens] beta-1,3-galactosyltransferase
11	2778782	g6691443		1.10E-167	[Mus musculus] GalNAc alpha-2, 6-sialyltransferase V Ikehara, Y. et al. (1999) A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1alpha synthase (ST6GalNAc V)(1) FEBS Lett. 463, 92-96
12	2715885	g2984156		0	[Aquifex aeolicus] rRNA methylase
12	2715885	g6458456		9.00E-25	[F1]Deinococcus radiodurans] RNA methyltransferase, TrmH family
13	1742628	g12274933		0	[F1][Homo sapiens] alanine:glyoxylate aminotransferase 2 homolog 1, splice form 1
14	2124971	g3954978		5.70E-253	[Mus musculus] acetylglucosaminyltransferase-like protein Peyrard, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:598-603
15	2258250	g2564249		2.40E-208	[Homo sapiens] serine palmitoyltransferase, subunit II Weiss, B. and Stoffel, W. (1997) Eur. J. Biochem. 249:239-247
16	2626035	g3954978		5.70E-253	[Mus musculus] acetylglucosaminyltransferase-like protein Peyrard, M. et al. <i>supra</i>

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank NO:	ID	Probability Score	GenBank Homolog
17	4831382	g10336504		2.70E-159	[Homo sapiens] UDP-GalNAc: polypeptide Toba, S. et al. (2000) Biochim. Biophys. Acta 1493:264-268
18	2122183	g5123545		8.30E-55	[Arabidopsis thaliana] arginine methyltransferase (pam1)
19	7484338	g2621120		3.50E-28	[Methanobacterium thermoautotrophicum] O-linked GlcNAc
20	8326588	g6690087		5.60E-175	[Homo sapiens] transglutaminase X Aeschlimann, D. et al. (1998) J. Biol. Chem. 273:3452-3460

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	168639CD1	314	S174 S240 S286 S287 S61 T201 T21 T78	N114 N178 N222 N257 N274	Signal peptide: M1-G46	SPSCAN
					NAD:arginine ADP-ribosyltransferase domain: P19-R313	HMMER_PFAM
					NAD:arginine ADP-ribosyl-transferase signature BL01291: F55-T84, P112-T127, F136-L169, G205-N257, L267-A285	BLIMPS_BLOCKS
					Arginine ADP-ribosyltransferase signature PR00970: D56-L77, D86-W104, M115-N129, F154-L170, I202-A217, T224-S240, K242-N257	BLIMPS_PRINTS
					TRANSFERASE ADPRIBOSYLTRANSFERASE NADP+ARGININE PRECURSOR GLYCOSYLTRANSFERASE SIGNAL NAD MONOADPRIBOSYLTRANSFERASE MONOADP RIBOSYLTRANSFERASE PD004385:P19-C290	BLAST_PRODROM
					RIBOSYLTRANSFERASE; NAD; ADP; ARGININE DM02464 A55461 1-312:W25-K284 DM02464 JC4367 1-300:L26-S299 DM02464 P52961 1-326:P19-P294 DM02464 P17982 1-274:D56-K311	BLAST_DOMO

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
2	2792817CD1	309	S132 S201 S211 S86 S95 T215 T246		TRANSFERASE URACIL PHOSPHORIBOSYLTRANSFERASE UMP PYROPHOSPHORYLASE UPRTASE GLYCOSYLTRANSFERASE PROTEIN URIDINE KINASE PD150187:L122-T215 PHOSPHORIBOSYL PYROPHOSPHATE BINDING DOMAIN DM02076 P18562 44-250:L112-T290 DM02076 P27515 287-500:L113-D309 DM02076 P43049 1-208:T124-I288 DM02076 P47276 1-205:I119-V293	BLAST_PRODOM BLAST_DOMO
3	3090127CD1	434	S100 S117 S138 S185 S231 S319 S344 S77 T113 T283 T373 T50 T57	N228 N308 N309 N343 N422 N430	Signal peptide: M1-G29	SPSCAN
					Signal peptide: M1-G29	HMMER
					Transmembrane domains: W14-G33, M143-I166	HMMER
					ACYLTRANSFERASE PUTATIVE LYSOPHOSPHATIDIC ACID TRANSFERASE R07E3.5 PROTEIN PD036247:L104-D292	BLAST_PRODOM
					ACYLTRANSFERASE PROTEIN PUTATIVE LYSOPHOSPHATIDIC ACID TRANSFERASE R07E3.5 M79.3 PD022151:K293-L406	BLAST_PRODOM
4	7480989CD1	402	S124 S351 S368 T205	N203	Signal peptide: M1-A27	SPSCAN

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
					Signal peptide: M1-A27	HMIMER
					Galactosyltransferase domain: E130-V363	HMIMER_PFAM
					PROTEIN TRANSFERASE GLYCOSYLTRANSFERASE UDPGAL:BETAGLCNAC BETA BETA1 3GALACTOSYLTRANSFERASE C54C8.3 E03H4.11 C47F8.6 PD004190:R131-L319	BLAST_PRODOM
5	2280673CD1	866	S352 S355 S474 S537 S588 S694 S759 S779 S813 S855 S92 T180 T229 T328 T399 T406 T534 T652 T729 T788 T789 T795 T801 Y26 Y260	N317 N586 N751	TPR Domain: W83-K111, Q377-P405	HMIMER_PFAM
					NTERMINAL ACETYLTRANSFERASE TRANSFERASE AMINOTERMINAL ALPHA AMINO ACYLTRANSFERASE ACETYLATION PD156409:L250-D631	BLAST_PRODOM

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
6	1517230CD1	828	S110 S136 S232 S321 S405 S414 S448 S459 S550 S565 S57 S670 S685 S688 S801 T30 T48 T504 T549 T657 T761 T763 Y364	N128 N135 N375 N450 N454 N741 N95	Transmembrane domain: R172-F197	HMMER
					Acyltransferase domain: K215-S412	HMMER_PFAM
					ACYLTRANSFERASE TRANSFERASE GLYCEROL3PHOSPHATE GPAT PHOSPHOLIPID BIOSYNTHESIS MITOCHONDRIAL PRECURSOR TRANSMEMBRANE MITOCHONDRION PD025192:N132-S414, I465-G600, Q675-G698, PD042760:F650-L828, PD037846:M1-S155, PD152739:R453-K649	BLAST_PRODROM
					GLYCEROL; ACYLTRANSFERASE DM083300 P44857 185-805:D153-L410, I465-K596 DM083300 P00482 205-826:E131-D442, I465-A587	BLAST_DOMO
7	5665262CD1	801	S125 S18 S369 S383 S390 S447 S589 S654 S66 S668 T204 T207 T3 T419 T55 T648 T666 T737 T758		ACYLTRANSFERASE TRANSFERASE GLYCEROL3PHOSPHATE GPAT PHOSPHOLIPID BIOSYNTHESIS MITOCHONDRIAL PRECURSOR TRANSMEMBRANE MITOCHONDRION PD025192:I104-W395, PD037846:L40-S125, PD042760:I629-I799, PD152739:D439-R627	BLAST_PRODROM
8	2119916CD1	349	S126 S127 S189 S74 T43 Y260	N234	Signal peptide: M1-I22	SPSCAN

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
					Signal peptide: M1-I22	HMMER
					Transmembrane domain: A2-K25	HMMER
					Glycosyl transferase family 8 domain: G63-A340	HMMER_PFAM
9	8186259CD1	555	S103 S115 S161 S184 S210 S469 S522 S534 S551 S75 T131 T132 T173 T376 T379 T386 T413 T416 T499 T52 T540	N159 N197 N272 N482 N497 N68	Signal peptide: M1-A42	SPSCAN
					Transmembrane domains: L216-W234, L326-G344, H348-I371, D433-W451	HMMER
					Glycosyl transferase signature PF00953A: L300-L326	BLIMPS-PFAM
					PROTEIN PRECURSOR PTM1 TRANSMEMBRANE SIGNAL MEMBRANE ISOLOG SIMILAR S CERVISIAE PD014374: I157-T479	BLAST_PRODOM
10	70250400CD1	401	S181 S395 T145 T216 T342 Y189 Y282	N214 N391 N88 N94	Rgd R131-D133	MOTIFS
					Signal_cleavage:M1-F25	SPSCAN
					Signal_peptide:M1-S28	HMMER
					Galactosyltransferase Galactosyl_D148-A383	HMMER_PFAM
					PROTEIN TRANSFERASE GLYCOSYLTRANSFERASE UDPGAL: BETAGLCNAC BETA BETA1 3GALACTOSYLTRANSFERASE C54C8.3 E03H4.11 C47F8.6 PD004190: R149-V339	BLAST_PRODOM

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
11	2778782CD1	336	S112 S140 S227 S293 T220 T299	N137 N161	Signal cleavage: M1-G30	SPSCAN
					Signal peptide: M1-S21	HMMER
					Sialyltransferase family domain: V25-W307	HMMER_PFAM
					TRANSFERASE GLYCOSYLTRANSFERASE ALPHANACETYLGALACTOSAMINIDE ALPHA2 6SIALYLTRANSFERASE ST6GALNACIII STY GLYCOPROTEIN TRANSMEMBRANE SIGNALANCHOR PD129520:W171-W320	BLAST_PRODOR
					do SIALYLTRANSFERASE; ALPHA-2; POLYSIALYLTRANSFERASE; III; DM03800 48686 44-369:P88-D273, DM03800 A56950 48-377:C96-E287, DM03800 I39169 44-369:P88-D273	BLAST_DOMO
					LUMENAL DOMAIN	BLAST_DOMO
					DM01020 P15907 67-393:C96-N306	
12	2715885CD1	353	S138 S177 S181 S203 S229 S242 S305 S327 S343 S47 T176 T312 T46 T6	N291	SpoU rRNA Methylase family SpoU_methylase:Q145 H301	HMMER_PFAM
					SpoU rRNA Methylase family PF00588: P156-A166	BLIMPS_PFAM
					METHYLTRANSFERASE; HI0860; DM01021 , P25270 207-409:L146-I303 Q06753 64-241:H116-I303 E64160 63-240:Q114-I303 P44906 63-240:Q114-I303	BLAST_DOMO

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
					METHYLTRANSFERASE TRANSFERASE TRNA/RRNA RRNA METHYLASE CONSERVED SPOU 23S ANTIBIOTIC PD001243:L146-H301	BLAST_PRODOM
13	1742628CD1	499	S111 S145 S203 S21 S302 S470 S479 S494 S6 S96 T10 T173 T291 T418 T434 T442 T98		Aminotransferases class-III pyridoxal-phosphate attachment site F243-G283	MOTIFS
					Aminotransferases class-III pyridoxal-phosphate attachment site aa_transfer_class_3.prf: E234-G303	PROFILESKAN
					Aminotransferases class-III pyridoxal- aminotran_3: Y40-H142 V165-M415	HMMER_PFAM
					Aminotransferases class BL00600: P271-G283, Y306-L324, D43-V66, K102-R127, V132-I147, I207-G220, Y228-G256	BLIMPS_BLOCKS
					AMINOTRANSFERASES CLASS-III PYRIDOXAL-PHOSPHATE ATTACHMENT DM00188 JC1497 1-417: P30-R431 P30268 2-430: I31-R431 P16932 2-426: D29-D421 DM00188 P07991 3-419: P30-L429	BLAST_DOMO
14	2124971CD1	721	S108 S290 S342 S520 S529 S545 S568 T278 T323 T362 T464 T558 T583 T681	N107 N231 N80	Signal cleavage: M1-G28	SPSCAN

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
					Glycosyl transferase family 8 Glyco_transf_8: R110-G368, HMM_score 19.6	HMMER_PFAM
					signal peptide signal_peptide: M1-G28	HMMER
					transmembrane domain: L10-F27	HMMER
					KIAA0609 PROTEIN PD139660: C33-V300, PROTEIN IBETA1	BLAST_PRODROM
					3NACETYLGLUCOSAMINYLTRANSFERASE	
					TRANSFERASE GLYCOSYLTRANSFERASE	
					KIAA0609 K09C8.4 PD134147: F584-F693, KIAA0609 K09C8.4 PD042423: I301-L581	
15	2258250CD1	552	S180 S197 S21 S414 S537 S54 T151 T155 T183 T198 T268 T295 T347 T515 T546 Y290	N149 N22 N254 N275 N28 N457	Aminotransferases class-II pyridoxal-phosphate attachment site T368-G377	MOTIFS
					Aminotransferases class-II pyridoxal-phosphate attachment site aa_transfer_class_2.prf: G350-S396	PROFILESKAN
					Aminotransferases class- BL00599: A196-A224, S250-V259, D336-G348, V362-T368	BLIMPS_BLOCKS
					Aminotransferases class-II aminotran_2: G203-R501	HMMER_PFAM
					AMINOTRANSFERASES CLASS-II PYRIDOXAL-PHOSPHATE ATTACHMENT S DM00464 P40970 111-523: R116-L529 S54046 111-523: R116-L529 Q09925 143-550: R116-G524 P48241 110-521: R116-L527	BLAST_DOMO

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
					TRANSFERASE SERINE PALMITOYLTRANSFERASE LONG CHAIN BASE BIOSYNTHESIS PROTEIN SPT ACYLTRANSFERASE PD009687: E53-R199 Signal cleavage: M1-F27	BLAST_PRODOM
16	2626035CD1	690	S259 S311 S489 S498 S514 S537 S77 T247 T292 T331 T433 T527 T552 T650	N200 N49 N76	Glycosyl transferase family 8 Glyco_transf_8: R79-G337 signal_peptide: M1-F27 transmembrane domain: L10-F27, KIAA0609 PROTEIN PD139660: R30-V269, PD042423: I270-L550 PROTEIN IBETA1 3NACETYLGLUCOSAMINYLTRANSFERASE TRANSFERASE GLYCOSYLTRANSFERASE KIAA0609 K09C8.4 PD134147: F553-F662 Signal cleavage: M1-G43	SPSCAN
						HMMER
						HMMER
						BLAST_PRODOM
17	4831382CD1	607	S127 S148 S152 S176 S196 S197 S214 S226 S428 S476 S568 S578 T4 T436 T449 T522 T59 Y470 Y545	N146 N195 N320		SPSCAN
					Glycosyl transferases Glycos_transf_2: S157-G345	HMMER_PFAM

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
					ACETYLGALACTOSAMINYLTRANSFERASE; POLYPEPTIDE; DM03891 Q07537 32-558: P107-V601 P34678 37-600: P107-L572 I37405 21-571: H76-T599	BLAST_DOMO
					NACETYLGALACTOSAMINYLTRANSFERASE TRANSFERASE POLYPEPTIDE ACETYLGALACTOSAMINYLTRANSFERASE UDPGALNAC: POLYPEPTIDE GLYCOSYLTRANSFERASE PROTEINUDP PROTEIN UDP N PD003162: E319-M461	BLAST_PRODOM
18	2122183CD1	375	S220 S242 S301 S55 T21 T39 T85 Y47		ARGININE NMETHYLTRANSFERASE TRANSFERASE METHYLTRANSFERASE PROTEIN INTERFERON RECEPTOR IBOUND ALTERNATIVE SPLICING PD011237: P186-T368	BLAST_PRODOM
19	7484338CD1	760	S102 S214 S325 S368 S581 S591 S684 S99 T110 T155 T25 T318 T455 T473 Y607	N516 N628 N97	Squalene and phytoene synthases: F149-L173	HMMER_PFAM
					TPR Domain: A705-A738, A501-Y534, V535-F568, A569-Y602, P603-H636, S637-D670, H671-A704	HMMER_PFAM
					F32D1.3 PROTEIN SIMILAR E NIDULANS BIMA GENE PRODUCT PD041324: Y304-L466 transmembrane domain: I131-F149 K459-R479 Aldo/keto reductase family putative active site signature L365-L380	BLAST_PRODOM HMMER MOTIFS

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
20	8326588CD1	710	S145 S16 S214 S232 S252 S267 S378 S430 S445 S452 S463 S53 S615 S689 S90 T117 T317 T328 T34 T542 T550 T593 T6	N200 N312 N421	Transglutaminase-like superfamily: V274-T363	HMMER_PFAM
					Transglutaminase family: V4-G125, E481-V706	HMMER_PFAM
					Transglutaminases active site DM00983 Q01841 10-693: L7-G483	BLAST_DOMO
					Transglutaminases active site DM00983 P51176 2-685: L7-S463	BLAST_DOMO
					Transglutaminases active site DM00983 Q08188 1-692: V4-V704	BLAST_DOMO
					Transglutaminases active site DM00983 P52181 3-682: V12-V704	BLAST_DOMO
					TRANSGLUTAMINASE TRANSFERASE ACYLTRANSFERASE PROTEIN GLUTAMINE GAMMA GLUTAMYL-TRANSFERASE CALCIUM BINDING TGASE TISSUE C MEMBRANE PD002491: K21-E454	BLAST_PRODOM
					TRANSGLUTAMINASE TRANSFERASE ACYLTRANSFERASE PROTEIN GLUTAMINE GAMMA GLUTAMYL-TRANSFERASE TGASE CALCIUM BINDING TISSUE C MEMBRANE PD002568: W547-V704	BLAST_PRODOM

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
					Transglutaminases proteins BL00547: I425-A462, L630-L650, N19-L45, F137-F187, R216-P253, V274-Y318, D332-Q366, P376-W407	BLIMPS_BLOCKS
					Transglutaminases active site: W256-T314	PROFILESKAN

Table 4

Polynucleotide SEQ ID NO:	Incyte ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
21	168639CB1	1096	1-266, 1050-1096	168639H1 (LIVRNOT01)	472	809
				7334682H1 (CONFTDN02)	1	563
				g681754	556	1096
				g1495420_CD	187	989
				4085481H1 (LIVRNOT06)	637	813
22	2792817CB1	1380	261-451, 1-51 1219-1380	6777156J1 (OVARDIR01)	604	1380
				3984662F7 (UTRSTUT05)	1	658
23	3090127CB1	2647	2626-2647, 1-532, 1737-1772	6637438J1 (KIDNNOR08)	1	427
				70853185V1	1142	1748
				4165641H1 (BRSTNOT32)	1006	1284
				4069256F6 (KIDNNOT26)	494	967
				6562333H1 (MCLDTXT04)	2099	2647
				6902902H1 (MUSLTDR02)	859	1277
				3324214T7 (PTHYNOT03)	1851	2455
				6830181J1 (SINTNOR01)	332	941
				70852505V1	1375	1875
24	7480989CB1	2117	605-973	7368215H1 (ADREFEC01)	287	860
				8092634H1 (EYERNOA01)	1	542
				7744738J1 (ADRETUE04)	1042	1688
				3209148H1 (BLADNOT08)	667	943
				1256113F6 (MENITUT03)	1637	2117
				71264921V1	879	1561
25	2280673CB1	3320	2081-2155, 3300-3320	6848571H1 (KIDNTMN03)	2665	2874
				6493021H1 (MIXDUNB01)	2689	3320
				1340719F6 (COLNTUT03)	968	1558
				5958093F9 (BRATNOT05)	2082	2703
				933894R1 (CERVNOT01)	1	517
				7375225H1 (ESOGTUE01)	410	1001
				6422836H1 (BRSTUNT01)	646	1116
				2509485F6 (CONUTUT01)	1686	2226
				7583735H1 (BRAIFEC01)	1219	1829
26	1517230CB1	3210	1-209, 870-1300	5522234H1 (LIVRDIR01)	878	988
				8121650H1 (HEAONOC01)	168	901
				5372875F7 (BRAINOT22)	2667	3210
				6706879H1 (HEAADIR01)	2627	3156
				7319987R8 (ADRETUE02)	1064	1668
				5522072F6.edit (LIVRDIR01)	1023	1419
				6840482H1 (BRSTNON02)	1	671
				7437552H1 (ADRETUE02)	1892	2527
				1517230F6 (PANCTUT01)	2415	2648
				GBI.g10186839_000001.edit	751	1131

Table 4

Polynucleotide SEQ ID NO:	Incyte ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
				5522072R6 (LIVRDIR01)	1632	2096
27	5665262CB1	2755	753-1876	70210554V2	1028	1651
				6851641H1 (BRAIFEN08)	114	712
				8211580H1 (FIBRTXC01)	685	1540
				3476033H1 (LUNGNOT27)	1	70
				70094757V1	1802	2337
				70096200V1	2240	2755
				70210579V2	1530	2148
				7591448H1 (LIVRNOC07)	10	514
28	2119916CB1	2008	916-969, 1-108	3133936F6 (SMCCNOT01)	1767	2008
				2614961F6 (GBLANOT01)	653	1323
				2824140F6 (ADRETUT06)	331	1042
				2614961T6 (GBLANOT01)	1423	1983
				2824140T6 (ADRETUT06)	1187	1973
				7982192H1 (URETTUC01)	1	454
29	8186259CB1	5205	1-3586, 5045-5205	1234960F1 (LUNGFET03)	4776	5044
				7241072H1 (PROSTMY01)	2377	2974
				1671024F6 (BMARNOT03)	4007	4525
				7260802H1 (UTRETMCO1)	1	560
				5305370F6 (MONOTXT02)	1139	1721
				1794133H1 (PROSTUT05)	4267	4527
				2824731F6 (ADRETUT06)	1697	2227
				3235738H1 (COLNUCT03)	1559	1802
				1733636F6 (BRSTTUT08)	3398	3985
				3236301T6 (COLNUCT03)	2148	2861
				5267925F6 (BRAFDIT02)	683	1307
				2228207T6 (SEMVNOT01)	1792	2381
				2824731T6 (ADRETUT06)	2848	3384
				2222302H1 (LUNGNOT18)	4890	5205
				1733636T6 (BRSTTUT08)	4355	5033
				1796831R6 (PROSTUT05)	3590	4137
				3420341F6 (UCMCNOT04)	362	1073
				2603614H1 (LUNGTUT07)	3150	3447
30	70250400CB1	1360	1-27, 136-199	7324412H1 (COLRTUE01)	531	1147
				7608721H1 (COLRTUE01)	245	667
				483037H1 (HNT2RAT01)	3	235
				7716277H1 (SINTFEE02)	813	1360
31	2778782CB1	2075	332-368, 1-49, 2023-2075	71383487V1	667	1312
				8136875J1 (PANHTUR01)	201	901
				71135361V1	1331	1996
				1336105H1 (COLNNOT13)	1	222
				5405901F8 (BRAMNOT01)	1509	2075
				71137687V1	1186	1940
32	2715885CB1	1828	545-1450	3908474T9 (LUNGNOT23)	679	1097

Table 4

Polynucleotide SEQ ID NO:	Incyte ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
				3464795F6 (293TF2T01)	686	1212
				7926352H1 (COLNTUS02)	1140	1740
				6314212H1 (NERDTDN03)	1428	1828
				8001379H1 (LNODTUC02)	1	474
				7292858F8 (BRAIFER06)	56	735
33	1742628CB1	2110	2083-2110, 762-1250	6439575H1 (BRAENOT02)	1196	1883
				284722R6 (CARDNOT01)	1395	1951
				6880673J1 (BRAHTDR03)	383	1155
				6978576H1 (BRAHTDR04)	1	618
				6884027H1 (BRAHTDR03)	629	1228
				5958523H1 (BRATNOT05)	1565	2109
34	2124971CB1	2481	1-265	2186857F6 (PROSNOT26)	571	1112
				6200872H1 (PITUNON01)	1726	2173
				7611233J1 (KIDCTME01)	1153	1752
				8266101J1 (MIXDUNL23)	1894	2481
				7755301H1 (SPLNTUE01)	1305	1767
				7699248H1 (KIDPTDE01)	657	1237
				7674234J1 (NOSETUE01)	1	598
				2186857T6 (PROSNOT26)	1803	2473
35	2258250CB1	1933	984-1261	71362564V1	939	1427
				6918305H1 (PLACFER06)	379	1043
				6609643H1 (EPIGTMC01)	1286	1933
				71002003V1 (SG0000344)	649	1288
				5646649F8 (BRAITUT23)	1	543
36	2626035CB1	2370	1-159	2626035H1 (PROSTUT12)	1	250
				6200872H1 (PITUNON01)	1620	2067
				7611233J1 (KIDCTME01)	1047	1646
				8266101J1 (MIXDUNL23)	1788	2370
				7693212H2 (LNODTUE01)	156	719
				6921814H1 (PLACFER06)	618	1153
				3673721F6 (PLACNOT07)	269	844
				2186857T6 (PROSNOT26)	1697	2367
37	4831382CB1	2534	1-33, 2514- 2534	3269987F6 (BRAINOT20)	1	536
				1393337F6 (THYRNOT03)	1647	2157
				71661886V1	1171	1757
				5720121H1 (PANCNOT16)	2010	2534
				577323R6 (BRAVTXT04)	540	1097
				71663195V1	924	1659
				71659191V1	437	1075
				1439103T6 (PANCNOT08)	1816	2503
38	2122183CB1	2599	2159-2599, 1- 20, 1024-1236	7077721H1 (BRAUTDR04)	1127	1723
				70844379V1	581	1070
				638214T6 (BRSTNOT03)	1942	2593
				70843761V1	1404	2034

Table 4

Polynucleotide SEQ ID NO:	Incyte ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
				3369434F6 (CONNTUT04)	1	688
				2316116H1 (OVARNOT02)	2341	2599
				1510019F6 (LUNGNOT14)	828	1363
39	7484338CB1	3745	2780-3745, 1-66, 2116-2225	4169101T6 (PANCNOT21)	1986	2547
				1546352H1 (PROSTUT04)	3544	3745
				2131004R6 (KIDNNOT05)	2859	3347
				2708511F6 (PONSAT01)	1676	2165
				6757006J1 (SINTFER02)	283	939
				70483292V1	1042	1636
				5408918H1 (BRAMNOT01)	2358	2615
				70483348V1	891	1401
				4406240H1 (PROSDIT01)	1	248
				8196327H1 (BRAINOR03)	1183	1874
				1691757T6 (PROSTUT10)	2965	3607
				5069485H1 (PANCNOT23)	2685	2950
				2499325F6 (ADRETUT05)	2445	2944
				7658334J1 (UTREDME06)	169	748
40	8326588CB1	2323	1-1002, 1128-1764	g7149446	1848	2313
				GNN.g9454509_000002_002.edi t	11	2133
				g1492141	1949	2323
				8326588T1 (BMARNOT03)	1401	1678
				GNN.g6067178_008.edit	1	439

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
21	168639CB1	LUNGNOT37
22	2792817CB1	OVARDIR01
23	3090127CB1	TYMTXT01
24	7480989CB1	MENITUT03
25	2280673CB1	TYMNOT05
26	1517230CB1	PANCTUT01
27	5665262CB1	HEAONOT02
28	2119916CB1	ADRETUT06
29	8186259CB1	PROSTUT05
30	70250400CB1	COLRTUE01
31	2778782CB1	PANHTUR01
32	2715885CB1	LUNGNOT23
33	1742628CB1	BRAFNOT01
34	2124971CB1	PLACNOT07
35	2258250CB1	PLACNOB01
36	2626035CB1	PLACNOT07
37	4831382CB1	THYRNOT03
38	2122183CB1	BRSTNOT07
39	7484338CB1	ADRETUT05
40	8326588CB1	BMARNOT03

Table 6

Library	Vector	Library Description
ADRETUT05	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 52-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated a pheochromocytoma.
ADRETUT06	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
BMARNOT03	pINCY	Library was constructed using RNA isolated from the left tibial bone marrow tissue of a 16-year-old Caucasian male during a partial left tibial osteotomy with free skin graft. Patient history included an abnormality of the red blood cells. Previous surgeries included bone and bone marrow biopsy, and soft tissue excision. Family history included osteoarthritis.
BRAFNOT01	pINCY	Library was constructed using RNA isolated from amygdala tissue and adjacent area removed from the brain of a 35-year-old Caucasian male who died from cardiac failure.
BRSTNOT07	pINCY	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
COLRTUE01	PSPORT1	This 5' biased random primed library was constructed using RNA isolated from rectum tumor tissue removed from a 50-year-old Caucasian male during closed biopsy of rectum and resection of rectum. Pathology indicated grade 3 colonic adenocarcinoma which invades through the muscularis propria to involve pericolonic fat. Tubular adenoma with low grade dysplasia was also identified. The patient presented with malignant rectal neoplasm, blood in stool, and constipation. Patient history included benign neoplasm of the large bowel, hyperlipidemia, benign hypertension, alcohol abuse, and tobacco abuse. Previous surgeries included above knee amputation and vasectomy. Patient medications included allopurinol, Zantac, Darvocet, Centrum vitamins, and an unspecified stool softener. Family history included congestive heart failure in the mother; and benign neoplasm of the large bowel and polypectomy in the sibling(s).
HEAONOT02	pINCY	Library was constructed using RNA isolated from aortic tissue removed from a 10-year-old Caucasian male, who died from anoxia.
LUNGNOT23	pINCY	Library was constructed using RNA isolated from left lobe lung tissue removed from a 58-year-old Caucasian male. Pathology for the associated tumor tissue indicated metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, and an acute duodenal ulcer with hemorrhage. Family history included prostate cancer, breast cancer, and acute leukemia.

Table 6

Library	Vector	Library Description
LUNGN0T37	pINCY	Library was constructed using RNA isolated from lung tissue removed from a 15-year-old Caucasian female who died from a closed head injury. Serology was positive for cytomegalovirus.
MENITUT03	pINCY	Library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer.
OVARDIR01	PCDNA2.1	This random primed library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, vaginal suspension and fixation, and incidental appendectomy. Pathology indicated stromal hyperthecosis of the right and left ovaries. Pathology for the matched tumor tissue indicated a dermoid cyst (benign cystic teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified. The cervix showed squamous metaplasia. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, and primary tuberculous complex.
PANCTUT01	pINCY	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology indicated an invasive grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, benign neoplasm in the large bowel, and a cataract. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
PANHTUR01	PBK-CMV	This random primed library was constructed RNA isolated from pancreatic tumor tissue removed from a 65-year-old female. Pathology indicated well-differentiated neuroendocrine carcinoma (islet cell tumor), nuclear grade 1, forming a dominant mass in the distal pancreas. Multiple smaller tumor nodules were immediately adjacent to the main mass. The liver showed metastatic grade 1 islet cell tumor, forming multiple nodules. Multiple (4) pericholedochal lymph nodes contained metastatic grade 1 islet cell tumor.
PLACNOB01	PBLUESCRIPT	Library was constructed using RNA isolated from placenta.
PLACNOT07	pINCY	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. Serology was positive for anti-CMV (cytomegalovirus).

Table 6

Library	Vector	Library Description
PROSTUT05	PSPORT1	Library was constructed using RNA isolated from prostate tumor tissue removed from a 69-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. Family history included congestive heart failure, multiple myeloma, hyperlipidemia, and rheumatoid arthritis.
THYRNOT03	pINCY	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.
TLYMNOT05	pINCY	Library was constructed using RNA isolated from nonactivated Th2 cells. These cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells.
TLYMTXT01	pINCY	The library was constructed using RNA isolated from activated allogenic T-lymphocyte tissue removed from an adult (40-50-year-old) Caucasian male. The cells were incubated for 6 hours in the presence of OKT3 mAb (1 microgram/ml OKT3), anti-CD28 mAb (1 ug/ml) and 5% human serum. The patient had no allergies.

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 Atwood, 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
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.

Table 7

Program	Description	Reference	Parameter Threshold
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 (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, 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.	

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:1-20,
- 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 biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and
- d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

15 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-20.

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

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

20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:21-40.

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

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

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

30 9. A method 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 has an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

5

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

12. An isolated polynucleotide selected from the group consisting of:

10

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40,
- 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:21-40,
- c) a polynucleotide complementary to a polynucleotide of a),
- 15 d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

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

20

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:

25

- 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.

30

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:

- 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.

5

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

18. A composition of claim 17, wherein the polypeptide has an amino acid sequence selected
10 from the group consisting of SEQ ID NO:1-20.

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

15

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.

20

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
25 functional TRNFR, comprising administering to a patient in need of such treatment a composition of claim 21.

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.

30

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 TRNFR, 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.

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.

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,
- 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.

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

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 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 TRNFR in a biological sample, the method comprising:

- 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.

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

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

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

33. A method of diagnosing a condition or disease associated with the expression of TRNFR 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.

35. A method of diagnosing a condition or disease associated with the expression of TRNFR

in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

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 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

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.

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 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

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.

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.

44. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 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
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 in the sample.

45. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 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 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

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

47. A method of generating a transcript image 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 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 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
5 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.

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

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

15 53. An array of claim 48, further comprising said target polynucleotide hybridized to a 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.

20 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
25 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.

30 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.

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.

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

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.

10

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.

15

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

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.

20

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.

25

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

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.

30

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

76. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:21.

77. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:22.
78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:23.
- 5 79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:24.
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.
- 10 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.
- 15 84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.
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.
- 20 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.
- 25 89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:34.
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.
- 30 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:37.
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.
95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:40.

<110> INCYTE GENOMICS, INC.
 LAL, Preeti G.
 TANG, Y. Tom
 YUE, Henry
 BURFORD, Neil
 GANDHI, Ameena R.
 WARREN, Bridget A.
 YAO, Monique G.
 TRIBOULEY, Catherine M.
 BAUGHN, Mariah R.
 LEE, Ernestine A.
 HAFALIA, April J.A.
 LU, Yan
 GRIFFIN, Jennifer A.
 SANJANWALA, Madhu S.
 DING, Li

<120> TRANSFERASES

<130> PI-0241 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/236,523; 60/238,481; 60/244,025; 60/246,001; 60/247,931;
 60/249,639; 60/252,819

<151> 2000-09-29; 2000-10-06; 2000-10-07; 2000-11-03; 2000-11-09;
 2000-11-16; 2000-11-21

<160> 40

<170> PERL Program

<210> 1

<211> 314

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 168639CD1

<400> 1

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

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

<210> 2

<211> 309

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2792817CD1

<400> 2

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

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

<210> 3

<211> 434

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3090127CD1

<400> 3

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

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

<210> 4

<211> 402

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7480989CD1

<400> 4

Met	Arg	Arg	Arg	Leu	Arg	Leu	Arg	Arg	Asp	Ala	Leu	Leu	Thr	Leu
1				5					10					15
Leu	Leu	Gly	Ala	Ser	Leu	Gly	Leu	Leu	Leu	Tyr	Ala	Gln	Arg	Asp
				20					25					30
Gly	Ala	Ala	Pro	Thr	Ala	Ser	Ala	Pro	Arg	Gly	Arg	Gly	Arg	Ala
				35					40					45
Ala	Pro	Arg	Pro	Thr	Pro	Gly	Pro	Arg	Ala	Phe	Gln	Leu	Pro	Asp

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

<210> 5

<211> 866

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2280673CD1

<400> 5

```

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

```

Ala Gly Asn Ile	Lys Glu Ala Ala Arg	Trp Met Asp Glu Ala Gln
410	415	420
425	430	435
Ala Leu Asp Thr	Ala Asp Arg Phe Ile	Asn Ser Lys Cys Ala Lys
440	445	450
Tyr Met Leu Lys	Ala Asn Leu Ile Lys	Glu Ala Glu Glu Met Cys
455	460	465
Ser Lys Phe Thr	Arg Glu Gly Thr Ser	Ala Val Glu Asn Leu Asn
470	475	480
Glu Met Gln Cys	Met Trp Phe Gln Thr	Glu Cys Ala Gln Ala Tyr
485	490	495
Lys Ala Met Asn	Lys Phe Gly Glu Ala	Leu Lys Lys Cys His Glu
500	505	510
Ile Glu Arg His	Phe Ile Glu Ile Thr	Asp Asp Gln Phe Asp Phe
515	520	525
His Thr Tyr Cys	Met Arg Lys Ile Thr	Leu Arg Ser Tyr Val Asp
530	535	540
Leu Leu Lys Leu	Glu Asp Val Leu Arg	Gln His Pro Phe Tyr Phe
545	550	555
Lys Ala Ala Arg	Ile Ala Ile Glu Ile	Tyr Leu Lys Leu His Asp
560	565	570
Asn Pro Leu Thr	Asp Glu Asn Lys Glu	His Glu Ala Asp Thr Ala
575	580	585
Asn Met Ser Asp	Lys Glu Leu Lys Lys	Leu Arg Asn Lys Gln Arg
590	595	600
Arg Ala Gln Lys	Lys Ala Gln Ile Glu	Glu Glu Lys Lys Asn Ala
605	610	615
Glu Lys Glu Lys	Gln Gln Arg Asn Gln	Lys Lys Lys Lys Asp Asp
620	625	630
Asp Asp Glu Glu	Ile Gly Gly Pro Lys	Glu Glu Leu Ile Pro Glu
635	640	645
Lys Leu Ala Lys	Val Glu Thr Pro Leu	Glu Glu Ala Ile Lys Phe
650	655	660
Leu Thr Pro Leu	Lys Asn Leu Val Lys	Asn Lys Ile Glu Thr His
665	670	675
Leu Phe Ala Phe	Glu Ile Tyr Phe Arg	Lys Glu Lys Phe Leu Leu
680	685	690
Met Leu Gln Ser	Val Lys Arg Ala Phe	Ala Ile Asp Ser Ser His
695	700	705
Pro Trp Leu His	Glu Cys Met Ile Arg	Leu Phe Asn Thr Ala Val
710	715	720
Cys Glu Ser Lys	Asp Leu Ser Asp Thr	Val Arg Thr Val Leu Lys
725	730	735
Gln Glu Met Asn	Arg Leu Phe Gly Ala	Thr Asn Pro Lys Asn Phe
740	745	750
Asn Glu Thr Phe	Leu Lys Arg Asn Ser	Asp Ser Leu Pro His Arg
755	760	765
Leu Ser Ala Ala	Lys Met Val Tyr Tyr	Leu Asp Pro Ser Ser Gln
770	775	780
Lys Arg Ala Ile	Glu Leu Ala Thr Thr	Leu Asp Glu Ser Leu Thr
785	790	795
Asn Arg Asn Leu	Gln Thr Cys Met Glu	Val Leu Glu Ala Leu Tyr
800	805	810
Asp Gly Ser Leu	Gly Asp Cys Lys Glu	Ala Ala Glu Ile Tyr Arg
815	820	825
Ala Asn Cys His	Lys Leu Phe Pro Tyr	Ala Leu Ala Phe Met Pro

	830		835		840
Pro Gly Tyr Glu	Glu Asp Met Lys Ile	Thr Val Asn Gly Asp	Ser		
	845		850		855
Ser Ala Glu Ala	Glu Glu Leu Ala Asn	Glu Ile			
	860		865		

<210> 6

<211> 828

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1517230CD1

<400> 6

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

9/50

Leu	Gln	Arg	Leu	Leu	Gly	Pro	Leu	Leu	Glu	Ala	Tyr	Ser	Ser	Ala
				725					730					735
Ala	Ile	Phe	Val	His	Asn	Phe	Ser	Gly	Pro	Val	Pro	Glu	Pro	Glu
				740					745					750
Tyr	Leu	Gln	Lys	Leu	His	Lys	Tyr	Leu	Ile	Thr	Arg	Thr	Glu	Arg
				755					760					765
Asn	Val	Ala	Val	Tyr	Ala	Glu	Ser	Ala	Thr	Tyr	Cys	Leu	Val	Lys
				770					775					780
Asn	Ala	Val	Lys	Met	Phe	Lys	Asp	Ile	Gly	Val	Phe	Lys	Glu	Thr
				785					790					795
Lys	Gln	Lys	Arg	Val	Ser	Val	Leu	Glu	Leu	Ser	Ser	Thr	Phe	Leu
				800					805					810
Pro	Gln	Cys	Asn	Arg	Gln	Lys	Leu	Leu	Glu	Tyr	Ile	Leu	Ser	Phe
				815					820					825
Val	Val	Leu												

<210> 7

<211> 801

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5665262CD1

<400> 7

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

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

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

Asn His His Ser

<210> 9

<211> 555

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8186259CD1

<400> 9

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

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

<210> 10
 <211> 401
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 70250400CD1

<400> 10

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

Gly	Phe	Lys	Thr	Phe	Gly	Ile	Ser	Arg	Asn	Arg	Asn	Ser	Arg	Met
				350					355					360
Asn	Lys	Glu	Pro	Cys	Phe	Phe	Arg	Ala	Met	Leu	Val	Val	His	Lys
				365					370					375
Leu	Leu	Pro	Pro	Glu	Leu	Leu	Ala	Met	Trp	Gly	Leu	Val	His	Ser
				380					385					390
Asn	Leu	Thr	Cys	Ser	Arg	Lys	Leu	Gln	Val	Leu				
				395					400					

<210> 11

<211> 336

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2778782CD1

<400> 11

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

	275		280		285
His Glu Arg Gly Arg Lys Gly Ser His		His Arg Phe Ile Thr Glu			
	290		295		300
Lys Arg Val Phe Lys Asn Trp Ala Arg Thr		Phe Asn Ile His Phe			
	305		310		315
Phe Gln Pro Asp Trp Lys Pro Glu Ser Leu		Ala Ile Asn His Pro			
	320		325		330
Glu Asn Lys Pro Val Phe					
	335				

<210> 12

<211> 353

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2715885CD1

<400> 12

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

```

Cys Gln Leu Leu Leu Thr Ile Leu Pro Arg Arg Gln Leu Pro Pro
      275      280      285
Gly Leu Glu Ser Leu Asn Val Ser Val Ala Ala Gly Ile Leu Leu
      290      295      300
His Ser Ile Cys Ser Gln Arg Lys Gly Phe Pro Thr Glu Gly Glu
      305      310      315
Arg Arg Gln Leu Leu Gln Asp Pro Gln Glu Pro Ser Ala Arg Ser
      320      325      330
Glu Gly Leu Ser Met Ala Gln His Pro Gly Leu Ser Ser Gly Pro
      335      340      345
Glu Lys Glu Arg Gln Asn Glu Gly
      350

```

<210> 13

<211> 499

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1742628CD1

<400> 13

```

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

```

	245		250		255
Gly Lys His Phe Trp	Ser Phe Gln Met Tyr	Gly Glu Asp Phe Val			
	260		265		270
Pro Asp Ile Val Thr	Met Gly Lys Pro Met	Gly Asn Gly His Pro			
	275		280		285
Val Ala Cys Val Val	Thr Thr Lys Glu Ile	Ala Glu Ala Phe Ser			
	290		295		300
Ser Ser Gly Met Glu	Tyr Phe Asn Thr Tyr	Gly Gly Asn Pro Val			
	305		310		315
Ser Cys Ala Val Gly	Leu Ala Val Leu Asp	Ile Ile Glu Asn Glu			
	320		325		330
Asp Leu Gln Gly Asn	Ala Lys Arg Val Gly	Asn Tyr Leu Thr Glu			
	335		340		345
Leu Leu Lys Lys Gln	Lys Ala Lys His Thr	Leu Ile Gly Asp Ile			
	350		355		360
Arg Gly Ile Gly Leu	Phe Ile Gly Ile Asp	Leu Val Lys Asp His			
	365		370		375
Leu Lys Arg Thr Pro	Ala Thr Ala Glu Ala	Gln His Ile Ile Tyr			
	380		385		390
Lys Met Lys Glu Lys	Arg Val Leu Leu Ser	Ala Asp Gly Pro His			
	395		400		405
Arg Asn Val Leu Lys	Ile Lys Pro Pro Met	Cys Phe Thr Glu Glu			
	410		415		420
Asp Ala Lys Phe Met	Val Asp Gln Leu Asp	Arg Ile Leu Thr Val			
	425		430		435
Leu Glu Glu Ala Met	Gly Thr Lys Thr Glu	Ser Val Thr Ser Glu			
	440		445		450
Asn Thr Pro Cys Lys	Thr Lys Met Leu Lys	Glu Ala His Ile Glu			
	455		460		465
Leu Leu Arg Asp Ser	Thr Thr Asp Ser Lys	Glu Asn Pro Ser Arg			
	470		475		480
Lys Arg Asn Gly Met	Cys Thr Asp Thr His	Ser Leu Leu Ser Lys			
	485		490		495
Arg Leu Lys Thr					

<210> 14

<211> 721

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2124971CD1

<400> 14

Met Leu Pro Arg Gly	Arg Pro Arg Ala Leu	Gly Ala Ala Ala Leu			
1	5	10			15
Leu Leu Leu Leu Leu	Leu Leu Gly Phe Leu	Leu Phe Gly Gly Asp			
	20	25			30
Leu Gly Cys Glu Arg	Arg Glu Pro Gly Gly	Arg Ala Gly Ala Pro			
	35	40			45
Gly Cys Phe Pro Gly	Pro Leu Met Pro Arg	Val Pro Pro Asp Gly			
	50	55			60
Arg Leu Arg Arg Ala	Ala Ala Leu Asp Gly	Asp Pro Gly Ala Gly			
	65	70			75

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

Gly	Pro	Leu	Tyr	Pro	Val	Asn	Gln	Leu	Arg	Asn	Val	Ala	Leu	Ala
				500					505					510
Gln	Ala	Leu	Thr	Pro	Tyr	Val	Phe	Leu	Ser	Asp	Ile	Asp	Phe	Leu
				515					520					525
Pro	Ala	Tyr	Ser	Leu	Tyr	Asp	Tyr	Leu	Arg	Ala	Ser	Ile	Glu	Gln
				530					535					540
Leu	Gly	Leu	Gly	Ser	Arg	Arg	Lys	Ala	Ala	Leu	Val	Val	Pro	Ala
				545					550					555
Phe	Glu	Thr	Leu	Arg	Tyr	Arg	Phe	Ser	Phe	Pro	His	Ser	Lys	Val
				560					565					570
Glu	Leu	Leu	Ala	Leu	Leu	Asp	Ala	Gly	Thr	Leu	Tyr	Thr	Phe	Arg
				575					580					585
Tyr	His	Glu	Trp	Pro	Arg	Gly	His	Ala	Pro	Thr	Asp	Tyr	Ala	Arg
				590					595					600
Trp	Arg	Glu	Ala	Gln	Ala	Pro	Tyr	Arg	Val	Gln	Trp	Ala	Ala	Asn
				605					610					615
Tyr	Glu	Pro	Tyr	Val	Val	Val	Pro	Arg	Asp	Cys	Pro	Arg	Tyr	Asp
				620					625					630
Pro	Arg	Phe	Val	Gly	Phe	Gly	Trp	Asn	Lys	Val	Ala	His	Ile	Val
				635					640					645
Glu	Leu	Asp	Ala	Gln	Glu	Tyr	Glu	Leu	Leu	Val	Leu	Pro	Glu	Ala
				650					655					660
Phe	Thr	Ile	His	Leu	Pro	His	Ala	Pro	Ser	Leu	Asp	Ile	Ser	Arg
				665					670					675
Phe	Arg	Ser	Ser	Pro	Thr	Tyr	Arg	Asp	Cys	Leu	Gln	Ala	Leu	Lys
				680					685					690
Asp	Glu	Phe	His	Gln	Asp	Leu	Ser	Arg	His	His	Gly	Ala	Ala	Ala
				695					700					705
Leu	Lys	Tyr	Leu	Pro	Ala	Leu	Gln	Gln	Pro	Gln	Ser	Pro	Ala	Arg
				710					715					720

Gly

<210> 15

<211> 552

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2258250CD1

<400> 15

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

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

				515						520					525
Leu	Leu	Gln	Leu	Lys	Tyr	Ser	Arg	His	Lys	Lys	Ser	Ala	Arg	Pro	
				530						535				540	
Glu	Leu	Tyr	Asp	Glu	Thr	Ser	Phe	Glu	Leu	Glu	Asp				
				545						550					

<210> 16

<211> 690

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2626035CD1

<400> 16

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

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

<210> 17

<211> 607


```

Leu Gly Ile Arg Val Trp Gln Cys Gly Gly Ser Val Glu Val Leu
      365                      370                      375
Pro Cys Ser Arg Ile Ala His Ile Glu Arg Ala His Lys Pro Tyr
      380                      385                      390
Thr Glu Asp Leu Thr Ala His Val Arg Arg Asn Ala Leu Arg Val
      395                      400                      405
Ala Glu Val Trp Met Asp Glu Phe Lys Ser His Val Tyr Met Ala
      410                      415                      420
Trp Asn Ile Pro Gln Glu Asp Ser Gly Ile Asp Ile Gly Asp Ile
      425                      430                      435
Thr Ala Arg Lys Ala Leu Arg Lys Gln Leu Gln Cys Lys Thr Phe
      440                      445                      450
Arg Trp Tyr Leu Val Ser Val Tyr Pro Glu Met Arg Met Tyr Ser
      455                      460                      465
Asp Ile Ile Ala Tyr Gly Val Leu Gln Asn Ser Leu Lys Thr Asp
      470                      475                      480
Leu Cys Leu Asp Gln Gly Pro Asp Thr Glu Asn Val Pro Ile Met
      485                      490                      495
Tyr Ile Cys His Gly Met Thr Pro Gln Asn Val Tyr Tyr Thr Ser
      500                      505                      510
Ser Gln Gln Ile His Val Gly Ile Leu Ser Pro Thr Val Asp Asp
      515                      520                      525
Asp Asp Asn Arg Cys Leu Val Asp Val Asn Ser Arg Pro Arg Leu
      530                      535                      540
Ile Glu Cys Ser Tyr Ala Lys Ala Lys Arg Met Lys Leu His Trp
      545                      550                      555
Gln Phe Ser Gln Gly Gly Pro Ile Gln Asn Arg Lys Ser Lys Arg
      560                      565                      570
Cys Leu Glu Leu Gln Glu Asn Ser Asp Leu Glu Phe Gly Phe Gln
      575                      580                      585
Leu Val Leu Gln Lys Cys Ser Gly Gln His Trp Ser Ile Thr Asn
      590                      595                      600
Val Leu Arg Ser Leu Ala Ser
      605

```

<210> 18

<211> 375

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2122183CD1

<400> 18

```

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

```

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

<210> 19

<211> 760

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7484338CD1

<400> 19

Met Ile Pro Asn Gln His Asn Ala Gly Ala Gly Ser His Gln Pro					
1	5		10		15
Ala Val Phe Arg Met Ala Val Leu Asp Thr Asp Leu Asp His Ile					
	20		25		30
Leu Pro Ser Ser Val Leu Pro Pro Phe Trp Ala Lys Leu Val Val					

				35					40				45	
Gly	Ser	Val	Ala	Ile	Val	Cys	Phe	Ala	Arg	Ser	Tyr	Asp	Gly	Asp
				50					55					60
Phe	Val	Phe	Asp	Asp	Ser	Glu	Ala	Ile	Val	Asn	Asn	Lys	Asp	Leu
				65					70					75
Gln	Ala	Glu	Thr	Pro	Leu	Gly	Asp	Leu	Trp	His	His	Asp	Phe	Trp
				80					85					90
Gly	Ser	Arg	Leu	Ser	Ser	Asn	Thr	Ser	His	Lys	Ser	Tyr	Arg	Pro
				95					100					105
Leu	Thr	Val	Leu	Thr	Phe	Arg	Ile	Asn	Tyr	Tyr	Leu	Ser	Gly	Gly
				110					115					120
Phe	His	Pro	Val	Gly	Phe	His	Val	Val	Asn	Ile	Leu	Leu	His	Ser
				125					130					135
Gly	Ile	Ser	Val	Leu	Met	Val	Asp	Val	Phe	Ser	Val	Leu	Phe	Gly
				140					145					150
Gly	Leu	Gln	Tyr	Thr	Ser	Lys	Gly	Arg	Arg	Leu	His	Leu	Ala	Pro
				155					160					165
Arg	Ala	Ser	Leu	Leu	Ala	Ala	Leu	Leu	Phe	Ala	Val	His	Pro	Val
				170					175					180
His	Thr	Glu	Cys	Val	Ala	Gly	Val	Val	Gly	Arg	Ala	Asp	Leu	Leu
				185					190					195
Cys	Ala	Leu	Phe	Phe	Leu	Leu	Ser	Phe	Leu	Gly	Tyr	Cys	Lys	Ala
				200					205					210
Phe	Arg	Glu	Ser	Asn	Lys	Glu	Gly	Ala	His	Ser	Ser	Thr	Phe	Trp
				215					220					225
Val	Leu	Leu	Ser	Ile	Phe	Leu	Gly	Ala	Val	Ala	Met	Leu	Cys	Lys
				230					235					240
Glu	Gln	Gly	Ile	Thr	Val	Leu	Gly	Leu	Asn	Ala	Val	Phe	Asp	Ile
				245					250					255
Leu	Val	Ile	Gly	Lys	Phe	Asn	Val	Leu	Glu	Ile	Val	Gln	Lys	Val
				260					265					270
Leu	His	Lys	Asp	Lys	Ser	Leu	Glu	Asn	Leu	Gly	Met	Leu	Arg	Asn
				275					280					285
Gly	Gly	Leu	Leu	Phe	Arg	Met	Thr	Leu	Leu	Thr	Ser	Gly	Gly	Ala
				290					295					300
Gly	Met	Leu	Tyr	Val	Arg	Trp	Arg	Ile	Met	Gly	Thr	Gly	Pro	Pro
				305					310					315
Ala	Phe	Thr	Glu	Val	Asp	Asn	Pro	Ala	Ser	Phe	Ala	Asp	Ser	Met
				320					325					330
Leu	Val	Arg	Ala	Val	Asn	Tyr	Asn	Tyr	Tyr	Tyr	Ser	Leu	Asn	Ala
				335					340					345
Trp	Leu	Leu	Leu	Cys	Pro	Trp	Trp	Leu	Cys	Phe	Asp	Trp	Ser	Met
				350					355					360
Gly	Cys	Ile	Pro	Leu	Ile	Lys	Ser	Ile	Ser	Asp	Trp	Arg	Val	Ile
				365					370					375
Ala	Leu	Ala	Ala	Leu	Trp	Phe	Cys	Leu	Ile	Gly	Leu	Ile	Cys	Gln
				380					385					390
Ala	Leu	Cys	Ser	Glu	Asp	Gly	His	Lys	Arg	Arg	Ile	Leu	Thr	Leu
				395					400					405
Gly	Leu	Gly	Phe	Leu	Val	Ile	Pro	Phe	Leu	Pro	Ala	Ser	Asn	Leu
				410					415					420
Phe	Phe	Arg	Val	Gly	Phe	Val	Val	Ala	Glu	Arg	Val	Leu	Tyr	Leu
				425					430					435
Pro	Ser	Val	Gly	Tyr	Cys	Val	Leu	Leu	Thr	Phe	Gly	Phe	Gly	Ala
				440					445					450
Leu	Ser	Lys	His	Thr	Lys	Lys	Lys	Lys	Leu	Ile	Ala	Ala	Val	Val

				455					460					465
Leu	Gly	Ile	Leu	Phe	Ile	Asn	Thr	Leu	Arg	Cys	Val	Leu	Arg	Ser
				470					475					480
Gly	Glu	Trp	Arg	Ser	Glu	Glu	Gln	Leu	Phe	Arg	Ser	Ala	Leu	Ser
				485					490					495
Val	Cys	Pro	Leu	Asn	Ala	Lys	Val	His	Tyr	Asn	Ile	Gly	Lys	Asn
				500					505					510
Leu	Ala	Asp	Lys	Gly	Asn	Gln	Thr	Ala	Ala	Ile	Arg	Tyr	Tyr	Arg
				515					520					525
Glu	Ala	Val	Arg	Leu	Asn	Pro	Lys	Tyr	Val	His	Ala	Met	Asn	Asn
				530					535					540
Leu	Gly	Asn	Ile	Leu	Lys	Glu	Arg	Asn	Glu	Leu	Gln	Glu	Ala	Glu
				545					550					555
Glu	Leu	Leu	Ser	Leu	Ala	Val	Gln	Ile	Gln	Pro	Asp	Phe	Ala	Ala
				560					565					570
Ala	Trp	Met	Asn	Leu	Gly	Ile	Val	Gln	Asn	Ser	Leu	Lys	Arg	Phe
				575					580					585
Glu	Ala	Ala	Glu	Gln	Ser	Tyr	Arg	Thr	Ala	Ile	Lys	His	Arg	Arg
				590					595					600
Lys	Tyr	Pro	Asp	Cys	Tyr	Tyr	Asn	Leu	Gly	Arg	Leu	Tyr	Ala	Asp
				605					610					615
Leu	Asn	Arg	His	Val	Asp	Ala	Leu	Asn	Ala	Trp	Arg	Asn	Ala	Thr
				620					625					630
Val	Leu	Lys	Pro	Glu	His	Ser	Leu	Ala	Trp	Asn	Asn	Met	Ile	Ile
				635					640					645
Leu	Leu	Asp	Asn	Thr	Gly	Asn	Leu	Ala	Gln	Ala	Glu	Ala	Val	Gly
				650					655					660
Arg	Glu	Ala	Leu	Glu	Leu	Ile	Pro	Asn	Asp	His	Ser	Leu	Met	Phe
				665					670					675
Ser	Leu	Ala	Asn	Val	Leu	Gly	Lys	Ser	Gln	Lys	Tyr	Lys	Glu	Ser
				680					685					690
Glu	Ala	Leu	Phe	Leu	Lys	Ala	Ile	Lys	Ala	Asn	Pro	Asn	Ala	Ala
				695					700					705
Ser	Tyr	His	Gly	Asn	Leu	Ala	Val	Leu	Tyr	His	Arg	Trp	Gly	His
				710					715					720
Leu	Asp	Leu	Ala	Lys	Lys	His	Tyr	Glu	Ile	Ser	Leu	Gln	Leu	Asp
				725					730					735
Pro	Thr	Ala	Ser	Gly	Thr	Lys	Glu	Asn	Tyr	Gly	Leu	Leu	Arg	Arg
				740					745					750
Lys	Leu	Glu	Leu	Met	Gln	Lys	Lys	Ala	Val					
				755					760					

<210> 20

<211> 710

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8326588CD1

<400> 20

Met	Asp	Gln	Val	Ala	Thr	Leu	Arg	Leu	Glu	Ser	Val	Asp	Leu	Gln
1				5					10					15
Ser	Ser	Arg	Asn	Asn	Lys	Glu	His	His	Thr	Gln	Glu	Met	Gly	Val
				20					25					30

Lys	Arg	Leu	Thr	Val	Arg	Arg	Gly	Gln	Pro	Phe	Tyr	Leu	Arg	Leu	35	40	45
Ser	Phe	Ser	Arg	Pro	Phe	Gln	Ser	Gln	Asn	Asp	His	Ile	Thr	Phe	50	55	60
Val	Ala	Glu	Thr	Gly	Pro	Lys	Pro	Ser	Glu	Leu	Leu	Gly	Thr	Arg	65	70	75
Ala	Thr	Phe	Phe	Leu	Thr	Arg	Val	Gln	Pro	Gly	Asn	Val	Trp	Ser	80	85	90
Ala	Ser	Asp	Phe	Thr	Ile	Asp	Ser	Asn	Ser	Leu	Gln	Val	Ser	Leu	95	100	105
Phe	Thr	Pro	Ala	Asn	Ala	Val	Ile	Gly	His	Tyr	Thr	Leu	Lys	Ile	110	115	120
Glu	Ile	Ser	Gln	Gly	Gln	Gly	His	Ser	Val	Thr	Tyr	Pro	Leu	Gly	125	130	135
Thr	Phe	Ile	Leu	Leu	Phe	Asn	Pro	Trp	Ser	Pro	Glu	Asp	Asp	Val	140	145	150
Tyr	Leu	Pro	Ser	Glu	Ile	Leu	Leu	Gln	Glu	Tyr	Ile	Met	Arg	Asp	155	160	165
Tyr	Gly	Phe	Val	Tyr	Lys	Gly	His	Glu	Arg	Phe	Ile	Thr	Ser	Trp	170	175	180
Pro	Trp	Asn	Tyr	Gly	Gln	Phe	Glu	Glu	Asp	Ile	Ile	Asp	Ile	Cys	185	190	195
Phe	Glu	Ile	Leu	Asn	Lys	Ser	Leu	Tyr	His	Leu	Lys	Asn	Pro	Ala	200	205	210
Lys	Asp	Cys	Ser	Gln	Arg	Asn	Asp	Val	Val	Tyr	Val	Cys	Arg	Val	215	220	225
Val	Ser	Ala	Met	Ile	Asn	Ser	Asn	Asp	Asp	Asn	Gly	Val	Leu	Gln	230	235	240
Gly	Asn	Trp	Gly	Glu	Asp	Tyr	Ser	Lys	Gly	Val	Ser	Pro	Leu	Glu	245	250	255
Trp	Lys	Gly	Ser	Val	Ala	Ile	Leu	Gln	Gln	Trp	Ser	Ala	Arg	Gly	260	265	270
Gly	Gln	Pro	Val	Lys	Tyr	Gly	Gln	Cys	Trp	Val	Phe	Ala	Ser	Val	275	280	285
Met	Cys	Thr	Val	Met	Arg	Cys	Leu	Gly	Val	Pro	Thr	Arg	Val	Val	290	295	300
Ser	Asn	Phe	Arg	Ser	Ala	His	Asn	Val	Asp	Arg	Asn	Leu	Thr	Ile	305	310	315
Asp	Thr	Tyr	Tyr	Asp	Arg	Asn	Ala	Glu	Met	Leu	Ser	Thr	Gln	Lys	320	325	330
Arg	Asp	Lys	Ile	Trp	Asn	Phe	His	Val	Trp	Asn	Glu	Cys	Trp	Met	335	340	345
Ile	Arg	Lys	Asp	Leu	Pro	Pro	Gly	Tyr	Asn	Gly	Trp	Gln	Val	Leu	350	355	360
Asp	Pro	Thr	Pro	Gln	Gln	Thr	Ser	Ser	Gly	Leu	Phe	Cys	Cys	Gly	365	370	375
Pro	Ala	Ser	Val	Lys	Ala	Ile	Arg	Glu	Gly	Asp	Val	His	Leu	Ala	380	385	390
Tyr	Asp	Thr	Pro	Phe	Val	Tyr	Ala	Glu	Val	Asn	Ala	Asp	Glu	Val	395	400	405
Ile	Trp	Leu	Leu	Gly	Asp	Gly	Gln	Ala	Gln	Glu	Ile	Leu	Ala	His	410	415	420
Asn	Thr	Ser	Ser	Ile	Gly	Lys	Glu	Ile	Ser	Thr	Lys	Met	Val	Gly	425	430	435
Ser	Asp	Gln	Arg	Gln	Ser	Ile	Thr	Ser	Ser	Tyr	Lys	Tyr	Pro	Glu	440	445	450


```

caatccagct gctgaggaaa gacagcatca tggagaatgg cactctgtgc tatgaggtgc 600
attataggac gaaggatgtc cactttaatg cctacacagg ggccaccatt cgatttggcc 660
aattcctctc cacatccctc ctgaaagaag aggcacagga gtttgggaac cagacactat 720
ttaccatatt cacctgcctg ggtgcacctg tacagtactt ctccctcaag aaggaagtct 780
tgatccctcc ctatgagctg tttaaagtta taaatatgag ctaccacca agaggagact 840
ggttgacgtt gaggtcaact gggaacctga gcacatataa ctgtcagctg ctaaaagctt 900
ccagcaagaa atgcatccct gatcctatag ctattgcac tctctccttt ttgaccagt 960
tcatcatctt ttccaaaagc agagtataaa gaaatcttgt ggctcctttt atttaaaaaa 1020
aatattttaa taaaatcttt tttattggca aaaaaaaaaa aaaaaaaaag atctttaatt 1080
aagcgggtcca agctta 1096

```

<210> 22

<211> 1380

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2792817CB1

<400> 22

```

acagccaggt tagatgttct gaggaggcgg gagcaaccga gagagcacgt gagcatctgt 60
cctttctacc cgttcctctt tatcttttagt gttcagtagc agcggggata gcccggggcc 120
cgggtgtatgg ccacggagtt acagtgtccg gactccatgc cctgtcacia ccagcaagta 180
aactctgcct caaccccaag tcccagacag ctgcgacctg gcgatctgat cctggaccac 240
gcagggggaa acagagcctc cagggccaag gtgattctcc tcacggggta cgccattct 300
agcctgccgg ccgagctgga ctctggggcc tgcggcggt ccagcctcaa ctgagaggc 360
aacagtggta gtggtgacag tagcagctat gacgcaccag ctggcaactc cttcctagag 420
gactgcgaac tctcccggca gatcggggcg cagcttaagc tgctgcctat gaatgatcag 480
atacgggagc tacagaccat catccgggac aagacagcca gtagagggtga cttcatgttt 540
tctgcggtac gtttgatcag acttggttggt gaagagggt tgaatcagct gccatataaa 600
gaatgcatgg tgaccactcc aacagggtac aagtatgaag gagtgaaatt tgagaaggga 660
aattgtgggg tcagcataat gagaagcggg gaggcaatgg aacaagggtt acgagactgc 720
tgtcgatcca tacgaattgg aaagatcctg attcagagt atgaggagac acaaagagcc 780
aaagtatat atgccaaatt cccccagac atttaccgga gaaaagtcct tctgatgtat 840
ccaatttct cactggaaa tactgtaatt gaagctgtaa aggttcttat agaacatgga 900
gttcaaccga gtgttatcat cctactcagt ctgttctcca ctccatcatg tgccaaatca 960
atcattcagg agtttccaga gatcacaatt ttaactactg aagttcatcc tgttgacact 1020
acacattttg gacagaaata ctttggaaca gactaagtta tttaagtaaa ataattgtct 1080
tatgtaatat tacaatcatg ttttgatttt ctatttgttt tactgattca cttgagggtg 1140
gcagagaaaa atgtgttaaa atgcttttta gttttggaag tgggtatat ttgaggttata 1200
tctcatttag ttatttgttt actgttggca ccgaattcaa caatgaagta tatgcaactc 1260
ttacaaaaca taaattttaa taatattcta atgcaaatta ctgaaccctc agtgcattaa 1320
aattatttcc taattaatga ctttcagca cactgcgccc gtatatagt agtgcggctc 1380

```

<210> 23

<211> 2647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3090127CB1

<400> 23

```

ggaaatcagg caccgggccc ggccgggttc tggctgcgct cgcgcgctct gcccgcgccc 60

```

```

cggtgtgcct ccgcttacct gcagctccga ccactggctc gcgctaccca ggtctccgca 120
cgccgcggtg gcttcagccc agacctgggc agccagcgga gaaagagtta actggcaggg 180
gcgaggagga gccagggag gaaggaagga tattgccgta attctgaaag tttttttcct 240
tcctctcttc ccttcgcaga ggtgagtgcc gggtcggcg ctctgctcct ggagctcccc 300
cgggactgcc tggggacagg gactgctgtg gcgtcggcc ctccactgcg gacctctcct 360
gagtgggtgc gccgagtcac ggagggcgca gactggccg ggaagatcct ttccacctgg 420
ctgacgtggt ttctcggctt catcctttta ccttcggtct tcggagtgtc tctgggcatc 480
tccgagatct acatgaagat cctagtgaac actttagagt gggccacaat acgaattgaa 540
aaaggaaccc caaaggagtc gattccttaa aactctgctt ctggttggtat tatccaaaga 600
gatgagtcac ccatggaaaa agggctctct ggtctacgag gaagggactt tgagctgtct 660
gacgtgtttt atttctccaa gaagggattg gaagccattg tagaagatga agtgaccag 720
aggttttcct cagaggagct agtgtcatgg aatctcctca caagaaccaa tgtaaatttc 780
cagtacatca gtctgcggct cactatgggt tgggtgctgg gcgtcatagt gcgctattgt 840
gtcctactgc ctctgagggg taccttggtt ttcatggga tcagtttgct ggttatagga 900
actacactgg ttgggcagct gccagacagc agcctcaaaa actggctgag tgaactggct 960
catctgactt gctgccggat ctgtgtgcga gccctctctg gtaccattca ttatcataac 1020
aagcagtaca gacccagaa gggaggcatt tgtgttgcca accatacttc cccattgat 1080
gttttaactc tgacaacgga tggatgttat gctatggttg gccaggttca tggcggcttg 1140
atgggaatta ttcagagagc tatggtcaag gcttgcctc atgtctggtt tgaacgctca 1200
gaaatgaagg atcgacacct gggtactaag agactaaaag aacatattgc tgataagaag 1260
aaactaccca tactaatttt tcctgaagga acttgcatca acaatacttc agtcatgatg 1320
tttaaaaagg ggagctttga aattggagga accatacatc cagttgcaat taagtataac 1380
cctcagttcg gtgatgcatt ttggaacagt agtaaataca acatggtgag ctacctgctt 1440
cgaatgatga ccagctgggc catcgtctgt gacgtgtggt acatgcccc catgaccaga 1500
gaggaaggag aagatgcagt ccagtttgct aacagggtta agtctgctat tgctatacaa 1560
ggaggcctga ctgaacttcc ctgggatgga ggactaaaga gagcaaagg gaaggacatc 1620
tttaaggaag agcagcagaa aaattacagc aagatgattg tgggcaatgg atctctcagc 1680
taagaggacg gatgacagcc ttttagatcta gaactagccc ttagaaatgg aatggctttt 1740
tttgttttgt tttgttttat tgttttgttt ttattattgt taatcttttc tacagaatga 1800
ttgtctctac ctctttatgc cagaggcaga acctacaggt gccctttttg gcttttgttg 1860
ttgttgtaac attagcccca tggattgtaa ggtgggtttac tgagttaaaa cagattctgc 1920
ttttgtaaaa tgatggcatc actgtggact gaatgaaata tttgtataga aaaaagtgc 1980
tgaaaagtgt gtttggaaact catcgatagg gtaattctcc aaaaatgcc aaactctctt 2040
tctgtaatta gccttgccac tttcttcagt cacttaaatg gtgagattac acatcagtgc 2100
aagatgacca ttatgggttat ggtctactgc aaggttgaaa ggaaaaatgg aggattgtat 2160
ttaggaaaag ggacaacttt gtggccacct gctctgaaag tcaaaaaggaa atgtaaatta 2220
gtgtcattag tgtgttgga gagaaatact attcagtaag cttcgccaaa gaaaagtga 2280
tcaaagttaa tgtgtgtgtg catttatatg taggcagctc gtagaccaca ttttagccag 2340
caactggtaa caaagagctt agttttcctt gtttgaatgc tgtagatctg tacctagtac 2400
ccctcccac tactgatttg tttgtttttg taaccaaaca cattttcaga tagaaggagc 2460
cttaaaaaaa aaaaaatcac attgagtaac ttcagtatga atgaatgaga gtgtgtggag 2520
ctacccctca ccctccaccc ctttgtgctt tttattcccg aattttccca gtctcttaaa 2580
cagaaaaatg actgatataa ttatcttttg gaaactgagc cttaattttt ttttagaggg 2640
gaaataa 2647

```

<210> 24

<211> 2117

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7480989CB1

<400> 24

acggggcagg ttgcatttgg ctggggcccg gtccatggcg gcccctcgga ccctgcgctg 60

```

agccccggag gccagggcgt cgggggctgc gccacttccg agggccgagt cgctgcctgg 120
tcccggcggg gcgacacggc cgggaggagg gagaacaacg caaggggggt acaaccgtcg 180
gtcgttgagg cccccccggg ggcgtggcct cccgccccct cagctgggga gggcggggct 240
cgctgcccc tgctgccgac tgcgaccctt acaggggagg gagggcgag tgcgcgcgg 300
agatgaggag gaggtgcgc ctacgcaggg acgcattgct cagctgctc cttggcgct 360
ccctgggcct cttactctat gcgcagcgcg acggcgcggc cccgacggcg agcgcgcgc 420
gagggcgagg gagggcgga ccgaggcca ccccggaacc ccgcgcttc cagttaccg 480
acggcgggtg agccccggc gcctacgaag gggacacacc ggcgcgcgcc acgcctacg 540
gaccctttga cttcgccgc tatttgcgcg ccaaggacca ggcggggtt ccactgctc 600
ttaaccagcc gcacaagtgc cgcggcgacg gcgcaccggg tggcgcccg gacctgcta 660
ttgctgtcaa gtcggtggca gaggacttcg agcggcgcca agcgtgcgc cagacgtggg 720
gcgcggaggg tcgctgcag gggcgctgg tgcgcgcgt gttcttgct ggcgtgcca 780
ggggcgaggg ctcgggggg gccgacgaag ttggggagg cgcggaacc cactggcgcg 840
ccctgctgcg ggccgagagc cttgcgtatg cggacatcct gctctgggc ttcgacgaca 900
ccttttttaa cctaacgctc aaggagatcc actttctagc ctgggcctc gctttctgcc 960
ccgacgtgcg cttcgttttt aaggcgacg cagatgtgtt cgtgaacgt ggaaatctcc 1020
tggagtccct ggcgcgcgg gaccggcg cagactgct tgcgtgtgac gtaattgtgc 1080
atgcgcggcc catccgcacg cgggctagca agtactacat ccccgaggcc gtgtacggcc 1140
tgcccgccca tccggcctac gggggcgcg gtggctttgt gctttccgg gccacgtgc 1200
accgcctggc tggcgctgt gcgcaggtcg agctcttccc catcgacgac gtctttctg 1260
gcatgtgtct gcagcgctg cggctcacgc ccgagcctca ccctgccttc cgcaccttg 1320
gcatccccc gccttcagcc ggcgcgcat tgagcacctt cgaccctgc ttttaccgtg 1380
agctggttgt agtgacggg ctctcgccg ctgacatctg gcttatgtg cgctgctgc 1440
acgggcccga tgggcccagc tgtgcgcat cacagcctgt cgctgcagg cccttccaat 1500
gggactccta gctccccact acagcccaa gctcctaact cagaccaga atggagccg 1560
tttcccagat tattgcgctg tatgtggtt tccctgatc accaggtgc tgtctccaca 1620
ggatcccagg ggatgggggt taagcttggc tctggcggt ccacctgct ggaaccagt 1680
gaaacccgtg taatggtgac ctttgagcg agccaaggct ggggtgtaga tgaccatct 1740
ttgtccaaca ggtcccagag cagtggatat gtctggctct cctagtagca cagaggtgt 1800
ttctggtgtg gtggcaggga cttagggaat cctaccactc tgctggattt ggaacccct 1860
aggctgacgc ggacgtatgc agaggctctc aaggccaggc cccacaggga ggtggaggg 1920
ctccggccgc cacagcctga attcatgaac ctggcaggca ctttgccata gctcatctga 1980
aaacagatat tatgcttccc acaacctctc ctgggccag gtgtggctga gcaccaggga 2040
tggagccaca cataagggac aaatgagtgc acgtcctac ctagtcttct ctcacttcct 2100
gactcacaca acaatgc 2117

```

<210> 25

<211> 3320

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2280673CB1

<220>

<221> unsure

<222> 3136

<223> a, t, c, g, or other

<400> 25

```

ctgcctctgt cggctctgtt agttaccacg tgaaccgccg acggagacct gtagtggggg 60
aggcggcggc agcgttaagt gagaaaggaa aaaagacaac gagggaaaag gaggtgtccg 120
ggtagggcaa cgcggcgaca cccgaggcct ggtgggtggc gcggatcgag atattcaagg 180
ctgaagcagc tacggaacgg cagcggcggc ggtcggacaa actgactgac cgagccgggt 240
ggtggcgagg gcagcgggag cagccggaac gatgccggcc gtgagcctcc cgccaaggga 300

```

```

gaatgcgctc ttcaagcggg tcttgagggt ttatgaacat aaacagtata gaaatggatt 360
gaaattctgt aaacaaatac tttctaatac caaatttgca gagcatggag aaaccttggc 420
tatgaaagga ttaacattga actgttttggg gaaaaaggaa gaagcttatg aattggttcg 480
tagaggtttg agaaatgact tgaagagtca tgtgtgttgg cacgtttatg gccttcttca 540
gaggtcagac aagaagtatg atgaagccat taagtgttac agaaatgcac taaaatggga 600
taaagacaat cttcaaactc taagggacct ttccttacta cagattcaaa tgcgagatct 660
tgagggttac agggaaacga ggtatcagtt acttcagctt cgacctgcgc agagagcatc 720
atggatttgt tatgctattg cttaccatct attagaagat tatgaaatgg cagcaaagat 780
tttagaagaa tttaggaaaa cacaacagac atcccctgac aagggtggatt atgaatatag 840
tgaactactc ttatatcaga atcaagttct tcgggaagca ggtctctata gagaagcttt 900
ggaacatctt tgtacctatg aaaagcagat ttgtgataaa cttgctgtag aagaaaccaa 960
aggggaactt ctgttgcaac tatgtcgttt ggaagatgct gcagatgttt atagaggatt 1020
gcaagagaga aatcctgaaa actgggccta ttacaaaggc ttggaaaaag cactcaagcc 1080
agctaatatg ttagaacggc taaaaattta tgaggaaaggc tggactaaat atcccagggg 1140
actggtgcc aagaaggctgc cgttaaactt tttatctggt gagaagttta aagaatgttt 1200
ggataagttc ctaaggatga atttcagcaa gggttgccca ccagtcttca atactttaag 1260
atcattatac aaagacaaag aaaaggtggc aatcatagaa gagttagtag taggttatga 1320
aacctctcta aaaagctgcc ggttatttta cccaatgat gatggaaagg aggaaccacc 1380
aaccacatta ctttgggtcc agtactactt ggcaacaacat tatgacaaaa ttggtcagcc 1440
atctattgct ttggagtaca taaatactgc tattgaaagt acacctacat taatagaact 1500
ctttctcgtg aaagctaaaa tctataagca tgctggaaat attaaagaag ctgcaagggtg 1560
gatggatgag gccaggcct tggacacagc agacagatct atcaactcca aatgtgcaaa 1620
atacatgcta aaagccaacc tgattaaaga agctgaagaa atgtgctcaa agtttacaag 1680
ggaaggaaca tcagcggtag agaatttgaa tgaaatgcag tgcattgtgt tccaaacaga 1740
atgtgcccag gcttataaag caatgaataa atttggtgaa gcacttaaga aatgtcatga 1800
gattgagaga cattttatag aaatcactga tgaccagttt gactttcata catactgtat 1860
gaggaagatt acccttagat catatgtgga cttattaaaa ctagaagatg tacttcgaca 1920
gcatccattt tacttcaagg cagcaagaat tgctatagag atctatttga agcttcatga 1980
caacccccct acagatgaga ataaagaaca cgaagctgat acagcaaaca tgtctgacaa 2040
agagctaaag aagctacgta ataaacaaag aagagctcaa aagaaagccc agatagaaga 2100
agagaaaaaa aatgcagaaa aagaaaagca gcagagaaat cagaaaaaga agaaggatga 2160
tgatgatgag gagataggag gtccaaaaga agaacttatt ccagagaaac tggccaaggt 2220
tgaaactcca ttggaagaag ctattaaatt tttaacaccg ttgaagaact tgggtgaaga 2280
caagatagag actcatcttt ttgcctttga gatttacttt aggaaagaaa agtttctttt 2340
gatgctacaa tcagtaaaga gggcatttgc tattgattct agtcatccct ggcttcatga 2400
gtgtatgatt cgtctcttta atactgcagt gtgtgaaagt aaagatttat ctgatacagt 2460
tagaacagta ttaaaacaag aaatgaatcg tctttttgga gcaacgaatc caaagaattt 2520
taatgaaact tttctgaaaa ggaattctga ttcatgcca cacagattat cagctgccaa 2580
aatggtatat tacttagatc cttctagtca gaagcgagct atagagtttg caacaacact 2640
tgatgaatct ctactaaca gaaacctcca gacatgtatg gaggtatttg aagccttgta 2700
tgatggtagc ctaggagact gtaaagaagc tgctgaaatt tatagagcaa attgtcataa 2760
gcttttccct tatgcttttg ctttcatgcc tcctggatat gaagaggata tgaagatcac 2820
agttaatgga gatagttctg cagaagctga agaactggcc aatgaaattt gaacatcact 2880
aaacaagcaa atggaatgac tttggaccat atctagtata taatattttt gtcacgcacc 2940
tgctgcattg ctctaactta cacagaatga gaggagtaaa tgttcttgcc tcaaatagt 3000
gttttacgtt ttttatcctg ctgaaaaagt atatataaaa tatctaacat tacaggatag 3060
aggttcagtt tcttaaaaaa ttaaagctgc taaaattgag tggttaaaaa agatacctta 3120
tactattact cccanccac ccatgttttt aaactaattt atatgaaatc tggaggctgt 3180
tacagctaac aaagcagggt tgtgggggaa atattacttt aaatttgtgc tgtgagattt 3240
tactatatct cagacagcat aaatgctggt gtagcactgg gttctttcac tgagcacaag 3300
aagtgtgggg ggcttagaac 3320

```

<210> 26

<211> 3210

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1517230CB1

<400> 26

```
gagaatttgc tgcctgcgggg ggactctttc tgaggttact gtggagcacc caaagtctgt 60
cagcctctgg ccgtgcaaac aggcacccag aggaaccaga ccttgcttat tcaccacacag 120
cctgggactg tcttctccag agtctccatc agctttgcta atcgactgat tggaaataat 180
tcctcaaaca ccaccaagtc aaggatacag gcagcagcgg ctcccctggt gtatggacat 240
tctgcacccg aaactgatag ctgagtcctg aagttttatg ttatgaaaca gaagaacttt 300
catcccagca catgatttgg gaattacact ttgtgacatg gatgaatctg cactgaccct 360
tggtacaata gatgtttctt atctgccaca ttcacagaa tacagtgttg gtcgatgtaa 420
gcacacaagt gaggaatggg gtgagtgtgg ctttagaccc accatcttca gatctgcaac 480
tttaaaatgg aaagaaaagg taatgagtcg gaaaaggcca tttgttggaa gatgttgtaa 540
ctcctgcact cccagagct gggacaaatt tttcaacccc agtatcccg ctttgggttt 600
gcggaatggt atttatatca atgaaactca cacaagacac cgcggatggc ttgcaagacg 660
cctttcttac gttcttttta ttcaagagcg agatgtgcat aaggcgatgt ttgccaccaa 720
tgtgactgaa aatgtgctga acagcagtag agtacaagag gcaattgcag aagtggctgc 780
tgaattaaac cctgatgggt ctgccagca gcaatcaaaa gccgttaaca aagtgaaaaa 840
gaaagctaaa aggattcttc aagaaatggg tgccactgtc tcaccggcaa tgatcagact 900
gactgggtgg gtgctgctaa aactgttcaa cagcttcttt tggaacattc aaattcacia 960
aggtcaactt gagatgggta aagctgcaac tgagacgaat ttgccgcttc tgtttctacc 1020
agttcataga tcccatattg actatctgct gctcactttc attctcttct gccataacat 1080
caaagcacca tacattgctt caggcaataa tctcaacatc ccaatcttca gtaccttgat 1140
ccataagctt gggggcttct tcatacgacg aaggctcgat gaaacaccag atggacggaa 1200
agatgttctc tatagagctt tgctccatgg gcatatagtt gaattacttc gacagcagca 1260
attcttgagg atcttctctg aaggcacacg ttctaggagt ggaaaaacct cttgtgctcg 1320
ggcaggactt ttgtcagttg tggtagatac tctgtctacc aatgtcatcc cagacatctt 1380
gataatacct gttggaatct cctatgatcg cattatcgaa ggtcactaca atggtgaaca 1440
actgggcaaa cctaagaaga atgagagcct gtggagtgtg gcaagagggt ttattagaat 1500
gttacgaaaa aactatgggt gtgtccgagt ggattttgca cagccatttt ccttaaagga 1560
atatttagaa agccaaagtc agaaaccggg gtctgctcta ctttccctgg agcaagcgtt 1620
gttaccagct atacttcctt caagaccag tgatgctgct gatgaaggta gagacacgtc 1680
cattaatgag tccagaaatg caacagatga atccctacga aggagggttg ttgcaaactt 1740
ggctgagcat attctattca ctgctagcaa gtccctgtgc attatgtcca cacacattgt 1800
ggcttgccct ctctctaca gacacaggca gggaattgat ctctccacat tggctgaaga 1860
cttctttgtg atgaaagagg aagtcctggc tcgtgatttt gacctggggt tctcaggaaa 1920
ttcagaagat gtagtaatgc atgccataca gctgctggga aattgtgtca caatcaccca 1980
cactagcagg aacgatgagt tttttatcac cccagcaca actgtcccat cagtcttcga 2040
actcaacttc tacagcaatg gggacttca tgtctttatc atggaggcca tcatagcttg 2100
cagcctttat gcagtctga acaagagggg actggggggg cccactagca cccacctaa 2160
cctgatcagc caggagcagc tgggtgcggaa ggcgccagc ctgtgctacc ttctctccaa 2220
tgaaggcacc atctcactgc cttgccagac attttaccaa gtctgccatg aaacagtagg 2280
aaagtttatc cagtatggca ttcttacagt ggcagagcac gatgaccagg aagatatcag 2340
tcctagtctt gctgagcagc agtgggacaa gaagcttcct gaacctttgt cttggagaag 2400
tgatgaagaa gatgaagaca gtgactttgg ggaggaacag cgagattgct acctgaagg 2460
gagccaatcc aaggagcacc agcagtttat caccttctta cagagactcc ttgggccttt 2520
gctggaggcc tacagctctg ctgccatctt tgttcacaa ttcagtgggt ctgttcaga 2580
acctgagtat ctgcaaaagt tgcacaaata cctaataacc agaacagaaa gaaatgttg 2640
agtattgctt gagagtgcga catattgtct tgtgaagaat gctgtgaaaa tgtttaagga 2700
tattgggggt ttcaaggaga ccaaacaaaa gagagtgtct gttttagaac tgagcagcac 2760
ttttctacct caatgcaacc gacaaaaact tctagaatat attctgagtt ttgtggtgct 2820
gtaggtaacg tgtggcactg ctggcaaatg aaggtcatga gatgagttcc ttgtaggtac 2880
cagcttctgg ctcaagagtt gaaggtgccg tcgcagggtc aggcctgcc tgtcccgaag 2940
tgatctcctg gaagacaagt gccttctccc tccatggatc tgtgatcttc ccagctctgc 3000
atcaacacag cagcctgcag ataacacttg gggggacctc agcctctatt cgcaactcat 3060
```

```

aatccgtaga ctacaagatg aaatctcaat aaattatttt tgagttttatt aaagattgac 3120
attttaagta caacttttaa ggactaatta ctgtgatgga cacagaaatg tagctgtgtt 3180
ctggaactga atcttacatg gtataacttag 3210

```

<210> 27

<211> 2755

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5665262CB1

<400> 27

```

gagcagtcga cgccttgtgg cggcttttgcg gagctgctgc tttggcgggg gttggaagct 60
ggtgtgaggt tctgtgggga gaaggagagt gccagaggtg actggttcat ggttcttcta 120
ggctctcatg gccacatgt tggaaggcag atgccaaact cagccaagga gcagccccag 180
tgcccgagag gctagcctgt ggtcgtcagg ctttgggatg aagctggagg ctgtcactcc 240
attcctgggc aagtatcgcc cctttgtggg tcgctgttgc cagacctgca cccccaagag 300
ctgggagtcc ctcttcacac gaagcataac ggacctaggc ttctgcaatg tgatcctggt 360
gaaggaggag aacacaaggt ttccgggctg gctgggttcg aggcctctgt attcctgtg 420
gtccctggag cagcacatcc ccccttgcca ggatgtccca cagaagatca tggaaagcac 480
cggggtgcag aacctcctct caggaggagg cccaggaggc actggggaag gccagggtgc 540
tgacctgtg aagaaggagg tacagcgcat cctgggtcac atccaggccc caccctgtcc 600
cttctgtgtc aggcctgttca gctgggctgc gctgaggttc ctgaactgcc tgttcttgaa 660
tgtgcagctc cacaagggtc agatgaagat ggtccagaag gccgcccagg caggcttgcc 720
gcttgtcctc ctctctactc acaaaaccct cctggatggg atcctgctgc cctttatgct 780
gctctcccag ggcctgggtg tgcttcgtgt ggcctgggac tcccgccctc gctccctgtc 840
cctcagagct ctgctgagga agcttggggg gcttttctct ccccagagg ccagcctctc 900
cctggacagc tctgaggggc tccttgccag ggctgtggtc caggcgggtc tagagcagct 960
gctggttagt gggcagcccc tgctcatctt cctggaggaa cctcctgggg ctctggggcc 1020
acggctgtca gccctgggac aggcctgggt ggggtttgtg gtgcaggcag tccagggtgg 1080
catcgtccca gatgctctgc tgggtaccag gcccgctacc tatgacctgg ttccggatgc 1140
accgtgtgac atagaccatg cctcggtccc cctggggctg tggacaggag ctctggctgt 1200
cctacgtagc ttgtggagcc gctggggctg cagccaccgg atctgctccc ggggtgcacct 1260
agctcagccc ttttccctgc aggaatacat cgtcagtgc agaagctgt ggggcggcag 1320
acagaccctg gagcagctac tgcagcccat cgtgctgggc caatgtactg ctgtcccaga 1380
cactgagaag gagcaggagt ggacccccat aactgggcct ctctggccc tcaaggaaga 1440
ggaccagctc ctggtcagga gactgagctg tcatgtcctg agtgccagt tagggagctc 1500
tgcggtgatg agcacggcca ttatggcaac gctgctgctc ttcaagcatc agaagggtgt 1560
gttctgtctg cagctcctgg gggagttctc ctggctgacg gaggagatac tgttgctgtg 1620
ctttgatgta ggcttctctg ggcagctgcg gagcctgctg cagcactcac tgagcctgct 1680
gcgggcgcac gtggccctgc tgcgcattcc tcagggtgac ttgctggtgg tgccgcagcc 1740
tggcccaggc ctcacacacc tggcacaact gagtgtgag ctgctgccc tcttctgag 1800
cgaggctgtg ggcgcctgtg cagtgcgggg gctgctggca ggcagagtgc cgccccagg 1860
gccttgggag ctgcagggca tattgtgtgt gagccagaat gagctgtacc gccagatcct 1920
gctgctgatg cacctgctgc cgcaagacct gctgctgcta aagccctgcc agtcttctta 1980
ctgctactgt caggaggtgc tggaccggct catccaatgc gggctcctgg ttgtgagga 2040
gacccaggc tcccgccag cctgtgacac agggcgacag cgattgagca gaaagctgct 2100
gtgaaaccg agtggggact ttactgatag tgacagtgat gacttcggag aggtgacgg 2160
ccgtacttcc aggcctcagcc agcagtcaca ctgcccagat ttctttcttt tctctgccc 2220
cctgctcagc ccgctgctca aggcctttgc acaggctgcc gccttctctc gccagggcca 2280
gctgcccgat actgagttgg gctacacaga gcagctgttc cagttcctgc aggccaccgc 2340
ccaggaagaa gggatcttct agtgtgcgga cccaaagctc gccatcagtg ctgtctggac 2400
cttcagagac ctagggttct tgcagcagac gccgagccct gcaggcccca ggctccacct 2460
gtcccctact tttgccagcc tggacaatca ggaaaaacta gaacagttca tccggcagtt 2520

```

```

catttgtagc tagaactgtg aggaggagcc tgtgctgaga cttctcagcc ccagaacaca 2580
gctgtgtcct agagccagaa gatggagagg aggctgcaaa cccttagctg ctctataaat 2640
ataatcattg aggcttgatt gtcccttgcc atctcttgct ttttcccttc tttgatgtga 2700
taaacaaggg gacgagacga gttgtctttt cccagcccca gcagcaaaaa aaaaa 2755

```

<210> 28

<211> 2008

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2119916CB1

<400> 28

```

aggaagtctg gcgcatccga gccggccagc cgagcacatc tggaagtggg ttctggggcc 60
gccccctctc gccagcgcaa ctctggggtt cccagcggct tcgcgagag gtggaagaaa 120
cccagagcgt tccgaagtca acgcaagcaa aggggagtgc gggtcgggga ggaatattct 180
tttggaacg taatatggc cttggggctc tccagccctt tgggacttcc aatgggatct 240
tagaagcagc cgaagcagcg tgagggcggc agcccagggc cagccacgat ttgaacgctc 300
tgccttgtag ctctctgga ccgaggagcc caaagcccta ccctcacatc tcaccaggctc 360
ctgtgggaag agcagcgtgg aggtgggctg aggttagaag gtgcagagcg tggagaaga 420
ttgtgagctg agtattggac atctgttctt gaatagtccc tgggcctgcc ataggaaagg 480
aagttctcca gggttacagt tcttatccgc gtgaatacac atggctctgt tacgaaaaat 540
taatcagggtg ctgctgttcc ttctgatcgt gacctctgt gtgattctgt ataagaaagt 600
tcataagggg actgtgcca agaatacgc agatgatgaa tccgagactc ctgaagaact 660
ggaagaagag attcctgtgg tgatttgtgc tgcagcaggg aggatgggtg ccactatggc 720
tgccatcaat agcatctaca gcaacactga cgccaacatc ttgttctatg tagtgggact 780
ccggaatact ctgactcgaa tacgaaaatg gattgaacat tccaaactga gagaaataaa 840
ctttaaaatc gtggaattca acccgatggg cctcaaaggg aagatcagac cagactcatc 900
gaggcctgaa ttgctccagc ctctgaactt tgttcgattt tatctccctc tacttatcca 960
ccaacacgag aaagtcatct atttggacga tgatgtaatt gtacaagggtg atatccaaga 1020
actgtatgac accaccttgg ccctgggcca cgcggcggct ttctcagatg actgcgattt 1080
gccctctgct caggacataa acagactcgt gggacttcag aacacatata tgggctatct 1140
ggactaccgg aagaaggcca tcaaggacct tggcatcagc cccagcacct gctctttcaa 1200
tcctgggtgtg attgttgcca acatgacaga atggaagcac cagcgcatca ccaagcaatt 1260
ggagaaatgg atgcaaaaaga atgtggagga aaacctctat agcagctccc tgggaggagg 1320
ggtggccacc tccccaatgc tgattgtgtt tcatgggaaa tattccacaa ttaacccct 1380
gtggcacata aggcacctgg gctggaatcc agatgccaga tattcggagc attttctgca 1440
ggaagctaaa ttactccact ggaatggaag acataaacct tgggacttcc ctagtgttca 1500
caacgactta tgggaaagct ggtttgttcc tgacctgca gggatattta aactcaatca 1560
ccatagctga tataactcta cccttaaaat attccctgta tagaaatgtg gaattgtccc 1620
tttgtagcca actataacat tgttctttat gaatattacc tttgatacat atgatccaca 1680
atataaaaac caaaaactac tgtgtgcaaa ttataccttg gaccatatag gcattgatta 1740
acttctttta gtacatgtga taactatgga aatcaagatt atgtgactga aaaacataaa 1800
ggaagagacc catctagata acagcaatca acctgcttaa ttctgaatga caattatata 1860
cacaaatctt taaaacttct acatgtatct ttcacatgaa gatctcctta acaggttgcc 1920
aaccttttct tttataaaac tattacattt aaaatatgga cgtctgaaaa ataaaaatatt 1980
catcattttt atgaaaaaaa aaaaaaaaa 2008

```

<210> 29

<211> 5205

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8186259CB1

<400> 29

```

agagccctaa gccctgcctc ccggtcctgg ccgggtttcc cagaactgca cggcgcctct 60
ccgcccaggc ccaagcgcgga gcccctcctc cacaccgag tccgagcccc gcgtccggat 120
tcggaccgcg ctgcctgggg cggtgctgca ccaggcgcg gtgtggcagg cgtctcggag 180
cgccagggtg agcttcctgg tcaagatggg ccgcgcctgc cgctcggtag ccgggctcct 240
gccacgcgcg cgccgctgct tccccgccc ggccccgctg ctgcgcgctc ccctctgcct 300
cctgtgctgg accccggcgg ctgtgcgcgc ggtccctgag ctcgggctct ggttagagac 360
agtcaacgac aaatcaggac ctttgatatt taggaaaact atgtttaact ctacagatat 420
caagttatct gttaagtcac tccattgttc tgggctctgt aagtttacca tagtgtggca 480
tttgaagtat catacctgtc acaatgagca ttctaactcg gaagagctgt tccaaaaaca 540
taaacttagt gttgatgaag acttttgcga ttatttgaag aatgacaact gttggacaac 600
aaaaaatgaa aacttagatt gcaacagtga ttcacagggt tttccctctt tgaataataa 660
agaactaata aatatcagaa atgtttcaaa ccaggaaaga tcaatggatg ttgtagccag 720
aacacaaaaa gatgggtttc atatctttat tgtttctatt aaaacggaga atacagatgc 780
aagctggaat ttgaatgttt ctctttctat gattgggcct catggatata tctctgcac 840
agattggccc ctaatgattt ttacatggg gatgtgtatt gtttatatat tatatggcat 900
actctggctg acgtggtctg cctgttattg gaaagatata ttaagaatcc agttctggat 960
tgcagctggt atttttttgg gaatgcttga aaaagcagtt ttttatagtg aataccaaaa 1020
catcagcaac actggactgt caacccaagg cttattgata tttgoggagt tgatttctgc 1080
gattaagagg acgttggctc gccttctcgt gatcattgtg agcctgggct atggcattgt 1140
gaagcctcgt ttaggaacag tcatgcaccg ggtgatcgga ctggggcttc tatacttaac 1200
ctttgcagct gttgaaggcg tgatgagagt cattgggggt tctaaccatt tagctgttgt 1260
tcttgatgac attatttttag cagttattga ctccattttt gtgtggttca tttttattag 1320
tttggcacia actatgaaga ccctaaggct aagaaagaac actgtgaaat tttcattata 1380
tagacatttt aaaaatactc tgatctttgc tgtgctggct tctatagtgt ttatggggtg 1440
gacaactaag acatttagaa ttgcaaaatg ccaatcagat tggatggaac gctgggttga 1500
cgatgcattt tggagcttcc ttttttcgct tatccttatt gtaatcatgt ttttgtggag 1560
accatcagca aacaatcaga gatatgcctt catgccctta atagatgatt ctgatgatga 1620
aattgaggaa ttcattgtaa cttctgaaaa tttaaccgaa ggaataaaat taagagcctc 1680
aaaatcagtt tccaatggaa cagctaagcc tgccacttct gagaactttg atgaagattt 1740
gaagtgggta gaagaaaata ttccctcttc attcacagat gtagctcttc cagtgttagt 1800
ggattcagat gaggaaatca tgaccagatc tgaaatggct gaaaaaatgt tctcttcaga 1860
aaagataatg tgattggaac ccgtataaga aatgtagtta agcctgaagg actatccttc 1920
atcaagactg aaagtgaagt ttgatttgat attgcctaaa aatttttatt gtgttatctt 1980
ggaagtctgt gtatcaaaat gaagaattca gatggttaga ggttctatag tccttttaaa 2040
gctgactctt gagtgtcagt tgaatatcca ttaaatgga tttggaaata acctgaggaa 2100
agtattatga taaagatctg cacagatgcc tcttagctga taggtggcag gcctgtgggt 2160
ttgggttctc cctcttttct ctggaacata tgacaattcc agattaaaga aaaatgtttt 2220
ttaataaata cccttgggtc ttcttctagt cacccttgag gtagatattg tgattttctg 2280
gagtatagta tatccgtgtc tctgtgtctt aggttacta gatgcaataa tacttctctt 2340
tgacatttgt actgaagtga tttgatatta agtaaaacag ttaatgtttg aatataggca 2400
tatttatagg ttttttccgc tcccccccaa ccacccttt ttaaaaaatc tatacaaacg 2460
ccttgtttga gtctcatcat gcacatcaaa tcatggagt aggtcttctc tgagctcagg 2520
ggaacacaag tgcacagaga gagatgtctt gagggctact accaaagaat taccctcatt 2580
gtccctcact caggccatgt gtacatgcga tgcgtctgag tgtgctgggg tgggtgggtg 2640
ccacgtggct ccccagagc acttccctaac tggcaagctg ggagacccat tactggtgaa 2700
ctttgtggaa attagaactg tatcttttac ataactcttg catattacat ttcataataa 2760
aaacatacat ttagttgcat gctacatcac tattgatttt ataattaatt tcttaagctt 2820
caaccatggt ttatacctta tttcgttaca tcatatatat gtaatgtgta atatgaaatc 2880
ttttgcttta atgtcttttt ttaaaatgta gaatgttcta aacttgaaag gcaattgaat 2940
gtagtatgat gaaaatgtga atgttttgct gctttcatga ccaaagatac agggctagt 3000
gacatttaga ataataatta aagctagagt cttgtatgtc ttttctttga aggagtctta 3060
accttgtaaa ttgagaatga cttcagagaa ttttgattaa gaaaacatta aaatcttaac 3120

```

```

cggcacaaac actccaattt ttttcaactgt gaagccgcaa gcaattttttt ttcttttttct 3180
ttcaaaagcc tacctttctga atttattttct tgtttactca tttcagagag ggtagtaaag 3240
aagatctatt tctggtagtc atatcgcttg aaaggtagtg gtaaagtgtg tttcagtcgt 3300
gaccatgtgg aaagtgaaca gtgttggcaa acattaccga gaaaatcatg cttttcaaga 3360
tgcccttgct ttgggatata cttcctaggg agaaaaaaaa aaagtagttt aacaattgtg 3420
aattccattt cttattttcag tttctgctgc agtaatgggt tcccacccac tataattccc 3480
agcatttatg ttctgttgta ttctccctt agcccagtaa cattttttatc taatacccca 3540
ttccccaagt tttgagacag attgaccccc tactcattat gtggctctag ttgaatttta 3600
aaatgtggaa tattgggctt gcaggcagta ggagctgcaa atctggtaga gtgggagtgt 3660
ggagttaatg gtgagtatgt taataaaggg aaactgtctc tgacagaatc tcagtaatgt 3720
ttacaaaaac atgtctttct acagctggta ggataaatga tgctaccctg tagctcagct 3780
acaggctgca gtgcaaactt ttcttccatc cagagaaagc agaattccct cctagtaacc 3840
tcattacaaa tactgttact agaagggcat gtgctgtctg tcaccttcag taatatttgt 3900
gccatctctt gatgactgat gacctggatc gagtattttct atgaagggtc ttcttaggcc 3960
ccttacatac gcaagagggg tgctctagtg ccatagctgt agttcacagg aaggacacca 4020
ggagaagtta tacctagggc tactgagcag ctcatcatcc ctgtttctgc acagtttctc 4080
gaaactggcc atcagggcct ctgaggcact caaatcagtt tacttttagc atgcccccat 4140
cagggtgggt ctcaactgtta gtgaggatac gggctctgggt tgatgttttt ctaggcaaaa 4200
tgcttaagtg ttctgggttat gccattcatt catacgaatg gtgaaatttg cttaaaaggg 4260
aattttcatg atttgattta gattagtatt taaatatctg ctttagatag caattaattt 4320
tattgtaaaa ataaggaaaa atatgtgaat atgtgaattt tttaagcctg agagatgata 4380
gaatgttccc atattttttct tgtaaagaaa ataataattt aacttacaca tcctgtagaa 4440
aataccacct tttcccttg tattacagta caatgtttac attactatac tgtcaagctg 4500
aaagtataaa aaatgtacat atacattttg agttatgtat ccttttttta aaaaaagtt 4560
cgagtctggt gcactaggct gtacatgact aaagttgaca gatgctatgc tagatttata 4620
atcactagtt ctggtacttg tgtctttgta tgatcaaagc atgcaataag caatacaaaa 4680
taccaagcct tatacttaaa agaagtttaa catattgggt aatatactgg ttaataact 4740
ggttaaacat attgaatgta tataagtggc aaaactagat ttttaaggaa gtgtacatta 4800
taatattgga gctcagtact gcatgaagag acttcattaa aactaagaaa acattttatt 4860
ggggagaaat tttaggcatt taagaacttg tatttttcta ttttaaaaag tttaaattatt 4920
ccgtaatttg gaagaagttt cggttgatgt aggacataac cgtttgagg gttttcattt 4980
gaaaaattga tgtattttgt gccttaatat tttgttcttt taataaaaat gctctgaatt 5040
tgaatgattg attcttgata gtattttatt gtgctagatt atataaatct gtaggacata 5100
tagacatata tagactccaa tagatggcag gacataaaat tcttaaaaca ggaggcacag 5160
ctaatacaac atgggaaaaa tgtcaacctc agtaagtggg aaaga 5205

```

<210> 30

<211> 1360

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 70250400CB1

<400> 30

```

cgccgctcgc ccctccgccg ctccggcccg ggccgccatg tcgctgtgga agaaaaccgt 60
ctaccggagt ctgtgcctgg ccctggccct gctcgtggcc gtgacggtgt tccaacgcag 120
tctcaccctt ggtcagtttc tgcaggagcc tccgccacc accctggagc cacagaaggc 180
ccagaagcca aatggacagc tgggtgaacc caacaacttc tggaagaacc cgaaagatgt 240
ggctgcgccc acgcccattg cctctcaggg gccccaggcc tgggacgtga ccaccactaa 300
ctgctcagcc aatatcaact tgacccacca gccctggttc caggctcctg agccgcagtt 360
ccggcagttt ctcttctacc gccactgccg ctacttcccc atgctgctga accaccgga 420
gaagtgcagg ggcgatgtct acctgctggg ggttgctcaag tcggtcatca cgcagcacga 480
ccgccgcgag gccatccgcc agacctgggg ccgcgagcgg cagtccgcgg gtgggggccg 540
aggcgccgtg cgcacctctt tcctgctggg cagggcctcc aagcaggagg agcgcacgca 600

```

```

ctaccagcag ctgctggcct acgaagaccg cctctacggc gacatcctgc agtggggctt 660
tctcgacacc ttcttcaacc tgacctcaa ggagatccac ttctcaagt ggctggacat 720
ctactgcccc cacatccccct tcattttcaa aggcgacgat gacgtcttcg tcaacccac 780
caacctgcta gaatttctgg ctgaccggca gccacaggaa aacctgttcg tgggcgatgt 840
cctgcagcac gctcgcccc ttcgcaggaa agacaacaaa tactacatcc cgggggccct 900
gtacggcaag gccagctatc cgcctatgc aggcggcggg ggcttctca tggccggcag 960
cctggccggg cgcctgcacc atgcctgcga caccctggag ctctaccga tcgacgacgt 1020
ctttctgggc atgtgcctgg aggtgctggg cgtgcagccc acggcccacg agggcttcaa 1080
gactttcggc atctcccgga accgcaacag ccgcatgaac aaggagccgt gctttttccg 1140
cgccatgctc gtggtgcaca agctgctgcc ccctgagctg ctgccatgt gggggctggt 1200
gcacagcaat ctacactgct ccgcaagct ccaggtgctc tgacccacgc cgggctacta 1260
ggacaggcca gggcacttgc tcctgagccc ccattggtatt ggggctggag ccacagtgcc 1320
caggcctagc ctttggtccc caaggggagg tggagggttg 1360

```

<210> 31

<211> 2075

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2778782CB1

<400> 31

```

ctgagagact acgaggggtcc ggttcagttt taattctgtc tctaattctt gcaacagccg 60
cgcttccccg gtcccgccgc tcccgcgcgc gatctgccgc ggccggctgc tgggcaaaaa 120
tcagagccgc ctccgcccc ttacccatca tggaaaccct ccaggaaaaa gtggcccccg 180
acgcgcgagc ctgaggattc tgcacaaaag aggtgcccaa aatgaagacc ctgatgcgcc 240
atggtctggc agtgtgttta gcgctacca ccattgtcac cagcttgttg ctagtgtaca 300
gcagcctcgg cggccagaag gagcggcccc cgcagcagca gcagcagcag cagcaacagc 360
agcagcaggc gtccggccacc ggcagctcgc agccggcggc ggagagcagc acccagcagc 420
gccccggggt ccccgcgga ccgcggccac tggacggata cctcggagtg gcggaccaca 480
agccccgtga aatgcactgc agggactgtg ccctgtgac cagctcaggg catctgctgc 540
acagtcggca aggtctccag attgaccaga cagagtgtgt catccgcagc aatgacgccc 600
ccacacgcgg ctatggcggt gacgtgggca atcgaccag cctgagggtc atcgcgcat 660
ccagcatcca gaggatcctc cgcaaccgcc atgacctgct caacgtgagc cagggcaccg 720
tgttcatctt ctggggcccc agcagctaca tgcggcgga cggcaagggc caggtctaca 780
acaacctgca tctcctgagc caggtgctgc cccggtgaa ggcttcatg attactcgcc 840
acaagatgct gcagtttgat gagctcttca agcaggagac tggcaaagac aggaagatat 900
ccaacacttg gctcagcact ggctggttta caatgacaat tgcactggag ctctgtgaca 960
ggatcaatgt ttatggcatg gtgccccag acttctgcag ggatcccaat cacccttcag 1020
taccttatca ttattatgac ccttttggac ctgatgaatg tacaatgtac ctctcccatg 1080
agcgaggagc caagggcagt catcaccgct ttatcacaga gaaacgagtc tttagaact 1140
gggcacggac attcaatatt cacttttttc aaccagactg gaaaccagaa tcacttgcta 1200
taaatcatcc tgagaataaa cctgtgttct aaggaatgag catgccagac tgtaatcca 1260
ggtattcact gcatcagaca ccgagacact gaacttctg agccaccaga caggaaagg 1320
tagcagaaaa cagcttcact cctcaggaag taccatggac agacgcctac caggggtgac 1380
aaagcagtgc agttggattg taaggaaaaa tcccggaatt aatgcacct aatgaatgtt 1440
gtccccctca atggtgttac cttaggagct gaacattcaa ttcagttaca ccactatgac 1500
taaaaacagt ttggatctct tagtattgcc tttgaaactg caacataagc aactcaacaa 1560
tattagtgtc attcctttat agacatacca tgtcaaagac gtttttctat caagttgtat 1620
tctttcctgt tctataacct ttgtcatctg ttagactctg tatgtgtgat ttgtaaaaag 1680
caggctgaaa ctatggacat gatttctgaa gagcacatct ccactgactt tcataaagca 1740
aatgtccaat atttatttat tgagagtttt ttagtgcaat ctgggcccagc atttttatag 1800
attatgatta tgtggttaatt tatccttcct aactctttaa tcctgaatga tgggtggaaa 1860
tggcctagaa ttaggttact ctgttcacaa tgctcattgt tagcatgcaa ttggtatttg 1920

```

acttgggaagt gttgtgttgtt atttttttgaa cccctaggct tcaggaaaac tgctctttttg 1980
taaaaaagaat agcgatgaca ttttctaatag tgcagaaatg ttccaaaagg acaaaattga 2040
aaacaaaaaa ctatgttatt aaaacaaaaa aatgc 2075

<210> 32

<211> 1828

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2715885CB1

<400> 32

cacgtgggag cctgggagcg ggtgggtcgta gctcggtagt ccagttgtgg gtaatcgggg 60
ctgtttgttc ctgtccgaga gagctcggcg gagacggctg tcgagtaccc ttcacctcgg 120
tggtgggagc ctgggagcga actgcggcgc gggttaccgc tcccggggac gcagcaaggg 180
gcatcgagtc cctggcgggg gctgcgccat ggcatgtctc tcgaccgtcc ggggcgcgac 240
ctgggggtcgc ctctgtcacc gtcattttct ccatgcagcg cggcatgggg agcggcctgg 300
tggtggaggag ctaagccgct tgctgtctgga tgacctgtg ccgacctctc ggctggagct 360
tctgtttggc atgaccccg gtctcctggc tctgcaggcc gcccgccgct ctgtggcccg 420
gctcctgtct caggcgggta aagctgggct gcaggggaag cgggcccagc tgctccggat 480
ggccgaggcg cgggacattc cagttctgct gcccagacgg cagaaactgg acacaatgtg 540
ccgctaccag gtccaccagg gtgtctgcat ggaggtgagc ccgctgcggc cccggccttg 600
gagagaggcc ggggagggca gcccagggca cgaccccag cagttgtggc tcgtcctcga 660
tggtgatccg gatccccgga attttggggc tgtgtctgct tccgcacact tcctcggagt 720
ggataaggtc atcaccagcc ggagaaacag ctgcccgtc actccagtag tcagcaagtc 780
cagcgcgggg gctatggagg tgatggacgt gttctccact gatgacctca ccgattttt 840
acagacaaaa gcccagcagg gctggctcgt ggccggcacg gtgggctgcc caagcacaga 900
ggatccccag tcctccgaga tccccatcat gagtgtcttg gagttcctct gggaacggcc 960
tactctcctt gtgtctgggga atgagggctc aggtctatcc caggaggtgc aggcctcctg 1020
ccagcttctc ctaccatcc tgccccggcg ccagctgcct cctggacttg agtccttgaa 1080
cgtctctgtg gctgcaggaa ttcttcttca ctccatttgc agccagagga agggtttccc 1140
cacagagggg gagagaaggc agcttctcca agaccccaa gaaccctcag ccaggtctga 1200
agggtcagc atggctcagc acccagggct gtcttcaggc ccagagaaag agaggcaaaa 1260
tgagggtgta cgtggactgt ccacagtgtt catgtgctgg agtcaggggac ggccgcacct 1320
gcctccgccc gctccagtgt ggggggagcc tctgcctgag tgtgcaccag gcccatgttt 1380
attgaccaca gtctgggggg gggggaaggg gactgcggtg gacaccagag gaagctgttt 1440
cctgttgtga tggtggacct gtagtaggac atggtgattt gttaatttcc atgggaagcc 1500
atgatggcct agcatggagg gaatctgttc ccaggccctg cctggaagtt gagggaaagt 1560
ttagacatct gcagagaggc aggcagccca gcccagggga cccgttctc ttgaaccagt 1620
cattgcctgt ggcaaagtgt tgtatgagaa tgtggggggg ggagggcggg gccctgatgt 1680
ggagtagaca gtgcgcacct caggcccaca cacggccccg ccctgggggc ttgagcgag 1740
gcctcatctt tctgtgccgc gggactccgc acctacctca cagggttgtt gtgaggctca 1800
aataaaacat cactcagcaa aaaaaaaa 1828

<210> 33

<211> 2110

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1742628CB1

<400> 33

```

gtgacctccg acgccccggg caagagaacg ccaggaggga taacgggagg aaggccggcc 60
ggggccgcca aggcagtcgc aggcctgcgc agggaggcgc cagaccttgc accttgcacc 120
ttcgagcgc cctgcacccc gccaccatgt gcgagctgta cagtaagcgg gacactctgg 180
ggctgaggaa gaagcacatc gggccctcat gcaaagtttt ctttgcacg gatcccatca 240
aaatagttag agcccagagg cagtacatgt ttgatgagaa cgggtgaacag tacttgact 300
gcatcaacaa tgttgcccat gtgggacact gtcaccagg agtgggtcaaa gctgccctga 360
aacagatgga actgctaaat acaaattctc gattcctcca cgacaacatt gttgagtatg 420
ccaaacgcct ttcagcaact ctgccggaga aactctctgt ttgttatttt acaaattcag 480
gatccgaagc caacgactta gccttacgcc tggctcggca gttcagaggc caccaggatg 540
tgatcactct tgaccatgct taccatggtc acctatcatc cttaattgag attagcccat 600
ataagtttca gaaaggaaaa gatgtcaaaa aagaatttgt acatgtggca ccaactccag 660
atacttacag aggaaaaatat agagaagacc atgcagactc agccagtgtc tatgcagatg 720
aagtgaagaa aatcattgaa gatgtcata acagtggaa gaagattgtc gcctttattg 780
ctgaatccat gcagagttgt ggcggacaaa taattcctcc agcaggctac ttccagaaag 840
tggcagaata tgtacacggg gcagggggtg tgtttatagc tgatgaagtt caagtgggct 900
ttggcagagt tgggaaacat ttctggagct tccagatgta tgggtgaagac tttgttccag 960
acatcgtcac aatgggaaaa ccgatgggca acggccaccc ggtggcatgt gtggttaaca 1020
ccaaagaaat tgcagaagcc ttcagcagct ctgggatgga atattttaat acgtatggag 1080
gaaatccagt atcttgtgct gttgggtttg ctgtcctgga tataattgaa aatgaagacc 1140
ttcaaggaaa tgccaagaga gtagggaatt atctactga gttactgaaa aacagaagg 1200
ctaaacacac tttgatagga gatattaggg gcattggcct ttttattgga attgatttag 1260
tgaaggacca tctgaaaagg acccctgcc cagctgaagc tcagcacatc atctacaaga 1320
tgaaagaaaa acgagtgtct ctcagtgcc atggacctca tagaaatgta cttaaaataa 1380
aaccacctat gtgcttctact gaagaagatg caaagttcat ggtggacca cttgatagga 1440
ttctaacagt tttagaagaa gctatgggaa ccaaaccga aagtgtgacc tctgagaata 1500
ctccatgcaa aacaaagatg ctgaaagaag ccacataga actgcttagg gacagacca 1560
ctgactccaa agaaaatccc agcagaaaga gaaatggaat gtgcacggat acacattcac 1620
tgctcagtaa gaggtcaag acatgactga tttgcatttt aaagcaagat gcgatgtcca 1680
gagttacaga gaatgagtag atgtgtctca tcggttaata gctctattat acctctaaag 1740
gtggaattgt cagttagat tcataaatga aaaggtaa at gagtaatcag aataaaccaa 1800
gtgataatca aaccatgtca agattattag ttcagactct agcctgttaa ttttcttagt 1860
tgattttctga agctacctga tttattctat taaattgtaa gcttgcaaac tcaaaataaa 1920
ttggcagatt tacctctcat gttttaatgt gtcaaattag agagcaaagt ataacagggtg 1980
ccttcacttt tgagacttag tgccctaaaa tatgtattct ataatgattt catatataaa 2040
agtatattta ttgactgtaa taaaataaaa tatgatgtaa acaaaaaaaaa aaaaaaaaaa 2100
aaaaaaaaaa 2110

```

<210> 34

<211> 2481

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2124971CB1

<400> 34

```

cagggcctgc gatggagcct gcagccccgg gtcgcgtccc tccctgagcg cccccgtcgg 60
cgcccatgct gccccgaggg cgcgcccggg cgctgggggc cgccgcgctg ttgctgctgc 120
tgctgctgct cggattcctc ctgttcgggtg gggacctggg gtgtgagcgc cgcgagcctg 180
gcgggcgagc gggggcccg ggatgcttcc ccggcccgct catgccacgt gtccccccag 240
acgggagggt gcggagagcc gccgccctgc acggagacc gggggccggc cccggggacc 300
acaaccgctc cgactcgggc ccgcagccgc cgcccgcc caagtgcgag ctcttgcatg 360
tggccatcgt gtgtgcgggg cataactcca ccgagacgt catcaccctg gtgaagtcca 420
tgctcttcta caggaaaaat ccactgcacc tccacttggg gactgacgcc gtggccagaa 480
acatcctgga gacgctcttc cacacatgga tgggtgcctgc tgtccgtgtc agcttttatc 540

```

```

atgccgacca gctcaagccc caggtctcct ggatcccca caagcactac tccggcctct 600
atgggctaata gaagctgggtg ctgcccagtg ccttgccctg tgagctggcc cgcgtcattg 660
tcctggacac ggatgtcacc ttgcctctg acatctcgga gctctgggcc ctctttgtct 720
acttttctga cacgcaggcg atcggctctt tggagaacca gactgactgg tacctgggca 780
acctctggaa gaaccacagg ccttggcctg ccttgggccc gggatttaac acaggtgtga 840
tcctgctgcg gctggaccgg ctccggcagg ctggctggga gcagatgtgg aggctgacag 900
ccaggcggga gctccttagc ctgcctgcc cctcactggc tgaccaggac atcttcaacg 960
ctgtgatcaa ggagaccccg gggctagtgc agcgtctgcc ttgtgtctgg aatgtgcagc 1020
tgtcagatca cacactggcc gagcgtgct actctgaggc gtctgacctc aaggtgatcc 1080
actggaactc accaaagaag cttcgggtga agaacaagca tgtggaattc ttccgcaatt 1140
tctacctgac cttcctggag tacgatggga acctgctgcg gagagagctc tttgtgtgcc 1200
ccagccagcc ccacctgggt gctgagcagt tgcagcaggc cctggcacia ctggacgagg 1260
aagacccctg ctttgagttc cggcagcagc agctcactgt gcaccgtgtg catgtcactt 1320
tcctgcccc tgaaccgcca ccccccgcc ctacgatgt cacccttgtg gccagctgt 1380
ccatggaccg gctgcagatg ttggaagccc tgtgcaggca ctggcctggc cccatgagcc 1440
tggccttgta cctgacagac gcagaagctc agcagttcct gcatttcgtc gaggcctcac 1500
cagtgtctgc tgcccgagc gacgtggcct accatgtggt gtaccgtgag gggcccttat 1560
accccgctca ccagcttcgc aacgtggcct tggccaggc cctcacgctc tacgtcttcc 1620
tcagtacat tgacttcctg cctgcctatt ctctctacga ctacctcagg gcctccattg 1680
agcagctggg gctgggcagc cggcgcaagg cagcactggg ggtgccggca tttgagacct 1740
tgcgctaccg cttcagcttc cccattcca aggtggagct gttggccttg ctggatgcgg 1800
gcactctcta caccttcagg taccacgagt ggccccgagg ccacgcaccc acagactatg 1860
cccgtgaggc ggaggctcag gccccgtacc gtgtgcaatg ggccggccaa tatgaaccct 1920
acgtgggtgt gccacgagac tgtccccgct atgacccctg ctttgtgggc ttccgctgga 1980
acaaagtggc ccacattgtg gagctggatg cccaggaata tgagctcctg gtgctgccc 2040
aggccttcac catccatctg cccacgctc caagcctgga catctccgc ttccgctcca 2100
gccccaccta tcgtgactgc ctccaggccc tcaaggacga attccaccag gacttgtccc 2160
gccaccatgg ggctgctgcc ctcaaatacc tcccagccct gcagcagccc cagagccctg 2220
cccagggtg aggtggggcc ggcgctgccc ctcatcttag cattgggcag acaccagggc 2280
aacctgccct ccgccatccc tgetatttaa attatttaag gtctctggga agggctgggg 2340
cagagcatct gtgggtggg gtcttcccct tgctgctatt gtatggctgg ggactggctc 2400
ctctctgccc cagccagttt ggggctggtt ccccatctt gaattgttta tccctttttc 2460
ataattaaag ttttaaaaca t 2481

```

<210> 35

<211> 1933

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2258250CB1

<400> 35

```

ggacaatctc ctttacagtt tcggaagcag gtttgttgcc atggagttca ctttttgacg 60
ggagttgaga agtataaagg taaccatttg ttttagtttc aacgatctga caaaaagata 120
ggctgttgct cttcttctgg aaaagcctga ttggtaagat tcctttaagg gctcagcccc 180
aaagagcttt atcccatccc ctgcagact gaaaactaaa gcctgcagag acctctgaag 240
gaaaacctgt cccgggctct gtcacttcac acccatggct aacctggag gtggtgctgt 300
ttgcaacggg aaacttcaca atcacaagaa acagagcaat ggctcacaaa gcagaaactg 360
cacaagaagt ggaatagtga aggaagccca gcaaaatggg aagccacatt tttatgataa 420
gctcattgtt gaatcgtttg aggaagcacc ccttcattgt atggttttca cttataggg 480
atatggaatt ggaacctgt ttggctatct cagagacttt ttaagaaact ggggaataga 540
aaaatgcaac gcagctgtgg aacgaaaaga acaaaaagat tttgtgccac tgtatcaaga 600
ctttgaaaat ttttatacaa gaaaccttta catgcgaatc agagacaact ggaaccggcc 660
catctgcagt gccccagggc ctctgtttga tgtgatggag agggatatcg acgactataa 720

```

```

ctggacgttt aggttttactg gaagagtcac caaagatgtc atcaacatgg gctcctataa 780
cttccttggc cttgcagcca agtatgatga gtctatgagg acaataaagg atgttttaga 840
ggtgtatggc acaggcgtgg ccagcaccag gcatgaaatg ggcaccttgg ataagcaca 900
ggagttggag gaccttgtgg ctaagttcct gaatgtggaa gcagctatgg tctttgggat 960
gggatttgca actaactcaa tgaatatccc agcattagtt ggaaagggat gcctcatttt 1020
aagtgatgag ttaaaccaca catcgcttgt gcttggggcc cgactctcag gtgcaaccat 1080
aagaatcttc aaacacaaca acacacaaag cctagagaag ctcttgagag atgctgtcat 1140
ctatggccag cctcgaaccc gcagagcttg gaaaaagatt ctcatcctgg tggaggggtgt 1200
ctacagcatg gaaggttcca tcgtgcatct gccccagatc atagctctaa agaagaaata 1260
caaggcttac ctctacatag atgaagctca cagtattggg gccgtgggccc caaccggccc 1320
gggtgtcacg gagttctttg gactagaccc tcatgaagtt gatgtgctca tgggcacatt 1380
caccaaaagt tttggagctt caggagggtta catagctgga aggaaggacc tcgtggatta 1440
tttacgggtt cactcgcata gtgctgttta tgcttcatcc atgagccac cgatagcaga 1500
gcaaatcatc agatcactaa aacttatcat gggactggat gggaccactc aagggtgca 1560
gagagtacag caacttgca aaaacacaag atacttcaga caaagactgc aggaaatggg 1620
attcattatc tatggcaatg agaatgcttc tgtgttcct ctgcttcttt acatgcctgg 1680
taaagtagcg gcttttgcaa ggcataatgt agagaaaaaa attggagtgg tggctgtggg 1740
atttcagacc actccctcgc cagaagctcg ggctcgggtt tgtgtttcag cggcacatac 1800
ccgggagatg tttagacagg ttttagaagc tcttgatgaa atgggtgatc tcttgcaact 1860
gaaatattcc cggcacaaga agtcagcacg tcctgagctc tatgatgaga cgagctttga 1920
actcgaagat taa
1933

```

<210> 36

<211> 2370

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2626035CB1

<400> 36

```

ggagcctgca gccccggggtc ggcgtccctcc ctgagcgccc ccgtcggcgg ccatgctgcc 60
ccgagggcgc ccccgggcgc tgggggcccgc cgcgtgttg ctgctgctgc tgctgctcgg 120
attcctcctg ttcgacggga ggctgctggag agccgccgcc ctgacggag acccgggggc 180
cggccccggg gaccacaacc gctccgactg cggcccgag ccgccgcgc cgcacaagt 240
cgagctcttg catgtggcca tcgtgtgtgc ggggcataac tccagccgag acgtcatcac 300
cctggtgaag tccatgctct tctacaggaa aaatccactg cactccact tggtgactga 360
cgccgtggcc agaaacatcc tggagacgct ctccacaca tggatggtgc ctgctgtccg 420
tgtcagcttt tatcatgccg accagctcaa gcccaggtc tcctggatcc ccaacaagca 480
ctactccggc ctctatgggc taatgaagct ggtgctgccc agtgccctgc ctgctgagct 540
ggcccgctgc attgtcctgg acacggatgt cacttcgcc tctgacatct cggagctctg 600
ggcctctttt gctcactttt ctgacacgca ggcgatcggc cttgtggaga accagagtga 660
ctggtagctg ggcaacctct ggaagaacca caggccctgg cctgccttgg gccggggatt 720
taacacaggt gtgatcctgc tgcggctgga ccggctccgg caggctggct gggagcagat 780
gtggaggctg acagccaggc gggagctcct tagcctgcct gccacctcac tggctgacca 840
ggacatcttc aacgctgtga tcaaggagca cccggggcta gtgcagcgtc tgccttgtgt 900
ctggaatgtg cagctgtcag atcacacact ggccgagcgc tgctactctg aggcgtctga 960
cctcaagggt atccactgga actcacaaa gaagcttcgg gtgaagaaca agcatgtgga 1020
attcttccgc aatttctacc tgaccttcct ggagtacgat gggaacctgc tgcggagaga 1080
gctcttttgt tgccccagcc agccccacc tgggtgctgag cagttgcagc aggcctggc 1140
acaactggac gaggaagacc cctgctttga gttccggcag cagcagctca ctgtgcaccg 1200
tgtgcatgtc actttcctgc cccatgaacc gccaccccc cggcctcacg atgtcacct 1260
tgtggccag ctgtccatgg accggctgca gatgttggaa gccctgtgca ggcactggcc 1320
tggcccatg agcctggcct tgtacctgac agacgcagaa gctcagcagt tcctgcattt 1380
cgtcagggcc tcaccagtgc ttgctgcccg gcaggacgtg gcctaccatg tgggtgtaccg 1440

```

```

tgagggggccc ctataccccg tcaaccagct tcgcaacgtg gccttgggccc aggccctcac 1500
gccttacgtc ttccctcagt acattgactt cctgcctgcc tattctctct acgactacct 1560
cagggcctcc attgagcagc tggggctggg cagccggcgc aaggcagcac tgggtggtgcc 1620
ggcatttgag accctgcgct accgcttcag ctccccccat tccaagggtg agctgttggc 1680
cttgctggat gcggggcactc tctacacctt caggtaccac gagtggcccc gaggccacgc 1740
accacagac tatgcccgct ggcgggaggg tcaggccccg taccgtgtgc aatgggcggc 1800
caactatgaa ccctacgtgg tgggtgccac agactgtccc cgctatgac ctcgctttgt 1860
gggcttcggc tggaaacaaag tggcccacat tgtggagctg gatgccagg aatatgagct 1920
cctggtgctg ccgagggcct tcaccatcca tctgccccac gctccaagcc tggacatctc 1980
ccgcttcgcg tccagcccca cctatcgtga ctgcctccag gccctcaagg acgaattcca 2040
ccaggacttg tcccgccacc atggggctgc tgccctcaaa tacctcccag ccctgcagca 2100
gccccagagc cctgcccag gctgaggctg ggccggcgct gccctcatc ttagcattgg 2160
gcagacacca gggcaacctg ccctccgcca tccctgctat ttaaattatt taaggctctc 2220
gggaagggct ggggcagagc atctgtgggg tggggtcttc cccttgctgc tattgtatgg 2280
ctggggactg gtctctctct gcccagcca gtttggggct gggtcccca tcttgaattg 2340
tttatccctt tttcataatt aaagttttaa 2370

```

<210> 37

<211> 2534

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4831382CB1

<400> 37

```

cgctgccgct gccgcggttc tcctagcagc gcgggcccgt cggaccgcca aggcagccgg 60
cgctggcgat gggaagcggc gtggccgccc acacaggcag tggcaaagtt tcccagacgt 120
acacatctgg acgcgcggct gccggctacc cgtgaccctt ctaggaaggg ttccaggatt 180
tttaatttgg aaaaaaatcc acctggtttc ctttgtcaag gtctctccgg gtggccacgc 240
gcaggagctg caaacttggtg cacggcggct acaccggcag cggaccgggc tttggagaac 300
ctcgggactc aggtgctgag gtgccagcg gctccggacg tgctacgggg tgcgagcgcg 360
ggggagttcg gggcgcacga caaggaaggg ccccggggag ctctatatgg aggaaggagc 420
ccagaatggt gtgcaccagg aagacaaaaa ctttggtgtc cacttgctgt atcctgagcg 480
gcatgactaa catcatctgc ctgctctacg tgggctgggt caccaactac atcgccagcg 540
tgtatgtgcy ggggcaggag cggcgcccg acaagaagct ggaggaagac aaaggggaca 600
ctctgaagat tattgagcgg ctggaccacc tggagaatgt catcaagcag cacattcaag 660
aggctcctgc caagcctgag gaggcagagg ccgagccctt cacagactcc tctctgtttg 720
cacactgggg ccaggagctc agccccgaag gccggcgcggt ggccctgaag caattccagt 780
actacggcta caacgcctac ctacgcgacc gcctgcccct ggaccggccc ctgcctgacc 840
tcagaccagc tgggtgcggt aacctctcat ttcctgacag cctgccagag gtgagcatcg 900
tgttcatctt cgtcaatgaa gcgctttcag tgctgctgcy ctccatccac tcggccatgg 960
aacgcacgcc ccacatctg ctcaaggaga tcattctggt ggatgacaac agcagtaacg 1020
aggaaactgaa ggagaagctg accgaatatg tggacaaggt gaacagccag aagccaggct 1080
tcatcaaagt cgtgcgtcac agcaagcagg aaggcctcat ccgctccagg gtcagtggct 1140
ggagggcggc cactgcccct gtggtggcac tctttgatgc ccacgtggag ttcaatgtgg 1200
gctgggctga acctgtactc acccgcatca aggagaaccg gaagcggatc atctcgccat 1260
cctttgataa catcaaatat gacaactttg agatagaaga gtaccgctg gctgcccagg 1320
gctttgactg ggagctgtgg tgccgctacc taaatcccc caaggcctgg tggaaagctgg 1380
agaactccac agcgccaatc aggagccctg ccctcattgg ctgcttcatt gtggaccggc 1440
agtacttcca ggagatcggc ctgctggacg aaggcatgga agtctacggg ggcgagaatg 1500
tggagcttgg gatcaggggtg tggcagtgtg gcgggagtgat ggaggtcctg cctcgtcac 1560
ggattgcccc cattgagcga gccacacaga cctacacaga ggacctcacc gccatgtcc 1620
gcaggaaacgc tctcaggggtg gctgaagtct ggatggatga atttaaaagc cacgtctaca 1680
tggcatggaa cataccgcag gaggactcag gaattgacat tggggacatc actgcaagga 1740

```

```

aggctctcag gaaacagctg cagtgcaga ccttccggtg gtacctgggtc agcgtgtacc 1800
cagagatgag gatgtactcc gacatcattg cctatggagt gctgcagaat tctctgaaga 1860
ctgattttgtg tcttgaccag gggccagata cagagaatgt ccccatcatg tacatctgcc 1920
atgggatgac gcctcagaac gtgtactaca cgagcagtca gcagatccat gtgggcattc 1980
tgagccccac cgtgggatgat gatgacaacc gatgcctggg ggacgtcaac agccggcccc 2040
ggctcatcga atgcagctac gccaaagcca agaggatgaa gcttcactgg cagttctctc 2100
agggaggacc catccagaac cgcaagtcta agcgtgtctt ggagctgcag gagaatagcg 2160
acctggagtt cggcttccag ctggtgttgc agaagtgtct gggccagcac tggagcatca 2220
ccaacgtcct gaggagcctc gcgtcctgac ccaccggggc cacttccggc tgcctctttg 2280
ctactgtgta gcacctgctg caacgttgcc tgctgtccac gtggggttgt ttggagtctg 2340
gggaaccagg ttagtgggcc cccaagaaga gctttttatt tcctattcaa ttttcatgga 2400
gtttatagaa agatgctgat tggtaggtga tggtagata tcaaactatt ttgcagttgt 2460
aaatagggga cagatggaaa atatttataa ctgacaataa aatattatta agaaaaggga 2520
aaaaaaaaaa aaaa 2534

```

<210> 38

<211> 2599

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2122183CB1

<400> 38

```

cgggcagagc ggccaagatg tgcagccca agaaaagaaa gcttgagtcg gggggcggcg 60
gcgaaggagg ggaggggaact gaagaggaag atggcgcgga gcgggaggcg gccctggagc 120
gaccccgag gactaagcgg gaacgggacc agctgtacta cgagtgtac tcggacgttt 180
cgggtccacga ggagatgac gcggaccgcg tccgcaccga tgccctaccg ctgggtatcc 240
ttcggaactg ggcagcactg cgaggcaaga cgttactgga cgtgggcgcg ggcaccggca 300
ttctgagcat cttctgtgcc caggccgggg cccggcgct gtacgcggta gaggccagcg 360
ccatctggca acaggcccg gaggtggtgc ggttcaacgg gctggaggac cgggtgcacg 420
tcctgccggg accagtggag actgtagagt tgccggaaca ggtggatgcc atcgtgagcg 480
agtggatggg ctacggactc ctgcacgagt ccatgctgag ctccgtctc cacgcgcgaa 540
ccaagtggct gaaggagggc ggtcttctcc tgccggcctc cgccgagctc ttcatagccc 600
ccatcagcga ccagatgctg gaatggcgcc tgggcttctg gagccagggt aagcagcact 660
atggtgtgga catgagctgc ctggagggct tcgccacgcg ctgtctcatg ggccactcgg 720
agatcgttgt gcagggattg tccggcgagg acgtgtggc ccggccgag cgctttgctc 780
agctagagct ctcccgcgcc ggttggagc aggagctgga ggccggagtg ggccggcgct 840
tcgctgcag ctgctatggc tcggcgccca tgcatggctt tgccatctgg ttccagtgta 900
ccttccttg aggggagtc gagaaacccc tgggtgctgtc cactctgcct tttcaccgg 960
ccactcactg gaaacaggcg ctctctacc tgaacgagcc ggtgcaagt gagcaagaca 1020
cggacgttcc aggagagatc acgtgtctgc cctccggga caacccccgt cgctgcgcg 1080
tgctgtcgcg ctacaaagt ggagaccagg aggagaagac caaagacttt gccatggagg 1140
actgagcgtt gccttttctc ccagctacct cccaaagcag cctgacctgc gtgggagagg 1200
cgtagcgagg tcggagggga aaggagatc ccacgtgcaa gtagggggaa tatctcccc 1260
ttttccctca tagcctctag ggaggagag tgacttcatt ctccatttga agagattctt 1320
ctggtgatgt ttacttaaaa agtgatcccc ctcaacaacg gatacagcgt gcttattatt 1380
gggcatttag cctcaaaagc atgtagtacc aagcacttgt atttccgtat attttgttcc 1440
gcgggggag gagggggaag aacacggatg aaaaatgtcag tttttgaagg gtccatgcac 1500
atccctgaca cctcacacct tatctaagtc tgaagctggg gagaaagggg ttcatttaga 1560
cttcatacat ttccagtag acttttagtat ctctccagag ccataatttc tcagtccgaa 1620
ttaattcccc ctccctaggt gcctgtaggc tatggtactt cttcctcatt gttttctagg 1680
taaacttcac tactggtaat taaggggaag gatagtagga agcagtttaa atagccctgt 1740
tctcattact ctgaccacat acatcatagg gtgctaaagt tgatgaacac attaatccgt 1800
taagtaaaat ggactttgta attgtacagc atacctaaga aactcagaag gtgcatttaa 1860

```

```

gagagagacc tgaagaaat agtatggatt tttaaaaatt cttgtctcta ctattataac 1920
caaaaaatat ttcttgatg tcccataaaa atatttgtgt aattcttatg aaacaggctg 1980
gtagaggagg tttctgagcc tagcccaagg gcttattcat caccatgggt aaattattta 2040
aactcactta attaaggaaa atattttccc agctagaaaa gtatactcat tctcatttaa 2100
actctctcat ttggagggat catgtgagtt ggcctactta caagtagtga aagttccctt 2160
ttcagttttg ttttgttttg ttttgttttt ctctttcact cagccaaatg tgaaagtgtg 2220
gaatttagga aaatcacttg taatgaagtg tgaatcttgt tatcaaattt atttctctga 2280
tgtttccttc cttatccttg tagccaataa aacattgaca ttctcacgtt ttatagatga 2340
ggtaaaaagt cttgtgtgct gtgagttata atgcctttgc ctttttaata ttattagttc 2400
ttaagtgtta cagcccttc agaataaac ttcaggacaa ttcaaactat gcttaatgta 2460
tgattttcga gcttctgtat gctaagaaaa taggtgtgaa aaactgggtg tctgaaatag 2520
cctaacattt attgtaattc tgaattttct gcccttttat tcattgcata ttaaagtatt 2580
agagtataaa aactaaaaa

```

<210> 39

<211> 3745

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7484338CB1

<400> 39

```

ggagtaattt ctttagctga catgtaagca tagaggttgg ggtctgtgtg ctaacctgtg 60
ttgtgtttgc agttgtaatt tagattcgag aagtggttta tcctttgact ggaaaagaaa 120
agtagctgca gtattcccc agcacttgct gagagcatgc cgtatgccag gctgtgaggc 180
tcgagagaca agcagtggaa gagttgcggc ctgtttcatc tctggattgt aaactctgagc 240
ctccttctgg ccctggaag gggacagcat caccatggaa tgattcctaa ccagcataat 300
gctggagccg ggagccacca acctgcagtt ttcagaatgg ccgtgttgga cactgatttg 360
gatcacattc ttccatcttc tgttcttctt ccattctggg ctaagttagt agtgggatcg 420
gttgccattg tgtgttttgc acgcagctat gatggagact ttgtctttga tgactcagaa 480
gctattgtta acaataagga cctccaagca gaaacgcccc tgggggacct gtggcatcat 540
gacttctggg gcagtagact gagcagcaac accagccaca agtctaccg gcctctcacc 600
gtcctgactt tcaggattaa ctactacctc tcgggaggct tccaccccg gggctttcac 660
gtggccaaca tcctcctgca cagtggcatc tctgtcctca tgggtggacgt cttctcgggt 720
ctgtttggcg gctgcagta caccagtaaa ggccggaggc tgcacctcgc cccagggcg 780
tcctgctggg ccgcgctgct gtttgcgtgc catcctgtgc acaccgagtg tgttgcgtgt 840
gttgctcgcc gtgcagacct cctgtgtgcc ctgttcttct tgttatcttt ccttggctac 900
tgtaaagcat ttagagaaag taacaaggag ggagcgcatt ctccacctt ctgggtgctg 960
ctgagtatct ttctgggagc agtggccatg ctgtgcaaag agcaaggat cactgtgctg 1020
ggtttaaatg cggtatattga catcttgggtg ataggcaaat tcaatgttct ggaaattgtc 1080
cagaaggtac tacataagga caagtcatta gagaatctcg gcatgctcag gaacgggggc 1140
ctcctcttca gaatgacctt gctcacctct ggaggggctg ggatgctcta cgtgcgctgg 1200
aggatcatgg gcacgggccc gccggccttc accgaggtgg acaaccggc ctcctttgct 1260
gacagcatgc tggtaggggc cgtaaaactac aattactact attcattgaa tgcttggtg 1320
ctgctgtgtc cctggtggct gtgttttgat tgggtcaatgg gctgcatccc cctcattaag 1380
tccatcagcg actggagggt aattgcactt gcagcactct ggttctgcct aattggcctg 1440
atatgccaa g cctgtgtctc tgaagacggc cacaagagaa ggatccttac tctgggcctg 1500
ggatttctcg ttatcccatt tctccccgcg agtaacctgt tcttccaggt gggcttcgtg 1560
gtcgcggagc gtgtcctcta cctccccagc gttgggtact gtgtgtgtgt gacttttgga 1620
ttcggagccc tgagcaaaca taccaagaaa aagaaactca ttgccgctgt cgtgtgggga 1680
atcttattca tcaacacgct gagatgtgtg ctgcgcagcg gcgagtggcg gagtaggaa 1740
cagcttttca gaagtgtct gtctgtgtgt cccctcaatg ctaaggttca ctacaacatt 1800
ggcaaaaacc tggctgataa aggcaaccag acagctgcca tcagatacta ccgggaagct 1860
gtaagattaa atcccaagta tgttcatgcc atgaataatc ttggaaatat cttaaaagaa 1920

```

```

aggaatgagc tacaggaagc tgaggagctg ctgtcttttg ctgttcaaat acagccagac 1980
tttgccgctg cgtggatgaa tctaggcata gtgcagaata gcctgaaacg gtttgaagca 2040
gcagagcaaa gttaccggac agcaattaaa cacagaagga aatacccaga ctgttactac 2100
aacctcgggc gtctgtatgc agatctcaat cgccacgtgg atgccttgaa tgcgtggaga 2160
aatgccaccg tgctgaaacc agagcacagc ctggcctgga acaacatgat tatactcctc 2220
gacaatacag gtaatttagc ccaagctgaa gcagttggaa gagaggcact ggaattaata 2280
cctaattgatc actctctcat gttctcgttg gcaaactgtc tggggaaatc ccagaaatac 2340
aaggaatctg aagctttatt cctcaaggca attaaagcaa atccaaatgc tgcaagttac 2400
catggtaatt tggctgtgct ttatcatcgt tggggacatc tagacttggc caagaaacac 2460
tatgaaatct ccttgcagct tgacccacg gcacaggaa ctaaggagaa ttacggtctg 2520
ctgagaagaa agctagaact aatgcaaaag aaagctgtct gatcctgttt ccttcatgtt 2580
ttgagtttga gtgtgtgtgt gcatgaggca tatcattaat agtatgtggt tacatttaac 2640
catttaaaag tcttagacat gttattttac tgattttttt ctatgaaaac aaagacatgc 2700
aaaaagatta tagcaccagc aatatactct tgaatgcgtg atatgatttt tcatgaaatc 2760
tgtatttttt cagacaactc aaatgtaatt ctaaaattcc aaaaatgtct tttttaatta 2820
aacagaaaaa gagaaaaaat tatcttgagc aacttttagt agaattgagc ttacatttgg 2880
gatctgagcc ttgtcgtgta tggactagca ctattaaact tcaattatga ccaagaaagg 2940
atacactggc ccttacaatt tgtataaata ttgaacatgt ctatatatta gcatttttat 3000
ttaatgacaa agcaaaattaa gtttttttat ctcttttttt taaaacaaca tactgtgaac 3060
tttgtaagga aatattttatt tgtattttta tgttttgaat agggcaaata atcgaaatgag 3120
gaatggaagt ttaacatag tatatctata tgcttttccc cataggaaga aattgactct 3180
tgcagttttt ggatgctctg acttgtgcaa tttcaataga caggagatta tgtaatgtaa 3240
tatttttcat aagcggttac tatcaattga aagttcaagc catgcttttag gcaagagcag 3300
gcagcctcac atcttttatt ttgttacatc caaggtgaag agggcaacac atctgtgtaa 3360
gctgcttttt agtgtgttta tctgaaggcc gttttccatt ttgcttaatg taactacaga 3420
cattatccag aaaatgcaaa attttctatc aaatggagcc acattcgggg aattcgtggt 3480
atttttaaga attgagttgt tctgtctgtt ttttatttga tccaaacaat gttttgtttt 3540
gttcttctct gtatgtctgt gacctaatga tttatgcaat ctctgtaat tcttatgcag 3600
taaaattact acacaaacta gcatgaaaat gtcataattgc cttcttaatc aattattttc 3660
aagtagtgaa ctttgtatcc tcctttacct taaaatgaaa tcaactgac caaatcatca 3720
tttatgtggc ttctgtgtga cttggg 3745

```

<210> 40

<211> 2323

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8326588CB1

<400> 40

```

atggatcagg tggcaacctt ggggcttgag tctgtcgacc tgcagagctc caggaacaac 60
aaggagcacc acacgcagga gatgggcgtc aagcggctca ctgtgcgccg cggccagccc 120
ttctacctcc ggctgagctt cagccgaccc ttccagtcctc agaacgacca catcaccttt 180
gtggctgaga ccggacccaa gccgtcagag ctgctgggga cccgagccac attcttcttc 240
acccgggtcc agcccgggaa tgtctggagc gcttctgatt tcaccattga ctccaactct 300
ctccaagttt ccctttttcac accagccaat gcagttattg gccattacac tctgaaaata 360
gagatctctc agggccaagg tcacagtgtg acttaccgcg tgggaacttt catcctactt 420
tttaacctt ggagtccaga ggacgacgtc tacctgccaa gtgaaatact gctgcaggag 480
tatatcatgc gagattatgg ctttgttttac aagggtcatg aaagattcat cacctcctgg 540
ccctggaact acgggcagtt tgaagaggac atcatagaca tctgctttga gatcctgaac 600
aagagcctgt atcacttaaa gaaccgggcc aaagactgtt cccagcggaa cgacgtggtg 660
tatgtgtgca ggggtggtgag tgccatgac aacagcaacg atgacaatgg cgtgctgcag 720
gggaactggg gcgaggacta ctccaaaggg gtcagtcctc tggagtggaa gggcagtggt 780
gccatcctac agcagtggct agccaggggc gggcagcctg tgaagtacgg acagtgtctg 840

```

```

gtcttcgcct ctgttatgtg caccgtaatg agatgcttag gtgttccaac ccgtgttggt 900
tccaatttcc gttccgcgca caacgtggat aggaacttga ccatcgatac gtactatgac 960
cgaaatgccg agatgctgtc aatcagaaa cgagacaaaa tatggaactt ccacgtcttg 1020
aatgagtgtc ggatgatccg gaaagatctc ccaccaggat acaacgggtg gcaggttctg 1080
gacccactc cccagcagac cagcagtggg ctgttctgct gtggccctgc ctctgtgaag 1140
gccatcaggg aaggggatgt ccacctggcc tatgacaccc cttttgtgta tgccgaggtg 1200
aacgccgatg aagtcatttg gctccttggg gatggccagg ccaggaaat cctggcccac 1260
aacaccagtt ccatcgggaa ggagatcagc actaagatgg tggggtcaga ccagcgccag 1320
agcatcacca gctcctacaa gtaccagaa ggatcccctg aggagagagc tgtcttcctg 1380
aaggcttctc ggaaaatgct gggcccccac agagcttctt tgcccttcct ggatctcctg 1440
gagtctgggg gtcttaggga tcagccagcg cagctgcagc ttcacctggc caggataccc 1500
gagtggggcc aggacctgca gctgctgctg cgtatccaga ggggtccaga cagcaccac 1560
cctcgggggc ccatcggaact ggtggtgcgc ttctgtgcac aggcctgct gcattgggg 1620
ggtaccaga agcccttctg gaggcacaca gtgcggatga acctggactt tgggaaggag 1680
acacagtggc cgctcctcct gccctacagc aattacagaa acaagctaac ggacgaaaag 1740
ctcatccgcy tgtctggcat cgcggagggt gaagagacag ggaggtccat gctggtccta 1800
aaagatatct gtctggagcc tccccacttg tctattgagg tgtctgagag ggctgaggtg 1860
ggcaaggcgc tgagagtcca tgtcaccctc accaacacct taatgggtggc tctgagcagc 1920
tgcacgatgg tgctggaagg aagcggcctc atcaatgggc agatagcaaa ggaccttggg 1980
actctggtgg ccggacacac cctccaaatt caactggacc tctaccgac caaagctgga 2040
ccccgccagc tccaggttct catcagcagc aacgaggtca aggagatcaa aggctacaag 2100
gacatcttgc tcaactgtggc tggggctccc tgagacccgc cctccagctg cctccctgg 2160
caccctgcc ccacctggct cctttctact cctggctatg tcgtcttggc tccacctctg 2220
tcctctctct agcctgctg ggaatgaatg aagctctgtt agaaacaccg tgtgctttgg 2280
gaagagacaa taaagatgtc tttatttatc aaaaaaaaaa aaa 2323

```