ABSTRACT

The present invention relates to a composition comprising a plurality of cDNAs which are differentially expressed in prostate cancer and which may be used in their entirety or in part as to diagnose, to stage to treat or to monitor the treatment of a subject with prostate cancer.
PROSTATE CANCER MARKERS

[0001] This application is a continuation of U.S. application Ser. No. 09/919,172, filed Jul. 30, 2001, which further claims the benefit of U.S. Provisional Application Ser. No. 60/222,469, filed Jul. 28, 2000.

FIELD OF THE INVENTION

[0002] The present invention relates to a composition comprising a plurality of cDNAs which are differentially expressed in prostate cancer and which may be used entirely or in part to diagnose, to stage, to treat, or to monitor the progression or treatment of prostate cancer.

BACKGROUND OF THE INVENTION

[0003] Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are used to detect the expression of a specific gene and its variants. When an expression profile is examined, arrays provide a platform for examining which genes are tissue-specific, carrying out housekeeping functions, parts of a signaling cascade, or specifically related to a particular genetic predisposition, condition, disease, or disorder.

[0004] The potential application of gene expression profiling is particularly relevant to improving diagnosis, prognosis, and treatment of disease. For example, both the levels and sequences expressed in tissue from subjects with prostate cancer may be compared with the levels and sequences expressed in normal tissue.

[0005] Prostate cancer is a common malignancy in men over the age of 50, and the incidence increases with age. In the US, there are approximately 132,000 newly diagnosed cases of prostate cancer and more than 33,000 deaths from the disorder each year.

[0006] Once cancer cells arise in the prostate, they are stimulated by testosterone to a more rapid growth. Thus, removal of the testes can indirectly reduce both rapid growth and metastasis of the cancer. Over 95% of prostatic cancers are adenocarcinomas which originate in the prostatic acini. The remaining 5% are divided between squamous cell and transitional cell carcinomas, both of which arise in the prostatic ducts or other parts of the prostate gland.

[0007] As with most cancers, prostate cancer develops through a multistage progression ultimately resulting in an aggressive, metastatic phenotype. The initial step in tumor progression involves the hyperplasia of normal luminal and/or basal epithelial cells that become hyperplastic and evolve into early-stage tumors. The early-stage tumors are localized in the prostate but eventually may metastasize, particularly to the bone, brain, or lung. About 80% of these tumors remain responsive to androgen treatment, an important hormone controlling the growth of prostate epithelial cells. However, in its most advanced state, cancer growth becomes androgen-independent and there is currently no known treatment for this condition.

[0008] A primary diagnostic marker for prostate cancer is prostate specific antigen (PSA). PSA is a tissue-specific serine protease almost exclusively produced by prostatic epithelial cells. The quantity of PSA correlates with the number and volume of the prostatic epithelial cells, and consequently, the levels of PSA are an excellent indicator of abnormal prostate growth. Men with prostate cancer exhibit an early linear increase in PSA levels followed by an exponential increase prior to diagnosis. However, as PSA levels are also influenced by factors such as inflammation, androgen and other growth factors, some scientists maintain that changes in PSA levels are not useful in detecting individual cases of prostate cancer.

[0009] Current areas of cancer research provide additional prospects for markers as well as potential therapeutic targets for prostate cancer. Several growth factors have been shown to play a critical role in tumor development, growth, and progression. The growth factors Epidermal Growth Factor (EGF), Fibroblast Growth Factor (FGF), and Tumor Growth Factor alpha (TGFα) are important in the growth of normal as well as hyperproliferative prostate epithelial cells, particularly at early stages of tumor development and progression, and affect signaling pathways in these cells in various ways (Lin J et al. (1999) Cancer Res. 59:2891-2897; Putz T et al. (1999) Cancer Res 59:227-233). The TGF-β family of growth factors are generally expressed at increased levels in human cancers and the high expression levels in many cases correlates with advanced stages of malignancy and poor survival (Gold L I (1999) Crit Rev Oncog 10:303-360). Finally, there are human cell lines representing both the androgen-dependent stage of prostate cancer (LNCap) as well as the androgen-independent, hormone refractory stage of the disease (PC3 and DU-145) that have proved useful in studying gene expression patterns associated with the progression of prostate cancer, and the effects of cell treatments on these expressed genes (Chung T D (1999) Prostate 15:199-207).

[0010] The present invention provides for a composition comprising a plurality of cDNAs for use in detecting changes in expression of genes encoding proteins that are associated with prostate cancer. Such a composition can be employed for the diagnosis, prognosis or treatment of prostate cancer and related disorders correlated with differential gene expression. The present invention satisfies a need in the art in that it provides a set of differentially expressed genes which may be used entirely or in part to diagnose, to stage, to treat, or to monitor the progression or treatment of a subject with prostate cancer.

SUMMARY

[0011] The present invention provides a composition comprising a plurality of cDNAs and their complements which are differentially expressed in prostate adenocarcinomas and which are selected from SEQ ID Nos:1-13, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75, 76, 78-86, 88-90, 92-97, 99-101 as presented in the Sequence Listing. In one embodiment, each cDNA is differentially regulated in metastatic versus non-metastatic tissue samples, SEQ ID Nos:1-3, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75; in another embodiment, each cDNA is differentially regulated at all stages of the disease, SEQ ID Nos:76, 78-86, 88-90, 92-97, 99-101. In one aspect, the composition is immobilized on a substrate. In another aspect, the composition is used to diagnose the presence and stage of prostate cancer in a
subject. The invention also provides proteins encoded by the cDNAs and which are selected from SEQ ID NOs:4, 7, 9, 16, 20, 22, 29, 31, 33, 37, 39, 41, 46, 51, 54, 57, 66, 69, 74, 77, 87, 91, 98 as presented in the Sequence Listing.

[0012] The invention also provides a high throughput method to detect differential expression of one or more of the cDNAs of the composition. The method comprises hybridizing the substrate comprising the composition with the nucleic acids of a sample, thereby forming one or more hybridization complexes, detecting the hybridization complexes, and comparing the hybridization complexes with those of a standard, wherein differences in the size and signal intensity of each hybridization complex indicates differential expression of nucleic acids in the sample. In one aspect, the sample is from a subject with prostate cancer and differential expression determines an early, mid, and late stage of the disorder.

[0013] The invention further provides a high throughput method of screening a library or a plurality of molecules or compounds to identify a ligand. The method comprises combining the substrate comprising the composition with a library or a plurality of molecules or compounds under conditions to allow specific binding and detecting specific binding, thereby identifying a ligand. The library or a plurality of molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acid molecules, mimetics, peptides, transcription factors, repressors, and other regulatory proteins.

[0014] The invention still further provides an isolated cDNA encoding the protein comprising the amino acid sequence of SEQ ID NO:37. The invention also provides an isolated cDNA comprising SEQ ID NO:36 as presented in the Sequence Listing. The invention also provides a vector comprising the cDNA, a host cell comprising the vector, and a method for producing a protein comprising culturing the host cell under conditions for the expression of a protein and recovering the protein from the host cell culture. The invention additionally provides a method for purifying a ligand, the method comprising combining a cDNA of the invention with a sample under conditions which allow specific binding, recovering the bound cDNA, and separating the cDNA from the ligand, thereby obtaining purified ligand.

[0015] The present invention provides a purified protein encoded and produced by a cDNA of the invention. The invention also provides a high-throughput method for using a protein to screen a library or a plurality of molecules or compounds to identify a ligand. The method comprises combining the protein or a portion thereof with the library or a plurality of molecules or compounds under conditions to allow specific binding and detecting specific binding, thereby identifying a ligand which specifically binds the protein. A library or a plurality of molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acid molecules, mimetics, peptides, antigens, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds, and pharmaceutical agents. The invention further provides for using a protein to purify a ligand. The method comprises combining the protein or a portion thereof with a sample under conditions to allow specific binding, recovering the bound protein, and separating the protein from the ligand, thereby obtaining purified ligand. The invention still further provides a pharmaceutical composition comprising the protein. The invention yet still further provides a method for using the protein to produce an antibody. The method comprises immunizing an animal with the protein or an antigenically-effective epitope under conditions to elicit an antibody response, isolating animal antibodies, and screening the isolated antibodies with the protein to identify an antibody which specifically binds the protein. The invention yet still further provides a method for using the protein to purify antibodies which bind specifically to the protein.

DESCRIPTION OF THE SEQUENCE LISTING AND TABLES

[0016] A portion of the disclosure of this patent document contains material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

[0017] The Sequence Listing is a compilation of cDNAs obtained by sequencing and extension of clone inserts. Each sequence is identified by a sequence identification number (SEQ ID NO) and by the template number (TEMPLATE ID) from which it was obtained.

[0018] Table 1 shows the differential expression of cDNAs of the present invention in metastatic versus non-metastatic prostate adenocarcinoma. Column 1 shows the Clone ID of each sequence represented on a microarray. Columns 2-6 show differential expression in adenocarcinomas derived from prostate tissue relative to primary prostate epithelium. Differential expression values are presented as log2(normal tissue/adenocarcinoma). Negative values represent an increase in expression. Column 7 shows the t-test statistic used to evaluate markers specific to metastatic versus non-metastatic prostate adenocarcinoma.

[0019] Table 2 shows the differential expression of cDNAs of the present invention in prostate adenocarcinoma versus normal prostate. Column 1 shows the Clone ID of each sequence represented on a microarray. Columns 2-6 show differential expression in adenocarcinomas derived from prostate tissue relative to primary prostate epithelium. Differential expression values are presented as log2(normal tissue/adenocarcinoma). Negative values represent an increase in expression.

[0020] Table 3 shows the region within a gene template of each cDNA encompassed by a clone identified in Tables 1 and 2. Columns 1 and 2 show the SEQ ID NO: and Template ID, respectively. Column 3 shows the Clone ID and columns 4 and 5 show the first residue (Start) and last residue (Stop) encompassed by the clone on the template.

[0021] Table 4 lists the functional annotation of the cDNAs of the present invention. Columns 1 and 2 show the SEQ ID NO and Template ID, respectively. Columns 3, 4, and 5 show the GeneBank hit (GI Number), probability score (E-value), and functional annotation, respectively, as determined by BLAST analysis (version 1.4 using default parameters; Altschul (1993) J Mol Evol 36: 290-300; Altschul et al. (1990) J Mol Biol 215:403-410) of the cDNA against GenBank (release 117; National Center for Biotechnology Information (NCBI), Bethesda Md.).
Table 5 shows Pfam annotations of the cDNAs of the present invention. Columns 1 and 2 show the SEQ ID NO and Template ID, respectively. Columns 3, 4, and 5 show the first residue (Start), last residue (Stop), and reading frame, respectively, for the segment of the cDNA identified by Pfam analysis. Columns 6, 7, and 8 show the Pfam Hit, Pfam Annotation, and E-value, respectively, corresponding to the polypeptide domain of the protein or encoded by the cDNA segment.

Table 6 shows signal peptide and transmembrane regions predicted within the cDNAs of the present invention. Columns 1 and 2 show the SEQ ID NO and Template ID, respectively. Columns 3, 4, and 5 show the first residue (Start), last residue (Stop), and reading frame, respectively, for a segment of the cDNA, and column 6 identifies the polypeptide encoded by the segment as either a signal peptide (SP) or transmembrane (TM) domain.

DESCRIPTION OF THE INVENTION

Definitions

“Array” refers to an ordered arrangement of at least two cDNAs on a substrate. At least one of the cDNAs represents a control or standard sequence, and the other, a cDNA of diagnostic interest. The arrangement of from about two to about 40,000 cDNAs on the substrate assures that the size and signal intensity of each labeled hybridization complex formed between a cDNA and a sample nucleic acid is individually distinguishable.

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

A “composition” comprises at least two sequences selected from the Sequence Listing. “cDNA” refers to a chain of nucleotides, an isolated nucleotide molecule, nucleic acid molecule, or any fragment or complement thereof. It may have originated recombinantly or synthetically, be double-stranded or single-stranded, coding and/or noncoding, an exon with or without an intron from a genomic DNA molecule, and purified or combined with carbohydrate, lipids, protein or inorganic elements or substances. Preferably, the cDNA is from about 4000 to about 5000 nucleotides.

The phrase “cDNA encoding a protein” refers to a nucleic acid sequence that closely aligns with sequences which encode conserved regions, motifs or domains that were 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) which provides identity within the conserved region. Brenner et al. (1998; Proc Natl Acad Sci 95:6073-6078) who analyzed 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 et al., page 6076, column 2).

“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 quosine or of an analog such as hypoxanthine. These substitutions are well known in the art. Derivatization of a protein involves the replacement of a hydrogen by an acetyl, acyl, alkyl, amino, formyl, or morpholino group. Derivative molecules retain the biological activities of the naturally occurring molecules but may confer advantages such as longer lifespan or enhanced activity.

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

“Disorder” refers to conditions, diseases or syndromes associated with prostate cancer.

“Fragment” refers to a chain of consecutive nucleotides from about 200 to about 700 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. Nucleic acids and their ligands identified in this manner are useful as therapeutics to regulate replication, transcription or translation.

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’. The degree of complementarity and the use of nucleotide analogs affect the efficiency and stringency of hybridization reactions.

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

“Oligonucleotide” refers to 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. Substantially equivalent terms are amplimer, primer, and oligomer.

“Portion” refers to any part of a protein used for any purpose; but especially, to an epitope for the screening of ligands or for the production of antibodies.

“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 molecule 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 retains at least one biological or antigenic characteristic of a native protein. An “oli-
"gopeptide" 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.

[0040] “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.

[0041] “Sample” is used in its broadest sense as containing nucleic acids, proteins, antibodies, and the like. A sample may comprise a bodily fluid; 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 print; a fingerprint, buccal cells, skin, or hair; and the like.

[0042] “Specific binding” refers to a special and 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, the hydrophobic bonding along the backbone between two single stranded nucleic acids, or the binding between an epitope of a protein and an agonist, antagonist, or antibody.

[0043] “Similarity” 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) orBLAST2 (Altschul et al. (1997) Nucleic Acids Res 25:3389-3402). BLAST2 may be used in a standardized and reproducible way to insert gaps in one of the sequences in order to optimize alignment and to achieve a more meaningful comparison between them.

[0044] “Substrate” refers to any rigid or semi-rigid support to which cDNAs or proteins are bound and includes membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, capillaries or other tubing, plates, polymers, and microparticles with a variety of surface forms including wells, trenches, pins, channels and pores.

[0045] “Variant” refers to molecules that are recognized variations of a cDNA or a protein encoded by the cDNA. 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.

THE INVENTION

[0046] The present invention provides for a composition comprising a plurality of cDNAs or their complements, SEQ ID NO: 1-3, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75, 76, 78-86, 88-90, 92-97, 99-101, which may be used on a substrate to diagnose, to stage, to treat or to monitor the progression or treatment of prostate cancer. These cDNAs represent known and novel genes differentially expressed in cells from non-metastatic and metastatic prostate tumors. The composition may be used in its entirety or in part, as subsets of cDNAs differentially regulated between non-metastatic and metastatic prostate cancer, SEQ ID NOs: 1-3, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75, or of cDNAs differentially regulated at all stages of prostate cancer, SEQ ID NOs: 76, 78-86, 88-90, 92-97, 99-101. SEQ ID NOs: 24, 36, 47, 60, 82, 88, 89, 92, 93, and 95 represent novel cDNAs associated with prostate cancer. Since the novel cDNAs were identified solely by their differential expression, it is not essential to know a priori the name, structure, or function of the gene or it’s encoded protein. The usefulness of the novel cDNAs exist in their immediate value as diagnostics for prostate cancer.

[0047] Table 1 shows the differential expression of cDNAs of the present invention in metastatic versus non-metastatic prostate adenocarcinoma. Column 1 shows the Clone ID of each sequence represented on a microarray. Columns 2-6 show the differential expression in adenocarcinomas derived from prostate tissue relative to primary prostate epithelium. Differential expression values are presented as log2 of the absolute expression in normal prostate tissue +the absolute expression in prostate adenocarcinoma. Negative values represent an increase in expression. Column 7 shows the t-test statistic used to evaluate markers specific to metastatic versus non-metastatic prostate adenocarcinoma. All of the cDNAs in Table 1 show significant differential regulation in metastatic cancer relative to non-metastatic cancer. Further, expression profiles between the metastatic cancer lines show a high degree of correlation (>0.48), as do the expression profiles between the non-metastatic lines (0.64). However, the expression profiles between the metastatic and non-metastatic lines show significantly less correlation (<0.3).

[0048] Table 2 shows the differential expression of cDNAs of the present invention in prostate adenocarcinoma versus normal prostate. Column 1 shows the Clone ID of each sequence represented on a microarray. Columns 2-6 show differential expression in adenocarcinomas derived from prostate tissue relative to primary prostate epithelium. Differential expression values are presented as log2 (normal tissue -adenocarcinoma). Negative values represent an increase in expression. The expression profile for the cDNAs identified in Table 2 show high correlation between all tumor lines (>0.5).

[0049] SEQ ID NO: 36 is a novel sequence differentially regulated between metastatic and non-metastatic prostate tumors. SEQ ID NO: 36 encodes SEQ ID NO: 37 which is 193 amino acids in length.

[0050] The cDNAs of the invention define a differential expression pattern against which to compare the expression pattern of biopsied and/or in vitro treated tissues. Experimentally, differential expression of the cDNAs can be evaluated by methods including, but not limited to, differential display by spatial immobilization or by gel electrophoresis, genome mismatch scanning, representational discriminate analysis, clustering, transcript imaging and array technologies. These methods may be used alone or in combination.

[0051] The composition may be arranged on a substrate and hybridized with tumor tissues from subjects to identify those sequences which are differentially expressed in both prostate cancer and tumors derived from other tissues. This allows identification of those sequences of highest diagnos-
tic and potential therapeutic value. In one embodiment, an additional set of cDNAs, such as cDNAs encoding signaling molecules, are arranged on the substrate with the composition. Such combinations may be useful in the elucidation of pathways which are affected in a particular cancer or to identify new, coexpressed, candidate, therapeutic molecules.

[0052] In another embodiment, the composition can be used for large scale genetic or gene expression analysis of a large number of novel, nucleic acid molecules. These samples are prepared by methods well known in the art and are from mammalian cells or tissues which are in a certain stage of development; have been treated with a known molecule or compound, such as a cytokine, growth factor, a drug, and the like; or have been extracted or biopsied from a mammal with a known or unknown condition, disorder, or disease before or after treatment. The sample nucleic acid molecules are hybridized to the composition for the purpose of defining a novel gene profile associated with that developmental stage, treatment, or disorder.

[0053] cDNAs and Their Uses

[0054] cDNAs can be prepared by a variety of synthetic or enzymatic methods well known in the art. cDNAs can be synthesized, in whole or in part, using chemical methods well known in the art (Carathers et al. (1980) Nucleic Acids Symp. Ser. (7):215-233). Alternatively, cDNAs can be produced enzymatically or recombinantly, by in vitro or in vivo transcription.

[0055] Nucleotide analogs can be incorporated into cDNAs by methods well known in the art. The only requirement is that the incorporated analog must base pair with native purines or pyrimidines. For example, 2,6-diaminopurine can substitute for adenine and form stronger bonds with thymidine than those between adenine and thymidine. A weaker pair is formed when hypoxanthine is substituted for guanine and base pairs with cytosine. Alternatively, cDNAs can include nucleotides that have been derivatized chemically or enzymatically.

[0056] cDNAs can be synthesized on a substrate. Synthesis on the surface of a substrate may be accomplished using a chemical coupling procedure and a piezoelectric printing apparatus as described by Baldeschwieger et al. (PCT publication WO95/251116). Alternatively, the cDNAs can be synthesized on a substrate surface using a self-addressable electronic device that controls when reagents are added as described by Heller et al. (U.S. Pat. No. 5,605,662). cDNAs can be synthesized directly on a substrate by sequentially dispensing reagents for their synthesis on the substrate surface or by dispensing preformed DNA fragments to the substrate surface. Typical dispensers include a microcapillary delivering solution to the substrate with a robotic system to control the position of the microcapillary with respect to the substrate. There can be a multiplicity of dispensers so that reagents can be delivered to the reaction regions efficiently.

[0057] cDNAs can be immobilized on a substrate by covalent means such as by chemical bonding procedures or UV irradiation. In one method, a cDNA is bound to a glass surface which has been modified to contain epoxide or aldehyde groups. In another method, a cDNA is placed on a polylysine coated surface and UV cross-linked to it as described by Shalon et al. (WO95/35505). In yet another method, a cDNA is actively transported from a solution to a given position on a substrate by electrical means (Heller, supra). cDNAs do not have to be directly bound to the substrate, but rather can be bound to the substrate through a linker group. The linker groups are typically about 6 to 50 atoms long to provide exposure of the attached cDNA. Preferred linker groups include ethylene glycol oligomers, diarnines, diacids and the like. Reactive groups on the substrate surface react with a terminal group of the linker to bind the linker to the substrate. The other terminus of the linker is then bound to the cDNA. Alternatively, polynucleotides, plasmids or cells can be arranged on a filter. In the latter case, cells are lysed, proteins and cellular components degraded, and the DNA is coupled to the filter by UV cross-linking.

[0058] The cDNAs may be used for a variety of purposes. For example, the composition of the invention may be used on an array. The array, in turn, can be used in high-throughput methods for detecting a related polynucleotide in a sample, screening a plurality of molecules or compounds to identify a ligand, diagnosing prostate cancer, or inhibiting or inactivating a therapeutically relevant gene related to the cDNA.

[0059] When the cDNAs of the invention are employed on a microarray, the cDNAs are arranged in an ordered fashion so that each cDNA is present at a specified location. Because the cDNAs are at specified locations on the substrate, the hybridization patterns and intensities, which together create a unique expression profile, can be interpreted in terms of expression levels of particular genes and can be correlated with a particular metabolic process, condition, disorder, disease, stage of disease, or treatment. Hybridization

[0060] The cDNAs or fragments or complements thereof may be used in various hybridization technologies. The cDNAs may be labeled using a variety of reporter molecules by either PCR, recombinant, or enzymatic techniques. For example, a commercially available vector containing the cDNA is transcribed in the presence of an appropriate polymerase, such as T7 or SP6 polymerase, and at least one labeled nucleotide. Commercial kits are available for labeling and cleanup of such cDNAs. Radioactive (Amersham Pharmacia Biotech (API), Piscataway N.J.), fluorescent (Operon Technologies, Alameda Calif.), and chemiluminescent labeling (Promega, Madison Wis.) are well known in the art.

[0061] A cDNA may represent the complete coding region of an mRNA or be designed or derived from unique regions of the mRNA or genomic molecule, an intron, a 3' untranslated region, or from a conserved motif. The cDNA is at least 18 contiguous nucleotides in length and is usually single stranded. Such a cDNA may be used under hybridization conditions that allow binding only to an identical sequence, a naturally occurring molecule encoding the same protein, or an allelic variant. Discovery of related human and mammalian sequences may also be accomplished using a pool of degenerate cDNAs and appropriate hybridization conditions. Generally, a cDNA for use in Southern or northern hybridizations may be from about 400 to about 6000 nucleotides long. Such cDNAs have high binding specificity in solution-based or substrate-based hybridizations. An oligonucleotide, a fragment of the cDNA, may be used to detect a polynucleotide in a sample using PCR.

[0062] The stringency of hybridization is determined by G+C content of the cDNA, salt concentration, and tempera-
ture. In particular, stringency is increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature. Hybridization may be performed with buffers, such as 5x saline sodium citrate (SSC) with 1% sodium dodecyl sulfate (SDS) at 60°C, that permit the formation of a hybridization complex between nucleic acid sequences that contain some mismatches. Subsequent washes are performed with buffers such as 0.2x SSC with 0.1% SDS at either 45°C (medium stringency) or 65-68°C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acid molecules are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide may be added to the hybridization solution to reduce the temperature at which hybridization is performed. Background signals may be added 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 et al. (1997, Short Protocols in Molecular Biology, John Wiley & Sons, New York N.Y., Units 2.8-2.11, 3.18-3.19 and 4.6-4.9).

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. Protein binding may be confirmed by raising antibodies against the protein and adding the antibodies to the gel retardation assay where specific binding will cause a supershift in the 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.

The cDNA may be used to purify a ligand from a sample. A method for using a cDNA to purify a ligand would involve combining the cDNA or a fragment thereof with a sample under conditions to allow specific binding, recovering the bound cDNA, and using an appropriate agent to separate the cDNA from the purified ligand.

[0069] Protein Production and Uses

[0070] The full length cDNAs or fragment thereof may be used to produce purified proteins using recombinant DNA technologies described herein and taught in Ausubel et al. (supra; Units 16.1-16.62). One of the advantages of producing proteins by these procedures is the ability to obtain highly-enriched sources of the proteins thereby simplifying purification procedures.

[0071] The proteins may contain amino acid substitutions, deletions or insertions made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. Such substitutions may be conservative in nature when the substituted residue has structural or chemical properties similar to the original residue (e.g., replacement of leucine with isoleucine or valine) or they may be nonconservative when the replacement residue is radically different (e.g., a glycine replaced by a tryptophan). Computer programs included in LASERGENE software (DNASTAR, Madison Wis.), MACVECTOR software (Genetics Computer Group, Madison Wis.) and RasMol software (www.amass.edu/microbio/rasmol) may be used to help determine which and how many amino acid residues in a particular portion of the protein may be substituted, inserted, or deleted without abolishing biological or immunological activity.

[0072] Expression of Encoded Proteins

[0073] Expression of a particular cDNA may be accomplished by cloning the cDNA into a vector and transforming this vector into a host cell. The cloning vector used for the construction of cDNA libraries in the LIFSEQ databases may also be used for expression. Such vectors usually contain a promoter and a polylinker useful for cloning, priming, and transcription. An exemplary vector may also contain the promoter for β-galactosidase, an amino-terminal methionine and the subsequent seven amino acid residues of β-galactosidase. The vector may be transformed into competent E. coli cells. Induction of the isolated bacterial strain with isopropylthiogalactoside (IPTG) using standard meth-
ods will produce a fusion protein that contains an N terminal methionine, the first seven residues of β-galactosidase, about 15 residues of linker, and the protein encoded by the cDNA.

[0074] The cDNA may be shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotides containing cloning sites and fragments of DNA sufficient to hybridize to stretches at both ends of the cDNA may be chemically synthesized by standard methods. These primers may then be used to amplify the desired fragments by PCR. The fragments may be digested with appropriate restriction enzymes under standard conditions and isolated using gel electrophoresis. Alternatively, similar fragments are produced by digestion of the cDNA with appropriate restriction enzymes and filled in with chemically synthesized oligonucleotides. Fragments of the coding sequence from more than one gene may be ligated together and expressed.

[0075] Signal sequences that dictate secretion of soluble proteins are particularly desirable as component parts of a recombinant sequence. For example, a chimeric protein may be expressed that includes one or more additional purification-facilitating domains. Such domains include, but are not limited to, metal-chelating domains that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the 5'FLAGS extension/affinity purification system (Immunex, Seattle Wash.). The inclusion of a cleavable linker sequence such as ENTEROKINASEMAX (Invitrogen, San Diego Calif.) between the protein and the purification domain may also be used to recover the protein.

[0076] Suitable host cells may include, but are not limited to, mammalian cells such as Chinese Hamster Ovary (CHO) and human 293 cells, insect cells such as Sf9 cells, plant cells such as Nicotiana tabacum, yeast cells such as Saccharomyces cerevisiae, and bacteria such as E. coli. For each of these cell systems, a useful vector may also include an origin of replication and one or two selectable markers to allow selection in bacteria as well as in a transformed eukaryotic host. Vectors for use in eukaryotic host cells may require the addition of 3' poly(A) tail if the cDNA lacks poly(A).

[0077] Additionally, the vector may contain promoters or enhancers that increase gene expression. Many promoters are known and used in the art. Most promoters are host specific and exemplary promoters includes SV40 promoters for CHO cells; T7 promoters for bacterial hosts; viral promoters and enhancers for plant cells; and PGH promoters for yeast. Adenoviral vectors with the rous sarcoma virus enhancer or retroviral vectors with long terminal repeat promoters may be used to drive protein expression in mammalian cell lines. Once homogeneous cultures of recombinant cells are obtained, large quantities of secreted soluble protein may be recovered from the conditioned medium and analyzed using chromatographic methods well known in the art. An alternative method for the production of large amounts of secreted protein involves the transformation of mammalian embryos and the recovery of the recombinant protein from milk produced by transgenic cows, goats, sheep, and the like.

[0078] In addition to recombinant production, proteins or portions thereof may be produced manually, using solid-phase techniques (Stewart et al. (1969) Solid-Phase Peptide Synthesis, W H Freeman, San Francisco Calif.; Merrifield (1963) J Am Chem Soc 5:2149-2154), or using machines such as the ABI 431A peptide synthesizer (Applied Biosystems, Foster City Calif.). Proteins produced by any of the above methods may be used as pharmaceutical compositions to treat disorders associated with null or inadequate expression of the genomic sequence.

[0079] Screening and Purification Assays

[0080] A protein or a portion thereof encoded by the cDNA may be used to screen a library or a plurality of molecules or compounds for a ligand with specific binding affinity or to purify a molecule or compound from a sample. The protein or portion thereof employed in such screening may be free in solution, affixed to an abiotic or biotic substrate, or located intracellularly. For example, viable or fixed prokaryotic host cells that are stably transformed with recombinant nucleic acids that have expressed and positioned a protein on their cell surface can be used in screening assays. The cells are screened against a library or a plurality of ligands and the specificity of binding or formation of complexes between the expressed protein and the ligand may be measured. The ligands may be DNA, RNA, or PNA molecules, agonists, antagonists, antibodies, immunoglobulins, inhibitors, peptides, pharmaceutical agents, proteins, drugs, or any other test molecule or compound that specifically binds the protein. An exemplary assay involves combining the mammalian protein or a portion thereof with the molecules or compounds under conditions that allow specific binding and detecting the bound protein to identify at least one ligand that specifically binds the protein.

[0081] 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 or oligopeptide or fragment thereof. One method for high throughput screening using very small assay volumes and very small amounts of test compound is described in U.S. Pat. No. 5,876,946. Molecules or compounds identified by screening may be used in a model system to evaluate their toxicity, diagnostic, or therapeutic potential.

[0082] The protein may be used to purify a ligand from a sample. A method for using a protein to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, recovering the bound protein, and using an appropriate chaotropic agent to separate the protein from the purified ligand.

[0083] Production of Antibodies

[0084] A protein encoded by a cDNA of the invention may be used to produce specific antibodies. Antibodies may be produced using an oligopeptide or a portion of the protein with inherent immunological activity. Methods for producing antibodies include: 1) injecting an animal, usually goats, rabbits, or mice, with the protein, or an antigenically-effective portion or an oligopeptide thereof, to induce an immune response; 2) engineering hybridomas to produce monoclonal antibodies; 3) inducing in vivo production in the lymphocyte population; or 4) screening libraries of recombinant immunoglobulins. Recombinant immunoglobulins may be produced as taught in U.S. Pat. No. 4,816,567.

[0085] Antibodies produced using the proteins of the invention are useful for the diagnosis of prepathologic
disorders as well as the diagnosis of chronic or acute diseases characterized by abnormalities in the expression, amount, or distribution of the protein. A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies specific for proteins are well known in the art. Immunoassays typically involve the formation of complexes between a protein and its specific binding molecule or compound and the measurement of complex formation. Immunoassays may employ a two-site, monoclonal-based assay that utilizes monoclonal antibodies reactive to two noninterfering epitopes on a specific protein or a competitive binding assay (Pound (1998) Immunochemical Protocols, Humana Press, Totowa N.J.).

Imunoassay procedures may be used to quantify expression of the protein in cell cultures, in subjects with a particular disorder or in model animal systems under various conditions. Increased or decreased production of proteins as monitored by immunoassay may contribute to knowledge of the cellular activities associated with developmental pathways, engineered conditions or diseases, or treatment efficacy. The quantity of a given protein in a given tissue may be determined by performing immunoassays on freeze-thawed detergent extracts of biological samples and comparing the slope of the binding curves to binding curves generated by purified protein.

Labeling of Molecules for Assay

A wide variety of reporter molecules and conjugation techniques are known by those skilled in the art and may be used in various cDNAs, polynucleotides, protein, peptide or antibody assays. Synthesis of labeled molecules may be achieved using commercial kits for incorporation of a labeled nucleotide such as \(^{32}\)P-dCTP, Cy3-dCTP or Cy5-dCTP or amino acid such as \(^{35}\)S-methionine. Polynucleotides, cDNAs, proteins, or antibodies may be directly labeled with a reporter molecule by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC (Molecular Probes, Eugene Oreg.).

The proteins and antibodies may be labeled for purposes of assay by joining them, either covalently or noncovalently, with a reporter molecule that provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported in the scientific and patent literature including, but not limited to U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

Diagnostics

The cDNAs, or fragments thereof, may be used to detect and quantify differential gene expression; absence, presence, or excess expression of mRNAs; or to monitor mRNA levels during therapeutic intervention in subjects with prostate-related disorders including prostate cancer. These cDNAs can also be utilized as markers of treatment efficacy against prostate cancer over a period ranging from several days to months. 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 altered gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

For example, the cDNA may be labeled by standard methods and added to a biological sample from a patient under conditions for hybridization complex formation. 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 the amount of label in the patient sample is significantly altered in comparison to the standard value, then the presence of the associated condition, disease or disorder is indicated.

In order to provide a basis for the diagnosis of a condition, disease or disorder associated with gene expression, a normal or standard expression profile is established. This may be accomplished by combining a biological sample taken from normal subjects, either animal or human, with a probe under conditions for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained using normal subjects with values from an experiment in which a known amount of a substantially purified target sequence is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a particular condition, disease, or disorder. Deviation from standard values toward those associated with a particular condition is used to diagnose that condition.

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

Gene Expression Profiles

A gene expression profile comprises a plurality of cDNAs and a plurality of detectable hybridization complexes, wherein each complex is formed by hybridization of one or more probes to one or more complementary sequences in a sample. The cDNA composition of the invention is used as elements on a microarray to analyze gene expression profiles. In one embodiment, the microarray is used to monitor the progression of prostate cancer. Researchers can assess and catalog the differences in gene expression between healthy and diseased tissues or cells. By analyzing changes in patterns of gene expression, prostate cancer can be diagnosed at earlier stages before the patient is symptomatic. The invention can be used to formulate a prognosis and to design a treatment regimen. The invention can also be used to monitor the efficacy of treatment. For treatments with known side effects, the microarray is employed to improve the treatment regimen. A dosage is established that causes a change in genetic expression patterns indicative of successful treatment. Expression patterns associated with the onset of undesirable side effects are avoided. This approach may be more sensitive and rapid than waiting for the patient to show inadequate improvement, or to manifest side effects, before altering the course of treatment.

In another embodiment, animal models which mimic a human disease can be used to characterize expres-
sion profiles associated with a particular condition, disorder or disease; or treatment of the condition, disorder or disease. Novel treatment regimens may be tested in these animal models using microarrays to establish and then follow expression profiles over time. In addition, microarrays may be used with cell cultures or tissues removed from animal models to rapidly screen large numbers of candidate drug molecules, looking for ones that produce an expression profile similar to those of known therapeutic drugs, with the expectation that molecules with the same expression profile will likely have similar therapeutic effects. Thus, the invention provides the means to rapidly determine the molecular mode of action of a drug.

[0098] Assays Using Antibodies

[0099] Antibodies directed against epitopes on a protein encoded by a cDNA of the invention may be used in assays to quantify the amount of protein found in a particular human cell. Such assays include methods utilizing the antibody and a label to detect expression level under normal or disease conditions. The antibodies may be used with or without modification, and labeled by joining them, either covalently or noncovalently, with a labeling moiety.

[0100] Protocols for detecting and measuring protein expression using either polyclonal or monoclonal antibodies are well known in the art. Examples include ELISA, RIA, and fluorescent activated cell sorting (FACS). Such immunoassays typically involve the formation of complexes between the protein and its specific antibody and the measurement of such complexes. These and other assays are described in Pound (supra). The method may employ a two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes, or a competitive binding assay. (See, e.g., Coligan et al. (1997) Current Protocols in Immunology, Wiley-Interscience, New York N.Y.; Pound, supra)

[0101] Therapeutics

[0102] The cDNAs and fragments thereof can be used in gene therapy. cDNAs can be delivered ex vivo to target cells, such as cells of bone marrow. Once stable integration and transcription and or translation are confirmed, the bone marrow may be reintroduced into the subject. Expression of the protein encoded by the cDNA may correct a cancer associated with mutation of a normal sequence, reduction or loss of an endogenous target protein, or overexpression of an endogenous or mutant protein. Alternatively, cDNAs may be delivered in vivo using vectors such as retrovirus, adenovi-

[0103] In addition, expression of a particular protein can be regulated through the specific binding of a fragment of a cDNA to a genomic sequence or an mRNA which encodes the protein or directs its transcription or translation. The cDNA can be modified or derivatized to any RNA-like or DNA-like material including peptide nucleic acids, branched nucleic acids, and the like. These sequences can be produced biologically by transforming an appropriate host cell with a vector containing the sequence of interest.

[0104] Molecules which regulate the activity of the cDNA or encoded protein are useful as therapeutics for prostate cancer. Such molecules include agonists which increase the expression or activity of the polynucleotide or encoded protein, respectively; or antagonists which decrease expression or activity of the polynucleotide or encoded protein, respectively. In one aspect, an antibody which specifically binds the protein may be used directly as an antagonist or indirectly as a delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express the protein.

[0105] Additionally, any of the proteins, or their ligands, or complementary nucleic acid sequences may be administered as pharmaceutical compositions or in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to affect the treatment or prevention of the conditions and disorders associated with an immune response. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects. Further, the therapeutic agents may be combined with pharmacologically-acceptable carriers including excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration used by doctors and pharmacists may be found in the latest edition of Remington’s Pharmaceutical Sciences (Maack Publishing, Easton Pa.).

[0106] Model Systems

[0107] 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 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 strains provide a convenient model for investigation of the physiological consequences of underexpression or overexpression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal inbred to overexpress a particular gene (for example, secreted in milk) may also serve as a convenient source of the protein expressed by that gene.

[0108] Transgenic Animal Models

[0109] Transgenic rodents that overexpress or underexpress a gene of interest may be inbred and used to model human diseases or to test therapeutic or toxic agents. (See, e.g., U.S. Pat. No. 5,175,838 and U.S. Pat. No. 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 such as the mouse 129/SvJ cell line are placed in a blastocyst from the C57BL/6 mouse strain, they resume normal development and contribute to tissues of the live-born animal. ES cells are preferred for use in the creation of experimental knockout and knockin animals. The method for this process is well known in the art and the steps are: the cDNA is introduced into a vector, the vector is transformed into ES cells, transformed cells are identified and microinjected into mouse cell blastocysts, blastocysts are surgically transferred to pseudopregnant dams. The resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

Knockout Analysis

In gene knockout analysis, a region of a gene is enzymatically modified to include a non-natural intervening sequence such as the neomycin phosphotransferase gene (neo; Capecci 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.

Knockin Analysis

ES cells can be used to create knockin humanized animals or transgenic animal models 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. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on the progression and treatment of the analogous human condition.

As described herein, the uses of the cDNAs, provided in the Sequence Listing of this application, and their encoded proteins are exemplary of known techniques and are not intended to reflect any limitation on their use in any technique that would be known to the person of average skill in the art. Furthermore, the cDNAs provided in this application may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known to the person of ordinary skill in the art, e.g., the triplet genetic code, specific base pair interactions, and the like. Likewise, reference to a method may include combining more than one method for obtaining or assembling full length cDNA sequences that will be known to those skilled in the art. It is also to be understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary. It is also understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims. The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

EXAMPLES

Construction of cDNA Libraries

RNA was purchased from Clontech Laboratories (Palo Alto Calif.) or isolated from various tissues. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL reagent (Life Technologies, Rockville Md.). The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated with either isopropanol or ethanol and sodium acetate, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In most cases, RNA was treated with DNase. For most libraries, poly(A) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OligoTEX latex particles (Qiagen, Valencia Calif.), or an OligoTEX mRNA purification kit (Qiagen). Alternatively, poly(A) RNA was isolated directly from tissue lysates using other kits, including the POLY(A)PURA mRNA purification kit (Ambion, Austin Tex.).

In some cases, Stratagene (La Jolla Calif.) was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies) using the recommended procedures or similar methods known in the art. (See Ausabel, supra, Units 5.1 through 6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S 1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (APB) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polyclinker of the PBUL/SCRIPT phagemid (Stratagene), PSPORI plasmid (Life Technologies), or PINCY plasmid (Incyte Pharmaceuticals). Recombinant plasmids were transformed into XL1-BLUE, XL1-BLUEMR, or SOLR competent E. coli cells (Stratagene) or DH5α, DH10B, or ELECTROMAX DH10B competent E. coli cells (Life Technologies).

In some cases, libraries were superinfected with a 5x excess of the helper phage, M13KO7, according to the method of Vieira et al. (1987, Methods Enzymol. 153:3-11) and normalized or subtracted using a methodology adapted from Soares (1994, Proc Natl Acad Sci 91:9228-9232), Swoop et al. (1991, Nucl Acids Res 19:1954), and Bandoni et al. (1996, Genome Research 6:791-806). The modified Soares normalization procedure was utilized to reduce the repetitive cloning of highly expressed high abundance cDNAs while maintaining the overall sequence complexity of the library. Modification included significantly longer hybridization times which allowed for increased gene discovery rates by biasing the normalized libraries toward those infrequently expressed low-abundance cDNAs which are poorly represented in a standard transcript image (Soares et al., supra).

II Isolation and Sequencing of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using one of the following: the Magic or WIZARD MINIPREPS DNA purification
system (Promega); the AGTC MINIPREP purification kit (Edge BioSystems, Gaithersburg Md.); the QIAWELL 8, QIAWELL 8 Plus, or QIAWELL 8 Ultra plasmid purification systems, or the REAL PREP 96 plasmid purification kit (QIAGEN, Valencia Calif.). Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

[0124] Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao (1994) Anal Biochem 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes) and a FLUOROSKAN II fluorescence scanner (Lab systems Oy, Helsinki, Finland).

[0125] cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 thermal cycler (Applied Biosystems) or the DNA ENGINE thermal cycler (MJ Research, Watertown Mass.) in conjunction with the HYDRA microdispenser (Robbins Scientific, Sunnyvale Calif.) or the MICROLAB 2200 system (Hamilton, Reno Nev.). cDNA sequencing reactions were prepared using reagents provided by APB or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE cycle sequencing kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled cDNAs were carried out using the MEGABACE 1000 DNA sequencing system (APB); the ABI PRISM 373 or 377 sequencing systems (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, supra, Unit 7.7).

[0126] III Extension of cDNA Sequences

[0127] Nucleic acid sequences were extended using the cDNA clones 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 Oligo 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

[0128] Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed. Preferred libraries are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred because they will contain more sequences with the 5’ and upstream regions of genes. A randomly primed library is particularly useful if an oligo d(T) library does not yield a full-length cDNA.

[0129] High fidelity amplification was obtained by PCR using methods well known in the art. 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 Mg²⁺, (NH₄)₂SO₄, 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 Pharmaceuticals): Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ (Stratagene) were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

[0130] The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN reagent (0.25% reagent in 1x TE, v/v; Molecular Probes) and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton Mass.) 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 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1% agarose mini-gel to determine which reactions were successful in extending the sequence.

[0131] The extended nucleic acids were desalted and concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison Wis.), and sonicated or sheared prior to religation into pUC18 vector (APB). For shotgun sequencing, the digested nucleic acids were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with AGARACE enzyme (Promega). Extended clones were religated using T4 DNA ligase (New England Biolabs, Beverly Mass.) into pUC18 vector (APB), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transformed into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carbenicillin liquid media.

[0132] The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (APB) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 25 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified using PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same 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).

[0133] IV Assembly and Analysis of Sequences

[0134] Component nucleotide sequences from chromatograms were subjected to PHRED analysis (Phil Green, University of Washington, Seattle Wash.) and assigned a quality score. The sequences having at least a required quality score were subject to various pre-processing algorithms to eliminate low quality 3' ends, vector and linker

[0135] Processed sequences were subjected to assembly procedures in which the sequences were assigned to bins, one sequence per bin. Sequences in each bin were assembled to produce consensus sequences, templates. Subsequent new sequences were added to existing bins using BLAST (Altschul supra; Altschul et al. supra; Karlin et al. (1988) Proc Natl Acad Sci 85:841-845), BLASTn (vers. 1.4, WashU), and CROSSMATCH software (Phil Green, supra). Candidate pairs were identified as all BLAST hits having a quality score greater than or equal to 150. Alignments of at least 82% local identity were accepted into the bin. The component sequences from each bin were assembled using PHRAP (Phil Green, supra). Bins with several overlapping component sequences were assembled using DEEP PHRAP (Phil Green, supra).

[0136] Bins were compared against each other, and those having local similarity of at least 82% were combined and reassembled. Reassembled bins having templates of insufficient overlap (less than 95% local identity) were re-split. Assembled templates were also subjected to analysis by STITCHER/EXON MAPPER algorithms which analyzed the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types, disease states, and the like. These resulting bins were subjected to several rounds of the above assembly procedures to generate the template sequences found in the LIFESEQ GOLD database (Incyte Genomics).

[0137] The assembled templates were annotated using the following procedure. Template sequences were annotated using BLASTn (vers. 2.0, NCBI) versus GBpr (GenBank vers. 116). "Hits" were defined as an exact match having from 95% local identity over 200 base pairs through 100% local identity over 100 base pairs, or a homolog match having an E-value equal to or greater than 1x10^-8. The "E-value" quantifies the statistical probability that a match between two sequences occurred by chance. The hits were subjected to frameshift FASTx versus GENPEPT (GenBank version 109). In this analysis, a homolog match was defined as having an E-value of 1x10^-8. The assembly method used above was described in U.S. Ser. No. 09/276,534, filed Mar. 25, 1999, and the LIFESEQ GOLD user manual (Incyte Genomics).


[0139] V. Selection of Sequences, Microarray Preparation and Use

[0140] Incyte clones represent template sequences derived from the LIFESEQ GOLD assembled human sequence database (Incyte Genomics). In cases where more than one clone was available for a particular template, the 5’-most clone in the template was used on the microarray. The HUMAN GENOME GEM series 1-3 microarrays (Incyte Pharmaceuticals) contain 28,626 array elements which represent 10,068 annotated clusters and 18,558 unannotated clusters. Tables 1 and 2 show the GenBank annotations for SEQ ID Nos:1-x of this invention as produced by BLAST analysis.

[0141] To construct microarrays, cDNAs were amplified from bacterial cells using primers complementary to vector sequences flanking the cDNA insert. Thirty cycles of PCR increased the initial quantity of cDNAs from 1-2 ng to a final quantity greater than 5 μg. Amplified cDNAs were then purified using SEPHAERYL-400 columns (APB). Purified cDNAs were immobilized on polymer-coated glass slides. Glass microscope slides (Corning, Corning N.Y.) were cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides were etched in 4% hydrofluoric acid (VWR Scientific Products, West Chester Pa.), washed thoroughly in distilled water, and coated with 0.05% aminopropyl silane (Sigma Aldrich) in 95% ethanol. Coated slides were cured in a 110° C. oven. cDNAs were applied to the coated glass substrate using a procedure described in U.S. Pat. No. 5,807,522. One microliter of the cDNA at an average concentration of 100 ng/µl was loaded into the open capillary printing element by a high-speed robotic apparatus which then deposited about 5 nl of cDNA per slide.

[0142] Microarrays were UV-crosslinked using a STRATAinker UV-crosslinker (Stratagene), and then washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites were blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (Tropix, Bedford Mass.) for 30 minutes at 60° C. followed by washes in 0.2% SDS and distilled water as before.

[0143] VI. Preparation of Samples

[0144] The following cell lines were obtained from American Type Culture Collection (Manassas Va.) and cultured in media according to the manufacturer’s protocols: PZ-HPV-7 was derived from epithelial cells cultured from normal tissue from the peripheral zone of the prostate. CA-HPV-10 was derived from cells of a prostatic adenocarcinoma of Gleason Grade 4/4. Both PZ cells were transformed by transfection with human papillomavirus (HPV)-18, and express keratins 5 and 8 and the early region 6 oncoprotein of HPV. PZ-HPV-7 and CA-HPV-10 are negative for prostate specific antigen (PSA). DU-145 is a prostate carcinoma cell line isolated from a 69 year-old man with widespread metastatic disease. DU-145 was isolated from a brain metastasis and has no detectable hormone sensitivity. Further, DU-145 is negative for PSA; PC-3 is a prostate adenocarcinoma cell line isolated from a 62 year-old male with grade IV prostate adenocarcinoma metastasized to the bone. PC-3 cells exhibit low acid phosphatase and testosterone-5-alpha reductase activities; LNCaP is a prostate carcinoma cell line isolated from a lymph node biopsy of a 50
year-old male with metastatic prostate carcinoma. LNCaP cells are responsive to 5-alpha-dihydrotestosterone and express androgen receptors.

[0145] PrEC, a primary prostate epithelial cell line isolated from a normal donor, was obtained from Clinomics Corporation (Walkersville Md.) and cultured in media according to the manufacturer’s protocols.

[0146] All cultures were maintained at 37° C. and 5% CO₂ for 3-5 passages.

[0147] Isolation and Labeling of Sample cDNAs

[0148] Cells were harvested when cultures were approximately 70% confluent and lysed in 1 ml of TRIZOL reagent (5x10⁶ cells/ml; Life Technologies). The lysates were vortexed thoroughly and incubated at room temperature for 2-3 minutes and extracted with 0.5 ml chloroform. The extract was mixed, incubated at room temperature for 5 minutes, and centrifuged at 15,000 rpm for 15 minutes at 4° C. The aqueous layer was collected and an equal volume of isopropanol was added. Samples were mixed, incubated at room temperature for 10 minutes, and centrifuged at 15,000 rpm for 20 minutes at 4° C. The supernatant was removed and the RNA pellet was washed with 1 ml of 70% ethanol, centrifuged at 15,000 rpm at 4° C, and resuspended in RNase-free water. The concentration of the RNA was determined by measuring the optical density at 260 nm.

[0149] Poly(A) RNA was prepared using an Oligotex mRNA kit (Qiagen) with the following modifications: oligotex beads were washed in tubes instead of on spin columns, resuspended in elution buffer, and then loaded onto spin columns to recover mRNA. To obtain maximum yield, the mRNA was eluted twice.

[0150] Each poly(A) RNA sample was reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo(dT) primer (21 mer), 1x first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, and 40 µM dATP, dGTP, dTTP, dCTP, Cy3 or dCTP, Cy5 (APB). The reverse transcription reaction was performed in a 25 µl volume containing 200 ng poly(A) RNA using the GE Bahrain kit (Incite Pharmaceuticals). Specific control poly(A) RNAs (YC8R06, YC8R45, YC8R67, YC8R85, YC8R43, YC8R22, YC8R23, YC8R25, YC8R44, YC8R26) were synthesized by in vitro transcription from non-coding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, control mRNAs (YC8R06, YC8R45, YC8R67, and YC8R85) at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng were diluted into reverse transcription reaction at ratios of 1:100,000, 1:10,000, 1:1000, 1:100 (w/w) to sample mRNA, respectively. To sample differential expression patterns, control mRNAs (YC8R43, YC8R22, YC8R23, YC8R25, YC8R44, YC8R26) were diluted into reverse transcription reaction at ratios of 1:3, 1:10, 10:1, 1:25, 25:1 (w/w) to sample mRNAs. Reactions were incubated at 37° C. for 2 hr, treated with 2.5 µl of 0.5M sodium hydroxide, and incubated for 20 minutes at 85° C. to stop the reaction and degrade the RNA.

[0151] cDNAs were purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech). Cy3- and Cy5-labeled reaction reactions were combined as follows: Aliquots of Cy3-labeled PrEC cDNA were individually mixed with Cy5 labeled cDNA from PZ-HPV-7, CA-HPV-10, DU-145, PC-3, and LNCaP cells. The mixtures were ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 µg of glycogen, and 300 µl of 100% ethanol, dried to completion using a SpeedVac system (Savant Instruments, Holbrook, N.Y.) and resuspended in 14 µl 5×SSC/0.2% SDS.

[0152] VII Hybridization and Detection

[0153] Hybridization reactions contained 9 µl of sample mixture containing 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5×SSC, 0.2% SDS hybridization buffer. The mixture was heated to 65° C. for 5 minutes and was aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The microarrays were transferred to a water bathed chamber having a cavity just slightly larger than a microscope slide. The chamber was kept at 100% humidity internally by the addition of 140 µl of 5× SSC in a corner of the chamber. The chamber containing the microarrays was incubated for about 6.5 hours at 60° C. The microarrays were washed for 10 min at 45° C. in low stringency wash buffer (1×SSC, 0.1% SDS), three times for 10 minutes each at 45° C. in high stringency wash buffer (0.1×SSC), and dried.

[0154] Reporter-labeled hybridization complexes were detected with a microscope equipped with an Incova 70 mixed gas 10 W laser (Coherent, Santa Clara Calif.) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light was focused on the microarray using a 2mms microscope objective (Nikon, Melville N.Y.). The slide containing the microarray was placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm×1.8 cm microarray used in the present example was scanned with a resolution of 20 micrometers.

[0155] In two separate scans, the mixed gas multline laser excited the two fluorophores sequentially. Emitted light was split, based on wavelength, into two photomultiplier tube detectors (PMT R1477; Hamamatsu Photonics Systems, Bridgewater N.J.) corresponding to the two fluorophores. Appropriate filters positioned between the microarray and the photomultiplier tubes were used to filter the signals. The emission maxima of the fluorophores used were 565 nm for Cy3 and 650 nm for Cy5. Each microarray was typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus was capable of recording the spectra from both fluorophores simultaneously.

[0156] The sensitivity of the scans was calibrated using the signal intensity generated by a cDNA control species. Samples of the calibrating cDNA were separately labeled with the two fluorophores and identical amounts of each were added to the hybridization mixture. A specific location on the microarray contained 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.

[0157] The output of the photomultiplier tube was digitized using 12-bit R-R835F analog-to-digital (A/D) conversion board (Analog Devices, Norwood, Mass.) installed in an IBM-compatible PC computer. The digitized data were displayed as an image where the signal intensity was mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high...
Signal). The data was also analyzed quantitatively. Where two different fluorophores were excited and measured simultaneously, the data were first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore’s emission spectrum.

[0159] A grid was superimposed over the fluorescence signal image such that the signal from each spot was centered in each element of the grid. The fluorescence signal within each element was then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis was the GEM-TOOLS gene expression analysis program (Incyte Pharmaceuticals). Significance was defined as signal to background ratio exceeding 2x and area hybridization exceeding 40%.

[0159] VIII Data Analysis and Results

[0160] Array elements that exhibited at least 2.5-fold change in expression at one or more time points, a signal intensity over 250 units, a signal-to-background ratio of at least 2.5, and an element spot size of at least 40% were identified as differentially expressed using the GEMTOOLS program (Incyte Genomics). Differential expression values were converted to log base 2 scale. Differential expression values were then compared between the cell lines to identify genes which discriminated between normal and cancerous and between non-metastatic and metastatic cancer. The student’s t-test and Pearson correlation statistics were used to test significant differences between the groups. The resulting cDNAs are shown in Tables 1 and 2. The cDNAs are identified by their Clone ID. Table 3 shows the sequence overlap between the clones identified in Tables 1 and 2 and gene templates. Columns 1-5 show the SEQ ID NO., Template ID, and Clone ID, respectively. Columns 4 and 5 show the start and stop nucleotides for the clone on the template. Table 4 shows a GenBank homolog and description associated with at least a fragment of each template ID. The descriptions were obtained using the sequences of the Sequence Listing and BLAST analysis. SEQ ID NOs:1-3, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75 are highly correlated with metastatic prostate cancer cells PC-3, LNCaP, and DU-145, and SEQ ID NOs:76, 78-86, 88-90, 92-97, 99-101 are differentially expressed at significant levels in all of the prostate cancer cell lines.

[0161] IX Other Hybridization Technologies and Analyses

[0162] Other hybridization technologies utilize a variety of substrates such as nylon membranes, capillary tubes, etc. Arranging cDNAs on polymer coated slides is described in Example V; sample cDNA preparation and hybridization and analysis using polymer coated slides is described in examples VI and VII, respectively. cDNAs are applied to a membrane 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 37°C 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 0.2xSSC for 10 min each. The membrane is then UV irradiated in a STRATALINKER UV-crosslinker (Stratagene).

[0163] 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.

[0164] Hybridization probes derived from cDNAs of the Sequence Listing are employed for screening cDNAs, mRNAs, or genomic DNA in membrane-based hybridizations. Probes are prepared by diluting the cDNAs to a concentration of 40-50 ng in 45 µl TE buffer, denaturing by heating to 100°C 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 microliters of [32P]dCTP is added to the tube, and the contents are incubated at 37°C for 10 min. The labeling reaction is stopped by adding 5 µl of 0.2M EDTA, and probe is purified from unincorporated nucleotides using a PROBEQUANT G-50 microcolumn (APB). The purified probe is heated to 100°C for five min and then snap cooled for two min on ice.

[0165] Membranes are pre-hybridized in hybridization solution containing 1% Sarkosyl and 1x high phosphoric buffer (0.5 M NaCl, 0.1 M Na2HPO4, 5 mM EDTA, pH 7) at 55°C for two hr. The probe, diluted in 15 µl fresh hybridization solution, is then added to the membrane. The membrane is hybridized with the probe at 55°C for 16 hr. Following hybridization, the membrane is washed for 15 min at 25°C in 1 mM Tris (pH 8.0), 1% Sarkosyl, and four times for 15 min each at 25°C in 1 mM Tris (pH 8.0). To detect hybridization complexes, XOMAT-AR film (Eastman Kodak, Rochester N.Y.) is exposed to the membrane overnight at ~70°C, developed, and examined.

[0166] X Further Characterization of Differentially Expressed cDNAs and Proteins

[0167] Clones were blasted against the LIFSEQ Gold 5.1 database (Incyte Genomics) and an Incyte template and its sequence variants were chosen for each clone. The template and variant sequences were blasted against GenBank database to acquire annotation. The nucleotide sequences were translated into amino acid sequence which was blasted against the GenPept and other protein databases to acquire annotation and characterization, i.e., structural motifs.

[0168] Percent sequence identity can be determined electronically for two or more amino acid or nucleic acid sequences using the MEGALIGN program (DNASTAR). The percent identity between two amino acid sequences is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue
matches between sequence A and sequence B, times one hundred. Gaps of low or of no homology between the two amino acid sequences are not included in determining percentage identity.

[0169] Sequences with conserved protein motifs may be searched using the BLOCKS search program. This program analyses sequence information contained in the Swiss-Prot and PROSITE databases and is useful for determining the classification of uncharacterized proteins translated from genomic or cDNA sequences (Bairoch et al. (supra); Attwood et al. (supra). PROSITE database is a useful source for identifying functional or structural domains that are not detected using motifs due to extreme sequence divergence. Using weight matrices, these domains are calibrated against the SWISS-PROT database to obtain a measure of the chance distribution of the matches.

[0170] The PRINTS database can be searched using the BLIMP search program to obtain protein family “fingerprints”. The PRINTS database complements the PROSITE database by exploiting groups of conserved motifs within sequence alignments to build characteristic signatures of different protein families. For both BLOCKS and PRINTS analyses, the cutoff scores for local similarity were: >1300=strong, 1000-1300=suggestive; for global similarity were: p<exp-3; and for strength (degree of correlation) were: >1300=strong, 1000-1300=weak.

[0171] X Expression of the Encoded Protein

[0172] Expression and purification of a protein encoded by a cDNA of the invention is achieved using bacterial or virus-based expression systems. For expression in bacteria, cDNA is subcloned into a vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the trp-tac (lac) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the lac operator regulatory element. Recombinant vectors are transformed into bacterial hosts, such as BL21 (DE3). Antibiotic resistant bacteria express the protein upon induction with IPTG. Expression in eukaryotic cells is achieved by infecting Spodoptera frugiperda (SF9) insect cells with recombinant baculovirus, Autographa californica nuclear polyhedrosis virus. The polyhedrin gene of baculovirus is replaced with the cDNA by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of transcription.

[0173] For ease of purification, the protein is synthesized as a fusion protein with glutathione-S-transferase (GST; APB) or a similar alternative such as FLAG. The fusion protein is purified on immobilized glutathione under conditions that maintain protein activity and antigenicity. After purification, the GST moiety is proteolytically cleaved from the protein with thrombin. A fusion protein with FLAG, an 8-amino acid peptide, is purified using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak, Rochester N.Y.).

[0174] XI Production of Specific Antibodies

[0175] A denatured protein from a reverse phase HPLC separation is obtained in quantities up to 75 mg. This denatured protein is used to immunize mice or rabbits following standard protocols. About 100 μg is used to immunize a mouse, while up to 1 mg is used to immunize a rabbit. The denatured protein is radioiodinated and incubated with murine B-cell hybridomas to screen for monoclonal antibodies. About 20 mg of protein is sufficient for labeling and screening several thousand clones.

[0176] In another approach, the amino acid sequence translated from a cDNA of the invention is analyzed using PROTEAN software (DNASTAR) to determine regions of high antigenicity, essentially antigenically-effective epitopes of the protein. The optimal sequences for immunization are usually at the C-terminus, the N-terminus, and those intervening, hydrophilic regions of the protein that are likely to be exposed to the external environment when the protein is in its natural conformation. Typically, oligopeptides about 15 residues in length are synthesized using an ABI 431 peptide synthesizer (Applied Biosystems) using Fmoc-chemistry and then coupled to keyhole limpet hemocyanin (KLH; Sigma Aldrich) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester. If necessary, a cysteine may be introduced at the N-terminus of the peptide to permit coupling to KLH. Rabbits are immunized with the oligopeptide-KLH complex in complete Freund’s adjuvant. The resulting antisera are tested for antipeptide activity by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG.

[0177] Hybridomas are prepared and screened using standard techniques. Hybridomas of interest are detected by screening with radioiodinated protein to identify those fusions producing a monoclonal antibody specific for the protein. In a typical protocol, wells of 96 well plates (FAST, Becton-Dickinson, Palo Alto Calif.) are coated with affinity-purified, specific rabbit-anti-mouse (or suitable anti-species Ig) antibodies at 10 μg/ml. The coated wells are blocked with 1% BSA and washed and exposed to supernatants from hybridomas. After incubation, the wells are exposed to radiolabeled protein at 1 mg/ml. Clones producing antibodies bind a quantity of labeled protein that is detectable above background.

[0178] Such clones are expanded and subjected to 2 cycles of cloning at 1 cell/3 wells. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from the ascitic fluid by affinity chromatography on protein A (APB). Monoclonal antibodies with affinities of at least 10^8 M^-1, preferably 10^10 to 10^12 M^-1 or stronger, are made by procedures well known in the art.

[0179] XII Purification of Naturally Occurring Protein Using Specific Antibodies

[0180] Naturally occurring or recombinant protein is substantially purified by immunoaffinity chromatography using antibodies specific for 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 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 and the protein or portions thereof are labeled with $^{32}$P-dCTP, Cy3-dCTP, Cy5-dCTP (APB), or BODIPY or FITC (Molecular Probes), respectively. Candidate molecules or compounds previously arranged on a substrate are incubated in the presence of labeled nucleic acid or amino acid. After incubation under conditions for either a cDNA or a protein, the substrate is washed, and any position on the substrate retaining label, which indicates specific binding or complex formation, is assayed. The binding molecule is identified by its arrayed position on the substrate. Data obtained using different concentrations of the nucleic acid or protein are used to calculate affinity between the labeled nucleic acid or protein and the bound molecule. High throughput screening using very small assay volumes and very small amounts of test compound is fully described in Burbaum et al. U.S. Pat. No. 5,876,946.

All patents and publications mentioned in the specification are incorporated herein by reference. 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 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.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 102
<210> SEQ ID NO 1
<211> LENGTH: 1645
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc feature
<223> OTHER INFORMATION: Incyte ID No.: 1382961.3
<225> SEQ ID NO: 1
<400 SEQUENCE: 1
ctttcctcaat ttaccgagac accttcctctt cactcagca acagtctccgct egotcactctc 60
cttctctcagct acagtctacgct cccccctcctt ctctcactcgc ccaatc 120
ctttctccttc ctgtctggggtgac gctttcctctt ccgacctctt cccccctcctt ccaatc 180
cctgcgtgggtgggt gcgggtggtgac gctttcctctt ccgacctctt cccccctcctt ccaatc 240
cctttcctcttc ctgtctggggtgac gctttcctctt ccgacctctt cccccctcctt ccaatc 300
dacacagcagc cttctcttcag cccccctcctt ccgacctctt cccccctcctt ccaatc 360
ggttggtttggtgac cccccctcctt ccgacctctt cccccctcctt ccaatc 420
gttgcacgtgc acacagcagc cccccctcctt ccgacctctt cccccctcctt ccaatc 480
ggacacccactg gccgaccccaag ctgcctccag cccccctcctt ccgacctctt cccccctcctt ccaatc 540
tgctctccgatgc acacacaccccc cccccctcctt ccgacctctt cccccctcctt ccaatc 600
tttctctatcag cccccctcctt ccgacctctt cccccctcctt ccaatc 660
gcccctcttcag cccccctcctt ccgacctctt cccccctcctt ccaatc 720
gggcgcgtgggtgac gctttcctctt ccgacctctt cccccctcctt ccaatc 780
ggttggtttggtgac cccccctcctt ccgacctctt cccccctcctt ccaatc 840
ggacacccactg gccgaccccaag ctgcctccag cccccctcctt ccgacctctt cccccctcctt ccaatc 900
tgctctccgatgc acacacaccccc cccccctcctt ccgacctctt cccccctcctt ccaatc 960
ggacacccactg gccgaccccaag ctgcctccag cccccctcctt ccgacctctt cccccctcctt ccaatc 1020
ggttggtttggtgac cccccctcctt ccgacctctt cccccctcctt ccaatc 1080
catactccag cccccctcctt ccgacctctt cccccctcctt ccaatc 1140
ggacacccactg gccgaccccaag ctgcctccag cccccctcctt ccgacctctt cccccctcctt ccaatc 1200
gttggcagc gtggagggc agctggcaca gctcgcctgc gagatggagc agcagacca 1260
gggatacaag atctgctggc acgtgaagac gcggctggag caggagatcg caaoctacgg 1320
cgctcgtcg gaggccgcgg acgcccaact ctctcccctt caggtcctct ctggactgca 1380
gtctacctca ggtctgactt ctctccggcg ccacaccgcgc ccacatgtcgc tgcatgtgcg 1440
cgcgggaag gttgctgctca ccacagagca gctcctggag ggacagacact gaggtggccg 1500
agccgcdttc agggctagaag ggccggccgcttg gtcgagcagc atccacagt cagatcctct 1560
cctctggccg agcactcctc agctggaccc tgtctacccct tccaccccct cctggcatcag 1620
atatacgctc attatacgag ttgca 1645
<210> SEQ ID NO 2
<211> LENGTH: 1051
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1382961.5
<400> SEQUENCE: 2
agacaatgcc ttcctctcgg atgtctctgg caagggcttc tggctatctc ccacacctg 60
actctccca ttttagagc gcatgtagtc ccaggaagag cttggagcggc tccacaggg 120
gctgtgaggg ccgctctcag gagatagagct gcggctggagc gcgcagacca gcagacact 180
ccctctacca tgcgcgtccgg gcgggtcttc ccctcccctc ccacagacagc gagcagaccc 240
tgtgatttta agggcccttc aagttcctagg gactagctcg gtgcgtgctc tgagagagc 300
cacgcgcgac cttgagagtg aatcgtctcg ctgtcagccag ctggccgctcg ctcgtagagc 360
cacgacacta agctgtcctc tcagacacct tgagggcctc aggacagaga ccctcctccg 420
cacagttggg aatgcagctg cttcctcaggt gactagcact ggcgtctggg cccggcctg 480
ctcttagcgc aagttagaga ccagctcttg cctctgcgag atgtggagag cagacactc 540
atggctgcgg caggtgcctg gcaggaactg ccctgagcgg agctgacccc agtgcagagc 600
tgaggcgctg aggagggcct gcggcggctgt gcagagagac cagggcggag cagatgaatgc 660
cctgtgaggg caggtgcctg agatgagcct tggagggctg gcagggctgt gcagagagac 720
cctgagccgc attcagcagc agatgccttg ccagctctcag aagatggcaag agaagacacc 780
caagagttgg gcagagttgg cttcctacaa gcagagggag ctggacccgcg agtggccacg 840
cacagacactgt cgtgcgagcg gcggagacat gcactgctcg gcggctggtg gcaggctcagc 900
gacagctg agtggagcct gcgtgagtg gcagctggtg gcagctcgtg gcgagctgagc 960
gacagagac caggattaca agctgcctgt gcagctggaag cgcggctcag agccgagatc 1020
cgcacaccgc ccgcctgcag tggagggcag g 1051
<210> SEQ ID NO 3
<211> LENGTH: 1930
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2852561C01
<400> SEQUENCE: 3
ccctaggggg gcggcgcgggc ccggggcgcg cttgacccttc cccagagccgc cgcctgcctg 60
tgagcgagtgcagtacgcca aactgctcgag aggacacgag gcccgggaga gcgttctggg 120
tccaggytcgagagctgg tggagccccactgcacgc aactgctcgc gctgctccgg 180
tgtcgcagcc acacagcgcc aagagggataa ggcttctggg aactgctcgc 240
tgtcgcagc ccacagcgcc gcccgggaga gcgttctggg aactgcacgc 300
ggagagac gaaatgctccg taagaactcgg ggtgctccg gccagagggaga 360
cgcgctacgtactcaggggag aacctgctcgc gctgctccgg 420
tgagcgagtgcagtacgcca aactgctcgc gctgctccgg 480
tgctgcgtgc gtcgagctcgg ttgtgtcactgc gcgcgtctccgc gcgcgtctccgc 540
gtgcgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 600
cgcgctacgtactcaggggaga aacctgctcgc gctgctccgg 660
ggcgctacgtactcaggggaga aacctgctcgc gctgctccgg 720
ggcgctacgtactcaggggaga aacctgctcgc gctgctccgg 780
taaaaag aacagtctcgc aacagtctcgc aacagtctcgc aacagtctcgc aacagtctcgc 840
ggcgctacgtactcaggggaga aacctgctcgc gctgctccgg 900
ggcgctacgtactcaggggaga aacctgctcgc gctgctccgg 960
cctgctctctgg ggtgctccgg aacacattcgc aacacattcgc aacacattcgc 1020
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1080
tgcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1140
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1200
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1260
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1320
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1380
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1440
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1500
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1560
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1620
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1680
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1740
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1800
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1860
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1920
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 1980
tgctgcgtgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc gcgcgtctccgc 2040
<210> SEQ ID NO 4
<211> LENGTH: 529
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2852561CD1
<400> SEQUENCE: 4
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Ser Gln Asp Thr Glu Val Asp Met Lys Glu Val Glu Leu Asn</td>
<td>1 5 10 15</td>
</tr>
<tr>
<td>Glu Leu Glu Pro Glu Lys Glu Pro Met Asn Ala Ala Ser Gly Ala</td>
<td>20 25 30</td>
</tr>
<tr>
<td>Ala Met Ser Leu Ala Gly Ala Glu Met Leu Ala Gly Val Lys Ile</td>
<td>35 40 45</td>
</tr>
<tr>
<td>Lys Val Ala Glu Asp Glu Ala Ala Ala Ala Ala Lys Phe</td>
<td>50 55 60</td>
</tr>
<tr>
<td>Thr Gly Leu Ser Lys Glu Glu Leu Leu Lys Val Ala Gly Ser Pro</td>
<td>65 70 75</td>
</tr>
<tr>
<td>Gly Trp Val Arg Thr Arg Trp Ala Leu Leu Leu Leu Phe Trp Leu</td>
<td>80 85 90</td>
</tr>
<tr>
<td>Gly Trp Leu Gly Met Leu Ala Gly Ala Val Val Ile Val Arg</td>
<td>95 100 105</td>
</tr>
<tr>
<td>Ala Pro Arg Cys Arg Glu Leu Pro Ala Gin Lys Trp Trp His Thr</td>
<td>110 115 120</td>
</tr>
<tr>
<td>Gly Ala Leu Tyr Arg Ile Gly Asp Leu Gin Ala Phe Gin Gly His</td>
<td>125 130 135</td>
</tr>
<tr>
<td>Gly Ala Gly Asn Leu Ala Gly Leu Lys Gly Arg Leu Asp Tyr Leu</td>
<td>140 145 150</td>
</tr>
<tr>
<td>Ser Ser Leu Lys Val Lys Gly Leu Val Leu Gly Pro Ile His Lys</td>
<td>155 160 165</td>
</tr>
<tr>
<td>Asn Gin Lys Asp Asp Val Ala Gin Thr Asp Leu Gin Ile Asp</td>
<td>170 175 180</td>
</tr>
<tr>
<td>Pro Aan Phe Gly Ser Lys Glu Asp Phe Asp Ser Leu Leu Gin Ser</td>
<td>185 190 195</td>
</tr>
<tr>
<td>Ala Lys Lys Ser Ile Arg Val Ile Leu Asp Leu Thr Pro Aan</td>
<td>200 205 210</td>
</tr>
<tr>
<td>Tyr Arg Gly Glu Asn Ser Trp Phe Ser Thr Gin Val Asp Thr Val</td>
<td>215 220 225</td>
</tr>
<tr>
<td>Ala Thr Lys Val Lys Asp Ala Leu Glu Phe Trp Leu Gin Ala Gly</td>
<td>230 235 240</td>
</tr>
<tr>
<td>Val Aap Gly Phe Gin Val Arg Asp Ile Glu Aen Leu Lys Asp Ala</td>
<td>245 250 255</td>
</tr>
<tr>
<td>Ser Ser Phe Leu Ala Glu Trp Gin Aan Ile Thr Lys Gly Phe Ser</td>
<td>260 265 270</td>
</tr>
<tr>
<td>Glu Asp Arg Leu Leu Ile Ala Gly Thr Aen Ser Ser Asp Leu Gin</td>
<td>275 280 285</td>
</tr>
<tr>
<td>Gin Ile Leu Ser Leu Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>290 295 300</td>
</tr>
<tr>
<td>Ser Ser Tyr Leu Ser Asp Ser Gly Ser Thr Gly Glu His Thr Lys</td>
<td>305 310 315</td>
</tr>
<tr>
<td>Ser Leu Val Thr Gin Tyr Leu Aen Ala Thr Gin Gin Gin Gin Gin</td>
<td>320 325 330</td>
</tr>
<tr>
<td>Ser Trp Ser Leu Ser Gin Ala Arg Leu Leu Thr Ser Phe Leu Pro</td>
<td>335 340 345</td>
</tr>
<tr>
<td>Ala Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>350 355 360</td>
</tr>
<tr>
<td>Gly Thr Pro Val Phe Ser Tyr Gly Asp Gin Gin Gin Gin Gin Gin</td>
<td>365 370 375</td>
</tr>
<tr>
<td>Ala Ala Leu Pro Gly Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>380 385 390</td>
</tr>
<tr>
<td></td>
<td>380</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>Asp</td>
<td>Glu</td>
</tr>
<tr>
<td>Phe</td>
<td>Pro</td>
</tr>
<tr>
<td>Pro</td>
<td>Asp</td>
</tr>
<tr>
<td>Val</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>2</td>
</tr>
<tr>
<td>Net</td>
<td>Thr</td>
</tr>
<tr>
<td>Gly</td>
<td>Gln</td>
</tr>
<tr>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Ala</td>
</tr>
<tr>
<td>Leu</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>Phe</td>
<td>His</td>
</tr>
<tr>
<td>Ser</td>
<td>Ala</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>Ser</td>
<td>2</td>
</tr>
<tr>
<td>Ser</td>
<td>Tyr</td>
</tr>
<tr>
<td>His</td>
<td>Trp</td>
</tr>
<tr>
<td>Aan</td>
<td>Glu</td>
</tr>
<tr>
<td>Leu</td>
<td>Val</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td>Leu</td>
<td>2</td>
</tr>
<tr>
<td>Glu</td>
<td>Arg</td>
</tr>
<tr>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>Phe</td>
<td>2</td>
</tr>
<tr>
<td>Pro</td>
<td>Tyr</td>
</tr>
</tbody>
</table>

**SEQ ID NO 5**
**LENGTH: 664**
**TYPE: DNA**
**ORGANISM: Homo sapiens**
**FEATURE: misc_feature**
**OTHER INFORMATION: Incyte ID No: 335942.2**

```plaintext
ccaaatgtga gggatatata cctaagcctt ctgtgcgcga aaggtttttata tgctaactcag 60
ggaggttaatt gttgataatt acctctgtgaa gcctggctca aagatgaaggt caaagcagaa 120
aatgttatttt tttatattatt attatatatt gatatttttatt atatatattata gataattata 180
atatcataatt attggyaga ttttctcttct tttgtattttg gaccggcact ctctcacaatt 240
agcggagctt gttttttgag ataaatgatt tattattcat taataacaggg cattttggtgc 300
ccccggtgtg tttccttattg gactactggg aagccttgaa 360
tcataataa attatatttt gatagcctgg tatttttaatt tttttattctt ctcctgactca actctctctat 420
tttgcatcatt aatttattttt gacacttctttg attaaactcctt ctcctgactca actctctctat 480
gctaaaaatt gcgagcacta acttggcacta cctgtgacacta gctgctctctat 540
tttctgatat taattaattttt tcttttatatt tttttttatttt tggattctcact ccataatgtt 600
acattttatt ccaactaattg gattgataga ggtatttactc ctaataattt aaaaagaccc 660
cacc 664
```

**SEQ ID NO 6**
**LENGTH: 1667**
**TYPE: DNA**
**ORGANISM: Homo sapiens**
**FEATURE: misc_feature**
**OTHER INFORMATION: Incyte ID No: 2403854CB1**

```plaintext
```
gtcttttcag tgctgtatga gtcagcaaaag aagtcagagat ggccasaggg tccagacatgt
60
ttgaagacct gaaagctgat tacaggaga atagaagaga cagttctctct attgatctct
120
tgctctgaa tcaagaaatc tctctactag tgaagactga cccactcctat gaagcctgca
180
tgcaactac ttgctctctg aatctctctct acctctctct acctctctct acctctctct
240
aggaagcact ggtggtgtaa gcacaccacg ggaagctgct gagaaggaga cgggttgtgt
300
tacccctgctg catcactgct gagctgctgct acggcgccag tcaagcttctgc gagaagtcac
360
tccactactg tagctctcaag ccttttgcct ctgctgctgct ctttttgcct ctttttgcct
420
ggactctac atcaggtttc atcagttcgt accagcctca atacagttcat atacagttcat
480
atcaggtttc atcagttcgt accagcctca atacagttcat atacagttcat atacagttcat
540
ttggtgtacct taaggactctg cttctctctct cttctctctct cttctctctct cttctctctct
600
aaacacctgt gatctgtgct gocacagagt gcgcacccgtg aacctgtgtgt cgcacccgtg
660
cagcagatct cccacccgct cccacccgct cccacccgct cccacccgct cccacccgct
720
aacacccgct cccacccgct cccacccgct cccacccgct cccacccgct cccacccgct
780
aacacccgct cccacccgct cccacccgct cccacccgct cccacccgct cccacccgct
840
aacacccgct cccacccgct cccacccgct cccacccgct cccacccgct cccacccgct
900
tgctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
960
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
gcagctagctct gccacaccag gcacaccag gcacaccag gcacaccag gcacaccag
1020
tataaggttg tttatttaag aacaccctct atttttcgtct ctacccctac tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<table>
<thead>
<tr>
<th>Gly</th>
<th>Cys</th>
<th>Met</th>
<th>Asp</th>
<th>Gln</th>
<th>Ser</th>
<th>Val</th>
<th>Ser</th>
<th>Leu</th>
<th>Ser</th>
<th>Ile</th>
<th>Ser</th>
<th>Glu</th>
<th>Thr</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Thr</td>
<td>Ser</td>
<td>Lys</td>
<td>Leu</td>
<td>Thr</td>
<td>Phe</td>
<td>Lys</td>
<td>Glu</td>
<td>Ser</td>
<td>Met</td>
<td>Val</td>
<td>Val</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Gly</td>
<td>Lys</td>
<td>Val</td>
<td>Leu</td>
<td>Lys</td>
<td>Lys</td>
<td>Arg</td>
<td>Arg</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>Ser</td>
<td>Ile</td>
<td>Thr</td>
<td>Asp</td>
<td>Asp</td>
<td>Leu</td>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
<td>Asn</td>
<td>Asp</td>
<td>Ser</td>
<td>Glu</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>95</td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>Ile</td>
<td>Ile</td>
<td>Lys</td>
<td>Pro</td>
<td>Arg</td>
<td>Ser</td>
<td>Ala</td>
<td>Pro</td>
<td>Phe</td>
<td>Ser</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110</td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Lys</td>
<td>Tyr</td>
<td>Asn</td>
<td>Phe</td>
<td>Met</td>
<td>Arg</td>
<td>Ile</td>
<td>Ile</td>
<td>Lys</td>
<td>Tyr</td>
<td>Glu</td>
<td>Phe</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>125</td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Asp</td>
<td>Ala</td>
<td>Leu</td>
<td>Asn</td>
<td>Gln</td>
<td>Ser</td>
<td>Ile</td>
<td>Ile</td>
<td>Arg</td>
<td>Ala</td>
<td>Asn</td>
<td>Asp</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>140</td>
</tr>
<tr>
<td>Tyr</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td>Ala</td>
<td>Leu</td>
<td>His</td>
<td>Asn</td>
<td>Leu</td>
<td>Asp</td>
<td>Ala</td>
<td>Glu</td>
<td>Ala</td>
<td>Val</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>155</td>
</tr>
<tr>
<td>Phe</td>
<td>Asp</td>
<td>Met</td>
<td>Gly</td>
<td>Ala</td>
<td>Tyr</td>
<td>Lys</td>
<td>Ser</td>
<td>Ser</td>
<td>Lys</td>
<td>Asp</td>
<td>Ala</td>
<td>Lys</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>170</td>
</tr>
<tr>
<td>Thr</td>
<td>Val</td>
<td>Ile</td>
<td>Leu</td>
<td>Arg</td>
<td>Ile</td>
<td>Ser</td>
<td>Lys</td>
<td>Thr</td>
<td>Gin</td>
<td>Leu</td>
<td>Tyr</td>
<td>Val</td>
<td>Thr</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>185</td>
</tr>
<tr>
<td>Gln</td>
<td>Asp</td>
<td>Glu</td>
<td>Asp</td>
<td>Gln</td>
<td>Pro</td>
<td>Val</td>
<td>Leu</td>
<td>Leu</td>
<td>Lys</td>
<td>Glu</td>
<td>Met</td>
<td>Pro</td>
<td>Glu</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>200</td>
</tr>
<tr>
<td>Pro</td>
<td>Lys</td>
<td>Thr</td>
<td>Ile</td>
<td>Thr</td>
<td>Gly</td>
<td>Ser</td>
<td>Thr</td>
<td>Asn</td>
<td>Leu</td>
<td>Leu</td>
<td>Phe</td>
<td>Phe</td>
<td>Thr</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>215</td>
</tr>
<tr>
<td>Glu</td>
<td>Thr</td>
<td>His</td>
<td>Gly</td>
<td>Thr</td>
<td>Lys</td>
<td>Asn</td>
<td>Tyr</td>
<td>Phe</td>
<td>Thr</td>
<td>Ser</td>
<td>Val</td>
<td>Ala</td>
<td>His</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>230</td>
</tr>
<tr>
<td>Asn</td>
<td>Leu</td>
<td>Phe</td>
<td>Ile</td>
<td>Ala</td>
<td>Thr</td>
<td>Lys</td>
<td>Gln</td>
<td>Asp</td>
<td>Tyr</td>
<td>Thr</td>
<td>Trp</td>
<td>Val</td>
<td>Cys</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>245</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Pro</td>
<td>Pro</td>
<td>Ser</td>
<td>Ile</td>
<td>Thr</td>
<td>Asp</td>
<td>Phe</td>
<td>Gln</td>
<td>Ile</td>
<td>Leu</td>
<td>Glu</td>
<td>Asn</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>260</td>
</tr>
</tbody>
</table>

Ala

*SEQ ID NO: 8*

**LENGTH:** 1511

**TYPE:** DNA

**ORGANISM:** Homo sapiens

**FEATURE:**

**NAME/KEY:** misc_feature

**OTHER INFORMATION:** Incyte ID No: 1454852CB1

**SEQUENCE:**

```
ccctctctc cccagccttg tctctgtgt gcgtgctcct ccctctctcc caccctgtcg 60
cctctctcc cccagccttg tctctgtgt gcgtgctcct ccctctctcc caccctgtcg 120
gcgtgctcct cccagccttg tctctgtgt gcgtgctcct ccctctctcc caccctgtcg 180
gcgtgctcct cccagccttg tctctgtgt gcgtgctcct ccctctctcc caccctgtcg 240
ttgctgtgtt gtgtgtggtat ggcgcagcgt tgggggttgc tgatggtgcg 300
ttgctgtgtt gtgtgtggtat ggcgcagcgt tgggggttgc tgatggtgcg 360
ttgctgtgtt gtgtgtggtat ggcgcagcgt tgggggttgc tgatggtgcg 420
ttgctgtgtt gtgtgtggtat ggcgcagcgt tgggggttgc tgatggtgcg 480
ttgctgtgtt gtgtgtggtat ggcgcagcgt tgggggttgc tgatggtgcg 540
```
<table>
<thead>
<tr>
<th>Met</th>
<th>Thr</th>
<th>Thr</th>
<th>Ser</th>
<th>Ile</th>
<th>Arg</th>
<th>Gln</th>
<th>Phe</th>
<th>Thr</th>
<th>Ser</th>
<th>Ser</th>
<th>Ser</th>
<th>Ser</th>
<th>Ser</th>
<th>Ile</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Gly</td>
<td>Ser</td>
<td>Arg</td>
<td>Thr</td>
<td>Ser</td>
<td>Cys</td>
<td>Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Ser</td>
<td>Cys</td>
<td>Arg</td>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Thr</td>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Tyr</td>
<td>Ser</td>
<td>Ser</td>
<td>Cys</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Tyr</td>
<td>Ser</td>
<td>Phe</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
<td>Tyr</td>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
<td>Phe</td>
<td>Gly</td>
<td>Gly</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Val</td>
<td>Asp</td>
<td>Gly</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Gly</td>
<td>Lys</td>
<td>Ala</td>
<td>Thr</td>
<td>Met</td>
<td>Gln</td>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Asp</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>Leu</td>
<td>Asp</td>
<td>Lys</td>
<td>Val</td>
<td>Arg</td>
<td>Ala</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>Ala</td>
<td>Asn</td>
<td>Thr</td>
<td>Glu</td>
<td>Leu</td>
<td>Glu</td>
<td>Val</td>
<td>Lys</td>
<td>Ile</td>
<td>Arg</td>
<td>Asp</td>
<td>Trp</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Arg</td>
<td>Gln</td>
<td>Ala</td>
<td>Pro</td>
<td>Gly</td>
<td>Pro</td>
<td>Ala</td>
<td>Arg</td>
<td>Asp</td>
<td>Tyr</td>
<td>Ser</td>
<td>Gln</td>
<td>Tyr</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Thr</td>
<td>Ile</td>
<td>Glu</td>
<td>Leu</td>
<td>Gln</td>
<td>Asn</td>
<td>Lys</td>
<td>Ile</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td>Thr</td>
<td>Val</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>140</td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 9
<211> LENGTH: 432
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE: 
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1454852CD1

<400> SEQUENCE: 9
Asp Asn Ala Asn Ile Leu Leu Gln Ile Asp Asn Ala Arg Leu Ala
155 160 165
Ala Asp Asp Phe Arg Thr Lys Phe Glu Thr Glu Gln Ala Leu Arg
170 175 180
Leu Ser Val Glu Ala Asp Ile Asn Gly Leu Arg Arg Val Leu Asp
185 190 195
Glu Leu Thr Leu Ala Arg Ala Asp Leu Glu Met Gln Ile Glu Asn
200 205 210
Leu Lys Glu Glu Leu Ala Tyr Leu Lys Asn His Glu Glu Glu
215 220 225
Met Asn Ala Leu Arg Gly Gin Val Gly Gly Gin Ile Asn Val Glu
230 235 240
Met Asp Ala Ala Pro Gly Val Asp Leu Ser Arg Ile Leu Asn Glu
245 250 255
Met Arg Asp Gin Tyr Glu Lys Met Ala Glu Lys Asn Arg Lys Asp
260 265 270
Ala Glu Asp Trp Phe Phe Ser Lys Thr Glu Glu Leu Asn Arg Glu
275 280 285
Val Ala Thr Asn Ser Glu Leu Val Gin Ser Gly Lys Ser Glu Ile
290 295 300
Ser Glu Leu Arg Arg Thr Met Gin Ala Leu Glu Ile Glu Leu Gin
305 310 315
Ser Gin Leu Ser Met Lys Ala Ser Leu Gin Gly Gin Asn Leu Ala Glu
320 325 330
Thr Glu Asn Arg Tyr Cys Val Gin Leu Ser Gin Ile Gin Gly Leu
335 340 345
Ile Gly Ser Val Glu Glu Leu Ala Gin Leu Arg Cys Glu Met
350 355 360
Glu Gin Gin Asn Gin Glu Tyr Gin Ile Leu Leu Gin Asp Val Lys Thr
365 370 375
Arg Leu Gin Gin Glu Gin Ala Thr Tyr Arg Gin Leu Gin Leu Gin
380 385 390
Glu Aep Ala His Leu Thr Gin Tyr Lys Glu Gin Pro Gin Val Thr Thr
395 400 405
Arg Gin Val Arg Thr Ile Val Gin Ala Gin Asp Gly Lys Val
410 415 420
Ile Ser Ser Arg Gin Val Gin His Gin Thr Thr Arg
425 430

<210> SEQ ID NO 10
<211> LENGTH: 309
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<223> OTHER INFORMATION: Incyte ID No: 353005.1
<220> FEATURE:
<223> OTHER INFORMATION: unsure
<222> LOCATION: 6, 10, 18, 24-25, 67, 76, 83, 98, 159, 290
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 10

ggtggntgn accatggnc tgttnatctc cccttcttcc ctttgtgcotc tgtggacggt
60
gccttncat ctgtcnccac ttnaatctc tgtctctnga ggaactagc tccaaccagc
-continued

ggatccaggt ttctactcag atttgtaagtt ctgggctnta tgcacaacctc gtgatctctc 180
cocatgggat tgcctggctgc ctgggatgtgc ttcagcgtgg ggccgacggc aggcacaacg 240
gcgcgcgcc atgtgtgtaga gatacgccgc aatgtgtatag aatgacatnt ttgggttccc 300
agatcttt 309

<210> SEQ ID NO: 11
<211> LENGTH: 176
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 376497.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 18, 30, 35, 39, 44, 52, 87, 93, 108, 112, 114, 151, 166, 168, 170
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 11

gcccaacgtc atcgggcgcg atgtgcaccc tctgtnesana ggcantacac gcanttccac 60
gctgtcaacc tgcacgggtg gctgggcccc tntgtnaagag atgagcgggcg antnccgggt 120
tcctttcacc gcgcacgtgtc nagggtggtg agaacanntn cattca 176

<210> SEQ ID NO: 12
<211> LENGTH: 3944
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 994684.9

<400> SEQUENCE: 12

cagtctcacc ttgttttctc ccactctcct ctccccgggag gcttttctgt gcttaagacc 60
tgctgtcaagt cagttcccc caagttcgac aaccagccccc atggagattg cctotttctag 120
gcaagtgcct aagcgctgtt ttatcctttt tttgtatcagc ctgctgtttc gttatagggc 180
cagttgagag atgtggccaa ccaagcccgcc aggcttggtta tcagacccag 240
tgtggcgcag tgcgcctgcg aagttgcccc ctggctgccg acctgcggca cccctactagc 300
gagatctctg atgatagccg gattgacagg aggttggagag ccaagattga caatgttcaag 360
acacagtgcg ccactctcgc gaaagccatt gggatgtcgcg acgcagctgg ggggtgctggc 420
tcctcaaggt gcacggccac gcctggtgcc atgggaggg gctttccagg gccggcgac 480
gacagtggcc gcgtgcgtgc tggatgacag gacatgctatc agaccaagtc ggccotgtagc 540
gttgagattgc ccactttcag caagtctctg ggggggggatt gcagacagct cagttgagaa 600
ggttgggagc agcaaatcatt cgataagttcg ctttacagag ctacaccacca cgcataatt 660
atgctgtata tttgtgcaga acttgaaccttgctgtttttct tgaatcacc cttttccccat 720
cacccatatg ggttatattg acataactcg aatttaattg tgggtgttgg gagggaagg 780
acagattgag aggccagcaactgcgcttta tgcctctctctc ttaaacgggt aagccagttca 840
ttttaatttt gttgcaactt gtccgctgg gttgctgtcg taacgtagc cccacagtaacg 900
cagttctcct ccaatcgtact cttgttctcttccccaga actacacctga gactcagct 960
cccattcgata cagttctctt ctcagatagc aaccagcttc cttatcact cccacccctt 1020
tttgggact tttaataaat gcagttcttt cggggccttg ttggtacttc tggaaaaaggg 1080
gtccagtaga gttcctcaac accacagagt caaaataatg ggcctatgcag gatcagcctg
1140
gcagagtggc ttctcgagtc ctctctcttt tccccggcag tyttgcctca agcagtttct
1200
cctctgggtta tggcagtgcc agtgggatag gggggccact gagggagaggt ctgggagggc
1260
gcctgggctg agtggctggcg ggcagacgca tgtggatctgct tctctcgacg agcagtcggg
1320
gttcagctg agtggcggg ctcaggtgcc ggctgctcag cagactgtca
1380
gagggcgag ggtagggttgg gcagctgccc gggtagcagc cctccagcttc aatattgtct
1440
tccacacctct ctctctctgg gggagtctag aagacgttga acgtatctgg gccacacgctgc
1500
ttcagacgtgc agcaacacag cccatgtgga ttcctcttcg tagcagtagc agcctctgcg
1560
ctctctctct gattggttgg gattggttgg gattggttgg agctggtctgct tctctctcct
1620
atgagctgctg agctggtctgct tccagctcgg aagcacggct ctctctctct tccagctcgt
1680
ggcctgctgg gcctgctgg gcctgctgg gcctgctgg gcctgctgg gcctgctgg
1740
ctctctctct ctctctctct tccacacgct gtcaagctgg gcacacgctgc ctctctctct
1800
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
1860
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
1920
agcagcagtc acacaaaagtc gccgttgctct tccatctgat gccgctctct gccgctctct
1980
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
2040
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2100
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2160
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
2220
atgagctgctg agctggtctgct tccagctcgg aagcacggct ctctctctct tccagctcgt
2280
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
2340
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2400
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
2460
agcagcagtc acacaaaagtc gccgttgctct tccatctgat gccgctctct gccgctctct
2520
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2580
atcagacgc gcagacgcgc gcagacgcgc gcagacgcgc gcagacgcgc gcagacgcgc
2640
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2700
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2760
agcagcagtc acacaaaagtc gccgttgctct tccatctgat gccgctctct gccgctctct
2820
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
2880
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
2940
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
3000
agcagcagtc acacaaaagtc gccgttgctct tccatctgat gccgctctct gccgctctct
3060
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
3120
agcagcagtc acacaaaagtc gccgttgctct tccatctgat gccgctctct gccgctctct
3180
ctctctctct ctctctctct tccacacgctgc ctctctctct tccacacgctgc ctctctctct
3240
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
3300
tccacacgct gcacacgctgc ctctctctct tccacacgctgc ctctctctct tccacacgctg
3360
-continued

gcggagtatc tacatcatgc ttccccatcag cttcaggtt ttccccagca tgggccccgc 3420
tgacgcgaca gccgcaattt ttccccatgc tatacataca aacagaga tcgacccgca 3480
tccaatattt ttcttttgtt ctaactatct ccagaacttg ttctataaaaa tgcctttata 3540
ttat
3544

<210> SEQ ID NO: 13
<211> LENGTH: 3000
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 985610.1

<400> SEQUENCE: 13

ggagttattt ctaaacagcc tctoacagta taccggtgca tgggtggctg gtttatattta 60
atccattcgg ctccgttttt cctctctttg cctctctttc cttcctctat cttacactac 120
atccacagct cttaacgcgc ggcgctaacat acccctttctt cttccactct cccctgggac 180
cctgagcctca gcgggcttcct tttccacagct ttaccgtgcc gcgccttcgg ctgcctcgcc 240
cgggtcgcgtgc ggcgtgtggc ggggtggttt cctcggttgg gcctgcaagc ggcgcatcgc 300
ctgatcattat tccacacagaa ccagatacg gcctgtccaa cagactcgct acgagagggcg 360
agcgagagaaaa ggagagaggt tttgagaggg gcaagagaaaa agggagaggg aagagagtta 420
attgcgctg cctgctctct ctgtttttact ccttgagacta gcggctccgg ctgcctcgcc 480
gagttcttcca caaccctttcc ccaccccttc ccaccccttc ataagagcccc cggcggttcc 540
ccaaagaagc gcggcggtggg gaaaaagaaaa aagagccct cctgctaatc tcgcgcaccc 600
ggocctttact agcgcgagtt tggacagcgc cttggccccc ccaagctggtg tggacgcccc 660
cgccacgcgc gcggccgcgc ctctcctcgg cccctcgctc gtaccagaga gggagctgtc 720
ctccagagctgg ggcgggggac aagagagagc cgggctcccc gcggactagc tttgagagggc 780
ctgggaggtg ttctctagtgt gctttactag attccacagca gaggccaggg gaggctgggt 840
ggagctcggttg gggaggactg acgcggcgag gagagcctgc gcgtgggccc ctggaggaag 900
ggagatacgag cgcagagtgg ggggtagggc tttccggcag cccctccggct gttggccccag 960
ccggagctgg gcacactccgy cggagacttc gcacactctgg cccctccggcgc gcgggagccag 1020
tttgaaactcg aacctacac ccggagcaggg gacggcagct cccgagccgc ggggagcaag 1080
cttgccctctcg gcggccgagt ccggcagctg ggggagctct gaaacctcttg 1140
cttgctcgct gccgactc ctggctatgg gcagctgtgc gcagctgctg gcagctgctg 1200
gacgactcgc ctcggcxztaactac gcgcgtaaactac gcagctgcgc gcagctgcgc 1260
ggagatacgg ctggactcgg cgcgcgctcc gcagatctg cccgcgctcc gcagatctg 1320
gggtggcgcc ctgcgcgcttg gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1380
cgcgcgcttg gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1440
cctgcgcgcttg gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1500
tggagatctg gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1560
cggagacgcttcgcggctcc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1620
cgcgcgctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc gcgcgtggctgc 1680
-continued

gaagcagccc gcacccgcgc gcgcggcaca gcgtgtcgcct caaccccgac tggcgctggc 1740
gagatccagg cgccgcccccg tacgagtcga tcggcccccctc ggtgtctcctt cocacccctc 1800
tccagcagc cagctcggccc aagcctcctcg ctgcgcaaga cctccacgcgc ttcgctcctgt 1860
cctcggttct ctgcgctcct ctcgaggagct ctcgcccgcgc gcggacgcgc gcggccctctg 1920
ttgctccatga ggagaaccgc cccacaccaac gcacgacact cgagaggggc caagagacag 1980
ggagggagc ctctgggaaac agagcgcggc ctctggaacaa agctcagagt 2040
cctggttcacc ttcgctggga ggcacagcga acacccctca cagcccaactgt gtcctcaaga 2100
gggtgcaacgt cccacacact caacagaacct acgggacggc ctcctccact cggagagact 2160
atctgtgtcg cagagggctc aagttgagga gttctgcaag ctgcagagcc ctcgcaacc 2220
cacgaaagtc cccacagcgc aagtcctcgcg acacccgagga caagtgcaag acgcacgcac 2280
acacgcgcttt gcgcgccgac agggagacg agctaaaaac acgtttttttt gctctgctg 2340
acacagttcc gcgcgggacg gcaagagaaa cgcggcccaca ggtcttttat cttttaaaac 2400
cgccagata ctcctctgctg gtcacagcag aggacacaa gcctttttct gcagacgact 2460
ttgcaagacg agccgcagag caacggaactc acagcccagc aacacaggct 2520
gtgaagaaaa gcggagagaa aagattcctt ctcacagaaaa ctgtccgcag acacacatct 2580
gaactcgttt caaatgacat afrcaaaagc acctcacaac ctgtgctgag ctgtgagact 2640
ggagagttta gcgtataactt aacatgctgct aacatggacct tgggcctaa aaagaactttt 2700
ttaccttcgc cttctttcctt tttctcttacct cgaatgtgcttt ctttttaaaac 2760
aatattaaga tttcacaact gttctttcgtc aataatgctc tttaaatgat aataaacttta 2820
atataaaactgt tataacggtt acaagagaatt tcaacattcag tataagctctcct tttttat 2880
atgatcactc aacacctttt tttttttttt aacacacctttc gttcttttac agtcttttctt 2940
ttttattttg tttttttaataaacta cttggacaata ttcctttggg ccacaaaaaa 3000

<210> SEQ ID NO 14
<211> LENGTH: 427
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Inocyte ID No: 417119.1
<400> SEQUENCE: 14

aaacaaagca sgcttgaaac caacagcaacct cctctcttta gttgaactta ttctttaaga 60
agatcctgtg gggtctcgac atttgccacct cctctcactct ttcctcgaggg ctttccactt 120
acatctggtc ccagatgctcg taccctctct ctgccgaagt gttatgcaca tcaacatttctt 180
acacttaagtt cctgctttct gcgcagagct gagcaagtac gacagctgacg tccgggaccc 240
ttgctcaatg tcgctgccgt ctgtttttctt ccaacccaccc tccacgcccttc acctctacttct 300
tgctgcggcgc gatcaaggatc gttttttttcaaatagcctcg ttacaccaaat ccaatctttttgt 360
caccctgcttc ctcgacggcgc tgcctctctgt gttttttttc taccaccaagt gtcctcact 420
ttcgaa

<210> SEQ ID NO 15
<211> LENGTH: 4108
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Inocyte Id No: 3615080CB1

<400> SEQUENCE: 15

tgcaagtct tcgacaccc cgctcggtyt ggctcaaccc aagcagggag tacccaggya 60
tgaggctgt tttaaacc cccttgagct cgggggagac cccacaccct aagaggagct 120
ttcgccgct ccggctgaag aacgcagcag ccaacacgca gggaggtcct ttctcgagga 180
tcaccacct gcgtgaagat gacacaccc ttcctcttgt gtgtccagct ggcggcagtc 240
tgctgtgccc aaccgcagct ctccgctggag gatgtcactc cagatgtgctt gggaggtcct 300
gatggagga ccctcccttcc ccgaatcttc tttactgtttg gcgtgagac cccagacacc 360
tactgccgct agtactcgtag cgggctgact gactgtgcct caggtgcgccttgaggtctgctt 420
cacaactact acaggtacac ccgtaggaact gttgtctccatt cttccgcgcc cctcctctgg 480
tgctgctcccc ccacactgct gcgcctgtgc ttcgggtcct cgggaccctg caggtctgctt 540
cgccttacag ggactgtcgag ggagcccgtc gcgcctgtgc cttctcgctgcc ggtggaccc 600
cgcctcctag acgccagtag gcgcctgtgc cttccgcgcc ggtggaccc 660
tcaccacct gcgtggctcg cccgggctgg ccctcgagct gcggagagtc tgcctgtgctt 720
tccggctcc acgcgcgtgc ccagcagcag cccacacgca gggaggtcct ttctcgagga 780
gatgtcactc cttccgcgcc ggtggaccc 840
cacaactact acaggtacac ccgtaggaact gttgtctccatt cttccgcgcc cctcctctgg 900
cgccttacag ggactgtcgag ggagcccgtc gcgcctgtgc cttctcgctgcc ggtggaccc 960
cgcctcctag acgccagtag gcgcctgtgc cttccgcgcc ggtggaccc 1020
tcaccacct gcgtggctcg cccgggctgg ccctcgagct gcggagagtc tgcctgtgctt 1080
cacaactact acaggtacac ccgtaggaact gttgtctccatt cttccgcgcc cctcctctgg 1140
gatgtcactc cttccgcgcc ggtggaccc 1200
tcaccacct gcgtggctcg cccgggctgg ccctcgagct gcggagagtc tgcctgtgctt 1260
gatgtcactc cttccgcgcc ggtggaccc 1320
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1380
tcaccacct gcgtggctcg cccgggctgg ccctcgagct gcggagagtc tgcctgtgctt 1440
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1500
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1560
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1620
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1680
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1740
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1800
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1860
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1920
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 1980
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 2040
acgccttcct ccagcgcctag gcgcctgtgc cttccgcgcc ggtggaccc 2100
cggcgcttc ctgcggcccc cgccgctcca gcggcggagg tggctccagg ggcgcctgcc 2160
actctctccc ccggcgccaa totcggcggc gtggctctgg agctggccct ggcggaggg 2220
aaggtgtccc ttggcaagaa ccctggagag tttggccgaa gctggcaasttg ttctcatctt 2280
atctatgga gcagggagga gcaagtggaa aaaaactagc gttgagatcc tcggcggagg 2340
ctttgctgct gtaagcagac ctaagcagct ctggccgcaag ggctgctgcag gttctcggc 2400
agcggcgtgc ttggagccac ggcggcagcc agcggcagag cgccgagagg gccggagagt 2460
cggagggac ggagggagac cccagcgagc ccaggaagtttg cgcccgctag cgtggaaggt 2520
tctctcttc gcctgguagt gcccagcctc aacaagagct cggcgagagc cagcgcag 2580
gccgggcac ccatactagt ccttgtgctag ctatgtcctcc cgaccaagttgc cagcggcattg 2640
gtccggccgcc gcagggggtgt cttccggcag gcgggtgggg gctctcggctag gcgggggagg 2700
gttggctgag cagtcggcggg cttcataagg gcacgctcag gggagggggc gatgatagg 2760
gcgcggcagag aatgctgctc acacagttcag tcgggtgccg cggcgctgca ggcggctgtg 2820
gcggaggact ttcaacccac cccgcagact gctgagccag ctatccagag gttggaagg 2880
gcgcggagc cttgcctcttc ggcggctagc acgcctagtg tttggcggagc gatgaaagag 2940
atccggccag ttgcggcagt caggtcggcag gcggcagaa cgggtgggtgc caggtcggcag 3000
atcgcggagc ttggcgcagt caggtcggcag gcggcagaa cgggtgggtgc caggtcggcag 3060
caacctgggc gcggcggcag cgggtgggtgc caggtcggcag gcggcagaa cgggtgggtgc 3120
gcggagggat gcggggcggc gcgggagggat gcggggcggc gcgggagggat gcggggcggc 3180
caggggctcc cgccgggcgtc ggctggtggt gcgaggaggt ggaatcgggt ggggagggat 3240
cgggttggtc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3300
cgggttggtc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3360
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3420
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3480
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3540
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3600
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3660
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3720
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3780
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3840
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3900
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 3960
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 4020
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 4080
gggtgggctc ctaagctggc atggcggcag gcggagggcg tgggtgcagc cggagggcggc 4100

<210> SEQ ID NO: 16
<211> LENGTH: 1172
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3615080CD1
<400> SEQUENCE: 16
Met Arg Pro Phe Phe Leu Leu Cys Phe Ala Leu Pro Gly Leu Leu 1 5 10 15
His Ala Gln Gln Ala Cys Ser Gly Ala Cys Tyr Pro Pro Val 20 25 30
Gly Asp Leu Leu Val Gly Arg Thr Arg Phe Leu Arg Ala Ser Ser 35 40 45
Thr Cys Gly Leu Thr Lys Pro Glu Thr Tyr Cys Thr Gln Tyr Gly 50 55 60
Glu Trp Gln Met Lys Cys Cys Asp Ser Arg Gln Pro His 65 70 75
Aan Tyr Tyr Ser His Arg Val Gln Val Ala Ser Ser Ser Gly 80 85 90
Pro Met Arg Trp Trp Gln Ser Gln Asp Val Aan Pro Val Ser 95 100 105
Leu Gln Leu Aep Leu Aep Arg Arg Phe Gln Leu Gln Glu Val Met 110 115 120
Met Glu Phe Glu Gly Pro Met Pro Ala Gly Met Leu Ile Glu Arg 125 130 135
Ser Ser Asp Phe Gly Lys Thr Trp Arg Val Tyr Gln Tyr Leu Ala 140 145 150
Ala Aep Cys Thr Ser Thr Phe Pro Arg Val Arg Gln Gly Arg Pro 155 160 165
Gln Ser Trp Gln Asp Val Arg Cys Gln Ser Leu Pro Gln Arg Pro 170 175 180
Aas Ala Arg Leu Aen Gly Gly Lys Val Gln Leu Aen Leu Met Asp 185 190 195
Leu Val Ser Gly Ile Pro Ala Thr Gln Ser Gln Lys Ile Gln Glu 200 205 210
Val Gly Glu Ile Thr Aen Leu Arg Val Aen Phe Thr Arg Leu Ala 215 220 225
Pro Val Pro Gln Arg Gly Tyr His Pro Pro Ser Ala Tyr Tyr Ala 230 235 240
Val Ser Gln Leu Arg Leu Gln Gly Ser Cys Phe Cys His Gly His 245 250 255
Ala Aep Arg Cys Ala Pro Lys Pro Gly Ala Ser Ala Gly Pro Ser 260 265 270
Thr Ala Val Gln Val His Asp Val Cys Val Cys Gln His Aen Thr 275 280 285
Ala Gly Pro Aen Cys Glu Arg Cys Ala Pro Phe Tyr Aen Aen Arg 290 295 300
Pro Trp Arg Pro Ala Gln Gly Aep Ala His Gln Gly Glu Arg 305 310 315
Cys Aep Cys Aen Gly His Ser Glu Thr Cys His Phe Aep Pro Ala 320 325 330
Val Phe Ala Ala Ser Gln Gly Ala Tyr Gly Gly Val Cys Aen 335 340 345
Cys Arg Aep His Thr Glu Gly Lys Aen Cys Glu Arg Cys Gln Leu 350 355 360
His Tyr Phe Arg Aen Arg Arg Pro Gly Ala Ser Ile Gln Glu Thr
-continued-

365 370 375

Cys Ile Ser Cys Glu Cys Asp Pro Asp Gly Ala Val Pro Gly Ala 380 385 390

Pro Cys Asp Pro Val Thr Gly Gln Cys Val Cys Lys Glu His Val 395 400 405

Gln Gly Glu Arg Cys Asp Leu Cys Lys Pro Gly Phe Thr Gly Leu 410 415 420

Thr Tyr Ala Asn Pro Gln Gly Cys His Arg Cys Asp Cys Asn Ile 425 430 435

Leu Gly Ser Arg Arg Asp Met Pro Cys Asp Glu Glu Ser Gly Arg 440 445 450

Cys Leu Cys Leu Pro Asn Val Val Gly Pro Lys Cys Asp Gln Cys 455 460 465

Ala Pro Tyr His Trp Lys Leu Ala Ser Gly Gln Gly Cys Glu Pro 470 475 480

Cys Ala Cys Asp Pro His Asn Ser Leu Ser Pro Gln Cys Asn Gln 485 490 495

Phe Thr Gly Gln Cys Pro Cys Arg Glu Gly Phe Gly Gly Gly Leu Met 500 505 510

Cys Ser Ala Ala Ala Ile Arg Cys Gln Cys Pro Asp Arg Thr Tyr Gly 515 520 525

Asp Val Ala Thr Gly Cys Arg Ala Cys Asp Asp Phe Arg Gly 530 535 540

Thr Glu Gly Pro Gly Cys Asp Lys Ala Ser Gly Arg Cys Leu Cys 545 550 555

Arg Pro Gly Leu Thr Gly Pro Arg Cys Asp Glu Gln Arg Cys Arg Gly 560 565 570

Tyr Cys Asn Arg Tyr Pro Val Cys Val Ala Cys His Pro Cys Phe 575 580 585

Gln Thr Tyr Asp Ala Asp Leu Arg Glu Gln Ala Leu Arg Phe Gly 590 595 600

Arg Leu Arg Asn Ala Thr Ala Ser Leu Trp Ser Gly Pro Gly Leu 605 610 615

Glu Asp Arg Gly Leu Ala Ser Arg Ile Leu Asp Ala Lys Ser Lys 620 625 630

Ile Gln Ile Arg Ala Val Leu Ser Ser Pro Ala Val Thr Glu 635 640 645

Gln Glu Val Ala Gln Val Ala Ser Ala Ile Leu Ser Leu Arg Arg 650 655 660

Thr Leu Gln Gly Leu Gln Leu Asp Leu Pro Leu Glu Glu Glu Thr 665 670 675

Leu Ser Leu Pro Arg Asp Leu Glu Ser Leu Asp Arg Ser Phe Asn 680 685 690

Gly Leu Leu Thr Met Tyr Gln Arg Lys Arg Glu Gln Phe Gly Lys 695 700 705

Ile Ser Ser Ala Asp Pro Ser Gly Ala Phe Arg Met Leu Ser Thr 710 715 720

Ala Tyr Glu Gln Ser Ala Gln Ala Ala Gln Gln Val Ser Asp Ser 725 730 735

Ser Arg Leu Leu Asp Glu Leu Arg Asp Ser Arg Arg Glu Ala Glu 740 745 750
Arg Leu Val Arg Gln Ala Gly Gly Gly Gly Gly Gly Thr Gly Ser Pro
755 760 765
Lys Leu Val Ala Leu Arg Leu Met Ser Ser Leu Pro Asp Leu
770 775 780
Thr Pro Thr Phe Asn Lys Leu Cys Gly Asn Ser Arg Gln Met Ala
785 790 795
Cys Thr Pro Ile Ser Cys Pro Gly Leu Cys Pro Gln Asp Asn
800 805 810
Gly Thr Ala Cys Gly Ser Arg Cys Arg Gly Val Leu Pro Arg Ala
815 820 825
Gly Gly Ala Phe Leu Met Ala Gly Gln Val Ala Glu Leu Arg
830 835 840
Gly Phe Asn Ala Gln Leu Gln Arg Thr Arg Gln Met Ile Arg Ala
845 850 855
Ala Glu Glu Ser Ala Ser Gln Ile Gln Ser Ser Ala Glu Arg Leu
860 865 870
Glu Thr Gln Val Ser Ala Ser Gln Met Gln Glu Asp Val
875 880 885
Arg Arg Thr Arg Leu Leu Gln Gln Val Arg Asp Phe Leu Thr
890 895 900
Asp Pro Asp Thr Asp Ala Thr Ile Gln Glu Val Ser Glu Ala
905 910 915
Val Leu Ala Leu Trp Leu Pro Thr Asp Ser Asp Thr Val Leu Gln
920 925 930
Lys Met Asn Gln Ile Gln Ala Ala Arg Leu Pro Asn Val
935 940 945
Asp Leu Val Leu Ser Gln Thr Lys Gln Asp Ile Ala Arg Ala Arg
950 955 960
Arg Leu Gln Ala Glu Ala Glu Ala Arg Ser Arg Ala His Ala
965 970 975
Val Gly Gly Gln Val Gln Asp Asp Val Glu Arg Gln Gly
980 985 990
Thr Val Ala Leu Gln Ala Gln Asp Thr Met Gln Gly Thr Ser
995 1000 1005
Arg Ser Leu Arg Leu Ile Gln Asp Arg Val Ala Glu Val Gln Gln
1010 1015 1020
Val Leu Arg Pro Ala Glu Lys Leu Val Thr Ser Met Thr Lys Gln
1025 1030 1035
Leu Gly Asp Phe Trp Thr Arg Met Gln Leu Arg His Gln Ala
1040 1045 1050
Arg Glu Gly Ala Gln Ala Gln Ala Glu Gly Gly Ala Val Gln Ala Glu
1055 1060 1065
Gly Ala Ser Glu Gln Ala Leu Ser Ala Gln Gly Phe Glu Arg
1070 1075 1080
Ile Lys Gln Lys Tyr Ala Glu Leu Asp Arg Leu Gly Gln Ser
1085 1090 1095
Ser Met Leu Gly Gln Gly Ala Arg Ile Gln Ser Val Lys Thr
1100 1105 1110
Glu Ala Glu Glu Leu Phe Gly Gly Thr Met Glu Met Met Asp Arg
1115 1120 1125
-continued

<table>
<thead>
<tr>
<th>Met Lys Asp Met Glu Leu Glu Leu Leu Arg Gly Ser Gln Ala Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1130 1135 1140</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Met Leu Arg Ser Ala Asp Leu Thr Gly Leu Glu Lye Arg Val Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td>1145 1150 1155</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gln Ile Arg Asp His Ile Aen Gly Arg Val Leu Tyr Tyr Ala Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1160 1165 1170</td>
</tr>
</tbody>
</table>

Cys Lys

<210> SEQ ID NO 17
<211> LENGTH: 795
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 331749.3

<400> SEQUENCE: 17

```
atttcata ctctgtatg ttattgcatt cctgggtagc aatctggttgc gcaccctata
  60
aattcagcct ttatctattg gccaacaaat ttgctattt cagacaatt ggtatactaa
  120
agactaagtg gcccataaggg ttgctaaat gttgggaaattg ttcacctacat cagacaatt
  180
aattcataatg cggctgagcg cttcttccag tttgggcaatg tggctaatg aatctggttgc
  240
ccttggtaggc tgaattcagc gagccttaag gggtggttagc aatctggttgc
  300
taatatatga cttctctacag tttggttcttt cccacactcc cccacactcc
  360
tcctagtct gctggttctgt gctggttctgt gcactggttgc aatctggttgc
  420
ttggctctgg ctggtggttctgt gcactggttgc aatctggttgc
  480
tgtatataag tttggttctgt gcactggttgc aatctggttgc
  540
taatatatga cttctctacag tttggttcttt cccacactcc cccacactcc
  600
```

<210> SEQ ID NO 18
<211> LENGTH: 2538
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 979243.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 1479-1784, 1933-2000, 2002
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 18

```
cgcgcagcgg cgcgcagcgg cgcgcagcgg cgcgcagcgg cgcgcagcgg cgcgcagcgg cgcgcagcgg
  60
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  120
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  180
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  240
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  300
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  360
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  420
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  480
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  540
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  600
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  660
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  720
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  780
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  840
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  900
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
  960
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1020
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1080
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1140
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1200
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1260
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1320
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1380
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1440
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1500
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1560
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1620
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1680
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1740
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1800
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1860
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1920
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 1980
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2040
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2100
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2160
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2220
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2280
tgtagagctgg tgggtggttctgt gcactggttgc aatctggttgc
 2340
```
aacttgcggg ctgccaacc cctgtaactg gctcagtgag cgtgagaccc cttgaaggg 420
cctggyccgt gcagaaagcc atctccacca ctttigagag gcagatcctc acacacgca 480
cctgtgctca cttcagactc acagagcagg gctoaactct gacaggtgct cagaggaag 540
tgctccttgg gcgcattcag cctcctcacc cctcctgttt ctgtggtatt gcctgaacag 600
accgggaagt gcagaagact tgcacacccc ctggtacttt tgggtgggtg gcgcacagac 660
agacacagcc tcagagagac gtatgcaacc ttcttgggaat gatagcactgt gcagagccag 720
cctcagacgt cactgctcct cttgtacgtgc ctgtcagagga ggcagaaaggt agtgttagggg 780
gagcactgtc gttcagcattt ccaaccacctt caggggtctcg cttgaggcac ccctcctcacc 840
cctgcaaggg gttggtgtgg taggcatatt gcacagcaacct tcatatggag TAGGGATT 900
ggcattagcc tgcacaccc gagcagcttg gtttagtcag ctcagacagt caacacggacc 960
ccggggyggg cagttctccc cttacataag gcagacacctgcag gacagtactg acgtcttct 1020
agacggtcct ggcactttga acatgtgtct ccatactacct gaaagcagctc cctcctggaa 1080
gcaccgccca tggataatgc cttcagaggac cctctgcccact cttcagcctg ggcctgtccc 1140
acacacaggt ctctcagaca gacgtgcgac cccgtttcgc tggcagacag caggtgtctg 1200
cgcctgacct ccggggtctcgt gttgtcagac ccgtggattcg tcagnacagct actaacacacc 1260
agcagtacct tttctcccatt tctgttctct tctgtttcttt ccacactctcctt ccttctttct 1320
gctgttcttt ccagacgtcag cctgacgccc ctagctgggtgt gttagtatta 1380
ggcaagcata cggggagggcct ttgcttccag ctcagacggt cagctgtaac 1440
cctcttgacag tggatagcagc cccctctctt cttctcmmnn nnnnnnnnnn nnnnnnnnnn 1500
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1560
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1620
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1680
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1740
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1800
agattcatcag ccaggtctgcc atctggttgc accttgcagc ggcgyggggcag 1860
agacaaaggg ccattaattgc ttagttcgct tcaagctcttg tttgtcag gctttccocca 1920
acggccgctc cmmnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 1980
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 2040
acgggggaga gacagagact caagagtcga acagggcccct ttcctaatag gcgtggagag 2100
agcgggtcct tattcaccacc ccaccccccct ccaccccccct ccagcgttcc agaaggtgaa 2160
acacattctc cccagggggt ccccttttccc cttgagacccc cttacagccacgcg 2220
gttcagggcag cccctattgc ccaggtctgg cccagcctgcg tggctggcagc 2280
cggccctttct cccagacggcct ccagatcagc cccagcctgcg ctgtgactagcctcttc 2340
ccggtggccc ctcggagagc ctggagccaccc cctaccccccct cccagcgttcc gcgtggagag 2400
gcggccaaagt gttgagactgc tggctggccct ccaccccccct cccagcctgcg 2460
cctgctgctgt cgggggtgct ccagaggtcct cttcagaggt ctcagaggtcct 2520
gttgggttct ggcacacagc 2538
---continued

`<210>` SEQ ID NO 19
`<211>` LENGTH: 1730
`<212>` TYPE: DNA
`<213>` ORGANISM: Homo sapiens
`<220>` FEATURE:
`<221>` NAME/KEY: misc_feature
`<223>` OTHER INFORMATION: Incoy ID No: 3189059C81

`<400>` SEQUENCE: 19

ggggccccgc ggtatcccaac ccagcccaac ccccccgggc gggaccgctg cagttgacct 60
ggatcccccg acgccggccc gcacgcaagc gttctctcct ctactttcgg ggtggtttgt 120
gcctgoggq gacgctgcgc ttgaaaaccc ttcgctgtgga cccccggcgc cggatgcgcc 180
ggtggggct tggggcgccc gctggacgct gttggggact cgcctcggcg tgtagccgacat 240
tttctctccg aggcctcacc ttcagctgct tttggggttgg agctgctcagttctcgctca 300
tgcgaaaga cttggttctcg cagtgcagct ttgcgaccc agttccacgg ttggaagata 360
ccagctgccg atggcgagct tccaccaagtt gttccagca gcacggtgccc aactcacaacc 420
catccatggg gctgtccttg ctggtctctc acocgcagcg cttggaattt ctcgccgcttt 480
cctctccgta ccagcttacgt cgcctctcgc gnctgggct ctgcggatgq ggctgctcaca 540
ttgtcagtt gcacacttct ctcacgggca atcaggaagac cagtcataact ctcagctgga 600
tgcggacaac cccactaatt atagaggta cccagctgct gttcgaagcc aagaaaggggg 660
agctgacaa gccctgcggct gcacactgca ctccagcaac ttttcggtc gctcgagcttg 720	
tatcctgaga aacctgagtt aacsgtaggg cccccagctg gggacactcc ggacacccca 780
tgcggacagtc agcctcacgc gctgtgcgc cagcagggga gccacacaag 840
agctcctgag ctgcagcttc aactcacaagtt ttcaggaacg cttcctctca 900
agctgagctg tgcggatcagtt tgcacacttg agggtttag gctggcacag tttgcagctc 960
ggatgagctc gctggtccat tgcagactgt ctttcccagc gttgtgaggg ccgacacaca 1020
ggacaccgct aatagggctct ccccagaggg cttgagaggg ccagacacact gttagcaagt 1080
tcagggacc atacactaat acctgtgcag gcctaaact ctttgcaggg ccaacacca 1140
tgcgtaaacg ctgcggccag ggtgagctca atacacaagtt ctctggctct ccgctgcgct 1200
tccgagacac gtcggcgccc ggcgagggc gcacagcgcc cttactttgg ggtggtgggg 1260
ggacactgt gcctgggttg cttctggtcg gggatcgctgt gttgcgctgcgc 1320
ggacaccctt tcaagggcag tcaagagcgg ctatggcgacg gttcagcagaa 1380
agccgccttc cccccggacc cacccgaccc cttcggcgagct cccgagacaag 1440
cgcgccgcag gcagcagcgc gcgccacagt gttggggact ccagggagag cggagggagg 1500
agggcggctt gggcgggagcc gcacgagcag aggtgagggcccc agaattccag 1560
agccgcgcag gcgctcctgtct gtcgctgact gccagctccag gcgacgctcag ggtgctacag 1620
ggcaccggctttggtcct cagctcgcct gccctgtcgt ggtgcttagc 1680
ttttctcaga ggcaggtgctttcgcatc aagagggctg tgcggctag 1730
<table>
<thead>
<tr>
<th>Seq</th>
<th>Amino Acid</th>
<th>Amino Acid</th>
<th>Amino Acid</th>
<th>Amino Acid</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Met</td>
<td>Ala</td>
<td>Arg</td>
<td>Met</td>
<td>Gly</td>
</tr>
<tr>
<td>5</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
<td>Ala</td>
</tr>
<tr>
<td>10</td>
<td>Gly</td>
<td>Arg</td>
<td>Met</td>
<td>Trp</td>
<td>Trp</td>
</tr>
<tr>
<td>15</td>
<td>Gly</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td>Phe</td>
</tr>
<tr>
<td>20</td>
<td>Gly</td>
<td>Thr</td>
<td>Phe</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>25</td>
<td>Gly</td>
<td>Val</td>
<td>Gly</td>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td>30</td>
<td>Asp</td>
<td>Ser</td>
<td>Met</td>
<td>Tyr</td>
<td>Gly</td>
</tr>
<tr>
<td>35</td>
<td>Tyr</td>
<td>Gly</td>
<td>Phe</td>
<td>Ile</td>
<td>Gly</td>
</tr>
<tr>
<td>40</td>
<td>Gly</td>
<td>Thr</td>
<td>Ile</td>
<td>Ser</td>
<td>Phe</td>
</tr>
<tr>
<td>45</td>
<td>Thr</td>
<td>Met</td>
<td>Ala</td>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>50</td>
<td>Phe</td>
<td>Ala</td>
<td>Arg</td>
<td>Val</td>
<td>Pro</td>
</tr>
<tr>
<td>55</td>
<td>Gly</td>
<td>Ala</td>
<td>Val</td>
<td>Pro</td>
<td>Ser</td>
</tr>
<tr>
<td>60</td>
<td>Val</td>
<td>Pro</td>
<td>Ser</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>65</td>
<td>Lys</td>
<td>Ile</td>
<td>Thr</td>
<td>Glu</td>
<td>Val</td>
</tr>
<tr>
<td>70</td>
<td>Thr</td>
<td>Trp</td>
<td>Glu</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>75</td>
<td>Glu</td>
<td>Lys</td>
<td>Ser</td>
<td>Thr</td>
<td>Asn</td>
</tr>
<tr>
<td>80</td>
<td>Asn</td>
<td>Val</td>
<td>Ala</td>
<td>Ile</td>
<td>Tyr</td>
</tr>
<tr>
<td>85</td>
<td>Tyr</td>
<td>Asn</td>
<td>Ser</td>
<td>Met</td>
<td>Gly</td>
</tr>
<tr>
<td>90</td>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
<td>Val</td>
<td>Val</td>
</tr>
<tr>
<td>95</td>
<td>Ala</td>
<td>Pro</td>
<td>Tyr</td>
<td>Arg</td>
<td>Glu</td>
</tr>
<tr>
<td>100</td>
<td>Tyr</td>
<td>Arg</td>
<td>Glu</td>
<td>Val</td>
<td>Phe</td>
</tr>
<tr>
<td>105</td>
<td>Glu</td>
<td>Val</td>
<td>Arg</td>
<td>Pro</td>
<td>Ser</td>
</tr>
<tr>
<td>110</td>
<td>Arg</td>
<td>Leu</td>
<td>Arg</td>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td>115</td>
<td>Leu</td>
<td>Ser</td>
<td>Arg</td>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td>120</td>
<td>Glu</td>
<td>Leu</td>
<td>Asp</td>
<td>Glu</td>
<td>Gly</td>
</tr>
<tr>
<td>125</td>
<td>Val</td>
<td>Tyr</td>
<td>Ile</td>
<td>Cys</td>
<td>Glu</td>
</tr>
<tr>
<td>130</td>
<td>Cys</td>
<td>Glu</td>
<td>Phe</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td>135</td>
<td>Glu</td>
<td>Pro</td>
<td>Thr</td>
<td>Gly</td>
<td>Asn</td>
</tr>
<tr>
<td>140</td>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
<td>Arg</td>
<td>Thr</td>
</tr>
<tr>
<td>145</td>
<td>Gln</td>
<td>Leu</td>
<td>Thr</td>
<td>Met</td>
<td>Ala</td>
</tr>
<tr>
<td>150</td>
<td>Leu</td>
<td>Thr</td>
<td>Val</td>
<td>Met</td>
<td>Ala</td>
</tr>
<tr>
<td>155</td>
<td>Glu</td>
<td>Thr</td>
<td>Gln</td>
<td>Ala</td>
<td>Val</td>
</tr>
<tr>
<td>160</td>
<td>Thr</td>
<td>Gln</td>
<td>Ala</td>
<td>Val</td>
<td>Leu</td>
</tr>
<tr>
<td>165</td>
<td>Ala</td>
<td>Pro</td>
<td>Ser</td>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td>170</td>
<td>Pro</td>
<td>Ser</td>
<td>Ala</td>
<td>Lys</td>
<td>Pro</td>
</tr>
<tr>
<td>175</td>
<td>Ala</td>
<td>Lys</td>
<td>Gly</td>
<td>Ala</td>
<td>Arg</td>
</tr>
<tr>
<td>180</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>Glu</td>
</tr>
<tr>
<td>185</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>Glu</td>
</tr>
<tr>
<td>190</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>Glu</td>
<td>Leu</td>
</tr>
<tr>
<td>195</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>Arg</td>
</tr>
<tr>
<td>200</td>
<td>Pro</td>
<td>Gly</td>
<td>Asp</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>205</td>
<td>Gly</td>
<td>Thr</td>
<td>Pro</td>
<td>Met</td>
<td>Ala</td>
</tr>
<tr>
<td>210</td>
<td>Pro</td>
<td>Ala</td>
<td>Pro</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>215</td>
<td>Arg</td>
<td>Tyr</td>
<td>Leu</td>
<td>Val</td>
<td>Pro</td>
</tr>
<tr>
<td>220</td>
<td>Arg</td>
<td>Ser</td>
<td>Asp</td>
<td>Arg</td>
<td>Ala</td>
</tr>
<tr>
<td>225</td>
<td>Ala</td>
<td>Cys</td>
<td>Ile</td>
<td>Val</td>
<td>Asn</td>
</tr>
<tr>
<td>230</td>
<td>Cys</td>
<td>Leu</td>
<td>Arg</td>
<td>Met</td>
<td>Asp</td>
</tr>
<tr>
<td>235</td>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
<td>Lys</td>
</tr>
<tr>
<td>240</td>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Val</td>
<td>Glu</td>
</tr>
<tr>
<td>245</td>
<td>Thr</td>
<td>Leu</td>
<td>Val</td>
<td>Glu</td>
<td>Val</td>
</tr>
<tr>
<td>250</td>
<td>Ala</td>
<td>Gly</td>
<td>Arg</td>
<td>Met</td>
<td>Asp</td>
</tr>
<tr>
<td>255</td>
<td>Asp</td>
<td>Val</td>
<td>Lys</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td>260</td>
<td>Lys</td>
<td>Ala</td>
<td>Asp</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td>265</td>
<td>Ala</td>
<td>Pro</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td>270</td>
<td>Gly</td>
<td>Leu</td>
<td>Thr</td>
<td>His</td>
<td>Trp</td>
</tr>
<tr>
<td>275</td>
<td>Leu</td>
<td>Asn</td>
<td>Gly</td>
<td>Ser</td>
<td>Lys</td>
</tr>
<tr>
<td>280</td>
<td>Lys</td>
<td>Val</td>
<td>Ala</td>
<td>Gln</td>
<td>Asn</td>
</tr>
<tr>
<td>285</td>
<td>Leu</td>
<td>Phe</td>
<td>Pro</td>
<td>Tyr</td>
<td>Thr</td>
</tr>
<tr>
<td>290</td>
<td>Phe</td>
<td>Gly</td>
<td>Pro</td>
<td>Ile</td>
<td>Asn</td>
</tr>
<tr>
<td>295</td>
<td>Gly</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Thr</td>
</tr>
<tr>
<td>300</td>
<td>Ala</td>
<td>Cys</td>
<td>Glu</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td>305</td>
<td>Cys</td>
<td>Glu</td>
<td>Ala</td>
<td>Pro</td>
<td>Ala</td>
</tr>
<tr>
<td>310</td>
<td>Ala</td>
<td>Gly</td>
<td>Thr</td>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td>315</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
<td>Arg</td>
</tr>
<tr>
<td>320</td>
<td>Glu</td>
<td>Val</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td>325</td>
<td>Val</td>
<td>Ala</td>
<td>Thr</td>
<td>Pro</td>
<td>Ser</td>
</tr>
<tr>
<td>330</td>
<td>Ala</td>
<td>Asn</td>
<td>Thr</td>
<td>Ala</td>
<td>Arg</td>
</tr>
<tr>
<td>335</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td>340</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ile</td>
<td>Gly</td>
</tr>
<tr>
<td>345</td>
<td>Ala</td>
<td>Gly</td>
<td>Arg</td>
<td>Val</td>
<td>Pro</td>
</tr>
<tr>
<td>350</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
<td>Pro</td>
<td>Thr</td>
</tr>
<tr>
<td>355</td>
<td>Ala</td>
<td>Ile</td>
<td>Ile</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>360</td>
<td>Val</td>
<td>Ala</td>
<td>Gly</td>
<td>Ser</td>
<td>Ile</td>
</tr>
</tbody>
</table>
Val Val Ala Leu Arg Arg Arg Arg His Thr Phe Lys Gly Asp Tyr
Ser Thr Lys His Val Tyr Gly Asn Gly Tyr Ser Lys Ala Gly
Ile Pro Glu His His Pro Met Ala Glu Asn Leu Glu Tyr Pro
Asp Asp Ser Asp Asp Glu Lys Ala Gly Pro Leu Gly Gly Ser
Ser Tyr Glu Glu Glu Glu Glu Gly Gly Gly Gly Glu Asp Glu Glu
Glu Arg Lys Val Gly Gly Pro His Pro Lys Tyr Asp Glu Asp Ala
Lys Arg Pro Tyr Phe Thr Val Asp Glu Ala Glu Ala Arg Glu Aap
Gly Tyr Gly Asp Arg Thr Leu Gly Tyr Gin Tyr Asp Pro Glu Gin
Leu Asp Leu Ala Glu Asn Met Val Ser Gin Asn Asp Gly Ser Phe
Ile Ser Leu Lys Glu Trp Tyr Val

<210> SEQ ID NO 21
<211> LENGTH: 1444
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1650519CB1

<400> SEQUENCE: 21

ggagaatttt aasaggctcc caaaggggac atacgaagag gataaactacc
365
tgctgctgta ggggagagtt ttggtttttta tcggagcttaa gaaatactccag ttcocagcog
370
acottagttt caaaggttat ttoattaattag agaacagaaa cttgatatca cttgagacag
375
ccttattgaa cagotgcttg 96.O
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 965
acggcagtca cctattgaa cagotgcttg 970
tttgtggc.cg to citggtgat totgatcgga t gttacatag ccatctgctc 975
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 980
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 985
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 990
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 995
tttgtggc.cgt cctcagtctg tttgtggc.tg cttggtgat totgatcgga t gttacatag ccatctgctc 1000
aaatccagca ggcaatcccc aagtcagtcg acggcgacgc gaaacataa ccagagcatc
agggttgttg tggctgtgta ttttacctcg atcacttycg cagaattgct
ttcactttta gtcacctaga cagggctttta gatgaacctgc caaacaataa ctctatattc
tycgaaagaa tcacactcttt ctgtgctgct gtyaatggttt gcctggtttc caataatttac
ttttctgtgtgagttctgtgg tgtgctgaa aatccaatat cagaccagg
agagaaagca tcagactact gcaagaagtytg aagaagtcgg aagtcggctat attatagat
tacagtatttg taggctcttt tttatttggg ttggaacgta tatgtaaa gttgatttac
tcag
<210> SEQ ID NO: 22
<211> LENGTH: 358
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> OTHER INFORMATION: Incyte ID No: 1650519CD1
<400> SEQUENCE: 22
Met Gly Phe Asn Leu Thr Leu Ala Lys Leu Pro Asn Asn Glu Leu
1  5 10 15
His Gly Gln Glu Ser His Asn Ser Gly Asn Arg Ser Asp Gly Pro
20 25 30
Gly Lys Asn Thr Thr Leu His Asn Glu Phe Asp Thr Ile Val Leu
35 40 45
Pro Val Leu Tyr Leu Ile Ile Phe Val Ala Ser Ile Leu Leu Asn
50 55 60
Gly Leu Ala Val Trp Ile Phe Phe His Ile Arg Asn Lys Thr Ser
65 70 75
Phe Ile Phe Tyr Leu Lys Asn Ile Val Val Ala Asp Leu Ile Met
80 85 90
Thr Leu Thr Phe Pro Phe Arg Ile Val His Asp Ala Gly Phe Gly
95 100 105
Pro Trp Tyr Phe Lys Phe Ile Leu Cys Arg Tyr Thr Ser Val Leu
110 115 120
Phe Tyr Ala Asn Met Tyr Thr Ser Ile Val Phe Leu Gly Leu Ile
125 130 135
Ser Ile Asp Arg Tyr Leu Lys Val Val Lys Pro Phe Gly Asp Ser
140 145 150
Arg Met Tyr Ser Ile Thr Phe Thr Lys Val Leu Ser Val Cys Val
155 160 165
Trp Val Ile Met Ala Val Leu Ser Leu Pro Asn Ile Ile Leu Thr
170 175 180
Asn Gly Gln Pro Thr Glu Asp Asn Ile His Asp Cys Ser Lys Leu
185 190 195
Lys Ser Pro Leu Gly Val Lys Trp His Thr Ala Val Thr Tyr Val
200 205 210
Asn Ser Cys Leu Phe Val Ala Val Leu Val Ile Leu Gly Cys
215 220 225
Tyr Ile Ala Ile Ser Arg Tyr Ile His Lys Ser Ser Arg Gln Phe
230 235 240
Ile Ser Glu Ser Ser Arg Lys Arg Lys His Asn Glu Ser Ile Arg 245 250 255
Val Val Val Ala Val Tyr Phe Thr Cys Phe Leu Pro Tyr His Leu 260 265 270
Cys Arg Met Pro Ser Thr Phe Ser His Leu Asp Arg Leu Leu Asp 275 280 285
Glu Ser Ala Gln Lys Ile Leu Tyr Tyr Cys Lys Glu Ile Thr Leu 290 295 300
Phe Leu Ser Ala Cys Asn Val Cys Leu Asp Pro Ile Ile Tyr Phe 305 310 315
Phe Met Cys Arg Ser Phe Ser Arg Trp Leu Phe Lys Ser Asn 320 325 330
Ile Arg Pro Arg Ser Glu Ser Ile Arg Ser Leu Glu Ser Val Arg 335 340 345
Arg Ser Glu Val Arg Ile Tyr Asp Tyr Thr Asp Val 350 355

<210> SEQ ID NO: 23
<211> LENGTH: 5933
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 474630.4
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2373-2407
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 23

cagagcgagc gcaacgacag agcagagcgt ggcggagaga gggagaaga ggatgccag 60
gcacgccccc agcactgtgg cagcggagcc tcgggtcagc tttcctcctc 120
tgggcacttg gcacgtaagc acctgtccac ttcagggctg gaaacttga ggtcaaca 180
tccagcctg agagaaggct gacgctcagc caggggctgg cagactagga gcaagggcg 240
tcgtctggct aagcaactg cagagtcata gaaaccagaa gaaaccagaa gaaaccagaa 300
agcggccgca cgaagccgct gcaggagagc acaatcttgg aagctacttgg ctgcct 360
tgtgataag aagcagcctg gatcagcact gtctggcagc agaagitggt acaggacggc 420
cagagcgagc cagcagacgag ctgggcaagc ggcggagaga tggagctgat cgggacgag 480
ccttccaaat ccaacagggc ccaagattga cccacacctg cgggctcagc agatgtcccc 540
cagagctgct cgggtctgtg tcgggcgccg ttcggcgcaag catttctgag tggagcttct 600
tgaggctcag gaggccggc tggactctgt aatccctctg gcttttctctactcactgcc 660
cagatgtctg gcacactca gaaagatgg gcaaacactg gtcgggtatg ctgagactg 720
cacgcagcag tacactattg gattggcag gttgtggcag aataaatgctg tcgggaatag 780
ggcagatgct ccctgagagc tggagcagct tcggcccaac agtttgccc cttctctctt 840
cgcaggatac atcgccgtta ctcaggaatt ggtatggttg ttgagcttac tgggagggct 900
gaggcttaa gcaacactctg atgtgctgact ggctggttc ggtctgacc tggcagcagc 960
tgtgtgtagc aggccagcct gtcggccgcc ggcggccagc aaccctctct gctttcctc 1020
cagttcagc ttcctacttg aggtctgtgg cggcagcctgc ttcgggtcag ttcgagccg 1080
cacgcagcag cagcaggaga ttcggagagc gtcgggtact tggcagcagc 1140
ctacccctgc gtcgccaccc tggtgacgct gtcgocaag cacaacatca toccacattt 1200
tgctgacgcc caaacttcgct tatacgctct acggaactct tccacactat tcccctgtctt 1260
cctacatggg ggttgacgag gagacctggc caaacatctg ggagcgtctg gaggacgct 1320
tocactgat gcgccacccg tccacactac gcgccacccg ggaaccacag cccgaccaccc 1380
cagagctcag tccacagagc tcgaagagca gcggagctcg gtcgacgctgt cctggcgccg 1440
ggggatccttt tataacccag ttgcaagcgc ggcacctggc gtcgacgctgt cctggcgccg 1500
tgcgcagct gcggggacgc sgacgggcac acaactcctaat gaaactcctaat ctgcacgct 1560
gctacaggc gctcggcgcag acctacccag tgcagctcgt cctggcgccg 1620
tctggtcgaga tgggcagacag ctgcagcagtc ctgtgtgcac gtcgacgctgt cctggcgccg 1680
agggtaggag ggcggacacac aacggtgctcct gcgtgtgcac gtcgacgctgt cctggcgccg 1740
gccggaagc ggcggacacac aacggtgctcct gcgtgtgcac gtcgacgctgt cctggcgccg 1800
tggtgtgctg cctgggacag gcgggtcactg gccggtcgacag gcgggtcactg gcgggtcactg 1860
tgcgagctcag tctgggacag gcgggtcactg gccggtcactg gcgggtcactg gcgggtcactg 1920
tggtgtgacag tgggtgacag gcgggtcactg gccggtcactg gcgggtcactg gcgggtcactg 1980
tccagcgca ggggagcccc ggtgtgtgccg gcgggtcactg gcgggtcactg gcgggtcactg 2040
gccgccgacg ggggagcccc ggtgtgtgccg gcgggtcactg gcgggtcactg gcgggtcactg 2100
gccgccgacg ggggagcccc ggtgtgtgccg gcgggtcactg gcgggtcactg gcgggtcactg 2160
agaagggcgc caacgctggg gattcgagag gcaggtgacag gcgggtcactg gcgggtcactg 2220
gccgccgacg ggggagcccc ggtgtgtgccg gcgggtcactg gcgggtcactg gcgggtcactg 2280
agtgaagacg ctggagtgaag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2340
agagaactc cccccgagcc gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2400
nnnnnnccg cccggtgctgc cgctgtgctg gcgggtcactg gcgggtcactg gcgggtcactg 2460
tggtgtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2520
tggtgtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2580
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2640
tcggtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2700
tggtgtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2760
tggtgtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2820
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2880
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 2940
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3000
tcggtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3060
tcggtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3120
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3180
tcggtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3240
tcggtgacag tgggtgacag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3300
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3360
agacaacagt ggtgggagag gcgggtcactg gcgggtcactg gcgggtcactg gcgggtcactg 3420
agctgccaga agtggacctc ctcttgcggg gcgcgcaggt ccgcgcgttcc caagtccagc 3480
tcgcaccca tcaggctggg gcgcgcctgg gcgcgcctcc ctccaccccttc ctccatccaa 3540
ggcgcgcagc tgaactgcaac cgaggctttca gcgcggtcag gtgcctctca gcgcgcctcc 3600
tcgcgcgcgca ctcgcgcgcg gcgcgcgctgc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3660
tcgcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3720
ttcgggtgca ctcgcaatcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3780
tcgcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3840
ggcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3900
cgcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 3960
cggcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4020
aacgcgcgcg aatgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4080
tgcgggtgca gacgcgcggt ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4140
cccgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4200
cgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4260
acgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4320
ccgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4380
cccgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4440
tgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4500
accgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4560
tgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4620
tgcgcgcgcgca ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4680
cccgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4740
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4800
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4860
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4920
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 4980
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5040
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5100
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5160
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5220
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5280
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5340
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5400
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5460
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5520
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5580
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5640
acgcgcgcgcgcg ctcgcgcgcg gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc gcgcgcgctcc 5700
-continued

cctctcccg aggctctca gctactcct catctgacc ccctgggggcc agctccacccg 5760
catgacagca gacggggcca ggtgtctct gggagcctat aggaggccaa ggttcgcttct 5820
cgtctgggcc aacactattt gtaaccagaa agctgggaga acgacaaagga cccagccttt 5880
gtttgtcagct taaaatcgtg tttctgctct gctaaaaaagaa aaaaaagc gccg 5933

<210> SEQ ID NO: 24
<211> LENGTH: 573
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 093496.1

<400> SEQUENCE: 24

aaaaagaaaaa gaaagaaaaa tggtgactt ggtgaaaco ttcaggtcca acaatggga 60
ttcagagc cagagacgac gtgggaatca tttctgtagag gttctcttggt gttacacctt 120
cctggtgct ctcactccct actctggtcc aagctgggggcc tggagagatacg tggatctgtt 180
cagctccagca gttccacatg tggagctgta acacatggg acaagctctc gtcagcttca 240
gttaggaata cccacaaagcc cagcagccca gtaagaagca gggagcctaa ccctctctga 300
gagagagaca tggagccag gagagcgaac gcatgctttg tgcgagatgc aatttctggg 360
cactctccac gacacaatctctctgcgct tgggctctgt gaagaaaactg tgggctctgag 420
agttgaggcc aagggggccc ggcacacact gacagtgtt gttgggggcttg gggggtctg 480
gggtgggaaac aagggggccc gacagtgtgct tgggagcctc ccctttgtttt ggaactaga 540
cattcgccct tgcgcttgacg tgcgctgact gaa 573

<210> SEQ ID NO: 25
<211> LENGTH: 269
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 1231633.4

<400> SEQUENCE: 25

aacaotgagac atcctgagctg ccgctatgggcc gcggcggcccg cccggttccgt aaggggtagt 60
agagacgcc gctgacccaa gtcctggcttc cgccaggttg tcctgtgact gggaggtgcc 120
cgataggtg cccataagct atcgtgasaa aaggtggcaca gatggctgctt ccatacggc 180
gttcggctcc aaccctccaa cgtctacgctc atcaacagag tggagctcttg tggagggctct 240
ggaggcagtc aacagggccat gggaggtg 269

<210> SEQ ID NO: 26
<211> LENGTH: 1743
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 988891.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 1562
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 26
<210> SEQ ID NO 27
<211> LENGTH: 391
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 98091.15
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 14
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 27

atatttttag agantgttct aaaaaattt tagttttagt acaatcct tttaaacagc 60
-continued

gcagggcacc aagtttttac acgccccca aaggagtttc tggcagccac tgtcaagatc 120
cgtaaagct ggtctccgga ctctacagca aaaaagctgg tyagtgcttg ggcataatgaa 180
cgtaaagct ttacatcctca ggccaggaat tgtggagccac aaggtggatt ctcctagttc 240
tttgtcagc gagagacagc agagcgacaa atggcagatc gacccagctt ctcaggttca 300
gatctgtac ctgctctgctc tgtcattaga ttgataaacg ccaacttacct atggggcaga 360
gtttagct ccccaaaaaa atttacccca a 391

<210> SEQ ID NO: 28
<211> LENGTH: 7045
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 3774181CB1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 103, 6960
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 28

tagtttgtctg gatctttggtg ccagccgagc aaaaagaga ccagctgta ttttgtgttt 60
tttcattagg ccacaaagag ttcagctctac ccagagcaat ctaaatcgtg catcagcgcc 120
agatcagct ggtgatagat agaagcctt ccaagagcgct aagctgcttg ggggcataatg 180
agtcagagct ttctagctccaa atgtccagct aaacctctct ctccagccatatggagaa 240
gccaaacaa taacttacttc gaagacaaag aaattaatcaggtcctg ccacccatcctgtt 300
gcctcagc ctcctctcgc ctaatcagat gacatctcag atggtaaggtc atttgcttatg 360
tgaggtttaag gcggagaagtt cctttggtttg gcgatagcc atggctccca aagtcctcca 420
gagaggtgt gcctgctgcc gcagccggcag cctcaggtcct ctaagccctca cttcctgctcag 480
tgcagaccc gcgagctgcg caagctctgc ctcgctatag cgcctcagcg gatcagtcgct 540
atataatgg cagagcagct gtcgatcatt ccctctcagc gaggattgta tggccatagct 600
gctgaggaag tgagcctctt ccagggagct ctcttcctgc gcactgctaga acggctcctag 660
aatataatga ctctctgatt gcagttttcg caagaggagc caagatccag gtcctttttttttc 720
ggcgtagct cagcactct ggcagggagc gcagcattag tgaagagat gtttataatggg 780
tctctcagct tgtgctgcc gcagcagct gtccttgca tattcagctgct gtccttgca gtttataatggg 840
gagctggtcc cggcagctcg ccgtctgtcg atctctcagc ctctctcagc cggcctctcagc 900
tctctcttgc tcgtgcgcct gcagttttcg gcagctgctgct gcagctgtca ggttccttcctcagc 960
aagctatgctt cggcctctac gcagttttcg gcagctgtca gggccttgcg gcagctgctgct 1020
ggagttgggt gtctctctct gcagctgtca gggccttgcg gcagctgctgct gcagctgtca 1080
aatgttgtg cggcagctcg ctctctctct gcagctgtca gggccttgcg gcagctgctgct 1140
tcagtttgc gcagctgtca gggccttgcg gcagctgctgct gcagctgctgct gcagctgctgct 1200
tgcctctcagct cttcctctgc cggccttgcg gcagctgctgct gcagctgctgct gcagctgctgct 1260
tctctcttgc cggccttgcg gcagctgctgct gcagctgctgct gcagctgctgct gcagctgctgct 1320
gcagtttgc gcagctgctgct gcagctgctgct gcagctgctgct gcagctgctgct gcagctgctgct 1380
aagctatgctt cggccttgcg gcagctgctgct gcagctgctgct gcagctgctgct gcagctgctgct 1440
-continued

tgcaaaatg ttcgtgctca gatgaaacac aggttacggg acttagaggg catggcata 
1500
tatctgagct atactaacaag cactctcct cctttagagt tttgaatgat cgaaggttggag 
1560
actacattaa gaaagctatg ggaattttcct cctgaaatata attaaacccct aagcacaacag 
1620
tggatataag aggagatct ggtgtaccta attaagatga aagagacagt aagatggagc 
1680
ttgcaaaat atgcaagacac gtactagcagt acagtgagact atactagattt aacacaagtt 
1740
acactacggg cctcttagata ttcaacaacct aatctccccag tgaaccgccc aagatgtc 
1800
agttcagcag atctcttatt tcaagagttc atggacctaag gggctcggta tcttgacctg 
1860
gtcaactctc tgcacaatac tatataatttt gctggtgatt cttgaagag cgtgagagac 
1920
gagggataa aaggtgtaaa gggagttctct gcacagtggg cacatttagac tctgttgcag 
1980
cytcgagag aacacagctct tgaagatagt cacccattcag gaaagataag tgaagtgag 
2040
agagattgtag atcaacttataa gaacaaacag ttctagatag agagaaacct tctcaaggtc 
2100
agggagcct cgaacataag attgaagag ccagacgacaa atgggagagat tattctctgt 
2160
cagagataaa ggcttgaagcg tgaagccagag cagtacgagca gcaagactgga aagactgtg 
2220
agaggagaga aagcgagcctg aagagactct ggcgggtggtg gcacgctcaac tcatagcgcc 
2280
gaggtttaaa gagctcagct gcagagagaac ctctgaaatatt tgcccaatcct gttgacgaa 
2340
aacaaccctca cgaagcgcac aagggagagt ctcttataaaaa gaaagattt aagcttcaaat 
2400
agtggagc aacaaaaaaa taattaagtt gaaggaatataa gaagaaagag aagcaagata 
2460
gagacactct tgaagctagat aagagcgattg gaaagacgcc ttgacattcct gaaagcgtta 
2520
gacagagaac atggagagaa aagacgaaata attgaaatgg aagcaagaga aaaaataact 
2580
gaaatctagt atacagttag agaagaagcag ttgccagcttg tgcagcctcc aacagctca 
2640
tcagcagcgg cagtcgacg gttcactgaa gacagacgca aacagcagag aagacactct 
2700
aacaaggtt gatagtaaat aacagctgccc actagagatg tgaacaaagta cagagagag 
2760
tgcacataag aacaaaaagc cgtucgcgtct gaaaaaccgct cattgagga aaggyctcgt 
2820
ttgttaaaag ataaactagaa gaaaaacaact ataaactagct ggtcgcttt ctggagctc 
2880
gaaaggaag agtccagcgg gcagagcagtt cttcaactac ttcagagcgt tggagccgaa 
2940
tgtagttaaa ccaacagaaa agctgaaagc gctcgtgagg tagtagtgaa cttcagaaaaa 
3000
atactggcct attatgatttt gagattagaaaa ctcttctataagc atggaaaggg gasactcaaa 
3060
agagaagtag atacagatcc aagggcaacag ctgctagctag agagaattat tocaattttaa 
3120
aattccaaa ttcgactctct tagatgactag aaaaattagtag atgacactca aatgtccgcc 
3180
agagaatctac gttccttctag aaaaaacctgct gagaaaacctct gttgagctcct 
3240
atctacacaa ccaacacataa aaggaagattg ggaacgctt ggttcatgatt gttaaaaataat 
3300
agtactagtg gtggagagatg gtttaaaaac ggaattagag acggctactag gcagatatt 
3360
gaaagccacta taagatagtc gaagaacttc gcgaaatctag aagagaaaat ttttgggaa 
3420
cacaatatcc acagagagat ggaaagacat tgaagactcctg gaaatcacttag ttaagtaaat 
3480
catcagcagca cctccataag gagcagaccc cagcagacagt gttaaaaaaa 
3540
aataataact atttgaagag cagacctggc aaaaacccaa attttgtaga tgaattttgag 
3600
ccaaagttgag accacaaacat ctatttaccata cagactacca acaaaaagtt cttatttctg 
3660
aatgagcacc tcaatgttttt cacagagagc agggagagag atgttcagcag 
3720
-continued

caacaagctc agyttggcaga ggtaataac agyttggaaa aagttcacaag cgaattacoc 3780
ttaagacaca taggaagagc gccagcagac agaagaggtg attttgttga caaagtaaatct 3840
ggttaattca aacatcagc agaggaattt cgaagagacca tggaaattatt attggagttcc 3900
aagctctaca tggtaattgta tattcacaag actatggtgg ctatagttgtc tatttcacaa 3960
gaaactccta gagcccaaga aatgtcataa cttttgtgaaa aacatcattaa aagacattgaa 4020
agacagcttc aacagtctagc tgaacataga caaggaaggc acgacatgga aagcaaatctc 4080
taccaaaaact gtcagagaccact ctcagaaaccagc agaagctgta ggttgcaaac 4140
cgtaaacga aagcaagacc aagcattccaa ggcgtaagac atcataataat tttgcttccag 4200
tgtgaanattc aanaaaaaag cacccgcaaan gcgcgtacata ccacacacaaa ttttgggagat 4260
agctgaagg agtggcagcag cttcggagag ctcctctctca gaaacctctgg cagcccacaac 4320
cacaacacca gatccaccttc gttgagatggt actccaagac ccagagcattt gggaagagag 4380
ctggcagacat gggatgtagt acaagctccca aagagaattg cattccagcc accaggggct 4440
ccacagcaga aagagagaag aacagcagtgt tccttgagatc atttttctca gaacagcaac 4500
gagtacagca taacctttgga tgaacagaaa ccacatcacaagcgcctgaagtaaattcgaag 4560
ataagagacc aagcctataa ccaatctgta ggtatcaccag taaaggtctat 4620
gaaatggaactgtgagagttttaaagcaagctgtagctcgaaagcctgaag 4680
agctcaacag aagcctacac ccataaaaaa caagacccag cgggtcagcag tctgggaaga 4740
tggtgcatgg aagggcagcag agcctgttgg ggcagcacaag aggggcttcct ctctagagaag 4800
gggttgacag cagagcaccctt ccagaactttt ggtgatgata atgtcatgctc acgtgagggat 4860
ggataatta aaaaacctcg ggtggagcag actcgagctg ccaagcactt ggtggaagtt 4920
aaccctctgg aacagcagcc aagctgagcag ctcgcagctgc gttcagagaa cgttggagag 4980
gtcgcaaaaa cttttatatg gttttgctag aagcctaccct ctcagcgcag accaatagcct 5040
gaaotanca aagagagat gtaacatcttc ctatcgccgg gaggagttat caaatctcaaa 5100
attggggctgtgccttctgtg gaaagtcagc ggtctcaagc ggttactgta actctcgctgccagt 5160
tgggtggca aatattcatttg tgaagagcgc gttttataat gttccatgttc 5220
agatagacggctacctctgag aagggggcat gtgtgtgatc attttatgattt gttttaagaca 5280
ttgggactt gttcaccctat gggaaagcat gacgcaatctg tttttttctccttctcctt 5340
tggagagccg aatgggcttag cccgggctct gtaagcggctg cttcgtttct 5400
cccacacgct tagtccagcag ggtggttctgg aaaaatgtcag aatttgctata cttattccag 5460
cccacacgct tagtccagcag ggtggttctgg aaaaatgtcag aatttgctata cttattccag 5520
gaatattzac gatggatgtg attttgtgta actttctctt hctttgggtt 5580
bagagacacct cttccttcct catcgctcaag aaaaacctata gaaattctgatt gtagatctct 5640
aaacacggct cgaattacag ctgtgagatg acgccccaga aacagcatttg tgaagagagat 5700
atatatoctg aacatctcag cttcttcaat ccaataacac agaagacttgc gttttatatc 5760
ttagggctct tttttttctg gtagatcctg aactaacagc gttttaacattt 5820
agaggttctg gggcgggacat aagcgctcag gctttggctgt aaaaatctaca ggaaggcctct 5880
atcacaactta cagacactgct cttgctctgtt ctggagcctgt taagcctccata gaaagacct 5940
cacttccgctt tttagctgatgtgtgggtg cagggacttct tgactataa 6000
<table>
<thead>
<tr>
<th>50</th>
<th>Val</th>
<th>Leu</th>
<th>Ala</th>
<th>Asn</th>
<th>Asn</th>
<th>Ser</th>
<th>His</th>
<th>Arg</th>
<th>Ala</th>
<th>Lys</th>
<th>Lys</th>
<th>Trp</th>
<th>Lys</th>
<th>Val</th>
<th>Ile</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>Cys</td>
<td>Val</td>
<td>Leu</td>
<td>Ala</td>
<td>Asn</td>
<td>Ser</td>
<td>His</td>
<td>Arg</td>
<td>Ala</td>
<td>Lys</td>
<td>Lys</td>
<td>Trp</td>
<td>Lys</td>
<td>Val</td>
<td>Ile</td>
<td>70</td>
</tr>
<tr>
<td>65</td>
<td>Ser</td>
<td>Pro</td>
<td>Thr</td>
<td>Gly</td>
<td>Asn</td>
<td>Ala</td>
<td>Met</td>
<td>Val</td>
<td>Pro</td>
<td>Ser</td>
<td>Val</td>
<td>Cys</td>
<td>The</td>
<td>Thr</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Ser</td>
<td>Pro</td>
<td>Thr</td>
<td>Gly</td>
<td>Asn</td>
<td>Ala</td>
<td>Met</td>
<td>Val</td>
<td>Pro</td>
<td>Ser</td>
<td>Val</td>
<td>Cys</td>
<td>The</td>
<td>Thr</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Val</td>
<td>Pro</td>
<td>Pro</td>
<td>Asn</td>
<td>Lys</td>
<td>Ala</td>
<td>Val</td>
<td>Asp</td>
<td>Ala</td>
<td>Asn</td>
<td>Arg</td>
<td>Ile</td>
<td>95</td>
<td>100</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Glu</td>
<td>Gln</td>
<td>Tyr</td>
<td>Gln</td>
<td>Asn</td>
<td>Val</td>
<td>Leu</td>
<td>Thr</td>
<td>Leu</td>
<td>Trp</td>
<td>His</td>
<td>Glu</td>
<td>Ser</td>
<td>His</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Glu</td>
<td>Gln</td>
<td>Tyr</td>
<td>Gln</td>
<td>Asn</td>
<td>Val</td>
<td>Leu</td>
<td>Thr</td>
<td>Leu</td>
<td>Trp</td>
<td>His</td>
<td>Glu</td>
<td>Ser</td>
<td>His</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>Ile</td>
<td>Asn</td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>His</td>
<td>Tyr</td>
<td>Leu</td>
<td>Asn</td>
<td>Glu</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Ile</td>
<td>Asn</td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>His</td>
<td>Tyr</td>
<td>Leu</td>
<td>Asn</td>
<td>Glu</td>
<td>130</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Ile</td>
<td>Asn</td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
<td>Val</td>
<td>Ser</td>
<td>Trp</td>
<td>His</td>
<td>Tyr</td>
<td>Leu</td>
<td>Asn</td>
<td>Glu</td>
<td>135</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO: 29
<211> LENGTH: 2125
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3774181CD1
<400> SEQUENCE: 29
Leu Pro Gly Glu His Gln Gln Val Leu Ser Asn Leu Gln Ser Arg
155 160 165
Phe Glu Asp Phe Leu Glu Asp Ser Gln Glu Ser Gln Val Phe Ser
170 175 180
Gly Ser Asp Ile Thr Gln Leu Glu Lys Glu Val Asn Val Cys Lys
185 190 195
Gln Tyr Tyr Gln Glu Leu Leu Lys Ser Ala Glu Arg Glu Glu Gin
200 205 210
Glu Glu Ser Val Tyr Asn Leu Tyr Ile Ser Glu Val Arg Asn Ile
215 220 225
Arg Leu Arg Leu Gln Asn Cys Glu Asp Arg Leu Ile Arg Glu Gin Ile
230 235 240
Arg Thr Pro Leu Glu Arg Asp Leu His Glu Ser Val Phe Arg
245 250 255
Ile Thr Glu Glu Lys Leu Lys Lys Glu Leu Arg Leu Lys
260 265 270
Asp Leu Glu Gly Thr Ile Thr Asn Lys Cys Glu Glu Phe Phe Ser
275 280 285
Gln Ala Ala Ala Ser Ser Ser Val Pro Thr Leu Arg Ser Glu Leu
290 295 300
Asn Val Val Leu Gin Asn Met Asn Gin Val Tyr Ser Met Ser Ser
305 310 315
Thr Tyr Ile Asp Lys Leu Lys Thr Val Asn Leu Val Leu Lys Asn
320 325 330
Thr Gln Ala Ala Glu Ala Leu Val Lys Leu Tyr Glu Thr Lys Leu
335 340 345
Cys Glu Glu Glu Ala Val Ile Ala Asp Lys Asn Asn Ile Glu Asn
350 355 360
Leu Ile Ser Thr Leu Lys Gin Trp Arg Ser Glu Val Asp Glu Lys
365 370 375
Arg Gin Val Phe His Ala Leu Glu Asp Glu Leu Gln Lys Ala Lys
380 385 390
Ala Ile Ser Asp Glu Met Phe Thr Tyr Lys Glu Arg Asp Leu
395 400 405
Asp Phe Asp Trp His Lys Glu Ala Asp Gin Leu Val Glu Arg
410 415 420
Trp Gin Asn Val His Val Gin Ile Asp Asn Arg Leu Arg Asp Leu
425 430 435
Glu Gly Ile Gly Lys Ser Leu Tyr Tyr Arg Asp Thr Tyr His
440 445 450
Pro Leu Asp Asp Thr Ile Glu Gin Val Glu Thr Thr Gin Arg Lys
455 460 465
Ile Gin Glu Asn Gin Pro Glu Asn Ser Lys Thr Leu Ala Thr Gin
470 475 480
Leu Asn Gin Gin Lys Met Leu Val Ser Gin Pro Gin Met Lys Gin
485 490 495
Ser Lys Met Asp Glu Cys Gin Lys Tyr Ala Gin Tyr Ser Ala
500 505 510
Thr Val Lys Asp Tyr Glu Leu Gin Thr Met Thr Tyr Arg Ala Met
515 520 525
Val Asp Ser Gln Gln Lys Ser Pro Val Lys Arg Arg Arg Met Gln 530 535 540
Ser Ser Ala Asp Leu Ile Ile Gln Glu Phe Met Asp Leu Arg Thr 545 550 555
Arg Tyr Thr Ala Leu Val Thr Leu Met Thr Gln Tyr Ile Lys Phe 560 565 570
Ala Gly Asp Ser Leu Lys Arg Leu Glu Glu Glu Ile Lys Arg 575 580 585
Cys Lys Glu Thr Ser Glu His Gly Ala Tyr Ser Asp Leu Leu Gln 590 595 600
Arg Gln Lys Ala Thr Val Leu Glu Asn Ser Lys Leu Thr Gly Lys 605 610 615
Ile Ser Glu Leu Glu Arg Met Val Ala Glu Leu Lys Gln Lys 620 625 630
Ser Arg Val Glu Glu Leu Pro Lys Val Arg Glu Ala Ala Glu 635 640 645
Asn Glu Leu Arg Lys Gln Gln Arg Asn Val Glu Asp Ile Ser Leu 650 655 660
Gln Lys Ile Arg Ala Glu Ser Glu Ala Lys Gln Tyr Arg Arg Glu 665 670 675
Leu Glu Thr Ile Val Arg Glu Lys Ala Ala Glu Arg Glu Leu 680 685 690
Glu Arg Val Arg Glu Leu Thr Ile Glu Ala Glu Ala Lys Arg Ala 695 700 705
Ala Val Glu Glu Asn Leu Leu Asn Phe Arg Asn Glu Leu Glu Glu 710 715 720
Asn Thr Phe Thr Arg Arg Thr Leu Glu Asp His Leu Lys Arg Lys 725 730 735
Asp Leu Ser Leu Asn Asp Leu Glu Gln Gln Lys Asn Lys Leu Met 740 745 750
Glu Glu Leu Arg Lys Arg Asp Asn Glu Glu Glu Leu Leu Glu Lys 755 760 765
Leu Ile Lys Gln Met Glu Lys Asp Leu Ala Phe Glu Lys Gln Val 770 775 780
Ala Glu Lys Gln Leu Lys Glu Gln Lys Ile Glu Leu Glu Ala 785 790 795
Arg Arg Lys Ile Thr Glu Ile Gln Tyr Thr Cys Arg Glu Asn Ala 800 805 810
Leu Pro Val Cys Pro Ile Thr Gin Ala Thr Ser Cys Arg Ala Val 815 820 825
Thr Gly Leu Gin Gin His Asp Lys Gin Lys Ala Glu Glu Leu 830 835 840
Lys Gin Gin Val Asp Gin Leu Thr Ala Ala Asn Arg Lys Ala Glu 845 850 855
Gln Asp Met Arg Glu Leu Thr Tyr Glu Leu Asn Ala Leu Gin Leu 860 865 870
Glu Lys Thr Ser Ser Glu Glu Lys Ala Arg Leu Leu Lys Asp Lys 875 880 885
Leu Asp Glu Thr Asn Asn Thr Leu Arg Cys Leu Lys Leu Glu Leu 890 895 900
Glu Arg Lys Asp Gin Ala Glu Lys Gly Tyr Ser Gin Gin Leu Arg
<table>
<thead>
<tr>
<th></th>
<th>905</th>
<th></th>
<th></th>
<th>910</th>
<th></th>
<th>915</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu Leu Gly Arg Gln Leu Asn Gln Thr Thr Gly Lys Ala Glu Glu</td>
<td>920</td>
<td></td>
<td></td>
<td>925</td>
<td></td>
<td>930</td>
</tr>
<tr>
<td>Ala Met Gln Glu Ala Ser Asp Leu Lys Ile Lys Arg Asn Tyr</td>
<td>935</td>
<td></td>
<td></td>
<td>940</td>
<td></td>
<td>945</td>
</tr>
<tr>
<td>Gln Leu Glu Leu Glu Ser Leu Asn His Glu Lys Gly Lys Leu Gln</td>
<td>950</td>
<td></td>
<td></td>
<td>955</td>
<td></td>
<td>960</td>
</tr>
<tr>
<td>Arg Glu Val Asp Arg Ile Thr Arg Ala His Ala Val Ala Glu Lys</td>
<td>965</td>
<td></td>
<td></td>
<td>970</td>
<td></td>
<td>975</td>
</tr>
<tr>
<td>Asn Ile Gln His Leu Asn Ser Gln Ile His Ser Phe Arg Asp Glu</td>
<td>980</td>
<td></td>
<td></td>
<td>985</td>
<td></td>
<td>990</td>
</tr>
<tr>
<td>Lys Glu Leu Glu Arg Leu Gln Ile Cys Gln Arg Lys Ser Asp His</td>
<td>995</td>
<td></td>
<td>1000</td>
<td></td>
<td>1005</td>
<td></td>
</tr>
<tr>
<td>Leu Lys Glu Gln Phe Glu Lys Ser His Gln Leu Leu Gln Aan</td>
<td>1010</td>
<td></td>
<td>1015</td>
<td></td>
<td>1020</td>
<td></td>
</tr>
<tr>
<td>Ile Lys Ala Glu Lys Glu Asn Asn Asp Lys Ile Gln Arg Leu Aan</td>
<td>1025</td>
<td></td>
<td>1030</td>
<td></td>
<td>1035</td>
<td></td>
</tr>
<tr>
<td>Glu Glu Leu Glu Lys Ser Asn Glu Cys Ala Glu Met Leu Lys Gln</td>
<td>1040</td>
<td></td>
<td>1045</td>
<td></td>
<td>1050</td>
<td></td>
</tr>
<tr>
<td>Lys Val Glu Glu Leu Thr Arg Glu Asn Glu Thr Lys Leu Met</td>
<td>1055</td>
<td></td>
<td>1060</td>
<td></td>
<td>1065</td>
<td></td>
</tr>
<tr>
<td>Met Glu Arg Ile Gln Ala Glu Ser Glu Asn Ile Val Leu Glu Lys</td>
<td>1070</td>
<td></td>
<td>1075</td>
<td></td>
<td>1080</td>
<td></td>
</tr>
<tr>
<td>Gln Thr Ile Gln Gln Arg Cys Glu Ala Leu Lys Ile Gln Ala Aas</td>
<td>1085</td>
<td></td>
<td>1090</td>
<td></td>
<td>1095</td>
<td></td>
</tr>
<tr>
<td>Gly Phe Lys Asp Gln Leu Arg Ser Thr Asn Glu His Leu His Lys</td>
<td>1100</td>
<td></td>
<td>1105</td>
<td></td>
<td>1110</td>
<td></td>
</tr>
<tr>
<td>Gln Thr Lys Thr Glu Gln Asp Phe Glu Arg Lys Ile Lys Cys Leu</td>
<td>1115</td>
<td></td>
<td>1120</td>
<td></td>
<td>1125</td>
<td></td>
</tr>
<tr>
<td>Glu Glu Asp Leu Ala Lys Ser Gln Asn Leu Val Ser Glu Phe Lys</td>
<td>1130</td>
<td></td>
<td>1135</td>
<td></td>
<td>1140</td>
<td></td>
</tr>
<tr>
<td>Gln Lys Cys Asp Gln Gln Asn Ile Ile Gln Asn Thr Lys Lys</td>
<td>1145</td>
<td></td>
<td>1150</td>
<td></td>
<td>1155</td>
<td></td>
</tr>
<tr>
<td>Glu Val Arg Asn Leu Asn Ala Glu Leu Aan Ala Ser Lys Glu Glu</td>
<td>1160</td>
<td></td>
<td>1165</td>
<td></td>
<td>1170</td>
<td></td>
</tr>
<tr>
<td>Lys Arg Arg Gly Glu Gln Val Gln Leu Gln Gln Ala Gln Val</td>
<td>1175</td>
<td></td>
<td>1180</td>
<td></td>
<td>1185</td>
<td></td>
</tr>
<tr>
<td>Gln Glu Leu Aan Arg Leu Gln Lys Val Gln Asp Glu Leu His</td>
<td>1190</td>
<td></td>
<td>1195</td>
<td></td>
<td>1200</td>
<td></td>
</tr>
<tr>
<td>Leu Lys Thr Ile Glu Glu Gln Met Thr His Arg Lys Met Val Leu</td>
<td>1205</td>
<td></td>
<td>1210</td>
<td></td>
<td>1215</td>
<td></td>
</tr>
<tr>
<td>Phe Glu Glu Glu Ser Gly Lys Phe Lys Gin Ser Ala Glu Glu Phe</td>
<td>1220</td>
<td></td>
<td>1225</td>
<td></td>
<td>1230</td>
<td></td>
</tr>
<tr>
<td>Arg Lys Lys Met Glu Lys Leu Met Glu Ser Lys Val Ile Thr Glu</td>
<td>1235</td>
<td></td>
<td>1240</td>
<td></td>
<td>1245</td>
<td></td>
</tr>
<tr>
<td>Asn Asp Ile Ser Gly Ile Arg Leu Asp Phe Val Ser Leu Gln Gln</td>
<td>1250</td>
<td></td>
<td>1255</td>
<td></td>
<td>1260</td>
<td></td>
</tr>
<tr>
<td>Glu Asn Ser Arg Ala Gln Glu Asn Ala Lys Leu Cys Glu Thr Asn</td>
<td>1265</td>
<td></td>
<td>1270</td>
<td></td>
<td>1275</td>
<td></td>
</tr>
<tr>
<td>Ile Lys Glu Leu Glu Arg Gin Leu Gin Tyr Arg Glu Gin Met</td>
<td>1280</td>
<td></td>
<td>1285</td>
<td></td>
<td>1290</td>
<td></td>
</tr>
</tbody>
</table>
Gln Gln Gly Gln His Met Glu Ala Asn His Tyr Gln Lys Cys Gln 1295 1300 1305
Lys Leu Glu Asp Glu Leu Ile Ala Gln Lys Arg Glu Val Glu Asn 1310 1315 1320
Leu Lys Gln Lys Met Asp Gln Gln Ile Lys Glu His Glu His Gln 1325 1330 1335
Leu Val Leu Leu Gln Cys Glu Ile Gln Lys Lys Ser Thr Ala Lys 1340 1345 1350
Asp Cys Thr Phe Lys Pro Asp Phe Glu Met Thr Val Lys Glu Cys 1355 1360 1365
Gln His Ser Gly Glu Leu Ser Ser Arg Asn Thr Gly His Leu His 1370 1375 1380
Pro Thr Pro Arg Ser Pro Leu Arg Trp Thr Gln Glu Pro Gln 1385 1390 1395
Pro Leu Glu Glu Lys Trp Gln His Arg Val Val Glu Gln Ile Pro 1400 1405 1410
Lys Glu Val Gln Phe Gln Pro Gly Ala Pro Leu Glu Lys Glu 1415 1420 1425
Lys Ser Gln Gln Cys Tyr Ser Glu Tyr Phe Ser Gln Thr Ser Thr 1430 1435 1440
Glu Leu Gln Ile Thr Phe Asp Glu Thr Asn Pro Ile Thr Arg Leu 1445 1450 1455
Ser Glu Ile Glu Lys Ile Arg Asp Gln Ala Leu Asn Asn Ser Arg 1460 1465 1470
Pro Pro Val Arg Tyr Gln Asn Ala Cys Glu Met Glu Leu Val 1475 1480 1485
Lys Val Leu Thr Pro Leu Glu Ile Ala Lys Asn Lys Gln Tyr Asp 1490 1495 1500
Met His Thr Gln Val Thr Thr Leu Lys Glu Glu Lys Asn Pro Val 1505 1510 1515
Pro Ser Ala Glu Glu Trp Met Leu Glu Gly Cys Arg Ala Ser Gly 1520 1525 1530
Gly Leu Lys Lys Gly Asp Phe Leu Lys Gly Leu Glu Pro Glu 1535 1540 1545
Thr Phe Gln Asn Phe Asp Gly Asp His Ala Cys Ser Val Arg Asp 1550 1555 1560
Asp Glu Phe Lys Phe Glu Gln Gly Leu Arg His Thr Val Thr Ala Arg 1565 1570 1575
Gln Leu Val Glu Ala Lys Leu Leu Asp Met Arg Thr Ile Glu Glu 1580 1585 1590
Leu Arg Leu Gly Leu Lys Trp Val Glu Glu Val Gln Lys Thr Leu 1595 1600 1605
Asn Lys Phe Leu Thr Lys Ala Thr Ser Ile Ala Gly Leu Tyr Leu 1610 1615 1620
Glu Ser Thr Lys Glu Lys Ile Ser Phe Ala Ser Ala Ala Glu Arg 1625 1630 1635
Ile Ile Ile Asp Lys Met Val Ala Leu Ala Phe Leu Glu Ala Gin 1640 1645 1650
Ala Ala Thr Gly Phe Ile Ile Asp Pro Ile Ser Gly Gln Thr Tyr 1655 1660 1665
<table>
<thead>
<tr>
<th>Residue</th>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Glu</td>
</tr>
<tr>
<td>Arg</td>
<td>Ile</td>
<td>Arg</td>
</tr>
<tr>
<td>Tyr</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Met</td>
<td>Leu</td>
<td>Asp</td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Pro</td>
<td>Glu</td>
<td>Ile</td>
</tr>
<tr>
<td>Gln</td>
<td>Phe</td>
<td>Leu</td>
</tr>
<tr>
<td>Pro</td>
<td>Asn</td>
<td>Asn</td>
</tr>
<tr>
<td>Cys</td>
<td>Val</td>
<td>Phe</td>
</tr>
<tr>
<td>Glu</td>
<td>Arg</td>
<td>Asn</td>
</tr>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Val</td>
</tr>
<tr>
<td>Ala</td>
<td>Phe</td>
<td>Gln</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Gln</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gly</td>
<td>His</td>
</tr>
<tr>
<td>His</td>
<td>Phe</td>
<td>Asn</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Val</td>
</tr>
<tr>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
</tr>
<tr>
<td>His</td>
<td>Ser</td>
<td>Pro</td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Val</td>
</tr>
<tr>
<td>Thr</td>
<td>Ala</td>
<td>Leu</td>
</tr>
<tr>
<td>Ile</td>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>His</td>
<td>Arg</td>
<td>Gly</td>
</tr>
<tr>
<td>Cys</td>
<td>Glu</td>
<td>Leu</td>
</tr>
<tr>
<td>Met</td>
<td>Met</td>
<td>Ser</td>
</tr>
<tr>
<td>Glu</td>
<td>Met</td>
<td>Gly</td>
</tr>
<tr>
<td>Leu</td>
<td>Ile</td>
<td>Glu</td>
</tr>
</tbody>
</table>
Leu Gln Val Gly Ile Ile Asp Val Leu Ile Ala Thr Lys Leu Lys
2045  2050  2055
Asp Gln Lys Ser Tyr Val Arg Asn Ile Ile Cys Pro Gln Thr Lys
2055  2060  2065
Arg Lys Leu Thr Tyr Lys Glu Ala Leu Glu Lys Ala Asp Phe Asp
2065  2070  2075
Phe His Thr Gly Leu Lys Leu Leu Glu Val Ser Glu Pro Leu Met
2075  2080  2085
Thr Gly Ile Ser Ser Leu Tyr Ser Ser
2086  2100  2105

<210> SEQ ID NO: 30
<211> LENGTH: 1708
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> OTHER INFORMATION: Incyte ID No: 1709387C8

<400> SEQUENCE: 30

cctgcacagca ttccttgggt tgtgtgagaa ttcaggggct ccaagttacct ggcctagacc 60
aaccaactt tcgacatcc tcctctctcac ttggtgggt tgtgtgtgtg aagggggttc
cc 120
tgctttcgtg gcggagaggg ctgtgggtgg ggtggtctct ctcggggaggg tgtgaagccc
agagtctcag ctcctcctcg taggtttggt ttcctcaggt caggaggaggt aatcgggggtt
240
gccctaggg tctgtggcct tgtgtggaggg gtccgtgtgt tgtgtgtggg aagtttttga
agg 300
ggggggcttg tgtggggttt tgtgttggtg tccttggtgct ggcgtgtgtgct tcctctctct
360
ggcctagaga aatctctcact gccacagcag cacagctagcccc ctgcctccct ccctgcagcag
420
gtcagttggc tcggagagcc caatgtgctac cttggagaggt gtatccgatt cgg 480
aagacagacc cagcagccag cagacagccc tccagccgatt cggctgctac cggctgtgcc
540
cctgccagcct tagatcagcag cccacatccc gccacagcag cggctacgct cggctgtgctc
600
aagacagcgc tggctgctgg accatctcag ctcctagtagg cgcagctagc gcggctgcagc
660
cggctgcttg aagctatcagc cagnccggtt cggccagctc cggctatgctg cggccagcag
720
agcgctagc gcggctgtcg ccagccggtt cggctgtctg ccagcggctc cggccagcag
780
aaccacgagc aagctggtcg cggccagcgc cggctgtctg ccagcggctc cggccagcag
840
atgagcagc ccagaggggtg ccggctgtgct cggctgctag cggccagctg cggctgctag
900
gagagaagag cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
960
agcgctagc ccagccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1020
aaccacgagc aagctggtcg cggccagcgc cggctgctag cggccagctg cggctgctag
1080
aaggtcggcg ccggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1140
caggctagc cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1200
gagggagag cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1260
agggctagc cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1320
agggctagc cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1380
tcagtctagc cggccagagc cggccagcgc cggctgctag cggccagctg cggctgctag
1440
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MetThrThrThrPheLeuGlnThrSerSerThrPheGlyGly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>GlySerThrArgGlyGlySerLeuAlaGlyGlyGlyGlyPhe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>GlyGlyGlySerLeuSerGlyGlyGlyGlySerArgSerIleSer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>AlaSerSerAlaArgPheValSerSerGlySerArgSerIleSer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>GlyGlyGlyMetArgValCysGlyPheGlyGlyGlyAlaGlySer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>100</td>
<td>105</td>
</tr>
<tr>
<td>LysIleThrMetGlnAsnLeuAsnAspArgLeuAlaSerTyrLeu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>115</td>
<td>120</td>
</tr>
<tr>
<td>AspLysValArgAlaLeuGluAlaAsnAlaAspLeuGluVal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>130</td>
<td>135</td>
</tr>
<tr>
<td>LysIleHisAspTyrGlnLysGlnThrProAlaSerProGlu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>145</td>
<td>150</td>
</tr>
<tr>
<td>CysAspTyrSerGlnTyrPheLysThrIleGluGluLeuArgAsp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>155</td>
<td>160</td>
<td>165</td>
</tr>
<tr>
<td>LysIleMetAlaThrIleAspAsnSerArgValIleLeuGlu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td>175</td>
<td>180</td>
</tr>
<tr>
<td>IleAspAsnAlaArgLeuAlaAlaAspPheArgLeuLysTyr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>190</td>
<td>195</td>
</tr>
<tr>
<td>GluAsnGluLeuAlaLeuArgGlnGlyValGluAlaAspIleAsn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>205</td>
<td>210</td>
</tr>
<tr>
<td>GlyGluArgValLeuAspLeuThrLeuAsaArgThrAsp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>215</td>
<td>220</td>
<td>225</td>
</tr>
<tr>
<td>AlaGluGlnGluThrGluAsnGluGluLeuAlaNitrLeu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>230</td>
<td>235</td>
<td>240</td>
</tr>
<tr>
<td>LysAsnHisGluGluGluGluGluMetAlaPheSerSerGlnLeu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>245</td>
<td>250</td>
<td>255</td>
</tr>
<tr>
<td>AlaGluGlnValAsnValGluGlnAspAlaAlaProGlyValAsp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>265</td>
<td>270</td>
</tr>
<tr>
<td>LeuThrArgValLeuAlaGluMetArgGluGlnTyrGluAlaMet</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ala Glu Lys Asn Arg Arg Asp Val Glu Ala Trp Phe Phe Ser Lys
Thr Glu Glu Leu Asn Lys Glu Val Ala Ser Asn Thr Glu Met Ile
Gln Thr Ser Lys Thr Glu Ile Thr Asp Leu Arg Arg Thr Met Gln
Glu Leu Glu Ile Glu Leu Gln Ser Gln Leu Ser Met Lys Ala Gly
Leu Glu Asn Ser Leu Ala Glu Thr Glu Cys Arg Tyr Ala Thr Gln
Leu Gln Gln Ile Gln Gly Leu Ile Gly Leu Glu Ala Gln Leu
Ser Glu Leu Arg Cys Glu Met Gln Ala Asn Gln Glu Tyr Lys
Met Leu Leu Asp Ile Lys Thr Arg Leu Glu Glu Ile Ala Thr
Tyr Arg Ser Leu Leu Glu Gly Asp Ala Lys Met Ala Gly Ile
Gly Ile Arg Glu Ala Ser Ser Gly Gly Gly Ser Ser Ser Asn
Phe His Ile Asn Val Glu Glu Ser Val Asp Gly Glu Val Val Ser
Ser His Lys Arg Glu Ile

<210> SEQ ID NO 32
<211> LENGTH: 1393
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1709118CB1

<400> SEQUENCE: 32

ggacaggct ggagcagggc ttcocaggtt gcggagagga aacaaccccgt aagctgtgcc 60
aagatgttg cagcagaggag gacoacaggc ctggtgtcgg acacagcctct tggggcttctt 120
aagccgctg tctggcgag ctgagccccc cgggtgtgct tccggctctc cttggggctc 180
ccgcgcaccc agggtttttg gttgtccag tggactagag attcctactcgt cgggtggaaa 240
gcggagacca agacgagccgt ctcgaccttct aactaatcagcct cgggtggtct 300
aacttgccag aacttgccag aacttgccag aacttgccag aacttgccag aacttgccag aacttgccag 360
cgagagcc aaccccagct gctgcagag aggccagctg aacccagagc cagccaggag 420
gaagatcagc agtagcagtt ctgacagctt cccttctact gactgttctgg gcgtgaagcc 480
gaagtctctg ctgctgtact cttggtggtgt cctacagagc tttttgtttt cttggtggat 540
ggctgactc cagccggtcc catccatagag gactgttctgg tggccagtcg cttcaggttg 600
cctggtttct ctgctgggag tccccaggtg actttctatcag attacaggtc 660
tactcttttg tccccaggtg gacagcgtgaa atggctcctgt accataagag ggatggtgc 720	tagtggtcgc tggagtttga gactgttctgg gcctacagctt cttctctcct ctctcttccag 780
aagagtctag aagctgtcag tgcggagtt aactctatcctg gcaagtgctt cttcaggttg 840
<210> SEQ ID NO 33
<211> LENGTH: 377
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURES:
<221> NAME/KEY: misc feature
<222> OTHER INFORMATION: Incyte ID No: 1709118CD

<400> SEQUENCE: 33

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

<210> SEQ ID NO: 34
<211> LENGTH: 2310
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 000513.49
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2307
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 34
cttcctctct ctctcagcct ctcacactc ttcacagcct tcacactcct tcacactcct tcagatcatc 60
gcagagctc ctccacactc ctcacagcct ctcacagcct ctcacagcct ctcacagcct ctcacagcct 120
ttcagagctc ttctctcctc ctcacagcct ctcacagcct ctcacagcct ctcacagcct ctcacagcct 180
caggggcagt gggtggcttg gggtgctttg ggtggtttgg ctcctcgcag ccagtctcctc ctcacagcct 240
gttgcctgc cgggtgttgc ggtgtggtgc ggtgtggtgc ggtgtggtgc ggtgtggtgc ggtgtggtgc 300
cgggtgcctgc gggtgcctgc gggtgcctgc ggtgtggtgc ggtgtggtgc ggtgtggtgc ggtgtggtgc 360
tttctcgct gcctgcctgc ctgctgcctgc ctgctgcctgc ctgctgcctgc ctgctgcctgc ctgctgcctgc 420
ttttgggttc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc 480
cagtgcgcgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc gcgtgtgtgc 540
cggtgctgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc 600
ggctgctgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc 660
ggctgctgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc 720
cagtgcgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc 780
cagtgcgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc gcgtgtgcgc 840
-continued

gcgcgcagc agggaatttg ttagctttgaa ggaaggatggt gtgtctggct cactgaacaa 900

gtgaatcag  taaacgcaac caaaccactt ccaacatcag atcaacctcc tggagagctt 960

gatgcacgca gatgcctgcc agatgcacag ccaactctca gcaacatctg tggagctgtc 1020

catggaacaa acacgccacc tggagctgga cagctatcag tggagctgtca aggcccataa 1080

tggagagtt gtagctagaa gcctgggttg ggtgaaggtgct cgtgtactca gcaagtcaag 1140

ggagctggag gtccagacag tgaacactgg ggcaagacct gccaacacca gcagggagat 1200

tgtagagac accagctgta tcacaggtgt ggaacctagg atcgaccaag tcaagaagca 1260

gtgcccaacctcgcagcgtctagctca tggagttgcag cgtggggaga tggcccctca 1320

ggattgcatc acaccctggt ggagggctgaa ggggcctgtc cagacagccg acagagacct 1380

ggcccggctg cggagagtct accgagctc accgtgagct ggtgaacttc aagctgcccc 1440

gatgcocac tccagcaagg cagctgtgag ggctggagcttg aagctgaagg ggtgaaaggt 1500

tggaaacgtg aactctcttg tggagcactg ccggctctcc gccgtcctat ggctggccag 1560

tggtagcggc aagttgggtcg ggtcctggttg ggaacgagctg tacaccttg gctagtgtctg 1620

tggctggtggct cggctgccag tggagctgac gttgggctgg gcggctggtct 1680

tggtagggc gcagagcccc cccaaaaggg cccctgaggt ttagtaaatag ggtgaaaggg 1740

catataacca taaggtgtcg cttgcagtcgcc gccctcaca agtcctcag cocctctcttg 1800

ggtcagac cgcctctctgt cggtagctt tgtgttttgt gcgtctcttg cgtcctcttc 1860

cacacacagc cgtgcgcgtag agtggggtag cggcccctct ttctctcttc ttctctctcg 1920

<210> SEQ ID NO: 35
<211> LENGTH: 493
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 047568.1

<400> SEQUENCE: 35

cctagctca attttataag aagacaatac atgttataat mtaacttaag atcttttatag 60

aatyyyyag gctatggaca cacattataag ggaagccctt gttctgttgaag gtgtatatgtct 120

gttggcacttg gttgttttgc ccacattact tcacgtgtcgg ggttatttggag gtttttctgc 180

catggaacaa acacgccacc tggagctgga cagctatcag tggagctgtca aggcccataa 240

cggcgcagc agggaatttg ttagctttgaa ggaaggatggt gtgtctggct cactgaacaa 300

agaaggtat ccagaaattttt acggaatatt getatggtaa castgtctct tttagagagt 360

cattttatc tatacttctt cacagacatac aagtttctca aatctcttctt aatctcttctt 420

ttatattata cattattata aatattataa aatattataa aatattataa aatattataa 480
ttgaaatagt tca

<210> SEQ ID NO: 36
<211> LENGTH: 983
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: 
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 312007DCB1

<400> SEQUENCE: 36

ggacacgct ccocgcagcc tctctcttct tgggggggg gcggcccgct cagggcacgc 60
tcctggtcc ggcgggggcc gcggggtgg cggggggtgg ggggggggg 120
gggccgggt cgggggggg ggggggggg ggggggggg ggggggggg 180
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 240
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 300
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 360
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 420
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 480
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 540
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 600
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 660
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 720
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 780
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 840
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 900
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 960
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1020
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1080
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1140
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1200
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1260
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1320
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1380
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1440
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1500
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1560
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1620
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1680
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1740
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1800
tgggggggg ggggggggg ggggggggg ggggggggg ggggggggg 1860
<210> SEQ ID NO: 38
<211> LENGTH: 1516
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1303705CB1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 1512
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 38
ctttgttttt ggacatagct gagccaggtg otctcaacag aagcagcag agataaact 60
ttcgttgtct aggtgttgtg tccttttttt tctcttttct tcagagag acagtctcag 120
tgccagtgtg aagatcttctg atctttagtt ctcttacag aagctcag 180
aaagacactt ttctaaaaat ggcaatgta tcagatctcc tcagaaggcc ctggtttatt
240
gaaatgyag agcaggaata tggcagacct gtyagctaat ccacaaggygg toccygatca
300
gcagttgagc cttagctctac cttaaatagct tcgtcttcgt togtgcttct gcataaggcc
360
atgttgatg aaggctgtgga tcgacaccaacct acattggaca ttcatattaa gcgaacacat
420
gcagcyttac acacgatcc aaagcagcat ctccagggaa cagagaaagcc cctggttgaa
480
acactgyaaga aggccttacat gyygctctct gyygctctct gyygctctct gyygctctct
540
ccgagaatcttgctcttctgacatgagtcgacagtagtgcc agggccttgg aactgtggtga
600
gattctctac ttcagatctcc gcacgcaagagt caggtcctct ccagaacgct caaacctgaac
660
ggcagctgag acgtccttct tcctcctccag ggtgctgctt gcgtgctctg ggtgctgctt
720
aattgtggtgc gcgtctttct gcgtctttct gcgtctttct gcgtctttct gcgtctttct
780
agtgggtctt aggcttacct ctacgtatcgt gttacgctcag ctagctctccacctctccacct
840
gggcagagct tccactctct ctcttcacccg aagsccagaag aagagagaga gctctgtggt
900
agagctcttt acaagatccaa cagcttcttt agccgctctct tcagccttctc tcacggtgtt
960
gaggttagg agctcttcgg acaagtctcc ccgatttgcc gtaaggtggtc gcagacccagaa
1020
ccagctttct tcctcttcagc ctctctctct gcagcttctc gcagcttcagc gcagcttctc
1080
gcattgtct ggtctgttct tccttccttc gcagcttctc gcagcttcagc gcagcttctc
1140
ttccagatc tgcagcttct ctctctctct gcagcttctc gcagcttcagc gcagcttctc
1200
tataagaaaa gatctgttcg totttgtgga gagaactaaa cattcccttg atggtcttca
1260
gtatgcctg gaaactctta attatatatt ccatccttcgcgcagcttctc gcagcttctc
1320
ctccagaggt attacagttg acgtacatc agtgctcagaa atagctgtaaatcctccattt
1380
ctttatatt cactctgtta atagatagatgctttt taaaagagtt tctcctcctcctcct
1440
cataaaaaacc tctacaggt gtttagttaa cattcaggtc gaaagagtytctgtctctg
1500
aataataaat gngtcg
1516

<210> SEQ ID NO 39
<211> LENGTH: 346
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1303785CD1
<400> SEQUENCE:
Met Ala Met Val Ser Glu Phe Leu Lys Gln Ala Trp Phe Ile Glu
1     5     10     15
Asn Glu Glu Glu Tyr Val Gln Thr Val Lys Ser Ser Lys Gly
20    25    30
Gly Pro Gly Ser Ala Val Ser Pro Tyr Pro Thr Phe Aen Pro Ser
35    40    45
Ser Asp Val Ala Leu His Lys Ala Ile Met Val Lys Gly Val
50    55    60
Asp Glu Ala Thr Ile Ile Asp Ile Leu Thr Lys Arg Aen Asn Ala
65    70    75
Gln Arg Gln Glu Ile Lys Ala Ala Tyr Leu Glu Glu Thr Gly Lys
80    85    90
<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>95</td>
<td>100</td>
<td>Pro Leu Asp Glu Thr Leu Lys Ala Leu Thr Gly His Leu Glu</td>
</tr>
<tr>
<td>110</td>
<td>115</td>
<td>Glu Val Val Leu Ala Leu Leu Thr Val Val Gln Phe Asp Ala</td>
</tr>
<tr>
<td>125</td>
<td>130</td>
<td>Asp Glu Leu Arg Ala Ala Met Lys Gly Leu Gly Thr Asp Glu Asp</td>
</tr>
<tr>
<td>140</td>
<td>145</td>
<td>Thr Leu Ile Glu Ile Leu Ala Ser Arg Thr Asn Lys Gly Ile Arg</td>
</tr>
<tr>
<td>155</td>
<td>160</td>
<td>Asp Ile Asn Arg Val Tyr Arg Glu Leu Lys Arg Asp Leu Ala</td>
</tr>
<tr>
<td>170</td>
<td>175</td>
<td>Lys Asp Ile Thr Ser Asp Thr Ser Gly Asp Phe Arg Asn Ala Leu</td>
</tr>
<tr>
<td>185</td>
<td>190</td>
<td>Leu Ser Leu Ala Lys Gly Asp Arg Ser Gly Asp Phe Gly Val Asn</td>
</tr>
<tr>
<td>200</td>
<td>205</td>
<td>Glu Asp Leu Ala Asp Ser Asp Ala Arg Ala Leu Tyr Glu Ala Gly</td>
</tr>
<tr>
<td>215</td>
<td>220</td>
<td>Glu Arg Arg Lys Gly Thr Asp Val Asn Val Phe Asn Thr Ile Leu</td>
</tr>
<tr>
<td>230</td>
<td>235</td>
<td>Thr Thr Arg Ser Tyr Pro Glu Leu Arg Arg Val Phe Glu Lys Tyr</td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>Thr Lys Tyr Ser Lys His Asp Met Asn Lys Val Leu Asp Leu Glu</td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td>Leu Lys Gly Asp Ile Glu Lys Cys Leu Thr Ala Ile Val Lys Cys</td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td>Ala Thr Ser Lys Pro Ala Phe Ala Glu Lys Leu His Glu Ala</td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td>Met Lys Gly Val Gly Thr Arg His Lys Ala Leu Arg Ile Met</td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td>Val Ser Arg Ser Glu Ile Met Asn Asp Ile Lys Ala Phe Tyr</td>
</tr>
<tr>
<td>320</td>
<td>325</td>
<td>Gln Lys Met Tyr Gly Ile Ser Leu Cys Glu Ala Ile Leu Asp Glu</td>
</tr>
<tr>
<td>335</td>
<td>340</td>
<td>Thr Lys Gly Tyr Glu Lys Ile Leu Val Ala Leu Cys Gly Gly</td>
</tr>
<tr>
<td></td>
<td>345</td>
<td>Asn</td>
</tr>
</tbody>
</table>

SEQ ID NO: 40
LENGTH: 2712
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: misc_feature
OTHER INFORMATION: Incyte ID No: 1798379CB!
-continued

ggagccgca gggcttttgag ggctggcaag ccgttctcag gtgggttgctt cggtygtgctt cggtyggctg 420

ggtggtgctt gcgggctgg agttcttttg ccgggctggag gggttttcttt gcggggctgg agtgttcagtg 480

ggtctcttttg ccggggcctct cctggctgca acctacgggt tgtgtcttcag 540
cctggctgcc tggctccctg cacaatggaa 600
cataacgttac ccggacacg aatcggacgc aatccggttac 660
cggctcttttg ccggctccac cccgggttgtt cgggtctgtc cgagaaaaat 720

eatggctggc cccgccccat tgcctgtgctt ctgggcctct ggggtattat cggagttgctt 780

cagccgatccata tggctggctt ggggtgtccg gggggggtggt 840
cctggctccag ggtgggctgg ggtgggtctgt cgggtgtgatt ctgggctgctt 900
ggtttttgtg ggggtgattc gatctatatcg agcggcagcggctgttgc gggacgtcgcctta 960
gggggtgttcag ggtggtcgctc gagacccgctg acgggtctttc ggaggggattttc 1020
ttggtatgcc catgatcattg cggccggtctt cgggtctgttgc ggggtttgctt 1080
ggggtgttgtg ggggtgtttggt cgggttttgcct ggggtgatttc ctgggttttgc 1140
ggggtgttgcc cgggtgatcag ctcgggtttgct ggggtgggtgctt 1200
ggggtgtgttg ggggtgtttcag ctcgggtttcag ctcgggtttcag 1260
ggggtgtttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1320
ttggtatccttg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1380
ggggtgtttgct ggggtgtttcag ctcgggtttcag ctcgggtttcag 1440
ggggtgttttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1500
ggggtgttttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1560
ggggtgttttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1620
ggggtgttttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1680
ggggtgttttctg ctcgggtttcag ctcgggtttcag ctcgggtttcag 1740
tttttttttttt tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<210> SEQ ID NO: 41
<211> LENGTH: 645
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<222> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1798379CD1

<400> SEQUENCE: 41

Met Ser Cys Gln Ile Ser Cys Ser Arg Gly Arg Gly Gly Gly
1      5     10     15
Gly Gly Gly Phe Arg Gly Phe Ser Ser Gly Ser Ala Val Val Ser
20     25     30
Gly Gly Ser Arg Ser Thr Ser Ser Phe Ser Cys Leu Ser Arg
35     40     45
50     55     60
Arg Ser Leu Val Gly Leu Gly Gly Thr Lys Ser Ile Ser Ile Ser
65     70     75
Val Ala Gly Gly Gly Phe Gly Ala Ala Gly Gly Phe Gly
80     85     90
95     100    105
Ser Gly Phe Gly Gly Gly Ser Gly Phe Ser Gly Gly Gly Phe Gly
110    115    120
125    130    135
140    145    150
Gly Tyr Pro Gly Gly Ile His Glu Val Ser Val Aan Gln Ser Leu
155    160    165
Leu Gln Pro Leu Aan Val Lys Val Asp Pro Glu Ile Gln Aan Val
170    175    180
Lys Ala Gln Glu Arg Glu Gln Ile Lys Thr Leu Aan Aen Lys Phe
185    190    195
Ala Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Aan Aen
200    205    210
Val Leu Gln Thr Lys Trp Glu Leu Glu Gln Met Aan Val Gly
215    220    225
Thr Arg Pro Ile Aan Leu Glu Pro Ile Phe Glu Gly Tyr Ile Aas
230    235    240
Ser Leu Lys Arg Tyr Leu Aas Gly Leu Thr Ala Glu Arg Thr Ser
245    250    255
Gln Aan Ser Glu Leu Aan Aen Met Gln Aas Leu Val Glu Asp Tyr
260    265    270
Lys Lys Tyr Glu Asp Glu Ile Aan Lys Arg Thr Ala Ala Glu
275    280    285
Asn Aas Phe Val Thr Leu Lys Aas Val Asp Aan Ala Tyr Met
290    295    300
Ile Lys Val Glu Leu Gln Ser Lys Val Asp Leu Leu Aen Gln Glu
<table>
<thead>
<tr>
<th>305</th>
<th>310</th>
<th>315</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile</td>
<td>Glu</td>
<td>Phe</td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td>Tyr</td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Gln</td>
</tr>
<tr>
<td>Ile</td>
<td>320</td>
<td>325</td>
</tr>
<tr>
<td>His</td>
<td>Gln</td>
<td>Ser</td>
</tr>
<tr>
<td>Val</td>
<td>Thr</td>
<td>Asp</td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Val</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>Met</td>
<td>Asp</td>
<td>Aan</td>
</tr>
<tr>
<td>335</td>
<td>340</td>
<td>345</td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Asn</td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>Ser</td>
</tr>
<tr>
<td>Ile</td>
<td>Ala</td>
<td>Gln</td>
</tr>
<tr>
<td>Val</td>
<td>Lys</td>
<td>Ala</td>
</tr>
<tr>
<td>350</td>
<td>355</td>
<td>360</td>
</tr>
<tr>
<td>Gln</td>
<td>Tyr</td>
<td>Glu</td>
</tr>
<tr>
<td>Glu</td>
<td>Ile</td>
<td>Ala</td>
</tr>
<tr>
<td>Gln</td>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td>Lys</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td>Glu</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>Glu</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>365</td>
<td>370</td>
<td>375</td>
</tr>
<tr>
<td>Leu</td>
<td>Tyr</td>
<td>His</td>
</tr>
<tr>
<td>Ser</td>
<td>Lys</td>
<td>Tyr</td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
</tr>
<tr>
<td>Gln</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>Val</td>
<td>Gly</td>
<td>Arg</td>
</tr>
<tr>
<td>380</td>
<td>385</td>
<td>390</td>
</tr>
<tr>
<td>His</td>
<td>Gly</td>
<td>Asp</td>
</tr>
<tr>
<td>Ser</td>
<td>Leu</td>
<td>Lys</td>
</tr>
<tr>
<td>Ile</td>
<td>Gln</td>
<td>Leu</td>
</tr>
<tr>
<td>Ile</td>
<td>Gln</td>
<td>Ile</td>
</tr>
<tr>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
</tr>
<tr>
<td>395</td>
<td>400</td>
<td>405</td>
</tr>
<tr>
<td>Asn</td>
<td>Arg</td>
<td>Val</td>
</tr>
<tr>
<td>Ile</td>
<td>Gln</td>
<td>Arg</td>
</tr>
<tr>
<td>Leu</td>
<td>Gln</td>
<td>Gly</td>
</tr>
<tr>
<td>Gln</td>
<td>Ile</td>
<td>Ala</td>
</tr>
<tr>
<td>His</td>
<td>Val</td>
<td>Lys</td>
</tr>
<tr>
<td>410</td>
<td>415</td>
<td>420</td>
</tr>
<tr>
<td>Lys</td>
<td>Gln</td>
<td>Cys</td>
</tr>
<tr>
<td>Lys</td>
<td>Asn</td>
<td>Val</td>
</tr>
<tr>
<td>Gln</td>
<td>Gln</td>
<td>Asp</td>
</tr>
<tr>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>Gln</td>
<td>425</td>
<td>430</td>
</tr>
<tr>
<td>Arg</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td>His</td>
<td>Ala</td>
<td>Leu</td>
</tr>
<tr>
<td>Lys</td>
<td>Asp</td>
<td>Ala</td>
</tr>
<tr>
<td>Arg</td>
<td>Asn</td>
<td>Lys</td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td>Aan</td>
</tr>
<tr>
<td>440</td>
<td>445</td>
<td>450</td>
</tr>
<tr>
<td>Leu</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td>Leu</td>
<td>Gln</td>
<td>Gln</td>
</tr>
<tr>
<td>Ala</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>Asp</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td>Arg</td>
<td>Leu</td>
<td>455</td>
</tr>
<tr>
<td>Leu</td>
<td>Arg</td>
<td>Asp</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gln</td>
<td>Glu</td>
</tr>
<tr>
<td>Leu</td>
<td>Met</td>
<td>Aan</td>
</tr>
<tr>
<td>Val</td>
<td>Lys</td>
<td>Leu</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Leu</td>
</tr>
<tr>
<td>470</td>
<td>475</td>
<td>480</td>
</tr>
<tr>
<td>Val</td>
<td>Glu</td>
<td>Ile</td>
</tr>
<tr>
<td>Ala</td>
<td>Thr</td>
<td>Tyr</td>
</tr>
<tr>
<td>Arg</td>
<td>Lys</td>
<td>Leu</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>Cys</td>
</tr>
<tr>
<td>485</td>
<td>490</td>
<td>495</td>
</tr>
<tr>
<td>Arg</td>
<td>Met</td>
<td>Ser</td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>Ser</td>
<td>Aan</td>
<td>Val</td>
</tr>
<tr>
<td>Thr</td>
<td>Val</td>
<td>Ser</td>
</tr>
<tr>
<td>Val</td>
<td>Thr</td>
<td>500</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>Thr</td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Ser</td>
<td>Asn</td>
<td>Lys</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>Phe</td>
</tr>
<tr>
<td>Gly</td>
<td>515</td>
<td>520</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Arg</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Tyr</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>530</td>
<td>535</td>
<td>540</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Gly</td>
<td>Arg</td>
<td>Gln</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Arg</td>
<td>545</td>
<td>550</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Tyr</td>
</tr>
<tr>
<td>Gly</td>
<td>560</td>
<td>565</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Arg</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>575</td>
<td>580</td>
</tr>
<tr>
<td>Lys</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Tyr</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>590</td>
<td>595</td>
</tr>
<tr>
<td>Lys</td>
<td>His</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Gly</td>
<td>605</td>
<td>610</td>
</tr>
<tr>
<td>Gly</td>
<td>Tyr</td>
<td>Gly</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Lys</td>
</tr>
<tr>
<td>Gly</td>
<td>620</td>
<td>625</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Glu</td>
<td>Ala</td>
<td>Phe</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>Phe</td>
<td>Ser</td>
<td>Phe</td>
</tr>
<tr>
<td>Arg</td>
<td>635</td>
<td>640</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 42
<211> LENGTH: 663
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<223> OTHER INFORMATION: Incyte ID No: 350650.1
-continued

cctggcactg aaaaagttcct ctagagtgaac ctggtagttct acgagcgttc tgctgctgtc 60
tcggtgctca gaaaactgaa cggcagcttt cgctgctttcg gggcaggtgg gtcgtgcctc 120
gggtgagact ttgctggagtc gtctgctgct ctgggtggtgct tggcagcttc 180
gaggaggtttg atctgagaac aacatgctg gctgcagctgct gcggtgtggc 240
ttgcaatatt gcttcggagaa cctccggtgc gctgctgcagct ggtgtggtgctg 300
tcttctggta ttgctcagct ttgctcagcttt ccagagacctt cctctccttt 360
gagtctcttc gtttcagcagc ggttcttcct ctttcttcct 420
gcagcttttc cagagccctt gcctctccct gggtctctccct ttgggtctctc 480
gcgagagat ggcacagtctt cagagagagat ggcacagtctt ccagagagagat 540
ggatctgctt tttctctctt ccagctctct ccagctctct 600
ttcacctttt ccagctctct ccagctctctt ctctctctct 660
ttg 663

<210> SEQ ID NO: 43
<211> LENGTH: 809
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID: 474630.24
<223> LOCATION: 5-11
<224> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 43
ccgagcactg gcagcagctg ccccctctgag cccccgtcag cccccagcctg cccccgctcag 60
tgcagcagcag cccccgctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 120
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 180
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 240
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 300
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 360
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 420
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 480
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 540
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 600
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 660
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 720
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 780
tgcagcagcact cccccctcag cccccctctgag cccccggctcag cccccctcag cccccgctcag 840
gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgggagctgg gcgtgtgtgc ttgg
ccctggcagct gcgtttaa gcgtttctgc agagttggcc tgggacggaga 1560
ggggctgaga gggtggggc tiggaaccCct ... togaatacact to atgttcaa 1680
gtattaag ac citatgcaata tttitttacitt ttctaataaa catgtttgtt aaaaacaaaa 1740
aaaa 1744

<210> SEQ ID NO 46
<211> LENGTH: 232
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3346307C01

<400> SEQUENCE: 46
Met Pro Ser Leu Trp Asp Arg Phe Ser Ser Ser Ser Thr Ser Ser Ser
1  5  10  15
Ser Pro Ser Ser Leu Pro Arg Thr Pro Thr Pro Asp Arg Pro Pro
20 25  30
Arg Ser Ala Trp Gly Ser Ala Thr Arg Glu Glu Phe Asp Arg
35 40  45
Ser Thr Ser Leu Glu Ser Ser Asp Cys Glu Ser Leu Asp Ser Ser
50 55  60
Asn Ser Gly Phe Gly Pro Glu Ala Tyr Leu Asp Gly
65 70  75
Val Ser Leu Pro Asp Phe Glu Leu Ser Asp Pro Glu Asp Glu
80 85  90
His Leu Cys Ala Ser Leu Met Glu Leu Leu Glu Glu Ser Leu Ala
95 100 105
Gln Ala Arg Leu Gly Ser Arg Arg Pro Ala Arg Leu Met Pro
110 115 120
Ser Glu Leu Val Ser Gln Val Val Lys Glu Leu Leu Arg Leu Ala
125 130 135
Tyr Ser Glu Pro Cys Gly Leu Arg Gly Ala Leu Leu Asp Val Cys
140 145 150
Val Glu Gln Gly Lys Ser Cys His Ser Val Gly Gln Leu Ala Leu
155 160 165
Asp Pro Ser Leu Val Pro Thr Phe Gln Leu Thr Leu Val Leu Arg
170 175 180
Leu Asp Ser Arg Leu Trp Pro Lys Ile Gln Gly Leu Phe Ser Ser
185 190 195
Ala Asn Ser Pro Phe Leu Pro Gly Phe Ser Glu Ser Leu Thr Leu
200 205 210
Ser Thr Gly Phe Arg Val Ile Lys Lys Leu Tyr Ser Ser Glu
215 220 225
Gln Leu Leu Ile Glu Glu Gly Cys
230

<210> SEQ ID NO 47
<211> LENGTH: 897
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
-continued

taccagacca agtccaggg aacctcaagtt tctgcccagc ttcatgggga caggtacgcg 1020
gaaacgaag acctcagctc tcgctcacc caagacagt acagggctca gacgaactc 1080
gagacactca agaacgaaag cgcagcctg cagcccgtca ctcactgtag tcgagcagtt 1140
 ggagctcgtg cctcaccacc ggtggcagcc aaggggtcag aactgagacgc ctcactgttgg 1200
atgccccagc agaacagtggc cccggtcgtg tgcagacacc agagccgtac gacgacgcc 1260
ccttcctcagc atggtgagat tygcaacttc acgcagcgtg tcacaggyoga ggcagcagg 1320
atgtctgaga agtccacacc cccagtcctt atcctctcag tggagagacgc ctcactgttgg 1380
tctggagagc ttcctgagag cttgggcgtc actctgagac ccctgctagag gcagcaggctg 1440
cctgacgtgc gctgcaacct cagcactgtc ggcagcaggc acactttctc cctgctagag 1500
aagacacagt tttttgttgc ctgtctgctg tctgctgctc aagcctaacc gctgctgctc 1560
aacacactga agaacgcgt tgcacgtgc agtgaagact cccacttaaa cccacttaaa 1620
cacgacgcc cctccctcag aaccacttgg gacacctccc cccacttaaa aacggtacct ctgaagaccc 1680
ggacacecc cttgctggtg ccacaggcc aaccggtacct aaccggtacct ctgaagaccc 1740
cataacagc tocccctcagc ctgtctgagc cttcctgaaag cctctagact ctgaagaccc 1800
aacaacagc atgtctgact cccacacc 1827

<210> SEQ ID NO 49
<211> LENGTH: 3936
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1088524.8
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2060-2170, 3796, 3799, 3816
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 49

taccagacca tgcgtgcagc aacactttca cgggaagggc ccctcgagcc caggaaggtc 60
tctctgcgct tggagagagc gtgcagccga aggcgtccca acggagaggc ttttgctggtt 120

tctgaaaccg acgagcttgc acacccgctc ctcagctctg caagcttggc acagagttgtc 180

cccacacccg cccagacgtc cccagcccgtc ctcctcgccg cctccacagg cagacggtctc 240

gtgcagccgc cccggtcctt tggagaggtc cccagctgtg ctcctcgccg cagacggtctc 300
cagccggtc cccggtcctt tggagaggtc cccagctgtg ctcctcgccg cagacggtctc 360

cctggttcgca gtgtccctgct ggcagagagc agggagcgcttc cagtcctgtg ctcctcgccg 420

gctgttgca ggcagagagc ctcctgtgct ggcagagagc agggagcgcttc cagtcctgtg ctcctcgccg 480

<210> SEQ ID NO 49
<211> LENGTH: 3936
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1088524.8
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2060-2170, 3796, 3799, 3816
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 49

taccagacca tgcgtgcagc aacactttca cgggaagggc ccctcgagcc caggaaggtc 60
tctctgcgct tggagagagc gtgcagccga aggcgtccca acggagaggc ttttgctggtt 120

tctgaaaccg acgagcttgc acacccgctc ctcagctctg caagcttggc acagagttgtc 180

cccacacccg cccagacgtc cccagcccgtc ctcctcgccg cctccacagg cagacggtctc 240

gtgcagccgc cccggtcctt tggagaggtc cccagctgtg ctcctcgccg cagacggtctc 300
cagccggtc cccggtcctt tggagaggtc cccagctgtg ctcctcgccg cagacggtctc 360

cctggttcgca gtgtccctgct ggcagagagc agggagcgcttc cagtcctgtg ctcctcgccg 420

gctgttgca ggcagagagc ctcctgtgct ggcagagagc agggagcgcttc cagtcctgtg ctcctcgccg 480

<210> SEQ ID NO 49
<211> LENGTH: 3936
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1088524.8
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2060-2170, 3796, 3799, 3816
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 49

taccagacca tgcgtgcagc aacactttca cgggaagggc ccctcgagcc caggaaggtc 60

-continued

tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 965
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 960
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1020
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1080
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1140
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1200
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1260
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1320
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1380
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1440
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1500
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1560
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1620
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1680
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1740
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1800
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1860
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1920
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 1980
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2040
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2100
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2160
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2220
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2280
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2340
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2400
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2460
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2520
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2580
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2640
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2700
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2760
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2820
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2880
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 2940
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 3000
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 3060
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 3120
tgagactgata gctctaggttg cctctggtgc cctctggtgc ccctctggtgc ccctctggtgc ccctctggtgc 3180
ggsccaaccg tyygtgacat cagcgtgccg cccaccccg ccagcagaa gcccaccaag 3240
cagcgtgccg cccaccccg ccagcagaa gcccaccaag 3300
cagcgtgccg cccaccccg ccagcagaa gcccaccaag 3360
tccacccct cccat gagcc tgcgtgccg cccaccccg ccagcagaa gcccaccaag 3420
ggtgcaccc cccgctgcg cccaccccg ccagcagaa gcccaccaag 3480
tggtgccca ccccgccggc cccaccccg ccagcagaa gcccaccaag 3540
ggcgcgcg ggcgcgcg ggcgcgcg cccaccccg ccagcagaa gcccaccaag 3600
ggcgcgcg ggcgcgcg cccaccccg ccagcagaa gcccaccaag 3660
ggcgcgcg ggcgcgcg cccaccccg ccagcagaa gcccaccaag 3720
cccacgctc cccgctgcg cccaccccg ccagcagaa gcccaccaag 3780
ttgccgctc cccgctgcg cccaccccg ccagcagaa gcccaccaag 3840
agccctgtc ttggcgtctc cccgctgcg cccaccccg ccagcagaa gcccaccaag 3900
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 3960

<210> SEQ ID NO 50
<211> LENGTH: 1114
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte Id No: 632664CE1

<400> SEQUENCE: 50
ggcgcgcgtc cccgctgcg cccaccccg ccagcagaa gcccaccaag 60
cccgctgcg cccaccccg ccagcagaa gcccaccaag 120
cccgctgcg cccaccccg ccagcagaa gcccaccaag 180
ggcgcgcgtc cccgctgcg cccaccccg ccagcagaa gcccaccaag 240
cccgctgcg cccgctgcg cccaccccg ccagcagaa gcccaccaag 300
ggcgcgcgtc cccgctgcg cccaccccg ccagcagaa gcccaccaag 360
ggcgcgcgtc cccgctgcg cccaccccg ccagcagaa gcccaccaag 420
ggcgcgcgtc cccgctgcg cccaccccg ccagcagaa gcccaccaag 480
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 540
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 600
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 660
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 720
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 780
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 840
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 900
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 960
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 1020
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 1080
tggtgccca cccgctgcg cccaccccg ccagcagaa gcccaccaag 1140
<210> SEQ ID NO 51
<211> LENGTH: 266
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 632664CD1

<400> SEQUENCE: 51

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

<210> SEQ ID NO 52
<211> LENGTH: 1189
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 457372.17

<400> SEQUENCE: 52

acaggtgtga gccacocacac caacagttt ttttaaggct acaaataatg aagaactgga 60
tttgtgacca gtcttgttgc acactaaata gagtgcctta caatttatgt gcagtctgca 120
tttgccaatgc cagcagacct ctttgaatct ctttccccct ctggcaaaaca gtaacccatat 180
cctactgcga acctctaaag tggtagttgag agatccaaag aatgaaaaat tttaaaatat 240
tgtaacatct atcctcaatttaa aagcttattgg ggcgaggtat gtctgactcag gatttgggt 300
tttttactct gcacccccat aagagctttg cagttgctccat cagaaagctt tatcattata 360
tccacaagct accataagggc tttgctracat gtatacctag acaaataact ttaacagaaac 420
tatatgtggc tcttttacaa cttttctctag aaggttcat cttaacatag agttgccagg 480
tttgaccttt ctccctcatct acattttgaa ttgtaacact tttgactatt ctaaatttatt 540
tgagtttatcttat ttttccaaag aatctttataa aatattagat tttttttcag 600
ctctttttgg ccctgctgagat aagttgaaat aaccatcatt ctaacacggt ttaaagctac 660
gatactgtgaa caaattgcaaa aagagatgct tttttagttg cttttaagct 720
cgttttattgt gtaacagata gtagctgtct tttttttttt tttttttttt tttttttttt 780
ggttttgcttt ttgtttttttt ctttccccctta taactttgtct ctttttctttta 840
tgcctctggaa aactccctcgg caatccaggt atttattttg cttaaagctg aagatgtgaa 900
gcagagtcttt ttttttttcttt aacctccccacttta ttccacagaa ggcgagatga 960
agacggaacc agggagccata accattattag ggtgtgcgag accatgattg tggagctttta 1020
aagggcagagtttgc ccacagctgg acgttttaag gtagctgacg actgctgtct 1080
tggaaatgtg tccttgggagg tacattgaggg aagaaaccaact gtagtgccag 1140
cccttacccctt ctaccccttag cttgagggag agatctattc 1189

<210> SEQ ID NO: 53
<211> LENGTH: 2539
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: Misc_feature
<222> OTHER INFORMATION: Incyte ID No: 2993696CB1
<400> SEQUENCE: 53

ctcagaagcg aacagcagcc agaaagagt tgctgatcggcag agtggtgag 60
gggacgacg cggcctgctat ttctgaccct gcctctgcc gccgtctgtg 120
cctgccggcct cggccagctt caaaggagaa agggggtgtct gctctgcgct cagcyctggt 180
ggggtgagctc gcgtctgctc ccacattggg tgcacaacagt gcgcctcccc cggcgccgctc 240
gttgtgtctg tgggttggagct cggccggtg ccagggagag gacaagcaggg aggggctggag 300
cacaagtggtc gcagctgaccc ttgggacccaga ctacttcgta gcggcgcgctca gcgcctccc 360
cagcgctgag aggctgaccc gagctgaccc atgggcagctt cagctgctt cggcctctcc 420
cactcgacca ggagagcgcg tggactgcca ccagcagccc gagaagctt cctctcccct 480
cggagacgcg ctctgggcag cccagccgct cctcgtccgc ccagcagctc ccagcctccc 540
gcgacgacag atcaagcttt gtcgctccct gcgtctgcagag aagaacatgg gagaagctct 600
tcaagtggtt attggtgagg gcgacccacc gcagttgtcttc gctgaaacagt ttcttgccct 660
ggtgcctcct caatgtgagag aacgcggactg ggttattggt ggaaagaagg ctgctctcgc 720
taggtggtact gttcagctgct tttattttta ggcgtccaccg cgaagcccag aagagctgg 780
-continued-

```
aactattgct ggcctaaagt ttatgaggt actctaaacgc gctcagccag ctgtctcttgcc 840
ttaggcctct gcaagagggg agggagaaa gaacacctcg gtccttgcacc tgtttggcggg 900
aaccctccgct gcccttccttc cacccattgaa caatggtgtc ttcgaagtttg tgggaccaaa 960
ctggagatct cacttgcggt gagaagactt tgcaccaagct gctatggaac actctoaaca 1020
actgtacaaa aagaaagaggg ccaagaagct aatagagctg tgcagaaac ttcggagaact 1080
cgcggcgcag tgtgaaacag ccgaacgggc cctgttcttc tacaatccag caagaaattga 1140
aatggagtcct ttcttgaag gagaaactct ttcggagacc ctgactcggg ccacacttga 1200
agagctcaac atggacagtct ttccgttcac ttcggaagacc gtcggaagac tgtgggaaga 1260
tctctgattg aagaaagtcgg atatgatgta atctctgtct ttttgtctgtgct gcctactgat 1320
tccaaaggt cgcgcactcg tttaaagatttt cttcaatggcg aaggaacctcg ccgcttggtgat 1380
aaaccaagat gacgctgtag ccgctaggct tgcctgctccag gctgtgctggc ctctctgtga 1440
tcaagatcag ggtgctgccg tgcctgtgca tgcctgctgca ccctacacggt ggatgaaac 1500
tggtggagaag tgtgtgaaacc aactgtgacct aagggccaagct ggtggcgctc caagaagcct 1560
tcgctatcttt tctctagacgct ttgtcctaag ccacactgtt cccactcagga tctctgaaga 1620
tggaacacc ctagcnaaag acacacactt tcttggtctt ccggtatgta ctggaaacctcc 1680
tctggtctct ctggggtcacc cacaagattc aggacaccttt gcagataagt tgaataatgtgat 1740
tctccgccg aacgctgag acatggcctt ggccgacaa aataagacac caaagcctca 1800
tgaggagat ctcgttggtgct ctagagatatt tgggtgatgtg tctgagattt 1860
tggtgaggaa gacagaaaag tcaagggagcg catttgactc aagactgctg tgggagatact 1920
tgctacatct ctaaagactc agatggagaa taagaaagat cgtggaggtta aacttcttcc 1980
tggaagatgcc cgagctgacc aagaaagtcg atgaagagggc cttggagaga 2040
tgaggagatc gcaccttgag aacctcagag taagaagagag gacgtaagaa aactttgctca 2100
accacttacto agacacctct atgggaagctgc agggacccacc cacactgtgtc gagaggtacc 2160
agcagaaaaat gcgacttggtg agacacctgt ctgctgctccgc tggtaatattt aasactcttg 2220
acccggcag ttcctgctgg aaaaaataag gaacatctta aacccgctatg gtcctggattt 2280
tctcactcag cacgtggattg gaaactgctag tggcctcttc gctcttccat 2340
tacgtagtctgcc tccgttgggg ggacgacca ggaagcgagaa ctggagccct ttctttaaaaa 2400
cgggaaaaaa actcggggta cyggtgccgt tctcctcagag atggctcct cttaaacaact 2460
ggtcatgtgcc atctctgtct ggaaggttctt tctccacactt gtcacccaa taagatatgtg 2520
cttatttccct tccgttggcc 2539
```

<210> SEQ_ID: 54
<211> LENGTH: 654
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2993696CD1
<400> SEQUENCE: 54

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

Ser Gly Asp Gln Asp Thr Gly Asp Leu Val Leu Leu Asp Val Val Cys 410 415 420
Pro Leu Thr Leu Gly Ile Glu Thr Val Gly Gly Val Met Thr Lys 425 430 435
Leu Ile Pro Arg Asn Thr Val Val Pro Thr Lys Lys Ser Gln Ile 440 445 450
Phe Ser Thr Ala Ser Asp Asn Gln Pro Thr Val Thr Ile Lys Val 455 460 465
Tyr Glu Gly Glu Arg Pro Leu Thr Lys Asn His Leu Leu Gly 470 475 480
Thr Phe Asp Leu Thr Gly Ile Pro Pro Ala Pro Arg Gly Val Pro 485 490 495
Gln Ile Glu Val Thr Phe Glu Ile Asp Val Asn Gly Ile Leu Arg 500 505 510
Val Thr Ala Glu Asp Lys Gly Thr Gly Asn Lys Asn Lys Ile Thr 515 520 525
Ile Thr Asn Asp Gln Asn Arg Leu Thr Pro Glu Ile Glu Arg 530 535 540
Met Val Asn Asp Ala Glu Lys Phe Ala Glu Glu Asp Lys Lys Leu 545 550 555
Lys Glu Arg Ile Asp Thr Arg Asn Glu Leu Glu Ser Tyr Ala Tyr 560 565 570
Ser Leu Lys Asn Glu Ile Gly Asp Lys Leu Gly Gly Gly Lys 575 580 585
Leu Ser Ser Glu Asp Lys Glu Thr Met Glu Lys Ala Val Glu Glu 590 595 600
Lys Ile Glu Trp Leu Glu Ser His Gln Asp Ala Asp Ile Glu Asp 605 610 615
Phe Lys Ala Lys Lys Glu Leu Glu Gly Ile Val Glu Pro Ile 620 625 630
Ile Ser Lys Leu Tyr Gly Ser Ala Gly Pro Pro Thr Gly Glu 635 640 645
Glu Aep Thr Ala Glu Lys Asp Glu Leu 650

<210> SEQ ID NO 55
<211> LENGTH: 5762
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 331106.6
<400> SEQUENCE: 55

ggcgcacgct cccccgggtg ggccggcgcg cagcgccgag aqaqgagcag ggtggtcgcg  60
gtacgcacgc gcggcgcacc tcgtaaccia gcgggagcg caggcgcggc gcggcagttc 120
cccgctcccc acggcgcctg tgcgcgtagc gcggcgcgcg ccgcgcgttg cttgccttac 180
cggcgtcgcc ggcgctcggt cggccttcgc gcgccttgcc accggcagaa cggcgcagac 240
cgcgctccgg cgacgtgcag gcgcgctgcc gcgccttggt gcgcgctgcg gcgcgctggcg 300
tgccacgtgc gcggagcggc gcggagcgag gcgcgctgcc gcgcgctgcg gcgcgctggcg 360
cggcgtgcag gcgcgctgcg gcgcgctgcg gcgcgctgcg gcgcgctgcg gcgcgctgcg 420
aacattcagt ggccaaaga aattgcaast gggaatgtt tggtttttttt ggtgaataga 2760

aatcocagat gatgaaaaa aagactctgt gacccaaaag aaggatatata ctcctgaaagctc 1730

carecgagat gtcctcacgc aagaaagaa aaaaatatgt gacgcacgaa ctctctctctctc 4800

catcagcgtt aatccatca ggtgtggctg ggtggatcat tactatatcggtactgcaactc 1730

ccttttactactactactact 4800
<table>
<thead>
<tr>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>tgcctgtgttt tggacagact ggaacagctt gctctatgccc aacctgtttc tgggggggcc 5040</td>
</tr>
<tr>
<td>ttcttctttt cgggacactt aasactacaa atacactgaa cggctttgttg gttgtacttt 5100</td>
</tr>
<tr>
<td>gggaccttata tctcactact tacaactatt tctgattaaac aggtactaaa 5160</td>
</tr>
<tr>
<td>asgtctttaaa gtcgaagctt tctgtctggg gaactgtgct ctcctatagt ttttagttgc 5220</td>
</tr>
<tr>
<td>taggtggctga aggacgtgac atcttgacag gctctctgta aagataaaag 5280</td>
</tr>
<tr>
<td>aagctgtctcc tggagaggggt ttgcctcaac aagaaacaaa gaagttatgg 5340</td>
</tr>
<tr>
<td>tggatagatt tgggtgtgat aaaaagctca tctcaagcgt aagctactgtt gttggtcgt 5400</td>
</tr>
<tr>
<td>ttcgactcttt tcagctcttt aagagtaactt caggaagacg tggacagcag aagta 5460</td>
</tr>
<tr>
<td>ggatctttttt atacagttgc gaaactattt caaatcatct tcttgctacc tgtacttca 5520</td>
</tr>
<tr>
<td>cagactgtaa tagtagaagc aacatatcta atataataag caaacctgga atatatattt 5580</td>
</tr>
<tr>
<td>caatctcaag aagctactgtt ttcatttttt tgggtctgct tgaatatttta gtaaagcaaa 5640</td>
</tr>
<tr>
<td>gggactatgt gccaaacgaag cggatctggg gcgtttttttt gtttttctgt ctatctgggt 5700</td>
</tr>
<tr>
<td>gactctcaag ttttttaattt ttctttactta tgattaaag aacactcaag gttttaaca 5760</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO: 56
<211> LENGTH: 2471
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATIOM: Incyte ID No: 1256895CBL
<400> SEQUENCE: 56

<table>
<thead>
<tr>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>cccgctgctc gctgctctgc gacgcagctg ctgacagccg cggcggcgcg agctctctct 60</td>
</tr>
<tr>
<td>ctcgcgtacg ccggcgccg cgcagccatt ttcgctgctg cttcgtgcttc 120</td>
</tr>
<tr>
<td>ttctttgtgc cacatctagt gacgccgggc ttctgcaagaa ggccggaagg ccctgcaggg 180</td>
</tr>
<tr>
<td>ggacacgctt ggccgacgac tctgcggggc agggccggcc gcgtctcagcc 240</td>
</tr>
<tr>
<td>tctcgggggg ggctgtctgg ggccggcgtc atggttgttg cttggggccag ctcctatgtgc 300</td>
</tr>
<tr>
<td>gccgcttgcc gcggcccctt gccggggctt ggccgctctgt gcgtgggggct 360</td>
</tr>
<tr>
<td>aggaggctgg gcacccacca cagctgaaa cagcgagctt cggccaaacc aacagtaa 420</td>
</tr>
<tr>
<td>aagctgctgg ctcgactgac gctggggggcag tttggggggc gcgtggggtc 480</td>
</tr>
<tr>
<td>gaaagctccg gacgacccgc atccataact ccgtagagta caatactact cgactcact 540</td>
</tr>
<tr>
<td>gtaaagacct gcacccacca cccaccaag tctatactgct gacgaccaca 600</td>
</tr>
<tr>
<td>accacacaa ctgcctgctac gacgacgctc aatcatactt cccaccaaaacc aacgacccaa 660</td>
</tr>
<tr>
<td>caacaggacca ggggaaac tccacccaga ccaacgtaaa gatgtggag gcggcgtggtt 720</td>
</tr>
<tr>
<td>aggagctctg ttcacccacg tggcaggg aatctccggc cctcttacag agaggctgca 780</td>
</tr>
<tr>
<td>gcagctctt cttctctctt ccaactcgtga tcctctctgt cttcctctt aotctctct 840</td>
</tr>
<tr>
<td>taggctgttg aggacgctgc tctgtcttttc ccacccttttc taatctttttt ccagcttgag 900</td>
</tr>
<tr>
<td>ggagggctga ccctctgtcgc goccccttttg gttgggtgtc tcaatcttttt ttototottt 960</td>
</tr>
<tr>
<td>gtcctggata ggtctataca cttcctgaggc acgtactggg gcaagaaacc atagagaata 1020</td>
</tr>
<tr>
<td>cctgagagcg tcgtgctgcc ccctttgttt gctcccttac tggcctgtgg caaatgctaa 1080</td>
</tr>
<tr>
<td>gccagtaaag gctataaacgc aataaacact tggtaatact gcacattatttg ggactattag 1140</td>
</tr>
</tbody>
</table>
aaaaaaa a

<210> SEQ ID NO 57
<211> LENGTH: 253
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1256895CD1
<400> SEQUENCE: 57

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

Lys Thr Arg Met Lys His Met Ala Gly Ala Ala Ala Gly Ala
110 115 120

Val Val Gly Gly Tyr Val Leu Gly Ser Ala Met Ser
125 130 135

Arg Pro Ile Ile His Phe Gly Ser Asp Tyr Glu Asp Arg Tyr Tyr
140 145 150

Arg Glu Asn Met His Arg Tyr Pro Asn Glu Val Tyr Tyr Arg Pro
155 160 165

Met Asp Glu Tyr Ser Asn Glu Asn Phe Val His Asp Cys Val
170 175 180

Asn Ile Thr Ile Lys Gln His Thr Val Thr Thr Thr Lys Gly
185 190 195

Glu Asn Phe Thr Glu Thr Asp Val Lys Met Met Glu Arg Val Val
200 205 210

Glu Gln Met Cys Ile Thr Gln Tyr Glu Arg Ser Glu Ala Tyr
215 220 225

Tyr Glu Arg Gly Ser Ser Met Val Leu Phe Ser Ser Ser Pro Pro Val
230 235 240

Ile Leu Leu Ile Ser Phe Leu Ile Phe Leu Ile Phe Val
245 250

<210> SEQ ID NO: 58
<211> LENGTH: 5681
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 474630.29

<400> SEQUENCE: 58

ccggcgcccg cccctgctgc gcggcgccgc tggagccggt cacgtgacgtg 60

ccggcgccag gcggcgccgc gcggcgccgc cccctgctgc gcggcgccgc gcggcgccgc 120

agccggcgcc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 180

ccggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 240

ccggcgcc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 300

cctgctgcgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 360

ccggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 420

agccggcgcc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 480

gctgctgcgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 540

cggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 600

tcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 660

tcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 720

agccggcgcc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 780

tcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 840

tcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 900

ccggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc gcggcgccgc 960
<table>
<thead>
<tr>
<th>ATGAGGCCTGA</th>
<th>TGGCACCACG</th>
<th>GTGCTGGCTG</th>
<th>GCATCTAGAGC</th>
<th>CCGCAACGAT</th>
<th>GACGTTGCC</th>
<th>1020</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACGTACACGAC</td>
<td>CAGCGCACC</td>
<td>CACCAAGCTG</td>
<td>AAGACACAAA</td>
<td>TCCACCCCAT</td>
<td>CTTGCTGCTC</td>
<td>1140</td>
</tr>
<tr>
<td>CCTCGTGCGC</td>
<td>CTGCTGCCG</td>
<td>AGCAAGGAA</td>
<td>TCACCCCTG</td>
<td>CTGTTGCTAC</td>
<td>ACCAAGTCT</td>
<td>1200</td>
</tr>
<tr>
<td>CTTTGTTGAC</td>
<td>TCTGACTG</td>
<td>ACCACATTTGC</td>
<td>TACCTGCTG</td>
<td>GGTTTACGCG</td>
<td>TGGGTGCTG</td>
<td>1260</td>
</tr>
<tr>
<td>CCAGAAGCGT</td>
<td>GCTCAACACG</td>
<td>GTGAGCTGCG</td>
<td>TGGAAGGAGCG</td>
<td>TCCATCGG</td>
<td>ACGCTCTCA</td>
<td>1320</td>
</tr>
<tr>
<td>ACTGAGTTA</td>
<td>CCGGGCCTCAG</td>
<td>GACGAGCGG</td>
<td>GGGCCTCCTG</td>
<td>GCAGAACAGG</td>
<td>ACTTTCAGA</td>
<td>1380</td>
</tr>
<tr>
<td>TGCCGAGAA</td>
<td>GAGCAGACG</td>
<td>GGTCTTCTC</td>
<td>ACATCCGGCG</td>
<td>GGTTGAAGGG</td>
<td>GGTATATTCC</td>
<td>1440</td>
</tr>
<tr>
<td>CGTACGCGC</td>
<td>GGGCCTCTT</td>
<td>GACCGCTGCG</td>
<td>ATGGGACCA</td>
<td>CAGTGCGGAC</td>
<td>TGGCGGAGA</td>
<td>1500</td>
</tr>
<tr>
<td>CCTCCGCTG</td>
<td>TGTACGCTG</td>
<td>ACTTCCGACG</td>
<td>TCAAGAGAAGA</td>
<td>GGTGCGCTCA</td>
<td>GCTCGTTCGA</td>
<td>1560</td>
</tr>
<tr>
<td>GTACGACG</td>
<td>AGACTTGGTG</td>
<td>TCACGAGAGT</td>
<td>GTGGTGCGG</td>
<td>CGAGGCCTGA</td>
<td>AGTGGGGCGA</td>
<td>1620</td>
</tr>
<tr>
<td>CTTGGAGCTG</td>
<td>GTCACCCCGG</td>
<td>CTGCTGAGT</td>
<td>ACACCCGCG</td>
<td>CTGCGTCCGG</td>
<td>GGGGCGAAG</td>
<td>1680</td>
</tr>
<tr>
<td>ACGTACGCGT</td>
<td>CCGGGCGGCT</td>
<td>GGGTGGGAG</td>
<td>AGTGGGGCGG</td>
<td>CTGCTGAGT</td>
<td>ACGGGTACG</td>
<td>1740</td>
</tr>
<tr>
<td>GCAGACACG</td>
<td>GGTAGCATT</td>
<td>GACAAGCACG</td>
<td>CTTGCACTGG</td>
<td>ACTACCCCGG</td>
<td>ACTCCCGGCTG</td>
<td>1800</td>
</tr>
<tr>
<td>TCTTTGAA</td>
<td>TGACCGAGAA</td>
<td>GCTGCTCA</td>
<td>TGGGCAAGTG</td>
<td>GTTTGCGGAG</td>
<td>CTGCGTGGGA</td>
<td>1860</td>
</tr>
<tr>
<td>CGGGCAAGC</td>
<td>GTTCGAGCTG</td>
<td>CCGCTCAGA</td>
<td>AGCCGACCTG</td>
<td>CTACGAGCAG</td>
<td>AAGTGGGGGA</td>
<td>1920</td>
</tr>
<tr>
<td>TCTTAAAGG</td>
<td>ACGTGCGGCA</td>
<td>TCTGGAGTGG</td>
<td>GCGTGGCCG</td>
<td>CTGGCAAGCG</td>
<td>AGTTCGCTT</td>
<td>1980</td>
</tr>
<tr>
<td>AGCAGCCGC</td>
<td>CACTGCGGAC</td>
<td>ACAATCAAAC</td>
<td>CCGGATCCTA</td>
<td>CCCGCGCTGCT</td>
<td>TGCCGAGC</td>
<td>2040</td>
</tr>
<tr>
<td>TAAGTCTCG</td>
<td>GTCTGAGTCG</td>
<td>CGGCGTGCCG</td>
<td>GGAGCGCGGA</td>
<td>GAAGAAGGGG</td>
<td>CGCACTGGTG</td>
<td>2100</td>
</tr>
<tr>
<td>AAGATGCGCA</td>
<td>CTGCTAGTGC</td>
<td>AAGAGGGTG</td>
<td>ACGAGTTAA</td>
<td>GAGAAGCGAG</td>
<td>GGGATGCTG</td>
<td>2160</td>
</tr>
<tr>
<td>TGCTGCTGTG</td>
<td>CGGCGGCGG</td>
<td>CAGGAGCGC</td>
<td>ATCACCACCT</td>
<td>CAGCTGCACG</td>
<td>AGTGGGCTG</td>
<td>2220</td>
</tr>
<tr>
<td>ACCTGGGACC</td>
<td>TGGGCACCG</td>
<td>AGCTGCTGC</td>
<td>TGGCGACGG</td>
<td>AGGACGAGAAGA</td>
<td>TGGCCTCGGG</td>
<td>2280</td>
</tr>
<tr>
<td>GCTCCTGTGG</td>
<td>GTGCTCTATT</td>
<td>CCGCTCAGT</td>
<td>TCCGCTCGG</td>
<td>GCACGCTGGG</td>
<td>GCGCTCCTGG</td>
<td>2340</td>
</tr>
<tr>
<td>TCTGCTGATAG</td>
<td>TGCGTCTGG</td>
<td>GCAGGGCGTGG</td>
<td>CTTGCGCTCT</td>
<td>CTGCGCTGGT</td>
<td>GACGCTCCTGG</td>
<td>2400</td>
</tr>
<tr>
<td>GCACGAGGAG</td>
<td>TCCATCATGG</td>
<td>GCCATTTGG</td>
<td>AGGAGACCTA</td>
<td>GAGTGGATGG</td>
<td>GCAGGCACTA</td>
<td>2460</td>
</tr>
<tr>
<td>TGCTGGTGCT</td>
<td>GAAGCTGACG</td>
<td>TGCGCTCCCG</td>
<td>GAGGTACGG</td>
<td>GGATCCTGAGG</td>
<td>GGGGCGAGG</td>
<td>2520</td>
</tr>
<tr>
<td>TGGACACAG</td>
<td>GACACAGTCC</td>
<td>ACGACCACTG</td>
<td>AGCTTGGTGG</td>
<td>ATGTTGACGAT</td>
<td>AAACAGACG</td>
<td>2580</td>
</tr>
<tr>
<td>CGAAAGTCCG</td>
<td>GCTGACGTTC</td>
<td>GACGGATTACG</td>
<td>GGTGCGTCTG</td>
<td>ACGGCGGCCG</td>
<td>GCGCTCGCGG</td>
<td>2640</td>
</tr>
<tr>
<td>AGAACTCGA</td>
<td>AAGTGCTTAC</td>
<td>AAGGCACTT</td>
<td>CGGTTGGCA</td>
<td>CACGTCCGCAG</td>
<td>CAGACAGGT</td>
<td>2700</td>
</tr>
<tr>
<td>GCCCGGCGC</td>
<td>GGCAAGCGCG</td>
<td>GGGAAAGG</td>
<td>TCTTTGGGAA</td>
<td>AGGAGCCG</td>
<td>AATGCTGTTT</td>
<td>2760</td>
</tr>
<tr>
<td>TGCGCCCGG</td>
<td>CCGGCGCGCA</td>
<td>CCGCGCGGC</td>
<td>TGAAGCTTACG</td>
<td>AGGGAAGGG</td>
<td>ATTGATCGG</td>
<td>2820</td>
</tr>
<tr>
<td>GGGCTTCTG</td>
<td>AGCTCCTGGG</td>
<td>GTCACCCCGG</td>
<td>GTTCACTGCC</td>
<td>CTTTCTCAGG</td>
<td>GCCAGGCCAG</td>
<td>2880</td>
</tr>
<tr>
<td>CCGGGCTCCT</td>
<td>GGCGGCTCCG</td>
<td>TTGAGAGAGTGG</td>
<td>GTTGTTGCACG</td>
<td>AGCGCCGACC</td>
<td>GTGGAGCCGA</td>
<td>2940</td>
</tr>
<tr>
<td>CCGGGCCGAT</td>
<td>GTGGGAGGCTTCGGAGG</td>
<td>GTGAGGGG</td>
<td>AGGAGCTGGTGG</td>
<td>ACGTCGCTC</td>
<td>GGGCCGAGCG</td>
<td>3000</td>
</tr>
<tr>
<td>TTCTGGCGC</td>
<td>TGAGGATCC</td>
<td>GAGCAGAGGAG</td>
<td>AGCTGTTGCTG</td>
<td>GAGGGCGGAT</td>
<td>GGTGGGCTGG</td>
<td>3060</td>
</tr>
<tr>
<td>CAGCGACTC</td>
<td>GACCGCCGCG</td>
<td>CGCGGCGCG</td>
<td>TCACCCCGG</td>
<td>CATCACGAGG</td>
<td>GCAGCCAGCA</td>
<td>3120</td>
</tr>
<tr>
<td>GAGCGCATG</td>
<td>TCTTCTGGG</td>
<td>GAGTGGTCTG</td>
<td>TCTTTGCG</td>
<td>CCGGGCGGC</td>
<td>CAGGCGCCG</td>
<td>3180</td>
</tr>
<tr>
<td>GCACGACTG</td>
<td>GTCTCGCTCCG</td>
<td>GCAGGAGGCCC</td>
<td>GCCGCTGCCG</td>
<td>GCAGGCGGCG</td>
<td>TGGGCTGAGC</td>
<td>3240</td>
</tr>
</tbody>
</table>
---continued---

eggstggcacc cyggcagggc aaacgggacct aacatcccgct ggagggtggg ctgtctgctcc 3300
agocgtgggg gcggcagggga gaagtcaggg ngacccctgg tgaagctcct gaggctggaa gaagtggact 3360
cctcccgtgc ggccgcccag gtcgcccgtt tccaaagctca gtcagccaa cctaaaggttgc 3420
gggccacact gcggcagcccc cctccacca ccctccactc cagggacccaa gatgaaactgcc 3480
acocgagcttc acocagctcag attgtgctct caacagcacc cacccacccgc gacotgggacg 3540
cocccagaa ccccaatgcgt aagggcgcttg ggtccagggaa gatccatcttc aactgtgccgctggc 3600
cctttctgg gcagccaatgc ggtcagcggg taaagtcactg gactcaggggt gacctcgagct 3660
cogaagccca cctcgccgac agcaggtgcc cctcaagtga gtcctacaccc cttgacccttg 3720
atggtggcata tggagctagaag gttgctggct acocggggccaga gggagccagga gctcatggcct 3780
ccttgccgctct gcggcagcca cacoaggggta gtcgaggcacg gacctggggct ctggtttcctca 3840
atgggtactt cctcaaggtgc acoocagctga gtcgggtgctg gcccctctgag acoaaagctg 3900
agatcagcgc ccccagcgggctc tgcctgg gccctgaacgc tgggctacggc cctatggccg 3960
ccagggagaa agoctccggtt gacaacctca aggacagcgg tgcggctctt gggacactttg 4020
ggggccgctgc gcgtcctgag ttacagggta agggagccaa gggggagggc tgggggctgg 4080
acggaggggc ccacccattcg tcacaggtca aagggagccaa gggggagggc tgggggctgg 4140
tccctgacat cccatgggct gaccgcagca gcgggagggc ctcacacagc cctcctgtgct 4200
acacagatgag agttttctgc ttcctccaggg aggccagggct gcccctttgg gcccctcccacc aacctgctcc 4260
cgtgacacct gctgctcagg ccctgtgacgt ctcctttcctt gggccgaccc aacgttccctc 4320
acggtgatac acccactcagag atggctcactt cctgaggcct cctagcggcct 4380
acgggtgccct gccgcagccg gagaagttgg gcgggagggc tgggggctgg 4440
aacccactca cctccagcagc ctgccaggggg aacaacagtca cacctcatacg aacggtcctcg 4500
agagcttccgg cctgctggtt ggtgctgcag caccgagccg cgcgggtctg ttccttgcc 4560
tggcgcccccg aacgttcgcct gctcttgctgg ggggcagggct ggtgctgtcg ggcctgtgctgc 4620
gtctacagtct gggagccagg cggccgcccc ctgctgcag cggccgcccc ctgctgcag cggccgcccc 4680
acacttgccca ggcctgctgct ctgctggggag actcctcgc cccatcctc tgccttgccct 4740
gcggggcagc ccagcagggg gaaagcaggg gctaggggttg tcaagctctt gctcaagctctt 4800
aactccaggt gcccagccagc agccacagtct gtcctccttgcc aggtctctccg ctcctcttgg 4860
gacacccagag tgcgggctgc tccctgctcc gggacagggct ggttctgtgct ggttctgtgct 4920
tggctgggtgg gggagccccag aagccatcctg ggtctctctc ggtctctctc ggtctctctc 4980
agatgcgcac gggaggggag cccagcaggttc ccttccggtct gggggagggc tgggggctgg 5040
ggcccacactag cggccgcccc ctgctgcag cccagcaggttc ccttccggtct gggggagggc tgggggctgg 5100
acaggggtac tccgttgccg cccccgagggc aacgttcgcct gctcttgcc 5160
ggaccaccccc ccagcaggttc ccttccggtct gggggagggc tgggggctgg 5220
atggtggcgct cccccgagggc aacgttcgcct gctcttgcc 5280
tggccctgggg cgcacccccct gggagccagg ctgctgcag cggccgcccc ctgctgcag cggccgcccc 5340
agttttgctg cccccgagggc aacgttcgcct gctcttgcc 5400
tccctctca acctccagac ccacccctaca gcagcagggc cggccgcccc ctgctgcag cggccgcccc 5460
aactccaggt gggaggggag cccagcaggttc ccttccggtct gggggagggc tgggggctgg 5520
-continued

cgcctgcgcc aagccagggg ctatgtgtct cctgggaggg atgaaggggg caaggyttcggt 5580

cctctgaggg ccaccacata tttagtcac caaagcctgg aagcagaaca aagccagggc 5640
tttttctgac actataacca tttttttttt actgtacaa a 5681

<210> SEQ ID NO: 59
<211> LENGTH: 1366
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1256295.18

<400> SEQUENCE: 59

cctttgtccto caggtgtggt tagccaggtg ctgggagcga gggcaccat tatttgtgctg 60

cggttcaggt gggtgctgggg agacgagacaa agcaagacgcc aggtcagggc agggagtggc 120

ggagagggag ggcgtgcgcag agggcggaga gcccgggcgg gcggagcggc gcggagcgc 180

gcacaggtct gaggctggcc agccggctct gctagtcagt gcagaaca gaaggccgcg 240
cctcagctg ccggcagcgc ggccgcgcgc cccgttcctt cccgtggtcag cctggtgctttt 300
gcataatgac ttgtggggct gattggcagaa gcggggcagat atggataagg 360
tggtggagtac cttccgggaa atgtcgccag tccagcgccag atcactgctg 420

gggtacgga gcgggcaag agtgcctacag cgggtccggcc tctgttcttt tcagtgctgc 480

tggtcgccag ggcagagacat gacccgttag tcgtccatttt cgtccgctttc aacgtcttcc 540

gagctttgigctcaggaac gacatacaac tctgctgct gcaacagcag gggggttcgg 600
cgcgtcgcttt gctctgctgt ggccaggccg gcgggagcgc gcgggagcgc 660

cccggcctct gctcagcagtgt cttgtggtgca atccacctct atctaatagt aaggtctgct 720
ccttaactgca actatttatttt ttgctggcgg gaaagctgctc cctgtgtcag tgggttcgagc 780

tgatgtacct cttcgggacgt tggggtcact ggtgtgaaac taactgcaac aaattgcaact 840

gagttttttt aaattctttt cgaggttactt gcacgttaac tctcaactt gatgaagaga 900

ttcacgagaag tgggtactag aggttactg caatgtcctgc ggagggcgag 960

tatactactg ttcagcaggg aagaggtgta aaatagaagaa ggaagtctgtt ttaagacaga 1020

<210> SEQ ID NO: 60
<211> LENGTH: 1432
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 444096.1

<400> SEQUENCE: 60

gaccccgctcg gggctggagt ggtccacatg ccgagagggag ggggctgtgg 60
agcctggtct ggtcctccac tgctgtaacag gtacagggagt gaaattcaaca agcactctgc 120
tgacagaaac gatttctgcg tgcctcaagaa ggtctgaaaa gcaagatcac agtgaccgat 180
ggactgctcg ggcaaaaggg ggaaacgggca cccagagatt gacctctgcg agcactcta 240
tgacagtgag cttgctcaag tgcagacgcc cggacagccaa ccaaatgctgc ggctgctgct 300
ggaaacacac agcaacctgg acctggacag cactctgccg gacctcaagc ccacgatag 360
gctgattgac ccagagacgc ggctgtgacc cgggctctgg taccagaccccg atagttgag 420
gtctgacgctg cttgctggag aagcagggga caaactgcgg gacaccaaga acgagctggc 480
tggtctcacc cggcagatcc acagacctca ggggaggcct gatgcaagca agaaagcgag 540
tcagctgccg ccaagccgca cttggagcaag cggagcgctg ttgaggactg cactctcagg 600
atgctcgcga gaaacgttgcc gctcttggtg tgcctctgcga cccagcaagc gaggacctga 660
cagccccggt ggcctgcact cagagcctga ggaaacgcctg ccagctgtgct gacgctgtgaa 720
tgctccagtt cccgaatctc cttgctggcc gtcgaggtgc gcgtgcaagcc gacggagcag 780
gtcctgcagcg cttcttcttg agtcgtgcct ccacccacgtg gtcgagctgg ggcctccacc 840
cctggctagtt ggtctcctcc ttgaggtcss gtttgacccc cccaagggggt gattcagcac 900
aaatggtgcc tattcagacg tcagacgagg gcagctctct ggcggtcagc ccaactgtcc 960
gcgacctct acggcctcaag cctgcctcag cgggtagatt ctgtcgcagct ctcggaagcc 1020
cctggccac atgcctcctc ccctcctcagc ccacctcgc cgtgccctcc agtcgtctct 1080
cggagccggc aacccagcca ggcacccgt cccccaggtat tttccaccaaga gcgtctttgc 1140
atgttggaa cgggtcctaat ccacccagct tttcttcttg gccttggcag ttcgggagag 1200
ggaggtatca ttgggcagct ttgcctggag acttggcttg gccgtctttga 1260
ccacccgagg ggcagctgcttt ccctgccagc cattggccgg gcagacgcttg ggcctgtcgg 1320
ccaccacggc ttcgacctcc cgctgcctcc aacccagccaa cctctctccct gcacgctacc 1380
anacctctct ctctgctgctg ttgctcctac(aa) tgttgtgcaac ctgcaaaaaa aa 1432

<210> SSEQ ID NO 61
<211> LENGTH: 4559
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 008942.10

<400> SEQUENCE: 61
agcgggccgc gcggcactct cgctctgggct ggcgtctcccc ggttgctccaa gcggcggcgg 60
ggcgcgaag cgctggaggg ggcggcggag aagcggcgcgc gctagcgccgg cgggagccgg 120
ggagggagca ggagcttgcc agagactggcctgc gcacagcactaa acgtctgcgc acgagctggg 180
cyccagcggc gggcagggcgc gcggctccag cgggaacact ccagctgtcgc acgagctggg 240
cgaagcagtgc ggtgccacat atggctcggcc gcagccacagg gcgggtgctca 300
ggagggagc ccgctgggctct gtcgctcttg gggttgctggc cgggctgtgg gcgtgtccga 360
tggctgggct gctctcagcg gcagctctgc ttcacacact ggc gaggtgcat 420
agctcctacg ctggagcgttcac cggcctctg ccctcagcct gcgtctgacg 480
tgccttcct gcagttccctc gcgcttgcctc gctgtctgctc gcgtctgctc 540
tcagctctgt ccctcacaatt gcggggcgcgc cggcagggcgc gcgggagct gcggcggccc 600
-continued

tgtaggtgct gctgctacgc gcgcggacgt gctacagcgt gagaaggccc aoccgggtcc 660
apagttgcct tcgcgtgcgg cagctgcttg cggcgtgcgt cggcgtgcgt 720
tccagatcc gcagagagtg tgtgtcaacc taattacaaa ttttctcttt tgtttaacc 780
aacattgtg gggctgtctg tgtctgtcgt cttcatagcc cccctcttgg gccttggtct 840
gcagtagtt gcatttgcct acgagagaaa tgcctacccc ttcagagaa cgtcccttg 900
dccagttgc ctctctcgcc atgcgacgct tgtgtactgt gctgacccac ctggtctact 960
tccagactg gctctgcagct ccagattgcg ccagtgaggg gcctttccgg 1020
atcatcttc gcggctcaag tctcttgcg cttcgctgg ccgcttgctg tctcttggct 1080
gctcgcctag tctcatcttc ccagatcctg cctctctctg cttctctcgcc tccgcttgctg 1140
gccagctgc ctctctcttg tctcttctg agacactgcc ctctctctcgcc cttctctctgg 1200
tctgctctag ctctcttcttg aagcctcttc ccagctctcc tctctctctcc cttctctctcgcc 1260
tcagctcttg cagctctctcc ccagctcttc ccagctctcc cttctctctcc cttctctctcgcc 1320
ctctctcttc ctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1380
tctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1440
ctctctcttc ctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1500
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1560
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1620
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1680
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1740
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1800
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1860
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1920
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 1980
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2040
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2100
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2160
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2220
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2280
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2340
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2400
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2460
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2520
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2580
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2640
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2700
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2760
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2820
gctctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc ccctctctcc 2880
coccoacaco aagaaagaccc cccccacagga tccccctgcga gctcggggg ctctgtcctc 2940
taacagcccc ggcttctcct tttcoccacgc cagaggtgaga acaagctggt caagccatcc 3000
tacagctgct gacctggctc gaggagctgt gctctgggca cagccttgag tggctcaca 3060
cagygcagcc cagagcggcag cagcgtctcct cttgacccgct gctgctcagc cccaaagctc 3120
tccgcaagtc ccacgctggc cccagcctgcc ccaagctggt gcaagccaccc ccttcaacctc 3180
cagyccgcc cagagacggccc gaggaacgtc tttgggacct gcccaacgat gtcgaactct 3240
cctgggcaag gatacgacac ctttatatcctt ccgacctgcc tggctcctccg ctagccatgg 3300
acaagcctta cagcgtctgc tctatagacc ccgatccct ctggggacga attcacacag 3360
tccagagagc agccagcggt cttggagcgc aagcccaacc gcaagctggt ggaacgcttg 3420
ttgccgcttc cttgcctcct ctgggcctgt tctcttggtc gcaaaataag ccagaggctg 3480
gttgacccaco tggcgtggct gatgctgtg gtaacgccgg aacagtggca atcgatggca 3540
gcagccggc cgggttacag aagagcggct cctgcggagc cgaacatct gccttacagct 3600
gaccttgcgc ggtccacgcag cggagccgcag ggtgccgcct tctggtcgcc cttagggcttc 3660
tcaggtgctgc ttcttgctgc aacatcctgc tctgggaacgt ggacatgcct caagyataca 3720
ggcggccggc ggtctctcct gggagcgacc ggtgctgttg tgtctggttg gttgggctgga 3780
gatgggggc tgcagcgcgc tgtgtgggac aagttggcttg aagtggatgg cctggggcgc 3840
tggggtgctt cctcatcgcct cttggtgctg tctggtgcct gcagcctggc tggaggcttc 3900
tctgggcaacc gggttgggag gggccacagt tctgatccct tgtgctgtg ggcocacaacct 3960
agccctggga cccccctctca gaaagcggcg ggcttctcccg cgcgtctgcct ctagccatgg 4020
tccacccct cccacctgct agtcgtgaga ttcacacgg acagggccac gcggagggga 4080
tgcttcagc agccagctgt ttttcgccc ttttttccca cccctcactg caacagtccc 4140
ccaggttgtt ccacccgctg gttctccagt ttttccacga caacccacca caaccccaac 4200
gtggactttt tcgtcgcttg ctttttttttg tgtctgctca accttactacaa ttaaataggt 4260
tctgggcaac agccagctgt tttttwcgtc aagttggcttg aagtggatgg cctggggcgc 4320
gcggagccac agccagctgt tccgagccac gcggagccat ccacagtgcac 4380
gtcggccag cagcgtgagc aacccctgtg ggtgttgcct gggacccctct tcatcccttct 4440
ctattcctt tttttcgtgc accttgctaca cttgctgagg tgggtgtggct tcatttggtt 4500
ctttgggttg gtgttttttttttaagacac accactactt cctaaaaaaa aaaaaaaa 4559

<210> SEQ ID NO 62
<211> LENGTH: 1756
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<223> NAME/KEY: Misc_feature
<225> OTHER INFORMATION: Incyte ID No: 008942.9

<400> SEQUENCE: 62
agtcggcacca ccgctcctcgag aagagcaacct agggccacca ctctgttcttcgcttgcccg 60
tgatatttttt tccgtcagag ggacagcccg aggctcgcgg ccaagtggcag cccctcatttt 120
cactgcagat gttcagctgc atcactccag ggcagacggc cagggagccca ccaagacaca 180
gccacccag gcgggtcaccc ccagctgcagc ttcgggctcca ttcgctctccg tcagggccac 240
-continued

cacgggccc taccattcat cacaccaagc acagaggttg ggccacagtg tgcaccgcctc 300
acatcagct gotcaagtct ggtctggagga gtcggccttg ggcacacgca ctgattgtcc 360
acagacgccg gacccacacgc atagccgaag caccagtccc aactacacgc aggctgcatc 420
aagcagcgaac gcctatcagc aagtgcaaca gttgagccct gcaacacaco cacgacctgg 480
cacacaggg cacccacatca caagtccagt ggcccagcgg gaaacactga tgtcctgttg 540
catagagac gacacacagc acgtcattac ttcctgcgcct ctcgcttttg aggccctctgg 600
gettcacccc agttccaccc ccgtctcgaga gggttagatc agggacacct cgtgctctgtt 660
gttacggtg atttttttcg ctatggaggg caccgacgggc ctggccacgtg cggccacatt 720
tocctgctg ctctgctgcc catgctgcaac gcacggcagc cccgacacaaa gaaacaccccgg 780
cagggcctc actgggcctg gggggtcttg ctttctccag ccccgggtcct cccttcccctc 840
cacggcaggg tgtggtgcccag tgtctcaagct gotaacatgt ggttgagggaa 900
gtgggtgctt gccgacggct tgggtgcccag acgggctggg gcaacacggg ccggacacgcg 960
tccactgaa cagggctgtct caagggcgaac gttctgcaag aagtggccgt gggcggctgg 1020
ggcgccgcct gggcagccgg aaccocctcag ctcacagcttg gggccagagac aacgcgtcag 1080
cagtgggggcc aactcagggc agatgtctgt ctcctttgtct tggagacctg agagctttat 1140
tocacctgg cccgctggct cccgctgcaca atggcacagtc ttcagacgct ctgctcatacg 1200
gaacctgcac ctccccgaggg aacaaaaccc ctcgctgacag gcggagccag ccggagacgggg 1260
gcggagccac acacggccag aagttgactgt acttggttgg gcaacacacct tcccctctgg 1320
gcggtctctc ttgctggcgc ataagagcac cagcgggtgtg cccacacgcgc tgggtgccttg 1380
cctgtgtagt ccagggccagc gttgacatcg tggacaacca cggcccccggg ttaagagga 1440
gttgctcctgg gcggacggcc caacctggctc ccacgcccgg gcgggctggc gcaccggc 1500
gtcgctctct cagcggggct gtcgctctgt ggtctggatg gtcgctctgt gttgctcaac 1560
tcgctctgc ggctcggcag accagctgacct ggtgctctgt gttgctcaac 1620
tctgccgct ggcagctcct gttgctctgt ggtcgccttt tcacccctctt cttgggttta 1680
citaatata taattttttt aacatctactt ttggtctgcc ctgaacacaca caatagctgt 1740
yttttaagagctggg 1756

<210> SEQ ID NO: 63
<211> LENGTH: 3304
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 3267, 3276, 3289-3290, 3297, 3299
<223> OTHER INFORMATION: t, c, g, or other
<400> SEQUENCE: 63

cggaactctg gaggtgagggt tgtgcagctt ggtagggatt ggggtcctct cccccgcccag 60
tcgccctggc gttgcttaaa tgtgctcttt ggtcctgctg ctgctcgtgac acacggtgac 120
agagggcccc ccgggctgcct cctcgccgcc cccccgcctct tgggccgcct ctcttgactg 180
ggggcccctgc tggcaggttc ccgcccggcg gcccgttctgg ccctggattcgc 240
cagcggctgcc ctcagctctc acggaggcc ctgtctcgag ggtgtctgga ctggaagaga 300
-continued

cctccgagat ggaggagtct tccctgtctc tctttgatgc tccctattgca cttgatggtta 360  
attgatcc tagaagagag gctagctgta gataaatagc ggtggagaa aagaatcgyga 420  
atcacatctcg tggctcscctt gataaccaco cccagcttctt cccggtgtgc tgggaaccto 480  
tctaaactg ggctgctcttc cggcgctgtt tgcgcccgcag gttgctgcag cttaaaeggc 540  
agcyctgtgg gcctacacttt tttttctctccttctccaaatttacgagcttcagaattgtgcgtag 600  
gagattttta aagaaggggc gggcataactc ttagatttac agggcgatat cctccttttc 660  
cctccgcttct ctctcggatt ctttgaccaag gccgccttact taggaaggtgc tctctatttg 720  
ggtgaacctcg gaaactcgtgt gagaaggtgg ggggctgccct cggagccgta 780  
tctataactcg tgcgcatcgcc cctggtgtta gccagcactc cttcanaagcc gcgcgagc 840  
cctctgtgg gcctacacttt cctcggcttt cttggctccttg cttcaggata ctctcgagtc 900  
agagaagaagc ttcgggcgggc cgtgcagcagca gatgcgagcc caggtcgtag 960  
tgtgagcgag cctcggtgctt cacagcgctt ccaagcagcc acgagcagc ttcacactt 1020  
tggctgtggtg ttgccgcaacat gtttacaggt gttgttgcttt acgtctctct 1080  
acccgggga cccagctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 1140  
ggtgctggtgg ctatcgctcc caagcctgagc cctatgttga caaagatgaag gtaagcggc 1200  
atgtcccaat ccagctgtgt ttcctcagacg ttgctaataag ttcgagctgaa cccagaactc 1260  
tggtgattaa cttatagggaa ccgctctctt ccctcgctcttg ctttacacatt ggcggacttg 1320  
acccagggcct cggcccgggtc cggcctcttct cccctgctccgt ccagcccccc ttctgcctcc 1380  
gttcctacc gcctcactttt ccagctccatt gggccccctt cgtccctccag 1440  
gttcctac caagctctgtg cttcagcttg ctgtcctgctt ccaagctctgtt atctgccctc 1500  
acccgggga acgagcactac atagccccctt aagatgtgat tagcactac cagagacttc 1560  
acccgggga cccagctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 1620  
gcactctcct caagcccccc cccacgtctcc caagccgaa cagagcttcc cagcagctcag 1680  
ctggaactgg cggccgagaa gattcggcttg gaaactccttg aaagttacct cccagctggtt 1740  
atcgagcaca ccagctctgtt ctcgctgca cggggaggtt cccgctctcttg 1800  
caagaacgtg cccaccaact ttccttcttc gttataattt gttagaattaatc cggcctctgtt 1860  
tggctgtggtg ttgccgcaacat gtttacaggt gttgttgcttt acgtctctct 1920  
acccgggga cccagctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 1980  
acccgggga cccagctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 2040  
cctgagcagc cggcctcttg cctgcttacct ccaagctctgtt acagctctgtt ccaagctctgtt 2100  
aacgcccccc caacgctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 2160  
gttcctacc gcctcactttt ccagctctctt cagagcttcc cagcagctcag 2220  
tttctcact cttctcttc acacgctgct tagaagggtt ccagctctgtt ccaagctctgtt 2280  
cctgagcagc cggcctcttg cctgcttacct ccaagctctgtt acagctctgtt ccaagctctgtt 2340  
tggctgtggtg ttgccgcaacat gtttacaggt gttgttgcttt acgtctctct 2400  
acccgggga cccagctggtt ctaactctgct ctcgctgca cggggaggtt cccgctctcttg 2460  
cctgagcagc cggcctcttg cctgcttacct ccaagctctgtt acagctctgtt ccaagctctgtt 2520  
gttcctacc gcctcactttt ccagctctctt cagagcttcc cagcagctcag 2580
<210> SEQ ID NO: 64
<211> LENGTH: 7231
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1399356.20
<220> FEATURE: unnamed
<221> NAME/KEY: unsure
<222> LOCATION: 5601, 5609, 7107
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCES: 64

cocgagctgtt ccgtgaggtt cagggcggtt gcgcgttccc ggggcaacgc tttttttttc ggggaggtgc ggggaggtgc 64
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 120
ccgaggtgtt cggccagagtt gttcgagttg ggccacgc cccggagcttt tggccgagctt ggggaggtgc 180
tttgttgtttt cggctcgagtt ctttgccggc cctctctctg cggccagagtt ggggaggtgc 240
tttgttgtttt cggccagagtt ggggaggtgc ggggaggtgc 300
tttttttttc ggggaggtgc 360
ccgtgagctgtt ccgtgaggtt cagggcggtt gcgcgttccc ggggcaacgc tttttttttc ggggaggtgc ggggaggtgc 420
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 480
ccgaggtgtt cggccagagtt gttcgagttg ggccacgc cccggagcttt tggccgagctt ggggaggtgc 540
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 600
ccgaggtgtt cggccagagtt gttcgagttg ggccacgc cccggagcttt tggccgagctt ggggaggtgc 660
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 720
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 780
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 840
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 900
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 960
agcgacagc tctccctcag ctcag ctctccag ccggaccc ggccacgtg ccgagagcctg cctctctctct ggggaggtgc 1020
caggggcttg cgccacccatg tgcaccacgt gccaaggacg atcocyceaaag tgcctggaaga 1080
geaacaagag gtcgccacat aqgtgacccg gctccacccata gcttctcaaca aocggagttca 1140
gtcgaagatat aggaaaggat gacgtgtgta gactgtgact gatgtgctact gtcgaagactc 1200
agttacactg gtaaaagaag tgcctgcccc caactacgcccc gtctgcttactc gcaacaacttc 1260
tgacctgaaac gctctgccttc gcctgggccc caacggacatct ggcgacagatg gcctgtctcc 1320
atgctcgagc gcctcctcct gccctcagag ccctcctgcat ggacgccccg agcgggacgcc 1380
cctcctgcttc acacctcaca aacgagatgca ggctctccct gtcgaagacaac ggaacccgaa 1440
cattcagcagt gttcaaggtg gactcgttta gcaacccggtt gccctccagag gccttccatg 1500
gttcactttgt ctgctactcc gttggtaggg tgcctactggtt gcgtctccgtt ctcgcacactc 1560
tccagcccc cagacgtacc ggaacccctg ggagccgagaa gccttctggttc ccacaagcttg 1620
caagaagagc gcctgcccca taattgaggc tggggtctct gccctacccat ggcacactctg 1680
tctcgtcacc tgtggagagg ggttcacgaa aagctgctcgt tctgcctaca aaccccaccgg 1740
ccagtttggag gcacaggtgac gcctgggttag tgcataccacca aacccagacgacacggc 1800
ggcctgccag atcttgccag tgcctcgacca tccctccctgt gcgcttctag aagctcgact 1860
cacccctctg ggcagcggtaaat ggtatgctgc tgcctacccc gtaatgcctac gaaccttgcac 1920
ccagtgccact gatgtcagat gtcacagaaag aagctggtagt gcgtctgccca aacgcacagtg 1980
agagcgcccc cgcgccgtaaa gcggoccccct gccagccctg ctcgcagctg cccacccttct 2040
cacccgtcaca cccgcctcgcc gcgaggggtag gcaagactgcc gcgcacccac ccacacgcttg 2100
cacccgctgc acacagctgcc gcgcagcagct aacgacagatc ccacgacacgca 2160
cctacgtggc caataactgcg aacccatgta cccgtgccag tgtcaacgcttg gcattactgg 2220
caagggcctc atctgtggccc gcacagggcaca cctgtgctgcc tggccacatgg gcagccctgt 2280
gtctgctgcccc aagctgcccc aacccagtaag tgcctggccct tccctgacct 2340
aggggaagca gatccacgcc aagctgctaat ggtctgctgcc tggctgtgagt aagctggagcc 2400
tgatgaattc cccaggttgcc ggcacactctg tccctggtcttc cacaccccagct ctcagctgtg 2460
tcagccacgc gatggatgag gcagccctgctct gcacacgtct gcctacccac ccacacgccca 2520
tcagccgacc cccagcccca atggggagag aagccgctgt gcggcagacat tgtatggaga 2580
cgytctcctc atctgcccag aacactgcccc cctgctctac atgtgaggcc agagcacgcc 2640
tgatatgac gcggtgggag acgtgtgctca caattgcccc tgggacacata atcocyceac 2700
gctgctgctc gcctgctgcc tccccctgaca caaactcagg atatgttagta 2760
agctgcctcc cgcagcctcc tggacacctgt tccctcaggtct gccacccccct gacagctcca 2820
ccacacacacat gcggatagag gcagctgctcc gcgttgccgct gcgtctgcccc ctggggagca 2880
tgacccagag ccacaggtgc ggcgctgcccc cctgctgctgct ccacagctgtg ccagcctgcca 2940
tggtcgaggt gcctgctgcc aagctgatgg tgcctgctgac gctgctgctgct gcacagcactg 3000
cctcttctct gcagagagct gacacgcttc gcggcacatcc gcaagtctgct gcacagctgctg 3060
tctgggccac aagggggacct cccacacgtag ccctacccct gcggaggcag atctagttg 3120
agacagctgc cgcagccctcc tgggtctgtgt gcgccctgtct cgggtgtgtct cgtagcactc 3180
tgcgtgagct cctgctgcc caacacccat ccacacccaa ccacagctgtg ccgccccagcc 3240
atgtgttcca gcggcctgctc ctggagctgctt gttgtctggag gcacgacaccagcgac 3300
-continued

tctcgctgcc tggagagccag gccagcagaga acgcaaggtg gctggtgttg ttttttctgg 3360
agctgcacacgtgacaggtggcag cgtgttggtg ttttttttcttttctctcgttcttctgcttcttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttctttc
<210> SEQ ID NO 65
<211> LENGTH: 961
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3732868C81
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 19
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 65

tgctctctga cgccgcgtnc tctgagcgcga gacgccggaga ggcgtgcctc agaagaaga  60
taatagcggc cccggagcag taaagcggtg ccagccgcgg agtctccaca tcgagagg  120
gcgctgcgg agcgctgcga aacgctggag tccgctggaga cacagggagcg gctgacgg  180
-continued

gccaatgca gctcgcgctg cctctgcgcc gccgcgtgcc actaagctca ctccgccctg 240  
cgagacgcc gacggcaagtt gaaagcttcgtg aagcgcgtaa tccccctgyc cagagagatg 300  
agggcagaga agcagcaagg agaaggatgg agcctcagcg tggccggca gctgtgggcc 360  
ccttgctggc tgtgctgctg cctgatggag aagtggctca gcoccttccac gcggcgcaga 420  
cgctcgccgg aagggagggc acgcggcggc acgcggcggc acgcggcggc 480  
cgctcgccgg aagggagggc acgcggcggc acgcggcggc acgcggcggc 540  
gcccttcgcc aagggagggc acgcggcggc acgcggcggc acgcggcggc 600  
agagctgtga gcacagagac gacgcagaaggg gacagagaggg gacgcgcgcg 660  
gcccttcgcc aagggagggc acgcggcggc acgcggcggc acgcggcggc 720  
tcgaagacaagt gctgacactcg ctaagacag cttcacaacag agaggttttg 780  
ggtctgtgtgtgc aagggagggc acgcggcggc acgcggcggc acgcggcggc 840  
tactcttaagtc gctgacacattgcaagggaggg sasaatctt ttatgttatt 900  
atatattaattactacag cccccatgtt ttttcttccaa ttttcttccaa 960  
t 961

<210> SEQ ID NO 66
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3732868CD1

<400> SEQUENCE: 66

Met Glu Thr Val Gln Glu Leu Ile Pro Leu Ala Lys Glu Met Met
  1  5  10  15

Ala Gln Lys Arg Lys Gly Lys Met Val Lys Leu Tyr Val Leu Gly
 20  25  30

Ser Val Leu Ala Leu Phe Gly Val Val Leu Gly Leu Met Glu Thr
 35  40  45

Val Cys Ser Pro Phe Thr Ala Ala Arg Arg Leu Arg Asp Glu Glu
 50  55  60

Ala Ala Val Ala Val Ala Leu Glu Ala Ala Leu Glu Arg Glu Ala Leu
 65  70  75

Gln Lys Glu Ala Leu Gln Gln Lys Gly Lys Gln Glu Asp Thr Val
 80  85  90

Leu Gly Gly Arg Ala Leu Ser Asn Arg Glu His Ala Ser
 95 100

<210> SEQ ID NO 67
<211> LENGTH: 2608
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1137894.1

<400> SEQUENCE: 67

ttcagcgctc caactgctcg cggacacccc tgcaggggca acgcgtcaca ccagctgctg 60  
actggcatct agaaggtggc tccccgggcc agcggacccc gggtcagcggc atgcgtgactg 120
-continued

cggagtcgcc acctctctag cggcgcgaac tcggacgccc gccggctggc gccgtgtgcc tcggtctctc
180
aactggggtc ccctagcgc gcggcgcgaac tcggacgccc gccggctggc gccgtgtgcc tcggtctctc
240
gctggtggtg tccagggtga aatggaacc cggcgcgaac tcggacgccc gcagctcattg tccaggtcag
300
gttggggcgc ctttgaagag ggaaacttgg agaatcatt gcagcagcgc gccgtgtgcc tcggtctctc
360
gccagagagc ctgggtctgcat ctgcgttttt gcgggggcc gcggggttgg tcggtctctc
420
cggctggcgcctt tcggggctc cccctcctgct gcggtgggctc gcggggttgg tcggtctctc
480
cacctgtgcct aagacagag aacccataa aacgacactgcc ccggacttac ccgggactctt ccgagctcc
540
tgcccctgtt ctgcgtcttc gcggagcggt atcccgctgag cgggctgcgc cggggctgcgc ctgcgtcttc
600
cggcggagcc cggcgcgaac tcggacgccc gcagcagcgc gcggggttgg tcggtctctc
660
tccagacgct cacttcgggt cctccttcgct gcagcagcgc gcggggttgg tcggtctctc
720
tcagacgcct cactctcgtt gcgggtcgcgc gcggggttgg tcggtctctc
780
gggctgctggt ggcgctggtgc ggcgctggtgc ggcgctggtgc ggcgctggtgc ggcgctggtgc ggcgctggtgc
840
cggtctgcgcctc cgggtgcctc cgggtgcctc cgggtgcctc cgggtgcctc cgggtgcctc cgggtgcctc
900
gcttcggtc ctcgggggag ttcggggtc ctcgggggag ttcggggtc ctcgggggag ttcggggtc
960
gggtctcttc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1020
cggcggagcct gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1080
cggcgcgcgc gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg
1140
tgggctggtgcgc gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg
1200
tccagacgtg gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg gcggggttgg
1260
gctggcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1320
cgggctggtgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1380
gggtggtgggt gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1440
tgggctggtgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1500
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1560
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1620
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1680
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1740
gggtgtgtgt gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1800
tccagacgtgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1860
gggtggtgggt gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1920
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
1980
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2040
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2100
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2160
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2220
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2280
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2340
tcgcgcgcgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc gcggagacgc
2400
tgctctctc tgcctctctct ctctctctct ctctctcctc tgaatattta aagcgtagcgt gataattata 2460
taatagatc ataatatatata tagtatataataatataat ttaaaggtc tactttctgt 2520
tttggcagc ttttataataa caactcagct gttgtctctg taggaacact cgattgga 2580
cagaggtcct ggtacotcaca aatacaac 2608

<210> SEQ ID NO 68
<211> LENGTH: 1527
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 1410671C61

<400> SEQUENCE: 68
gttctctgag cgcgtggcg cggcgctgag cggcgctgag gggcgccgag gacgcagcag 60
cgcgtggcg cgcgtggcg cggcgctgag gggcgccgag gacgcagcag gaggcagcg 120
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 180
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 240
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 300
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 360
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 420
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 480
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 540
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 600
ggcctggtggat ctttgtagctg gggcgccgag gacgcagcag gaggcagcg 660

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 720
cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 780
cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 840

tggccccacat cgtgctgcttc tgcctggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 900

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 960

taggggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1020

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1080

tggccccacat cgtgctgcttc tgcctggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1140

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1200

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1260

tggccccacat cgtgctgcttc tgcctggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1320

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1380

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1440

cgtctgggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1500

tggccccacat cgtgctgcttc tgcctggtcg cggccgctgag gggcgccgag gacgcagcag gaggcagcg 1560

<210> SEQ ID NO 69
<211> LENGTH: 353
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1418671CD1

<400> SEQUENCE: 69

Met Glu Asp Gly Val Leu Lys Glu Gly Phe Leu Val Lys Arg Gly
1 5 10 15
His Ile Val His Asn Trp Lys Ala Arg Trp Phe Ile Leu Arg Gln
20 25 30
Asn Thr Leu Val Tyr Tyr Lys Leu Glu Gly Gly Arg Arg Val Thr
35 40 45
Pro Pro Lys Gly Arg Ile Leu Leu Asp Gly Cys Thr Ile Thr Cys
50 55 60
Pro Cys Leu Glu Tyr Glu Asn Arg Pro Leu Leu Ile Lys Leu Lys
65 70 75
Thr Gln Thr Ser Thr Glu Tyr Phe Leu Glu Ala Cys Ser Arg Glu
80 85 90
Glu Arg Asp Ala Trp Ala Phe Glu Ile Thr Gly Ala Ile His Ala
95 100 105
Gly Gln Pro Gly Lys Val Gln Gln Leu His Ser Leu Arg Asn Ser
110 115 120
Phe Lys Leu Pro Pro His Ile Ser Leu His Arg Ile Val Asp Lys
125 130 135
Met His Asp Ser Asn Thr Gly Ile Arg Ser Ser Pro Asn Met Glu
140 145 150
Gln Gly Ser Thr Tyr Lys Lys Thr Phe Leu Gly Ser Ser Leu Val
155 160 165
Asp Trp Leu Ile Ser Asn Ser Phe Thr Ala Ser Arg Leu Glu Ala
170 175 180
Val Thr Leu Ala Ser Met Leu Met Glu Asn Phe Leu Arg Pro
185 190 195
Val Gly Val Arg Ser Met Gly Ala Ile Arg Ser Gly Asp Leu Ala
200 205 210
Glu Gln Phe Leu Asp Arg Ser Thr Ala Leu Tyr Thr Phe Ala Glu
215 220 225
Ser Tyr Lys Lys Lys Ile Ser Pro Lys Glu Glu Ile Ser Leu Ser
230 235 240
Thr Val Glu Leu Ser Gly Thr Val Val Lys Glu Gly Tyr Leu Ala
245 250 255
Lys Gln Gly His Lys Arg Lys Asn Trp Lys Val Arg Arg Phe Val
260 265 270
Leu Arg Lys Asp Pro Ala Phe Leu His Tyr Tyr Asp Pro Ser Lys
275 280 285
Glu Glu Asn Arg Pro Val Gly Gly Phe Ser Leu Arg Gly Ser Leu
290 295 300
Val Ser Ala Leu Glu Asp Asn Gly Val Pro Thr Gly Val Lys Gly
305 310 315
Asn Val Gln Gly Asn Leu Phe Lys Val Ile Thr Lys Asp Asp Thr
320 325 330
His Tyr Tyr Ile Glu Ala Ser Ser Lys Ala Glu Arg Ala Glu Trp
335 340 345
Ile Glu Ala Ile Lys Lys Leu Thr
<210> SEQ ID NO 70
<211> LENGTH: 5648
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<223> OTHER INFORMATION: Incyte ID No: 464689.64

<400> SEQUENCE: 70

```
gtgtggtgtg gctgtcgcc acaatocccg gacgctgcctct ccgggtcgcc cgagcccctg 60
cctcgtcctc ctcctcccttc cgtacccggg acgggcccgg ccagatcaac 120
cctcgtcctc tggggtgcgg gcaaatcctag ttcggggttat aacgagacag 180
gtggggtggt gcaagagttg gttctcttac acgtcgtgg ccgacccggg 240
cgtgggggtc cccgtccccg gttggtgag cgggtgttgg gatggccggc 300
cgggagccgg gttggtgag cgggtgttgg cgggtgttgg gatggccggc 360
tgtggggtggt ctgggtgcgg gctgtcgcc acaatocccg gacgctgcctct ccgggtcgcc 420
tggtgtgggtg gcaagagttg gttctcttac acgtcgtgg ccgacccggg 480
acatgtcttc ggtgtgtggt ctgggtgcgg gctgtcgcc acaatocccg gacgctgcctct 540
tggtgtgggtg gcaagagttg gttctcttac acgtcgtgg ccgacccggg 600
ccgggggtggt ctgggtgcgg gctgtcgcc acaatocccg gacgctgcctct ccgggtcgcc 660
ggttttgtcc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 720
cgtgggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 780
gtgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 840
cgtgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 900
tgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 960
tgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 1020
tgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 1080
tgggggtggt ctcgggggggtc cccggtggtc ccggggtggt ctgggtgcgg gctgtcgcc 1140
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1200
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1260
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1320
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1380
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1440
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1500
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1560
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1620
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1680
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1740
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1800
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1860
ttggtgtggt ccggggtggt ccggggtggt ctgggtgcgg gctgtcgcc acaatocccg 1920
```
```
---continued

gttgattatct ttcacgcccaac gctttttcota gattgaatg cagaacaact tcattgaca
1980
cacaagctcttta aagcagatgta aatcggcag gagaagaatag gataagtctg aatacgaaga
2040
tgaggggaagc gagaacagag gaactcttccgc ttcgtacaggt ctatacaacct gcctaccaaa
2100
tgagaacct ctggagccac acatacaggct ttctcttttc acatgcgtct gcggtgaaa
2160
gtggcagggc ttgagtctcct gatttaccagc atgcagcacc atatagttga ggcaggttag
2220
aaagaagacoctttaaatttaagtttg atcagatccaa cggatattctgct tgggggagaagc gttttocttc
2280
cttttttctt atcacaagga gattttcttag ttcttatatac agaatgttac gaaagttgccg
2340
gttccacgaa caggctccag taagatattgtg agtatccacag gatgtggccta
2400
ttcagctctttc tcctgtgctctg ttcgatggcatt tttggtttcctc coccttcatacag
2460
atatagatttgattcagcaataacac cttttatttttg aagaggtattt aagaggtgta
2520
atcggccacta aasataatata atatatataatatacat gattagatgctc tataagttta
2580
agaatatcctt ttcacgcaag aagagtctg gaaaaacctgt tggccagagag ggaagataag
2640
tttggagacac atctctcttcttt ctactgctgctg acctacagtg gagacggtcag
2700
catagggctttatatcctct cgaacctacg atggaaggg atctgtagag acgtagccag
2760
ggtgcatagcct gcagtaatggc gggatcttaa ctaaatataat ctaaatataata
2820
agtagccttcttt caggtttttt cttgactagta gtaaaaaactga gagaacggtct
2880
cctccacgagggt ggatctgttt ctcaacattt tttttgggat agaaaggggct cttcttaaac
2940
actagcactaatatacaagc tctcctcggag ggcaggtttac gaaatcttcct ctcataccttct
3000
attctgtattctctgtcct gctgtctcctc cctcctttttgt ttcacagatc ttttcctcctc
3060
catggatcttctctgcctct cttgacagcg aggtctctct gcggcttattt agaggtcttct
3120
agatgtttatgttca gcaatgtggcag ctctctcagtc gctccataagtc agtagccttc
3180
tgtgctttct cggctactg atgtagagtc gataatctcag tcccctccag tggcctccag tggctcctc
3240
catacactctc tgcgtctctc cggatgagtc gactaagccag ataacctgct ctcgattttc
3300
cagaaaaataatgttgccag aagatcttctg ttcctctcttc ctatactaa gacaacagoc
3360
agacagcagc tcgctgttctg cctgctgctcc gacacactaca gcaagacggg cttttttttc
3420
gctcctctgctcttattcata gaccagagagc atgtgcatat ggcctgggccc gctttttttc
3480
agggtcctcag cctgagcttcc tggccttcag tggagctggca atcggatgtga aatagaggg
3540
aaggagttctt ccagctgctcta gcccagggaca gaaagcctat gcctgtttttg gggatctcag
3600
aggagttctt ccagctgctcta gcccagggaca gaaagcctat gcctgtttttg gggatctcag
3660
ttcacgctctc ccaggttctgt ctcgagcacc atcagatccaa cccattttt cacagcagc
3720
tgctgactaga gataaagaaaaaaa ctgcttcata gatgagctgcta
3780
gactagcattc tggcctttctca gataagattc ctcgatagttt gtcagatctgt ctttatttac
3840
aactttgaactctgactggca cgtgtcacc gcctctcagctctgctaca gatcctgcagctgagttcag
3900
aaggagttctt ccagctgctcta gcccagggaca gaaagcctat gcctgtttttg gggatctcag
3960
actactactgc aacagtagctg gtcagctcctg cggagatct gggatctcag
4020
tcagttggttc agcttttttc gattagcctttct gattagctgtt gtagtattgtaaaaaagtttttatttttttttttttctatct
4080
ttcagctctactgcagccag ctgcttttctgctcttattcata gaccagagagc atgtgcatat ggcctgggccc gctttttttc
4140
tcagttggttc agcttttttc gattagcctttct gattagctgtt gtagtattgtaaaaaagtttttatttttttttttttctatct
4200
```
ggttggatgg caccctcaggc tggggggccc aagtatgtg tggtcttgaggt tgtggtggg  4260
agtggttcctg ctgagtacaa gacactatt tctacagatt taaactctt ctaaatgaca  4320
catgtaagtt aacactagat ctgtaacaagag tggcaggtt tcaacatcct cttggttggg  4380
ggtcggtctg aaccctcaggc ttcacatcct ttcacataa ctggttcgatc ttc  4440
tgagctgctg ccacactcctt ctaagatgt acttggccca ctgtggggac  4500
tggctgagaa aagggcctgag gttgaggatt atctgaatac cyattggca gattccttcc  4560
tggttctctccttt gtttaaaggt tgcttttttt tactggtggc cattttggtg  4620
aggttgaa aatgttattg gtattttgaatt ttaaatggttt tttttttaaa tgggttaaaa  4680
aagttggttg aagttttcata atctgtaaat ttcctctttt tttatgtaag cctggttggg  4740
attctctgta gtagtagctc aatataggta ctaaaggaag cggccttggg  4800
tgtgtatggt gtggtaagg ccagggcttc tcaagacctt gtgcacatgc tttgtgggat  4860
gacccgctctc aacactcactt aacgtaaaag cgcctgctgc aggctcttgg  4920
gataataata cctagctcct tcaagggccg cgggtggacg atcagtaatg atttttgga  4980
agctttttgg cacacccgtcc cacagagatt ttcagaccct cctttttttt ttaaatagga  5040
atgcggtggtg tgcacatgc tttgtgggat tctacgtttg gttcataagg  5100
tctccagaaa gctatggatg ttcattggag aacggtggga taacctcaggc gacaaaattt  5160
gagcataacgaa caccctcaggc tggcaggtt tcaacatcct cttggttggg  5220
gggtggttta tgtttttttt ttatggttagg atcctgtgata ctggttggg  5280
tggtggttta tctagttctt tcaacctcgg aacagttttt acaatcctg acgttgttg  5340
aggtggtacag aagtggagtt tgcacggtcc agacagctat atgtcactac ccatatatcc  5400
agacagcgcg ggtgggtagg cgtcagcggg ctcagggccc taggtactgt cagtggtgctc  5460
atgdgtcattg ggtggagttt tgcacggtcc agacagctat atgtcactac ccatatatcc  5520
tttttttttt ttacctggaag aaggtatgta gacacgtatg tccctctgag tgggtgttta  5580
caccaagagtt ccaagtctcg ctcgggctgc ccacttattgc agatcgtgct gtgggggag  5640
aattgtcc  5648

<210> SEQ ID NO: 71
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 053959.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 2, 13, 20, 32, 41, 47
<222> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 71

gcggaccgct gaccccttcg cccacgcttc taccctggtc cagacgctt ctgag

<400> SEQUENCE: 72
cttgggtgat ggggccagga agaggttac aggccacggg accagcttt ctaacccttt  
  60
ggacaccaag tgaagaagt tcatactcgg agtcgccaggg ttacctgtgc tccagaacaa  
  120
catagggacc tccaacggga cgagtgggga cttctttgtg agygcceagag acctctgtac  
  180
cattgaagtg tggctggaaat ctttgcctt cggccttcag tggggttgga ccagtgggaac  
  240
gaggatgaag agcacccctc tggaccaaat gttaactcca aggatgtccag aggcaaaaa  
  300
cagttttctc ccaaaaaagga gactccagtt ctggtcgaag ttcggtccga tccttcagag  
  360
sattttgtgt ttggacgtat ctcctttgggc aggcagacta cttggttgtt cggagatgtac  
  420
aaacgctgca cagaagcaag aagctttaac aggtcgaag aaacccctga taagggccaa  
  480
tggacagagc tccaacaaaag tggagatggag gactccagca ttgtaattcg tggacaccct  
  540
ttcggtggtt gggagggag gcacatccag gcttccagccc  
  600

<210> SEQ ID NO: 73
<211> LENGTH: 2572
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: 
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: InCyte ID No: 021667CB1

<400> SEQUENCE: 73
gttgcgtaaca cactcaacga cagaacgtgt tcgctaatct tgcagcattc agcaccagca gttgaacgc  
  60
tgcccctct ggcgcggcgg tgcagcgcgt cccggttcct agggtggag cactaccgtaca  
  120
cacagcggca gctccagaggg tgcagcactt ttttccacaa agaagacctc ctaacacacga  
  180
tccacattt ttgcttctcct ccccagccgg gcaacggctgt gatgattgca atacagccgg  
  240
tcattgatag cccagctgcc aaaaagagaag cgcagtctta tgggcaatt tgggttttct  
  300
gcagcggctc atgacggtgc ctacacagaa cccctctcct cgcacagcgc ctgctgggata  
  360
aaccactgct tggctccagta ccaacaccgc ggcttgccct atccatgctg agagyctaagag  
  420
ggayacgacag ccttccactc cccgcttggtt ctgcaagcgc cgcaggcaac atacagcgg  
  480
gagagctgto cggatgattt ccccacccct acttccatcc gatcagagct ctggagacacg  
  540
ggatagcacc atacagagct ccggacgagc ttcagaccag cccggtttctg atccgtggctg  
  600
tgcagcattc aaaaaaaagct tgcagcactt ctaacagagc ttcctccctg ctggagcgagc  
  660
ttcacctgtg ttcgcagcag gtttcagcttt ttttcccag cttcagcagc cgtgtggtttg  
  720
tccacagctg ccaggcgtct ctcgctggct ctcagctgct cgcagcgtct tgggctgctgct  
  780
atctgcagct atgcagcggct atctgcagct ccggcgtct cgcagcgtct cgcagcgtct  
  840
gggcgcgcgc cggcgtctct cgcagcgtct ttttcccag cttcagcagc cgtgtgggagc  
  900
gagcgtttact gcggctgata cgcagcgtct cggcgtctct cgcagcgtct cgcagcgtct  
  960
ttccagcgcg gacacggctt cggagagtc ccggcgtctct cgcagcgtct cgcagcgtct  
 1020
tgattgcttta ccaggagcct cgcagcgtct cggagagtc ccggcgtctct cgcagcgtct  
 1080
tgcagcattc ccccagagag cggagagtc ccggcgtctct cgcagcgtct cgcagcgtct  
 1140
tgcagcattc ccccagagag cggagagtc ccggcgtctct cgcagcgtct cgcagcgtct  
 1200
tgcagcattc ccccagagag cggagagtc ccggcgtctct cgcagcgtct cgcagcgtct  
 1260
<table>
<thead>
<tr>
<th>Position</th>
<th>Amino Acid 1</th>
<th>Amino Acid 2</th>
<th>Amino Acid 3</th>
<th>Amino Acid 4</th>
<th>Amino Acid 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Met Pro Ser</td>
<td>Ser Thr Asn</td>
<td>Gln Thr Ala</td>
<td>Ala Met Asp</td>
<td>Thr Leu Asn</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Val Ser Met</td>
<td>Ser Ala Ala</td>
<td>Ala Gly Leu</td>
<td>Asn Thr His</td>
<td>Ser Thr Ser</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Ala Val Pro</td>
<td>Gly Thr Ala</td>
<td>Val Lys Gln</td>
<td>Phe Gln Gly</td>
<td>Met Pro Pro</td>
</tr>
<tr>
<td>26</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Cys Thr Tyr</td>
<td>Thr Met Pro</td>
<td>Ser Gly Leu</td>
<td>Pro Gly Ala</td>
<td>Thr Ala Thr</td>
</tr>
<tr>
<td>32</td>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Tyr Phe Pro</td>
<td>Pro Ser Pro</td>
<td>Ser Pro Ser</td>
<td>Glu Pro Gly</td>
<td>Ser Pro Asp</td>
</tr>
<tr>
<td>38</td>
<td>65</td>
<td>70</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Arg Gln Ala</td>
<td>Glu Met Leu</td>
<td>Glu Asn Leu</td>
<td>Thr Pro Pro</td>
<td>Ser Pro Tyr</td>
</tr>
<tr>
<td>44</td>
<td>80</td>
<td>85</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Ala Ala Thr</td>
<td>Ile Ala Ser</td>
<td>Lys Leu Ala</td>
<td>Ile His Asn</td>
<td>Pro Asn Leu</td>
</tr>
<tr>
<td>50</td>
<td>95</td>
<td>100</td>
<td>105</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
--continued

<table>
<thead>
<tr>
<th>Pro</th>
<th>Thr</th>
<th>Thr</th>
<th>Leu</th>
<th>Pro</th>
<th>Val</th>
<th>Asn</th>
<th>Ser</th>
<th>Gln</th>
<th>Asn</th>
<th>Ile</th>
<th>Gln</th>
<th>Pro</th>
<th>Val</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Asn</td>
<td>Arg</td>
<td>Arg</td>
<td>Ser</td>
<td>Asn</td>
<td>Pro</td>
<td>Asp</td>
<td>Leu</td>
<td>Glu</td>
<td>Lys</td>
<td>Arg</td>
<td>Arg</td>
<td>Ile</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Cys</td>
<td>Asp</td>
<td>Tyr</td>
<td>Pro</td>
<td>Gly</td>
<td>Cys</td>
<td>Thr</td>
<td>Lys</td>
<td>Val</td>
<td>Tyr</td>
<td>Thr</td>
<td>Lys</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Leu</td>
<td>Lys</td>
<td>Ala</td>
<td>His</td>
<td>Leu</td>
<td>Arg</td>
<td>Thr</td>
<td>His</td>
<td>Thr</td>
<td>Gly</td>
<td>Glu</td>
<td>Lys</td>
<td>Pro</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Cys</td>
<td>Thr</td>
<td>Trp</td>
<td>Glu</td>
<td>Gly</td>
<td>Cys</td>
<td>Asp</td>
<td>Trp</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Arg</td>
<td>Ser</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Leu</td>
<td>Thr</td>
<td>Arg</td>
<td>His</td>
<td>Tyr</td>
<td>Arg</td>
<td>Lys</td>
<td>His</td>
<td>Thr</td>
<td>Gly</td>
<td>Lys</td>
<td>Asa</td>
<td>Lys</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Cys</td>
<td>Gly</td>
<td>Val</td>
<td>Cys</td>
<td>Asn</td>
<td>Ser</td>
<td>Phe</td>
<td>Ser</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>Ser</td>
<td>Asp</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>His</td>
<td>Met</td>
<td>Lys</td>
<td>Arg</td>
<td>His</td>
<td>Gln</td>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 75
<211> LENGTH: 5325
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 224855.4
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 1500-1699
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 75

acagcagact ctcagtttcca ccgctgatac ttggagacca cccggtgccct cagtocgacg 60
tccagggcga gttggcagcga ggagggccgg ccggagcggcc ggagccagcgc 120
cggggagtga acgtgagggg cggccggcga ggcggagagt ccggagtgac gcgggtgcgg 180
tgctggaaac gaggagccca ccaagtgcgtg ctaggcagcgg caagtgccac 240
gggctggctgt ggagagacac ggagagacgc actggagcgc taocctggagccc gacggagcgc 300
asagctgcac acacagcgcg caggagcgc ggctcttac caaggcggttt gccttgacaag 360
gagaggggtcc gagaagactc ggacagagc gcctgcttcag gcctggacccc gccacggcag 420
tgggctgtagg acacagcagc tggccttggag gacgttgttg ggacggaccc gctggagcac 480
tgcagccaga ctggaggtta cagggggttg octgagggg gacaaaagag gagaagcaag 540
tctggggcttc gcgcagccag cagatcgtgc cgtgctggc acagagcga ctggagcgc 600
cgcctgctcc gcggagccgt gcggagcgc gaacacggcag aacggggtgg cggagagcgc 660
ggtgggactt tcggggactt cgtggtgccgc tggggtgccgc gggagggggt cggggggggg 720
cgcggggcgg ccagctggcgt cggaggtggt cggaggtgccg cggaggtggt cggaggtggt 780
tgggagggggt gtttcggagc agcggtgcg agcggtgcg cggaggtggt cggaggtggt 840
ggtggggtc cgcggagcag gcggagcag gcggagcag gcggagcag gcggagcag 900
cggaggtggt agggggttc cggaggtggt cggaggtggt cggaggtggt cggaggtggt 960
cgcggagcag gcggagcag gcggagcag gcggagcag gcggagcag gcggagcag 1020
cgcggagcag gcggagcag gcggagcag gcggagcag gcggagcag gcggagcag 1080
ggcggggtc cgcggagcag gcggagcag gcggagcag gcggagcag gcggagcag 1140
ctgggtgtg gccaagcagc gcccacaaca ctggagagag accaagacct caagttcttt 1200
gacacacg agatcttgtyc agaagctgca gccccctgtt acccagcggc ctctctgcca 1260
agattcagg tcatctetat tttggtggag gattttcccag accctcagag gcaagaagaag 1320
cctctaggct ctcctctgct cccctgcgga ggggcccaca gatatatttt ttgctcaca 1380
aaccagtgac atttctctag ggggtacgat ttcacagaa acatacgcag aaccagagat 1440
tacacagat ctatccggca tttctctatt caaagtgaac aaattctcaag atcagtgycn 1500

nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnnn 1560
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnnn 1620
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnnn 1680
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnnn 1740

ggcctctta ctcgttggtag gaaatcagaa gttgagattt cagtcagcgg gttgaggaag 1800
attgggccg gacacagcg gaaatcagat gcacotctgat tgtcagatga gttatatttc 1860
aaagggaggg ttgctcttgcga ttcttttggt ttctgctagc ttgctctattt gattgacact 1920

gtgaatatt accaagatgt gtggttcatc ttatcgaaa aggcattttta atttgctgta 1980

gattggtcgt gatgctgtgc gttgtcgtat tttcagggga agtagagctg aaatctgaag 2040

gatgctgct gatgctgtgc tttgtcgtat ttctctggga agtagagctg aaatctgaag 2100

gccgggtgc gttgcctcgc ttcgggttagg gatttttttt ttctgtctctt tttggggaac 2160

gtggctcgt gttggcctcgt ttctctggtg gatttttttt ttctgtctctt tttggggaac 2220

ttcgtcggct ttctctggctt ttctctggctt ttctctggtg gatttttttt ttctgtctctt 2280

tttggccttt ctttctctctt ttttcttttt ttctgtctctt tttggggaac 2340

tttgcttatttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2400

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2460

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2520

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2580

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2640

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2700

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2760

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2820

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2880

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 2940

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3000

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3060

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3120

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3180

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3240

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3300

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3360

ttttctttttt ttttcttttt ttttcttttt ttttcttttt ttttcttttt 3420
cctgtgcttg gacgtccact gcacagtcac cagtaggggt tgcaccttca ttgacagatg 3480
aacatgttga gtaatgtgga agcagactgt gttctgtgaa ctgcaagatg gtatacattc 3540
tataagacga aagtattttg ggttatatta tcttaaggga agataaagaat gatattaaga 3600
acttgcttct caggggggccc tatacttctga ccctctttttgc tggagamataat ttaacccccc 3660
accagcattc cagaagagct gttctggaag tctgtotcag gagcaocctg tttotottaat 3720
tccoeaagc cyagctoccat tccaatcgtg tttgaccttc tttctctttttt ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
ggctctctcc cacgacctcg agccctctcg cgccgccgcc cccagggtcc gcttggaggc 60
gcggagcc caagcgcata accccaaacgc cggacccccc ggccccgcgg ccgcgcgcgc 120
cctccccgcc cccggccggc agacacgatg atggggtggccc tccgacagaca 180
tctggtgggg tggcacaacc ccctgcagac gctggcggaa gatcgtctgt cgcgctcttct 240
taatatgtg ccctatatac ctgctttgtgct gtcgacagaga ggttgagggag gtgagcaggg 300
cgagtctttcctg tcaacaccac ctgcaagccag gctgaacgaa cgtggtgtcgt gcataactact 360
tcagcctcct ccacactcgg ctatggtcgc tctcgtgcac acogccagcgc 420
tcgtatgcg cgctagctcg ggaatccgga ggctggaag ttcataaggg 480
ggctagaaaa ggtggtatgt ggagacatcg aggtagcataa aacccgagaa gctgcacatg 540
aaggttccttc tgtgtggaccc tcaacaaccgcc gctttccttt cctgggtcatc tgtgacagcg 600
cctcagactg atgtcttctt gcctagtcag gcggtctctoct ctagcagcgg gctgagttgg 660
gccaogcgct gctgtgcccc aacgtgtgytc aatgttggct tctgggggcgg acgggagaga 720
cggtctcttc gcttcctcag attgctgtgt ctggatgctcctct ctagcactgt 780
aatgtgtta tttctcataa atagatgttt ctggtggactc aaaaagccca gtttacgagc 840
tgcgcaggt gttatgatct gaaatagaca gcagcagggag gatggagccag ctggcgttcg 900
ggctgtaag ctcagcagctg acgtcttccc acacccagaa tctgtagctt aatgtccacc 960
atggcaccgcc atgtttccgg tctgctgtaa acacagagtc caagcagctg acggtctggc 1020
tctagcctt cacsagacg tgtactcactt cttgcttccct aatattttctt cactttctgtc 1080
aggttcagcg aggccccaggg cctgttggaggt taatggtttg aagttacgtt gatatttttaa 1140
acagcgaga tggcattttg ttccttcttc tggaggcaac aqaaaagcag caggttccac 1200
agggcagac gcgaaggtggt gggtcctctt ctgggggcttc ttggccacac tttccccccc 1260
ttacaggttt aacctgctgct ttcatttgcg ttggaagtt ttaatcota acagtgccga 1320
aagttacacgg gcggcagctc tctgtacact ttttggagag tgaanaatttt ctgataaggat 1380
aggttatatt ggatgtagaa tgccttgggt atcatatatc gtctcccctg ttcctagggc 1440
tatgatgtg atatataaat ggtattacgt ctgcctttctt gtttatattt caaatatttct 1500
catgggttga tgaatactcg gcacagacgg tggagggcgt tgtgtgtat taatgggtt 1560
catagcactt aacactacag tggctctcg cgggcgtgag gcagcagtaa gttctatgtg 1620
tggcttactg atgcaatgtg cccatagctcg caaatattaga tgaatattttg tgaagaataa 1680
acagctgcgg ctcagcagcc aactacagct gtaatggcg ggggtcctcc cccatctcttcc 1740
cttctctgat cctggtatag cagcctctagc aagcaacgcgg attaattagg gctgctttggtgg 1800
aatttcatgg acctgcctcg aatttaattgg ggacacaaaa ttggccaggg gagggagaga 1860
tttctgttgg taaaaaagcct tttggagaaa ttcggactct aacctgtgtt gaatagaag 1920
agagtgttct atctccaaag ggtttgttct ccccccttt gctttagcttt ttttatttgtg 1980
aaatatataaacat acatagtgtg gaaagagaaa gaggagtaggg attaggaataat tattgacgaa 2040
attatatgatta gataataggt gattataggg caaatagggtaga aatggtttgct caa 2100
gratccacoa taccctgagt aggytaagtt tttaattggt gtcaggaata gctttgttaaa 2160
agatatttatt aatatttaaa aataattacttg tattatgtgct gtaactaaa ataatatttttgt 2220
aataattaanga aataaatctaga ataatattc gccagttaaag aataatttattt aataattttgtg 2278
<210> SEQ ID NO 77
<211> LENGTH: 226
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1518310CDI

<400> SEQUENCE: 77

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

<210> SEQ ID NO 78
<211> LENGTH: 445
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 098533.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 406, 413
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 78

ggctagggcc cacgtgctgg cgtgtgctgg cgtgtgccttt acttcgtgc agctcggccc
  60
ggctgcctct cacacaagacctgaccc cgctcaccct ggtgctaccc 120
cggcccgacc ttcctgtgct ccgaaacgc cccccttgctt aaggtcttctt tcccccaggg 180
tygttcctt tctctgtgct ttttccccoc cttttgttta ctggccaagc tyaatccttt 240
cottaaacca tggcctaaac tcagttcaca taacagtagaa cttgactcaaa 300
asaaaaaaaa acgtgaaact tatacttgct gcctgaacss aasatctttaa atttacgcgaa 360
attcaacac tcagtagcagc atttgagttt attggnatac egnnttttga 420
aaacocttaga ttgtagtaaat aagc 445

<210> SEQ ID NO: 79
<211> LENGTH: 5227
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc.feature
<223> OTHER INFORMATION: Incyte ID No: 410785.1
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 4928, 4934, 4939, 4944, 4973, 4992
<223> OTHER INFORMATION: a, c, g, or other

<400> SEQUENCE: 79

cacaagacgt ggaaattgca cagccccaca tatacattac agcagagat aaggtcaactg 60
ggccctggtc ttctactgtgc atcaactgctg tgcctgggtt cttctcagttt ggtatagca 120
tgggtggtac caaagcagct taatactaaa tataatctta tgtttggtgtg 180
ttcacactga tggcgcgggg aactaagca cccagcttac gatcaactcg 240
cccaatctt ctactcaatgg aacoccaaac cccgctcttg ggtgyaggga gagaactyggg 300
cagctgtgata ccacccacac atccctctgt ggcctctctt acoccttttt agcagtggtg 360
gagggactgc atcgtctcttt tgtggggtgc attggggagc atgggggggc 420
tgttgtgtgac ccccttgtggt tgtgggtctgt ggtgtgcctt ggttgtgtgtt ttaaagtttg 480
gaacccctgta tatactatat ttgctggaaa gcgctgtac agagatataat tgtggggctgaa 540
ttccagccgt ggtctctgtct tatactgttg aastgtctccc aocgctctc agggggacac 600
tggcgactct tccaagctgag gcggctgccct ggggccctct cttattggtgttc 660
tgggattatat ctggggcaact tagcatgctg gggactctgt cgggtgtgttc gtcgggtgtg 720

gagccgccat tcatctctctt ctctctcaga aagccccaca taccctttaca 780
taagtgatct ccgagacaaa ccagctggtc gagaactcaga ggtatgtag 840
tagcccacaa agatataat gaatagctgg aacccagagc agagctcagc agtggggcaca 900
aagctctat atgctggagct tttccaaagt cggctcctgc acagcttatt ctatgcgaaa 960
tgattgtagca tgggtggctct cattcttccc gcacaaatgg cattttttact cccaaactatga 1020
gaacacgtggtc cccctctcctt ccctggtggg aacccatcagc cggctggcctg 1080
taaacactgg taccactgtc tgtctctgtg tgtgctggga gggcgaggc cgaagctccc 1140
ttttcctaat tggcagtgctt gggagtctttt ttgctctat gtcacgctgc cggctggcttg 1200
tgctgtagga aacctctctc tcagctgctc atctctgcc aagctctgcttgatactcctcctc 1260
ttgtcagctc ttgcaatgtt gggcgaggtcc gttccccggt ttcagctttgg cttgttttttt 1320
cggccagga cccgctcctgc tgggtttagct atttgcatac tcctcagatcc ggcgctcga 1380
attcagtagc atctcactgt ctggcgaatgcc ttcagcttttt 1440
ccactaccca aagatgaaaa tggaaagaata atattcgcaca cagtggtacct ctctggccaca 480
tyggggctca tggagaagtq taagatgoca ggttgcgccga ggtgatcaggg ggtytcaac 540
ttcacgtca ggccagcga gatgatcctc aacaaagcgcag aatcgaagta caagcctgctc 600
tgcacgcagc tggaaagtcc tccctacgcgc aacgagccgca aacgagcaggg ttcctgcaag 660
ttcacgcagc gttttrgrgt gcgcgagcag ggttcggggaaa ccaacgccaca taaaactatgg 720
gtggacccca acccaccgcgt tcttgggag gacccagtcct tttgctcctt acctagaaaa 780
cacacacga cccacgcccct gattgcgcctg cystacccgc tgcagccttg gcgtggtgtc 840
cctggccagc ggtgatcagc ggccagcgcg agagaggcga tcctaggtttgg tgaattcgcc 900
tgcacacgc agggatatga aagttctcag tgcctcacaag aqaaattlato gatatgtggtt 960
catgtgatttc cttatagggcg atctaggttaca ttaattttaca gtataattcgcc gcatagacgg 1020
gtcctgagc cctcagccgc caggcgccttc gttggactgt gcgtgacgaca ggtatcctct 1080
tagcgtggtg gtcggacgcc gcctctctgt ttaaactcct cccctcctcgc tgcacgctct 1140
gatcagcgcct atatctcag cgtcgcagcg gcaacdacac gcaaaatacttctgctggtgagt 1199

<210> SEQ ID NO: 81
<211> LENGTH: 807
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> OTHER INFORMATION: Incyte ID No: 333453.6
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 32, 35, 166
<223> OTHER INFORMATION: a, t, c, g, or other
<400> SEQUENCE: 81

gaaacagact cgyccaaaaa aaaaaaatt gnttnggcac ctgtttctgg aatogycttg 60
aatcttccaa gcccctataa ttaattcctgt ctgcttctgg gccaacttac ttaattcttct 120
cttaaaatgc gggccatstac cttctcttccc ttaattcctgt cttgtiingaggt ttaattgcaag 180
ttccttttcgc gcctctctgt ctgcttcttg gcaacttac ttaattcttct 240
caatttacttc gatttacttg gacacttacg caataaatttc aataaaatttc aataaaggga 300
cacgacgct csggaggaga agtaaaagagc ccaatttactgc ggyccatttaa tccaatttct 360
cttctctctggt ttccttttct gttaattcctgt cttgttctct gcgctaaacta acctggaact 420
atctctctct gcctctctct gttaattcctgt cttgttctct gcgctaaacta acctggaact 480
atgtcgcgtg ggcctctcttct ctggcctctgt ctcgctgttc cgggggtgct gcttgattttata 540
agaactcgag ttcggtcgtt ctggcctctgt ctcgctgttc cgggggtgct gcttgattttata 600
cctttggttc gcaacgcgctg gctacttgcgt ttcgctcttct gcgctaaacta acctggaact 660
gaacagcgct ttcgctcttct gcgctaaacta acctggaact 720
gaccttaaac gttttgtgct gcgctaaacta acctggaact 780
cctaacagct acatctggaac gctaac 807
<216> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 365070.1

<400> SEQUENCE: 83

acctggctcg ccaactcttcgccgtctgtc aacaacatgaggtcctgggtagctgtcag 60
actctgtggct gccgcgcgtgcc gaaacggttgc ggctggcagtc gatgtctgaga 120
catcgtgtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgcttgccagctgct 180
tgtgccgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 240
gcccgccggc gcgttctgtcctgtctgtgcc ggtgtcgctgtgctgc 300
tgtctgtcctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 360
tgtctgtcctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 420
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 480
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 540
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 600
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 660
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 720
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 784

tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 840
tctgtctgctgctgc gctggtctgcc gcgtctgtcctgtctgtgcc ggtgtcgctgtgctgc 900
caccagaagg cctgtatggag tgagagagaa gtcttctgtgy tgtgcaacct taataaaacc 960
ttttaaatg tgcagcagca cagatattga ccagctctct tttttgtaag ggtctcaag 1020
agcgtgca acctgtcagc agaatagagag agggtagaaac gggaactcta tgcctagtct 1080
gagatagagag tggctcttgg cagcctcctca actaaccacc cactattcag tagatagctct 1140
ttttcatct gttatatatta cagactctatgtgctacagaatgtacagtgtaaatgac 1200
aabattcatgt gttaggtgca tttgtttttc attgttttaag gnaaaaaaaa assgctgtta 1260
cctatcactgt gattgtcctg tattgtcaga gtatcctgtt gttgctcttg ccatacgacag 1320
tcttg 1325

<210> SEQ ID NO: 84
<211> LENGTH: 3663
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 413921.2
<400> SEQUENCE: 84

gtgcgctag ctccggacctcc gcggcgcag actggaacctccccagctgctc ccaagggcc 60
ccaagccgc cggattccgagc agaatctgg gttatatcttg ggcagaggga gtaataaaaa 120
gaccagagc atatttttact cagtttgagag ctatatttct attttatttt tcaaggcaga 180
aacagaca gcctcgctact ctaaaaaaga gacttttttt ttaattacta cactaagctg 240
aacatattcat gttggttgtag cagaggaag aatacagagag ggttttattg atgttttcta 300
aatcagaygtgt tgggatatag agtattgtgc atctctctgtg caccataatg 360
tccattctag gttggtctagt atgctaatc atctagatcag tttgtgactc gtcacaaac 420
cagggctcctgtgagataagtttttaa tattacagc atatatttttt ttaaccggag 480
taatattcatc ctaatactc ggttagatcag agatattctag atgttcttagaa aatatttaaa 540
attagacgcc ggtgtagaa ataaatcact ctcttgagaa ctattttttaag aagactaacc 600
tccacagc aaaaaaga agcagagc ccccccacattcactagag aagacaagaa 660
taatattgaa aatatttttgctgatattgtgt ccttttcccag gcacagcagag cagactctt 720
ctattctag agagagagac agaatctcct tttaagaaag aagatattgact atgggtagag 780
agccagagt aatattttttt gtaattattta tattacagct cggggaagaa 840
atccacacc ttactactact attttatttt ttgctggaag tttatatatt gcaagcaga 900
gcaatcctc cggcagagag ataaaagag tttttttttg atagttactt ccaagcacc 960
aggtttagct cacttttctc ctttttttaag gatattttttag gatagctctc cgtttttttttt 1020
gcattttcat ataaagaaaa cacacacacct tcacaaagag tattacctcagctgttttcg 1080
ttttttggc tccattctag agatatttaa atatattttaag ccacaactcag cagcactttzza 1140
caccagacct gctacttccc tttttttgatggagata ccaagcacaac cggcattcctat 1200
ccagtattcag aatgagaagta cttttttttttag ggtttttttag aatagagc 1260
tggctttttgg ggtttttttttag cgttctgcaag atggagacac cagcctctctc atcgatccaa 1320
aggtattcag ggagagagac tatttttttta gacatattttaa aagagcacta aatattttttg 1380
gagcgctgac cccaggatgtactcattttgt ttttttttttt tagtttttttag tgccttttttag 1440
attcattttt actgagtctag cgggagagtttg cttatttttttag gatagctttttttttatttttttt 1500
agtccttcca agtagcagct tacgtgccag ctagcttgca gatggtgttg aagggagtgg
1560
gtagctgtgg agaasacactt tacccacagc agatotggtg gccagaaactgtgt tgatagaag
1620
tgaggccgga gttgtaaaaac taotgatatt tggaaatgcc aaggtatgttc tggatgtac
1680
gtaaaacagtt ttcctgctgt ctaagtttac ttgtaagtttg tggccacttg eagtgtttac
1740
attacagcgc ttcacgaccc aacactagtt tcaaggtttgg gttgtttttaa tgggtaagtt
1800
aacaccgcc ggcgaatgct cttttaaaaa ttacacactt ttagaatagg ctaacatagtt
1860
tactagggc cccgacgctct aacgccgaaa gggtgacgct cactgtaytg ataggtgtagt
1920
acgggtgagttg cccagcagaa acccagagg aagcgtctct tttccagtac tgctctgcac
1980
aataagagaa ctaagtttga ttaagggaaa tcttggaaga taatgcagtt tggacagcagt
2040
acgcttcgct acaggtgagt ttgcttttga cagacacctt tttttaaagt gttttaaggac
2100
acctccacttg aagctacactggt ctgctttga ccaaaatcatat gctgttgacc tggaggttgc
2160
aagctttgact gctgactctg ccaccagact gagaatcagct gttgacgttct cgggtgactgt
2220
tgataagtgt gattaaaccact acccttctctg accttctctg ccaacacagt gttgagcagtt
2280
attcagcagag caccagcttg aagcagcttg agccgctttc tttggtcatt cacaataaat tggaggttgc
2340
cagaccttg cagagtggct gacccgaggg cagagctttat ttaagaggtc cagccgaggg
2400
agtgctttgc caccagcactg caccagcttg tgccttttga cagacacctt tttttaaagt cgggtaaaat
2460
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
2520
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
2580
agtgtattat atatatataa ctatatattta tatatatattat agaasacacgg
2640
acgacgttttt tacatcatgg ttaatcagga gttcactctt gtaggttgag
2700
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
2760
agacaccgag caccagcactg caccagcttg tgccttttga cagacacctt tttttaaagt cgggtaaaat
2820
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
2880
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
2940
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3000
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3060
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3120
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3180
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3240
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3300
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3360
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3420
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3480
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3540
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3600
agctgccctaccttcttggt cccagacctt aagcagcttg gttgaggttg caggtgtttgc
3660
aaa
3663
---continued---

<210> SEQ ID NO: 85
<211> LENGTH: 1344
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<221> NAME/KEY: Misc_feature
<223> OTHER INFORMATION: Incyte ID No: 336615.1

<400> SEQUENCE: 85

GGAGGAGAGA CACGTCAGG TGATATAGGA GCCAATCTAA AAGGGCAAAG CCAGGGGAT

(...)

GGAGGAGG AGGGAGG AGGGAGG CCCCCTCCC CCCCTGGAGG TGGACGACC

GGGGCTT CAGGGT GAGGGATAGT TACCTCAGC TACCTGAGG TGGACCGGA

GAAGAAGT TCCCGCTCG CTCGCGGCAG ACCGAGCGC CAGGAAAGCA CACCAGCTT

(...)

GTACGTTGACT GTTGGCAGCA GCGGGCGGAA GTTGGTGCAG CTGGCCGCTC

GAGGAGAGG ATTTACAGG TCCGCGGCAG ACCGAGCGC CAGGAAAGCA CACCAGCTT

CAGGGATGG GCGGCTGCTC GCAGGAAAGC CAGGCTGCTC GCAGGAAAGC

(...)

GAGGGGGT ATTGGAGGG GCCAGGGCG CAGGCGGCA TACCAGAAAT AAGGGGAGG

(...)

GCTTCGCAGA AAAAAAAAA AAGG
tttaacctatt taaatgtcacc aangagaagag tatttttcag maaatataag attttaatac 360
cgcaacaagt gaagaagaatt aataacacca aaaaaacaag aagacatata gaaaaagcag 420
atgtaatacgt gactgactgct tatgaggcaco atggagatgta tccataacaac ccacataatc 480
gagggctgtgg aagagggaga aagatccttt ctatttccctc tgaattgata 540
acttaaagcag tggctacgga taacaggcgc ggcggtttgt ccaatgaaat ggcgaacccct 600
gttgggctgt gtgtggagtgt tataaaaccag acaaaccttt ctacataaat gggaaacagg 660
aaatataagt gacagacgtgt tcatctgcaac tccagggct  tttttgtgtg gaaaaaggtc 720
cctgcccctca aaaaaacgt cattgatttg aagagagggc accttattct 780
aacaagcagc cacaaacttg cattggtttc gcaaaagttt ctctgtgctg 840
aataattcac cctgacagtgg aacgacaacg aatgagcctac cagtattgca 900
gcogacggcag tcggcctgagt gatcactgta gtttcggcagc accttcccaac agccttccc 960
tgaaccgcttg aagacggcagctcaactgct ctctgaggct gtgtgaacag 1020
tggctgtgttt agtgctgtgag tggctccagca aagagggcag ggtgctacacag tctctccac 1080
tacacacag tgcgacaaa ctttcttgctg cccatcactc tctctgcrgc 1140
tgcagcagtctg gacgacacca cacgacgctg ccccacacat tctctacatc 1200
atgactgataa gggttgtggg tcatatagc cccactcttg ctgtgattgaa ccagactcag 1260
gacctttgg gtagagggac aagagttgag agtcgctgta aacatggaat ggaagagctt 1320
atgactgtcc gatgatatatta gtcgcacagcg ggttgatatt gttcttggtg aatggtcacc 1380
cagcagctatg cggcgagttg tcaacacttcc acctcgttgc cgggggcttc ctcgagcc 1440
cacacactgca ggaattcccag cgtcttacag gagggttaaa gcttttcttg ccagatatat 1500
caactacccag tcatctgattg ggtctttcag ttttgattct gccagacttt 1560
tccaacccag ttagctttgag ttagttcttg ccataaactg ccacccagctg 1620
acaactcag cctgactacgt gataatagcct ccagacaactg tgggaggtggc 1680
agggcattt gttgcctgag attatattt tttgatcgctg cggagcgtcttc aagagcagc 1740
aatatttat tccactacta acctttcggc cgggatcttt cttgacttcc gagaacgcct 1800
agcccgggctatcgagagt aacgtaacat gcaaacttict ccacaaagtcc gcggcctagag 1860
tggcgtgcag ccttctgccc tccacactcag ctggggccct gcccttacttg 1920
tgggaagga cagccctcctttcctcag cttgctgagg tgggggttag tggcagcagc 1980
gattttactcgccatctgattt gtcctgttct gcgtgaacct gtcgacacgag cccgctgtggc 2040
cgtctagt gtagctttcct cgttagtgaac aggtgtgtcag cgtgttttaa acatggctaa 2100
ctactctccag cttcatttcttg tccacagttc gagtttagtgc gtttctttcag 2160
 tacaaaaacag tggcctgcgt cttgggttctt gtagctttgc ctggaggatc 2220
acgtcctagt tcaagcttactg aacgtaata tccccagttct ggtgttcttg gcaaatagctg 2280
agggggtgga aagagggcga aacaggggct ttttgagttct ccaagcctgct ggctcttctt 2340
caggtctggg cctgagacctt ggcgccactgc cttgtgttgt tcacactgag aaaaatatag 2400
acccggaccg ctgtaaacat gtagggagtct cctgacatcct cttgggcagc 2460
acctttactgg ggccgagcct aacagttctg aataataaat ggtgttcttg ctctcaatatg 2520
tcagagaca cttaataact gctttttttt ttaaacatac aacgcaaat ctcgagcag 2580
-continued-

Cys Thr Phe Ile Tyr Asn Ser Thr Gln Asn Ala Thr Ala Ser Ile 230 235 240
Met Phe Met Gln Ser Tyr Leu Cys Gly Glu Ile Cys Asn Ala Ser 245 250 255
Thr His Asn Glu Ala Pro Asn Leu Glu Asn Glu Met Cys Ser 260 265 270
Leu Arg Ser Ala Trp Asp Val Ile Thr Asp Ser Ala Asp Phe His 275 280 285
His Ser Phe Pro Met Asn Gly Thr Glu Leu Pro Pro Pro Pro Thr 290 295 300
Phe Ser Leu Val Glu Ala Gly Asp Lys Val Val Cys Leu Val Leu 305 310 315
Asp Val Ser Ser Lys Met Ala Gly Ala Asp Arg Leu Leu Gln Leu 320 325 330
Gln Gln Ala Ala Gly Phe Tyr Leu Met Gln Ile Val Glu Ile His 335 340 345
Thr Phe Val Gly Ile Ala Ser Phe Asp Ser Lys Gly Glu Ile Arg 350 355 360
 Ala Gln Leu His Gln Ile Asn Ser Asn Asp Asp Arg Lys Leu Leu 365 370 375
Val Ser Tyr Leu Pro Thr Thr Val Ser Ala Lys Thr Asp Ile Ser 380 385 390
Ile Cys Ser Gly Leu Lys Lys Gly Phe Glu Val Val Glu Lys Leu 395 400 405
Asn Gly Lys Ala Tyr Gly Ser Val Met Ile Leu Val Thr Ser Gly 410 415 420
Asp Asp Lys Leu Leu Gly Asn Cys Leu Pro Thr Val Leu Ser Ser 425 430 435
Gly Ser Thr Ile His Ser Ile Ala Leu Gly Ser Ser Ala Asa Pro 440 445 450
Asn Leu Glu Glu Leu Ser Arg Leu Thr Gly Gly Leu Lys Phe Phe 455 460 465
Val Pro Asp Ile Ser Asn Ser Asn Ser Met Ile Asp Ala Phe Ser 470 475 480
Arg Ile Ser Ser Gly Thr Gly Asp Ile Phe Gin Gin His Ile Gin 485 490 495
Leu Glu Ser Thr Gly Glu Val Lys Pro His His Gln Leu Lys 500 505 510
Asn Thr Val Thr Val Asp Asn Thr Val Gly Asn Asp Thr Met Phe 515 520 525
Leu Val Thr Trp Gln Ala Ser Gly Pro Pro Glu Ile Ile Leu Phe 530 535 540
Asp Pro Asp Gly Arg Lys Tyr Thr Asn Asn Phe Ile Thr Asn 545 550 555
Leu Thr Phe Arg Thr Ala Ser Leu Thr Ile Pro Gly Thr Ala Lys 560 565 570
Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His Ser Leu 575 580 585
Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn Ser Ala 590 595 600
Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser Leu
605       610                 615
His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly
620       625                 630
Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro
635       640                 645
Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala
650       655                 660
Gly Ala Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe
665       670                 675
Phe Ser Phe Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val
680       685                 690
Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly
695       700                 705
Ser His Ala Met Tyr Val Pro Gly Tyr Thr Ala Asn Gly Asn Ile
710       715                 720
Gln Met Asn Ala Pro Arg Lys Ser Val Gly Arg Asn Glu Glu Glu
725       730                 735
Arg Lys Trp Gly Phe Ser Arg Val Ser Ser Gly Gly Ser Phe Ser
740       745                 750
Val Leu Gly Val Pro Ala Gly Pro His Pro Asp Val Phe Pro Pro
755       760                 765
Cys Lys Ile Ile Asp Leu Glu Ala Val Lys Val Glu Glu Leu
770       775                 780
Thr Leu Ser Trp Thr Ala Pro Gly Glu Asp Phe Asp Gln Gly Gln
785       790                 795
Ala Thr Ser Tyr Glu Ile Arg Met Ser Lys Ser Leu Gln Asn Ile
800       805                 810
Gln Asp Phe Asn Asn Ala Ile Leu Val Asn Thr Ser Lys Arg
815       820                 825
Asn Pro Gln Gln Ala Gly Ile Arg Glu Ile Phe Thr Phe Ser Pro
830       835                 840
Gln Ile Ser Thr Asn Gly Pro Glu His Gln Pro Asn Gly Glu Thr
845       850                 855
His Glu Ser His Arg Ile Tyr Val Ala Ile Arg Ala Met Asp Arg
860       865                 870
Asn Ser Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu
875       880                 885
Phe Ile Pro Pro Ser Asp Pro Val Pro Ala Arg Asp Tyr Leu
890       895                 900
Ile Leu Lys Gly Val Leu Thr Ala Met Gly Leu Ile Gly Ile Ile
905       910                 915
Cys Leu Ile Ile Val Val Thr His His Thr Leu Ser Arg Lys Lys
920       925                 930
Arg Ala Asp Lys Lys Glu Asn Gly Thr Lys Leu Leu
935       940

<210> SEQ ID NO 88
<211> LENGTH: 1121
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 399161.1
<220> FEATURE: 
<221> NAME/KEY: unsure 
<222> LOCATION: 1070 
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 88

```
caggcggcgg agccctgccc ggccgagatg gttttttaaa ttcacacctgt tcataattttt 60
taataagat acaacaaa acagttttg gctacacgga acttacacga taaggyaggc 120
catttccaa tgataaaggg ctattctcac caaagaggca gcacaactta cgttttttgt 180
tatatttga agagtcccaaa tyatatattaa gcacaasat aaactacaac agagaagat 240
acacacacctt gctcaaaagtg agggatttttc accacacctcg tagtaactctg tggaaaggt 300
cattgaaa taatgaggag aatagaagat tggacagca caaacaatcg cctagaggaat 360
catttagat gtggcctcgc atgottaaga atcataatct ttttaaaag aaaaacccaga 420
acacgcctgag agggaagata ctcatctccat gacggttgtc ttoccaagac agggttatttc 480
cattcgttt cggagtgtcgt gcggctgtggt tttccoaac agtggagaca cttgacgtca 540
taatggtttc tagggtgcct gttcctctcc ggttttttat atttttaaaa 600
aaaaaactt attaaagcgg cagaaactta aaaaaatttg acataaaaact ggctattttaa 660
gtaaatggcag acaaatctca gagggtggaa atgatacagc gatgttttgc tggacacagt 720
acacgtaaoc tagaatta acaaaaagat aacagggatt acggctgtgat attgcatcacc 780
tctgaatc ccctgggtat agaagaatg tctaaaagta atcatgattat 840
agaacaata cattatatga acgtctgtaga atccagcatt ttaaatgcaat atttttaaaa 900
gagggagaac tgtggatact cgggcaaaag cttccacctc caaagaattt aaaaagagaat 960
tagggacag gaaattttat tttttaaagaa gcacaaatga acaacattca tttattagaa 1020
aagaaagtgt cattaggaag gatacaaaaa gtagaagct gttattttna aagacttctg 1080
aataaaggtga atcataaaggt aacgcgaagta gttgaanagg g 1121
```
ctgcaggacc acacctggaag gtacgggacg ctgccttacca gaaagaaatg gacctggagt 540  
ccttgtcacc aagggcagag aaggaataat gagcagagcy gaaaagagt 600  
aggaagaagac gactcttttt ttagcttgag atttcaccag attttatata 660  
taatgcatgg taleggagca ttaaagagtgc gaaaggagac gaaaggggag 720  
gg 721

<210> SEQ ID NO 90
<211> LENGTH: 538
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<221> NAME/KEY: Incyte ID No: 697785CD1
<400> SEQUENCE: 90

cccagccta ctgctgagtc ttcgagcagc ttcgagctct gcggcgggac atctttcttg 60
actaactct gcggcgggct ttcgagcagc gaaacgtgga atctttcttg gaggagttggc 120
cttcagtgcg atcggcgggt gcgcctgacg atctttcttg ctgcctgggc ctggcgggag 180
acggcgaacc cctgtgctct cacatcaccgc ttgtctctca gcccgggggc gcggcggaga 240
ctagcttcag caccagagag gcggcgggct aagggggtgc aaggggggc gcggcggaga 300
atgttcotta gcggcgggag gcgcctgacg atctttcttg gcgcctgacg 360
cgcgcgctg ggcgcagagc ttagacttacagtccccca cccgtctcagc ctgagggcct 420
tagacttacag gcggcgggat gcggcgggaag atctttcttg gcgcctgacg 480
gcgcctgacg tggccggcggc ttagacttacagtccccca cccgtctcagc gcgcctgacg 538

<210> SEQ ID NO 91
<211> LENGTH: 135
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE: misc_feature
<221> NAME/KEY: Incyte ID No: 697785CD1
<400> SEQUENCE: 91

Met Ala Cys Gly Leu Val Ala Ser Aan Leu Aen Leu Leu Lys Pro Gly
1   5  10  15
Glu Cys Leu Arg Val Arg Gly Glu Val Ala Pro Asp Ala Lys Ser
20  25  30
Phe Val Leu Aen Leu Gly Asp Ser Aen Aen Leu Cys Leu His
35  40  45
Phe Aan Pro Arg Phe Aan Ala His Gly Asp Aan Thr Ile Val
50  55  60
Cys Aen Ser Lys Aep Gly Gly Ala Trp Gly Thr Glu Gln Arg Glu
65  70  75
Ala Val Phe Pro Phe Glu Pro Gly Ser Val Alga Val Cys Ile
80  85  90
Thr Phe Asp Gln Ala Aen Leu Thr Val Lys Leu Pro Aep Gly Tyr
95 100 105
Glu Phe Lys Phe Pro Aen Arg Leu Aen Leu Glu Ala Ile Aen Tyr
110 115 120
Net Ala Ala Asp Gly Asp Phe Lys Ile Lys Cys Val Ala Phe Asp
<210> SEQ ID NO 92
<211> LENGTH: 866
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 399785.1
<220> FEATURE:
<221> NAME/KEY: unsure
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 92
agaagagccg caccgctgag ggctggggag aacccgcag cttgctgctg gagaagggag
60
gtctgctgct gcacagccag cagcagcag ctgggagac ggcgcggcag actggctcct
120
ccttactag gcacagcag ctgggctgct gcacagccag gtcctttgttttta ctctgcctttta
180
cacgagagc ctctgctgcct ttcctctcttt tgctgctgct gctggcctctct ctctgagcgcctt
240
caccctgct cctctgctgcct ctgggctgct gctggctgcct gggctctctc ctctgagcgcctt
300
ctctgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
360
cgcgcaccgg gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
420
gcgcgcagg gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
480
tgactagcgc gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
540
cacagcagc ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
600
tgctctggct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
660
gagcagcag gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
720
gagcagcag gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
780
tgactagcgc gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
840
tgactagcgc gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
866

<210> SEQ ID NO 93
<211> LENGTH: 1274
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 002455.1

<400> SEQUENCE: 93
acacacgacac gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
60
cccggggtg gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
120
caccggccgg gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
180
ccacgagcag gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
240
gcgcgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
300
gcgcgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
360
gcgcgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
420
gcgcgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
480
gcgcgctgct gcacagcag ctctgctgcct ctgggctgct gcacagccag gtccttttta ctctgcctttta
540
-continued

<223> OTHER INFORMATION: Incyte ID No: 334749.1
<400> SEQUENCE: 95

gagcgcctcc astgacacaa cagatttccc actgtccccc cttctatccaa gtgaacacac 60
gccactgtta tatattgtgag cagggagagc tctgtttagc ttatagcttg aacotaagaag 120
aatattatt gacactttcc tggatctgtgc taattctcttg aataagtacg tgaaccacagt 180
agagagtaag atatataaat accttgcttgc agttccaga cacaaccaac caaatatat 240
tgtattttg cagactgtagc aagagcgaa tttttgtgct ctatgtatga ctgttatttg 300
tactatttt atatatagtc ttggtctgta ttaaatgttg gaaatgtttg tgtgtctttt 360
atatttacct ttatatataa atgtaatcagct tattatagtga tcatataaat accaatccta 420
acttgggact ctaaacataag ggaacatgg actctagggt tgaataaata aagcoagtgtg 480
ttatccttcag gctgagggcg aatacaacaa agctaaaaag cgtgtaggtg aaaaaaaa 540
aacaoagct caacacacc ctatctatag ccaagtttac ttttstgttg ggttcsttst 600
caatcaagag gtgcaggtatg atatataccca acac 634

<210> SEQ ID NO 96
<211> LENGTH: 579
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: Misc_feature
<223> OTHER INFORMATION: Incyte ID No: 041764.1

<400> SEQUENCE: 96

gaaaacacct ataatgaggg aagcccttgc cccaaactgg accatgctaa ccaagcagc 60
gtgataggg tcgccaaccg cccagctgac ggatcacagt accatagccg cacagagacc 120
tgagcactct cacggtacac ctttctgtaaat atgacggagc accttgggaga gggcagctaa 180
acccgaaag ctagttctgc cgcagatcata cgcagttggc cttctgtgaca cttggaggtg 240
taaccttctc acccagaga gatgccccct tcggttaaat accatgaccc 300
gagagcctgc tgcattttat cctggtgtac cttggccctt ccgagctttgctg 360
agcagcttttc ctcggtctag acgtgtcctcc atacacacac cagctggtgc aagagcagctc 420
tgacaccaag cttggtgagaa aaacccgctt aacaccatca aaaaatttttc ccaacatatc 480
atatgaaa cagagctgag tacagcctcag ctgcagttgaa aacagtttgc ttggtcattg 540
atataaaa cccaggtggttc tttgtgctgta aagctgggc 579

<210> SEQ ID NO 97
<211> LENGTH: 5432
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: Misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2700132C81

<400> SEQUENCE: 97

tggctgaccg atgtggtccttg cgggctggtat ctgccccaaagt gaaagtgttg aaaaatttgctt 60
tcggctctcc cccaggtgtc ataatggctgc ttcctctgct gttataactc ccctcactgg 120
ttagaaaaag tggccacag ggaaacgggtc agcagccgctg ttaaatcctgaa accgaggg 180
cactttttc ctcgagtcct gcaccgctgc ggtttttttcc gttttcattc gtgacactcgg 240
gtcgcacacc tccgagacct tgtgagcagaa tgtgtcttctc aatgcaaccag atgcagcagaa 2580
acagcgtct cataaactct ctgcaagccct ttcctctaag aagctgtgtta ttagagaaaa 2640
tggaacagg gcaaaaaagcc caaaagacac ctcaaaaaatg acctctcttg agcaaaaaac 2700
tctgagatct ggcgagacct ctctaaaacat atctatcccttg gatcagaaaa 2760
tcagatagtc agaatataac aagaagctacct tgtgaagact aaagggagag 2820
agaaattgtc agtgctctcc tctaaagagct tgtgagccagca acatactacct aacaagagag 2880
agagagacag atgaagagaaaa ttagaagacct tttagagaca tataaggaaat gtattgatt 2940
aaaaaagaca gtaaaacaatt gaagagatca aaaaatgaggg gagaaagatta 3000
tgcaacagct tgtcagcctga cagacctctga ggcggtctct gcatcagacac tcgtgaaa 3060
cgcgcagct ggccgagacct ccctcaaact atcagatctcc gccagcgccc caaaggtgaa 3120
gaagagcctta atcctcaaaat tgccctgcoca gtcaccacaa caaagaccaaa aaaaccc 3180
accacaccac aaccacagct tgaagccctc ccctcgggaaag tgtggtgtgaga agaagagctg 3240
cctgagctg ggccagcttca cccagacgct aagggagaaaae cccagcagcagc agagagacoc 3300
agagagaggt gcagaggggag tggagaaaaa ggtggagagc gcaaatggccg 3360
agagcagcctt gtactgagag tgaagaaggtg gcagaggggag ctaaaggaag aggccccctg 3420
actgagacag ctggctgcgct ctaagaagctg tctcagacac cccagctgctct ctgagcatt 3480
actgactggct gagaaactag cccacattagct gtcaaatccct ctccaccaggag acaagtgga 3540
cctcagcaca aagcagaaaaat aggccctaaa gttgagagc aggcccctgag 3600
agatattctta gcataagacgaa aactcaacccca atcaagggcc aagcagctgc ttaagcoca 3660
accacgagga ggctgaagag taaacatgag aagcactattgc ggaacctcaagt tggagaaaaa 3720
ggagctgggca ggaacctcagc ctggcagccaa aagagcacta cagactctcta aagggagaga 3780
ccagcagctta ggaacacctgt ctgcttttgaag agagctcttc gtcacagacag 3840
ggatttggtc gttcttggttca aaaaacactaa aaaaactctcg gacctctccag acctgagag 3900
agtgagaccc ccaccaacca gcacaagagc aagacagaga aagctcgaga aagagactgt 3960
agagagagaac ctctattaggt gcaagagact aagcattcct ggcagccagc cctgagcaca 4020
gcctaaaca tacgtgagtgt aaggaagagaa cattaaacatt ttgctggggaa ctctagcgaa 4080
gaacctgacgc tgcagagcaga acctacccgg cagcacagctc gggccacaag ctctatggaa 4140
agagcgcagg gctctggagag aactgcagcggt ctttaaaagaag cttcttccagt cctctggctga 4200
acagcagaaaaa cacgtgctgct ctgcaagccct ttcctctaag aagctgtgtta ttagagaaaa 4260
agactgcag gacaccccca cacaccccaag aagggagccc aagacoccttg tgggagaaaa 4320
ggagctagaag aagctctgctc cagocctgaa gaagtcacaa cagacatcaag gggaaacacc 4380
acacagactaat aagcatacaag ggtggagagaa aaaaagacaacata aagcgattaa gggaaactg 4440
aacaagaaaaa cttgagacccag cagaaagtgct gtaggtagac agagggccac caaaacttaa 4500
ggagaaggcacc caccctcctgag aagcctgtgct ctgctggagaag agaacacccag 4560
atccagccag cagacccaca cttcagcagaa aacactccaaatgtgctgctgaa gatacacaacc 4620
agcagctctg gtcagacaac ccagccctcc caagatacct gtagagagtgct 4680
ggacgtagaa gagaattact tctgaactcaag gaaatggaacaa ceatcagcaag ctaaaggccat 4740
gcaacacca aaccacagag taaagcttga gaaatggaacaa ceatcagcaag ctaaaggccat 4800
---continued---

tagcagaaag ctagacccag cagcaccac tctggcacc cctgactgya cagacccgg acacactgct 4860
taagagaaa aacctgagct gctgggtcct gacacccag tcacagtctg ttaacagtct 4920
agctcaagt gggccagcc gtaacctgc atcacagcccc aacagatcag gcaacgtctt 4980
acacacacag ctagacacac gcacagccg gctagattgc ctcaagagc tgacagtccgg 5040
gagaagggg gctgacgaaa gcacagtacg agtagtgac gctttgctag ttaagagctg 5100
gactacacag acacacagc aagccaggg gagacacgct acctgacag cagagctctc 5160
gcttcacac gtaggtttgg atctggtccg gtcagcagct ggcacgctct cagtttgttcg 5220
agctttctgc ggaagacacg agctctggc ccctgccagt gcacgtcctg gctttcttc 5280
agacacacag ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 5340
agctttcagc ctagacagc gtaagagag aacagactgct ttaagagagag 5400
ttaagagag aacagacagt gctagcttgg ctgaagagac gtaagacag 5460
agctttcagc ctagacagc gtaagagag aacagacagt gctagcttgg ctgaagagac 5520
ttagctttgg ctgactacag gctacagcag gctacagcag gctacagcag gctacagcag 5580
gcttcacac gtaggtttgg atctggtctg gtcagcagct ggcacgctct cagtttgttcg 5640
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 5700
ctgcacacag gacacagacg ccacacacag ctgatcagag cagcactctg 5760
agaaggctc aacacagacg acagacagc atcaagacag tgaagagag aacagactgct 5820
atcaagacag gacacagacg ccacacacag ctgatcagag cagcactctg 5880
caaccttgg ttagcactgc gaggtggtctg gaaatttcag gctgacgctg atcttgctgg 5940
gagcagctcc cagcactctg aacacagacg cacccttgcag gacacagacg ctgacgctg 6000
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6060
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6120
actacactgtc ccacacacag ctgacgctcg gacacagacg ccacacacag ctgacgctcg 6180
ccacacacag ctgacgctcg gacacagacg ccacacacag ctgacgctcg gacacagacg 6240
ccacacacag ctgacgctcg gacacagacg ccacacacag ctgacgctcg gacacagacg 6300
gagcagctcc cagcactctg aacacagacg cacccttgcag gacacagacg ctgacgctg 6360
ctgcacacag gacacagacg ccacacacag ctgacgctcg gacacagacg ccacacacag 6420
taacaacat cagactccag acagactccag ctgacgctcg gacacagacg ccacacacag 6480
gagcagctcc cagcactctg aacacagacg cacccttgcag gacacagacg ccacacacag 6540
ccacacacag ctgacgctcg gacacagacg ccacacacag ctgacgctcg gacacagacg 6600
ctgcacacag gacacagacg ccacacacag ctgacgctcg gacacagacg ccacacacag 6660
gagcagctcc cagcactctg aacacagacg cacccttgcag gacacagacg ccacacacag 6720
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6780
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6840
gcttcacag ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6900
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 6960
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 7020
agctttcagc ctagacagc atcaagacag tgaagagag aacagactgct ttaagagagag 7080
| CTCTCCAGG CTCTCCAGG CACCGGGAC TCAGAAGGCG AGACTGATG AGAAAGAC | 7140 |
| CAATAAGCG TGCAATCTGCA AACCAGGCG ACCATGGCGG ACCCAAGCAAC | 7200 |
| AGGYCGCAGG AAGAACTCA AGGAAGGAGG CTAGGAGGCA GAATTTTCTG CATCAGGGA | 7260 |
| AGCAACCC ACACTGAGCA ACGGGATGAAC CACCCACACT CCAAGCTTTA TGGATGAGAA | 7320 |
| AAATAACGCA ACTATGGTG AAAGCTCAG GCGAAAAGTC GAAGCTCTAC AGAAAATACC | 7380 |
| TGCCACAGG AAGCAAGGGC AGAAGCTCA GGAAAGGGT GAGGGCTCAG AGGGCCTTG | 7440 |
| TGGCTTCAAGG AAGCTTTCCG AGACGACG TCTCCAGTGG GTAGCTAGT CTGGGACAA | 7500 |
| AATCCAGGAGC TATCTTGAGA AATCTTCCACG GCACGGATCA TCTCAAGCTT CAAAGAGTC | 7560 |
| CACAGGAGG CTCAAGGCTG CCGTGAGGAA AGTGGAAGAG AAGAAGAGCC GCTCAGCAGT | 7620 |
| CGACGCTGACCGGAACAGGAGATGGGAAG TGCAGCAGA CACCAAGAGA CACCAAGGAA | 7680 |
| TAGTAAACGATC ACAACAGGTT CTCAGGACG TGACGTGACG CACGGCGAAG | 7740 |
| TGTAATGCTG ACACCGGGG AGCTGGACG CGTATAGGAA AAGGCCTGTG GCTTGCAGAG | 7800 |
| CCGTGTGCC TCCTCCAGACGC ACCAGTGGC ACTACAGAGA CAAGTCTTACT | 7860 |
| TGCCACAGG ACCAAGAATT CTGGCGACT CTCGCGCTTCACACT GCACGGAAGAC | 7920 |
| GACAGAAGG AGATGCCAGA AGACAGTCTG CAGGAAAGGA GTAAAGAAGG AGCTGTCAG | 7980 |
| AGTGGAGGG CTACGCAAGA ATGCGGGA AACCAGCCAG ACACCAAGAAG AACCAGGAAG | 8040 |
| CGTGTAGGG GGGCTCAAGG TTTAGGAC AGTCTGGGGG AAGAAAGACG ACGCCGTAAG | 8100 |
| AAGGAGGGCG ATGGGCCAGG GCGCAAGGGG ACCATAGGGA AAGGCCACCT CTGGCGGAAA | 8160 |
| CCGTGCGGTC GCCCTCCAGGCC TCTGCTAGGC ATCAAGTCGAC ACGCTGCTAC CTGGACTGC | 8220 |
| TGCCAGCG ACTAATAATG CTGGCGATG TCACCCACCA GAGTGGTACG ACACGAAGCG | 8280 |
| AAGCCACAAA GGGCATGGG GGAAGCTGTG GGAAGAGTGA TTTGGTAAAG AAGAGCTTCC | 8340 |
| AGCTGCTAGG TCCACACAAA GATCGGGGAA AACCAGGTTT GCAACAAGG AAGCCGAAGG | 8400 |
| TGGAAGTAAAG GGCAGCAGT GATTTGAGGA ATGTGGCAAGG AGAAGCGCTAG ACGCAGGTC | 8460 |
| AAGTATACCT GCAGGCAAGG GACACAGGAG GAAAGTGCCC AAGGCGATAGA | 8520 |
| AGCACGTGG GTCCTCAGGG ACCAAGGCGG AGCTGCCGAC CGAACGGTAA GCAGCTGTTA | 8580 |
| CAAAGGGCT CAAATCTCC GCAAGATCA ACGAACGAGG GGAAAACCG GTCAACAGCT | 8640 |
| AAGGAACGG TCACCGAGAAC CGGCCCAAGA TGCTGGATG AAGGAAGG GCTGTTAGCT | 8700 |
| TGCCAGCTCC ACCAAGAGGT CAGGGGAAC CGCGCAACG GAAAGAGGG CTGCTAGTGA | 8760 |
| BGCAACGGG ACAGAGCCGC TTAACCGGAC GCAGCAGCGG GCGCTGAGTT CAAAGATGG | 8820 |
| AAATTGGGACG AGGACAGCCGG CAGGACGCT TAAAGGAGG GCCACCCGG TGGAACGTCT | 8880 |
| GGCACTCCCC AGTCCAGCTG CTCAGAAGC CGCAAGAGAT GAGAAGCTTG CAAAGGTTG | 8940 |
| TGCTATAGGC TCCACAGCAG CTCACAGCCA AAGACCTGAG AGTGAAAGAC CTTAGAAAAA | 9000 |
| ATACAGGAGA GTCTCTCCTGG CCGCCAAAGG AGAACCTGTG GGGACAGTGG TAAAGGCCAG | 9060 |
| AGGACGTGC AATCACACAA GCAAGAAGCA AATCTGCTGT CCCCCACTGC CTTCAAGAG | 9120 |
| GGAGGCTGCG AAGAGCTGGA GGGTCGCAGG ACCAAGAGG GTGCTGGATG TGCCACACG | 9180 |
| AGAGAAGATT GGTGGGAGAG TGGCAAGCCG CAAAGAGCGA AGGGTGTGTT CCCCAGCAAG | 9240 |
| AGGCAACACA TGCCAGCACG TGTTGTGATG GAAGAAGAT TGGAAGCTG CTTGCAAAAG | 9300 |
| AAATTGAACTG GGCAAGGGCG CAACAGACA ACCACCAAG AAGGAACCAC | 9360 |
-continued

atttacaaggt tcgctccgt gaaaaatgg gattacacgt ggttccagag gccaattaa
9420
gactgaggca gcacgcaga taacgaggt cttttgtta tcgcagtaaga tcgaaaaa atggaaat
9480
cagaataaga aaagagcctc tgaagatcct cccagagaat gcatctcaga atcaagatga
9540
tgtaggcccc aaccacctac ctgagacaaa agctactgag aacaagaagtt gctctgaggct
9600
tgctgaggat atagaggtct cccagcctaa ggtggaagag gagaagcggag ggcaagaaag
9660
tgcaaggttc tctctgcaag atccagaaggt gcaggggaaag gcgaatttcc acaaactcct
9720
gtgcctgagc tcagagagaga caaaagccca gcttgcagca agcaacctgg agaqaaactc
9780
tgtgagagga gaagcgaggag gtgcacagag gtgctgagag aatccagaaag aagctgagga
9840
cgatgtgtgt gctcatagaa taagagcagc agctcttagg gcctgaagaa atattgcaaa
9900
ggaaagatca aactggagaat atataataag gatagttgctt tcgagaatcttc aagtctagt
9960
ttttgctaaat attacagtg aatcttgtaa gtaaggtgtc cagctctgtctt aaggggaaga
10020
aaaccttggat gctcgggtgt tgaatggtggt cttaaactctc cactgtgagggt cactctggggc
10080
tcctggaactcg aagatgtgct caacccgggggc gtctttgagga ggactgtcggg ccagaagttg
10140
cctcaggtt cgtcagatag acgcttggag ttcgctccaca cccagacgcag cctcaacagca
10200
gcattaactcg gctcaggttc cccactgtgtg cgttctttgt ttcgcttacgt cctcagggg
10260
gactggaagc gacaactatcg cctcgctttgc ccacaacctg aagagcagct cgtaaacaac
10320
gagatcagag atagacgccag ggtgtaacctcg aagttacaga tcctgtacctc caatggccgg
10380
ggagataggt ctcctccatt aaggtctttg gcctctgcga cagagcagtc ta
10432

<210> SEQ ID: NO 98
<211> LENGTH: 3256
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2700132CD1

<400> SEQUENCE: 98

Met Trp Pro Thr Arg Arg Leu Thr Ile Lys Arg Ser Gly Val 1
5
10
15
Asp Gly Pro His Phe Pro Leu Ser Leu Ser Thr Cys Leu Phe Gly 20
25
30
Arg Gly Ile Glu Cys Asp Ile Arg Ile Gin Leu Pro Val Val Ser 35
40
45
Lys Gin His Cys Ile Gin Leu His Glu Gin Glu Ala Ile Leu 50
55
60
His Asn Phe Ser Ser Thr Asn Pro Thr Gin Val Asn Gly Ser Val 65
70
75
Ile Asp Glu Pro Val Arg Leu Lys His Gly Asp Val Ile Thr Ile 80
85
90
Ile Asp Arg Ser Phe Arg Tyr Gin Aaa Gin Ser Leu Gin Ser Gly 95
100
105
Arg Lys Ser Thr Gin Phe Pro Arg Lys Ile Arg Gin Glu Gin Glu 110
115
120
Ala Arg Arg Val Arg Ser Arg Ser Phe Ser Ser Asp Pro Asp Glu 125
130
135
Lys Ala Gin Asp Ser Lys Ala Tyr Ser Lys Ile Thr Gin Gly Lys 140
145
150
Val Ser Gly Asn Pro Gln Val His Ile Lys Asn Val Lys Glu Asp 155 160 165
Ser Thr Ala Asp Ser Lys Asp Ser Val Ala Gln Gly Thr Thr 170 175 180
Asn Val His Ser Ser Glu His Ala Gly Arg Asn Gly Arg Asn Ala 185 190 195
Ala Asp Pro Ile Ser Gly Asp Phe Lys Glu Ile Ser Ser Val Lys 200 205 210
Leu Val Ser Arg Tyr Gly Glu Leu Lys Ser Val Pro Thr Thr Gln 215 220 225
Cys Leu Asp Ser Lys His Asn Glu Ser Pro Phe Trp Lys Leu 230 235 240
Tyr Glu Ser Val Lys Glu Leu Asp Val Lys Ser Gln Lys Glu 245 250 255
Asn Val Leu Gln Tyr Cys Arg Lys Ser Gly Leu Gln Thr Asp Tyr 260 265 270
Ala Thr Glu Lys Gly Ser Ala Asp Gly Leu Gln Gly Glu Thr Gln 275 280 285
Leu Leu Val Ser Arg Lys Ser Arg Pro Lys Ser Gly Ser Gly 290 295 300
His Ala Val Ala Glu Pro Ala Ser Pro Gln Glu Leu Asp Gln 305 310 315
Asn Lys Gly Lys Gly Arg Asp Val Glu Ser Val Gln Thr Pro Ser 320 325 330
Lys Ala Val Gly Ala Ser Phe Pro Leu Tyr Glu Pro Ala Lys Met 335 340 345
Lys Thr Pro Val Gln Tyr Ser Gln Gln Gin Asn Ser Pro Gin Lys 350 355 360
His Lys Asn Lys Asp Leu Tyr Thr Thr Gly Arg Arg Glu Ser Val 365 370 375
Asn Leu Gly Lys Ser Glu Gly Phe Lys Ala Gly Asp Lys Thr Leu 380 385 390
Thr Pro Arg Lys Leu Ser Thr Arg Asn Thr Thr Pro Ala Lys Val 395 400 405
Glu Asp Ala Ala Asp Ser Ala Thr Lys Pro Glu Asn Leu Ser Ser 410 415 420
Lys Thr Arg Gly Ser Ile Pro Thr Asp Val Glu Val Leu Pro Thr 425 430 435
Glu Thr Glu Ile His Asn Glu Pro Phe Leu Thr Leu Trp Leu Thr 440 445 450
Gln Val Glu Arg Lys Ile Gin Lys Asp Ser Leu Ser Lys Pro Glu 455 460 465
Lys Leu Gly Thr Ala Gly Gin Met Cys Ser Gly Leu Pro Gly 470 475 480
Leu Ser Ser Val Asp Ile Asn Phe Gly Asp Ser Ile Asn Glu 485 490 495
Ser Glu Gly Ile Pro Leu Lys Arg Arg Val Ser Phe Gly Gly 500 505 510
His Leu Arg Pro Glu Leu Phe Asp Glu Asn Leu Pro Pro Asn Thr 515 520 525
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Leu Lys Arg Gly Glu Ala Pro Thr Lys Arg Lys Ser Leu Val</td>
<td>530</td>
<td>535</td>
<td>540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met His Thr Pro Pro Val Leu Lys Lys Ile Ile Lys Glu Gln Pro</td>
<td>545</td>
<td>550</td>
<td>555</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln Pro Ser Gly Lys Gln Glu Ser Gly Ser Glu Ile His Val Glu</td>
<td>560</td>
<td>565</td>
<td>570</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Lys Ala Gln Ser Leu Val Ile Ser Pro Pro Ala Pro Ser Pro</td>
<td>575</td>
<td>580</td>
<td>595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg Lys Thr Pro Val Ala Ser Asp Gln Arg Arg Arg Ser Cys Lys</td>
<td>590</td>
<td>595</td>
<td>600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Ala Pro Ala Ser Ser Ser Lys Ser Gin Thr Glu Val Pro Lys</td>
<td>605</td>
<td>610</td>
<td>615</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg Gly Gly Glu Arg Val Ala Thr Cys Leu Gin Lys Arg Val Ser</td>
<td>620</td>
<td>625</td>
<td>630</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile Ser Arg Ser Gin His Asp Ile Leu Gin Met Ile Cys Ser Lys</td>
<td>635</td>
<td>640</td>
<td>645</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg Arg Ser Gly Ala Ser Ala Ass Leu Ile Val Ala Lys Ser</td>
<td>650</td>
<td>655</td>
<td>660</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp Ala Asp Val Val Lys Leu Gly Ala Lys Gin Thr Gin Thr Lys</td>
<td>665</td>
<td>670</td>
<td>675</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Ile Lys His Gly Pro Gin Arg Ser Met Ass Lys Arg Gin Arg</td>
<td>680</td>
<td>685</td>
<td>690</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg Pro Ala Thr Pro Lys Pro Val Gly Glu Val His Ser Gin</td>
<td>695</td>
<td>700</td>
<td>705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe Ser Thr Gly His Ala Ass Ser Pro Cys Thr Ile Ile Ile Gly</td>
<td>710</td>
<td>715</td>
<td>720</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys Ala His Thr Gin Lys Val Gin Val Gin Gin Thr Gin Gin Ass Arg</td>
<td>725</td>
<td>730</td>
<td>735</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Leu Asn Asn Phe Ile Ser Asn Gin Lys Met Gin Phe Lys Gin</td>
<td>740</td>
<td>745</td>
<td>750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp Leu Ser Gly Ile Ala Glu Met Phe Lys Thr Pro Val Lys Glu</td>
<td>755</td>
<td>760</td>
<td>765</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gin Pro Gin Leu Thr Ser Thr Cys His Ile Ala Ile Ser Gin Ass</td>
<td>770</td>
<td>775</td>
<td>780</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu Gin Leu Gin Gin Lys Gin Gin Gin Gin Thr Gin Gin Ass Gin</td>
<td>785</td>
<td>790</td>
<td>795</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu Gin Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>800</td>
<td>805</td>
<td>810</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe Gin Leu Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>815</td>
<td>820</td>
<td>825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala Ser Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>830</td>
<td>835</td>
<td>840</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Asn Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>845</td>
<td>850</td>
<td>855</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>860</td>
<td>865</td>
<td>870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>875</td>
<td>880</td>
<td>885</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gin Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td>890</td>
<td>895</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Gln Arg Arg Glu Gly Glu Met Lys Glu Ile Glu Arg Pro Phe Glu
Thr Tyr Lys Glu Asn Ile Glu Leu Lys Glu Asn Asp Glu Lys Met
Lys Ala Met Lys Arg Ser Arg Thr Trp Gly Gin Lys Asp Ala Pro
Met Ser Asp Leu Thr Asp Leu Lys Ser Leu Pro Asp Thr Glu Leu
Met Lys Asp Thr Ala Arg Gly Gin Asn Leu Leu Gin Thr Gin Asp
His Ala Lys Ala Pro Lys Ser Glu Lys Gin Lys Ile Thr Lys Met
Pro Cys Gin Ser Leu Gin Pro Glu Pro Ile Asn Thr Pro Thr His
Thr Lys Gin Gin Leu Lys Ala Ser Leu Gin Gly Val Gin Val Lys
Glu Glu Leu Leu Ala Val Gly Lys Phe Thr Arg Thr Ser Gly Glu
Thr Thr His Thr His Arg Glu Pro Ala Gly Asp Gly Lys Ser Ile
Arg Thr Phe Lys Glu Ser Pro Lys Gin Ile Leu Asp Pro Ala Ala
Arg Val Thr Gly Met Lys Lys Trp Pro Arg Thr Pro Lys Glu Glu
Ala Gln Ser Leu Glu Asp Leu Ala Asp Lys Gin Leu Phe Glu
Thr Pro Gly Pro Ser Glu Ser Met Thr Asp Glu Lys Thr Thr
Lys Ile Ala Cys Lys Ser Pro Pro Pro Glu Ser Val Asp Thr Pro
Thr Ser Thr Lys Gin Trp Pro Lys Arg Ser Leu Arg Lys Ala Asp
Val Glu Glu Glu Phe Leu Ala Leu Arg Lys Leu Thr Pro Ser Ala
Gly Lys Ala Met Leu Thr Pro Lys Pro Ala Gly Gin Asp Glu Lys
Asp Ile Lys Ala Phe Met Gin Pro Val Gin Lys Leu Asp Leu
Ala Gly Thr Leu Pro Gly Ser Lys Arg Gin Leu Gin Thr Pro Lys
Glu Lys Ala Glu Ala Leu Glu Asp Leu Ala Gly Phe Lys Glu Leu
Phe Gin Thr Pro Gly His Thr Glu Leu Val Ala Ala Gly Lys
Thr Thr Lys Ile Pro Cys Asp Ser Pro Gin Ser Asp Pro Val Asp
Thr Pro Thr Ser Thr Lys Gin Arg Pro Lys Arg Ser Ile Arg Lys
Ala Asp Val Glu Gly Glu Leu Leu Ala Cys Arg Asn Leu Met Pro
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser Ala Gly Lys Ala Met His Thr Pro Lys Pro Ser Val Gly Glu</td>
<td>1295</td>
<td>1305</td>
</tr>
<tr>
<td>Glu Lys Asp Ile Ile Ile Phe Val Gly Thr Pro Val Glu Lys Leu</td>
<td>1310</td>
<td>1320</td>
</tr>
<tr>
<td>Asp Leu Thr Glu Asn Leu Thr Gly Ser Lys Arg Arg Pro Glu Thr</td>
<td>1325</td>
<td>1335</td>
</tr>
<tr>
<td>Pro Lys Glu Glu Ala Gln Ala Leu Glu Asp Leu Thr Gly Phe Lys</td>
<td>1340</td>
<td>1350</td>
</tr>
<tr>
<td>Glu Leu Phe Gln Thr Pro Gly His Thr Glu Glu Ala Val Ala Ala</td>
<td>1355</td>
<td>1365</td>
</tr>
<tr>
<td>Gly Lys Thr Thr Lys Met Pro Cys Glu Ser Ser Pro Pro Glu Ser</td>
<td>1370</td>
<td>1380</td>
</tr>
<tr>
<td>Ala Asp Thr Pro Thr Ser Thr Arg Arg Gin Pro Lys Thr Pro Leu</td>
<td>1385</td>
<td>1395</td>
</tr>
<tr>
<td>Glu Lys Arg Asp Val Gln Lys Glu Leu Ser Ala Leu Lys Lys Leu</td>
<td>1400</td>
<td>1410</td>
</tr>
<tr>
<td>Thr Gln Thr Ser Gly Glu Thr Thr His Thr Asp Lys Val Pro Gly</td>
<td>1415</td>
<td>1425</td>
</tr>
<tr>
<td>Gly Glu Asp Lys Ser Ile Asn Ala Phe Arg Glu Thr Ala Lys Glu</td>
<td>1430</td>
<td>1440</td>
</tr>
<tr>
<td>Lys Leu Asp Pro Ala Ala Ser Val Thr Gly Ser Lys Arg His Pro</td>
<td>1445</td>
<td>1455</td>
</tr>
<tr>
<td>Lys Thr Lys Glu Ala Gln Pro Leu Glu Asp Leu Ala Gly Trp</td>
<td>1460</td>
<td>1470</td>
</tr>
<tr>
<td>Lys Glu Leu Phe Gln Thr Pro Val Cys Thr Asp Lys Pro Thr Thr</td>
<td>1475</td>
<td>1485</td>
</tr>
<tr>
<td>His Glu Lys Thr Thr Lys Ile Ala Cys Arg Ser Gin Pro Asp Pro</td>
<td>1490</td>
<td>1500</td>
</tr>
<tr>
<td>Val Asp Thr Pro Thr Ser Ser Lys Pro Gin Ser Lys Arg Ser Leu</td>
<td>1505</td>
<td>1515</td>
</tr>
<tr>
<td>Arg Lys Val Asp Val Glu Glu Phe Phe Ala Leu Arg Lys Arg</td>
<td>1520</td>
<td>1530</td>
</tr>
<tr>
<td>Thr Pro Ser Ala Gly Lys Ala Met His Thr Pro Lys Pro Ala Val</td>
<td>1535</td>
<td>1545</td>
</tr>
<tr>
<td>Ser Gly Glu Lys Asn Ile Tyr Ala Phe Met Gly Thr Pro Val Glu</td>
<td>1550</td>
<td>1560</td>
</tr>
<tr>
<td>Lys Leu Asp Leu Thr Glu Asn Leu Thr Gly Ser Lys Arg Arg Leu</td>
<td>1565</td>
<td>1575</td>
</tr>
<tr>
<td>Gln Thr Pro Lys Glu Lys Ala Gln Ala Leu Glu Asp Leu Ala Gly</td>
<td>1580</td>
<td>1590</td>
</tr>
<tr>
<td>Phe Lys Glu Leu Phe Gln Thr Arg Gly His Thr Glu Glu Ser Met</td>
<td>1595</td>
<td>1605</td>
</tr>
<tr>
<td>Thr Asn Asp Lys Thr Ala Lys Val Ala Cys Lys Ser Ser Gin Pro</td>
<td>1610</td>
<td>1620</td>
</tr>
<tr>
<td>Asp Leu Asp Lys Asn Pro Ala Ser Ser Lys Arg Arg Leu Lys Thr</td>
<td>1625</td>
<td>1635</td>
</tr>
<tr>
<td>Ser Leu Gly Lys Val Gly Val Lys Glu Leu Leu Ala Val Gly</td>
<td>1640</td>
<td>1650</td>
</tr>
<tr>
<td>Lys Leu Thr Gln Thr Ser Gly Glu Thr Thr His Thr His Thr Glu</td>
<td>1655</td>
<td>1665</td>
</tr>
</tbody>
</table>
Pro Thr Gly Asp Gly Lys Ser Met Lys Ala Phe Met Glu Ser Pro
1670 1675 1680
Lys Gln Ile Leu Asp Ser Ala Ala Ser Leu Thr Gly Ser Lys Arg
1685 1690 1695
Gln Leu Arg Thr Pro Lys Gly Lys Ser Glu Val Pro Glu Asp Leu
1700 1705 1710
Ala Gly Phe Ile Glu Leu Phe Gln Thr Pro Ser His Thr Lys Glu
1715 1720 1725
Ser Met Thr Asn Glu Lys Thr Thr Lys Val Ser Tyr Arg Ala Ser
1730 1735 1740
Gln Pro Asp Leu Val Asp Thr Pro Thr Ser Ser Lys Pro Gln Pro
1745 1750 1755
Lys Arg Ser Leu Arg Lys Ala Asp Thr Glu Glu Glu Phe Leu Ala
1760 1765 1770
Phe Arg Lys Gln Thr Pro Ser Ala Gly Lys Ala Met His Thr Pro
1775 1780 1785
Lys Pro Ala Val Gly Glu Lys Asp Ile Asn Thr Phe Leu Gly
1790 1795 1800
Thr Pro Val Gin Lys Leu Asp Gin Pro Gly Asn Leu Pro Gly Ser
1805 1810 1815
Asn Arg Asp Leu Gin Thr Arg Lys Glu Ala Gin Ala Leu Glu
1820 1825 1830
Glu Leu Thr Gly Phe Arg Glu Leu Phe Gin Thr Pro Cys Thr Asp
1835 1840 1845
Asn Pro Thr Thr Asp Glu Lys Thr Thr Lys Ile Leu Cys Lys
1850 1855 1860
Ser Pro Gin Ser Asp Pro Ala Asp Thr Pro Thr Asn Thr Lys Gin
1865 1870 1875
Arg Pro Lys Arg Ser Leu Lys Ala Asp Val Glu Glu Glu Phe
1880 1885 1890
Leu Ala Phe Arg Lys Leu Thr Pro Ser Ala Gly Lys Ala Met His
1895 1900 1905
Thr Pro Lys Ala Ala Val Gly Glu Lys Asp Ile Asn Thr Phe
1910 1915 1920
Val Gly Thr Pro Val Glu Lys Leu Asp Leu Leu Gly Asn Leu Pro
1925 1930 1935
Gly Ser Lys Arg Arg Pro Gin Thr Pro Lys Glu Lys Ala Lys Ala
1940 1945 1950
Leu Glu Asp Leu Ala Gly Phe Lys Glu Leu Phe Gin Thr Pro Gly
1955 1960 1965
His Thr Glu Glu Ser Met Thr Asp Asp Asp Lys Ile Thr Glu Val Ser
Cys Lys Ser Pro Gin Pro Asp Pro Val Lys Thr Pro Thr Ser Ser
1985 1990 1995
Lys Gin Arg Leu Lys Ile Ser Leu Gly Lys Val Gly Val Lys Glu
2000 2005 2010
Glu Val Leu Pro Val Gly Lys Leu Thr Gin Thr Ser Gly Lys Thr
2015 2020 2025
Thr Gin Thr His Arg Glu Thr Ala Gly Asp Gly Lys Ser Ile Lys
2030 2035 2040
Ala Phe Lys Glu Ser Ala Lys Gin Met Leu Asp Pro Ala Asn Tyr
-continued

Ala Glu Ala Leu Glu Asp Leu Val Gly Phe Lys Glu Leu Phe Gln
2435  2440  2445

Thr Pro Gly His Thr Glu Glu Ser Met Thr Asp Asp Lys Ile Thr
2450  2455  2460

Glu Val Ser Cys Lys Ser Ser Pro Gln Pro Glu Ser Phe Lys Thr Ser
2465  2470  2475

Arg Ser Ser Lys Gln Arg Leu Lys Ile Pro Leu Val Lys Val Asp
2480  2485  2490

Met Lys Glu Glu Pro Leu Ala Val Ser Lys Leu Thr Arg Thr Ser
2495  2500  2505

Gly Glu Thr Thr Gln Thr His Thr Glu Pro Thr Gly Asp Ser Lys
2510  2515  2520

Ser Ile Lys Ala Phe Lys Glu Ser Pro Lys Gln Ile Leu Asp Pro
2525  2530  2535

Ala Ala Ser Val Thr Gly Ser Arg Arg Glu Leu Arg Thr Arg Lys
2540  2545  2550

Glu Lys Ala Arg Ala Leu Glu Asp Leu Val Asp Phe Lys Glu Leu
2555  2560  2565

Phe Ser Ala Pro Gly His Thr Glu Glu Ser Met Thr Ile Asp Lys
2570  2575  2580

Asn Thr Lys Ile Pro Cys Lys Ser Pro Pro Pro Glu Leu Thr Asp
2585  2590  2595

Thr Ala Thr Ser Thr Lys Arg Cys Pro Lys Thr Arg Leu Arg Lys
2600  2605  2610

Glu Val Lys Glu Glu Leu Ser Ala Val Glu Arg Leu Thr Gln Thr
2615  2620  2625

Ser Gly Gln Ser Thr His Thr His Lys Glu Pro Ala Ser Gly Asp
2630  2635  2640

Glu Gly Ile Lys Val Leu Lys Gln Arg Ala Lys Lys Lys Pro Asn
2645  2650  2655

Pro Val Glu Glu Glu Pro Ser Arg Arg Arg Pro Arg Ala Pro Lys
2660  2665  2670

Glu Lys Ala Gln Pro Leu Glu Asp Leu Ala Gly Phe Thr Glu Leu
2675  2680  2685

Ser Glu Thr Ser Gly His Thr Gln Glu Ser Leu Thr Ala Gly Lys
2690  2695  2700

Ala Thr Lys Ile Pro Cys Glu Ser Pro Pro Leu Glu Val Val Asp
2705  2710  2715

Thr Thr Ala Ser Thr Lys Arg His Leu Arg Thr Arg Val Glu Lys
2720  2725  2730

Val Gly Val Lys Glu Glu Pro Ser Ala Val Lys Phe Thr Gln Thr
2735  2740  2745

Ser Gly Glu Thr Thr Asp Ala Asp Lys Glu Pro Ala Gly Glu Asp
2750  2755  2760

Lys Gly Ile Lys Ala Leu Lys Glu Ser Ala Lys Gln Thr Pro Ala
2765  2770  2775

Pro Ala Ala Ser Val Thr Gly Ser Arg Arg Arg Pro Arg Ala Pro
2780  2785  2790

Arg Glu Ser Ala Gln Ala Ile Glu Asp Leu Ala Gly Phe Lys Asp
2795  2800  2805
--continued--

Pro Ala Ala Gly His Thr Glu Glu Ser Met Thr Asp Asp Lys Thr
2810 2815 2820

Thr Lys Ile Pro Cys Lys Ser Ser Pro Glu Leu Glu Asp Thr Ala
2825 2830 2835

Thr Ser Ser Lys Arg Arg Pro Arg Thr Arg Ala Gln Lys Val Glu
2840 2845 2850

Val Lys Glu Glu Leu Leu Ala Val Gly Lys Leu Thr Gln Thr Ser
2855 2860 2865

Gly Glu Thr Thr His Thr Asp Lys Glu Pro Val Gly Glu Gly Lys
2870 2875 2880

Gly Thr Lys Ala Phe Lys Gln Pro Ala Lys Arg Lys Leu Asp Ala
2885 2890 2895

Glu Asp Val Ile Gly Ser Arg Gln Pro Arg Ala Pro Lys Glu
2900 2905 2910

Lys Ala Gln Pro Leu Glu Asp Leu Ala Ser Phe Gln Glu Leu Ser
2915 2920 2925

Gln Thr Pro Gly His Thr Glu Leu Ala Asn Gly Ala Ala Asp
2930 2935 2940

Ser Phe Thr Ser Ala Pro Lys Gln Thr Pro Asp Ser Gly Lys Pro
2945 2950 2955

Leu Lys Ile Ser Arg Arg Val Leu Arg Ala Pro Lys Val Glu Pro
2960 2965 2970

Val Gly Asp Val Val Ser Thr Arg Asp Pro Val Lys Ser Gln Ser
2975 2980 2985

Lys Ser Asn Thr Ser Leu Pro Pro Leu Pro Phe Lys Arg Gly Gly
2990 2995 3000

Gly Lys Asp Gly Ser Val Thr Gly Thr Lys Arg Leu Arg Cys Met
3005 3010 3015

Pro Ala Pro Glu Ile Val Glu Glu Leu Pro Ala Ser Lys Lys
3020 3025 3030

Gln Arg Val Ala Pro Arg Ala Arg Gly Lys Ser Ser Glu Pro Val
3035 3040 3045

Val Ile Met Lys Arg Ser Leu Arg Thr Ser Ala Lys Arg Ile Glu
3050 3055 3060

Pro Ala Glu Glu Leu Asn Ser Asn Asp Met Lys Thr Aen Lys Glu
3065 3070 3075

Glu His Lys Leu Gln Asp Ser Val Pro Glu Aen Lys Gly Ile Ser
3080 3085 3090

Leu Arg Ser Arg Arg Gln Asn Thr Glu Ala Glu Gln Gin Ile
3095 3100 3105

Thr Glu Val Phe Val Leu Ala Glu Arg Ile Glu Ile Aen Arg Aen
3110 3115 3120

Glu Lys Pro Met Lys Thr Ser Pro Glu Met Asp Ile Gin Aen
3125 3130 3135

Pro Asp Asp Gly Ala Arg Lys Pro Ile Pro Arg Asp Lys Val Thr
3140 3145 3150

Glu Aen Lys Arg Cys Leu Arg Ser Ala Arg Gin Aen Glu Ser Ser
3155 3160 3165

Gln Pro Lys Val Ala Glu Glu Ser Gly Gin Gin Lys Ser Ala Lys
3170 3175 3180

Val Leu Met Gin Aen Gin Lys Gly Lys Gly Glu Ala Gly Aen Ser
<210> SEQ ID NO: 99
<211> LENGTH: 826
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 211881.1

<400> SEQUENCE: 99

```
ctttaaacct gacatcttc atgtgctaa acctttataa ttcttttggg tggactaat
60
tttagatttt ttttattaatttt attttgcata gcaatggag tattaaagt
120
agctatattt tattcagat gtaatttttt aaactacaa gaacatatt acattttttt
180
catatatata ctttgaata gatatttacaa gtaacttaag ttttttattt ctacccatg
240
gtttaagttt atgttcaata tattggttaa atagttcttc cttgacactc ggtatattta
300
tttttttattac gtttttttcct tgtaatttt tttctagcgtg tgcacaccttc
360
tgcacatcattgacca ggcacatcct cccacactt ctaataact ttaacatttaa aatattttt
420
ccttcccaact tttgatccaa aggcacactt atccaacagc tatttatttt tatttattttat
480
cctttggcctta ctttttcca ctaagcttgc ctacagtctc cctcacttct
540
cacaaaggtat cagactata aaatctcaag gccttttattt ctttttctagct ctttttatttt
600
gtttatttt gatgtaatc ctttaaccac ctaagcttgc ctacagtctc cctcacttct
660
attttttttg atcattattgct ctttttttattt ctttaaccac ctaagcttgc ctacagtctc cctcacttct
720
cattacgtg gacagagaga tgtaoccaaa ggtggccttt acatctattg cactctgctg
780
gtggcgcct ctgcaaccttc cccacacttt ggacacgccc ctcgtacttct
826
```

<210> SEQ ID NO: 100
<211> LENGTH: 1498
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 409895.2

<400> SEQUENCE: 100

```
agcgtatgg tcagataa ctaacgctcc ggcacggtct tgtgcctcgt tcagcctcct
60
cctcgagagcg tgtgcgaacgc cagacagcgc ggtgacagag cactagtgcgc gccgtagggc
120
agggctgggc tgtggtcgtgc tgtgcctccc ttcggtcccgc ttcggtgggcc gcggcgcgccgc
180
cagctgctg cttttttattt acagcgagac tgtgcgaact tgcagtcctt ccttaaccct ttaaaaagttg
240
tgaatagggc ggcagcttgc cctacacttc ttcttttgcct attgaattttg
300
```
<210> SEQ ID NO: 101
<211> LENGTH: 860
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1422432C81
<220> FEATURE:
<222> NAME/KEY: unsure
<223> OTHER INFORMATION: un

<400> SEQUENCE: 101

agagcagaga tgtgagcatt tctcagaaaa cagccatca acgccactac tgtaaagcac tttcacaatc 60
cagacacac ccccgccaca aactaactaat cctcgaaga acaattccaa 120
ttcgtctaa aacacttccc aggtatttca cctccattca gatactgaa gggagctcaac 180
aatttctta gcccgcagaa gacagatctc tctccgactc gctgagctgt 240
cgtgagcgc tttgtgctca cccacagca gtcaagctc tgtgagcaag 300
tgctgagcgttgtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 360
cgtcgagcgttgtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 420
tgctgagcgttgtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 480
cgtgagcgttgtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 540
cgtgagcgttgtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 600
ccagcagcag ctggtgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 660

cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 720
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 780
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 840
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 900
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 960
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1020
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1080
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1140
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1200
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1260
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1320
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1380
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1440
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1500
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1560

cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1620
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1680
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1740
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1800
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1860
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1920
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 1980
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2040
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2100
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2160
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2220
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2280
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2340
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2400
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2460
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2520
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2580
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2640
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2700
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2760
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2820
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2880
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 2940
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3000
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3060
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3120
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3180
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3240
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3300
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3360
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3420
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3480
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3540
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3600
cggagcaggt cagcctgctgtc tgtgagcgtgtgtgctgtc tgtgagcgtgtgtgctgtc 3660
What is claimed is:


2. The composition of claim 1, wherein each of the cDNAs is differentially regulated between non-metastatic and metastatic prostate cancer and is selected from SEQ ID NOs: 1-3, 5, 6, 8, 10-15, 17-19, 21, 23-28, 30, 32, 34-36, 38, 40, 42-45, 47-50, 52, 53, 55, 56, 58-65, 67, 68, 70-73, 75.

3. The composition of claim 1, wherein each of the cDNAs is differentially regulated between prostate cancer and normal prostate and is selected from SEQ ID NOs: 76, 78-86, 88-90, 92-97, 99-101.

4. The composition of claim 1, wherein the cDNAs are immobilized on a substrate.

5. A high throughput method for detecting differential expression of one or more cDNAs in a sample containing nucleic acids, the method comprising:

(a) hybridizing the substrate of claim 4 with nucleic acids of the sample, thereby forming one or more hybridization complexes;

(b) detecting the hybridization complexes; and

(c) comparing the hybridization complexes with those of a standard, wherein differences between the standard and sample hybridization indexes indicate differential expression of cDNAs in the sample.

6. The method of claim 5, wherein the nucleic acids of the sample are amplified prior to hybridization.

7. The method of claim 5, wherein the sample is from a subject with prostate cancer and comparison with a standard defines an early, mid, or late stage of that disease.

8. A high throughput method of screening a plurality of molecules or compounds to identify a ligand which specifically binds a cDNA, the method comprising:

(a) combining the composition of claim 1 with the plurality of molecules or compounds under conditions to allow specific binding; and

(b) detecting specific binding between each cDNA and at least one molecule or compound, thereby identifying a ligand that specifically binds to each cDNA.

9. The method of claim 8 wherein the plurality of molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acid molecules, mimetics, peptides, transcription factors, repressors, and regulatory proteins.

10. An isolated cDNA encoding a protein comprising the amino acid sequence of SEQ ID NO: 36.

11. A vector containing the cDNA of claim 10.
12. A host cell containing the vector of claim 11.
13. A method for producing a protein, the method comprising the steps of:
   (a) culturing the host cell of claim 12 under conditions for expression of protein; and
   (b) recovering the protein from the host cell culture.
15. A high-throughput method for using a protein to screen a plurality of molecules or compounds to identify at least one ligand which specifically binds the protein, the method comprising:
   (a) combining the protein of claim 14 with the plurality of molecules or compounds under conditions to allow specific binding; and
   (b) detecting specific binding between the protein and a molecule or compound, thereby identifying a ligand which specifically binds the protein.
16. The method of claim 15 wherein the plurality of molecules or compounds is selected from DNA molecules, RNA molecules, peptide nucleic acid molecules, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds, and pharmaceutical agents.
17. A method of using a protein to produce an antibody, the method comprising:
   a) immunizing an animal with the protein of claim 14 under conditions to elicit an antibody response;
   b) isolating animal antibodies; and
   c) screening the isolated antibodies with the protein, thereby identifying an antibody which specifically binds the protein.
18. A method of purifying an antibody, the method comprising:
   a) combining the protein of claim 14 with a sample under conditions to allow specific binding;
   b) recovering the bound protein; and
   c) separating the protein from the antibody, thereby obtaining purified antibody.