

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 03/012037 A2

- (51) International Patent Classification⁷: **C12N**
- (21) International Application Number: PCT/US02/23684
- (22) International Filing Date: 24 July 2002 (24.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/917,372 27 July 2001 (27.07.2001) US
- (71) Applicant (for all designated States except US): **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US).
WARREN, Bridget, A. [US/US]; 1810 S. El Camino Real, B103, Encinitas, CA 92024 (US).
- (74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: TNF RECEPTOR 2 RELATED PROTEIN VARIANT

```

      11      20      29      38      47      56
5'GC CCC GCG GCC AGC TCG CTC CAC TCC CAC TTC CTG AGC TCC GCC ATG GGA GCC CTG

      65      74      83      92     101     110
GAG GCC CGG CCT GGC CGC TCC CGG CCC TGG GGT GCA CAT CGG CCC TGA GTC CCG

      119     128     137     146     155     164
TCC CAG GCT CTG GGC TCG GGC AGC CGC CGC CAC CGC TGC CCA GGA CGT CGG GCC

      173     182     191     200     209     218
TCC TGC CTT CCT CCC AGG CCC CCA CGT TGC TGG CCG CCT GGC CGA GTG GCC GCC

      227     236     245     254     263     272
ATG CTC CTG CCT TGG GCC ACC TCT GCC CCC GGC CTG GCC TGG GGG CCT CTG GTG
  M  L  L  P  W  A  T  S  A  P  G  L  A  W  G  P  L  V

      281     290     299     308     317     326
CTG GGC CTC TTC GGG CTC CTG GCA GCA TCG CAG CCC CAG GCG GTG CCT CCA TAT
  L  G  L  F  G  L  L  A  A  S  Q  P  Q  A  V  P  P  Y

      335     344     353     362     371     380
GCG TCG GAG AAC CAG ACC TGC AGG GAC CAG GAA AAG GAA TAC TAT GAG CCC CAG
  A  S  E  N  Q  T  C  R  D  Q  E  K  E  Y  Y  E  P  Q

```

(57) Abstract: The invention provides a cDNA which encodes a TNFR2PV. It also provides for the use of the cDNA, fragments, complements, and variants thereof and of the encoded protein, portions thereof and antibodies thereto for diagnosis and treatment of cancer and immune disorders, particularly cancer of the prostate, ovary, breast, colon, and brain, or an inflammatory disorder, including rheumatoid arthritis, asthma, and ulcerative colitis. The invention additionally provides expression vectors and host cells for the production of the protein and a transgenic model system.

WO 03/012037 A2



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.

TNF RECEPTOR 2 RELATED PROTEIN VARIANT

5

TECHNICAL FIELD

This invention relates to a cDNA which encodes a TNF receptor 2 related protein variant and to the use of the cDNA and the encoded protein in the diagnosis and treatment of cancers and immune disorders.

10

BACKGROUND OF THE INVENTION

Phylogenetic relationships among organisms have been demonstrated many times, and studies from a diversity of prokaryotic and eukaryotic organisms suggest a more or less gradual evolution of molecules, biochemical and physiological mechanisms, and metabolic pathways. Despite different evolutionary pressures, the proteins of nematode, fly, rat, and man have common chemical and structural features and generally perform the same cellular function. Comparisons of the nucleic acid and protein sequences from organisms where structure and/or function are known accelerate the investigation of human sequences and allow the development of model systems for testing diagnostic and therapeutic agents for human conditions, diseases, and disorders.

15

Cytokines are a diverse class of extracellular signaling molecules that regulate survival, growth, differentiation, and effector function in a variety of cells. Cytokines include families of signaling molecules such as growth factors, colony-stimulating factors, interleukins, lymphokines, monokines, and interferons. Cytokines are produced by cells that are widespread throughout the body and act in a paracrine or autocrine manner to carry out roles which often involve fighting infection and mediating tissue healing.

20

Tumor necrosis factor (TNF) is a pleiotropic cytokine that mediates immune regulation and inflammatory responses. *In vivo*, TNF possesses anti-tumor activity against certain solid tumors, acts as an immunomodulator, induces septic shock and fever, plays a role in multiple sclerosis, participates in autoimmunity, and is involved in response to bacterial, parasitic, and viral infections. (Aggarwal, B.B. and Reddy, S, in Nicola, NA (1994) Guidebook to Cytokines and Their Receptors, Oxford University Press, Oxford, UK, pp 103-104.) The cellular responses triggered by TNF are initiated through its interaction with two distinct cell surface receptors, TNF-R1 and TNF-R2. (Tartaglia, LA and Goeddel, DV (1992) *Immunol. Today* 13:151-153.) Both TNF receptors are part of the TNF receptor (TNFR) superfamily, whose members include the Fas antigen, the p75 subunit of the NGF receptor, Salf19R, CD27, CD30, and CD40. Members of the TNFR superfamily share the TNFR/NGFR family cysteine-rich region signature, which consists of cysteine-rich pseudo-repeats in the extracellular domains. (ExPASy PROSITE document PDOC00561; Bairoch et al. (1997) *Nucl. Acids. Res.* 25:217-221; and Smith et al. (1994) *Cell*

25

30

35

76:959-962.)

Lymphotoxin α (LT α) is also a soluble, secreted cytokine which, like TNF, mediates immune regulation and inflammatory responses. A cell surface form of LT α , is formed by a complex of LT α with LT β , a type II membrane protein of the TNF family. This cell surface LT α -LT β complex is also assumed to take part in immunological reactions by cell-cell contact, but binds to a receptor that is distinct from that for soluble LT α , designated TNF receptor 2 related protein, or LT β -R (Crowe *et al.* (1994) *Science* 264:707-710). The human and mouse forms of LT β -R share 66% identity, an N-terminal, signal peptide sequence, and a single hydrophobic, transmembrane domain in the middle of the protein that separates the extracellular, amino-terminal, from the intracellular, carboxy-terminal, domains of the receptor (Nakamura *et al.* (1995) *Genomics* 30:312-319).

TNF activity is normally regulated by natural inhibitors of the TNF signaling cascade. Two such inhibitors are soluble forms of TNF receptors, p75 and p55, produced by proteolytic cleavage of the extracellular domain of the cell-bound receptor. The soluble forms of the TNF receptors (sTNF-R), also known as TNF-binding proteins (TNF-BP), antagonize TNF cell signaling by competition with the membrane-bound form of the receptor. sTNF-Rs are implicated in control of numerous pathobiological conditions associated with excessive TNF signaling, including septic shock, cancer, AIDS, transplantation allograft rejection, multiple sclerosis, diabetes, and rheumatoid arthritis. (See Tracey, KJ and A Cerami (1994) *Annu Rev Med* 45:491-503 and references therein). High circulating levels of sTNF-Rs are also found in association with various inflammatory conditions and other TNF associated disorders (Van zee *et al.* (1992) *Proc Natl Acad Sci* 89:4845-4849; Glenn *et al.* (2000) *Human Molecular Genetics* 9:1943-1949; Oliveira *et al.* (2001) *Transplantation* 71:1835-1839). An effective treatment for rheumatoid arthritis using an sTNF-R formulation has recently been approved by the FDA (Baumgartner, SW (2000) *Southern Medical Journal* 93:753-759).

The discovery of a cDNA encoding a TNF receptor 2 related protein variant satisfies a need in the art by providing compositions which are useful in the diagnosis and treatment of cancers and immune disorders, particularly cancers of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, including rheumatoid arthritis, asthma, and ulcerative colitis.

SUMMARY OF THE INVENTION

The invention is based on the discovery of a cDNA encoding TNF receptor 2 related protein variant (TNFR2PV) which is useful in the diagnosis and treatment of cancers and immune disorders, particularly cancers of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.

The invention provides an isolated cDNA comprising a nucleic acid sequence encoding a

protein having the amino acid sequence of SEQ ID NO:1. The invention also provides an isolated cDNA or the complement thereof selected from the group consisting of a nucleic acid sequence of SEQ ID NO:2, a fragment of SEQ ID NO:2, or the complement thereof, selected from SEQ ID NOs:3-12, and a variant of SEQ ID NO:2, or the complement thereof, selected from SEQ ID NOs:13-

18. The invention additionally provides a composition, a substrate, and a probe comprising the cDNA, or the complement of the cDNA, encoding TNFR2PV. The invention further provides a vector containing the cDNA, a host cell containing the vector and a method for using the cDNA to make TNFR2PV. The invention still further provides a transgenic cell line or organism comprising the vector containing the cDNA encoding TNFR2PV. In one aspect, the invention provides a substrate containing at least one of these fragments or variants or the complements thereof. In a second aspect, the invention provides a probe comprising a cDNA or the complement thereof which can be used in methods of detection, screening, and purification. In a further aspect, the probe is a single-stranded complementary RNA or DNA molecule.

The invention provides a method for using a cDNA to detect the differential expression of a nucleic acid in a sample comprising hybridizing a probe to the nucleic acids, thereby forming hybridization complexes and comparing hybridization complex formation with a standard, wherein the comparison indicates the differential expression of the cDNA in the sample. In one aspect, the method of detection further comprises amplifying the nucleic acids of the sample prior to hybridization. In another aspect, the method showing differential expression of the cDNA is used to diagnose a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis. In another aspect, the cDNA or a fragment or a variant or the complements thereof may comprise an element on an array.

The invention additionally provides a method for using a cDNA or a fragment or a variant or the complements thereof to screen a library or plurality of molecules or compounds to identify at least one ligand which specifically binds the cDNA, the method comprising combining the cDNA with the molecules or compounds under conditions allowing specific binding, and detecting specific binding to the cDNA, thereby identifying a ligand which specifically binds the cDNA. In one aspect, the molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides, transcription factors, repressors, and regulatory molecules.

The invention provides a purified protein or a portion thereof selected from the group consisting of an amino acid sequence of SEQ ID NO:1, a variant having at least 95% identity to the amino acid sequence of SEQ ID NO:1, an antigenic epitope of SEQ ID NO:1, and a biologically active portion of SEQ ID NO:1. The invention also provides a composition comprising the purified

protein and a pharmaceutical carrier. The invention further provides a method of using the TNFR2PV to treat a subject with a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis comprising administering to a patient in need of such treatment the composition containing the purified protein. The invention still further provides a method for using a protein to screen a library or a plurality of molecules or compounds to identify at least one ligand, the method comprising combining the protein with the molecules or compounds under conditions to allow specific binding and detecting specific binding, thereby identifying a ligand which specifically binds the protein. In one aspect, the molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs. In another aspect, the ligand is used to treat a subject with a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.

The invention provides a method of using a protein to screen a subject sample for antibodies which specifically bind the protein comprising isolating antibodies from the subject sample, contacting the isolated antibodies with the protein under conditions that allow specific binding, dissociating the antibody from the bound-protein, and comparing the quantity of antibody with known standards, wherein the presence or quantity of antibody is diagnostic of a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.

The invention also provides a method of using a protein to prepare and purify antibodies comprising immunizing a animal with the protein under conditions to elicit an antibody response, isolating animal antibodies, attaching the protein to a substrate, contacting the substrate with isolated antibodies under conditions to allow specific binding to the protein, dissociating the antibodies from the protein, thereby obtaining purified antibodies.

The invention provides a purified antibody which binds specifically to a protein which is expressed in a cancer or immune disorder. The invention also provides a method of using an antibody to diagnose a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis comprising combining the antibody comparing the quantity of bound antibody to known standards, thereby establishing the presence of the disease.

The invention provides a method for inserting a heterologous marker gene into the genomic DNA of a mammal to disrupt the expression of the endogenous polynucleotide. The invention also provides a method for using a cDNA to produce a mammalian model system, the method comprising constructing a vector containing the cDNA selected from SEQ ID NOs:2-18,

transforming the vector into an embryonic stem cell, selecting a transformed embryonic stem cell, microinjecting the transformed embryonic stem cell into a mammalian blastocyst, thereby forming a chimeric blastocyst, transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric offspring containing the cDNA in its germ line, and breeding the
5 chimeric mammal to produce a homozygous, mammalian model system.

BRIEF DESCRIPTION OF THE FIGURES AND TABLE

Figures 1A, 1B, 1C, 1D, 1E, and 1F show the TNF receptor 2 related protein variant (TNFR2PV; SEQ ID NO:1) encoded by the cDNA (SEQ ID NO:2). The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA).

10 Figures 2A, 2B, and 2C demonstrate the conserved chemical and structural similarities among the sequences/motifs/domains of TNFR2PV (7497867CD1; SEQ ID NO:1), human TNF receptor 2 related protein/LTR β (g339762; SEQ ID NO:19), and mouse lymphotoxin-beta receptor (g600223; SEQ ID NO:20). The alignment was produced using the MEGALIGN program of LASERGENE software (DNASTAR, Madison WI).

15 Figures 3A and 3B show the hydrophobicity plots for human TNF receptor 2 related protein/LTR β (Figure 3A) versus TNFR2PV (Figure 3B). The positive Y axis represents hydrophilicity, and the negative Y axis, hydrophobicity. The X axis represents amino acid residue number. The plot was produced using MACDNASIS PRO software.

DESCRIPTION OF THE INVENTION

20 It is understood that this invention is not limited to the particular machines, materials and methods described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments and is not intended to limit the scope of the present invention which will be limited only by the appended claims. As used herein, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. For example, a
25 reference to "a host cell" includes a plurality of such host cells known to those skilled in the art.

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. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which
30 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

"Antibody" refers to intact immunoglobulin molecule, a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a recombinant antibody, a humanized antibody, single chain
35 antibodies, a Fab fragment, an F(ab')₂ fragment, an Fv fragment, and an antibody-peptide fusion

protein.

"Antigenic determinant" refers to an antigenic or immunogenic epitope, structural feature, or region of an oligopeptide, peptide, or protein which is capable of inducing formation of an antibody that specifically binds the protein. Biological activity is not a prerequisite for

5 immunogenicity.

"Array" refers to an ordered arrangement of at least two cDNAs, proteins, or antibodies on a substrate. At least one of the cDNAs, proteins, or antibodies represents a control or standard, and the other cDNA, protein, or antibody is of diagnostic or therapeutic interest. The arrangement of at least two and up to about 40,000 cDNAs, proteins, or antibodies on the substrate assures that the
10 size and signal intensity of each labeled complex, formed between each cDNA and at least one nucleic acid, each protein and at least one ligand or antibody, or each antibody and at least one protein to which the antibody specifically binds, is individually distinguishable.

A "bispecific molecule" has two different binding specificities and can be bound to two different molecules or two different sites on a molecule concurrently. Similarly, a "multispecific
15 molecule" can bind to multiple (more than two) distinct targets, one of which is a molecule on the surface of an immune cell. Antibodies can perform as or be a part of bispecific or multispecific molecules.

"TNFR2PV" refers to a TNF receptor 2 related protein that is exactly or highly homologous (>85%) to the amino acid sequence of SEQ ID NO:1 obtained from any species including bovine,
20 ovine, porcine, murine, equine, and preferably the human species, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The "complement" of a cDNA of the Sequence Listing refers to a nucleic acid molecule which is completely complementary over its full length and which will hybridize to a nucleic acid molecule under conditions of high stringency.

25 "cDNA" refers to an isolated polynucleotide, nucleic acid molecule, or any fragment thereof that contains from about 400 to about 12,000 nucleotides. It may have originated recombinantly or synthetically, may be double-stranded or single-stranded, may represent coding and noncoding 3' or 5' sequence, and generally lacks introns.

The phrase "cDNA encoding a protein" refers to a nucleic acid whose sequence closely
30 aligns with sequences that encode conserved regions, motifs or domains identified by employing analyses well known in the art. These analyses include BLAST (Basic Local Alignment Search Tool; Altschul (1993) J Mol Evol 36:290-300; Altschul et al. (1990) J Mol Biol 215:403-410) and BLAST2 (Altschul et al. (1997) Nucleic Acids Res 25:3389-3402) which provide identity within the conserved region. Brenner et al. (1998; Proc Natl Acad Sci 95:6073-6078) who analyzed
35 BLAST for its ability to identify structural homologs by sequence identity found 30% identity is a

reliable threshold for sequence alignments of at least 150 residues and 40% is a reasonable threshold for alignments of at least 70 residues (Brenner, page 6076, column 2).

A "composition" refers to the polynucleotide and a labeling moiety; a purified protein and a pharmaceutical carrier or a heterologous, labeling or purification moiety; an antibody and a labeling moiety or pharmaceutical agent; and the like.

"Derivative" refers to a cDNA or a protein that has been subjected to a chemical modification. Derivatization of a cDNA can involve substitution of a nontraditional base such as queosine or of an analog such as hypoxanthine. These substitutions are well known in the art. Derivatization of a cDNA or a protein can also involve the replacement of a hydrogen by an acetyl, acyl, alkyl, amino, formyl, or morpholino group (for example, 5-methylcytosine). Derivative molecules retain the biological activities of the naturally occurring molecules but may confer longer lifespan or enhanced activity.

"Differential expression" refers to an increased or upregulated or a decreased or downregulated expression as detected by absence, presence, or at least two-fold change in the amount of transcribed messenger RNA or translated protein in a sample.

"Disorder" refers to conditions, diseases or syndromes in which the cDNAs and TNF receptor 2 related protein variant are differentially expressed. Such a disorder includes cancer and immune disorders, particularly cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.

An "expression profile" is a representation of gene expression in a sample. A nucleic acid expression profile is produced using sequencing, hybridization, or amplification technologies and mRNAs or cDNAs from a sample. A protein expression profile, although time delayed, mirrors the nucleic acid expression profile and uses antibody arrays, enzyme-linked immunosorbent assays, fluorescence-activated cell sorting, 2D-PAGE and scintillation counting, protein arrays, radioimmunoassays, and western analysis and labeling moieties to detect protein expression in a sample. The nucleic acids, proteins, or antibodies may be used in solution or attached to a substrate, and their detection is based on methods and labeling moieties well known in the art.

"Fragment" refers to a chain of consecutive nucleotides from about 50 to about 5000 base pairs in length. Fragments may be used in PCR or hybridization technologies to identify related nucleic acid molecules and in binding assays to screen for a ligand. Such ligands are useful as therapeutics to regulate replication, transcription or translation.

"Guilt-by-association" (GBA) is a method for identifying cDNAs or proteins that are associated with a specific disease, regulatory pathway, subcellular compartment, cell type, tissue type, or species by their highly significant co-expression with known markers or therapeutics.

A "hybridization complex" is formed between a cDNA and a nucleic acid of a sample when

the purines of one molecule hydrogen bond with the pyrimidines of the complementary molecule, e.g., 5'-A-G-T-C-3' base pairs with 3'-T-C-A-G-5'. Hybridization conditions, degree of complementarity and the use of nucleotide analogs affect the efficiency and stringency of hybridization reactions.

5 "Identity" as applied to sequences, refers to the quantification (usually percentage) of nucleotide or residue matches between at least two sequences aligned using a standardized algorithm such as Smith-Waterman alignment (Smith and Waterman (1981) J Mol Biol 147:195-197), CLUSTALW (Thompson *et al.* (1994) Nucleic Acids Res 22:4673-4680), or BLAST2 (Altschul (1997, *supra*). BLAST2 may be used in a standardized and reproducible way to insert
10 gaps in one of the sequences in order to optimize alignment and to achieve a more meaningful comparison between them. "Similarity" uses the same algorithms but takes conservative substitution of residues into account. In proteins, similarity exceeds identity in that substitution of a valine for a leucine or isoleucine, is counted in calculating the reported percentage. Substitutions which are considered to be conservative are well known in the art.

15 "Isolated or "purified" refers to any molecule or compound that is separated from its natural environment and is from about 60% free to about 90% free from other components with which it is naturally associated.

"Labeling moiety" refers to any reporter molecule including radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, substrates, cofactors, inhibitors, or magnetic
20 particles than can be attached to or incorporated into a polynucleotide, protein, or antibody. Visible labels and dyes include but are not limited to anthocyanins, β glucuronidase, biotin, BIODIPY, Coomassie blue, Cy3 and Cy5, 4,6-diamidino-2-phenylindole (DAPI), digoxigenin, fluorescein, FITC, gold, green fluorescent protein, lissamine, luciferase, phycoerythrin, rhodamine, spyro red, silver, streptavidin, and the like. Radioactive markers include radioactive forms of hydrogen,
25 iodine, phosphorous, sulfur, and the like.

"Ligand" refers to any agent, molecule, or compound which will bind specifically to a polynucleotide or to an epitope of a protein. Such ligands stabilize or modulate the activity of polynucleotides or proteins and may be composed of inorganic and/or organic substances including minerals, cofactors, nucleic acids, proteins, carbohydrates, fats, and lipids.

30 "Oligonucleotide" refers a single-stranded molecule from about 18 to about 60 nucleotides in length which may be used in hybridization or amplification technologies or in regulation of replication, transcription or translation. Equivalent terms are amplicon, amplimer, primer, and oligomer.

A "pharmaceutical agent" may be an antibody, an antisense molecule, a bispecific
35 molecule, a multispecific molecule, a peptide, a protein, a radionuclide, a small drug molecule, a

cytospecific or cytotoxic drug such as abrin, actinomyosin D, cisplatin, croton, doxorubicin, 5-fluorouracil, methotrexate, ricin, vincristine, vinblastine, or any combination of these elements.

"Post-translational modification" of a protein can involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and the like. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cellular location, cell type, pH, enzymatic milieu, and the like.

"Probe" refers to a cDNA that hybridizes to at least one nucleic acid in a sample. Where targets are single-stranded, probes are complementary single strands. Probes can be labeled with reporter molecules for use in hybridization reactions including Southern, northern, in situ, dot blot, array, and like technologies or in screening assays.

"Protein" refers to a polypeptide or any portion thereof. A "portion" of a protein refers to that length of amino acid sequence which would retain at least one biological activity, a domain identified by PFAM or PRINTS analysis or an antigenic determinant of the protein identified using Kyte-Doolittle algorithms of the PROTEAN program (DNASTAR, Madison WI). An "oligopeptide" is an amino acid sequence from about five residues to about 15 residues that is used as part of a fusion protein to produce an antibody.

"Sample" is used in its broadest sense as containing nucleic acids, proteins, and antibodies. A sample may comprise a bodily fluid such as ascites, blood, cerebrospinal fluid, lymph, semen, sputum, urine and the like; the soluble fraction of a cell preparation, or an aliquot of media in which cells were grown; a chromosome, an organelle, or membrane isolated or extracted from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; a cell; a tissue, a tissue biopsy, or a tissue print; buccal cells, skin, hair, a hair follicle; and the like.

"Specific binding" refers to a precise interaction between two molecules which is dependent upon their structure, particularly their molecular side groups. For example, the intercalation of a regulatory protein into the major groove of a DNA molecule or the binding between an epitope of a protein and an agonist, antagonist, or antibody.

"Substrate" refers to any rigid or semi-rigid support to which polynucleotides, proteins, or antibodies are bound and includes magnetic or nonmagnetic beads, capillaries or other tubing, chips, fibers, filters, gels, membranes, plates, polymers, slides, wafers, and microparticles with a variety of surface forms including channels, columns, pins, pores, trenches, and wells.

A "transcript image" (TI) is a profile of gene transcription activity in a particular tissue at a particular time. TI provides assessment of the relative abundance of expressed polynucleotides in the cDNA libraries of an EST database as described in USPN 5,840,484, incorporated herein by reference.

"Variant" refers to molecules that are recognized variations of a protein or the

polynucleotides that encode it. Splice variants may be determined by BLAST score, wherein the score is at least 100, and most preferably at least 400. Allelic variants have a high percent identity to the cDNAs and may differ by about three bases per hundred bases. "Single nucleotide polymorphism" (SNP) refers to a change in a single base as a result of a substitution, insertion or deletion. The change may be conservative (purine for purine) or non-conservative (purine to pyrimidine) and may or may not result in a change in an encoded amino acid or its secondary, tertiary, or quaternary structure.

THE INVENTION

The invention is based on the discovery of a cDNA which encodes TNFR2PV and on the use of the cDNA, or fragments thereof, and protein, or portions thereof, directly or as compositions in the characterization, diagnosis, and treatment of cancer and immune disorders.

The table below shows Incyte cDNA clones used to assemble the full-length cDNA sequence of SEQ ID NO:2. The first column shows the SEQ ID NO from the Sequence Listing, the second column the Incyte cDNA clone number, the third column the length of the cDNA, the fourth column the alignment of the cDNA with the full-length sequence of SEQ ID NO:2, and the fifth column the cDNA library of the clone.

	SEQ ID	cDNA	Length	Nt _H Alignment	Library
	3	8113313H1	392	1-392	OSTEUNC01
	4	8235763H1	526	33-558	OSTEUNC01
20	5	4048821H1	301	494-794	SINTNOT18
	6	2105134H1	135	538-672	BRAITUT03
	7	7716364H1	651	629-1277	SINTFEE02
	8	8234468H1	574	804-1378	OSTEUNC01
	9	7716340H1	425	853-1277	SINTFEE02
25	10	697459H1	219	1179-1387	SYNORAT03
	11	3321983H1	279	1207-1485	PTHYNOT03
	12	8576918T1	862	1176-1973	PENIFEC01

It is particularly noteworthy that cDNA clones 7716364H1, 8234468H1, and 7716340H1 (SEQ ID NOs:7-9, respectively) all substantially overlap the the boundary representing the deleted transmembrane domain, which occurs at approximately nucleotides 887-888 of SEQ ID NO:2 (see Figure 1C).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1 as shown in Figures 1A, 1B, 1C, 1D, and 1E. TNFR2PV is 399 amino acids in length and has two potential N-glycosylation sites at N40 and N177; seven potential casein kinase II phosphorylation sites at T42, T82, S87, T140, T152, T189, and S287; seven potential protein kinase C phosphorylation sites at T42, S69, T117, S118, S182, T209, and S387; and one potential tyrosine kinase phosphorylation site at Y50. A signal peptide sequence is predicted from M1 to P30, and two potential TNF/NGFR family cysteine signatures are predicted between C43 and C80 and from C83 to C124. Pfam analysis predicts the two TNF/NGFR cysteine rich domains

between C43 to C80 and C83 to C124. BLIMPS analysis also indicates the presence of two TNF/NGFR family cysteine signatures between C58 to V68 and C116 to V126. BLOCKS DOMO analysis shows a lymphotoxin-beta receptor chain signature between G223 and G383, and two TNF/NGFR cysteine rich domains from K119 to T203, and from E39 to S118. As shown in Figures 2A, 2B and 2C, TNFR2PV has chemical and structural similarity with human TNF receptor 2 related protein/LT β -R (g339762; SEQ ID NO:19) and mouse lymphotoxin-beta receptor (g600223; SEQ ID NO:20). In particular, TNFR2PV and the human LT β -R are identical with the exception of a deleted region in TNFR2PV that corresponds to the transmembrane domain identified in both the human LT β -R and mouse Lt β -R, indicating that TNFR2PV is most likely a soluble form of the human LT β -R. Hydrophobicity plots shown in Figures 3A and 3B confirm the absence of the hydrophobic transmembrane domain from TNFR2PV identified in human LT β -R (see arrow, Figure 3A).

Northern analysis shows overexpression (transcript abundance >1) of TNFR2PV in various cancerous tissue libraries including prostate (PROSTUTO2), ovary (OVRTUT05, OVRTUT01), gallbladder (GBLATUT01), breast (BRSTTUT030), brain (BRAITUT02), liver (LIVRTUT04), and colon (COLNTUT03), and in immune disorders, including rheumatoid arthritis (SYNORAT03), asthma (LUNGAST01), ulcerative colitis (COLNUCT03), and inflamed adenoid (ADENINB01). A useful antigenic epitope of TNFR2PV extends from about amino acid residue P216 to about P235, and which encompasses the deleted transmembrane region in TNFR2PV. An antibody which specifically binds TNFR2PV is useful in a diagnostic assay to identify cancers, such as a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.

Mammalian variants of the cDNA encoding TNFR2PV were identified using BLAST2 with default parameters and the ZOOSEQ databases (Incyte Genomics, Inc., Palo Alto CA). These preferred variants have from about 85% to about 92% identity as shown in the table below. The first column shows the SEQ IDvar for variant cDNAs; the second column, the clone number for the variant cDNAs; the third column, the percent identity to the human cDNA; the fourth column, the alignment of the variant cDNA to the human cDNA; and the fifth column, the species of the variant cDNA.

SEQ ID _{var}	cDNA _{var}	Identity	Nt _H Alignment	Species
13	700302531H1	85%	1126-1285	Rat
14	702152066H1	85%	474-609	Rat
15	702022948H1	88%	715-823	Rat
16	702245091H1	88%	1126-1386	Dog
17	703193780J1	87%	1411-1579	Monkey
18	703678967J1	92%	1499-1579	Monkey

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of cDNAs encoding TNFR2PV, some bearing minimal similarity to the cDNAs of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of cDNA 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 encoding naturally occurring TNFR2PV, and all such variations are to be considered as being specifically disclosed.

The cDNAs of SEQ ID NOs:2-18 may be used in hybridization, amplification, and screening technologies to identify and distinguish among SEQ ID NO:2 and related molecules in a sample. The mammalian cDNAs, SEQ ID NOs:2-18, may be used to produce transgenic cell lines or organisms which are model systems for human cancers and immune disorders and upon which the toxicity and efficacy of potential therapeutic treatments may be tested. Toxicology studies, clinical trials, and subject/patient treatment profiles may be performed and monitored using the cDNAs, proteins, antibodies and molecules and compounds identified using the cDNAs and proteins of the present invention.

Characterization and Use of the Invention

cDNA libraries

In a particular embodiment disclosed herein, mRNA is isolated from mammalian cells and tissues using methods which are well known to those skilled in the art and used to prepare the cDNA libraries. The Incyte cDNAs were isolated from mammalian cDNA libraries prepared as described in the EXAMPLES. The consensus sequences are chemically and/or electronically assembled from fragments including Incyte cDNAs and extension and/or shotgun sequences using computer programs such as PHRAP (P Green, University of Washington, Seattle WA), and AUTOASSEMBLER application (Applied Biosystems, Foster City CA). After verification of the 5' and 3' sequence, at least one representative cDNA which encodes TNFR2PV is designated a reagent.

Sequencing

Methods for sequencing nucleic acids are well known in the art and may be used to practice any of the embodiments of the invention. These methods employ enzymes such as the Klenow fragment of DNA polymerase I, SEQUENASE, Taq DNA polymerase and thermostable T7 DNA polymerase (Amersham Pharmacia Biotech (APB), Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 system (Hamilton, Reno NV) and the DNA ENGINE thermal cycler (MJ Research, Watertown MA). Machines commonly used for sequencing include

the ABI PRISM 3700, 377 or 373 DNA sequencing systems (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (APB), and the like. The sequences may be analyzed using a variety of algorithms well known in the art and described in Ausubel *et al.* (1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7) and in Meyers (1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

Shotgun sequencing may also be used to complete the sequence of a particular cloned insert of interest. Shotgun strategy involves randomly breaking the original insert into segments of various sizes and cloning these fragments into vectors. The fragments are sequenced and reassembled using overlapping ends until the entire sequence of the original insert is known.

Shotgun sequencing methods are well known in the art and use thermostable DNA polymerases, heat-labile DNA polymerases, and primers chosen from representative regions flanking the cDNAs of interest. Incomplete assembled sequences are inspected for identity using various algorithms or programs such as CONSED (Gordon (1998) *Genome Res* 8:195-202) which are well known in the art. Contaminating sequences, including vector or chimeric sequences, or deleted sequences can be removed or restored, respectively, organizing the incomplete assembled sequences into finished sequences.

Extension of a Nucleic Acid Sequence

The sequences of the invention may be extended using various PCR-based methods known in the art. For example, the XL-PCR kit (Applied Biosystems), nested primers, and commercially available cDNA or genomic DNA libraries may be used to extend the nucleic acid sequence. For all PCR-based methods, primers may be designed using commercially available primer analysis software to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to a target molecule at temperatures from about 55C to about 68C. When extending a sequence to recover regulatory elements, it is preferable to use genomic, rather than cDNA libraries.

Hybridization

The cDNA and fragments thereof can be used in hybridization technologies for various purposes. A probe may be designed or derived from unique regions such as the 5' regulatory region or from a nonconserved region (i.e., 5' or 3' of the nucleotides encoding the conserved catalytic domain of the protein) and used in protocols to identify naturally occurring molecules encoding the TNFR2PV, allelic variants, or related molecules. The probe may be DNA or RNA, may be single-stranded, and should have at least 50% sequence identity to any of the nucleic acid sequences, SEQ ID NOs:2-18. Hybridization probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of a reporter molecule. A vector containing the cDNA or a fragment thereof may be used to produce an mRNA probe *in vitro* by addition of an RNA polymerase and labeled nucleotides. These procedures may be conducted using commercially available kits such as those provided by APB.

The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. Hybridization can be performed at low stringency with buffers, such as 5xSSC with 1% sodium dodecyl sulfate (SDS) at 60C, which permits the formation of a hybridization complex between nucleic acid sequences that contain some mismatches. Subsequent washes are performed at higher stringency with buffers such as 0.2xSSC with 0.1% SDS at either 45C (medium stringency) or 68C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acids are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of detergents such as Sarkosyl or TRITON X-100 (Sigma-Aldrich, St. Louis MO) and a blocking agent such as denatured salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (*supra*) and Sambrook *et al.* (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.

Arrays may be prepared and analyzed using methods well known in the art. Oligonucleotides or cDNAs may be used as hybridization probes or targets to monitor the expression level of large numbers of genes simultaneously or to identify genetic variants, mutations, and single nucleotide polymorphisms. Arrays may be used to determine gene function; to understand the genetic basis of a condition, disease, or disorder; to diagnose a condition, disease, or disorder; and to develop and monitor the activities of therapeutic agents. (See, e.g., Brennan *et al.* (1995) USPN 5,474,796; Schena *et al.* (1996) Proc Natl Acad Sci 93:10614-10619; Heller *et al.* (1997) Proc Natl Acad Sci 94:2150-2155; and Heller *et al.* (1997) USPN 5,605,662.)

Hybridization probes are also useful in mapping the naturally occurring genomic sequence. The probes may be hybridized to a particular chromosome, a specific region of a chromosome, or an artificial chromosome construction. Such constructions include human artificial chromosomes (HAC), yeast artificial chromosomes (YAC), bacterial artificial chromosomes (BAC), bacterial P1 constructions, or the cDNAs of libraries made from single chromosomes.

Expression

Any one of a multitude of cDNAs encoding TNFR2PV may be cloned into a vector and used to express the protein, or portions thereof, in host cells. The nucleic acid sequence can be engineered by such methods as DNA shuffling (USPN 5,830,721) and site-directed mutagenesis to create new restriction sites, alter glycosylation patterns, change codon preference to increase expression in a particular host, produce splice variants, extend half-life, and the like. The expression vector may contain transcriptional and translational control elements (promoters, enhancers, specific initiation signals, and polyadenylated 3' sequence) from various sources which

have been selected for their efficiency in a particular host. The vector, cDNA, and regulatory elements are combined using in vitro recombinant DNA techniques, synthetic techniques, and/or in vivo genetic recombination techniques well known in the art and described in Sambrook (supra, ch. 4, 8, 16 and 17).

5 A variety of host systems may be transformed with an expression vector. These include, but are not limited to, bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems transformed with baculovirus expression vectors; plant cell systems transformed with expression vectors containing viral and/or bacterial elements, or animal cell systems (Ausubel supra, unit 16).

10 For example, an adenovirus transcription/translation complex may be utilized in mammalian cells. After sequences are ligated into the E1 or E3 region of the viral genome, the infective virus is used to transform and express the protein in host cells. The Rous sarcoma virus enhancer or SV40 or EBV-based vectors may also be used for high-level protein expression.

Routine cloning, subcloning, and propagation of nucleic acid sequences can be achieved using the multifunctional PBLUESCRIPT vector (Stratagene, La Jolla CA) or PSPOINT1 plasmid (Life Technologies). Introduction of a nucleic acid sequence into the multiple cloning site of these vectors disrupts the lacZ gene and allows colorimetric screening for transformed bacteria. 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.

20 For long term production of recombinant proteins, the vector can be stably transformed into cell lines along with a selectable or visible marker gene on the same or on a separate vector. After transformation, cells are allowed to grow for about 1 to 2 days in enriched media and then are transferred to selective media. Selectable markers, antimetabolite, antibiotic, or herbicide resistance genes, confer resistance to the relevant selective agent and allow growth and recovery of cells which successfully express the introduced sequences. Resistant clones identified either by survival on selective media or by the expression of visible markers may be propagated using culture techniques. Visible markers are also used to estimate the amount of protein expressed by the introduced genes. Verification that the host cell contains the desired cDNA is based on DNA-DNA or DNA-RNA hybridizations or PCR amplification techniques.

30 The host cell may be chosen for its ability to modify a recombinant protein in a desired fashion. Such modifications include acetylation, carboxylation, glycosylation, phosphorylation, lipidation, acylation and the like. Post-translational processing which cleaves a "prepro" form may also be used to specify protein targeting, folding, and/or activity. Different host cells available from the ATCC (Manassas VA) which have specific cellular machinery and characteristic mechanisms for post-translational activities may be chosen to ensure the correct modification and processing of the recombinant protein.

Recovery of Proteins from Cell Culture

Heterologous moieties engineered into a vector for ease of purification include glutathione S-transferase (GST), 6xHis, FLAG, MYC, and the like. GST and 6-His are purified using commercially available affinity matrices such as immobilized glutathione and metal-chelate resins, respectively. FLAG and MYC are purified using commercially available monoclonal and polyclonal antibodies. For ease of separation following purification, a sequence encoding a proteolytic cleavage site may be part of the vector located between the protein and the heterologous moiety. Methods for recombinant protein expression and purification are discussed in Ausubel (*supra*, unit 16) and are commercially available.

Chemical Synthesis of Peptides

Proteins or portions thereof may be produced not only by recombinant methods, but also by using chemical methods well known in the art. Solid phase peptide synthesis may be carried out in a batchwise or continuous flow process which sequentially adds α -amino- and side chain-protected amino acid residues to an insoluble polymeric support via a linker group. A linker group such as methylamine-derivatized polyethylene glycol is attached to poly(styrene-co-divinylbenzene) to form the support resin. The amino acid residues are N- α -protected by acid labile Boc (t-butyloxycarbonyl) or base-labile Fmoc (9-fluorenylmethoxycarbonyl). The carboxyl group of the protected amino acid is coupled to the amine of the linker group to anchor the residue to the solid phase support resin. Trifluoroacetic acid or piperidine are used to remove the protecting group in the case of Boc or Fmoc, respectively. Each additional amino acid is added to the anchored residue using a coupling agent or pre-activated amino acid derivative, and the resin is washed. The full length peptide is synthesized by sequential deprotection, coupling of derivitized amino acids, and washing with dichloromethane and/or N, N-dimethylformamide. The peptide is cleaved between the peptide carboxy terminus and the linker group to yield a peptide acid or amide. (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego CA pp. S1-S20). Automated synthesis may also be carried out on machines such as the ABI 431A peptide synthesizer (Applied Biosystems). A protein or portion thereof may be purified by preparative high performance liquid chromatography and its composition confirmed by amino acid analysis or by sequencing (Creighton (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY).

Preparation and Screening of Antibodies

Various hosts including, but not limited to, goats, rabbits, rats, mice, and human cell lines may be immunized by injection with TNFR2PV or any portion thereof. Adjuvants such as Freund's, mineral gels, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemacyanin (KLH), and dinitrophenol may be used to increase immunological response. The oligopeptide, peptide, or portion of protein used to induce antibodies should consist of at least about five amino acids, more preferably ten amino acids, which

are identical to a portion of the natural protein. Oligopeptides may be fused with proteins such as KLH in order to produce antibodies to the chimeric molecule.

Monoclonal antibodies may be prepared using any technique which provides for the production of antibodies 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 *et al.* (1975) *Nature* 256:495-497; Kozbor *et al.* (1985) *J. Immunol Methods* 81:31-42; Cote *et al.* (1983) *Proc Natl Acad Sci* 80:2026-2030; and Cole *et al.* (1984) *Mol Cell Biol* 62:109-120.)

Alternatively, techniques described for antibody production may be adapted, using methods known in the art, to produce epitope-specific, single chain antibodies. Antibody fragments which contain specific binding sites for epitopes of the protein 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 *et al.* (1989) *Science* 246:1275-1281.)

The TNFR2PV or a portion thereof may be used in screening assays of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoassays 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 the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may also be employed (Pound (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TNFR2PV. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HSPDE10A-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 epitopes, represents the average affinity, or avidity, of the antibodies for TNFR2PV. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular 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 protein-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 TNFR2PV,

preferably in active form, from the antibody (Catty (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell and Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley, New York NY).

Labeling of Molecules for Assay

5 A wide variety of reporter molecules and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid, amino acid, and antibody assays. Synthesis of labeled molecules may be achieved using commercially available kits (Promega, Madison WI) for incorporation of a labeled nucleotide such as ³²P-dCTP (APB), Cy3-dCTP or Cy5-dCTP (Operon Technologies, Alameda CA), or amino acid such as ³⁵S-methionine (APB). Nucleotides and amino
10 acids may be directly labeled with a variety of substances including fluorescent, chemiluminescent, or chromogenic agents, and the like, by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC (Molecular Probes, Eugene OR).

DIAGNOSTICS

Nucleic Acid Assays

15 The cDNAs, fragments, oligonucleotides, complementary RNA and DNA molecules, and PNAs and may be used to detect and quantify differential gene expression for diagnosis of a disorder. Similarly antibodies which specifically bind TNFR2PV may be used to quantitate the protein. Disorders associated with differential expression include cancers, including cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as
20 rheumatoid arthritis, asthma, and ulcerative colitis. The diagnostic assay may use hybridization or amplification technology to compare gene expression in a biological sample from a patient to standard samples in order to detect differential gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

 For example, the cDNA or probe may be labeled by standard methods and added to a
25 biological sample from a patient under conditions for the formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization complexes, is quantified and compared with a standard value. If complex formation in the patient sample is significantly altered (higher or lower) in comparison to either a normal or disease standard, then differential expression indicates the presence of a disorder.

30 In order to provide standards for establishing differential expression, normal and disease expression profiles are established. This is accomplished by combining a sample taken from normal subjects, either animal or human, with a cDNA under conditions for hybridization to occur. Standard hybridization complexes may be quantified by comparing the values obtained using normal subjects with values from an experiment in which a known amount of a purified sequence is
35 used. Standard values obtained in this manner may be compared with values obtained from samples from patients who were diagnosed with a particular condition, disease, or disorder. Deviation from

standard values toward those associated with a particular disorder is used to diagnose that disorder.

Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies or in clinical trials or to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostic
5 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 a 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 years.

Protein Assays

10 Detection and quantification of a protein using either labeled amino acids or specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include two-dimensional polyacrylamide gel electrophoresis, enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). These assays and their quantitation against purified, labeled standards are well known in the art (Ausubel,
15 supra, unit 10.1-10.6). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may be employed. (See, e.g., Coligan et al. (1997) Current Protocols in Immunology, Wiley-Interscience, New York NY; and Pound, supra.)

20 THERAPEUTICS

Chemical and structural similarity, in particular the TNF/NGFR cysteine rich domains, exists between regions of TNFR2PV (SEQ ID NO:1) and the GenBank homologs shown in Figures 2A, 2B, and 2C for SEQ ID NOs:19 and 20. In addition, the transmembrane domain associated with the membrane bound, cell signalling form of the human TNF receptor 2 related protein
25 (g339762) is notably absent in TNFR2PV, indicating a potential soluble form of the receptor. Differential expression is also highly associated with cancers of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis. TNFR2PV clearly plays a role in these disorders, possibly as a natural antagonist of TNF signaling associated with these disorders.

30 Therefore, in the treatment of conditions associated with decreased expression of the protein or where further increased expression of the protein is desirable, such as in cancers of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis, it is desirable to increase expression or protein activity. In one embodiment, the protein, an agonist or enhancer may be administered to a subject
35 to treat a condition associated with decreased expression or activity. In another embodiment, a pharmaceutical composition comprising the protein, an agonist or enhancer and a pharmaceutical

carrier may be administered to a subject to treat a condition associated with the decreased expression or activity of the endogenous protein. In an additional embodiment, a vector expressing cDNA may be administered to a subject to treat the disorder.

Any of the cDNAs, complementary molecules, or fragments thereof, proteins or portions thereof, vectors delivering these nucleic acid molecules or expressing the proteins, and their ligands may be administered in combination with other therapeutic agents. Selection of the agents for use in combination therapy may be made by one of ordinary skill in the art according to conventional pharmaceutical principles. A combination of therapeutic agents may act synergistically to affect treatment of a particular disorder at a lower dosage of each agent.

10 Modification of Gene Expression Using Nucleic Acids

Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the gene encoding TNFR2PV. Oligonucleotides designed to inhibit transcription initiation are preferred. Similarly, inhibition can be achieved using triple helix base-pairing which inhibits the binding of polymerases, transcription factors, or regulatory molecules (Gee et al. In: Huber and Carr (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library or plurality of cDNAs may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

20 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 at sites such as GUA, GUU, and GUC. Once such sites are identified, an oligonucleotide with the same sequence may be evaluated for secondary structural features which would render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing their hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, in vitro or in vivo, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be modified to increase intracellular stability and half-life by addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. Either the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, methyl-, thio- groups renders the molecule less available to endogenous endonucleases.

35 Screening and Purification Assays

The cDNA encoding TNFR2PV may be used to screen a library or a plurality of molecules or compounds for specific binding affinity. The libraries may be DNA molecules, RNA molecules, PNAs, peptides, proteins such as transcription factors, enhancers, or repressors, and other ligands which regulate the activity, replication, transcription, or translation of the endogenous gene. The assay involves combining a polynucleotide with a library or plurality of molecules or compounds under conditions allowing specific binding, and detecting specific binding to identify at least one molecule which specifically binds the single-stranded or double-stranded molecule.

In one embodiment, the cDNA of the invention may be incubated with a plurality of purified molecules or compounds and binding activity determined by methods well known in the art, e.g., a gel-retardation assay (USPN 6,010,849) or a reticulocyte lysate transcriptional assay. In another embodiment, the cDNA may be incubated with nuclear extracts from biopsied and/or cultured cells and tissues. Specific binding between the cDNA and a molecule or compound in the nuclear extract is initially determined by gel shift assay and may be later confirmed by recovering and raising antibodies against that molecule or compound. When these antibodies are added into the assay, they cause a supershift in the gel-retardation assay.

In another embodiment, the cDNA may be used to purify a molecule or compound using affinity chromatography methods well known in the art. In one embodiment, the cDNA is chemically reacted with cyanogen bromide groups on a polymeric resin or gel. Then a sample is passed over and reacts with or binds to the cDNA. The molecule or compound which is bound to the cDNA may be released from the cDNA by increasing the salt concentration of the flow-through medium and collected.

In a further embodiment, the protein or a portion thereof may be used to purify a ligand from a sample. A method for using a protein or a portion thereof to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and using a chaotropic agent to separate the protein from the purified ligand.

In a preferred embodiment, TNFR2PV may be used to screen a plurality of molecules or compounds in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate (e.g. borne on a cell surface), or located intracellularly. For example, in one method, viable or fixed prokaryotic host cells that are stably transformed with recombinant nucleic acids that have expressed and positioned a peptide on their cell surface can be used in screening assays. The cells are screened against a plurality or libraries of ligands, and the specificity of binding or formation of complexes between the expressed protein and the ligand can be measured. Depending on the particular kind of molecules or compounds being screened, the assay may be used to identify DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies,

immunoglobulins, inhibitors, and drugs or any other ligand, which specifically binds the protein.

In one aspect, this invention contemplates a method for high throughput screening using very small assay volumes and very small amounts of test compound as described in USPN

5 5,876,946, incorporated herein by reference. This method is used to screen large numbers of molecules and compounds via specific binding. In another aspect, this invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound capable of binding to the protein. Molecules or compounds identified by screening may be used in a mammalian model system to evaluate their
10 toxicity, diagnostic, or therapeutic potential.

Pharmacology

Pharmaceutical compositions contain active ingredients in an effective amount to achieve a desired and intended purpose and a pharmaceutical carrier. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, the therapeutically
15 effective dose may be estimated initially either in cell culture assays or in animal models. The animal model is also used to achieve a desirable concentration range and route of administration. Such information may then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or inhibitor which
20 ameliorates the symptoms or condition. Therapeutic efficacy and toxicity of such agents may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it may be expressed as the ratio, LD_{50}/ED_{50} . Pharmaceutical compositions which exhibit large
25 therapeutic indexes are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for human use.

Model Systems

Animal models may be used as bioassays where they exhibit a phenotypic response similar to that of humans and where exposure conditions are relevant to human exposures. Mammals are
30 the most common models, and most infectious agent, cancer, drug, and toxicity studies are performed on rodents such as rats or mice because of low cost, availability, lifespan, reproductive potential, and abundant reference literature. Inbred and outbred rodent strains provide a convenient model for investigation of the physiological consequences of under- or over-expression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal
35 inbred to over-express a particular gene (for example, secreted in milk) may also serve as a convenient source of the protein expressed by that gene.

Toxicology

Toxicology is the study of the effects of agents on living systems. The majority of toxicity studies are performed on rats or mice. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic processes, and lethality in the rats or mice are used to generate a toxicity profile and to assess potential consequences on human health following exposure to the agent.

Genetic toxicology identifies and analyzes the effect of an agent on the rate of endogenous, spontaneous, and induced genetic mutations. Genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when chromosomal aberrations are transmitted to progeny. Toxicological studies may identify agents that increase the frequency of structural or functional abnormalities in the tissues of the progeny if administered to either parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because their short reproductive cycle allows the production of the numbers of organisms needed to satisfy statistical requirements.

Acute toxicity tests are based on a single administration of an agent to the subject to determine the symptomology or lethality of the agent. Three experiments are conducted: 1) an initial dose-range-finding experiment, 2) an experiment to narrow the range of effective doses, and 3) a final experiment for establishing the dose-response curve.

Subchronic toxicity tests are based on the repeated administration of an agent. Rat and dog are commonly used in these studies to provide data from species in different families. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose concentrations for periods of three to four months will reveal most forms of toxicity in adult animals.

Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity or the carcinogenic potential of an agent. When studies are conducted on rats, a minimum of three test groups plus one control group are used, and animals are examined and monitored at the outset and at intervals throughout the experiment.

Transgenic Animal Models

Transgenic rodents that over-express or under-express a gene of interest may be inbred and used to model human diseases or to test therapeutic or toxic agents. (See, e.g., USPN 5,175,383 and USPN 5,767,337.) In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal or postnatal development. Expression of the transgene is monitored by analysis of phenotype, of tissue-specific mRNA expression, or of serum and tissue protein levels in transgenic animals before, during, and after challenge with experimental drug therapies.

Embryonic Stem Cells

Embryonic (ES) stem cells isolated from rodent embryos retain the potential to form embryonic tissues. When ES cells are placed inside a carrier embryo, they resume normal development and contribute to tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors used to produce a transgenic strain contain a disease gene candidate and a marker gen, the latter serves to identify the presence of the introduced disease gene. The vector is transformed into ES cells by methods well known in the art, and 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.

ES cells derived from human blastocysts may be manipulated in vitro to differentiate into at least eight separate cell lineages. These lineages are used to study the differentiation of various cell types and tissues in vitro, and they include endoderm, mesoderm, and ectodermal cell types which differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes.

Knockout Analysis

In gene knockout analysis, a region of a mammalian gene is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (neo; Capecchi (1989) Science 244:1288-1292). The modified gene is transformed into cultured ES cells and integrates into the endogenous genome by homologous recombination. The inserted sequence disrupts transcription and translation of the endogenous gene. Transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines which lack a functional copy of the mammalian gene. In one example, the mammalian gene is a human gene.

Knockin Analysis

ES cells can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human 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 the analogous human condition. These methods have been used to model several human diseases.

Non-Human Primate Model

The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in evaluating the effects of therapeutic agents on non-human primates as they can be related to human

health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures under similar conditions. Cynomolgus and Rhesus monkeys (Macaca fascicularis and Macaca mulatta, respectively) and Common Marmosets (Callithrix jacchus) are the most common non-human primates (NHPs) used in these investigations. Since
5 great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be classified as a range of phenotypes from “extensive metabolizers” to “poor metabolizers” of these agents.

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

15 EXAMPLES

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention. The preparation of the human wrist synovial membrane library (SYNORAT03) library will be described.

20 I cDNA Library Construction

SYNORAT03

The SYNORAT03 library was derived from wrist synovial membrane tissue isolated from a 56-year-old female with rheumatoid arthritis. The frozen tissue was homogenized and lysed using a POLYTRON homogenizer (Brinkmann Instruments, Westbury NJ). The reagents and extraction
25 procedures were used as supplied in the RNA Isolation kit (Stratagene). The lysate was centrifuged over a 5.7 M CsCl cushion using an SW28 rotor in an L8-70M ultracentrifuge (Beckman Coulter, Fullerton CA) for 18 hr at 25,000 rpm at ambient temperature. The RNA was extracted twice with phenol chloroform, pH 8.0, and once with acid phenol, pH 4.0; precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol; resuspended in water; and treated with DNase for 15 min at
30 37C. The RNA was isolated with the OLIGOTEX kit (Qiagen, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERSCRIP^T plasmid system (Life Technologies) which contains a NotI primer-adaptor designed to prime the first strand cDNA synthesis at the poly(A) tail of mRNAs. Double stranded cDNA was blunted,
35 ligated to EcoRI adaptors and digested with NotI (New England Biolabs, Beverly MA). The cDNAs were fractionated on a SEPHAROSE CL4B column (APB), and those cDNAs exceeding

400 bp were ligated into pSPORT plasmid (Incyte Genomics). The plasmid pSPORT was subsequently transformed into DH5 α competent cells (Life Technologies).

II Construction of pINCY Plasmid

The plasmid was constructed by digesting the pSPORT1 plasmid (Life Technologies) with
5 EcoRI restriction enzyme (New England Biolabs, Beverly MA) and filling the overhanging ends
using Klenow enzyme (New England Biolabs) and 2'-deoxynucleotide 5'-triphosphates (dNTPs).
The plasmid was self-ligated and transformed into the bacterial host, *E. coli* strain JM109.

An intermediate plasmid, pSPORT 1- Δ RI, which showed no digestion with EcoRI, was
digested with Hind III (New England Biolabs); and the overhanging ends were filled in with
10 Klenow and dNTPs. A linker sequence was phosphorylated, ligated onto the 5' blunt end, digested
with EcoRI, and self-ligated. Following transformation into JM109 host cells, plasmids were
isolated and tested for preferential digestibility with EcoRI, but not with Hind III. A single colony
that met this criteria was designated pINCY plasmid.

After testing the plasmid for its ability to incorporate cDNAs from a library prepared using
15 NotI and EcoRI restriction enzymes, several clones were sequenced; and a single clone containing
an insert of approximately 0.8 kb was selected from which to prepare a large quantity of the
plasmid. After digestion with NotI and EcoRI, the plasmid was isolated on an agarose gel and
purified using a QIAQUICK column (Qiagen) for use in library construction.

III Isolation and Sequencing of cDNA Clones

20 Plasmid DNA was released from the cells and purified using either the MINIPREP kit
(Edge Biosystems, Gaithersburg MD) or the REAL PREP 96 plasmid kit (Qiagen). A kit consists
of a 96-well block with reagents for 960 purifications. The recommended protocol was employed
except for the following changes: 1) the bacteria were cultured in 1 ml of sterile TERRIFIC
BROTH (BD Biosciences, Sparks MD) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after
25 inoculation, the cells were cultured for 19 hours and then lysed with 0.3 ml of lysis buffer; and 3)
following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled
water. After the last step in the protocol, samples were transferred to a 96-well block for storage at
4C.

The cDNAs were prepared for sequencing using the MICROLAB 2200 system (Hamilton)
30 in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were
sequenced by the method of Sanger and Coulson (1975; J Mol Biol 94:441-448) using an ABI
PRISM 377 sequencing system (Applied Biosystems) or the MEGABACE 1000 DNA sequencing
system (APB). Most of the isolates were sequenced according to standard ABI protocols and kits
(Applied Biosystems) with solution volumes of 0.25x-1.0x concentrations. In the alternative,
35 cDNAs were sequenced using solutions and dyes from APB.

IV Extension of cDNA Sequences

The cDNAs were extended using the cDNA clone and oligonucleotide primers. One primer was synthesized to initiate 5' extension of the known fragment, and the other, to initiate 3' extension of the known fragment. The initial primers were designed using commercially available primer analysis software 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 68C to about 72C. Any stretch of nucleotides that would result in hairpin structures and primer-primer dimerizations was avoided.

Selected cDNA libraries were used as templates to extend the sequence. If more than one extension was necessary, additional or nested sets of primers were designed. Preferred libraries have been size-selected to include larger cDNAs and random primed to contain more sequences with 5' or upstream regions of genes. Genomic libraries are used to obtain regulatory elements, especially extension into the 5' promoter binding region.

High fidelity amplification was obtained by PCR using methods such as that taught in USPN 5,932,451. PCR was performed in 96-well plates using the DNA ENGINE thermal cycler (MJ Research). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (APB), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B (Incyte Genomics): Step 1: 94C, three min; Step 2: 94C, 15 sec; Step 3: 60C, one min; Step 4: 68C, two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C, five min; Step 7: storage at 4C. In the alternative, the parameters for primer pair T7 and SK+ (Stratagene) were as follows: Step 1: 94C, three min; Step 2: 94C, 15 sec; Step 3: 57C, one min; Step 4: 68C, two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C, five min; Step 7: storage at 4C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% reagent in 1x TE, v/v; Molecular Probes) and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning, Acton MA) and allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy) 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 minigel to determine which reactions were successful in extending the sequence.

The extended clones were desalted, 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 pUC18 vector (APB). For shotgun sequences, the digested nucleotide sequences were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and the agar was digested with AGARACE enzyme (Promega). Extended clones were religated using T4 DNA ligase (New England Biolabs) into pUC18 vector (APB), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into E. coli

competent cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37C in 384-well plates in LB/2x carbenicillin liquid media.

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

V Homology Searching of cDNA Clones and Their Deduced Proteins

The cDNAs of the Sequence Listing or their deduced amino acid sequences were used to query databases such as GenBank, SwissProt, BLOCKS, and the like. These databases that contain previously identified and annotated sequences or domains were searched using BLAST or BLAST2 to produce alignments and to determine which sequences were exact matches or homologs. The alignments were to sequences of prokaryotic (bacterial) or eukaryotic (animal, fungal, or plant) origin. Alternatively, algorithms such as the one described in Smith and Smith (1992, Protein Engineering 5:35-51) could have been used to deal with primary sequence patterns and secondary structure gap penalties. All of the sequences disclosed in this application have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

As detailed in Karlin and Altschul (1993; Proc Natl Acad Sci 90:5873-5877), BLAST matches between a query sequence and a database sequence were evaluated statistically and only reported when they satisfied the threshold of 10^{-25} for nucleotides and 10^{-14} for peptides. Homology was also evaluated by product score calculated as follows: the % nucleotide or amino acid identity [between the query and reference sequences] in BLAST is multiplied by the % maximum possible BLAST score [based on the lengths of query and reference sequences] and then divided by 100. In comparison with hybridization procedures used in the laboratory, the stringency for an exact match was set from a lower limit of about 40 (with 1-2% error due to uncalled bases) to a 100% match of about 70.

The BLAST software suite (NCBI, Bethesda MD; <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>), includes various sequence analysis programs including "blastn" that is used to align nucleotide sequences and BLAST2 that is used for direct pairwise comparison of either nucleotide or amino acid sequences. BLAST programs are commonly used with gap and other parameters set to default settings, e.g.: 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; and Filter: on. Identity is measured over the entire length of a sequence. Brenner *et al.* (1998; Proc Natl Acad Sci 95:6073-6078, incorporated herein by reference) analyzed BLAST for its ability to identify structural homologs by sequence identity and found 30% identity is a reliable threshold for sequence alignments of at least 150 residues and 40%, for alignments of at least 70 residues.

The cDNAs of this application were compared with assembled consensus sequences or templates found in the LIFESEQ GOLD database (Incyte Genomics). Component sequences from cDNA, extension, full length, and shotgun sequencing projects were subjected to PHRED analysis and assigned a quality score. All sequences with an acceptable quality score were subjected to various pre-processing and editing pathways to remove low quality 3' ends, vector and linker sequences, polyA tails, Alu repeats, mitochondrial and ribosomal sequences, and bacterial contamination sequences. Edited sequences had to be at least 50 bp in length, and low-information sequences and repetitive elements such as dinucleotide repeats, Alu repeats, and the like, were replaced by "Ns" or masked.

Edited sequences were subjected to assembly procedures in which the sequences were assigned to gene bins. Each sequence could only belong to one bin, and sequences in each bin were assembled to produce a template. Newly sequenced components were added to existing bins using BLAST and CROSSMATCH. To be added to a bin, the component sequences had to have a BLAST quality score greater than or equal to 150 and an alignment of at least 82% local identity. The sequences in each bin were assembled using PHRAP. Bins with several overlapping component sequences were assembled using DEEP PHRAP. The orientation of each template was determined based on the number and orientation of its component sequences.

Bins were compared to one another, and those having local similarity of at least 82% were combined and reassembled. Bins having templates with less than 95% local identity were split. Templates were subjected to analysis by STITCHER/EXON MAPPER algorithms that determine the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types or disease states, and the like. Assembly procedures were repeated periodically, and templates were annotated using BLAST against GenBank databases such as GBpri. An exact match was defined as having from 95% local identity over 200 base pairs through 100% local identity over 100 base pairs and a homolog match as having an E-value (or probability score) of $\leq 1 \times 10^{-8}$. The templates were also subjected to frameshift FASTx against GENPEPT, and homolog match was defined as having an E-value of $\leq 1 \times 10^{-8}$. Template analysis and assembly was described in USSN 09/276,534, filed March 25, 1999.

Following assembly, templates were subjected to BLAST, motif, and other functional analyses and categorized in protein hierarchies using methods described in USSN 08/812,290 and USSN 08/811,758, both filed March 6, 1997; in USSN 08/947,845, filed October 9, 1997; and in

USSN 09/034,807, filed March 4, 1998. Then templates were analyzed by translating each template in all three forward reading frames and searching each translation against the PFAM database of hidden Markov model-based protein families and domains using the HMMER software package (Washington University School of Medicine, St. Louis MO; <http://pfam.wustl.edu/>). The cDNA was further analyzed using MACDNASIS PRO software (Hitachi Software Engineering), and LASERGENE software (DNASTAR) and queried against public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote databases, SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

VI Chromosome Mapping

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 are used to determine if any of the cDNAs presented in the Sequence Listing have been mapped. Any of the fragments of the cDNA encoding TNFR2PV that have been mapped result in the assignment of all related regulatory and coding sequences to the same location. The genetic map locations are described as ranges, or intervals, of human chromosomes. The map position of an interval, in cM (which is roughly equivalent to 1 megabase of human DNA), is measured relative to the terminus of the chromosomal p-arm.

VII Hybridization Technologies and Analyses

Immobilization of cDNAs on a Substrate

The cDNAs are applied to a substrate by one of the following methods. A mixture of cDNAs is fractionated by gel electrophoresis and transferred to a nylon membrane by capillary transfer. Alternatively, the cDNAs are individually ligated to a vector and inserted into bacterial host cells to form a library. The cDNAs are then arranged on a substrate by one of the following methods. In the first method, bacterial cells containing individual clones are robotically picked and arranged on a nylon membrane. The membrane is placed on LB agar containing selective agent (carbenicillin, kanamycin, ampicillin, or chloramphenicol depending on the vector used) and incubated at 37C for 16 hr. The membrane is removed from the agar and consecutively placed colony side up in 10% SDS, denaturing solution (1.5 M NaCl, 0.5 M NaOH), neutralizing solution (1.5 M NaCl, 1 M Tris, pH 8.0), and twice in 2xSSC for 10 min each. The membrane is then UV irradiated in a STRATALINKER UV-crosslinker (Stratagene).

In the second method, cDNAs are amplified from bacterial vectors by thirty cycles of PCR using primers complementary to vector sequences flanking the insert. PCR amplification increases a starting concentration of 1-2 ng nucleic acid to a final quantity greater than 5 µg. Amplified nucleic acids from about 400 bp to about 5000 bp in length are purified using SEPHACRYL-400 beads (APB). Purified nucleic acids are arranged on a nylon membrane manually or using a dot/slot blotting manifold and suction device and are immobilized by denaturation, neutralization, and UV

irradiation as described above. Purified nucleic acids are robotically arranged and immobilized on polymer-coated glass slides using the procedure described in USPN 5,807,522. Polymer-coated slides are prepared by cleaning glass microscope slides (Corning, Acton MA) by ultrasound in 0.1% SDS and acetone, etching in 4% hydrofluoric acid (VWR Scientific Products, West Chester PA),
5 coating with 0.05% aminopropyl silane (Sigma Aldrich) in 95% ethanol, and curing in a 110C oven. The slides are washed extensively with distilled water between and after treatments. The nucleic acids are arranged on the slide and then immobilized by exposing the array to UV irradiation using a STRATALINKER UV-crosslinker (Stratagene). Arrays are then washed at room temperature in 0.2% SDS and rinsed three times in distilled water. Non-specific binding sites are blocked by
10 incubation of arrays in 0.2% casein in phosphate buffered saline (PBS; Tropic, Bedford MA) for 30 min at 60C; then the arrays are washed in 0.2% SDS and rinsed in distilled water as before.

Probe Preparation for Membrane Hybridization

Hybridization probes derived from the cDNAs of the Sequence Listing are employed for screening cDNAs, mRNAs, or genomic DNA in membrane-based hybridizations. Probes are
15 prepared by diluting the cDNAs to a concentration of 40-50 ng in 45 μ l TE buffer, denaturing by heating to 100C for five min, and briefly centrifuging. The denatured cDNA is then added to a REDIPRIME tube (APB), gently mixed until blue color is evenly distributed, and briefly centrifuged. Five μ l of [³²P]dCTP is added to the tube, and the contents are incubated at 37C for 10 min. The labeling reaction is stopped by adding 5 μ l of 0.2M EDTA, and probe is purified from
20 unincorporated nucleotides using a PROBEQUANT G-50 microcolumn (APB). The purified probe is heated to 100C for five min, snap cooled for two min on ice, and used in membrane-based hybridizations as described below.

Probe Preparation for Polymer Coated Slide Hybridization

Hybridization probes derived from mRNA isolated from samples are employed for
25 screening cDNAs of the Sequence Listing in array-based hybridizations. Probe is prepared using the GEMbright kit (Incyte Genomics) by diluting mRNA to a concentration of 200 ng in 9 μ l TE buffer and adding 5 μ l 5x buffer, 1 μ l 0.1 M DTT, 3 μ l Cy3 or Cy5 labeling mix, 1 μ l RNase inhibitor, 1 μ l reverse transcriptase, and 5 μ l 1x yeast control mRNAs. Yeast control mRNAs are synthesized by in vitro transcription from noncoding yeast genomic DNA (W. Lei, unpublished).
30 As quantitative controls, one set of control mRNAs at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng are diluted into reverse transcription reaction mixture at ratios of 1:100,000, 1:10,000, 1:1000, and 1:100 (w/w) to sample mRNA respectively. To examine mRNA differential expression patterns, a second set of control mRNAs are diluted into reverse transcription reaction mixture at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, and 25:1 (w/w). The reaction mixture is mixed and incubated at 37C for two
35 hr. The reaction mixture is then incubated for 20 min at 85C, and probes are purified using two successive CHROMA SPIN+TE 30 columns (Clontech, Palo Alto CA). Purified probe is ethanol

precipitated by diluting probe to 90 μ l in DEPC-treated water, adding 2 μ l 1mg/ml glycogen, 60 μ l 5 M sodium acetate, and 300 μ l 100% ethanol. The probe is centrifuged for 20 min at 20,800xg, and the pellet is resuspended in 12 μ l resuspension buffer, heated to 65C for five min, and mixed thoroughly. The probe is heated and mixed as before and then stored on ice. Probe is used in high density array-based hybridizations as described below.

Membrane-based Hybridization

Membranes are pre-hybridized in hybridization solution containing 1% Sarkosyl and 1x high phosphate buffer (0.5 M NaCl, 0.1 M Na_2HPO_4 , 5 mM EDTA, pH 7) at 55C for two hr. The probe, diluted in 15 ml fresh hybridization solution, is then added to the membrane. The membrane is hybridized with the probe at 55C for 16 hr. Following hybridization, the membrane is washed for 15 min at 25C in 1mM Tris (pH 8.0), 1% Sarkosyl, and four times for 15 min each at 25C in 1mM Tris (pH 8.0). To detect hybridization complexes, XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the membrane overnight at -70C, developed, and examined visually.

Polymer Coated Slide-based Hybridization

Probe is heated to 65C for five min, centrifuged five min at 9400 rpm in a 5415C microcentrifuge (Eppendorf Scientific, Westbury NY), and then 18 μ l is aliquoted onto the array surface and covered with a 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 5xSSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hr at 60C. The arrays are washed for 10 min at 45C in 1xSSC, 0.1% SDS, and three times for 10 min each at 45C in 0.1xSSC, and dried.

Hybridization reactions are performed in absolute or differential hybridization formats. In the absolute hybridization format, probe from one sample is hybridized to array elements, and signals are detected after hybridization complexes form. Signal strength correlates with probe mRNA levels in the sample. In the differential hybridization format, differential expression of a set of genes in two biological samples is analyzed. Probes from the two samples are prepared and labeled with different labeling moieties. A mixture of the two labeled probes is hybridized to the array elements, and signals are examined under conditions in which the emissions from the two different labels are individually detectable. Elements on the array that are hybridized to equal numbers of probes derived from both biological samples give a distinct combined fluorescence (Shalon WO95/35505).

Hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, 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, 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

with a resolution of 20 micrometers. In the differential hybridization format, the two fluorophores are sequentially excited by the laser. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Filters positioned between the array and the photomultiplier tubes are used to separate the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. The sensitivity of the scans is calibrated using the signal intensity generated by the yeast control mRNAs added to the probe mix. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000.

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

VIII Electronic Analysis

BLAST was used to search for identical or related molecules in the GenBank or LIFESEQ databases (Incyte Genomics). The product score for human and rat sequences was calculated as follows: the BLAST score is multiplied by the % nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences), such that a 100% alignment over the length of the shorter sequence gives a product score of 100. The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and with a product score of at least 70, the match will be exact. Similar or related molecules are usually identified by selecting those which show product scores between 8 and 40.

Electronic northern analysis was performed at a product score of 70. All sequences and cDNA libraries in the LIFESEQ database were categorized by system, organ/tissue and cell type. The categories included cardiovascular system, connective tissue, digestive system, embryonic structures, endocrine system, exocrine glands, female and male genitalia, germ cells, hemic/immune system, liver, musculoskeletal system, nervous system, pancreas, respiratory system, sense organs, skin, stomatognathic system, unclassified/mixed, and the urinary tract. For each category, the

number of libraries in which the sequence was expressed were counted and shown over the total number of libraries in that category. In a non-normalized library, significant expression may reflect presence or absence or differential expression of the cDNA.

IX Complementary Molecules

- 5 Molecules complementary to the cDNA, from about 5 (PNA) to about 5000 bp (complement of a cDNA insert), are used to detect or inhibit gene expression. Detection is described in Example VII. To inhibit transcription by preventing promoter binding, the complementary molecule is designed to bind to the most unique 5' sequence and includes nucleotides of the 5' UTR upstream of the initiation codon of the open reading frame.
- 10 Complementary molecules include genomic sequences (such as enhancers or introns) and are used in "triple helix" base pairing to compromise the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. To inhibit translation, a complementary molecule is designed to prevent ribosomal binding to the mRNA encoding the protein.
- 15 Complementary molecules are placed in expression vectors and used to transform a cell line to test efficacy; into an organ, tumor, synovial cavity, or the vascular system for transient or short term therapy; or into a stem cell, zygote, or other reproducing lineage for long term or stable gene therapy. Transient expression lasts for a month or more with a non-replicating vector and for three months or more if elements for inducing vector replication are used in the
- 20 transformation/expression system.

Stable transformation of dividing cells with a vector encoding the complementary molecule produces a transgenic cell line, tissue, or organism (USPN 4,736,866). Those cells that assimilate and replicate sufficient quantities of the vector to allow stable integration also produce enough complementary molecules to compromise or entirely eliminate activity of the cDNA encoding the

25 protein.

X Expression of TNFR2PV

- Expression and purification of the protein are achieved using either a mammalian cell expression system or an insect cell expression system. The pUB6/V5-His vector system (Invitrogen, Carlsbad CA) is used to express TNFR2PV in CHO cells. The vector contains the
- 30 selectable bsd gene, multiple cloning sites, the promoter/enhancer sequence from the human ubiquitin C gene, a C-terminal V5 epitope for antibody detection with anti-V5 antibodies, and a C-terminal polyhistidine (6xHis) sequence for rapid purification on PROBOND resin (Invitrogen). Transformed cells are selected on media containing blasticidin.

- Spodoptera frugiperda (Sf9) insect cells are infected with recombinant Autographica californica nuclear polyhedrosis virus (baculovirus). The polyhedrin gene is replaced with the
- 35 cDNA by homologous recombination and the polyhedrin promoter drives cDNA transcription. The

protein is synthesized as a fusion protein with 6xhis which enables purification as described above. Purified protein is used in the following activity and to make antibodies

XI Production of Antibodies

TNFR2PV is purified using polyacrylamide gel electrophoresis and used to immunize mice
5 or rabbits. Antibodies are produced using the protocols below. Alternatively, the amino acid sequence of TNFR2PV is analyzed using LASERGENE software (DNASTAR) to determine regions of high antigenicity. An antigenic epitope, usually found near the C-terminus or in a hydrophilic region is selected, synthesized, and used to raise antibodies. Typically, epitopes of about 15 residues in length are produced using an ABI 431A peptide synthesizer (Applied
10 Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase antigenicity.

Rabbits are immunized with the epitope-KLH complex in complete Freund's adjuvant. Immunizations are repeated at intervals thereafter in incomplete Freund's adjuvant. After a minimum of seven weeks for mouse or twelve weeks for rabbit, antisera are drawn and tested for
15 antipeptide activity. Testing involves binding the peptide to plastic, blocking with 1% bovine serum albumin, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. Methods well known in the art are used to determine antibody titer and the amount of complex formation.

XII Purification of Naturally Occurring Protein Using Specific Antibodies

Naturally occurring or recombinant protein is purified by immunoaffinity chromatography
20 using antibodies which specifically bind the protein. An immunoaffinity column is constructed by covalently coupling the antibody to CNBr-activated SEPHAROSE resin (APB). Media containing the protein is passed over the immunoaffinity column, and the column is washed using high ionic strength buffers in the presence of detergent to allow preferential absorbance of the protein. After
25 coupling, the protein is eluted from the column using a buffer of pH 2-3 or a high concentration of urea or thiocyanate ion to disrupt antibody/protein binding, and the protein is collected.

XIII Screening Molecules for Specific Binding with the cDNA or Protein

The cDNA, or fragments thereof, or the protein, or portions thereof, are labeled with ³²P-dCTP, Cy3-dCTP, or Cy5-dCTP (APB), or with BIODIPY or FITC (Molecular Probes, Eugene
30 OR), respectively. Libraries of candidate molecules or compounds previously arranged on a substrate are incubated in the presence of labeled cDNA or protein. After incubation under conditions for either a nucleic acid or amino acid sequence, the substrate is washed, and any position on the substrate retaining label, which indicates specific binding or complex formation, is assayed, and the ligand is identified. Data obtained using different concentrations of the nucleic
35 acid or protein are used to calculate affinity between the labeled nucleic acid or protein and the bound molecule.

XIV Two-Hybrid Screen

A yeast two-hybrid system, MATCHMAKER LexA Two-Hybrid system (Clontech Laboratories, Palo Alto CA), is used to screen for peptides that bind the protein of the invention. A cDNA encoding the protein is inserted into the multiple cloning site of a pLexA vector, ligated, and transformed into *E. coli*. cDNA, prepared from mRNA, is inserted into the multiple cloning site of a pB42AD vector, ligated, and transformed into *E. coli* to construct a cDNA library. The pLexA plasmid and pB42AD-cDNA library constructs are isolated from *E. coli* and used in a 2:1 ratio to co-transform competent yeast EGY48[p8op-lacZ] cells using a polyethylene glycol/lithium acetate protocol. Transformed yeast cells are plated on synthetic dropout (SD) media lacking histidine (-His), tryptophan (-Trp), and uracil (-Ura), and incubated at 30C until the colonies have grown up and are counted. The colonies are pooled in a minimal volume of 1x TE (pH 7.5), replated on SD/-His/-Leu/-Trp/-Ura media supplemented with 2% galactose (Gal), 1% raffinose (Raf), and 80 mg/ml 5-bromo-4-chloro-3-indolyl β -d-galactopyranoside (X-Gal), and subsequently examined for growth of blue colonies. Interaction between expressed protein and cDNA fusion proteins activates expression of a LEU2 reporter gene in EGY48 and produces colony growth on media lacking leucine (-Leu). Interaction also activates expression of β -galactosidase from the p8op-lacZ reporter construct that produces blue color in colonies grown on X-Gal.

Positive interactions between expressed protein and cDNA fusion proteins are verified by isolating individual positive colonies and growing them in SD/-Trp/-Ura liquid medium for 1 to 2 days at 30C. A sample of the culture is plated on SD/-Trp/-Ura media and incubated at 30C until colonies appear. The sample is replica-plated on SD/-Trp/-Ura and SD/-His/-Trp/-Ura plates. Colonies that grow on SD containing histidine but not on media lacking histidine have lost the pLexA plasmid. Histidine-requiring colonies are grown on SD/Gal/Raf/X-Gal/-Trp/-Ura, and white colonies are isolated and propagated. The pB42AD-cDNA plasmid, which contains a cDNA encoding a protein that physically interacts with the protein, is isolated from the yeast cells and characterized.

XV TNFR2PV Assay

An assay for TNFR2PV activity measures the inhibition of TNF α -induced cytotoxicity in vitro as described by Van Zee et al., supra. Briefly, varying concentrations of TNFR2PV are added to normal human plasma to which 1.5 ng of recombinant TNF α was added, and the mixtures allowed to incubate at room temperature for 30 minutes. The incubations are then assayed for TNF α -induced cytotoxicity using the WEHI 164 clone 13 fibroblast assay (Van Zee et al., supra). Cytotoxicity is expressed as the percent activity relative to a control incubation without TNFR2PV, and the activity of TNFR2PV in the sample is proportional to the decrease in cytotoxicity relative to the control incubation.

All patents and publications mentioned in the specification are incorporated by reference herein. Various modifications and variations of the described method and system 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 specific preferred
5 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 that are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

What is claimed is:

1. An isolated cDNA, or the complement of the cDNA, comprising a nucleic acid sequence encoding a polypeptide selected from:
 - a) an amino acid sequence of SEQ ID NO:1;
 - 5 b) a biologically active portion of SEQ ID NO:1;
 - c) an antigenic epitope of the protein of SEQ ID NO:1, and
 - d) an amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:1.
2. An isolated cDNA comprising a nucleic acid sequence selected from:
 - 10 a) SEQ ID NO:2 or the complement thereof;
 - b) a fragment of SEQ ID NO:2, or the complement thereof; and
 - c) a variant of SEQ ID NO:2 having at least 85% identity to SEQ ID NO:2, or the complement thereof.
3. A probe consisting of the cDNA of claim 2.
- 15 4. A composition comprising the cDNA of claim 2 and a labeling moiety.
5. An array element comprising the cDNA of claim 2.
6. A substrate upon which the cDNA of claim 2 is immobilized.
7. A vector comprising the cDNA of claim 1.
8. A host cell comprising the vector of claim 7.
- 20 9. A method for using a cDNA to produce a protein, the method comprising:
 - a) culturing the host cell of claim 8 under conditions for protein expression; and
 - b) recovering the protein from the host cell culture.
10. A method for using a cDNA to detect expression of a nucleic acid in a sample comprising:
 - 25 a) hybridizing the composition of claim 4 to nucleic acids of the sample under conditions to form at least one hybridization complex; and
 - b) detecting hybridization complex formation, wherein complex formation indicates expression of the nucleic acid in the sample.
11. The method of claim 10 further comprising amplifying the nucleic acids of the sample prior to hybridization.
- 30 13. The method of claim 10 wherein the composition is attached to a substrate.
14. The method of claim 10 wherein the sample is from prostate, ovary, gallbladder, breast, brain, liver, or colon.
15. The method of claim 10 wherein complex formation is compared to standards and is diagnostic of a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, and inflammatory disorders, such as rheumatoid arthritis, asthma, and ulcerative colitis.
- 35 16. A method of using a cDNA to screen a plurality of molecules or compounds, the method

comprising:

- a) combining the cDNA of claim 2 with a plurality of molecules or compounds under conditions to allow specific binding; and
- b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the cDNA.

17. The method of claim 16 wherein the molecules or compounds are selected from antisense molecules, artificial chromosome constructions, DNA molecules, enhancers, peptide nucleic acids, peptides, proteins, repressors, RNA molecules, and transcription factors.

18. A method for using a cDNA to assess efficacy of a molecule or compound, the method

comprising:

- a) treating a sample containing nucleic acids with the molecule or compound;
- b) hybridizing the nucleic acids of the sample with the cDNA of claim 2 under conditions for hybridization complex formation;
- c) determining the amount of complex formation; and
- d) comparing the amount of complex formation in the treated sample with the amount of complex formation in an untreated sample, wherein a difference in complex formation indicates efficacy of the molecule or compound.

19. A method for using a cDNA to assess toxicity of a molecule or compound, the method comprising:

- a) treating a sample containing nucleic acids with the molecule or compound;
- b) hybridizing the nucleic acids with the cDNA of claim 2 under conditions for hybridization complex formation;
- c) determining the amount of complex formation; and
- d) comparing the amount of complex formation in the treated sample with the amount of

complex formation in an untreated sample, wherein a difference in complex formation indicates the toxicity of the molecule or compound.

20. A purified protein comprising a polypeptide selected from:

- a) an amino acid sequence of SEQ ID NO:1;
- b) a biologically active portion of SEQ ID NO:1;
- c) an antigenic epitope of the protein of SEQ ID NO:1, and
- d) an amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:1.

21. A composition comprising the protein of claim 20 and a labeling moiety.

22. A composition comprising the protein of claim 20 and a pharmaceutical carrier.

23. A substrate upon which the protein of claim 20 is immobilized.

24. An array element comprising the protein of claim 20.

25. A method for detecting expression of a protein having the amino acid sequence of SEQ ID NO:1 in a sample, the method comprising:

- a) performing an assay to determine the amount of the protein of claim 20 in a sample; and
- b) comparing the amount of protein to standards, thereby detecting expression of the protein

5 in the sample.

26. The method of claim 25 wherein the assay is selected from antibody arrays, enzyme-linked immunosorbent assays, fluorescence-activated cell sorting, two dimensional-polyacrylamide gel electrophoresis and scintillation counting, radioimmunoassays, and western analysis.

27. The method of claim 25 wherein the protein is differentially expressed when compared with the standard and is diagnostic of a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, rheumatoid arthritis, asthma, or ulcerative colitis.

28. A method for using a protein to screen a plurality of molecules and compounds to identify at least one ligand, the method comprising:

- a) combining the protein of claim 20 with a plurality of molecules and compounds under conditions to allow specific binding; and
- b) detecting specific binding, thereby identifying a ligand that specifically binds the protein.

29. The method of claim 28 wherein the molecules and compounds are selected from agonists, antagonists, bispecific molecules, DNA molecules, small drug molecules, immunoglobulins, inhibitors, mimetics, multispecific molecules, peptides, peptide nucleic acids, pharmaceutical agent, proteins, and RNA molecules.

30. A method for using a protein to identify an antibody that specifically binds the protein having the amino acid sequence of SEQ ID NO:1 comprising:

- a) contacting a plurality of antibodies with the protein of claim 20 under conditions to allow specific binding, and
- b) detecting specific binding between an antibody and the protein, thereby identifying an antibody that specifically binds the protein.

31. The method of claim 30, wherein the plurality of antibodies are selected from a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a recombinant antibody, a humanized antibody, a single chain antibody, a Fab fragment, an F(ab')₂ fragment, an Fv fragment; and an antibody-peptide fusion protein.

32. A method of using a protein to prepare and purify a polyclonal antibody comprising:

- a) immunizing a animal with a protein of claim 20 under conditions to elicit an antibody response;
- b) isolating animal antibodies;
- c) attaching the protein to a substrate;
- d) contacting the substrate with isolated antibodies under conditions to allow specific binding

to the protein; and

e) dissociating the antibodies from the protein, thereby obtaining purified polyclonal antibodies.

33. A method of using a protein to prepare a monoclonal antibody comprising:

- 5 a) immunizing a animal with a protein of claim 20 under conditions to elicit an antibody response;
- b) isolating antibody-producing cells from the animal;
- c) fusing the antibody-producing cells with immortalized cells in culture to form monoclonal antibody producing hybridoma cells;
- 10 d) culturing the hybridoma cells; and
- e) isolating from culture monoclonal antibody that specifically binds the protein.

34. A method for using a protein to diagnose a disease or condition associated with increased TNF signaling in a sample, comprising:

- a) performing an assay to quantify the expression of the protein of claim 20 in a sample; and
- 15 b) comparing the expression of the protein to standards, wherein the comparison indicates increased expression of the protein in the sample.

35. The method of claim 34 wherein increased expression is diagnostic of a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, rheumatoid arthritis, asthma, or ulcerative colitis.

36. A method for testing a molecule or compound for effectiveness as an agonist comprising:

- 20 a) exposing a sample comprising the protein of claim 20 to the molecule or compound; and
- b) detecting agonist activity in the sample.

37. A method for testing a molecule or compound for effectiveness as an antagonist, the method comprising:

- a) exposing a sample comprising the protein of claim 20 to a molecule or compound; and
- 25 b) detecting antagonist activity in the sample.

38. An isolated antibody that specifically binds a protein having the amino acid sequence of SEQ ID NO:1.

39. A polyclonal antibody produced by the method of claim 32.

40. A monoclonal antibody produced by the method of claim 33.

30 41. A method for using an antibody to detect expression of a protein in a sample, the method comprising:

- a) combining the antibody of claim 38 with a sample under conditions which allow the formation of antibody:protein complexes; and
- b) detecting complex formation, wherein complex formation indicates expression of the
- 35 protein in the sample.

42. The method of claim 41 wherein complex formation is compared with standards and is

diagnostic of a cancer of a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, rheumatoid arthritis, asthma, or ulcerative colitis.

43. A method for using an antibody to immunopurify a protein comprising:

- a) attaching the antibody of claim 38 to a substrate;
 - 5 b) exposing the antibody to a sample containing protein under conditions to allow antibody:protein complexes to form;
 - c) dissociating the protein from the complex; and
 - d) collecting the purified protein.
44. A composition comprising an antibody of claim 38 and a labeling moiety.
- 10 45. A kit comprising the composition of claim 44.
46. An array element comprising the antibody of claim 38.
47. A substrate upon which the antibody of claim 38 is immobilized.
48. A composition comprising an antibody of claim 38 and a pharmaceutical agent
49. A method for treating a disease or condition associated with increased TNF signaling
- 15 comprising administering to a subject in need of therapeutic intervention the composition of claim 22.
50. The method of claim 49, wherein the disease or condition is selected from a cancer of the prostate, ovary, gallbladder, breast, brain, liver, or colon, rheumatoid arthritis, asthma, and ulcerative colitis.
- 20 51. An agonist that specifically binds the protein of claim 1.
52. A composition comprising an agonist of claim 51 and a pharmaceutical carrier.
53. An antagonist that specifically binds the protein of claim 1.
54. A composition comprising the antagonist of claim 53 and a pharmaceutical carrier.
55. A pharmaceutical agent that specifically binds the protein of claim 1.
- 25 56. A composition comprising the pharmaceutical agent of claim 55 and a pharmaceutical carrier.

1/10

```

11      20      29      38      47      56
5'GC CCC GCG GCC AGC TCG CTC CAC TCC CAC TTC CTG AGC TCC GCC ATG GGA GCC CTG

65      74      83      92     101     110
GAG GCC CGG CCT GGC GGC TCC CGG CCC TGG GGT GCA CAT CGG CCC TGA GTC CCG

119     128     137     146     155     164
TCC CAG GCT CTG GGC TCG GGC AGC CGC CGC CAC CGC TGC CCA GGA CGT CGG GCC

173     182     191     200     209     218
TCC TGC CTT CCT CCC AGG CCC CCA CGT TGC TGG CCG CCT GGC CGA GTG GCC GCC

227     236     245     254     263     272
ATG CTC CTG CCT TGG GCC ACC TCT GCC CCC GGC CTG GCC TGG GGG CCT CTG GTG
M   L   L   P   W   A   T   S   A   P   G   L   A   W   G   P   L   V

281     290     299     308     317     326
CTG GGC CTC TTC GGG CTC CTG GCA GCA TCG CAG CCC CAG GCG GTG CCT CCA TAT
L   G   L   F   G   L   L   A   A   S   Q   P   Q   A   V   P   P   Y

335     344     353     362     371     380
GCG TCG GAG AAC CAG ACC TGC AGG GAC CAG GAA AAG GAA TAC TAT GAG CCC CAG
A   S   E   N   Q   T   C   R   D   Q   E   K   E   Y   Y   E   P   Q

```

FIGURE 1A

CAC H	389	CGC	ATC	TGC	TGC	TGC	CGC	CGC	CGC	CGC	CCA	GGC	ACC	416	TAT	GTC	TCA	GCT	AAA	434
		ATC	TGC	TGC	CGC	CGC	CGC	CCA	GGC	ACC	TAT	GTC	TCA	GCT	AAA	TGT				
AGC S	443	CGC	ATC	CGG	GAC	ACA	GTT	TGT	TGC	ACA	TGT	GCC	470	GAG	AAT	TCC	TAC	AAC	488	
		ATC	CGG	GAC	ACA	GTT	TGT	TGC	ACA	TGT	GCC	AAT	TCC	TAC	AAC	GAG				
CAC H	497	TGG	AAC	TAC	CTG	ACC	ATC	TGC	CTG	TGC	CGC	524	CCC	TGT	GAC	CCA	GTG	ATG	542	
		AAC	TAC	CTG	ACC	ATC	TGC	CTG	TGC	CGC	CCC	TGT	GAC	CCA	GTG	ATG				
GGC G	551	CTC	GAG	GAG	ATT	GCC	CCC	TGC	ACA	AGC	AAA	CGG	578	AAG	ACC	CAG	TGC	CGC	596	
		CTC	GAG	GAG	ATT	GCC	CCC	TGC	ACA	AGC	AAA	CGG	AAG	ACC	CAG	TGC	CGC	TGC		
CAG Q	605	CCG	GGA	ATG	TTC	TGT	GCT	GCC	TGG	GCC	CTC	632	TGT	ACA	CAC	TGC	GAG	CTA	650	
		GGA	ATG	TTC	TGT	GCT	GCC	TGG	GCC	CTC	GAG	TGT	ACA	CAC	TGC	GAG	CTA			
CTT L	659	TCT	GAC	TGC	CCG	CCT	GGC	ACT	GAA	GCC	GAG	686	AAA	GAT	GAA	GTT	GGG	AAG	704	
		TCT	GAC	TGC	CCG	CCT	GGC	ACT	GAA	GCC	GAG	CTC	AAA	GAT	GAA	GTT	GGG	AAG		
GGT G	713	AAC	AAC	CAC	TGC	GTC	CCC	TGC	AAG	GCA	GGG	740	TTC	CAG	AAT	ACC	TCC	TCC	758	
		AAC	CAC	TGC	GTC	CCC	TGC	AAG	GCA	GGG	CAC	TTC	CAG	AAT	ACC	TCC	TCC	TCC		

FIGURE 1B

3/10

767	776	785	794	803	812
CCC AGC GCC CGC TGC CAG CCC CAC ACC AGG TGT GAG AAC CAA GGT CTG GTG GAG					
P S A R C Q					
821	830	839	848	857	866
GCA GCT CCA GGC ACT GCC CAG TCC GAC ACA ACC TGC AAA AAT CCA TTA GAG CCA					
A A P G T A Q					
875	884	893	902	911	920
CTG CCC CCA GAG ATG TCA GGA TCG CTG CTC AAG AGG CGT CCG CAG GGA GAG GGA					
L P P E M S G S L L K R R P Q G E G					
929	938	947	956	965	974
CCC AAT CCT GTA GCT GGA AGC TGG GAG CCT CCG AAG GCC CAT CCA TAC TTC CCT					
P N P V A G S S W E P P K A H P Y F P					
983	992	1001	1010	1019	1028
GAC TTG GTA CAG CCA CTG CTA CCC ATT TCT GGA GAT GTT TCC CCA GTA TCC ACT					
D L V Q P L L L P I S G D V S P V S T					
1037	1046	1055	1064	1073	1082
GGG CTC CCC GCA GCC CCA GTC GTC GTC GTC GTC GTC CCG CAA CAG CAG AGT CCT					
G L P A A P V L E A A G V P Q Q S P					
1091	1100	1109	1118	1127	1136
CTG GAC CTG ACC AGG GAG CCG CAG TTG GAA CCC GGG GAG CAG AGC CAG GTG GCC					
L D L T R E P Q L E P G E Q S Q V A					

FIGURE 1C

1145	1154	1163	1172	1181	1190
CAC GGT ACC AAT GGC ATT CAT GTC ACC GGC GGG TCT ATG ACT ATC ACT GGC AAC					
H G T N G I H V T G G S M T I T G N					
1199	1208	1217	1226	1235	1244
ATC TAC ATC TAC AAT GGA CCA GTA CTG GGG GGA CCA CCG GGT CCT GGA GAC CTC					
I Y I Y N G P V L G G P P G P G D L					
1253	1262	1271	1280	1289	1298
CCA GCT ACC CCC GAA CCT CCA TAC CCC ATT CCC GAA GAG GGG GAC CCT GGC CCT					
P A T P E P P Y P I P E E G D P G P					
1307	1316	1325	1334	1343	1352
CCC GGG CTC TCT ACA CCC CAC CAG GAA GAT GGC AAG GCT TGG CAC CTA GCG GAG					
P G L S T P P H Q E D G K A W H L A E					
1361	1370	1379	1388	1397	1406
ACA GAG CAC TGT GGT GCC ACA CCC TCT AAC AGG GGC CCA AGG AAC CAA TTT ATC					
T E H C G A T P S N R G P R N Q F I					
1415	1424	1433	1442	1451	1460
ACC CAT GAC TGA CGG AGT CTG AGA AAA GGC AGA AGA AGG GGC GCA CAA GGG CAC					
T H D					

FIGURE 1D

1469 1478 1487 1496 1505 1514
CTT CTC CCT TGA GGC TGC CCT GCC CAC GTG GGA TTC ACA GGG GCC TGA GTA GGG

1523 1532 1541 1550 1559 1568
CCC GGG GAA GCA GAG CCC TAA GGG ATT AAG GCT CAG ACA CCT CTG AGA GCA GGT

1577 1586 1595 1604 1613 1622
GGG CAC TGG CTG GGT ACG GTG CCC TCC ACA GGA CTC TCC CTA CTG CCT GAG CAA

1631 1640 1649 1658 1667 1676
ACC TGA GGC CTC CCG GCA GAC CCA CCC ACC CCC TGG GGC TGC TCA GCC TCA GGC

1685 1694 1703 1712 1721 1730
ACG GAC AGG GCA CAT GAT ACC AAC TGC TGC CCA CTA CAG CAC GCC GCA CCG GAG

1739 1748 1757 1766 1775 1784
CAC GGC ACC GAG GGA GCC GCC ACA CGG TCA CCT GCA AGG ACG TCA CGG GCC CCT

FIGURE 1E

1793 1802 1811 1820 1829 1838
CTA AAG GAT TCG TGG TGC TCA TCC CCA AGC TTC AGA GAC CCT TTG GGG TTC CAC

1847 1856 1865 1874 1883 1892
ACT TCA CGT GGA CTG AGG TAG ACC CTG CAT GAA GAT GAA ATT ATA GGG AGG ACG

1901 1910 1919 1928 1937 1946
CTC CTT CCC TCC CCT CCT AGA GGA GAG GAA AGG GAG TCA TTA ACA ACT AGG GGG

1955 1964 1973 1982 3 '
TTG GGT AGG ATT CCT AGG TAT GGG GAA GAG TTT TGG

FIGURE 1F

1	M	L	L	P	W	A	T	S	A	P	G	L	A	W	G	P	L	V	L	G	L	F	G	L	L	A	A	S	Q	P	SEQ ID NO1
1	M	L	L	P	W	A	T	S	A	P	G	L	A	W	G	P	L	V	L	G	L	F	G	L	L	A	A	S	Q	P	g339762
1	M	R	L	P	R	A	S	S	P	C	G	L	A	W	G	P	L	L	L	G	L	S	G	L	L	V	A	S	Q	P	g600223
31	Q	A	V	P	P	Y	A	S	E	N	Q	T	C	R	D	Q	E	K	E	Y	Y	E	P	Q	H	R	I	C	C	S	SEQ ID NO1
31	Q	A	V	P	P	Y	A	S	E	N	Q	T	C	R	D	Q	E	K	E	Y	Y	E	P	Q	H	R	I	C	C	S	g339762
31	Q	L	V	P	P	Y	R	I	E	N	Q	T	C	W	D	Q	D	K	E	Y	Y	E	P	M	H	D	V	C	C	S	g600223
61	R	C	P	P	G	T	Y	V	S	A	K	C	S	R	I	R	D	T	V	C	A	T	C	A	E	N	S	Y	N	E	SEQ ID NO1
61	R	C	P	P	G	T	Y	V	S	A	K	C	S	R	I	R	D	T	V	C	A	T	C	A	E	N	S	Y	N	E	g339762
61	R	C	P	P	G	E	F	V	F	A	V	C	S	R	S	Q	D	T	V	C	K	T	C	P	H	N	S	Y	N	E	g600223
91	H	W	N	Y	L	T	I	C	Q	L	C	R	P	C	D	P	V	M	G	L	E	E	I	A	P	C	T	S	K	R	SEQ ID NO1
91	H	W	N	Y	L	T	I	C	Q	L	C	R	P	C	D	P	V	M	G	L	E	E	I	A	P	C	T	S	K	R	g339762
91	H	W	N	H	L	S	T	C	Q	L	C	R	P	C	D	I	V	L	G	F	E	E	V	A	P	C	T	S	D	R	g600223
121	K	T	Q	C	R	C	Q	P	G	M	F	C	A	A	W	A	L	E	C	T	H	C	E	L	-	-	L	S	D	C	SEQ ID NO1
121	K	T	Q	C	R	C	Q	P	G	M	F	C	A	A	W	A	L	E	C	T	H	C	E	L	-	-	L	S	D	C	g339762
121	K	A	E	C	R	C	Q	P	G	M	S	C	V	Y	L	D	N	E	C	V	H	C	E	E	E	R	L	V	L	C	g600223
149	P	P	G	T	E	A	E	L	K	D	E	V	G	K	G	N	N	H	C	V	P	C	K	A	G	H	F	Q	N	T	SEQ ID NO1
149	P	P	G	T	E	A	E	L	K	D	E	V	G	K	G	N	N	H	C	V	P	C	K	A	G	H	F	Q	N	T	g339762
151	Q	P	G	T	E	A	E	V	T	D	E	I	M	D	T	D	V	N	C	V	P	C	K	P	G	H	F	Q	N	T	g600223

FIGURE 2A

179	S	S	P	S	A	R	C	Q	P	H	T	R	C	E	N	Q	G	L	V	E	A	A	P	G	T	A	Q	S	D	T	SEQ ID NO1
179	S	S	P	S	A	R	C	Q	P	H	T	R	C	E	N	Q	G	L	V	E	A	A	P	G	T	A	Q	S	D	T	g339762
181	S	S	P	R	A	R	C	Q	P	H	T	R	C	E	I	Q	G	L	V	E	A	A	P	G	T	S	Y	S	D	T	g600223
209	T	C	K	N	P	L	E	P	L	P	P	E	M	S	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SEQ ID NO1
209	T	C	K	N	P	L	E	P	L	P	P	E	M	S	G	T	M	L	M	L	A	V	L	L	P	L	A	F	F	L	g339762
211	I	C	K	N	P	P	E	P	-	-	-	-	-	-	G	A	M	L	L	A	I	L	L	S	L	V	L	F	L	g600223	
224	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	L	L	K	R	R	P	Q	G	-	-	-	-	-	SEQ ID NO1
239	L	L	A	T	V	F	S	C	I	W	K	S	H	P	S	L	C	R	K	L	G	S	L	L	K	R	R	P	Q	G	g339762
235	L	F	T	T	V	L	A	C	A	W	M	R	H	P	S	L	C	R	K	L	G	T	L	L	K	R	H	P	E	G	g600223
233	E	G	P	N	P	V	A	G	S	W	E	P	P	K	A	H	P	Y	F	P	D	L	V	Q	P	L	L	P	I	S	SEQ ID NO1
269	E	G	P	N	P	V	A	G	S	W	E	P	P	K	A	H	P	Y	F	P	D	L	V	Q	P	L	L	P	I	S	g339762
265	E	E	S	P	P	C	P	A	-	-	-	P	R	A	D	P	H	F	P	D	L	A	E	P	L	L	P	M	S	g600223	
263	G	D	V	S	P	V	S	T	G	L	P	A	A	P	V	L	E	A	G	V	P	Q	Q	Q	S	P	L	D	L	T	SEQ ID NO1
299	G	D	V	S	P	V	S	T	G	L	P	A	A	P	V	L	E	A	G	V	P	Q	Q	Q	S	P	L	D	L	T	g339762
291	G	D	L	S	P	S	P	A	G	P	P	T	A	P	S	L	E	V	V	L	Q	Q	Q	S	P	L	V	Q	A	g600223	
293	R	E	P	Q	L	E	P	G	E	Q	S	Q	V	A	H	G	T	N	G	I	H	V	T	G	G	S	M	T	I	T	SEQ ID NO1
329	R	E	P	Q	L	E	P	G	E	Q	S	Q	V	A	H	G	T	N	G	I	H	V	T	G	G	S	M	T	I	T	g339762
321	R	E	L	E	A	E	P	G	E	H	G	Q	V	A	H	G	A	N	G	I	H	V	T	G	G	S	V	T	V	T	g600223

FIGURE 2B

323	G	N	I	Y	I	Y	N	G	P	V	L	G	G	P	P	G	P	G	D	L	P	A	T	P	E	P	P	Y	P	I	SEQ ID NO1
359	G	N	I	Y	I	Y	N	G	P	V	L	G	G	P	P	G	P	G	D	L	P	A	T	P	E	P	P	Y	P	I	g339762
351	G	N	I	Y	I	Y	N	G	P	V	L	G	G	T	R	G	P	G	D	P	P	A	P	P	E	P	P	Y	P	T	g600223
353	P	E	E	G	D	P	G	P	P	G	L	S	T	P	H	Q	E	D	G	K	A	W	H	L	A	E	T	E	H	-	SEQ ID NO1
389	P	E	E	G	D	P	G	P	P	G	L	S	T	P	H	Q	E	D	G	K	A	W	H	L	A	E	T	E	H	-	g339762
381	P	E	E	G	A	P	G	P	S	E	L	S	T	P	Y	Q	E	D	G	K	A	W	H	L	A	E	T	E	T	L	g600223
382	-	C	G	A	T	P	S	N	R	G	P	R	N	Q	F	I	T	H	D												SEQ ID NO1
418	-	C	G	A	T	P	S	N	R	G	P	R	N	Q	F	I	T	H	D												g339762
411	G	C	Q	D	L																										g600223

FIGURE 2C

10/10

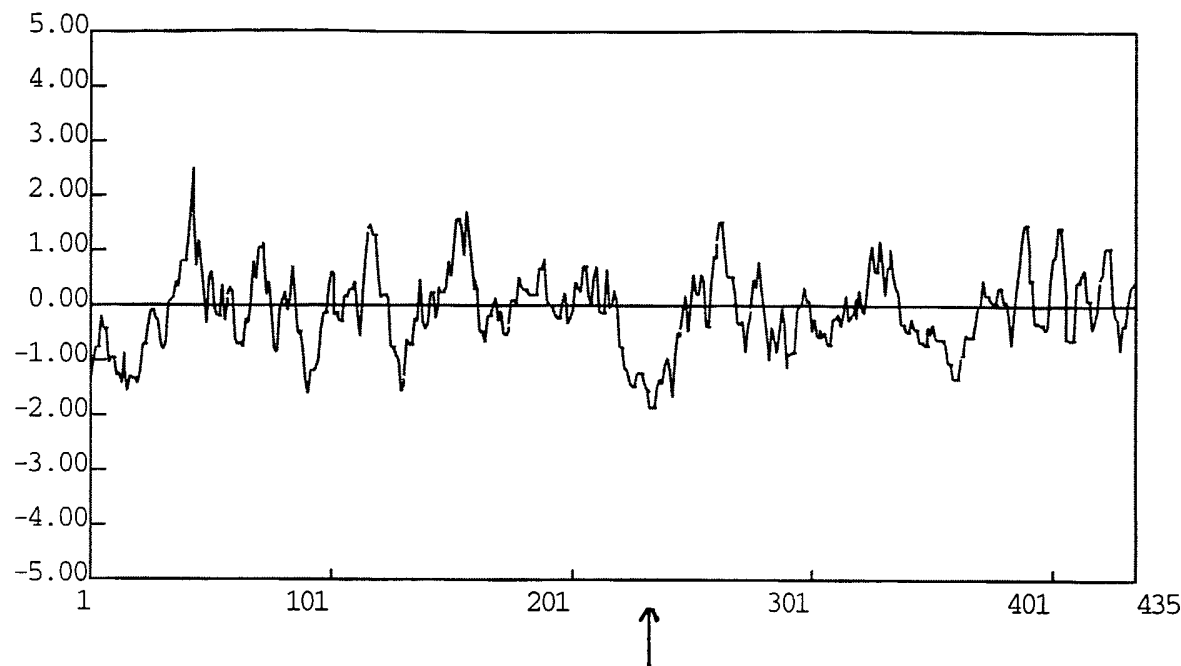


FIGURE 3A

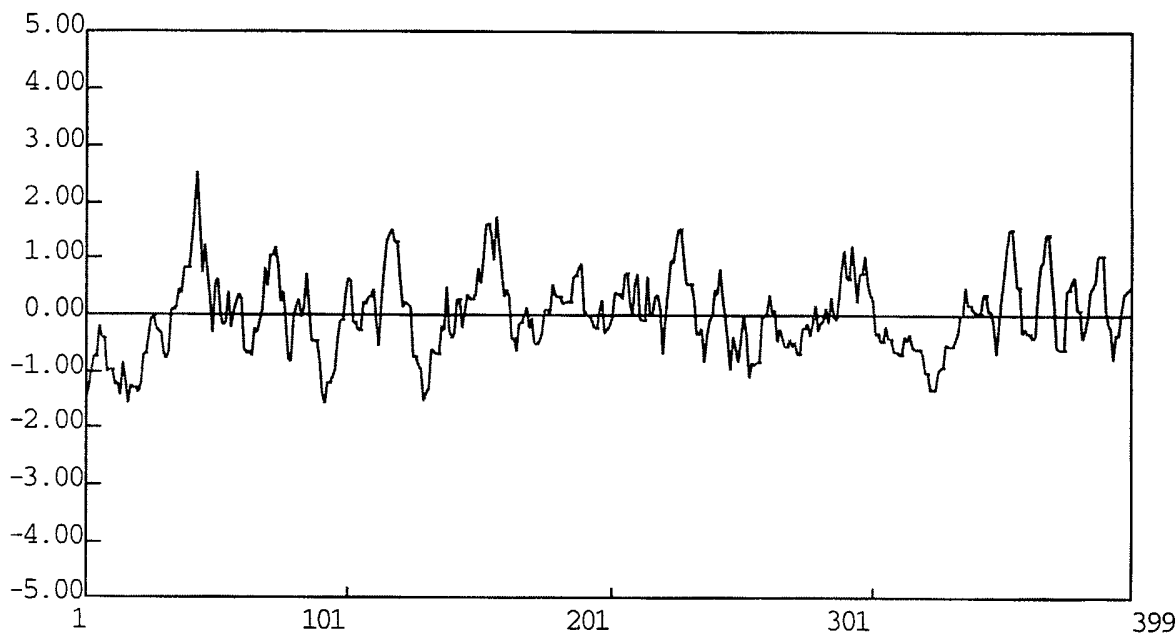


FIGURE 3B

<110> INCYTE GENOMICS, INC.
 LAL, Preeti G.
 WARREN, Bridget A.

<120> TNF RECEPTOR 2 RELATED PROTEIN VARIANT

<130> PC-0050 PCT

<140> To Be Assigned

<141> Herewith

<150> 09/917,372

<151> 2001-07-27

<160> 20

<170> PERL Program

<210> 1

<211> 399

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7497867CD1

<400> 1

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

Gln	Pro	Leu	Leu	Pro	Ile	Ser	Gly	Asp	Val	Ser	Pro	Val	Ser	Thr
				260					265					270
Gly	Leu	Pro	Ala	Ala	Pro	Val	Leu	Glu	Ala	Gly	Val	Pro	Gln	Gln
				275					280					285
Gln	Ser	Pro	Leu	Asp	Leu	Thr	Arg	Glu	Pro	Gln	Leu	Glu	Pro	Gly
				290					295					300
Glu	Gln	Ser	Gln	Val	Ala	His	Gly	Thr	Asn	Gly	Ile	His	Val	Thr
				305					310					315
Gly	Gly	Ser	Met	Thr	Ile	Thr	Gly	Asn	Ile	Tyr	Ile	Tyr	Asn	Gly
				320					325					330
Pro	Val	Leu	Gly	Gly	Pro	Pro	Gly	Pro	Gly	Asp	Leu	Pro	Ala	Thr
				335					340					345
Pro	Glu	Pro	Pro	Tyr	Pro	Ile	Pro	Glu	Glu	Gly	Asp	Pro	Gly	Pro
				350					355					360
Pro	Gly	Leu	Ser	Thr	Pro	His	Gln	Glu	Asp	Gly	Lys	Ala	Trp	His
				365					370					375
Leu	Ala	Glu	Thr	Glu	His	Cys	Gly	Ala	Thr	Pro	Ser	Asn	Arg	Gly
				380					385					390
Pro	Arg	Asn	Gln	Phe	Ile	Thr	His	Asp						
				395										

<210> 2

<211> 1982

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7497867CB1

<400> 2

```

gccccgcggc cagctcgcct cactcccact tcctgagctc cgccatggga gccctggagg 60
cccgccctgg ccgctcccgg ccctgggggtg cacatcggcc ctgagtccecg tcccaggctc 120
tgggctcggg cagccgcggc caccgctgcc caggacgtcg ggccctcctgc cttcctccca 180
ggccccccacg ttgctggccg cctggccgag tggccgccat gctcctgect tggggccact 240
ctgcccccg gctggcctgg gggcctctgg tgctgggcct cttcgggctc ctggcagcat 300
cgcagcccca ggcgggtgcct ccatatgcgt cggagaacca gacctgcagg gaccaggaaa 360
aggaatacta tgagccccag caccgcatct gctgctcccc ctgcccccca ggcacctatg 420
tctcagctaa atgtagccgc atccgggaca cagtttgtgc cacatgtgcc gagaattcct 480
acaacgagca ctggaactac ctgaccatct gccagctgtg ccgcccctgt gaccagtgta 540
tgggcctcga ggagattgcc ccctgcacaa gcaaaccgaa gacctagtgc cgctgccagc 600
cggaatggt ctgtgctgcc tgggcccctc agtgtagaca ctgcgagcta ctttctgact 660
gcccgcctgg cactgaagcc gagctcaaag atgaagtgg gaagggtaac aaccactgcg 720
tcccctgcaa ggcagggcac ttccagaata cctcctcccc cagcgccgcg tgccagcccc 780
acaccaggtg tgagaaccaa ggtctggtgg aggcagctcc aggcactgcc cagtccgaca 840
caacctgcaa aatccatta gagccactgc cccagagat gtcaggatcg ctgctcaaga 900
ggcgctccga gggagaggga cccaatcctg tagctggaag ctgggagcct ccgaaggccc 960
atccatactt ccctgacttg gtacagccat ttctggagat gtttcccag 1020
tatccactgg gctccccgca gcccagttt tggaggcagg ggtgccgcaa cagcagagtc 1080
ctctggacct gaccagggag ccgcagttgg aacccgggga gcagagccag gtggcccacg 1140
gtaccaatgg cattcatgtc accggcgggt ctatgactat cactggcaac atctacatct 1200
acaatggacc agtactggg ggaccaccgg gtccctggaga cctcccagct acccccgaac 1260
ctccataccc cattcccgaa gagggggacc ctggccctcc cgggctctct acaccccacc 1320
aggaagatgg caaggcttgg cacctagcgg agacagagca ctgtggtgcc acaccctcta 1380
acagggggccc aaggaaccaa tttatcacc atgactgact gactctgaga aaaggcagaa 1440
gaagggggggc acaagggcac cttctccctt gaggctgcc tgcccacgtg ggattcacag 1500
gggcctgagt agggcccggg gaagcagagc cctaagggat taaggctcag acacctctga 1560
gagcaggtgg gcactggctg ggtacgggtc cctccacagg actctcccta ctgcctgagc 1620
aaacctgagg cctcccggca gaccaccca cccctgggg ctgctcagcc tcaggcacgg 1680
acagggcaca tgataccaac tgctgccac tacagcacgc cgcaccggag cagggcaccg 1740
agggagccgc cacacggtca cctgcaagga cgtcacgggc ccctctaaag gattcgtggt 1800
gtcatcccc aagcttcaga gaccctttgg ggttccacac ttcacgtgga ctgaggtaga 1860
ccctgcatga agatgaaatt ataggaggga cgctccttcc ctcctctct agaggagagg 1920

```

aaagggagtc attaacaact aggggggttg gtaggattcc taggtatggg gaagagtttt 1980
gg 1982

<210> 3
<211> 392
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8113313H1

<400> 3
gccccgcggc cagctcgctc cactcccact tcctgagctc cgccatggga gccctggagg 60
cccggcctgg ccgctcccgg ccctggggtg cacatcggcc ctgagtcocg tcccaggctc 120
tgggctcggg cagccggcgc caccgctgcc caggacgtcg ggccctcctgc ctccctccca 180
ggccccacg ttgctggccg cctggccgag tggccgccat gctcctgcct tgggccacct 240
ctgcccccg cctggcctgg gggcctctgg tgctgggcct ctccgggctc ctggcagcat 300
cgagcccca ggcggtgcct ccatatgcgt cggagaacca gacctgcagg gaccaggaaa 360
aggaatacta tgagccccag caccgcatct gc 392

<210> 4
<211> 526
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8235763H1

<400> 4
ctgagctccg ccatgggagc cctggaggcc cggcctggcc gctcccggcc ctgggggtgca 60
catcgccctt gactcccgct ccaggctctg ggctcgggca gccgccgcca ccgctgcca 120
ggacgtcggg cctcctgcct tcctcccagg ccccccacgtt gctggccgcc tggccgagtg 180
gccgccatgc tcctgccttg ggccacctct gcccccgcc tggcctgggg gcctctggtg 240
ctgggcctct tcgggctcct ggcagcatcg cagccccagg cggctgcctcc atatgcgtcg 300
gagaaccaga cctgcaggga ccaggaaaag gaatactatg agccccagca ccgcatctgc 360
tgctcccgtt gcccgccagg cactatgtc tcagctaaat gtagccgcat ccgggacaca 420
gtttgtgcca catgtgccga gaattcctac aacgagcact ggaactacct gaccatctgc 480
cagctgtgcc gccctgtga cccagtgtat ggcctcgagg agattg 526

<210> 5
<211> 436
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4048821H1

<400> 5
gaactacctg accatctgcc agctgtgccg cccctgtgac ccagtgatgg gcctcgagga 60
gattgcccc tgcacaagca aacggaagac ccagtgccgc tgccagccgg gaatgttctg 120
tgctgcctgg gccctcagat gtacacactg cgagctactt tctgactgcc cgcctggcac 180
tgaagccgag ctcaaagatg aagttgggaa gggtacaac cactgcgtcc cctgcaaggc 240
agggcacttc cagaatacct cctccccag cgctcgctgc cagccccaca ccagggtgtga 300
gtgatggggt tcgaggagat tgccccctgc acaagcaaac ggaagacca gtgccgctgc 360
cagccgggaa tgttctgtgc tgctggggcc ctcgagtgtg cacactgcga gctactttct 420
gactgcccgc ctggca 436

<210> 6
<211> 135
<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2105134H1

<400> 6

```
tgatgggcct cgaggagatt gccccctgca caagcaaacg gaagaccag tgccgctgcc 60
agccgggaat gttctgtgct gcctgggccc tcgagtgtac aactgcgag ctactttctg 120
actgcccgcc tggca                                     135
```

<210> 7

<211> 651

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7716364H1

<400> 7

```
gggaatgggg tataggaggt tcgggggtag ctagggaggt ctccaggacc cggtgggtccc 60
cccagtactg gtccattgta gatgtagatg ttgccagtga tagtcataga cccgccggtg 120
acatgaatgc cattggtacc gtgggccacc tggctctgct ccccggttc caactgcggc 180
tccctgggtca ggtccagagg actctgctgt tgcggcacc ctgcctccaa aactggggct 240
gcggggagcc cagtggatac tggggaaaaca tctccagaaa tgggtagcag tggctgtacc 300
aagtcaggga agtatggatg ggccttcgga ggctcccagc ttccagctac aggattgggt 360
ccctctccct gcggacgcct cttgagcagc gatcctgaca tctctggggg cagtggctct 420
aatggatttt tgcaggttgt gtcggactgg gcagtgcctg gagctgcctc caccagacct 480
tggtttctac acctggtgtg gggctggcag cagggcgctg ggggaggagg tattctggaa 540
gtgccctgcc ttgcagggga cgcagtgggt gttacccttc ccaacttcat ctttgagctc 600
ggctacagtg gcaggcgggc agtcagaaaag tagctcgcag gtgtacactc g 651
```

<210> 8

<211> 574

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8234468H1

<400> 8

```
ctgggtggagg cagctccagg cactgcccag tccgacacaa cctgcaaaaa tccattagag 60
ccactgcccc cagagatgtc aggatcgctg ctcaagaggc gtccgcaggg agagggacct 120
aatcctgtag ctggaagctg ggagcctccg aaggccccatc catacttccc tgacttggtg 180
cagccactgc taccatttc tggagatgtt tccccagtat ccaactgggt ccccgagcc 240
ccagttttgg aggcaggggt gccgcaacag cagagtccctc tggacctgac caggagaccg 300
cagttggaac ccggggagca gagccagggtg gcccacggta ccaatggcat tcatgtcacc 360
ggcgggtcta tgactatcac tggcaacatc tacatctaca atggaccagt actgcgggga 420
ccaccgggtc ctggagacct cccagctacc cccgaacctc catacccat tcccgaagag 480
ggggaccctg gccctcccgg gctctctaaa ccccaccagg aagatggcaa ggcttgggac 540
ctagcggaga cagagcactg tgggtgcacac cctc 574
```

<210> 9

<211> 425

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7716340H1

```

<400> 9
gggaatgggg tatggaggtt cgggggtagc tgggaggtct ccaggaccg gtgggtcccc 60
cagtactggt ccattgtaga tgtagatggt gccagtgata gtcataagacc cgccgggtgac 120
atgaatgcca ttggtaccgt gggccacctg gctctgctcc ccgggttcca actgagggtc 180
cctgggtcagg tccagaggac tctgctgttg cggcaccctt gcctccaaaa ctgggggtgc 240
ggggagccca gtggatactg gggaaacatc tccagaaatg ggtagcagtg gctgtacca 300
gtcaggaag tatggatggg ccttcggagg ctcccagctt ccagctacag gattgggtcc 360
ctctcctgc ggacgcctct tgagcagcga tcctgacatc tctgggggca gtgggtctaa 420
tggat 425

```

```

<210> 10
<211> 219
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 697459H1

```

```

<220>
<221> unsure
<222> 76, 131
<223> a, t, c, g, or other

```

```

<400> 10
cacgcgtccg atcactggca acatctacat ctacaatgga ccagtactgg ggggaccacc 60
gggtcctgga gacctnccag ctacccccga acctccatac cccattcccg aagaggggga 120
ccctggccct nccgggtctt ctacacccca ccaggaagat ggcaaggctt ggcacctagc 180
ggagacagag cactgtggtg ccacaccctc taacagggg 219

```

```

<210> 11
<211> 279
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3321983H1

```

```

<400> 11
gaccagtact ggggggacca ccgggtcctg gagacctccc agctaccccc gaacctccat 60
acccatttcc cgaagagggg gacctgggcc ctcccgggct ctctacaccc caccaggaag 120
atggcaaggc ttggcaccta gcggagacag agcactgtgg tgccacaccc tctaacaggg 180
gccaaggaa ccaatttatc acccatgact gactgagtct gagaaaaggc agaagaaggg 240
gggcacaagg gcaccttctc cttgaggct gccctgccc 279

```

```

<210> 12
<211> 862
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 8576918T1

```

```

<400> 12
ttcccctata cctaggaatc ctaccaaac ccctaagttt gttaatgact ccctttcctc 60
tcctctagga ggggagggaa ggagcgtcct ccctataatt tcattctcat gcagggtcta 120
cctcagtcca cgtgaagtgt ggaaccccaa aggggtctctg aagcttgggg atgagcacca 180
cgaatccttt aaggggcccg tgacgtcctt gcaggtgacc gtgtggcggc tccctcggtt 240
gccgtgctcc ggtgcggcgt gctgtagtgg gcagcagttg gtatcatgtg ccctgtccgt 300
gcctgaggct gagcagcccc aggggtgggt gggctctgccg ggaggcctca ggtttgccta 360
ggcagtaggg agagtctgtg ggagggcacc gtaccagacc agtgcccacc tgctctcaga 420

```

```

ggtgtctgag ccttaatccc ttagggctct gcttccccgg gccctactca ggccccctgtg 480
aatccacagt gggcagggca gcctcaaggg agaaggtgcc cttgtgcccc ccttcttctg 540
ccttttctca gactcagtca gtcatgggtg ataaattggt tccttggggc cctgttagag 600
ggtgtggcac cacagttgct ctgtctccgc taggtgccaa agccttgccc atcttctctg 660
tggggtgtag agagcccggg aggccagggt ccctctcttc ggggaatggg gtatggaagg 720
gtccggggta gcctgggagg tctccaggac ccggtggtct cccccagtac ttggtcctt 780
gttgaatttc gacttttgcc cgtgattgtt cttgaaccgc cggtagctta ttgccttggg 840
taccgggggc cactggctct gc 862

```

<210> 13

<211> 206

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc_feature

<223> Incyte ID No: 700302531H1

<400> 13

```

cctgatccag gccagggagc tggaggctga gcctggggaa catggccagg tggccacagg 60
tgcgaatggc attcacgtga ccggaggctc tgtgactgtc accggcaata tctacatata 120
caatgggcca gtgctggggg gaacacgggg ccctggagac cctccagctc cccctgagcc 180
tccatacccg actcccgaag agggag 206

```

<210> 14

<211> 548

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc_feature

<223> Incyte ID No: 702152066H1

<400> 14

```

gcatcaagtc tgctgtctcc gctgtccccc aggcaagttt gtccatgctg tctgcagccc 60
cagccaagac acggttttgc agacttgcc ccataattcc tacaacgaac actggaacca 120
tctcttctcc tgccagctgt gccgccctg tgactctgtg ctgggcttcg aggagattgc 180
cccctgcacc agcgatcgga aacccgagtg ccgctgcaag ccggggatgt cctgcgtgta 240
tttggacaat gagtgtgtgc actgtgagga ggagcggtt gtactctgcc ggccctggcac 300
agaagctgag gtcacagatg aaattatgga tactgaagtc aactgtgtcc cctgtaagcc 360
aggacacttc cagaacacgt cctccccag agcccgctgt caaccccaca ccagggtgta 420
gagccagggc ctgggtggagg cagcttcagg tactctgtac tctgacacca tctgtaaaaa 480
tccacccgag gcagcaggaa caatgctgct actagccatc ctgctgtcgc tgggtcttctt 540
tctgcttt 548

```

<210> 15

<211> 471

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc_feature

<223> Incyte ID No: 702022948H1

<400> 15

```

agacacgggt tgcaagactt gcctcaataa ttcctacaac gaacactgga accatctctt 60
ctcctgccag ctgtgccgcc ctgtgactct gtgctgggct tcgaggagat tgccccctgc 120
accagcgatc ggaacccgga gtgccgctgc aagccgggga tgtcctgcgt gtatttggac 180
aatgagtgtg tgcaactgtg ggaggagcgg cttgtactct gacggcctgg cacagaagct 240
gaggtcacag atgaaattat ggatactgaa gtcaactgtg tccctgtaa gccaggacac 300
ttccagaaca cgtcctcccc cagagcccg tgtcaacccc acaccagggtg tgagagccag 360
ggcctggtgg aggcagcttc aggtacctcg tactctgaca acatctgtaa aaatccaacc 420
gaggaagaag gaacaatgct gctactagcc atcctgctgt cgctggtctt c 471

```

<210> 16
 <211> 371
 <212> DNA
 <213> *Canis familiaris*

<220>
 <221> misc_feature
 <223> Incyte ID No: 702245091H1

<400> 16
 ccagactccc agcaccacca agtctggagg aagagatgct acagcagcag agtcctctca 60
 gccaggccag agagctggag cctgagctcc cagaacaagg ccagggtggcc caggtacca 120
 atggcattca cgtcaccggc gggctctgtga ctgtgactgg caacatctac atctacaatg 180
 ggccagtact ggggggagca cgaggccctg gagaccccc tgcttcccca gagcctccat 240
 accccacccc tgaagagggg gcccctggcc ctccctgggt ctctacaccc taccaggagg 300
 atggcaaagc ttggcacctg gctgaaacag agaccctggg gtgccatgcc ctctaacagg 360
 gacaaggacc c 371

<210> 17
 <211> 618
 <212> DNA
 <213> *Macaca fascicularis*

<220>
 <221> misc_feature
 <223> Incyte ID No: 703193780J1

<400> 17
 atttgcatgt ggttacaata taaatacact cgccattttc ctccccttcc aaaactcttc 60
 cccacaccta ggaatcctac ccaacccta gttgttaatg actccccttc ctctcctcta 120
 ggaggggagg gaaggagcgt cctccctata atttcacct catgcagggt ctacctcagt 180
 ccatgtgaag tgtggaacc caaagggcct ctgaagcttg gggataagca ccacgaatcc 240
 tttagagggc ccggtgacgt ccttgagggt gaccatgtgg cggtccctc agtgccgtgc 300
 tccggtgceg cgtgccatag tgggcagcag ttggtatcat gtgccctgtc cgtgcccgag 360
 gctgagcagc cccagtgggt gggcggtct gccaggaggc ctccaggttg ctccaggcagt 420
 agggagagtc ctgtggagg catcgtaccc agccagtgcc caccctgctc cgggggtgtc 480
 tgagctttcg tcccttaggg ctctgcctcc ccgggcctac tcaggccccc tgtgaatccc 540
 acatgggcag ggtggcctca aggagaaggt gccctgttgc ctccaatctt ctggcttttc 600
 tcagacgcca tcagtcac 618

<210> 18
 <211> 536
 <212> DNA
 <213> *Macaca fascicularis*

<220>
 <221> misc_feature
 <223> Incyte ID No: 703678967J1

<400> 18
 tattttgcatg tggttacaat ataaatacac tcgccatttt cctccccttc caaaactctt 60
 cccacacct agaatccta ccaaccct agttgttaat gactcccctt cctctcctct 120
 aggaggggag ggaaggagcg tcctccctat aatttcaccc tcatgcaggg tctacctcag 180
 tccatgtgaa gtgtggaacc ccaaagggcc tctgaagctt ggggataagc accactaatc 240
 ctttagaggg cccggtgacg tccttgaggg tgaccatgtg gcggctccct cagtgccgtg 300
 ctccggtgcg gcgtgccata gtgggcagca gttgttatca tgtgccctgt ccgtgcccga 360
 ggctgagcag ccccagtggg tgggcgggtc tgccaggagg cctcaggttt gctcaggcag 420
 tagggagagt cctgtggagg gcacgtacc cagccagtgc ccacctgtc tcgggggtgt 480
 ctgagctttc gtcccttagg gctctgcctc cccgggccct actcaggccc cctgtg 536

<210> 19
 <211> 435
 <212> PRT

Thr Pro Ser Asn Arg Gly Pro Arg Asn Gln Phe Ile Thr His Asp
 425 430 435

<210> 20

<211> 415

<212> PRT

<213> Mus musculus

<220>

<221> misc_feature

<223> Incyte ID No: g600223

<400> 20

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

				365					370					375
Pro	Pro	Tyr	Pro	Thr	Pro	Glu	Glu	Gly	Ala	Pro	Gly	Pro	Ser	Glu
				380					385					390
Leu	Ser	Thr	Pro	Tyr	Gln	Glu	Asp	Gly	Lys	Ala	Trp	His	Leu	Ala
				395					400					405
Glu	Thr	Glu	Thr	Leu	Gly	Cys	Gln	Asp	Leu					
				410					415					