

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
4 March 2004 (04.03.2004)

PCT

(10) International Publication Number
WO 2004/018633 A2

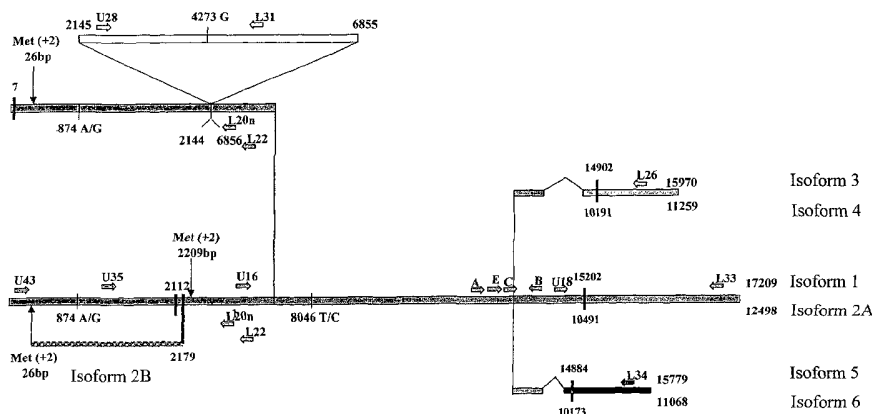
- (51) International Patent Classification⁷: **C12N**
- (21) International Application Number:
PCT/US2003/026335
- (22) International Filing Date: 19 August 2003 (19.08.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/405,006 20 August 2002 (20.08.2002) US
60/454,502 12 March 2003 (12.03.2003) US
- (71) Applicants (for all designated States except US): **THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; National Institutes Of Health, Office Of Technology Transfer, 6011 Executive Boulevard, Suite 325, Rockville, MD 20852-3804 (US). **AVENTIS PHARMA SA** [FR/FR]; 20, Avenue Raymond Aron, F-92165 Antony, Croix-de-Berny (FR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **DEAN, Micheal, C.** [US/US]; 1362 Hitchingpost Lane, Frederick, MD 21703 (US). **ARNOULD-REGUIGNE, Isabelle** [FR/FR]; 112,

Rue De Bry, F-94430 Chennevières Sur Marne (FR). **PRADES, Catherine** [FR/FR]; 30 Avenue Du General De Gaulle, F-94320 Thiais (FR). **ROSIER-MONTUS, Marie-Francoise** [FR/FR]; 21, rue des Baconnets, F-92160 Antony (FR). **DENEFFLE, Patrice** [FR/FR]; 45 Avenue Des Fusilles De Chateau, F-94100 Saint Maur (FR). **SHULENIN, Sergey** [RU/US]; 130 Stonegate Drive, Frederick, MD 21702 (US). **ANNILO, Tarmo** [EE/US]; 146 Willowdale Drive, #34, Frederick, MD 32702 (US). **TRIUNFOL, Marcia, L.** [BR/US]; c/o American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005 (US).

- (74) Agent: **HARDING, Tanya, M.**; Klarquist, Sparkman, LLP, One World Trade Center, Suite 1600, 121 SW Salmon Street, Portland, OR 97204 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: ABCA13 NUCLEIC ACIDS AND PROTEINS, AND USES THEREOF



(57) Abstract: This disclosure provides seven isoforms of a novel transporter protein, referred to as ABCA13, nucleic acid molecules encoding the disclosed isoforms, variants and fragments thereof, and methods of making and using these molecules. Also provided are methods of ameliorating, treating, detecting, prognosing, and diagnosing diseases and conditions believed to be associated with altered expression of ABCA13, such as hypercholesterolemia, drug resistance, retinal degeneration, or neurological disease. Kits are also provided.

WO 2004/018633 A2



ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,
SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv)) for US only*

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

-1-

ABCA13 NUCLEIC ACIDS AND PROTEINS, AND USES THEREOF**CROSS REFERENCE TO RELATED APPLICATIONS**

This application claims the benefit of U.S. Provisional Application No. 60/405,006, filed
5 August 20, 2002, and No. 60/454,502, filed March 12, 2003. Both of these applications are
incorporated herein in their entirety.

FIELD OF THE DISCLOSURE

The present disclosure is related to extra- and intra-cellular transport, including the
10 mechanisms controlling transport, diseases that arise from defects in such mechanisms, and methods
of influencing (either inhibiting or enhancing or otherwise changing) transport.

BACKGROUND OF THE DISCLOSURE

The ABC (ATP-binding cassette transporter) gene superfamily encodes active transporter
15 proteins and constitutes a family of proteins that have been extremely well conserved through
evolution, from bacteria to humans (Ames and Lecar, *FASEB J.*, 6: 2660-2666, 1992). The ABC
proteins are involved in extra- and intracellular membrane transport of various substrates, for
example ions, metals, amino acids, lipids, peptides, sugars, vitamins, or steroid hormones, across
membranes. Among the forty characterized human ABC genes, eleven members have been described
20 as associated with human disease, such as *ABCA1*, *ABCA4* (*ABCR*) and *ABCC7*, (*CFTR*) which are
thought to be involved in Tangier disease (Bodzioch *et al.*, *Nat. Genet.* 22(4): 347-351, 1999;
Brooks-Wilson *et al.*, *Nat. Genet.* 22(4): 336-345, 1999; Rust *et al.*, *Nat. Genet.* 22: 352-355, 1999),
the Stargardt disease (Lewis *et al.*, *Am. J. Hum. Genet.* 64: 422-434, 1999), and Cystic Fibrosis
(Riordan *et al.*, *Science* 245: 1066-1073, 1989), respectively. Another ABC transporter, *mdr1*
25 (multidrug resistance) gene, encodes the protein *mdr1a*, also called P-glycoprotein (P-gp). This
protein functions as a drug-efflux transmembrane protein pump. P-glycoprotein was first identified
over 20 years ago in chemotherapeutic drug-resistant tumor cells, and is now known to be a major
cause of multidrug resistance in many cancers (Van Asperen *et al.*, *J. Pharmaceut. Sci.* 86: 881-884,
1997; Tsuji, *Therap. Drug Monitor.* 20: 588-590, 1998). These implications reveal the importance of
30 the functional role of the ABC gene family. The discovery of a new family of ABC gene members is
expected to provide new insights into the physiopathology and treatment of human diseases.

The prototype ABC protein binds ATP and uses the energy from ATP hydrolysis to drive
the transport of various molecules across cell membranes. The functional protein contains two ATP-
binding domains (nucleotide binding fold, NBF) and multiple transmembrane (TM) domains. The
35 genes are typically organized as full transporters containing two of each domain, or half transporters
with only one of each domain. Most full transporters are arranged in a TM-NBF-TM-NBF fashion
(Dean *et al.*, *Curr Opin Genet* 5: 79-785, 1995).

Analysis of amino acids sequence alignments of the ATP-binding domains has allowed the
ABC genes to be separated into seven sub-families (Allikmets *et al.*, *Hum Mol Genet* 5: 1649-1655,

-2-

1996). Currently, according to the recent HUGO classification, seven ABC gene sub-families named ABC (A to G) have been described in the human genome (ABC1, CFTR/MRP, MDR, ABC8, ALD, GCN20, OABP) with all except one (OABP) containing multiple members. Among the ABC sub-families, the ABCA gene subfamily is probably the most evolutionary complex. The ABCA subfamily consists exclusively of full transporter genes. The *ABCA* genes and *OABP* represent the only two sub-families of ABC genes that do not have identifiable orthologs in the yeast genome (Decottignies and Goffeau, *Nat. Genet.* Feb;15(2):137-145,1997; Michaelis and Berkower, *Cold Spring Harb. Symp. Quant. Biol.* 60:291-307, 1995). There is, however, at least one *ABCA*-related gene in *C. elegans* (*ced-7*) and several in *Drosophila*. Thus, the ABCA genes appear to have diverged after eukaryotes became multicellular and developed more sophisticated transport requirements.

ABCA1 was demonstrated to be the gene responsible for Tangier disease, a disorder characterized by high levels of cholesterol in peripheral tissues, and a very low level of HDLs, and familial hypoalphalipoproteinemia (FHD) (Bodzioch *et al.*, *Nat Genet* 22: 347-351, 1999; Brooks-Wilson *et al.*, *Nat. Genet.* 4: 336-345, 1999; Rust *et al.*, *Nat Genet* 22: 352-355, 1999; Marcil *et al.*, *The Lancet* 354: 1341-1346, 1999). The ABCA1 protein is proposed to function in the reverse transport of cholesterol from peripheral tissues via an interaction with the apolipoprotein 1 (ApoA-1) of HDL tissues (see Wang *et al.*, *JBC* 275(42): 33053-33058, 2000).

The *ABCA2* gene is highly expressed in the brain, and *ABCA3* in the lung, but no function has been ascribed to these loci. The *ABCA4* gene is exclusively expressed in the rod photoreceptors of the retina and mutations thereof are responsible for several pathologies of human eyes, such as retinal degenerative disorders (Allikmets *et al.*, *Science* 277: 1805-1807, 1997; Allikmets *et al.*, *Nat. Genet.* 15: 236-246, 1997; Sun *et al.*, *J. Biol. Chem.* 8269-8281, 1999; Weng *et al.*, *Cell* 98: 13-23, 1999; Cremers *et al.*, *Hum. Mol. Genet.* 7: 355-362, 1998; Martinez-Mir *et al.*, *Genomics* 40: 142-146, 1997). *ABCA4* is believed to transport retinal and/or retinal-phospholipid complexes from the rod photoreceptor outer segment disks to the cytoplasm, facilitating phototransduction.

Characterization of new genes from the ABCA subfamily is likely to yield biologically important transporters, which may have translocase activity for membrane lipid transport or the transport of other substances and which may play a role in human pathologies.

30

SUMMARY OF THE DISCLOSURE

Disclosed herein are seven novel isoforms of a new gene belonging to the ABCA protein sub-family, which has been designated *ABCA13*. The protein is believed to be involved in the energy-dependent transport of one or a variety of substances, for example ions, metals, amino acids, lipids, peptides, sugars, vitamins and steroid hormones. The newly discovered gene shows similar gene organization and considerable conservation of the amino acid sequence with other ABCA transporter genes, particularly within the transmembrane domains (TM) and two ATP-binding domains (nucleotide binding domain, NBD1 and NBD2). The present disclosure includes novel

35

-3-

nucleic acids encoding the isoforms of ABCA13 (SEQ ID NO: 1, 18, 20, 22, 24, 26, and 28) and their predicted amino acid sequences (SEQ ID NO: 2, 19, 21, 23, 25, 27, and 29).

Also disclosed are methods of using these molecules in detecting biological conditions associated with mutation, altered expression, or duplication of *ABCA13* in a subject and methods of screening for agents that modulate ABCA13 transporter activity, such as specific transport inhibitors. Oligonucleotides for use in examples of such methods are also provided.

Also disclosed herein are protein specific binding agents, such as antibodies, that bind specifically to at least one epitope of an ABCA13 variant protein preferentially compared to wildtype ABCA13, and methods of using such antibodies in diagnosis and screening.

Kits are also provided for carrying out the methods described herein.

The foregoing and other features and advantages will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a series of diagrams showing the various transcripts of the *ABCA13* gene and corresponding isoforms.

FIG 1A shows a physical map of the 7p12.3 region containing the *ABCA13* gene. The location of the boundaries of BAC clones AC073927 is shown as an inset portion of the second Nucleotide Binding Domain region (NBD2) of *ABCA13*. Gene orientation is indicated by the arrows and the size and location of the corresponding transcripts is shown in the map. Analysis was performed by combination of several gene-finding programs such as GENSCAN (Burge and Karlin, *J Mol Biol.*; 268(1):78-94, 1997), FGENEH/FEXH (Solovyev and Salamov, *Ismb*; 5:294-302, 1997), and XPOUND (Thomas and Skolnick, *J Math Appl Med Biol.*, 11(1):1-16, 1994).

FIG 1B is a diagram showing the six different *ABCA13* transcripts that have been identified and are described herein. Termination codons are indicated by thick upright black lines. The large arrows symbolize the primers. The Taqman reagents are symbolized by a light grey star. Two *ABCA13* proteins are depicted by dark grey lines. The small protein that can be predicted in the 5' part of transcripts 2, 4, and 6 is indicated at the bottom left, by a black line.

FIG 1C is a first diagram of the predicted proteins from the six ABCA13 transcripts. Amino acid numbering of the different protein domains is given. Key: SP, signal peptide; TM, predicted transmembrane domain, NBD, nucleotide binding domain; TMA, additional TM domain; C-term, C-terminal regions due to alternative splicing.

FIG 1D is an alignment of the Isoforms of ABCA13 with the full length ABCA13. The STOP position corresponds to last base pair of the codon triplet. For each 3' end, the base pair positions above or below each bar correspond to the respective Isoform. Isoform 2B is depicted at the bottom left, between base pairs 26 and 2179.

Figure 2 is phylogenetic linkage tree showing the relationship of the ABCA gene family. The phylogenetic tree was constructed with the alignments of the N- and C-terminal ATP-binding

-4-

domains' sequence by the Neighbor joining method. The numbers indicate the bootstrap values (confidence interval for each branch after 100 conducted iterations, see Kumar *et al.*, *Bioinformatics* 17(12): 1244-1245, 2001) between the bracketed genes.

5 **Figure 3** is an alignment of the N-terminal region of ABCA subfamily proteins. Amino acid number one corresponds to the first position in all proteins, otherwise, the numbering is according to the ABCA1 sequence. Dashes indicate gaps.

Figure 4 is a diagram of the predicted amino-acid sequence of ABCA13 Isoform 1 (SEQ ID NO: 2). The predicted transmembrane domains (*e.g.*, residues 20-42, 3571-3793, 3414-3636, 3646-3668, 3677-3697, 3751-3773, 4226-4247, 4456-4473, 4508-4530, 4535-4556, 4571-4532, and 4654-10 4675 of SEQ ID NO: 2) and the hydrophobic domain at residues 3707-3729 of SEQ ID NO: 2 are shaded, large extracellular domains are underlined and the NBDs (*e.g.*, residues 3868-4050 and 4747-4932 of SEQ ID NO: 2) are boxed.

Figure 5 illustrates RNA expression analysis of human and mouse *ABCA13*. FIG 5A shows human *ABCA13* expression. This radial graph shows the expression profile that we have determined 15 using the set of probe and primers specific for transcripts 1 and 2. The arbitrary unit has been calculated after normalization to β 2-microglobulin.

FIG 5B is a blot that shows mouse *Abca13* expression. PCR was performed on the panel of mouse tissue-specific cDNAs, as described in the examples. Amplification products of 255 bp (fragment of exons one, two and three) and 381 bp (fragment of exons 17 and 18) were seen in a 20 reaction containing cDNA from kidney. A control reaction was performed with primers specific to G3PDH (glyceraldehyde 3-phosphate dehydrogenase).

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

25 The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three-letter code for amino acids, as defined in 37 C.F.R. § 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. In the accompanying sequence listing:

30 SEQ ID NO: 1 shows a cDNA encoding human *ABCA13* (Isoform 1) and the predicted amino acid sequence.

SEQ ID NO: 2 shows the amino acid sequence of the human ABCA13 Isoform 1 protein encoded by *ABCA13* (SEQ ID NO: 1).

SEQ ID NO: 3 through 15 show the nucleic acid sequences of primers used to amplify particular regions within the several isoforms of *ABCA13*.

35 SEQ ID NO: 16 and 17 show primers that may be used to amplify mouse *ABCA13*.

SEQ ID NO: 18 shows a cDNA encoding human *ABCA13* Isoform 2A and the predicted amino acid sequence.

SEQ ID NO: 19 shows the amino acid sequence of the human ABCA13 Isoform 2A protein encoded by *ABCA13* (SEQ ID NO: 18).

-5-

SEQ ID NO: 20 shows a cDNA encoding human *ABCA13* Isoform 2B and the predicted amino acid sequence.

SEQ ID NO: 21 shows the amino acid sequence of the human *ABCA13* Isoform 2B protein encoded by *ABCA13* (SEQ ID NO: 20).

5 SEQ ID NO: 22 shows a cDNA encoding human *ABCA13* Isoform 3, and the predicted amino acid sequence.

SEQ ID NO: 23 shows the amino acid sequence of the human *ABCA13* Isoform 3 protein encoded by *ABCA13* (SEQ ID NO: 22).

10 SEQ ID NO: 24 shows a cDNA encoding human *ABCA13* Isoform 4 and the predicted amino acid sequence.

SEQ ID NO: 25 shows the amino acid sequence of the human *ABCA13* Isoform 4 protein encoded by *ABCA13* (SEQ ID NO: 24).

SEQ ID NO: 26 shows a cDNA encoding human *ABCA13* Isoform 5 and the predicted amino acid sequence.

15 SEQ ID NO: 27 shows the amino acid sequence of the human *ABCA13* Isoform 5 protein encoded by *ABCA13* (SEQ ID NO: 26).

SEQ ID NO: 28 shows a cDNA encoding human *ABCA13* Isoform 6 and the predicted amino acid sequence.

20 SEQ ID NO: 29 shows the amino acid sequence of the human *ABCA13* Isoform 6 protein encoded by *ABCA13* (SEQ ID NO: 28).

SEQ ID NOs: 30-41 show a portion of the N-terminal amino acid sequence of the *ABCA1*, *ABCA2*, *ABCA3*, *ABCA4*, *ABCA5*, *ABCA6*, *ABCA7*, *ABCA8*, *ABCA9*, *ABCA10*, *ABCA11*, *ABCA12*, and *ABCA13* proteins, respectively.

25 SEQ ID NOs: 42-56 show nucleic acid sequences of primers used for RACE amplification of *ABCA13* sequences, and/or validation of such amplification.

SEQ ID NOs: 57-65 show nucleic acid sequences of primers used in real-time CPR analysis of *ABCA13*.

SEQ ID NOs: 66-72 show nucleic acid sequences of primers used in quantification of *ABCA13* sequences.

30 SEQ ID NOs: 73-75 show nucleic acid sequences of additional primers used to amplify particular regions within isoforms of *ABCA13*.

DETAILED DESCRIPTION

35 I. Abbreviations

40	ABC:	ATP-binding cassette transporter
	BAC:	bacterial artificial chromosome
	bp	base pair(s)
	DNA:	deoxyribonucleic acid
	gDNA:	genomic DNA

-6-

pfu: plaque forming unit
 ELISA: enzyme-linked immunosorbant assay
 PCR: polymerase chain reaction
 TAE: Tris acetate EDTA

5

II. Terms

Unless otherwise noted, technical terms are used according to conventional usage.

Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

In order to facilitate review of the various embodiments of the invention, the following explanations of specific terms are provided:

15 **Altered expression:** Expression of a nucleic acid (*e.g.*, mRNA or protein) in a subject or biological sample from a subject that deviates from expression in a subject or biological sample from a subject having normal characteristics for the biological condition associated with the nucleic acid. Normal expression can be found in a control, a standard for a population, etc. For instance, where the altered expression manifests as a transporter disease condition, such as deficient extra- or intracellular
 20 transport, characteristics of normal expression might include an individual who is not suffering from the transport disorder, a population standard of individuals believed not to be suffering from the disease, etc. For instance, certain altered expression, such as altered *ABCA13* nucleic acid or *ABCA13* protein expression, can be described as being associated with the biological conditions of altered (*e.g.*, reduced) transporter function and tendency to develop a transporter deficiency.

25 Likewise, altered expression may be associated with a disease. The term "associated with" includes an increased risk of developing the disease as well as the disease itself.

Altered protein expression, such as altered *ABCA13* protein expression, also refers to expression of a protein that is in some manner different to expression of the protein in a normal (wild type) situation. This includes but is not necessarily limited to: (1) a mutation in the protein such that
 30 one or more of the amino acid residues is different; (2) a short deletion or addition of one or a few amino acid residues to the sequence of the protein; (3) a longer deletion or addition of amino acid residues, such that an entire protein domain or sub-domain is removed or added; (4) expression of an increased amount of the protein, compared to a control or standard amount; (5) expression of an decreased amount of the protein, compared to a control or standard amount; (6) alteration of the
 35 subcellular localization or targeting of the protein; (7) alteration of the temporally regulated expression of the protein (such that the protein is expressed when it normally would not be, or alternatively is not expressed when it normally would be); and (8) alteration of the localized (*e.g.*, organ or tissue specific) expression of the protein (such that the protein is not expressed where it would normally be expressed or is expressed where it normally would not be expressed), each
 40 compared to a control or standard.

-7-

Animal: Living multi-cellular vertebrate organisms, a category that includes for example, mammals and birds.

Amplification: When used in reference to a nucleic acid, this refers to a collection of techniques that increase the number of copies of a nucleic acid molecule in a sample or specimen.

5 An example of amplification is the polymerase chain reaction, in which a biological sample collected from a subject is contacted with a pair of oligonucleotide primers, under conditions that allow for the hybridization of the primers to nucleic acid template in the sample. The primers are extended under suitable conditions, dissociated from the template, and then re-annealed, extended, and dissociated to amplify the number of copies of the nucleic acid. The product of *in vitro* amplification can be
10 characterized by electrophoresis, restriction endonuclease cleavage patterns, oligonucleotide hybridization or ligation, and/or nucleic acid sequencing, using standard techniques. Other examples of *in vitro* amplification techniques include strand displacement amplification (*see* U.S. Patent No. 5,744,311); transcription-free isothermal amplification (*see* U.S. Patent No. 6,033,881); repair chain reaction amplification (*see* WO 90/01069); ligase chain reaction amplification (*see* EP-A-320 308);
15 gap filling ligase chain reaction amplification (*see* U.S. Patent No. 5,427,930); coupled ligase detection and PCR (*see* U.S. Patent No. 6,027,889); and NASBA™ RNA transcription-free amplification (*see* U.S. Patent No. 6,025,134).

Antisense, Sense, and Antigene: Double-stranded DNA (dsDNA) has two strands, a 5' -> 3' strand, referred to as the plus strand, and a 3' -> 5' strand (the reverse complement), referred to as
20 the minus strand. Because RNA polymerase adds nucleic acids in a 5' -> 3' direction, the minus strand of the DNA serves as the template for the RNA during transcription. Thus, the RNA formed will have a sequence complementary to the minus strand and identical to the plus strand (except that U is substituted for T).

Antisense molecules are molecules that are specifically hybridizable or specifically
25 complementary to either RNA or the plus strand of DNA. Sense molecules are molecules that are specifically hybridizable or specifically complementary to the minus strand of DNA. Antigene molecules are either antisense or sense molecules directed to a dsDNA target.

Binding or stable binding: An oligonucleotide binds or stably binds to a target nucleic acid if a sufficient amount of the oligonucleotide forms base pairs or is hybridized to its target nucleic
30 acid, to permit detection of that binding. Binding can be detected by either physical or functional properties of the target:oligonucleotide complex. Binding between a target and an oligonucleotide can be detected by any procedure known to one skilled in the art, including both functional and physical binding assays. Binding can be detected functionally by determining whether binding has an observable effect upon a biosynthetic process such as expression of a gene, DNA replication,
35 transcription, translation and the like.

Physical methods of detecting the binding of complementary strands of DNA or RNA are well known in the art, and include such methods as DNase I or chemical footprinting, gel shift and affinity cleavage assays, Northern blotting, dot blotting and light absorption detection procedures. For example, one method that is widely used, because it is so simple and reliable, involves observing

a change in light absorption of a solution containing an oligonucleotide (or an analog) and a target nucleic acid at 220 to 300 nm as the temperature is slowly increased. If the oligonucleotide or analog has bound to its target, there is a sudden increase in absorption at a characteristic temperature as the oligonucleotide (or analog) and the target disassociate from each other, or melt.

5 The binding between an oligomer and its target nucleic acid is frequently characterized by the temperature (T_m) at which 50% of the oligomer is melted from its target. A higher (T_m) means a stronger or more stable complex relative to a complex with a lower (T_m).

Biological condition: Designates a condition of a subject that can be assessed through observation or through the analysis of a biological sample, *e.g.*, altered expression level of ABCA13
10 protein in comparison to a control expression level, or ability of cells from a subject to transport cholesterol.

Biological sample: Any sample in which the presence of a protein and/or ongoing expression of a protein may be detected. Suitable biological samples include samples containing genomic DNA or RNA (including mRNA), obtained from body cells of a subject, such as but not
15 limited to those present in peripheral blood, urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material.

cDNA (complementary DNA): A piece of DNA lacking internal, non-coding segments (introns) and transcriptional regulatory sequences. cDNA can also contain untranslated regions (UTRs) that are responsible for translational control in the corresponding RNA molecule. cDNA is synthesized
20 in the laboratory by reverse transcription from messenger RNA extracted from cells.

DNA (deoxyribonucleic acid): A long chain polymer that comprises the genetic material of most living organisms (some viruses have genes comprising ribonucleic acid (RNA)). The repeating units in DNA polymers are four different nucleotides, each of which comprises one of the four bases, adenine, guanine, cytosine and thymine bound to a deoxyribose sugar to which a
25 phosphate group is attached. Triplets of nucleotides (referred to as codons) code for each amino acid in a polypeptide. The term codon is also used for the corresponding (and complementary) sequences of three nucleotides in the mRNA into which the DNA sequence is transcribed.

 Unless otherwise specified, any reference to a DNA molecule is intended to include the reverse complement of that DNA molecule. Except where single-strandedness is required by the text
30 herein, DNA molecules, though written to depict only a single strand, encompass both strands of a double-stranded DNA molecule. Thus, a reference to the nucleic acid molecule that encodes a specific protein, or a fragment thereof, encompasses both the sense strand and its reverse complement. Thus, for instance, it is appropriate to generate probes or primers from the reverse complement sequence of the disclosed nucleic acid molecules.

35 **Deletion:** The removal of a sequence of DNA, the regions on either side being joined together.

Encode: A polynucleotide is said to "encode" a polypeptide if, in its native state or when manipulated by methods well known to those skilled in the art, it can be transcribed and/or translated

to produce the mRNA for and/or the polypeptide or a fragment thereof. The anti-sense strand is the complement of such a nucleic acid, and the encoding sequence can be deduced therefrom.

Functional fragments and variants of a polypeptide: Included are those fragments and variants that maintain at least one function of the parent polypeptide. It is recognized that the gene or cDNA encoding a polypeptide can be considerably mutated without materially altering one or more the polypeptide's functions. First, the genetic code is well known to be degenerate, and thus different codons encode the same amino acids. Second, even where an amino acid substitution is introduced, the mutation can be conservative and have no material impact on the essential functions of a protein (see Stryer, *Biochemistry* 4th Ed., W. Freeman & Co., New York, NY, 1995). Third, part of a polypeptide chain can be deleted without impairing or eliminating all of its functions. *e.g.*, sequence variants a protein, such as a 5' or 3' variant, may retain the full function of an entire protein. Fourth, insertions or additions can be made in the polypeptide chain for example, adding epitope tags, without impairing or eliminating its functions (Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Intersciences, 1998). Other modifications that can be made without materially impairing one or more functions of a polypeptide include, for example, *in vivo* or *in vitro* chemical and biochemical modifications or the incorporation of unusual amino acids. Such modifications include, for example, acetylation, carboxylation, phosphorylation, glycosylation, ubiquitination, labeling, *e.g.*, with radionucleides, and various enzymatic modifications, as will be readily appreciated by those well skilled in the art. A variety of methods for labeling polypeptides and labels useful for such purposes are well known in the art, and include radioactive isotopes such as ³²P, ligands that bind to or are bound by labeled specific binding partners (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands. Functional fragments and variants can be of varying length. For example, a fragment may consist of 10 or more, 25 or more, 50 or more, 75 or more, 100 or more, or 200 or more amino acid residues.

A functional fragment or variant of ABCA13 is defined herein as a polypeptide that is capable of transporter activity, including any polypeptide six or more amino acid residues in length that is capable of transporter activity. Fragments of ABCA13 that contain one or more domains as described herein, whether in the native arrangement or order, or in a different arrangement or order, are specifically contemplated.

Heterologous: A type of sequence that is not normally (*i.e.* in the wild-type sequence) found adjacent to a second sequence. In one embodiment, the sequence is from a different genetic source, such as a virus or organism, than the second sequence.

Hybridization: Oligonucleotides and their analogs hybridize by hydrogen bonding, which includes Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary bases. Generally, nucleic acid consists of nitrogenous bases that are either pyrimidines (cytosine (C), uracil (U), and thymine (T)) or purines (adenine (A) and guanine (G)). These nitrogenous bases form hydrogen bonds between a pyrimidine and a purine, and the bonding of the pyrimidine to the purine is referred to as "base pairing." More specifically, A will hydrogen bond

-10-

to T or U, and G will bond to C. "Complementary" refers to the base pairing that occurs between two distinct nucleic acid sequences or two distinct regions of the same nucleic acid sequence.

Specifically hybridizable and specifically complementary are terms that indicate a sufficient degree of complementarity such that stable and specific binding occurs between the oligonucleotide (or its analog) and the DNA or RNA target. The oligonucleotide or oligonucleotide analog need not be 100% complementary to its target sequence to be specifically hybridizable. An oligonucleotide or analog is specifically hybridizable when binding of the oligonucleotide or analog to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide or analog to non-target sequences under conditions where specific binding is desired, for example under physiological conditions in the case of *in vivo* assays or systems. Such binding is referred to as specific hybridization.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (especially the Na⁺ concentration) of the hybridization buffer will determine the stringency of hybridization, though wash times also influence stringency. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed by Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, chapters 9 and 11, herein incorporated by reference.

The following are exemplary hybridization conditions and are not meant to be limiting.

Very High Stringency (detects sequences that share 90% sequence identity)

25	Hybridization:	5x SSC at 65°C for 16 hours
	Wash twice:	2x SSC at room temperature (RT) for 15 minutes each
	Wash twice:	0.5x SSC at 65°C for 20 minutes each

High Stringency (detects sequences that share 80% sequence identity or greater)

30	Hybridization:	5x-6x SSC at 65°C-70°C for 16-20 hours
	Wash twice:	2x SSC at RT for 5-20 minutes each
	Wash twice:	1x SSC at 55°C-70°C for 30 minutes each

Low Stringency (detects sequences that share greater than 50% sequence identity)

35	Hybridization:	6x SSC at RT to 55°C for 16-20 hours
	Wash at least twice:	2x-3x SSC at RT to 55°C for 20-30 minutes each.

Isoform: As used herein, this term refers to are forms of a protein that have the same or similar biological function, but differ in the amino acid sequence comprising the protein (see *e.g.*, Chen *et al.*, *J. Exp. Biol.* 205(Pt 17): 2677-2686, 2002). Isoforms of a protein may be produced by different genes or by alternative splicing of RNA transcripts from the same gene. The amino acid sequences of seven isoforms of the ABCA13 protein are disclosed herein as SEQ ID NO: 2, 19, 21, 23, 25, 27, and 29.

Isolated: A biological component (such as a nucleic acid molecule, protein or organelle) that has been substantially completely separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *i.e.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have
5 been **isolated** include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

Labeled: A biomolecule attached covalently or noncovalently to a detectable label or reporter molecule. Typical labels include radioactive isotopes, enzyme substrates, co-factors, ligands, chemiluminescent or fluorescent agents, haptens, and enzymes. Methods for labeling and guidance in
10 the choice of labels appropriate for various purposes are discussed, *e.g.*, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989 and Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Intersciences, 1998. For example, ATP can be labeled in any one of its three phosphate groups with radioisotopes such as ^{32}P or ^{33}P , or
15 in its sugar moiety with radioisotopes such as ^{35}S .

Mammal: This term includes both human and non-human mammals. Similarly, the term **subject** includes both human and veterinary subjects.

Modulator: An agent that increases or decreases (modulates) the activity of a protein as measured by the change in an experimental parameter. A modulator can be essentially any
20 compound, such as a chemotherapeutic agent, a polypeptide, a hormone, a nucleic acid, a sugar, a lipid and the like.

Mutation: Any change of the DNA sequence within a gene or chromosome. In some instances, a mutation will alter a characteristic or trait (phenotype), but this is not always the case. Types of mutations include base substitution point mutations (*e.g.*, transitions or transversions),
25 deletions, and insertions. Missense mutations are those that introduce a different amino acid into the sequence of the encoded protein; nonsense mutations are those that introduce a new stop codon. In the case of insertions or deletions, mutations can be in-frame (not changing the frame of the overall sequence) or frame shift mutations, which may result in the misreading of a large number of codons (and often leads to abnormal termination of the encoded product due to the presence of a stop codon
30 in the alternative frame).

This term specifically encompasses variations that arise through somatic mutation, for instance those that are found only in disease cells, but not constitutionally, in a given individual. Examples of such somatically-acquired variations include the point mutations that frequently result in altered function of various genes that are involved in development of cancers. This term also
35 encompasses DNA alterations that are present constitutionally, that alter the function of the encoded protein in a readily demonstrable manner, and that can be inherited by the children of an affected individual. In this respect, the term overlaps with "polymorphism," as defined below, but generally refers to the subset of constitutional alterations that have arisen within the past few generations in a kindred and that are not widely disseminated in a population group. In particular embodiments, the

-12-

term is directed to those constitutional alterations that have major impact on the health of affected individuals.

Nucleotide: This term includes, but is not limited to, a monomer that includes a base linked to a sugar, such as a pyrimidine, purine or synthetic analogs thereof, or a base linked to an amino acid, as in a peptide nucleic acid (PNA). A nucleotide is one monomer in a polynucleotide. A nucleotide sequence refers to the sequence of bases in a polynucleotide.

Oligonucleotide: A plurality of joined nucleotides joined by native phosphodiester bonds, between about 6 and about 300 nucleotides in length. An oligonucleotide analog refers to moieties that function similarly to oligonucleotides but have non-naturally occurring portions. For example, oligonucleotide analogs can contain non-naturally occurring portions, such as altered sugar moieties or inter-sugar linkages, such as a phosphorothioate oligodeoxynucleotide. Functional analogs of naturally occurring polynucleotides can bind to RNA or DNA, and include peptide nucleic acid (PNA) molecules.

Particular oligonucleotides and oligonucleotide analogs can include linear sequences up to about 200 nucleotides in length, for example a sequence (such as DNA or RNA) that is at least 6 bases, for example at least 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100 or even 200 bases long, or from about 6 to about 50 bases, for example about 10-25 bases, such as 12, 15 or 20 bases.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

Open reading frame: A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide.

Ortholog: Two nucleic acid or amino acid sequences are orthologs of each other if they share a common ancestral sequence and diverged when a species carrying that ancestral sequence split into two species. Orthologous sequences are also homologous sequences.

Polypeptide: A polymer in which the monomers are amino acid residues that are joined together through amide bonds. When the amino acids are alpha-amino acids, either the L-optical isomer or the D-optical isomer can be used, the L-isomers being preferred. The term **polypeptide** or **protein** as used herein encompasses any amino acid sequence and includes modified sequences such as glycoproteins. The term **polypeptide** is specifically intended to cover naturally occurring proteins, as well as those that are recombinantly or synthetically produced.

The term **polypeptide fragment** refers to a portion of a polypeptide that exhibits at least one useful epitope. The phrase "**functional fragments of a polypeptide**" refers to all fragments of a polypeptide that retain an activity, or a measurable portion of an activity, of the polypeptide from which the fragment is derived. Fragments, for example, can vary in size from a polypeptide fragment as small as an epitope capable of binding an antibody molecule to a large polypeptide capable of

participating in the characteristic induction or programming of phenotypic changes within a cell. An **epitope** is a region of a polypeptide capable of binding an immunoglobulin generated in response to contact with an antigen. Thus, smaller peptides containing the biological activity of insulin, or conservative variants of the insulin, are thus included as being of use.

5 The term **soluble** refers to a form of a polypeptide that is not inserted into a cell membrane.

Conservative amino acid substitution tables providing functionally similar amino acids are well known to one of ordinary skill in the art. The following six groups are examples of amino acids that are considered to be conservative substitutions for one another:

- 1) Alanine (A), Serine (S), Threonine (T);
- 10 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

15 Variations in the cDNA sequence that result in amino acid changes, whether conservative or not, are usually minimized in order to preserve the functional and immunologic identity of the encoded protein. The immunologic identity of the protein may be assessed by determining whether it is recognized by an antibody; a variant that is recognized by such an antibody is immunologically conserved. A cDNA sequence variant will preferably introduce no more than thirty, in some
20 instances no more than twenty, and preferably fewer than ten amino acid substitutions into the encoded polypeptide. Variant amino acid sequences may, for example, be 80, 90 or even 95% or 98% identical to the native amino acid sequence. Programs and algorithms for determining percentage identity can be found at the NCBI website.

Polymorphism: Variant in a sequence of a gene, usually carried from one generation to
25 another in a population. Polymorphisms can be those variations (nucleotide sequence differences) that, while having a different nucleotide sequence, produce functionally equivalent gene products, such as those variations generally found between individuals, different ethnic groups, geographic locations. The term polymorphism also encompasses variations that produce gene products with altered function, for instance variants in the gene sequence that lead to gene products that are not
30 functionally equivalent. This term also encompasses variations that produce no gene product, an inactive gene product, or increased or increased activity gene product.

Polymorphisms can be referred to, for instance, by the nucleotide position at which the variation exists, by the change in amino acid sequence caused by the nucleotide variation, or by a change in some other characteristic of the nucleic acid molecule or protein that is linked to the
35 variation (*e.g.*, an alteration of a secondary structure such as a stem-loop, or an alteration of the binding affinity of the nucleic acid for associated molecules, such as polymerases, RNases, and so forth).

Probes and primers: Nucleic acid probes and primers can be readily prepared based on the nucleic acid molecules provided in this disclosure. A probe comprises an isolated nucleic acid attached

to a detectable label or reporter molecule. Typical labels include radioactive isotopes, enzyme substrates, co-factors, ligands, chemiluminescent or fluorescent agents, haptens, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, *e.g.*, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and Ausubel
5 *et al.* (In *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Intersciences, 1992).

Primers are short nucleic acid molecules, preferably DNA oligonucleotides 10 nucleotides or more in length. More preferably, longer DNA oligonucleotides can be about 15, 17, 20, or 23 nucleotides or more in length. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then the
10 primer extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, *e.g.*, by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods known in the art.

Methods for preparing and using probes and primers are described, for example, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989), Ausubel *et al.* (In
15 *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Intersciences, 1998), and Innis *et al.* (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). One of ordinary skill in the art will appreciate that the
20 specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 30 consecutive nucleotides of ABCA13 encoding nucleotide will anneal to a target sequence, such as an ABCA13 gene homolog from the gene family contained within a human genomic DNA library, with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers can be selected that comprise at least 17, 20, 23,
25 25, 30, 35, 40, 45, 50 or more consecutive nucleotides of ABCA13 nucleotide sequences.

The disclosure thus includes isolated nucleic acid molecules that comprise specified lengths of the disclosed ABCA13 cDNA sequences. Such molecules can comprise at least 17, 20, 23, 25, 30, 35, 40, 45 or 50 or more consecutive nucleotides of these sequences, and can be obtained from any region of the disclosed sequences. By way of example, the ABCA13 cDNA sequences can be
30 apportioned into halves, thirds or quarters based on sequence length, and the isolated nucleic acid molecules can be derived from the first or second halves of the molecules, from any of the three thirds or any of the four quarters. By way of example, the human ABCA13 cDNA, ORF, coding sequence and gene sequences can be apportioned into about halves, thirds or quarters based on sequence length, and the isolated nucleic acid molecules (*e.g.*, oligonucleotides) can be derived from
35 the first or second halves of the molecules, from any of the three thirds, or any of the four quarters. The cDNA also could be divided into smaller regions, *e.g.* about eighths, sixteenths, twentieths, fiftieths and so forth, with similar effect.

Another mode of division, provided by way of example, is to divide an ABCA13 encoding sequence based on the regions of the sequence that are relatively more or less homologous to other

members of the ABC transporter family. Thus, nucleic acid molecules, for instance to be used as hybridization probe molecules, may be selected from the region encoding the N-terminal nucleotide binding domain (NBD1) region (*e.g.*, about residues 3868-4050, or a fragment thereof), or from the region encoding the C-terminal nucleotide binding domain (NBD2) (*e.g.*, about residues 4747-4932 or a fragment thereof), or the or the region encoding the N-terminal transmembrane domain (TM1) region (*e.g.*, about residues 22-3774, or a fragment thereof) or the region encoding the C-terminal transmembrane domain (TM2) region (*e.g.*, about residues 4225-4675) of the amino acid sequence encoding the human ABCA13 cDNA shown in SEQ ID NO: 1 (SEQ ID NO: 2). More specifically, the residues corresponding to the hydrophobic regions that compose the transmembrane domains of ABCA13 are residues 20-42, 3571-3793, 3414-3636, 3646-3668, 3677-3697, 3751-3773, 4226-4247, 4456-4473, 4508-4530, 4535-4556, 4571-4532, and 4654-4675 of SEQ ID NO: 2 (see Figure 4). It should be noted however, that these regions are predicted using computer programs; thus, prediction using a different computer program may generate slightly different regions than those disclosed herein.

Another mode of division is to select the 5' (upstream) and/or 3' (downstream) region associated with an ABCA13 gene, or to select an intron or portion thereof.

Purified: In a more pure form than is found in nature. The term purified does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified protein preparation is one in which the protein referred to is more pure than the protein in its natural environment within a cell.

The term **substantially purified** as used herein refers to a molecule (*e.g.*, a nucleic acid, polypeptide, oligonucleotide, etc.) that is substantially free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated. In one embodiment, the molecule is a polypeptide that is at least 50% free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated. In another embodiment, the polypeptide is at least at least 80% free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated. In yet other embodiments, the polypeptide is at least 90% or at least 95% free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated.

Recombinant: A nucleic acid that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques.

Sequence identity: The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs or orthologs of the ABCA13 protein, and the corresponding cDNA sequence, will possess a relatively high degree of sequence identity when aligned using standard methods. This homology will be more significant when the orthologous proteins or cDNAs are derived from species that are

more closely related (e.g., human and chimpanzee sequences), compared to species more distantly related (e.g., human and *C. elegans* sequences).

Typically, ABCA13 orthologs are at least 50% identical at the nucleotide level and at least 50% identical at the amino acid level when comparing ABCA13 to an orthologous ABCA13.

5 Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman *J. Mol. Biol.* 147(1): 195-197, 1981; Needleman and Wunsch *J. Mol. Biol.* 48: 443-453, 1970; Pearson and Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444-2448, 1988; Higgins and Sharp *Gene*, 73: 237-244, 1988; Higgins and Sharp *CABIOS* 5: 151-153, 1989; Corpet *et al. Nuc. Acids Res.* 16, 10881-10890, 1988; Huang *et al. Computer Appls. in the Biosciences* 8, 155-165, 1992; and Pearson *et al. Meth. Mol. Bio.* 24: 307-331, 10 1994. Furthermore, Altschul *et al. (J. Mol. Biol.* 215:403-410, 1990) present a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al. J. Mol. Biol.* 215: 403-410, 1990) is available from several sources, including the National Center for Biotechnology 15 Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. The Search Tool can be accessed at the NCBI website, together with a description of how to determine sequence identity using this program.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence- 20 dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5° C to 20° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence remains hybridized to a perfectly matched probe or complementary strand. Conditions for nucleic acid hybridization and calculation of stringencies can be 25 found in Sambrook *et al. (1989) Molecular Cloning: A Laboratory Manual*, CSHL, New York and Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* Part I, Chapter 2, Elsevier, New York. Nucleic acid molecules that hybridize under stringent conditions to a human ABCA13 gene sequence will typically hybridize to a probe based on either an entire human ABCA13 gene or selected portions of the gene under wash conditions of 2x 30 SSC at 50° C. A more detailed discussion of hybridization conditions is presented herein.

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

35 **Specific binding agent:** An agent that binds substantially only to a defined target. Thus an ABCA13 protein-specific binding agent binds substantially only an ABCA13 protein. As used herein, the phrase ABCA13 **protein-specific binding agent** includes anti-ABCA13 protein antibodies and other agents (such as soluble receptors) that bind substantially only to an ABCA13 protein, such as the ABCA13 proteins of the disclosure, or conservative variants thereof.

Anti-ABCA13 protein antibodies can be produced using standard procedures described in a number of texts, including Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). The determination that a particular agent binds substantially only to the ABCA13 protein can readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane, *Antibodies, A Laboratory Manual*, CSHL, New York, 1988). Western blotting can be used to determine that a given ABCA13 protein binding agent, such as an anti-ABCA13 protein monoclonal antibody, binds substantially only to the ABCA13 protein.

A phosphor-specific binding agent specifically binds to a peptide containing a phosphorylated residue.

Shorter fragments of antibodies can also serve as specific binding agents. For instance, Fabs, Fvs, and single-chain Fvs (SCFvs) that bind to ABCA13 would be ABCA13-specific binding agents. These antibody fragments are defined as follows: (1) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain; (2) Fab', the fragment of an antibody molecule obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule; (3) (Fab')₂, the fragment of the antibody obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; (4) F(ab')₂, a dimer of two Fab' fragments held together by two disulfide bonds; (5) Fv, a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains; and (6) single chain antibody (SCA), a genetically engineered molecule containing the variable region of the light chain, the variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Methods of making these fragments are routine.

Subject: Living multi-cellular vertebrate organisms, a category that includes both human and non-human mammals.

Target sequence: "Target sequence" is a portion of ssDNA, dsDNA or RNA that, upon hybridization to a therapeutically effective oligonucleotide or oligonucleotide analog, results in the inhibition of expression. For example, hybridization of therapeutically effective oligonucleotide to an ABCA13 target sequence results in inhibition of ABCA13 expression. Either an antisense or a sense molecule can be used to target a portion of dsDNA, as both will interfere with the expression of that portion of the dsDNA. The antisense molecule can bind to the plus strand, and the sense molecule can bind to the minus strand. Thus, target sequences can be ssDNA, dsDNA, and RNA.

Transfected: A process by which a nucleic acid molecule is introduced into cell, for instance by molecular biology techniques, resulting in a **transfected cell**. As used herein, the term transfection encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transduction with viral vectors, transfection with plasmid vectors, and introduction of DNA by electroporation, lipofection, and particle gun acceleration.

-18-

Treating a disease: Includes inhibiting or preventing the partial or full development or progression of a disease, for example in a person who is known to have a predisposition to a disease. An example of a person with a known predisposition is someone with a history of diabetes in his or her family, or who has been exposed to factors that predispose the subject to a condition, such as lupus or rheumatoid arthritis. Furthermore, treating a disease refers to a therapeutic intervention that ameliorates at least one sign or symptom of a disease or pathological condition, or interferes with a pathophysiological process, after the disease or pathological condition has begun to develop.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transfected host cell. Recombinant DNA vectors are vectors having recombinant DNA. A vector can include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector can also include one or more selectable marker genes and other genetic elements known in the art. Viral vectors are recombinant DNA vectors having at least some nucleic acid sequences derived from one or more viruses.

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The singular terms "a," "an," and "the" include plural unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Thus "A or B" means "A" or "B" or "A and B." "Comprises" means "includes." It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

III. Description of Several Specific Embodiments

Provided herein are amino acid sequences encoding seven isoforms of human ABCA13, as SEQ ID NOs: 2, 19, 21, 23, 25, 27, and 29. Specific embodiments include purified proteins having high identity (*e.g.*, 85%, 90%, 95%) to amino acid residues 2278-4942 of ABCA13 Isoform 1 (SEQ ID NO: 2). Specific embodiments disclose proteins having high identity to a conserved region shared by all isoform proteins, residues 2278-4942 of SEQ ID NO: 2. The provided ABCA13 proteins have ABCA13 protein biological activity, for instance in that they can complement an *ABCA13* null phenotype. Specific embodiments include ABCA13 proteins having biological activity as extra- and/or intracellular membrane transport proteins that can complement *ABCA13* null phenotypes by ameliorating the respective transport deficiencies.

Also provided are nucleic acid molecules encoding the seven Isoform proteins, having nucleotide sequences as set forth as SEQ ID NOs: 1, 18, 20, 22, 24, 26 and 28. Recombinant polynucleotides encoding these sequences, including polynucleotides encoding the coding regions of SEQ ID NOs: 1, 18, 20, 22, 24, 26 and 28 are provided herein.

5 Methods to diagnose and detect defects or alterations in ABCA13 expression are provided, as are methods for screening for specific binding agents of ABCA13.

In specific embodiments, the methods are used to detect hypercholesterolemia, drug resistance, retinal degeneration, or neurological disease. Other methods are used to detect chemotherapy resistant cells, or employ primers having sequences identical to at least 10 contiguous
10 nucleotides of disclosed sequences, including SEQ ID NOs: 1, 18, 20, 22, 24, 26 or 28.

Antibodies specific to ABCA13, and the use of such antibodies (*e.g.*, in Western blot to ELISA assays) are disclosed.

Kits for using the ABCA13 Isoform proteins are disclosed, including kits for screening defects in ABCA13 biological activity, and kits used to assay particular defects in individuals with
15 defective ABCA13, including altered extra- or intracellular transport.

IV. ABCA13 Nucleic Acids and Proteins

The *ABCA13* gene encodes a protein proposed to be involved in the energy-dependent transport of one or a variety of substances, for example ions, metals, amino acids, lipids, peptides,
20 sugars, vitamins and steroid hormones. In addition, *ABCA13* is located in a locus genetically linked to the Shwachmann-Diamond syndrome, a disorder of the pancreas, and a locus involved in T cell tumorigenesis. Thus, *ABCA13* is a candidate gene for these phenotypes. Furthermore, the gene is also highest expressed in certain tumor cell lines from leukemia, prostate and CNS, suggesting that it may play a role in these cancers.

25 *ABCA13* cDNA sequences are used to pinpoint the location of *ABCA13* in the human genome. *ABCA13* was localized by amplification of monochrome hybrids and radiation hybrids using the Polymerase Chain Reaction (PCR) (see Morten *et al.*, *Hum. Genet.* 88(2): 200-203, 1991). Subsequently, the position of *ABCA13* on the draft human map was determined using chromosomal assignment using somatic cell hybrids (CASH) (see Ryu *et al.*, *Mol Cells* 10(5): 598-600, 2000).

30 RT-PCR analysis shows that ABCA13 is expressed in brain, lung, skeletal muscle and ovary. In a study of sixty tumor cell lines, the expression of ABCA13 was detected in SR leukemia, SNB-19 CNS tumor, and DU-145 prostate tumor cell lines. The *ABCA13* gene maps to chromosome 7p12.3, a region that contains an inherited disorder affecting the pancreas as well as a locus involved in tumorigenesis. *ABCA13* is therefore a positional candidate for these pathologies.

35 The *ABCA13* gene is not present in zebrafish, indicated that it is a recently evolved gene specific to mammals.

This disclosure provides ABCA13 transporter proteins and isoforms and variants thereof, and nucleic acid molecules encoding these proteins, including cDNA sequences. In specific embodiments, these sequences are used for ameliorating, treating, detecting, prognosing, and

-20-

diagnosing diseases and conditions believed to be associated with altered *ABCA13* expression (based on homology to other ABCA family members), such as hypercholesterolemia, drug resistance, retinal degeneration, or neurological disease. In one embodiment, an *ABCA13* nucleic acid has altered expression (*e.g.*, increased or decreased expression, such as altered transcription of *ABCA13* mRNA, a mutated or deleted expression product, improper subcellular localization of a nucleic acid, etc.) as compared to a control nucleic acid (*e.g.*, a nucleic acid amplified, using positive control sequences, from a subject not suffering from the biological condition). In some embodiments, expression of an *ABCA13* nucleic acid is more than 50%, more than 75%, more than 100%, more than 200%, or more than 300% different when compared to a suitable control. Suitable controls include a known control, a known sample, or a standard value as assigned by one of ordinary skill in the art as a suitable standard value.

It will be apparent to one skilled in the art that either the cDNAs disclosed herein or sequences derived from these clones may be utilized in applications, including but not limited to, studies of the expression of the *ABCA13* gene, studies of the function of the *ABCA13* protein, the generation of antibodies to the *ABCA13* protein, diagnosis and therapy of *ABCA13* deleted or mutated in subjects to prevent or treat the defects in cell and tissue development, such as transporter deficiency. Descriptions of applications describing the use of *ABCA13* cDNA, or fragments thereof, are therefore intended to comprehend the use of the *ABCA13* nucleic acids and conservative variants thereof.

A nucleic acid molecule encoding full-length *ABCA13* (Isoform 1) is shown in SEQ ID NO: 1; the nucleic acid molecule encodes a protein of 5058 amino acids in length (SEQ ID NO: 2). Splice variants of the *ABCA13* gene are also provided (*e.g.*, SEQ ID NOs: 18, 20, 22, 24, 26, and 28); these encode various isoforms of *ABCA13* (SEQ ID NOs: 19, 21, 23, 25, 27, and 29, respectively).

The full-length *ABCA13* gene is encoded by 62 exons and spans over 450 kb, making *ABCA13* the largest ABC gene described to date, including the largest number of exons. Exon and intron positions and accompanying data for full-length *ABCA13* are summarized in Table 1.

In SEQ ID NO: 2, amino acids 20-42, 3571-3793, 3414-3636, 3646-3668, 3677-3697, 3751-3773, 4226-4247, 4456-4473, 4508-4530, 4535-4556, 4571-4532, and 4654-4675 of SEQ ID NO: 2 are predicted transmembrane domains (see Figure 5). Amino acids 9-15 compose a conserved ABCA family motif. Amino acids 3868-4050 and 4747-4932 are predicted nucleotide binding fold domains. Amino acids 708-2911 compose the large extracellular domain exon. Amino acids 3707-3729 compose a hydrophobic domain, and amino acids 4152-4156 and 5026-5030 are potential phosphorylation sites.

The following sequences homologous to *ABCA13* Isoform 1 were extracted from GenBank (version 126 and 127): AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; AC073927 (g13992793), *Homo sapiens* chromosome 7 BAC, clone RP11-604B16, complete sequence, 185676 base pairs; and AC091770

-21-

(g15638906), *Homo sapiens* chromosome 7, BAC clone RP11-655M5, complete sequence, 93543 base pairs.

Table 1. ABCA13 Isoform 1 Positional Data

GenBank Genomic sequence	Brin d'ADN codant	Exon or Intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21699	313	464	152	21700	23470	1771
AC073424	+	4	23471	23489	465	493	29	23500	28455	5955
AC073424	+	5	28456	29619	494	657	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46898	47705	808
AC073424	+	11	47706	47806	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	73980	2692
AC073424	+	16	73981	78759	2146	6924	4779	78760	80287	1528
AC073424	+	17	80288	82114	6925	8751	1827	82115	83536	1422
AC073424	+	18	83537	83646	8752	8861	110	83647	80153	6507
AC073424	+	19	90154	90272	8862	8980	119	90273	97893	7621
AC073424	+	20	97894	98057	8981	9144	164	98058	99425	1368
AC073424	+	21	99426	99505	9145	9224	80	99506	100559	1054
AC073424	+	22	100560	100681	9225	9346	122	100682	112140	11459
AC073424	+	23	112141	112335	9347	9541	195	112336	115260	2925
AC073424	+	24	115261	115425	9542	9706	165	115426	116425	1000
AC073424	+	25	116426	116603	9707	9884	178	116604	119350	2747
AC073424	+	26	119351	119490	9885	10024	140	119491	137615	18125
AC073424	+	27	137616	137729	10025	10138	114	137730	140558	2829
AC073424	+	28	140559	140649	10139	10229	91	140650	152836	12187
AC073424	+	29	152837	153013	10230	10406	177	153014	154374	1361
AC073424	+	30	154375	154681	10407	10713	307	154682	169987	15306
AC073424	+	31	169988	170102	10714	10828	115	170103	174361	4259
AC073424	+	32	174362	174691	10829	11158	330	174692	176540	1849
AC073424	+	33	176541	176610	11159	11228	70	176611	178634	2024
AC073424	+	34	178635	178766	11229	11360	132	178767	190015	11249
AC073424	+	35	190016	190153	11361	11498	138	190154		
AC095039	+	36	2093	2273	11499	11679	181	2274	4973	2700
AC095039	+	37	4974	5192	11680	11898	219	5193	16735	11543
AC095039	+	38	16736	16932	11899	12095	197	16933	23572	6640
AC095039	+	39	23573	23730	12096	12253	158	23731	25405	1675
AC095039	+	40	25406	25636	12254	12484	231	25637	40818	15182
AC095039	+	41	40819	40924	12485	12590	106	40925	68089	27165
AC095039	+	42	68090	68339	12591	12840	250	68340		
AC073927	+	43	11593	11682	12841	12930	90	11683	16166	4484
AC073927	+	44	16167	16236	12931	13000	70	16237	25672	9436
AC073927	+	45	25673	25791	13001	13119	119	25792	27712	1921
AC073927	+	46	27713	27800	13120	13207	88	27801	33872	6072
AC073927	+	47	33873	33981	13208	13316	109	33982	50971	16990
AC073927	+	48	50972	51026	13317	13371	55	51027	52507	1481
AC073927	+	49	52508	52685	13372	13549	178	52686	55719	3034
AC073927	+	50	55720	55835	13550	13665	116	55836	61360	5525
AC073927	+	51	61361	61517	13666	13822	157	61518	64676	3159
AC073927	+	52	64677	64930	13823	14076	254	64931	68883	3953
AC073927	+	53	68884	69076	14077	14269	193	69077	72871	3795
AC073927	+	54	72872	72981	14270	14379	110	72982	124859	51678
AC073927	+	55	124860	125010	14380	14530	151	125011	131789	6779
AC073927	+	56	131790	131824	14531	14665	135	131825	139345	7421
AC073927	+	57	139346	139449	14666	14769	104	139450	159920	20471
AC073927	+	58	159921	160013	14770	14862	93	160014		
AC091770	-	59	89296	89191	14863	14968	106	89190	87969	1223
AC091770	-	60	87967	87830	14969	15103	135	87829	87167	663
AC091770	-	61	87167	85059	15107	17188	2080	85058		

5

Isoform 2A is a protein encoded by a nucleotide sequence of 12,498 base pairs (SEQ ID NO: 18). The protein sequence of Isoform 2A (SEQ ID NO: 19) is 2760 amino acids in length. Isoform 2B is encoded by a nucleotide sequence of 12,498 base pairs (SEQ ID NO: 20) and its predicted amino acid sequence (SEQ ID NO: 21) is 717 amino acids in length. The following sequences homologous to Isoforms 2A and 2B were extracted from GenBank: AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; AC073927 (g13992793), *Homo sapiens* chromosome 7, BAC clone RP11-604B16, complete sequence, 185676 base pairs; and AC091770 (g15638906), *Homo sapiens* chromosome 7,

10

-22-

BAC clone RP11-655M5, complete sequence, 93543 base pairs. Relevant exon and intron positions and accompanying data are summarized in Table 2.

Table 2. ABCA13 Isoform 2A Positional Data

GenBank Genomic sequence	Brin d'ADN codant	Exon or intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21699	313	464	152	21700	23470	1771
AC073424	+	4	23471	23499	465	493	29	23500	29455	5956
AC073424	+	5	29456	29619	494	657	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46898	47705	808
AC073424	+	11	47706	47806	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	78691	7403
AC073424	+	16	78692	78759	2146	2213	68	78760	80287	1528
AC073424	+	17	80288	82114	2214	4040	1827	82115	83536	1422
AC073424	+	18	83537	83646	4041	4150	110	83647	90153	6507
AC073424	+	19	90154	90272	4151	4269	119	90273	97893	7621
AC073424	+	20	97894	98057	4270	4433	164	98058	99425	1368
AC073424	+	21	99426	99505	4434	4513	80	99506	100559	1054
AC073424	+	22	100560	100681	4514	4635	122	100682	112140	11459
AC073424	+	23	112141	112335	4636	4830	195	112336	115260	2925
AC073424	+	24	115261	115425	4831	4995	165	115426	116425	1000
AC073424	+	25	116426	116603	4996	5173	178	116604	119350	2747
AC073424	+	26	119351	119490	5174	5313	140	119491	137615	18125
AC073424	+	27	137616	137729	5314	5427	114	137730	140558	2829
AC073424	+	28	140559	140649	5428	5518	91	140650	152836	12187
AC073424	+	29	152837	153013	5519	5695	177	153014	154374	1361
AC073424	+	30	154375	154681	5696	6002	307	154682	169987	15306
AC073424	+	31	169988	170102	6003	6117	115	170103	174361	4259
AC073424	+	32	174362	174691	6118	6447	330	174692	176540	1849
AC073424	+	33	176541	176610	6448	6517	70	176511	178634	2024
AC073424	+	34	178635	178766	6518	6649	132	178767	180015	11249
AC073424	+	35	190016	190153	6650	6787	138	190154		
AC095039	+	36	2093	2273	6788	6968	181	2274	4973	2700
AC095039	+	37	4974	5192	6969	7187	219	5193	16735	11543
AC095039	+	38	16736	16932	7188	7384	197	16933	23572	6640
AC095039	+	39	23573	23730	7385	7542	158	23731	25405	1675
AC095039	+	40	25406	25636	7543	7773	231	25637	40818	15182
AC095039	+	41	40819	40924	7774	7879	106	40925	68089	27165
AC095039	+	42	68090	68339	7880	8129	250	68340		
AC073927	+	43	11593	11682	8130	8219	90	11683	16166	4484
AC073927	+	44	16167	16236	8220	8289	70	16237	25672	9436
AC073927	+	45	25673	25791	8290	8408	119	25792	27712	1921
AC073927	+	46	27713	27800	8409	8496	88	27801	33872	6072
AC073927	+	47	33873	33981	8497	8605	109	33982	50971	16990
AC073927	+	48	50972	51026	8606	8660	55	51027	52507	1481
AC073927	+	49	52508	52685	8661	8838	178	52686	55719	3034
AC073927	+	50	55720	55835	8839	8954	116	55836	61360	5525
AC073927	+	51	61361	61517	8955	9111	157	61518	64676	3159
AC073927	+	52	64677	64930	9112	9365	254	64931	68883	3953
AC073927	+	53	68884	69076	9366	9558	193	69077	72871	3795
AC073927	+	54	72872	72981	9559	9668	110	72982	124659	51878
AC073927	+	55	124860	125010	9669	9819	151	125011	131789	6779
AC073927	+	56	131790	131924	9820	9954	135	131925	139345	7421
AC073927	+	57	139346	139449	9955	10058	104	139450	159920	20471
AC073927	+	58	159921	160013	10059	10151	93	160014		
AC091770	-	59	89296	89191	10152	10257	106	89190	87968	1223
AC091770	-	60	87967	87830	10258	10395	138	87829	87167	663
AC091770	-	61	87167	85089	10396	12475	2080	85088		

5

Isoform 3 (SEQ ID NO: 23) is 4958 amino acids in length and encoded by a nucleotide sequence of 15,970 base pairs (SEQ ID NO: 22). The following sequences homologous to Isoform 3 were extracted from GenBank: AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; and AC073927 (g13992793), *Homo sapiens* chromosome 7, BAC, clone RP11-604B16, complete sequence, 185676 base pairs. Relevant exon and intron positions and accompanying data are summarized in Table 3.

10

Table 3. ABCA13 Isoform 3 Positional Data

GenBank Genomic sequence	Brin d'ADN codant	Exon or intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21699	313	464	152	21700	23470	1771
AC073424	+	4	23471	23499	465	493	29	23500	29455	5956
AC073424	+	5	29456	29619	494	657	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46898	47705	808
AC073424	+	11	47706	47805	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	73980	2692
AC073424	+	16	73981	78759	2146	6924	4779	78760	80287	1526
AC073424	+	17	80288	82114	6925	8751	1827	82115	83536	1422
AC073424	+	18	83537	83646	8752	8861	110	83647	90153	6507
AC073424	+	19	90154	90272	8862	8980	119	90273	97893	7621
AC073424	+	20	97894	98057	8981	9144	164	98058	99425	1368
AC073424	+	21	99426	99505	9145	9224	80	99506	100569	1054
AC073424	+	22	100560	100681	9225	9346	122	100682	112140	11469
AC073424	+	23	112141	112335	9347	9541	195	112336	115260	2925
AC073424	+	24	115261	115425	9542	9706	165	115426	116425	1000
AC073424	+	25	116426	116603	9707	9884	178	116604	119350	2747
AC073424	+	26	119351	119490	9885	10024	140	119491	137615	18125
AC073424	+	27	137616	137729	10025	10138	114	137730	140558	2829
AC073424	+	28	140559	140649	10139	10229	91	140650	152836	12187
AC073424	+	29	152837	153013	10230	10406	177	153014	154374	1361
AC073424	+	30	154375	154681	10407	10713	307	154682	169987	16306
AC073424	+	31	169988	170102	10714	10828	115	170103	174361	4259
AC073424	+	32	174362	174691	10829	11158	330	174692	176540	1849
AC073424	+	33	176541	176610	11159	11228	70	176611	178634	2024
AC073424	+	34	178635	178766	11229	11360	132	178767	190015	11249
AC073424	+	35	190016	190153	11361	11498	138	190154		
AC095039	+	36	2093	2273	11499	11679	181	2274	4973	2700
AC095039	+	37	4974	5192	11680	11898	219	5193	16735	11543
AC095039	+	38	16736	16932	11899	12095	197	16933	23572	6640
AC095039	+	39	23573	23730	12096	12253	158	23731	25405	1675
AC095039	+	40	25406	25636	12254	12484	231	25637	40818	15182
AC095039	+	41	40819	40924	12485	12590	106	40925	68089	27165
AC095039	+	42	68090	68339	12591	12840	250	68340		
AC073927	+	43	11593	11682	12841	12930	90	11683	16166	4484
AC073927	+	44	16167	16236	12931	13000	70	16237	25672	9436
AC073927	+	45	25673	25791	13001	13119	119	25792	27712	1921
AC073927	+	46	27713	27800	13120	13207	88	27801	33872	6072
AC073927	+	47	33873	33981	13208	13316	109	33982	50971	16990
AC073927	+	48	50972	51026	13317	13371	55	51027	52507	1481
AC073927	+	49	52508	52685	13372	13549	178	52686	55719	3034
AC073927	+	50	55720	55835	13550	13665	116	55836	61360	5525
AC073927	+	51	61361	61517	13666	13822	157	61518	64676	3159
AC073927	+	52	64677	64930	13823	14076	254	64931	68883	3953
AC073927	+	53	68884	69076	14077	14269	193	69077	72871	3795
AC073927	+	54	72872	72981	14270	14379	110	72982	124859	51878
AC073927	+	55	124860	125010	14380	14530	151	125011	131789	6779
AC073927	+	56	131790	131924	14531	14665	135	131925	139345	7421
AC073927	+	57	139346	139449	14666	14769	104	139450	159920	20471
AC073927	+	58	159921	161095	14770	15946	1177	161096		

Isoform 4 of the *ABCA13* gene is 11259 base pairs long (SEQ ID NO: 24) with an open reading frame of 8079 base pairs, corresponding to 2660 amino acids (SEQ ID NO: 25). The following sequences homologous to Isoform 4 were extracted from GenBank: AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; and AC073927 (g13992793), *Homo sapiens* chromosome 7, BAC clone RP11-604B16, complete sequence, 185676 base pairs. Relevant exon and intron positions and accompanying data are summarized in Table 4.

Table 4. ABCA13 Isoform 4 Positional Data

GenBank Genomic sequence	Brin d'ADN codant	Exon or intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21699	313	464	152	21700	23470	1771
AC073424	+	4	23471	23499	465	493	29	23500	29455	5956
AC073424	+	5	29456	29619	494	667	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46898	47705	808
AC073424	+	11	47706	47806	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	78691	7403
AC073424	+	16	78692	78759	2146	2213	68	78760	80287	1528
AC073424	+	17	80288	82114	2214	4040	1827	82115	83536	1422
AC073424	+	18	83537	83646	4041	4150	110	83647	90153	6507
AC073424	+	19	90154	90272	4151	4269	119	90273	97893	7621
AC073424	+	20	97894	98057	4270	4433	164	98058	99425	1368
AC073424	+	21	99426	99505	4434	4513	80	99506	100559	1054
AC073424	+	22	100560	100681	4514	4635	122	100682	112140	11459
AC073424	+	23	112141	112335	4636	4830	195	112336	115260	2925
AC073424	+	24	115261	115425	4831	4995	165	115426	116425	1000
AC073424	+	25	116426	116603	4996	5173	178	116604	119350	2747
AC073424	+	26	119351	119490	5174	5313	140	119491	137615	18125
AC073424	+	27	137616	137729	5314	5427	114	137730	140558	2829
AC073424	+	28	140559	140649	5428	5518	91	140650	152836	12187
AC073424	+	29	152837	153013	5519	5695	177	153014	154374	1361
AC073424	+	30	154375	154681	5696	6002	307	154682	169987	15306
AC073424	+	31	169988	170102	6003	6117	115	170103	174361	4259
AC073424	+	32	174362	174691	6118	6447	330	174692	176540	1849
AC073424	+	33	176541	176610	6448	6517	70	176611	178634	2024
AC073424	+	34	178635	178766	6518	6649	132	178767	190015	11249
AC073424	+	35	190016	190153	6650	6787	138	190154		
AC095039	+	36	2093	2273	6788	6968	181	2274	4973	2700
AC095039	+	37	4974	5192	6969	7187	219	5193	16735	11543
AC095039	+	38	16736	16932	7188	7384	197	16933	23572	6640
AC095039	+	39	23573	23730	7385	7542	158	23731	25405	1675
AC095039	+	40	25406	25636	7543	7773	231	25637	40818	15182
AC095039	+	41	40819	40924	7774	7879	106	40925	68089	27165
AC095039	+	42	68090	68339	7880	8129	250	68340		
AC073927	+	43	11593	11682	8130	8219	90	11683	16166	4484
AC073927	+	44	16167	16236	8220	8289	70	16237	25672	9436
AC073927	+	45	25673	25791	8290	8408	119	25792	27712	1921
AC073927	+	46	27713	27800	8409	8496	88	27801	33872	6072
AC073927	+	47	33873	33981	8497	8605	109	33982	50971	16990
AC073927	+	48	50972	51026	8606	8660	55	51027	52507	1481
AC073927	+	49	52508	52685	8661	8838	178	52686	55719	3034
AC073927	+	50	55720	55835	8839	8954	116	55836	61360	5525
AC073927	+	51	61361	61517	8955	9111	157	61518	64676	3159
AC073927	+	52	64677	64930	9112	9365	254	64931	68883	3953
AC073927	+	53	68884	69076	9366	9558	193	69077	72871	3795
AC073927	+	54	72872	72981	9559	9668	110	72982	124659	51878
AC073927	+	55	124860	125010	9669	9819	151	125011	131789	6779
AC073927	+	56	131790	131924	9820	9954	135	131925	139345	7421
AC073927	+	57	139346	139449	9955	10058	104	139450	159920	20471
AC073927	+	58	159921	161095	10059	11235	1177	161096		

Isoform 5 of the *ABCA13* gene is 15779 base pairs long (SEQ ID NO: 26) with an open reading frame of 14877 base pairs, corresponding to 4952 amino acids (SEQ ID NO: 27). The following sequences homologous to Isoform 5 were extracted from GenBank: AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; and AC073927 (g13992793), *Homo sapiens* chromosome 7, BAC clone RP11-604B16, complete sequence, 185676 base pairs. Relevant exon and intron positions and accompanying data are summarized in Table 5.

Table 5. ABCA13 Isoform 5 Positional Data

Nom de la séquence génomique de référence	Brin d'ADN codant	Exon or intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21899	313	464	152	21700	23470	1771
AC073424	+	4	23471	23499	465	493	29	23500	29455	5956
AC073424	+	5	29456	29619	494	657	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46898	47705	808
AC073424	+	11	47706	47806	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	73980	2692
AC073424	+	16	73981	78759	2146	6924	4779	78760	80287	1528
AC073424	+	17	80288	82114	6925	8751	1827	82115	83536	1422
AC073424	+	18	83537	83646	8752	8861	110	83647	90153	6607
AC073424	+	19	90154	90272	8862	8980	119	90273	97893	7621
AC073424	+	20	97894	98057	8981	9144	164	98058	99425	1368
AC073424	+	21	99426	99505	9145	9224	80	99506	100559	1054
AC073424	+	22	100560	100681	9225	9346	122	100682	112140	11459
AC073424	+	23	112141	112335	9347	9541	195	112336	115260	2925
AC073424	+	24	115261	115425	9542	9706	165	115426	116425	1000
AC073424	+	25	116426	116603	9707	9884	178	116604	119350	2747
AC073424	+	26	119351	119490	9885	10024	140	119491	137615	18125
AC073424	+	27	137616	137729	10025	10138	114	137730	140558	2829
AC073424	+	28	140559	140649	10139	10229	91	140650	152836	12187
AC073424	+	29	152837	153013	10230	10406	177	153014	154374	1361
AC073424	+	30	154375	154681	10407	10713	307	154682	169987	15306
AC073424	+	31	169988	170102	10714	10828	115	170103	174361	4259
AC073424	+	32	174362	174691	10829	11158	330	174692	176540	1849
AC073424	+	33	176541	176610	11159	11228	70	176611	178634	2024
AC073424	+	34	178635	178766	11229	11360	132	178767	190015	11249
AC073424	+	35	190016	190153	11361	11498	138	190154		
AC095039	+	36	2093	2273	11499	11679	181	2274	4973	2700
AC095039	+	37	4974	5192	11680	11898	219	5193	16735	11543
AC095039	+	38	16736	16932	11899	12095	197	16933	23572	6640
AC095039	+	39	23573	23730	12096	12253	158	23731	25405	1675
AC095039	+	40	25406	25636	12254	12484	231	25637	40818	15182
AC095039	+	41	40819	40924	12485	12590	106	40925	68089	27165
AC095039	+	42	68090	68339	12591	12840	250	68340		
AC073927	+	43	11593	11682	12841	12930	90	11683	16166	4484
AC073927	+	44	16167	16236	12931	13000	70	16237	25672	9436
AC073927	+	45	25673	25791	13001	13119	119	25792	27712	1921
AC073927	+	46	27713	27800	13120	13207	88	27801	33872	6072
AC073927	+	47	33873	33981	13208	13316	109	33982	50971	16990
AC073927	+	48	50972	51026	13317	13371	55	51027	52507	1481
AC073927	+	49	52508	52665	13372	13549	178	52666	55719	3034
AC073927	+	50	55720	55835	13550	13665	116	55836	61360	5525
AC073927	+	51	61361	61517	13666	13822	157	61518	64676	3159
AC073927	+	52	64677	64930	13823	14076	254	64931	68883	3953
AC073927	+	53	68884	69076	14077	14269	193	69077	72871	3795
AC073927	+	54	72872	72981	14270	14379	110	72982	124859	51878
AC073927	+	55	124860	125010	14380	14530	151	125011	131789	6779
AC073927	+	56	131790	131924	14531	14665	135	131925	139345	7421
AC073927	+	57	139346	139449	14666	14769	104	139450	159920	20471
AC073927	+	58	159921	160013	14770	14862	93	160014	171237	11224
AC073927	+	59	171238	172131	14863	15756	894	172132		

Isoform 6 of the *ABCA13* gene is 11068 base pairs long (SEQ ID NO: 28) with an open reading frame of 8061 base pairs, corresponding to 2654 amino acids (SEQ ID NO: 29). The following sequences homologous to Isoform 6 were extracted from GenBank: AC073424 (g15321569), *Homo sapiens* chromosome 7, clone RP11-653O17, complete sequence, 191141 base pairs; AC095039 (g16974283), *Homo sapiens* chromosome 7, clone RP11-12G8, complete sequence, 70416 base pairs; and AC073927 (g13992793), *Homo sapiens* chromosome 7, BAC clone RP11-604B16, complete sequence, 185676 base pairs. Relevant exon and intron positions and accompanying data are summarized in Table 6.

Table 6. ABCA13 Isoform 6 Positional Data

Nom de la séquence génomique de référence	Brin d'ADN codant	Exon or Intron number	Exon start in genomic fragment	Exon stop in genomic fragment	Exon start in mRNA	Exon stop in mRNA	Length of exon	Intron start in genomic fragment	Intron stop in genomic fragment	Length in intron
AC073424	+	2	440	563	189	312	124	564	21547	20984
AC073424	+	3	21548	21699	313	464	152	21700	23470	1771
AC073424	+	4	23471	23499	465	493	29	23500	29455	5956
AC073424	+	5	29456	29619	494	657	164	29620	32018	2399
AC073424	+	6	32019	32149	658	788	131	32150	36211	4062
AC073424	+	7	36212	36345	789	922	134	36346	41434	5089
AC073424	+	8	41435	41599	923	1087	165	41600	43060	1461
AC073424	+	9	43061	43260	1088	1287	200	43261	46769	3509
AC073424	+	10	46770	46897	1288	1415	128	46998	47705	808
AC073424	+	11	47706	47806	1416	1516	101	47807	48056	250
AC073424	+	12	48057	48224	1517	1684	168	48225	50432	2208
AC073424	+	13	50433	50638	1685	1890	206	50639	51405	767
AC073424	+	14	51406	51545	1891	2030	140	51546	71173	19628
AC073424	+	15	71174	71288	2031	2145	115	71289	78691	7403
AC073424	+	16	78692	78759	2146	2213	68	78760	80287	1528
AC073424	+	17	80288	82114	2214	4040	1827	82115	83536	1422
AC073424	+	18	83537	83646	4041	4150	110	83647	90153	6507
AC073424	+	19	90154	90272	4151	4269	119	90273	97893	7621
AC073424	+	20	97894	98057	4270	4433	164	98058	99425	1368
AC073424	+	21	99426	99505	4434	4513	80	99506	100559	1054
AC073424	+	22	100560	100681	4514	4635	122	100682	112140	11459
AC073424	+	23	112141	112335	4636	4830	195	112336	115260	2925
AC073424	+	24	115261	115425	4831	4995	165	115426	116425	1000
AC073424	+	25	116426	116603	4996	5173	178	116604	119350	2747
AC073424	+	26	119351	119490	5174	5313	140	119491	137615	18125
AC073424	+	27	137616	137729	5314	5427	114	137730	140558	2829
AC073424	+	28	140559	140649	5428	5518	91	140650	152836	12187
AC073424	+	29	152837	153013	5519	5695	177	153014	154374	1361
AC073424	+	30	154375	154681	5696	6002	307	154682	169987	15306
AC073424	+	31	169988	170102	6003	6117	115	170103	174361	4259
AC073424	+	32	174362	174691	6118	6447	330	174692	176540	1849
AC073424	+	33	176541	176610	6448	6517	70	176611	178634	2024
AC073424	+	34	178635	178766	6518	6649	132	178767	190015	11249
AC073424	+	35	190016	190153	6650	6787	138	190154		
AC095039	+	36	2093	2273	6788	6968	181	2274	4973	2700
AC095039	+	37	4974	5192	6969	7187	219	5193	16735	11543
AC095039	+	38	16736	16932	7188	7384	197	16933	23572	6640
AC095039	+	39	23573	23730	7385	7542	158	23731	25405	1675
AC095039	+	40	25406	25636	7543	7773	231	25637	40818	15182
AC095039	+	41	40819	40924	7774	7879	106	40925	68089	27165
AC095039	+	42	68090	68339	7880	8129	250	68340		
AC073927	+	43	11593	11682	8130	8219	90	11683	16166	4484
AC073927	+	44	16167	16236	8220	8289	70	16237	25672	9436
AC073927	+	45	25673	25791	8290	8408	119	25792	27712	1921
AC073927	+	46	27713	27800	8409	8496	88	27801	33872	6072
AC073927	+	47	33873	33981	8497	8605	109	33982	50971	16990
AC073927	+	48	50972	51026	8606	8660	55	51027	52507	1481
AC073927	+	49	52508	52685	8661	8838	178	52686	55719	3034
AC073927	+	50	55720	55835	8839	8954	116	55836	61360	5525
AC073927	+	51	61361	61517	8955	9111	157	61518	64676	3159
AC073927	+	52	64677	64930	9112	9365	254	64931	68883	3953
AC073927	+	53	68884	69076	9366	9558	193	69077	72871	3795
AC073927	+	54	72872	72981	9559	9668	110	72982	124859	51878
AC073927	+	55	124860	125010	9669	9819	151	125011	131789	6779
AC073927	+	56	131790	131924	9820	9954	135	131925	139345	7421
AC073927	+	57	139346	139449	9955	10058	104	139450	159920	20471
AC073927	+	58	159921	160013	10059	10151	93	160014	171237	11224
AC073927	+	59	171238	172131	10152	11045	894	172132		

V. Production of ABCA13 Nucleic Acids

With the provision herein of the sequence of the full length ABCA13 cDNA (SEQ ID NO: 1), as well as the sequences encoding several isoforms, *in vitro* nucleic acid amplification [such as polymerase chain reaction (PCR)] may be utilized as a simple method for producing ABCA13 encoding sequences. The following provides representative techniques for preparing cDNA in this manner.

Total RNA is extracted from human cells by any one of a variety of methods well known to those of ordinary skill in the art (Sambrook *et al.*, in *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989; and Ausubel *et al.*, in *Current Protocols in Molecular Biology*, Greene Publ. Assoc. and Wiley-Intersciences, 1992 provide descriptions of methods for RNA isolation).

-27-

In one embodiment, primary cells are obtained from normal tissues. In another embodiment cells are obtained from tissues from subjects exhibiting the effects of ABCA13 transporter deficiency. In yet another embodiment cell lines, derived from normal or transporter-deficient tissues, are used as a source of such RNA. The extracted RNA is then used, for example, as a template for performing
5 reverse transcription (RT)-PCR amplification of cDNA. Methods and conditions for RT-PCR are described in Kawasaki *et al.*, (In *PCR Protocols, A Guide to Methods and Applications*, Innis *et al.* (eds.), 21-27, Academic Press, Inc., San Diego, California, 1990).

The selection of amplification primers will be made according to the portion(s) of the cDNA that is to be amplified. In one embodiment, primers may be chosen to amplify a segment of a cDNA
10 or, in another embodiment, the entire cDNA molecule. Variations in amplification conditions may be required to accommodate primers and amplicons of differing lengths and composition; such considerations are well known in the art and are discussed for instance in Innis *et al.* (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990). By way of example, the coding portion of the human ABCA13 cDNA molecule (approximately 15,198 base
15 pairs) and the full-length human ABCA13 Isoform 1 cDNA, as shown in SEQ ID NO: 1, may be amplified using primers, such as SEQ ID NO: 3 through 15 of the disclosure.

Similarly, the primers set forth as SEQ ID NO: 3, 4, 6-8, and 10-15 can be used to amplify a human ABCA13 Isoform 2A cDNA, as shown in SEQ ID NO: 18; the primers set forth as SEQ ID NO: 3-5, 7-15, and 73 can be used to amplify a human ABCA13 Isoform 3 cDNA, as shown in SEQ
20 ID NO: 22; the primers set forth as SEQ ID NO: 3, 4, 7, 8, 10-15 and 73 can be used to amplify a human ABCA13 Isoform 4 cDNA, as shown in SEQ ID NO: 24; the primers set forth as SEQ ID NO: 3-5, 7-15, 74, and 75 can be used to amplify a human ABCA13 Isoform 5 cDNA, as shown in SEQ ID NO: 26; and the primers set forth as SEQ ID NO: 3, 4, 7, 8, 10-15, 74, 75 can be used to amplify a human ABCA13 Isoform 6 cDNA, as shown in SEQ ID NO: 28.

25 These primers are illustrative only; one skilled in the art will appreciate that many different primers may be derived from the provided cDNA or gene sequence in order to amplify particular regions of the ABCA13 cDNA isoforms, as well as the complete sequence of the human ABCA13 cDNA (Isoform 1, SEQ ID NO: 1).

Re-sequencing of PCR products obtained by amplification procedures optionally can be
30 performed to facilitate confirmation of the amplified sequence and provide information about natural variation of this sequence in different populations or species. Oligonucleotides derived from the provided ABCA13 sequences may be used in such sequencing methods.

Orthologs of human ABCA13 can be cloned in a similar manner, where the starting material consists of cells taken from a non-human species. The mouse gene ortholog to human ABCA13
35 displays 54% amino acid identity and 69% similarity in the large extracellular domain. In one embodiment, orthologs will generally share at least 65% sequence identity with the disclosed human ABCA13 cDNA. Where the non-human species is more closely related to humans, the sequence identity will in general be greater. In other embodiments, closely related orthologous ABCA13

molecules may share at least 70%, at least 75%, at least 80% at least 85%, at least 90%, at least 91%, at least 93%, at least 95%, or at least 98% sequence identity with the disclosed human sequences.

Oligonucleotides derived from the human *ABCA13* cDNA isoform sequences (e.g., SEQ ID NO: 1, 18, 20, 22, 24, 26, and 28), or fragments of these cDNAs, are encompassed within the scope of the present disclosure. In one embodiment, such oligonucleotides may comprise a sequence of at least 10 consecutive nucleotides of the *ABCA13* nucleic acid sequence. If these oligonucleotides are used with an *in vitro* amplification procedure (such as PCR), lengthening the oligonucleotides may enhance amplification specificity. Thus, in other embodiments, oligonucleotide primers comprising at least 15, 20, 25, 30, 35, 40, 45 or 50 consecutive nucleotides of these sequences may be used. These primers for instance may be obtained from any region of the disclosed sequences. By way of example, the human *ABCA13* cDNA, ORF and gene sequences may be apportioned into about halves, thirds or quarters based on sequence length, and the isolated nucleic acid molecules (e.g., oligonucleotides) may be derived from the first or second halves of the molecules, from any of the three thirds, or from any of the four quarters. For example, the full-length human *ABCA13* cDNA, shown in SEQ ID NO: 1, can be used to illustrate this. The human *ABCA13* cDNA is 17209 nucleotides in length, and the open reading frame begins at position 5 and ends at position 15202. In one specific embodiment, the open reading frame may be hypothetically divided into about halves (nucleotides 5-7600 and 7599-15202), in another specific embodiment, in about thirds (nucleotides 5-5060, 5061-10120, 10121-15202) or in yet another specific embodiment, in about quarters (nucleotides 5-3800, 3801-7600, 7601-11400 and 11401-15202). In related embodiments, the cDNA sequences of the additional Isoforms 2A, 2B, and 3-6, can be similarly divided.

VI. *ABCA13* Sequence Variants

With the provision of several human *ABCA13* isoform protein (SEQ ID NO: 2, 19, 21, 23, 25, 27, and 29) and corresponding nucleic acid sequences (SEQ ID NO: 1, 18, 20, 22, 24, 26, and 28) herein, the creation of variants of these sequences is now enabled.

In one embodiment, variant *ABCA13* proteins include proteins that differ in amino acid sequence from the human *ABCA13* sequences disclosed but that share at least 65% amino acid sequence identity with the provided human *ABCA13* protein. In other embodiments, other variants will share at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% amino acid sequence identity. A *ABCA13* encoding sequence may be isolated by routine procedures, such as those provided in Example 1. For instance, a *ABCA13* sequence may be isolated by homology screening using the cDNA sequence and the BLAST program. Direct sequencing, using the "long-distance sequence method," of one or more BAC or PAC clones that contain the *ABCA13* sequence can be employed. Manipulation of the nucleotide sequence of the *ABCA13* isoforms using standard procedures, including in one specific, non-limiting, embodiment, site-directed mutagenesis or in another specific, non-limiting, embodiment, PCR, can be used to produce such variants. The simplest modifications involve the substitution of one or more amino acids for amino acids having similar

biochemical properties. These so-called conservative substitutions are likely to have minimal impact on the activity of the resultant protein.

In another embodiment, more substantial changes in transporter function or other protein features may be obtained by selecting amino acid substitutions that are less conservative than conservative substitutions. In one specific, non-limiting, embodiment, such changes include
5 changing residues that differ more significantly in their effect on maintaining polypeptide backbone structure (e.g., sheet or helical conformation) near the substitution, charge or hydrophobicity of the molecule at the target site, or bulk of a specific side chain. The following specific, non-limiting, examples are generally expected to produce the greatest changes in protein properties: (a) a
10 hydrophilic residue (e.g., seryl or threonyl) is substituted for (or by) a hydrophobic residue (e.g., leucyl, isoleucyl, phenylalanyl, or valyl); (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain (e.g., lysyl, arginyl, or histadyl) is substituted for (or by) an electronegative residue (e.g., glutamyl or aspartyl); or (d) a residue having a bulky side chain (e.g., phenylalanine) is substituted for (or by) one lacking a side chain (e.g.,
15 glycine).

In other embodiments, changes in transporter activity or other protein features may be obtained by mutating, substituting or deleting regions of ABCA13 that have a known function, or regions where the function is yet to be determined. For instance, a NBD (nucleotide binding domain) motif of ABCA13 (corresponding to residues 3863 to 4044 or 4742 to 4894 of SEQ ID NO: 2) can be
20 deleted, substituted with the NBD of another protein or a synthetic NBD, or residues within the NBD motif can be mutated. In other embodiments, residues within any of the twelve ABCA13 transmembrane domain (TM) (e.g., residues 20-42, 3571-3793, 3414-3636, 3646-3668, 3677-3697, 3751-3773, 4226-4247, 4456-4473, 4508-4530, 4535-4556, 4571-4532, and 4654-4675 of SEQ ID NO: 2, see Figure 4) are mutated or deleted, or the TM is substituted with a transmembrane domain
25 of another protein or a synthetic TM. In related embodiments, mutations may be made within the ABCA13 transmembrane domain or NBD motif of the sequences encoding ABCA 13 Isoforms 2A, 2B, and 3-6. In yet other embodiments, one or more domain from one ABCA13 isoform can be exchanged with the domain of another isoforms.

Variant ABCA13 encoding sequences may be produced by standard DNA mutagenesis
30 techniques. In one specific, non-limiting, embodiment, M13 primer mutagenesis is performed. Details of these techniques are provided in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 15., CSHL, New York, 1989). By the use of such techniques, variants may be created that differ in minor ways from the human ABCA13 sequences disclosed. In one embodiment, DNA molecules and nucleotide sequences that are derivatives of those specifically disclosed herein, and
35 which differ from those disclosed by the deletion, addition, or substitution of nucleotides while still encoding a protein that has at least 65% sequence identity with the human ABCA13 encoding sequence disclosed (SEQ ID NO: 1), are comprehended by this disclosure. In some embodiments, at least one or more, at least 5 or more, at least 10 or more, at least 15 or more, at least 20 or more, or at least 25 or more nucleotides are deleted, added, or substituted while still encoding a protein that has

at least 65% sequence identity with the human ABCA13 encoding sequence disclosed (SEQ ID NO: 1). In other embodiments, more closely related nucleic acid molecules that share at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% nucleotide sequence identity with the disclosed ABCA13 sequences are comprehended by this disclosure. In one
5 embodiment, such variants may differ from the disclosed sequences by alteration of the coding region to fit the codon usage bias of the particular organism into which the molecule is to be introduced.

In other embodiments, the coding region may be altered by taking advantage of the degeneracy of the genetic code to alter the coding sequence such that, while the nucleotide sequence is substantially altered, it nevertheless encodes a protein having an amino acid sequence substantially
10 similar to the disclosed human ABCA13 protein sequences. For example, because of the degeneracy of the genetic code, four nucleotide codon triplets – (GCT, GCG, GCC and GCA) - code for alanine. The coding sequence of any specific alanine residue within the human ABCA13 protein, therefore, could be changed to any of these alternative codons without affecting the amino acid composition or characteristics of the encoded protein. Based upon the degeneracy of the genetic code, variant DNA
15 molecules may be derived from the cDNA and gene sequences disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of DNA sequences. Thus, this disclosure also encompasses nucleic acid sequences that encode an ABCA13 protein, but which vary from the disclosed nucleic acid sequences by virtue of the degeneracy of the genetic code.

In one embodiment, variants of the ABCA13 isoform proteins may also be defined in terms
20 of their sequence identity with the prototype human ABCA13 protein. As described above, human ABCA13 proteins share at least 65%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% amino acid sequence identity with the human ABCA13 isoform 1 protein (SEQ ID NO: 2). In some embodiments, at least one or more, at least 5 or more, at least 10 or more, at least 15 or more, at least 20 or more, or at least 25 or more amino acids are deleted, added, or substituted
25 while still encoding a protein that has at least 65% sequence identity with the ABCA13 isoform 1 encoding sequence (SEQ ID NO: 2). Nucleic acid sequences that encode such proteins/fragments readily may be determined simply by applying the genetic code to the amino acid sequence of an ABCA13 protein or fragment, and such nucleic acid molecules may readily be produced by assembling oligonucleotides corresponding to portions of the sequence.

30 Nucleic acid molecules that are derived from the human *ABCA13* cDNA nucleic acid sequences include molecules that hybridize under low stringency, high stringency, or very high stringency conditions to the disclosed prototypical *ABCA13* nucleic acid molecules, and fragments thereof.

Human *ABCA13* nucleic acid encoding molecules (including the cDNA shown in SEQ ID
35 NOs: 1, 18, 20, 22, 24, 26, 28 and nucleic acids comprising these sequences), and orthologs and homologs of these sequences, may be incorporated into transformation or expression vectors.

It will also be apparent to one skilled in the art that homologs of the *ABCA13* nucleic acid may now be cloned from other species, such as the rat or a monkey, by standard cloning methods. Such homologs will be useful in the production of animal models demonstrating the formation and

progression of a variety of tumors. In one embodiment, such orthologous *ABCA13* molecules will share at least 65% sequence identity with the human *ABCA13* nucleic acid disclosed herein; and in other embodiments, more closely related orthologous sequences will share at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% sequence identity with this sequence.

5

VII. Expression of *ABCA13* Protein

With the provision of human *ABCA13* encoding sequences (such as the cDNA shown in SEQ ID NOs: 1, 18, 20, 22, 24, 26, and 28), the expression and purification of the *ABCA13* proteins by standard laboratory techniques is now enabled. Purified human *ABCA13* protein may be used for functional analyses, antibody production, diagnostics, and patient therapy.

For instance, the DNA sequence of the full-length *ABCA13* cDNA (*e.g.*, SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28) can be manipulated in studies to understand the expression of the gene and the function of its product. In other embodiments, Isoforms 2A, 2B and 3-6 and/or mutant forms of the human *ABCA13* may be isolated based upon information contained herein, and may be studied in order to detect alteration in expression patterns in terms of relative quantities, cellular localization, tissue specificity and functional properties of the encoded mutant *ABCA13* protein. In yet other embodiments, partial or full-length cDNA sequences, which encode for the subject protein, may be ligated into bacterial expression vectors.

Methods for expressing large amounts of protein from a cloned gene introduced into *Escherichia coli* (*E. coli*) may be utilized for the purification, localization and functional analysis of proteins. By way of example, fusion proteins consisting of amino terminal peptides encoded by a portion of the *E. coli lacZ* or *trpE* gene linked to *ABCA13* proteins may be used to prepare polyclonal and monoclonal antibodies against these proteins. Thereafter, these antibodies may be used in other embodiments to purify proteins by immunoaffinity chromatography, in diagnostic assays to quantitate the levels of protein and to localize proteins in tissues and individual cells by immunofluorescence. Such antibodies may be specific for epitope tags, which can be added to the expression construct for identification and/or purification purposes.

Intact native protein also may be produced in *E. coli* or other cell culture systems in large amounts for functional studies. Methods and plasmid vectors for producing fusion proteins and intact native proteins in bacteria are described in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989). Such fusion proteins may be made in large amounts, are easy to purify, and can be used to elicit antibody response. In one embodiment, native proteins can be produced in bacteria by placing a strong, regulated promoter and an efficient ribosome binding site upstream of the cloned gene. If low levels of protein are produced, additional steps may be taken to increase protein production; if high levels of protein are produced, purification is relatively easy. Suitable methods are presented in Sambrook *et al.* and are well known in the art. In one embodiment, proteins expressed at high levels are found in insoluble inclusion bodies. Methods for extracting proteins from these aggregates are described by Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989).

Vector systems suitable for the expression of *lacZ* fusion genes include the pUR series of vectors (Ruther and Muller-Hill, *EMBO J.* 2:1791-1794, 1983), pEX1-3 (Stanley and Luzio, *EMBO J.* 3(6):1429-1434, 1984) and pMR100 (Gray *et al.*, *Proc. Natl. Acad. Sci. USA* 79(21):6598-6602, 1982). Vectors suitable for the production of intact native proteins include pKC30 (Belfort *et al.*, *J. Biol. Chem.* 258(3):2045-2051, 1983), pKK177-3 (Amann and Brosius, *Gene* 40(2-3):183-190, 1985) and pET-3 (Studier and Moffatt, *J. Mol. Biol.* 189(1):113-130, 1986). In one embodiment, ABCA13 fusion proteins may be isolated from protein gels, lyophilized, ground into a powder and used as an antigen. In other embodiments, the DNA sequence can also be transferred from its existing context to other cloning vehicles, such as other plasmids, bacteriophages, cosmids, animal viruses and yeast artificial chromosomes (YACs) (Burke *et al.*, *Science* 236:806-812, 1987). These vectors may then be introduced into a variety of hosts including, but not limited to, somatic cells, and simple or complex organisms, such as, but not limited to, bacteria, fungi (Timberlake and Marshall, *Science* 244:1313-1317, 1989), invertebrates, plants, and animals (Pursel *et al.*, *Science* 244:1281-1288, 1989), which cells or organisms are rendered transgenic by the introduction of the heterologous ABCA13 cDNA.

For expression in mammalian cells, the cDNA sequence is ligated to heterologous promoters. In one specific, non-limiting embodiment it may be ligated to the simian virus (SV) 40 promoter in the pSV2 vector (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981), and introduced into cells, such as monkey COS-1 cells (Gluzman, *Cell* 23:175-182, 1981), to achieve transient or long-term expression. In one embodiment, the stable integration of the chimeric gene construct may be maintained in mammalian cells by biochemical selection, such as neomycin (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341, 1982) and mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981). In another embodiment, cell lines expressing native ABCA13 are created. In yet another embodiment, cell lines expressing a mutant ABCA13 are created.

DNA sequences can be manipulated with standard procedures such as restriction enzyme digestion, fill-in with DNA polymerase, deletion by exonuclease, extension by terminal deoxynucleotide transferase, ligation of synthetic or cloned DNA sequences, site-directed sequence-alteration via single-stranded bacteriophage intermediate or with the use of specific oligonucleotides in combination with nucleic acid amplification. These techniques are known to those of ordinary skill in the art.

The ABCA13 cDNA sequence (or portions derived from it) or a mini gene (a cDNA with an intron and its own promoter) may be introduced into eukaryotic expression vectors by conventional techniques. These vectors are designed to permit the transcription of the cDNA in eukaryotic cells by providing regulatory sequences that initiate and enhance the transcription of the cDNA and ensure its proper splicing and polyadenylation. Vectors containing the promoter and enhancer regions of the SV40 or long terminal repeat (LTR) of the Rous Sarcoma virus and polyadenylation and splicing signal from SV40 are readily available (Mulligan *et al.*, *Proc. Natl. Acad. Sci. USA* 78:1078-2076, 1981; Gorman *et al.*, *Proc. Natl. Acad. Sci. USA* 78:6777-6781, 1982). The level of expression of the

cDNA can be manipulated with this type of vector by using promoters that have different activities (for example, the baculovirus pAC373 can express cDNAs at high levels in *S. frugiperda* cells (Summers and Smith, in *Genetically Altered Viruses and the Environment*, Fields *et al.* (Eds.) 22:319-328, CSHL Press, Cold Spring Harbor, New York, 1985) or by using vectors that contain
5 promoters amenable to modulation, for example, a recombinant adenoviral vector containing a nuclear lacZ gene driven by a human ventricular/slow muscle myosin light chain 1 promoter (Shi *et al.*, *Hum Gene Ther.* 8(4):403-410, 1997). The expression of the cDNA can be monitored in the recipient cells 24 to 72 hours after induction (transient expression).

Some vectors contain selectable markers such as the *gpt* (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981) or *neomycin* (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341, 1982) bacterial genes. These selectable markers permit selection of transfected cells that exhibit
10 stable, long-term expression of the vectors (and therefore the cDNA). By way of example, the vectors can be maintained in the cells as episomal, freely replicating entities by using regulatory elements of viruses, such as papilloma (Sarver *et al.*, *Mol. Cell Biol.* 1:486-496, 1981) or Epstein-
15 Barr (Sugden *et al.*, *Mol. Cell Biol.* 5:410-413, 1985). In another embodiment, one can produce cell lines that have integrated the vector into genomic DNA. Both of these types of cell lines produce the gene product on a continuous basis. Alternatively, one can produce cell lines that have amplified the number of copies of the vector (and therefore of the cDNA as well) to create cell lines that can produce high levels of the gene product (Alt *et al.*, *J. Biol. Chem.* 253: 1357-1370, 1978).

The vectors may contain an internal ribosomal entry site (IRES) between the cDNA and a
20 marker gene, such as neomycin or enhanced green fluorescent protein (EGFP). The IRES allows for the simultaneous expression of the two elements from a single isoform. Ribosomes bind the isoform at both the 5' end to translate the cDNA and at the IRES to translate, in one specific, non-limiting embodiment, the antibiotic resistance marker, or in another specific, non-limiting embodiment, the
25 fluorescent marker. The bicistronic expression via the IRES sequence provides a high degree of correlation between the antibiotic resistance and stable expression of the cDNA. Alternatively, only cells expressing the cDNA will show green fluorescence. Thus, the use of expression vectors containing an IRES is an efficient way to select for cells expressing the cDNA of interest.

The transfer of DNA into eukaryotic, in particular human or other mammalian cells, is now
30 a conventional technique. Recombinant expression vectors can be introduced into the recipient cells as pure DNA (transfection) by, for example, precipitation with calcium phosphate (Graham and vander Eb, *Virology* 52(2):456-467, 1973) or strontium phosphate (Brash *et al.*, *Mol. Cell Biol.* 7(5):2031-2035, 1987), electroporation (Neumann *et al.*, *EMBO J* 1(7):841-845, 1982), lipofection (Felgner *et al.*, *Proc. Natl. Acad. Sci USA* 84(21):7413-7417, 1987), DEAE dextran (Schenborn *et al.*,
35 *Methods Mol. Biol.* 130:147-153, 2000), microinjection (Mueller *et al.*, *Cell* 15(2):579-585, 1978), protoplast fusion (Schafner, *Proc. Natl. Acad. Sci. USA* 77:2163-2167, 1980), or gene guns (Nishitani *et al.*, *Cancer Gene Ther.* 9(2):156-163, 2002). In another embodiment, the ABCA13 isoform cDNA sequences, or fragments thereof, can be introduced by infection with virus vectors. Systems are developed that use, for example, retroviruses (Huszar, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*

82(24):8587-8591, 1985), adenoviruses (Schaack *et al.*, *Virology*. 291(1):101-109, 2001), or Herpes virus (Spaete *et al.*, *Cell* 30(1):295-304, 1982). Techniques of use in packaging long transcripts can be found in Kochanek *et al.* (*Proc. Natl. Acad. Sci. USA* 93:5731-5739, 1996), Parks *et al.* (*Proc. Natl. Acad. Sci. USA* 93:13565-13570, 1996) and Parks and Graham (*J. Virol.* 71:3293-3298, 1997).

- 5 In yet another embodiment, ABCA13 encoding sequences can be delivered to target cells *in vitro* via non-infectious systems, for instance liposomes.

These eukaryotic expression systems can be used for studies of *ABCA13* encoding nucleic acids and mutant forms of these molecules, the *ABCA13* isoform proteins and mutant forms of these proteins. Regulatory elements located in the 5' region of the *ABCA13* gene on genomic clones can be
10 isolated from human genomic DNA libraries using the information contained herein. In other embodiments, the eukaryotic expression systems also may be used to study the function of the normal complete protein, specific portions of the protein, or of naturally occurring or artificially produced mutant proteins.

In several embodiments, using the above techniques, expression vectors containing the
15 *ABCA13* isoform gene sequences or cDNAs, or fragments or variants or mutants thereof, can be introduced into human cells, mammalian cells from other species or non-mammalian cells, as desired. The choice of cell is determined by the purpose of the treatment. For example, in one specific, non-limiting embodiment monkey COS cells (Gluzman, *Cell* 23:175-182, 1981) that produce high levels of the SV40 T antigen and permit the replication of vectors containing the SV40 origin of replication
20 are used. In other embodiments, Chinese hamster ovary (CHO), mouse NIH 3T3 fibroblasts, or human fibroblasts or lymphoblasts are used.

Embodiments described herein thus encompass recombinant vectors that comprise all or part of an *ABCA13* encoding sequence, such as the *ABCA13* isoform genes or cDNAs or variants thereof, for expression in a suitable host. In one embodiment, an *ABCA13* DNA is operatively linked in a
25 vector to an expression control sequence in the recombinant DNA molecule so that the *ABCA13* polypeptide can be expressed. In another embodiment, the expression control sequence may be selected from the group consisting of sequences that control the expression of genes of prokaryotic or eukaryotic cells and their viruses and combinations thereof. The expression control sequence may be specifically selected from the group consisting of the *lac* system, the *trp* system, the *tac* system, the
30 *trc* system, major operator and promoter regions of phage lambda, the control region of coat protein, the early and late promoters of SV40, promoters derived from polyoma, adenovirus, retrovirus, baculovirus and simian virus, the promoter for 3-phosphoglycerate kinase, the promoters of yeast acid phosphatase, the promoter of the yeast alpha-mating factors and combinations thereof.

In another embodiment, the host cell, which may be transfected with a vector, may be
35 selected from the group consisting of *E. coli*, *Pseudomonas*, *Bacillus subtilis*, *Bacillus stearothermophilus* or other bacilli; other bacteria; yeast; fungi; insect; mouse or other animal; or plant hosts; or human tissue cells.

It is appreciated that for mutant or variant *ABCA13* DNA sequences (e.g., *ABCA13* isoforms encoded by SEQ ID NO: 18, 20, 22, 24, 26, or 28), similar systems are employed to express and produce the mutant product.

5 **VIII. *Suppression of ABCA13 Protein Expression***

A diseased condition is in some instances attributed to overexpression of a protein (e.g., HER-2/neu is overexpressed in breast cancer, *see Kaya et al., Pathol. Oncol. Res.* 7(4): 279-283, 2001). In such instances, reversal of the over-expression (e.g., suppression) is a potential method for ameliorating the diseased condition. In one embodiment of the disclosure, suppression of *ABCA13*
10 is achieved in a subject having a transporter malfunction due to overexpression of *ABCA13*.

Reduction of *ABCA13* protein expression in a transgenic cell may be obtained for instance by introducing into cells an antisense construct based on an *ABCA13* encoding sequence, including the human *ABCA13* cDNAs (SEQ ID NO: 1, 18, 20, 22, 24, 26, and 28) or gene sequence or flanking regions thereof. In one specific, non-limiting embodiment, a nucleotide sequence from an *ABCA13*
15 encoding sequence, e.g. all or a portion of one of the *ABCA13* cDNA isoform sequences, is arranged in reverse orientation relative to the promoter sequence in the transformation vector. Other aspects of the vector may be chosen as discussed above (*see* Section VII).

The introduced sequence need not be the full-length human *ABCA13* cDNA (SEQ ID NO: 1) or gene, and need not be exactly homologous to the equivalent sequence found in the cell type to be
20 transfected. In one embodiment, portions or fragments of the human *ABCA13* isoform cDNA (SEQ ID NO: 1, 18, 20, 22, 24, 26, and 28) could be used to knock out expression of the human *ABCA13* gene. Generally, however, where the introduced sequence is of shorter length, a higher degree of identity to the native *ABCA13* sequence will be needed for effective antisense suppression. In other embodiments, the introduced antisense sequence in the vector may be at least 15 nucleotides in
25 length, and improved antisense suppression typically will be observed as the length of the antisense sequence increases. In yet other embodiments, the length of the antisense sequence in the vector advantageously may be greater than 100 nucleotides, and can be up to about the full length of the human *ABCA13* cDNA or gene. In another embodiment, for suppression of the *ABCA13* gene itself, transcription of an antisense construct results in the production of RNA molecules that are the reverse
30 complement of mRNA molecules transcribed from the endogenous *ABCA13* gene in the cell.

Although the exact mechanism by which antisense RNA molecules interfere with gene expression has not been elucidated, it is believed that antisense RNA molecules bind to the endogenous mRNA molecules and thereby inhibit translation of the endogenous mRNA.

In another embodiment, suppression of endogenous *ABCA13* expression can be achieved
35 using ribozymes. Ribozymes are synthetic RNA molecules that possess highly specific endoribonuclease activity. The production and use of ribozymes are disclosed in U.S. Patent No. 4,987,071 to Cech and U.S. Patent No. 5,543,508 to Haselhoff. In one embodiment, the inclusion of ribozyme sequences within antisense RNAs may be used to confer RNA cleaving activity on the

antisense RNA, such that endogenous mRNA molecules that bind to the antisense RNA are cleaved, which in turn leads to an enhanced antisense inhibition of endogenous gene expression.

In yet another embodiment, dominant negative mutant forms of ABCA13 may be used to block endogenous ABCA13 activity.

5

IX. Production of an Antibody to ABCA13 Protein

Monoclonal or polyclonal antibodies may be produced to either the native ABCA13 protein or variant (*e.g.*, isoform) forms of this protein. In one embodiment, antibodies raised against an ABCA13 protein would specifically detect the ABCA13 protein. That is, such antibodies would
10 recognize and bind the ABCA13 protein, or fragments thereof, and would not substantially recognize or bind to other proteins found in human cells. In other embodiments, antibodies against the human ABCA13 protein may recognize ABCA13 from other species (*e.g.*, murine ABCA13), and vice versa.

Monoclonal or polyclonal antibodies to the protein can be prepared as follows:

15

A. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibodies to epitopes of the ABCA13 isoform proteins can be prepared from murine hybridomas according to the classical method of Kohler and Milstein (*Nature* 256:495-497, 1975) or derivative methods thereof. In one specific, non-limiting embodiment, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks.
20 The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused with mouse myeloma cells using polyethylene glycol, and the excess, non-fused, cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). Successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate, where growth of the culture is continued. Antibody-producing clones are identified by detection of
25 antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall (*Enzymol.* 70(A): 419-439, 1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988).

30

B. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein (for instance, expressed using a method described herein), which, in one specific, non-limiting embodiment, can be modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors
35 related both to the antigen and the host species. In one embodiment, small molecules may tend to be less immunogenic than others and may require the use of carriers and adjuvant, examples of which are known. In another embodiment, host animals may vary in response to site of inoculations and dose, with either inadequate or excessive doses of antigen resulting in low titer antisera. In one specific, non-limiting embodiment, a series of small doses (ng level) of antigen administered at

multiple intradermal sites may be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis *et al. J. Clin. Endocrinol. Metab.* 33: 988-991, 1971.

In one embodiment, booster injections will be given at regular intervals, and antiserum harvested when antibody titer thereof begins to fall, as determined semi-quantitatively (for example, by double immunodiffusion in agar against known concentrations of the antigen). See, for example, Ouchterlony *et al.* (In *Handbook of Experimental Immunology*, Wier, D. (ed.) chapter 19, Blackwell, 1973). In one specific, non-limiting embodiment the plateau concentration of antibody is usually in the range of about 0.1 to 0.2 mg/ml of serum (about 12 μ M). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher (*Manual of Clinical Immunology*, Ch. 42, 1980).

C. Antibodies Raised against Synthetic Peptides

A third approach to raising antibodies against ABCA13 proteins is to use synthetic peptides synthesized on a commercially available peptide synthesizer based upon the predicted amino acid sequence of the ABCA13 protein. Polyclonal antibodies can be generated by injecting such peptides into, for instance, rabbits.

D. Antibodies Raised by Injection of ABCA13 Encoding Sequence

In one embodiment, antibodies may be raised against an ABCA13 protein by subcutaneous injection of a recombinant DNA vector that expresses the ABCA13 protein into laboratory animals, such as mice. In one specific, non-limiting embodiment, delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford *et al.*, *Particulate Sci. Technol.* 5:27-37, 1987), as described by Tang *et al.* (*Nature* 356:152-154, 1992). In other embodiments, expression vectors suitable for this purpose may include those that express the ABCA13 encoding sequence under the transcriptional control of either the human β -actin promoter or the cytomegalovirus (CMV) promoter.

Antibody preparations prepared according to these protocols are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample.

X. ABCA13 as a Transporter Protein

The ABCA13 isoform proteins and their functional variants are believed to participate in the regulation of energy-dependent transport of a wide variety of substrates, including ions, metals, amino acids, lipids, peptides, sugars, vitamins, and steroid hormones across membranes. In one embodiment, the lack of expression of ABCA13 in a cell leads to a defect in transporter activity. In one specific, non-limiting embodiment, the administration of an ABCA13 recombinant nucleotide to a subject suffering from a transporter deficiency ameliorates the effects of the deficiency.

The expression of mutant transporter proteins can yield important information about the importance of each amino acid in the protein as well as the details of the mechanism of action of these proteins. Cells may have ABCA13 null mutations, ABCA13 missense mutations, or inactivation

of ABCA13. In one embodiment, a mutant ABCA13 is expressed in a cell but is incapable of localizing to the correct subcellular location. In another embodiment, a mutant ABCA13 is incapable of binding to its intracellular binding partners. In yet another embodiment, a mutation in the upstream regulatory region of the *ABCA13* gene abrogates the expression of the protein. In another
5 embodiment, a mutant ABCA13 is incapable of transporting target substances.

It is possible that mutations in the *ABCA13* gene that may lead to tumor formation or progression are not included in the cDNA but rather are located in other regions of the *ABCA13* gene. Mutations located outside of the ORF that encode the ABCA13 isoform proteins are not likely to affect the functional activity of the proteins, but rather are likely to result in altered levels of the
10 proteins in the cell. For example, mutations in the promoter region of the *ABCA13* gene may prevent transcription of the gene and therefore lead to the complete absence of the ABCA13 protein, or absence of certain transcripts of the protein, in the cell.

Additionally, mutations within introns in the genomic sequence may also prevent expression of the ABCA13 protein. Following transcription of a gene containing introns, the intron sequences
15 are removed from the RNA molecule, in a process termed splicing, prior to translation of the RNA molecule that results in production of the encoded protein. When the RNA molecule is spliced to remove the introns, the cellular enzymes that perform the splicing function recognize sequences around the intron/exon border and in this manner recognize the appropriate splice sites. If a mutation exists within the sequence of the intron near and exon/intron junction, the enzymes may not recognize
20 the junction and may fail to remove the intron. If this occurs, the encoded protein will likely be defective. Thus, mutations inside the intron sequences within the *ABCA13* gene (termed "splice site mutations") may also lead to defects in transporter activity. However, knowledge of the exon structure and intronic splice site sequences of the *ABCA13* gene is required to define the molecular basis of these abnormalities. The provision herein of *ABCA13* isoform cDNA sequences enables the
25 cloning of the entire *ABCA13* gene (including the promoter and other regulatory regions and the intron sequences) and the determination of its nucleotide sequence. With this information in hand, diagnosis of a genetic predisposition to transporter deficiency based on DNA analysis will comprehend all possible mutagenic events at the *ABCA13* locus.

Compounds that modulate the expression or activity of ABCA13 can be used to regulate
30 transporter activity. For instance, in some cases it may be determined that ABCA13 is expressed at low levels in a subject suffering from a transporter deficiency that arose as the result of the inefficient expression of ABCA13. Administration, to the subject, of an agent that up-regulates ABCA13 expression can improve transporter activity or ameliorate the transporter deficiency.

Changes in transporter activity can be assessed using a variety of *in vitro* and *in vivo* assays.
35 These assays can be used to study the effect of ABCA13 on specific cell types or the effect of particular mutations on ABCA13 transporter activity. By systematically introducing mutant ABCA13 constructs into cells and assessing their ability to transport substances across membranes, the importance of each amino acid for the protein's transporter activity can be determined. In

addition, these assays can be used to screen for modulators of ABCA13 activity. The modulators identified in this way can then be used to alter ABCA13 expression in cells *in vitro* or *in vivo*.

In one embodiment, cells expressing wild-type ABCA13 are assessed for their ability to transport substances *in vitro*, as compared to cells that express a mutant ABCA13. In this
5 embodiment, ABCA13 is transfected into cells, and levels of the transported substance in these cells can be assayed and compared between wild-type and transfected cells for potential differences. Methods of analysis include, for instance, immobilization of the protein on columns to search for compounds that bind the immobilized protein (*see Wainer et al., J. Chromatogr. B. Biomed. Sci. Appl.* 724(1): 65-72, 1999) and disruption of the *ABCA13* gene by antisense or RNAi technology (*see*
10 Paddison *et al., Proc. Natl. Acad. Sci. U. S. A.* 99(3): 1443-1448, 2002).

In yet another embodiment, samples or assays that are treated with a test compound that potentially modulates ABCA13 are compared to control samples that are not treated with the test compound, to examine the extent of modulation. In one embodiment, the compounds to be tested are present in the range from 0.1 nM to 10 mM. Control samples (untreated with modulators) are
15 assigned a relative ABCA13 activity value of 100%. In one embodiment, inhibition of ABCA13 is achieved when the ABCA13 activity value relative to the control is about 90%. In other embodiments, inhibition of ABCA13 is achieved when the ABCA13 activity value relative to the control is about 75%, about 50%, about 25%, or about 5%. In another embodiment, activation of ABCA13 is achieved when the ABCA13 activity value relative to the control is about 110% (*e.g.*
20 10% more than the control). In other embodiments, activation of ABCA13 is achieved when the ABCA13 activity value relative to the control is about 150%, about 175% or about 200%.

The effect of test compounds upon ABCA13 activity can be assessed using the assays described above. Such assays include, but are not limited to, the ability to transport substances across membranes. In one embodiment, the compounds tested as modulators of ABCA13 are any small
25 chemical compound, or biological entity, such as a polypeptide, sugar, nucleic acid or lipid. Furthermore, the ability of cells transfected with ABCA13 to transport fluorescent compounds can be assessed (*see Gripar et al., J. Membr. Biol.* 173(3): 203-214, 2000).

In another embodiment, the modulator is a genetically altered version of ABCA13. In other embodiments, the effect of potential modulators on ABCA13 protein or mRNA levels, transcriptional
30 activation or repression of a reporter gene is measured.

XI. DNA-Based Diagnosis

The *ABCA13* sequence information presented herein can be used in the area of genetic testing for predisposition to transporter deficiency owing to defects in *ABCA13*, such as deletion,
35 duplication or mutation of the *ABCA13* gene, or a portion thereof. The gene sequence of the *ABCA13* gene, including intron-exon boundaries is also useful in such diagnostic methods. Whether an individual is carrying mutations in the *ABCA13* gene (or a portion thereof), or has a duplication(s) or heterozygous or homozygous deletion(s) of the *ABCA13* gene, may be detected at the DNA level with the use of a variety of techniques. For such a diagnostic procedure, a biological sample of the

-40-

subject, which biological sample contains either DNA or RNA derived from the subject, is assayed for a mutated, duplicated or deleted *ABCA13* gene. Suitable biological samples include samples containing genomic DNA or RNA obtained from body cells, such as those present in peripheral blood, urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material.

5 Biological samples can be obtained from normal, healthy subjects or from subjects who are predisposed to or who are suffering from any one of a variety of the effects of transporter deficiencies such as, but not limited to, hypercholesterolemia, or defective transport of hydrophobic compounds such as lipids, sterols or fatty acids.

The detection in the biological sample of either a mutant *ABCA13* gene, a mutant *ABCA13* RNA, or a duplicated or homozygously or heterozygously deleted *ABCA13* gene, may be performed
10 by a number of methodologies, examples of which are discussed below.

One embodiment of such detection techniques for the identification of unknown mutations is the amplification (*e.g.*, polymerase chain reaction amplification) of reverse transcribed RNA (RT-PCR) isolated from a subject, followed by direct DNA sequence determination of the products. The
15 presence of one or more nucleotide differences between the obtained sequence and the prototypical *ABCA13* cDNA sequence, and especially, differences in the open reading frame (ORF) portion of the nucleotide sequence, are taken as indicative of a potential *ABCA13* gene mutation.

Alternatively, DNA extracted from a biological sample may be used directly for amplification. Direct amplification from genomic DNA would be appropriate for analysis of the
20 entire *ABCA13* gene including regulatory sequences located upstream and downstream from the open reading frame, or intron/exon borders. Reviews of direct DNA diagnosis have been presented by Caskey (*Science* 236:1223-1228, 1989) and by Landegren *et al.* (*Science* 242:229-237, 1989).

Other mutation scanning techniques appropriate for detecting unknown mutations within amplicons derived from DNA or cDNA could also be performed. These techniques include direct
25 sequencing (without sequencing), single-strand conformational polymorphism analysis (SSCP) (for instance, see Hongyo *et al.*, *Nucleic Acids Res.* 21:3637-3642, 1993), chemical cleavage (including HOT cleavage) (Bateman *et al.*, *Am. J. Med. Genet.* 45:233-240, 1993; reviewed in Ellis *et al.*, *Hum. Mutat.* 11:345-353, 1998), denaturing gradient gel electrophoresis (DGGE), ligation amplification mismatch protection (LAMP), and enzymatic mutation scanning (Taylor and Deeble, *Genet. Anal.*
30 14:181-186, 1999), followed by direct sequencing of amplicons with putative sequence variations.

If studies of *ABCA13* genes/coding sequences isolated from biological samples reveal particular mutations, genomic amplifications, or deletions, which occur at a high frequency within a population of individuals, DNA diagnostic methods can be designed to specifically detect the most common, or most closely disease-linked, *ABCA13* defects.

35 The detection of specific DNA mutations may be achieved by methods such as hybridization using allele specific oligonucleotides (ASOs) (Wallace *et al.*, *CSHL Symp. Quant. Biol.* 51:257-261, 1986), direct DNA sequencing (Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81:1991-1995, 1988), the use of restriction enzymes (Flavell *et al.*, *Cell* 15:25-41, 1978; Geever *et al.*, 1981), discrimination on the basis of electrophoretic mobility in gels with denaturing reagent (Myers and

Maniatis, Cold Spring Harbor Symp. Quant. Biol. 51:275-284, 1986), RNase protection (Myers *et al.*, *Science* 230:1242-1246, 1985), chemical cleavage (Cotton *et al.*, *Proc. Natl. Acad. Sci. USA* 85:4397-4401, 1985), and the ligase-mediated detection procedure (Landegren *et al.*, *Science* 241:1077-1080, 1988).

5 Oligonucleotides specific to normal or mutant sequences are chemically synthesized using commercially available machines. These oligonucleotides are optionally labeled radioactively with isotopes (such as ³²P) or non-radioactively, with tags such as biotin (Ward and Langer, *Proc. Natl. Acad. Sci. USA* 78:6633-6657, 1981), and hybridized to individual DNA samples immobilized on membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. These
10 specific sequences are visualized by methods such as autoradiography or fluorometric (Landegren *et al.*, *Science* 242:229-237, 1989) or colorimetric reactions (Gebeyehu *et al.*, *Nucleic Acids Res.* 15:4513-4534, 1987). Using an ASO specific for a normal allele, the absence of hybridization would indicate a mutation in the particular region of the gene, or deleted *ABCA13* gene. In contrast, if an ASO specific for a mutant allele hybridizes to a clinical sample, this would indicate the presence of a
15 mutation in the region defined by the ASO.

 Sequence differences between normal and mutant forms of the *ABCA13* gene may also be revealed by the direct DNA sequencing method of Church and Gilbert (*Proc. Natl. Acad. Sci. USA* 81:1991-1995, 1988). Cloned DNA segments may be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with nucleic acid
20 amplification, *e.g.*, PCR (Wrichnik *et al.*, *Nucleic Acids Res.* 15:529-542, 1987; Wong *et al.*, *Nature* 330:384-386, 1987; Stoflet *et al.*, *Science* 239:491-494, 1988). In this approach, a sequencing primer that lies within the amplified sequence is used with double-stranded PCR product or single-stranded template generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotides or by automatic sequencing procedures with fluorescent
25 tags.

 Sequence alterations may occasionally generate fortuitous restriction enzyme recognition sites or may eliminate existing restriction sites. Changes in restriction sites are revealed by the use of appropriate enzyme digestion followed by conventional gel-blot hybridization (Southern, *J. Mol. Biol.* 98:503-517, 1975). DNA fragments carrying the restriction site (either normal or mutant) are
30 detected by their reduction in size or increase in corresponding restriction fragment numbers. Genomic DNA samples may also be amplified by PCR prior to treatment with the appropriate restriction enzyme; fragments of different sizes are then visualized under UV light in the presence of ethidium bromide after gel electrophoresis.

 Genetic testing based on DNA sequence differences may be achieved by detection of
35 alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing reagent. Small sequence deletions and insertions can be visualized by high-resolution gel electrophoresis. For example, a PCR product with small deletions is clearly distinguishable from a normal sequence on an 8 % non-denaturing polyacrylamide gel (WO 91/10734; Nagamine *et al.*, *Am. J. Hum. Genet.* 45:337-339, 1989). DNA fragments of different sequence compositions may be distinguished on denaturing

-42-

formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific "partial-melting" temperatures (Myers *et al.*, *Science* 230:1242-1246, 1985). Alternatively, a method of detecting a mutation comprising a single base substitution or other small change could be based on differential primer length in a PCR. For example, an invariant primer could be used in addition to a primer specific for a mutation. The PCR products of the normal and mutant genes can then be differentially detected in acrylamide gels. Another method, single-strand conformation polymorphism (SSCP), is based on the fact that a single-base substitution alters the conformation of single-stranded DNA under non-denaturing conditions. Altered conformation affects the migration velocity of single-stranded DNA, which is detected as shifted or new bands on a non-denaturing gel. The mutations underlying the shifted or new bands are then characterized by sequencing.

In addition to conventional gel-electrophoresis and blot-hybridization methods, DNA fragments may also be visualized by methods in which the individual DNA samples are not immobilized on membranes. The probe and target sequences may be both in solution, or the probe sequence may be immobilized (Saiki *et al.*, *Proc. Nat. Acad. Sci. USA* 86: 6230-6234, 1989). A variety of detection methods, such as autoradiography involving radioisotopes, direct detection of radioactive decay (in the presence or absence of scintillant), spectrophotometry involving calorogenic reactions and fluorometry involved fluorogenic reactions, may be used to identify specific individual genotypes.

If multiple mutations are encountered frequently in the *ABCA13* gene, a system capable of detecting such multiple mutations likely will be desirable. For example, a nucleic acid amplification reaction with multiple, specific oligonucleotide primers and hybridization probes may be used to identify all possible mutations at the same time (Chamberlain *et al.*, *Nucl. Acids Res.* 16:1141-1155, 1988). The procedure may involve immobilized sequence-specific oligonucleotide probes (Saiki *et al.*, *Proc. Nat. Acad. Sci. USA* 86:6230-6234, 1989).

Expression levels of the *ABCA13* gene can also be determined by methods such as Northern or Southern blot analysis using labeled oligonucleotides specific to normal or mutant sequences. These oligonucleotides are labeled radioactively with isotopes (such as ^{32}P) or non-radioactively, with tags such as biotin (Ward and Langer, *Proc. Natl. Acad. Sci. USA* 78: 6633-6657, 1981), and hybridized to individual DNA samples immobilized on membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. Quantitative or semi-quantitative PCR can also be used to measure the amount of *ABCA13* cDNA in a sample using *ABCA13* oligonucleotide primers. Visualization methods such as autoradiography or fluorometric (Landegren *et al.*, *Science* 242: 229-237, 1989) or colorimetric reactions (Gebeyehu *et al.*, *Nucleic Acids Res.* 15: 4513-4534, 1987) can be used to detect a signal and the signals quantitated using, for instance, a spectrophotometer, a scintillation counter, a densitometer or a Phosphorimager (Amersham Biosciences). The Phosphorimager is able to analyze both DNA and protein samples from blots and gels using autoradiographic, direct fluorescence or chemifluorescence detection. Because the Phosphorimager is more sensitive than ordinary x-ray film, exposure times can be reduced up to ten-fold and signal

quantitation of both weak and strong signals on the same blot is possible. Images can be visualized and evaluated with the aid of computer programs such as ImageQuant™.

XII. Qualitative and Quantitative Detection of ABCA13 Protein

5 The *ABCA13* isoform sequence information presented herein is useful in detecting the presence or absence of ABCA13 in cultured cells or primary cells. Quantitative and qualitative methods of detection of proteins are well-known in the art, and are discussed herein. Quantitative detection is useful for diagnosing over- or underexpression of ABCA13 proteins in a subject, while qualitative information gives, for example, information regarding tissue types in which ABCA13
10 isoforms may be expressed. For such qualitative or quantitative assessment, a biological sample of the subject, which biological sample contains either DNA or RNA derived from the subject, is assayed for the presence or absence of ABCA13. Suitable biological samples include samples containing genomic DNA or RNA obtained from body cells, such as those present in peripheral blood, urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material.
15 Biological samples can be obtained from normal, healthy subjects or from subjects who are predisposed to or who are suffering from any one of a variety of the effects of transporter deficiencies such as, but not limited to, hypercholesterolemia, or defective transport of hydrophobic compounds such as lipids, sterols or fatty acids.

 Antibodies can be used to assess the presence or absence of ABCA13 proteins in cultured
20 cells or primary cells. The determination whether an antibody specifically detects ABCA13 proteins is made by any one of a number of standard immunoassay methods; for instance, the Western blotting technique (Sambrook *et al.*, In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989).

 In one embodiment, it is determined whether a given antibody preparation (such as one produced in a mouse) specifically detects ABCA13 proteins by Western blotting. In one specific,
25 non-limiting embodiment total cellular protein is extracted from human cells (for example, lymphocytes) and electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel.

 In another embodiment, the cellular protein is extracted from a biological sample. The proteins are then transferred to a membrane (for example, nitrocellulose or PVDF) by Western blotting, and the antibody preparation is incubated with the membrane. After washing the membrane
30 to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of (by way of example) an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase. Application of an alkaline phosphatase substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium results in the production of a dense blue compound by immunolocalized alkaline phosphatase. Antibodies that specifically detect ABCA13 proteins will, by
35 this technique, be shown to bind to ABCA13 protein bands (which will be localized at a given position on the gel determined by its molecular weight, the full-length isoform of which is approximately 28 kDa based on its deduced amino acid sequence). Non-specific binding of the antibody to other proteins may occur and may be detectable as a weak signal on the Western blot. The non-specific nature of this binding will be recognized by one skilled in the art by the weak signal

obtained on the Western blot relative to the strong primary signal arising from the specific antibody-ABCA13 protein binding.

In one embodiment, substantially pure ABCA13 protein suitable for use as an immunogen is isolated from the transfected cells as described above. In one specific, non-limiting embodiment the concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon (Millipore, Bedford, Massachusetts) or similar filter device, to the level of a few micrograms per milliliter.

In other embodiments, antibodies against ABCA13 proteins are used to localize ABCA13 to specific cell types or to specific subcellular locations in immunohistochemical or immunofluorescence assays. In one embodiment, the cells are selected from a variety of cell lines. In other embodiments, primary cells are isolated from a subject and are maintained in culture or the sample is sectioned and the sections are prepared directly for immunohistochemistry or immunofluorescence. In one specific, non-limiting embodiment, the cells are fixed, incubated in a blocking medium, incubated with the antibody directed against ABCA13 followed by a second incubation with a secondary antibody that is conjugated to a fluorescent probe or a colorimetric agent. Cells that express an ABCA13 protein that is recognized by the antibody exhibit a color or are fluorescent when viewed under a light or fluorescence microscope, respectively.

An alternative method of diagnosing *ABCA13* gene deletion, amplification, or mutation is to quantitate the level of ABCA13 protein in the cells of a subject. In one embodiment, this diagnostic tool would be useful for detecting reduced levels of ABCA13 protein that result from, for example, mutations in the promoter regions of the *ABCA13* gene or mutations within the coding region of the gene that produce truncated, non-functional or unstable polypeptides, as well as from deletions of the entire *ABCA13* gene. In another embodiment, duplications of the *ABCA13* gene may be detected as an increase in the expression level of this protein. The determination of reduced or increased ABCA13 protein levels would be an alternative or supplemental approach to the direct determination of *ABCA13* deletion, duplication or mutation status, for instance by the methods described herein.

The availability of antibodies specific to ABCA13 proteins will facilitate the quantitation of cellular ABCA13 proteins by one of a number of immunoassay methods, which are well known in the art and are presented herein and in, for instance, Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). Many techniques are commonly known in the art for the detection and quantification of antigen. In one specific, non-limiting embodiment, the purified antigen will be bound to a substrate, the antibody of the sample will bind via its Fab portion to this antigen, the substrate will then be washed and a second, labeled antibody will then be added, which will bind to the Fc portion of the antibody that is the subject of the assay. The second, labeled antibody will be species specific, *i.e.*, if the serum is from a rabbit, the second, labeled antibody will be anti-rabbit-IgG antibody. The specimen will then be washed and the amount of the second, labeled antibody that has been bound will be detected and quantified by standard methods.

Examples of methods for the detection of antibodies in biological samples, including methods employing dip strips or other immobilized assay devices, are disclosed for instance in the

following patents: U.S. Patents No. 5,965,356 (Herpes simplex virus type specific seroassay);
6,114,179 (Method and test kit for detection of antigens and/or antibodies); 6,077,681 (Diagnosis of
motor neuropathy by detection of antibodies); 6,057,097 (Marker for pathologies comprising an auto-
immune reaction and/or for inflammatory diseases); and 5,552,285 (Immunoassay methods,
5 compositions and kits for antibodies to oxidized DNA bases).

For the purposes of quantitating the ABCA13 protein, a biological sample of the subject,
which sample includes cellular proteins, can be used. Such a biological sample may be obtained from
body cells, such as those present in peripheral blood, urine, saliva, tissue biopsy, amniocentesis
samples, surgical specimens and autopsy material. Biological samples can be obtained from normal,
10 healthy subjects or from subjects who are predisposed to or who are already suffering from any one
of a variety of transporter deficiencies, such as, but not limited to, defective transport of cholesterol,
fatty acids, or lipids, or the effects of defective steroid hormones derived from the cholesterol of
individuals with mutant ABCA13.

Quantitation of ABCA13 protein can be achieved for instance by immunoassay and
15 compared to levels of the protein found in healthy cells (*e.g.*, cells from a subject known not to suffer
from a transporter deficiency). In one embodiment, a significant (*e.g.*, 10% or greater, for instance,
20%, 25%, 30%, 50% or more) reduction in the amount of ABCA13 protein in the cells of a subject
compared to the amount of ABCA13 protein found in normal human cells would be taken as an
indication that the subject may have deletions or mutations in the *ABCA13* locus, whereas in another
20 embodiment, a significant (*e.g.*, 10% or greater, for instance, 20%, 25%, 30%, 50% or more) increase
would indicate that a duplication or enhancing mutation had occurred.

XIII. ABCA13 Knockout and Overexpression Transgenic Animals

Mutant organisms that under-express or over-express ABCA13 proteins are useful for
25 research. Such mutants allow insight into the physiological and/or pathological role of ABCA13 in a
healthy and/or pathological organism via study using a mutant model system (*e.g.*, a transgenic
ABCA13 knockout mouse). These mutants are "genetically engineered," meaning that information in
the form of nucleotides has been transferred into the mutant's genome at a location, or in a
combination, in which it would not normally exist. Nucleotides transferred in this way are said to be
30 "non-native." In one embodiment, a non-*ABCA13* promoter inserted upstream of a native *ABCA13*
gene would be non-native. In other embodiments, an extra copy of an *ABCA13* gene or other
encoding sequence on a plasmid, transfected into a cell, would be non-native, whether that extra copy
was *ABCA13* derived from the same, or a different species.

Mutants may be, for example, produced from mammals, such as mice, that either over-
35 express or under-express ABCA13 protein, or that do not express ABCA13 at all. In one
embodiment, over-expression mutants are made by increasing the number of *ABCA13*-encoding
sequences (such as genes) in the organism. In other embodiments, over-expression mutants are made
by introducing an *ABCA13*-encoding sequence into the organism under the control of a constitutive or
inducible or viral promoter such as the mouse mammary tumor virus (MMTV) promoter or the whey

acidic protein (WAP) promoter or the metallothionein promoter. In yet other embodiments, mutants that under-express ABCA13 may be made by using an inducible or repressible promoter, or by deleting the *ABCA13* gene, or by destroying or limiting the function of the *ABCA13* gene, for instance by disrupting the gene by transposon insertion.

5 In another embodiment, antisense genes may be engineered into the organism, under a constitutive or inducible promoter, to decrease or prevent ABCA13 expression, as discussed above.

A gene is "functionally deleted" when genetic engineering has been used to negate or reduce gene expression to negligible levels. When a mutant is referred to in this application as having the *ABCA13* gene altered or functionally deleted, this refers to the *ABCA13* gene and to any ortholog of
10 this gene. When a mutant is referred to as having "more than the normal copy number" of a gene, this means that it has more than the usual number of genes found in the wild-type organism, *e.g.*, in the diploid mouse or human.

In one embodiment, a mutant mouse over-expressing ABCA13 may be made by constructing a plasmid having the *ABCA13* gene driven by a promoter, such as the mouse mammary
15 tumor virus (MMTV) promoter or the whey acidic protein (WAP) promoter. In one specific, non-limiting embodiment, this plasmid may be introduced into mouse oocytes by microinjection. The oocytes are implanted into pseudopregnant females, and the litters are assayed for insertion of the transgene. Multiple strains containing the transgene are then available for study.

WAP is quite specific for mammary gland expression during lactation, and MMTV is
20 expressed in a variety of tissues including mammary gland, salivary gland and lymphoid tissues. In other embodiments, other promoters might be used to achieve various patterns of expression, *e.g.*, the metallothionein promoter.

In another embodiment, an inducible system may be created in which the subject expression construct is driven by a promoter regulated by an agent that can be fed to the mouse, such as
25 tetracycline. Such techniques are well known in the art.

In yet another embodiment, a mutant knockout animal (*e.g.*, mouse) from which the *ABCA13* gene is deleted or otherwise disabled can be made by removing coding regions of the *ABCA13* gene from embryonic stem cells. The methods of creating deletion mutations by using a targeting vector have been described (see, for instance, Thomas and Capecch, *Cell* 51: 503-512,
30 1987).

In one embodiment, knockout mice are used as hosts to test the effects of various ABCA13 constructs on cell growth. In other embodiments, transgenic mice with the endogenous *ABCA13* gene knocked-out can be used in an assay to screen for compounds that modulate the ABCA13 activity. A transgenic mouse that is heterozygous or homozygous for integrated transgenes that have functionally
35 disrupted the endogenous *ABCA13* gene can be used as a sensitive *in vivo* screening assay for the ABCA13 ligands and modulators of ABCA13 activity.

XIV. Nucleic Acid-Based ABCA13 Therapy

Medical genetic approaches for combating ABCA13-mediated defects in subjects, such as a transporter deficiency, are now made possible.

In one embodiment, retroviruses are a preferred vector for experiments in medical genetics, as they yield a high efficiency of infection and stable integration and expression (Orkin *et al.*, *Prog. Med. Genet.* 7:130-142, 1988). In one specific, non-limiting embodiment, the full-length *ABCA13* gene or cDNA can be cloned into a retroviral vector and driven from either its endogenous promoter or, for instance, from the retroviral LTR (long terminal repeat). In other embodiments, viral transfection systems may also be utilized for this type of approach, including adenovirus, adeno-associated virus (AAV) (McLaughlin *et al.*, *J. Virol.* 62:1963-1973, 1988), Vaccinia virus (Moss *et al.*, *Annu. Rev. Immunol.* 5:305-324, 1987), Bovine Papilloma virus (Rasmussen *et al.*, *Methods Enzymol.* 139:642-654, 1987) or members of the herpesvirus group such as Epstein-Barr virus (Margolskee *et al.*, *Mol. Cell. Biol.* 8:2837-2847, 1988).

Medical genetic techniques include the use of RNA-DNA hybrid oligonucleotides, as described by Cole-Strauss, *et al.* (*Science* 273:1386-1389, 1996). This technique may allow for site-specific integration of cloned sequences, thereby permitting accurately targeted gene replacement.

In addition to delivery of *ABCA13* to cells using viral vectors, it is possible to use non-infectious methods of delivery. In one embodiment, lipidic and liposome-mediated gene delivery will be used for transfection of various genes (for reviews, see Templeton and Lasic, *Mol. Biotechnol.* 11:175-180, 1999; Lee and Huang, *Crit. Rev. Ther. Drug Carrier Syst.* 14:173-206; and Cooper, *Semin. Oncol.* 23:172-187, 1996). In another embodiment, cationic liposomes will be used as a viable alternative to the viral vectors (de Lima *et al.*, *Mol. Membr. Biol.* 16:103-109, 1999). In yet other embodiments, cationic liposomes can be targeted to specific cells through the inclusion of, for instance, monoclonal antibodies or other appropriate targeting ligands (Kao *et al.*, *Cancer Gene Ther.* 3:250-256, 1996).

XV. Kits

Kits are provided which contain the necessary reagents for determining *ABCA13* gene copy number, for determining altered expression of *ABCA13* mRNA or *ABCA13* protein, or for detecting polymorphisms in *ABCA13* alleles. Instructions provided in the diagnostic kits can include calibration curves, diagrams, illustrations, or charts or the like to compare with the determined (*e.g.*, experimentally measured) values or other results.

Kits are also provided that contain cells that serve as either positive or negative controls. These control cells can be compared to experimental samples containing similar cells, for instance cells of unknown gene activity, mutational state, protein expression level, and so forth.

A. Kits for Detection of ABCA13 Genomic Sequences

The nucleotide sequences disclosed herein, and fragments thereof, can be supplied in the form of a kit for use in detection of *ABCA13* genomic sequences, for instance in order to diagnose a

-48-

deficiency in transporter activity. In one embodiment of such a kit, an appropriate amount of one or more of the *ABCA13*-specific oligonucleotide primers is provided in one or more containers. In other embodiments, the oligonucleotide primers may be provided suspended in an aqueous solution or as a freeze-dried or lyophilized powder, for instance. The container(s) in which the oligonucleotide(s) are supplied can be any conventional container that is capable of holding the supplied form, for instance, microfuge tubes, ampoules, or bottles. In other embodiments, pairs of primers may be provided in pre-measured single use amounts in individual, typically disposable, tubes or equivalent containers. In one specific, non-limiting embodiment, the sample to be tested for the presence of *ABCA13* genomic amplification can be added to the individual tubes and *in vitro* amplification carried out directly.

The amount of each oligonucleotide primer supplied in the kit can be any appropriate amount, depending for instance on the market to which the product is directed. In one embodiment, the kit is adapted for research or clinical use and the amount of each oligonucleotide primer provided is an amount sufficient to prime several *in vitro* amplification reactions. Those of ordinary skill in the art know the amount of oligonucleotide primer that is appropriate for use in a single amplification reaction. General guidelines may for instance be found in Innis *et al.* (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990), Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989), and Ausubel *et al.* (In *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

In one embodiment, a kit may include more than two primers, in order to facilitate the PCR *in vitro* amplification of *ABCA13* sequences, for instance the *ABCA13* gene, specific exon(s) or other portions of the gene, or the 5' or 3' flanking region thereof.

In some embodiments, kits may also include the reagents necessary to carry out PCR *in vitro* amplification reactions, including, for instance, DNA sample preparation reagents, appropriate buffers (*e.g.*, polymerase buffer), salts (*e.g.*, magnesium chloride), and deoxyribonucleotides (dNTPs). Instructions may also be included.

In other embodiments, kits may include either labeled or unlabeled oligonucleotide probes for use in detection of the *in vitro* amplified *ABCA13* sequences. In one specific, non-limiting embodiment, the appropriate sequences for such a probe will be any sequence that falls between the annealing sites of the two provided oligonucleotide primers, such that the sequence the probe is complementary to is amplified during the *in vitro* amplification reaction.

In yet another embodiment, the kit provides one or more control sequences for use in the amplification reactions. The design of appropriate positive control sequences is well known to one of ordinary skill in the appropriate art.

B. Kits for Detection of *ABCA13* mRNA Expression

Kits similar to those disclosed above for the detection of *ABCA13* genomic sequences can be used to detect *ABCA13* mRNA expression levels. One embodiment of such a kit may include an appropriate amount of one or more of the oligonucleotide primers for use in reverse transcription

-49-

amplification reactions, similarly to those provided above, with art-obvious modifications for use with RNA.

In some embodiments, kits for detection of *ABCA13* mRNA expression levels may also include the reagents necessary to carry out RT-PCR *in vitro* amplification reactions, including, for instance, RNA sample preparation reagents (including *e.g.*, an RNase inhibitor), appropriate buffers (5 *e.g.*, polymerase buffer), salts (*e.g.*, magnesium chloride), and deoxyribonucleotides (dNTPs). Instructions also may be included.

In other embodiments, kits may include either labeled or unlabeled oligonucleotide probes for use in detection of the *in vitro* amplified target sequences. In one specific, non-limiting 10 embodiment, the appropriate sequences for such a probe will be any sequence that falls between the annealing sites of the two provided oligonucleotide primers, such that the sequence the probe is complementary to is amplified during the PCR reaction.

In another embodiment, the kit provides one or more control sequences for use in the RT-PCR reactions. The design of appropriate positive control sequences is well known to one of 15 ordinary skill in the appropriate art.

In yet other embodiments, kits may be provided with the necessary reagents to carry out quantitative or semi-quantitative Northern analysis of *ABCA13* mRNA. Such kits include, for instance, at least one *ABCA13*-specific oligonucleotide for use as a probe. This oligonucleotide may be labeled in any conventional way, including with a selected radioactive isotope, enzyme substrate, 20 co-factor, ligand, chemiluminescent or fluorescent agent, hapten, or enzyme.

C. Kits for Detection of *ABCA13* Protein or Peptide Expression

In some embodiments, kits for the detection of *ABCA13* protein expression include for instance at least one target protein specific binding agent (*e.g.*, a polyclonal or monoclonal antibody or antibody fragment) and may include at least one control. In another embodiment, the *ABCA13* 25 protein specific binding agent and control may be contained in separate containers. In other embodiments, the kits may also include means for detecting *ABCA13*:agent complexes, for instance the agent may be detectably labeled. If the detectable agent is not labeled, it may be detected by second antibodies, or protein A for example, which may also be provided in some kits in one or more separate containers. Such techniques are well known.

30 In another embodiment, the kits include instructions for carrying out the assay. Instructions will allow the tester to determine whether *ABCA13* expression levels are altered, for instance in comparison to a control sample. In other embodiments, reaction vessels and auxiliary reagents such as cells, chromogens, buffers, media, enzymes, etc. also may be included in the kits.

In one specific, non-limiting embodiment, an effective and convenient immunoassay kit 35 such as an enzyme-linked immunosorbent assay can be constructed to test anti-*ABCA13* antibody in human serum, as reported for detection of non-specific anti-ovarian antibodies (Wheatcroft *et al*, *Clin. Exp. Immunol.* 96:122-128, 1994; Wheatcroft *et al*, *Hum. Reprod.* 12:2617-2622, 1997). In one embodiment, expression vectors can be constructed using the human *ABCA13* cDNA to produce the recombinant human *ABCA13* protein in either bacteria or baculovirus (as described herein). In

-50-

another embodiment, affinity purification is used to generate unlimited amounts of pure recombinant ABCA13 protein.

In one embodiment, an assay kit could provide the recombinant protein as an antigen and enzyme-conjugated goat anti-human IgG as a second antibody as well as the enzymatic substrates.

5 Such kits can be used to test if the sera from a subject contain antibodies against human ABCA13.

D. Kits for Detection of Homozygous versus Heterozygous Allelism

Also provided are kits that allow differentiation between individuals who are homozygous versus heterozygous for a polymorphism of *ABCA13*. In one embodiment such kits provide the materials necessary to perform oligonucleotide ligation assays (OLA), for instance as described at
10 Nickerson *et al.* (*Proc. Natl. Acad. Sci. USA* 87:8923-8927, 1990). In specific embodiments, these kits contain one or more microtiter plate assays, designed to detect allelism in the *ABCA13* sequence of a subject, as described herein.

In one embodiment, additional components in some of these kits may include instructions for carrying out the assay. Instructions will allow the tester to determine whether an *ABCA13* allele is
15 homozygous or heterozygous. In other embodiments, reaction vessels and auxiliary reagents such as chromogens, buffers, enzymes, etc. may also be included in the kits.

In another embodiment, the kit may provide one or more control sequences for use in the OLA reactions. The design of appropriate positive control sequences is well known to one of ordinary skill in the appropriate art.

20 E. Kits for Identifying Modulators of ABCA13 activity

Also provided are kits that permit the identification of modulators of ABCA13 activity. In one embodiment, such kits provide the materials necessary to assess the activity of ABCA13 *in vitro*. In one embodiment, this kit contains aliquots of isolated ABCA13 and cultured cells. In another embodiment, the kit contains cell lines that express either wildtype or mutant ABCA13. In yet
25 another embodiment, additional components in some of these kits may include instructions for carrying out the assay. In other embodiments, reaction vessels and auxiliary reagents such as chromogens, buffers, media, enzymes, etc. may also be included in the kits.

The disclosure is illustrated by the following non-limiting Examples.

30

Example 1

Identification of human ABCA13 in genomic databases.

This example provides a description of the methods by which the *ABCA13* genomic
35 sequence was first identified.

EST / genomic and protein database searches

To identify new ABCA subfamily genes, public and commercial databases were screening with conserved domains identified from alignments of the known ABC genes, and a putative new

gene sequence was identified. After a first round of 5' elongation of this fragment a 1296 bp sequence was obtained. The human bacterial artificial chromosome (BAC) clones AC095039, AC021250, AC073424, AC073927, AC011635 were found to be overlapping and to contain exons of a putative ABC gene with a high degree of similarity with other ABCA subfamily transporters. After several rounds of exon prediction and alignment, the complete cDNA sequence was obtained by RT-PCR using primers specific for the gene (see Example 5, below). Sequence analysis of this PCR product revealed that the two NBDs are separated by a putative transmembrane domain. Additional 5' and 3' RACE reactions on human lung cDNA using primers specific to the new cDNA sequence led to the identification of a full-length cDNA, and several isoforms, which we named *ABCA13*.

Expressed sequence tags (EST) of ABCA1-like genes as described by Allikmets *et al.* (*Hum. Mol. Genet.* 5: 1649-1655, 1996) were used to search Genbank and UniGene nucleotide sequence databases using BLAST2 (Altschul *et al.*, *Nucleic Acids Res.*, 25: 3389-3402, 1997). The main protein sequences databases screened were GenBank, Incyte, Swissprot, TrEMBL, Genpept and PIR.

Sequence analysis

The genomic DNA analysis was performed by combination of several gene-finding programs such as GENSCAN (Burge and Karlin *J. Mol. Biol.* 268(1): 78-94, 1997), FGENEH/FEXH (Solovyev and Salamov *ISMB* 5: 294-302, 1997), and XPOUND (Thomas and Skolnick *J. Math. Appl. Med. Biol.* 11(1): 1-16, 1994). The combination of different tools yield increased sensitivity and specificity. The second step in the genomic DNA analysis was homology searching in the EST and protein databases. Combination of software performing database searching and software for exon/intron prediction gave the most sensitive and specific results. Sequence assembly and analysis were performed using the Genetics Computer Group (GCG) sequence analysis software package, Autoassembler software (ABI Prism, Perkin Elmer Applied Biosystems, Palo Alto, CA), and the LFASTA program (see Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85(8):2444-2448, 1988).

For phylogenetic analysis, ATP-binding domains defined by the Pfam signature ABCA tran (PF00005) were aligned using CLUSTALW (BCM Search Launcher, Baylor College of Medicine website). The phylogenetic tree was obtained using MEGA version 2.1 (Kumar *et al.*, *Bioinformatics* 17(12): 1244-1245, 2001). The GenBank accession numbers of the ABCA proteins were: ABCA1, O95477; ABCA2, NP_001597; ABCA3, NP_001080; ABCA4, NP_000341; ABCA5, NP_061142; ABCA6, NP_525023; ABCA7, NP_061985; ABCA8, NP_009099; ABCA9, NP_525022; ABCA10, NP_525021; ABCA12, NP_056472.

Multiple alignments were generated by GAP software from GCG package and the Dialign2 program (Morgenstern *et al.*, *Proc. Natl. Acad. Sci. USA* 93(22): 12098-12103, 1996), the FASTA3 package (Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85(8): 2444-2448, 1988) and SIM4 (Florea *et al.*, *Genome Res.* 8(9):967-974, 1998). The specific ABCA motifs used to search the databases were the TMN, TMC, NBF1 and NBF2 described in the literature (Broccardo *et al.*, *Biochim. Biophys. Acta* 1461(2):395-404, 1999). This corresponds in ABCA1 to residues 630-846 for the N terminal (TMN = exon 14-16) and from 1647-1877 for the C terminal set of membrane

-52-

spanners (TMC = exon 36-40). The NBD corresponds to the extended nucleotide binding domain, *i.e.* in ABCA1 it spans from amino acids 885-1152 for the N-terminal one (NBD1 = exon 18-22) and 1918-2132 for the C-terminal one (NBD2 = exon 42-47).

Expressed sequence tags (EST) of ABCA1-like genes (see Allikmets *et al.* *Hum. Mol. Genet.* 5: 1649-1655, 1996) were used to search the GenBank and UniGene nucleotide sequence databases using the BLAST 2 alignment computer program (see Altschul *et al.*, *Nucleic Acids Res.* 25(17): 3389-3402, 1997). With the known ABCA specific amino acid sequences, four positional matrices based on the specific motif definitions described above were built. Each is specific for NBD1, NBD2, TMN, and TMC, respectively. With the help of these matrices, the major public protein databases (Swissprot, TrEMBL, Genpept and PIR) were screened, and significant hits were extracted.

Example 2

Prediction of *ABCA13* Gene Structure.

This example provides a description of the methods by which the *ABCA13* genomic structure was predicted.

Gene structure analysis

Contigs were assembled from the following BAC clones: Ø AC073424 (g15321569) *Homo sapiens* chromosome 7 clone RP11-653O17, complete sequence, 191141 bp in length; AC095039 (g16974283) *Homo sapiens* chromosome 7 clone RP11-12G8, complete sequence, 70416 bp in length; AC073927 (g13992793) *Homo sapiens* BAC clone RP11-604B16 from chromosome 7, complete sequence, 185676 bp in length; AC091770 (g 15638906) *Homo sapiens* BAC clone RP11-655M5 from chromosome 7, complete sequence, 93543 bp in length. These contigs from 7p12.3 were analyzed using the Affy Genes, Ensembl Genes, and Fgenesh++ Genes dropdown combo box on the 'dense' setting at the University of California at Santa Cruz web site.

Genomic DNA analysis was performed using several gene predictors such as Genscan, FGENEH, FEXH on fragment 15 from BAC AC073424/gi119229 (Dec.2000). This prediction was combined using GENEWISE and PSI-BLAST algorithms and HMMER packages.

Physical mapping

The mouse *Abca13* gene was mapped using the T31 mouse/hamster radiation hybrid panel (Research Genetics). PCR was performed using primers Ma13F2, CAGCACCTCACAAATTGCCACC and Ma13R21, GCATGGGTAGGGTGCTGCCTG with Amplitaq Gold Taq polymerase. Samples were heated at 94°C for 10 minutes followed by 35 cycles at 94°C for 30 seconds, 65°C for 15 seconds, 72°C for 30 seconds, and extended at 72°C for 5 minutes. Reaction products were resolved on a 1.2% TAE agarose gel and data was submitted to The Jackson Laboratory Mouse Radiation Hybrid Database.

Results

The *ABCA13* gene is unusual in that the putative extracellular domain between the first and second predicted transmembrane segments is unusually large (about 3500 amino acids, 70% of the protein) and is encoded in part by two very large exons of 4.7 and 1.8 kb. *ABCA13* is the only described gene encoding a large exon that encodes an extracellular domain.

The mouse *ABCA13* gene contains two large exons of similar size to human *ABCA13*, which encode an amino acid sequence 54% identical and 69% similar to human *ABCA13*. These sequences have no recognizable homology to any known human proteins and contain no identifiable domains by MOTIF or PFAM searches.

Most eukaryotic exons are between 100-200 bp in length, and very few human exons are larger than 300 bp. There are a small number of human genes with large exons including *TTN*, *BRCA1* and *BRCA2*. *TTN*, which encodes a long muscle filament protein titin, has the largest described exon at 17.1 kb (Labeit and Kolmerer, *Science* 270:293-296, 1995). *BRCA1* and *BRCA2* contain exons of 3.4 and 4.9 kb, respectively.

Example 3**Isolation of cDNA of Human ABCA13 Isoforms.**

This example describes the isolation and identification of cDNA molecules encoding the full-length and additional isoforms of the human *ABCA13* protein.

Primers

Primers were designed using Oligo4 primer analysis software (Molecular Biology Insights, Cascade, CO) and ordered from Life Technologies, Invitrogen, Carlsbad, CA). The primers used to amplify regions of SEQ ID NO: 1 are listed in Table 7.

Table 7. Primers used to amplify regions of SEQ ID NO: 1

Primer	Primer Sequence	Region of SEQ ID NO: 1 amplified	SEQ ID NO:
132286L20new	CTAAGAGGGGTTTCAGGACGTGAC	7514 to 7491	3
132286_L22	AAACTGTGGACATCTTCCCCGAAAC	7643 to 7619	4
132286_L31	GTTGATACTCATCCTTACTTGC	4540 to 4519	5
132286_L33	ACATATTTCTTTTGACAACTAGAC	17089 to 17065	6
132286_U16	CCCTCTAAGAAACAGAGCAAC	7417 to 7437	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	14537 to 14558	8
132286_U28	TTGACTCCTCCATTGTTCCC	2271 to 2290	9
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	12719 to 12736	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	14854 to 14834	13
132286_C	TACCCTCCCCTCAGACTCAC	12782 to 12801	14
132286_E	GGGCTTGTTTCATGGTGAG	12748 to 12765	15

-54-

The primers used to amplify regions of SEQ ID NO: 18 are listed in Table 8.

Table 8. Primers used to amplify regions of SEQ ID NO: 18

Primer	Primer Sequence	Region of SEQ ID NO: 18 amplified	SEQ ID NO:
132286_L20new	CTAAGAGGGGTTTCAGGACGTGAC	2803 to 2780	3
132286_L22	AAACTGTGGACATCTTCCCCGAAAC	2932 to 2908	4
132286_L33	ACATATTTCTTTTGACAACTAGAC	12378 to 12354	6
132286_U16	CCCTCTAAGAAACAGAGCAAC	2706 to 2726	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	9826 to 9847	8
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	8008 to 8025	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	10143 to 10123	13
132286_C	TACCCTCCCCTCAGACTCAC	8071 to 8090	14
132286_E	GGGCTTGTTTCATGGTGAG	8037 to 8054	15

5 The primers used to amplify regions of SEQ ID NO: 22 are listed in Table 9.

Table 9. Primers used to amplify regions of SEQ ID NO: 22

Primer	Primer Sequence	Region of SEQ ID NO: 22 amplified	SEQ ID NO:
132286_L20new	CTAAGAGGGGTTTCAGGACGTGAC	7514 to 7491	3
132286_L22	AAACTGTGGACATCTTCCCCGAAAC	7643 to 7619	4
132286_L26	TGACTATTTACCAAATCTCTGCTTGAG	15667 to 15641	73
132286_L31	GTTGATACTCATCCTTACTTGC	4540 to 4519	5
132286_U16	CCCTCTAAGAAACAGAGCAAC	7417 to 7437	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	14537 to 14558	8
132286_U28	TTGACTCCTCCATTGTTCCC	2271 to 2290	9
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	12719 to 12736	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	14854 to 14834	13
132286_C	TACCCTCCCCTCAGACTCAC	12782 to 12801	14
132286_E	GGGCTTGTTTCATGGTGAG	12748 to 12765	15

The primers used to amplify regions of SEQ ID NO: 24 are listed in Table 10.

Table 10. Primers used to amplify regions of SEQ ID NO: 24

Primer	Primer Sequence	Region of SEQ ID NO: 24 amplified	SEQ ID NO:
132286_L20new	CTAAGAGGGGTTTCAGGACGTGAC	2803 to 2780	3
132286_L22	AAACTGTGGACATCTTCCCCGAAAC	2932 to 2908	4
132286_L26	TGACTATTTACCAAATCTCTGCTTGAG	10956 to 10930	73
132286_U16	CCCTCTAAGAAACAGAGCAAC	2706 to 2726	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	9826 to 9847	8
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	8008 to 8025	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	10143 to 10123	13
132286_C	TACCCTCCCCTCAGACTCAC	8071 to 8090	14
132286_E	GGGCTTGTTTCATGGTGAG	8037 to 8054	15

The primers used to amplify regions of SEQ ID NO: 26 are listed in Table 11.

5 Table 11. Primers used to amplify regions of SEQ ID NO: 26

Primer	Primer Sequence	Region of SEQ ID NO: 26 amplified	SEQ ID NO:
132286_L20new	CTAAGAGGGGTTTCAGGACGTGAC	7514 to 7491	3
132286_L22	AAACTGTGGACATCTTCCCCGAAAC	7643 to 7619	4
132286_L24	TATTCCTCCGTCTCAACCATAC	15234 to 15213	74
132286_L31	GTTGATACTCATCCTTACTTGC	4540 to 4519	5
132286_L34	ATTCAACAAAGAGTAAAGTGTGTAG	15654 to 15630	75
132286_U16	CCCTCTAAGAAACAGAGCAAC	7417 to 7437	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	14537 to 14558	8
132286_U28	TTGACTCCTCCATTGTTCCC	2271 to 2290	9
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	12719 to 12736	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	14854 to 14834	13
132286_C	TACCCTCCCCTCAGACTCAC	12782 to 12801	14
132286_E	GGGCTTGTTTCATGGTGAG	12748 to 12765	15

The primers used to amplify regions of SEQ ID NO: 28 are listed in Table 12.

Table 12. Primers used to amplify regions of SEQ ID NO: 28

Primer	Primer Sequence	Region of SEQ ID NO: 28 amplified	SEQ ID NO:
132286_L20new	CTAAGAGGGGTTTCAGGACGTGAC	2803 to 2780	3
132286_L22	AAACTGTGGACATCTTCCCGAAAC	2932 to 2908	4
132286_L24	TATTCCTCCGTCTCAACCATAC	10523 to 10502	74
132286_L34	ATTCAACAAAGAGTAAAGTGTGTAG	10943 to 10919	75
132286_U16	CCCTCTAAGAAACAGAGCAAC	2706 to 2726	7
132286_U18	GGAGACCTCATCAGGCGCTTAC	9826 to 9847	8
132286_U35	CTCAGAAATGCGATAGCTCAG	1376 to 1396	10
132286_U43	GACTCAGGAACCCGGTCC	81 to 98	11
132286_A	CTGCCAGTCCTCTTCGTG	8008 to 8025	12
132286_f38B	GATGTGCTGAGGAGAACCAAG	10143 to 10123	13
132286_C	TACCCTCCCCTCAGACTCAC	8071 to 8090	14
132286_E	GGGCTTGTTTCATGGTGAG	8037 to 8054	15

5' and 3' Rapid Amplification of cDNA Ends (RACE)

To determine the 5' or 3' ends of the *ABCA13* mRNA, 5' and 3' rapid amplification of cDNA ends (RACE) was performed using the SMART RACE cDNA amplification kit from Clontech (Palo Alto, CA). For 5' RACE amplification, *ABCA13* specific primers L1 (SEQ ID NO: 42) and L2 (SEQ ID NO: 43) were used to perform the first amplification on human lung polyA+ RNA (Clontech, Palo Alto, CA). A second, specific RACE amplification was performed with nested primers L7 (SEQ ID NO: 44) and L8 (SEQ ID NO: 45). Two others 5' RACE were carried out with specific primers L19 (SEQ ID NO: 46), L21 (SEQ ID NO: 47) and L20n (SEQ ID NO: 3), L22 (SEQ ID NO: 4) respectively. 3' RACE amplification was carried out similarly with U17 (SEQ ID NO: 48) as the first amplification primer, and U18 (SEQ ID NO: 8) as nested primer. The 5' and 3' RACE products were cloned, and the amplified products were analyzed by agarose gel electrophoresis, purified on Sephadex P100 columns and sequenced as described below.

Additional 5' and 3' RACE on human lung cDNA using primers specific to the new cDNA sequence led to the identification of a full-length cDNA, which was named *ABCA13*.

As shown in Figure 1, the principal transcript of the *ABCA13* gene is 17209 base pairs long with an open reading frame of 15195 base pairs (FIG. 1A and 1B), corresponding to 5058 amino acids (FIG. 1C and 1D). The predicted amino acid sequence of the protein is provided as SEQ ID NO: 2. Analysis of the *ABCA13* predicted protein disclosed features that are typical of an ABCA transporter. These include twelve predicted transmembrane domains, with two large extracellular domains and two nucleotide binding fold domains (NBD1 and NBD2) (see Figure 4). The position of the conserved nucleotide binding folds as well as that of the predicted transmembrane domains, are indicated in the Table 13. Each of these conserved regions is composed of hydrophobic segments.

Table 13. Genomic structure of the human *ABCA13* gene.

Exon	Position on cDNA	Splice acceptor	Exon size (bp)	Splice donor	Intron size (kb)
1	1-94		94	CAGGAACCCG gtagtgctt	n.d.
2	95-188	aattttctag GTCCTTTTCC	94	AGAGACATTT gtaagttca	n.d.
3	189-312	ctatttctag GTTATTTGCA	124	ATCATTTTCG gtaagagaaa	20.9
4	313-464	ctgtttaaag TTTGTCTAGG	152	AACACTCCAG gcaagtaaac	1.8
5	465-493	tttattatag ATTCTTCTTA	29	TTTTTTTACA gtaagtatct	5.9
6	494-657	ttcccatcag ATGGATCTCA	164	CCATTTTGAA gtagtgaaa	2.4
7	658-788	ttggttctag TTCCTTAATA	131	GCAGTCACCG gtatgggtgc	4.1
8	789-922	attccaacag AGCCAGTTTA	134	ACTGAAGGAG gtacacatgc	5.1
9	923-1087	ttgtgtcag ATTCCCACAG	135	CTACCAACAG gtgctgtccc	1.5
10	1088-1287	gaattgtag GTGTTTGTTC	200	TTCCAAAGAT gtaagtcgca	3.5
11	1288-1415	ttgccctcag ATTACAGCAT	128	TTTGTCCAAG gtaagctagc	0.8
12	1415-1516	tactttgcag AAGTCCTCAT	101	GCTGAAACAG gtaaagcaca	0.3
13	1517-1684	aatcttttag ATGTTGGCGA	168	AGAGGATGCT gtaagtattc	2.2
14	1685-1890	tcttgtaaag GATCGTATTT	206	TCTCCACAGT gtaagtacat	0.8
15	1891-2030	ttgattgcag GATATTTTCAT	140	AGACTATTGG gtaagtcagt	19.6
16	2031-2145	ttttatttag CTTTTCCTGA	115	CCATATCCAG gtaagttatc	2.7
17	2146-6924	aatattacag GGCTTTAAAT	4779	CAGAGATGAG gtaggtatac	1.5
18	6925-8751	atatttacag TTTTGTCCCA	1827	CAGATCAAAG gtaattaa	1.4
19	8752-8861	tttgctaag TGTTGTTGAG	110	ATCGTTGCAG gtgggtgct	6.5
20	8862-8980	cctctggcag AAAACCCTTC	119	CCACTTCCAG gttgtcgtc	7.6
21	8981-9144	tgttttctag GAAATTGAAA	164	ACTTGTACAT gtagtctctg	1.4
22	9145-9224	gccatttttag GTTGCCAAA	80	GTAAGTGAGG gtaagtattg	1.1
23	9225-9346	tactttgcag ATGTAAAAAT	122	CCACTCCAAG gtgtggtgct	11.5
24	9347-9541	tttgaacag GTTCTCTTCA	195	CATTTACAAG gtagtgagag	2.9
25	9542-9706	tgctcttag ACTCTGATGC	165	CTTTCAACAG gtgtgtgtt	1.0
26	9707-9884	attttctcag GTTTCACAAA	178	GAAGATTCAA gtaagacagt	2.7
27	9885-10024	attgcaacag CACCGTTTTG	140	CATTCAAAAG gtaagttaa	18.1
28	10025-10138	catittgcag GCTAATTACA	114	CCAGCTGCAG gtagtggtt	2.8
29	10139-10229	ttttcttag GAGGCCCTGA	91	CAGACATACG gtaagtgtgc	12.2
30	10230-10406	cttcctcag GAGGGCTGCT	177	TTCTTGCCA gtaagtactc	1.4
31	10407-10713	ctgccactag GTATCATTTT	307	CCAGCGACCT gtagtagcc	15.3
32	10714-10828	ccccttacag ATTCCTGAAC	115	GATAGAAGAG gtaaatatcc	4.3
33	10829-11158	ttttggtag TATATGCGGA	330	GACATTTCTG gtaagtaagt	1.8
34	11159-11228	ctgtgggcag TGCCTTCTTT	70	CAAGAGACAG gtaagagcat	2.0
35	11229-11360	ttcctgtag GGATTCAATG	132	TTGATTCCTG gtaagaattt	11.3
36	11361-11498	catitttttag GAACATTTGG	138	GACAATAAAG gttgtaagg	1.1
37	11499-11679	tctctcacag GGTCATCACT	181	CCACTATCAT gtgggtccca	2.7
38	11680-11898	tctctggcag ATCCATGTTG	219	AAGTCAATCA gtagtaa	11.5
39	11899-12095	ctaatttcag AACTCTTCAG	197	TACCGAGAAG gtaggactg	6.6
40	12096-12253	ttggaccag GTCGTACGAT	158	CACGAGGCAG gtaaggagt	1.7
41	12254-12484	ttatggatag CCTTCTGTTC	231	CTTAGAAGAG gtactagaa	15.2
42	12485-12590	tctccgaaag GTGTTTTTGA	106	TCTGGCTACT gtaagtacag	27.2
43	12591-12840	cgtcctgcag GTGGCTCCCT	250	ACTTTTTCAG gtaagttgtt	11.7
44	12841-12930	tgaacagcag CAGTGGGGGC	90	ACCCACGCCA gtaagtgca	4.5
45	12931-13000	tgtttttag GAAGAATTCT	70	TGGCTGCCTG gtaggttct	9.4

Table 13 (cont.).

Exon	Position on cDNA	Splice acceptor	Exon size (bp)	Splice donor	Intron size (kb)
46	13001-13119	acaatttcag AAGTGTCCTAA	119	AAAAACCAAG gtgtgttcaa	1.9
47	13120-13207	ttgttaacag GCTTGGAGGT	88	TCTGGCAAAG gtaatcatat	6.1
48	13208-13316	cttttttag GTGTGGTATA	109	AGACAATACG gtaatgttat	17.0
49	13317-13371	tcttcacag GAATAACACT	55	AGGACAAGAT gtgagttgat	1.5
50	13372-13549	ccaccgcag CCTGGAGAGC	178	ATATGACATG gtaggatttg	3.0
51	13550-13665	atcttctcag CTCTTTTACT	116	CACTTTTCGG gtatgtgatg	5.5
52	13666-13822	tacttttcag ATATGCAACT	157	CAAAGCTAAG gtcagtagct	3.2
53	13823-14076	cactgtgcag AATTTACAGA	254	GATGGCCAAG gtgggttctg	3.9
54	14077-14269	atctcccag GGGTCATTCT	193	AAAAGGAGAG gtaatgaagc	3.8
55	14270-14379	ttcctcttag TGCTTTGGAC	110	CTCCCATGGG gtaagataca	51.9
56	14380-14530	ctctctgcag AGACGCCGTG	151	CATCCCTGAG gtaaatctcc	6.8
57	14531-14665	tctcttcag GTTGCTGGAG	135	TCTTTTATTG gtgagtagaa	7.4
58	14666-14769	ttctctcag GATGAGCCCA	104	CCTCCCACAG gtgagttcca	20.5
59	14770-14862	ttctttacag CATGGAGGAG	93	TCAAAAATAG gtgcgttgaa	27.9
60	14863-14968	taactctcag GTTTGGTGAT	106	TCAGTTCAAG gtagtgctaa	1.2
61	14969-15106	ttgcttttag GGACAGCACC	138	TTTGGAGCAG gtatagtatc	0.7
62	15107-	ttttttcag GTATTTATTA	>2079		

n.d.; not determined because of the gaps in the genomic sequence.

The *ABCA13* gene is encoded by 62 exons and spans over 450 kb, making this the largest ABC gene described to date with the largest number of exons. The most unusual feature is the presence of two large exons of 4779 and 1827 bp (exons 17 and 18). All the remaining coding exons are less than 300 bp.

This principal transcript codes for a 5058 amino acid protein (SEQ ID NO: 2). However, five additional transcripts were characterized (Figure 1A and 1B). Transcript 2 is 12,498 bp long (SEQ ID NO: 18) and codes for a 2760 amino acid protein (SEQ ID NO: 19). Transcript 3 is 15,970 bp long (SEQ ID NO: 22) and encodes a 4958 amino acid protein (SEQ ID NO: 23). Transcript 4 is 11,259 bp long (SEQ ID NO: 24) and codes for a 2660 amino acid protein (SEQ ID NO: 25). Transcript 5 is 15,779 bp long (SEQ ID NO: 26) and encode a 4952 amino acid protein (SEQ ID NO: 27), and transcript 6 is 11,068 bp long (SEQ ID NO: 28) and its putative protein contains 2654 amino acids (SEQ ID NO: 29). These alternative transcripts result in the loss of the first transmembrane domain or in alternative C-termini, and can potentially encode functional transporters.

In the ORF coding for transcripts 2, 4 and 6, another putative protein of 717 amino acids is predicted (SEQ ID NO: 21), due to an internal alternative splicing of the 4779 bp exon (giving rise to the transcript shown in SEQ ID NO: 20). This putative protein may act as a regulatory factor for *ABCA13* or other protein.

DNA Sequencing

The PCR product was subjected to DNA cycle sequencing after purification with the Microcon-100 microconcentrators (Amicon, Inc., Charlotte, NC). The sequences were obtained

-59-

using the ABI Prism BigDye terminator cycle sequencing kit (Perkin Elmer Applied Biosystems, Palo Alto). Sequencing reactions were resolved on an ABI 377 DNA sequencer (Perkin Elmer Applied Biosystems, Palo Alto, USA) according to the manufacturer's instructions.

5 Results

The complete sequence of the human *ABCA13* cDNA is shown in SEQ ID NO: 1.

Example 4

Reverse Transcriptase PCR (RT-PCR) Analysis of Human *ABCA13* cDNA.

10 This example provides a description of how human *ABCA13* cDNA was procured and analyzed.

The complete cDNA sequence for human *ABCA13* was obtained by RT-PCR using primers specific for the gene.

15 cDNA templates for amplification were synthesized by reverse transcribing 500 ng of mRNA poly(A)+ (Clontech, Palo Alto, CA) using 200 units of reverse transcriptase Superscript II (Life Technologies, Invitrogen, Carlsbad, CA) with 500 ng of oligodT primer (Life Technologies, Invitrogen, Carlsbad, CA). The reactions were denatured at 70°C for 10 minutes, and 10 units of RNasin (Life Technologies, Invitrogen, Carlsbad, CA) were added, followed by incubation for 45
20 minutes at 42°C.

PCR was performed with AmpliTaq Gold DNA polymerase (Perkin Elmer Applied Biosystems, Palo Alto, CA), using 50 ng of DNA or about 25 ng of cDNA. Reactions were carried out for 30 PCR cycles in a 9700 thermal cycler (Perkin Elmer Applied Biosystems, Palo Alto, CA) in 96-well microtiter plates. After an initial denaturation at 94°C for 10 minutes, each cycle consisted
25 of a denaturation step of 30 seconds (94°C), a hybridization step of 30 seconds (64°C for 2 cycles, 61°C for 2 cycles, 58°C for 2 cycles and 55°C for 28 cycles), and an elongation step of 1 minute/kb (72°C). PCR was concluded with a final 72°C extension, which lasted 7 minutes.

Several primers have been designed to validate electronically predicted sequences or products of RACE. The initial fragment was validated using specific primers E (SEQ ID NO: 15) and
30 F (SEQ ID NO: 49). The second partial sequence derived from analysis of a BAC fragment was analyzed with primers A (SEQ ID NO: 50) and B (SEQ ID NO: 13). The link with the initial fragment was confirmed with primers C (SEQ ID NO: 51), B and E. First RACE was validated with primers U3 (SEQ ID NO: 52), U5 (SEQ ID NO: 53), and L1. Second 5' RACE was validated with primer U16 (SEQ ID NO: 7) used with B followed by nested PCR with E and B. Third 5' RACE was
35 analyzed with primers U28 (SEQ ID NO: 9) and L20n, U36 (SEQ ID NO: 54) and L20n or L31 (SEQ ID NO: 5), U38 (SEQ ID NO: 55) and L31, U43 (SEQ ID NO: 11) and L41 (SEQ ID NO: 56).

The unique 3' RACE was validated with specific primers L33 (SEQ ID NO: 6), L34 (SEQ ID NO: 75), and L26 (SEQ ID NO: 73), all used with U18.

-60-

PCR product was subjected to DNA cycle sequencing after purification with the Microcon-100 microconcentrators (Amicon, Inc., Beverly). Sequences were obtained using the ABI Prism BigDye terminator cycle sequencing kit (Perkin Elmer Applied Biosystems, Palo Alto). Sequencing reactions were resolved on an ABI 377 DNA sequencer (Perkin Elmer Applied Biosystems, Palo Alto, USA). DNA sequences have been submitted to GenBank under accession # AY204751.

Sequence analysis of this PCR product revealed that the two putative NBDs are separated by a putative transmembrane domain. The order and sequence of these 3 putative protein motifs are indeed characteristic of an ABCA gene.

Example 5

Sequence Analysis of Human *ABCA13* cDNA.

This example provides a description of how the *ABCA13* sequence was analyzed.

Sequence assembly and analysis was performed using the Autoassembler software (ABI Prism, Perkin Elmer Applied Biosystems, Palo Alto, CA), the Genetics Computer Group (GCG) (Accelrys Corp., San Diego, CA) sequence analysis software package (Devereux *et al.*, *Nucleic Acids Res.* 12:387-395, 1984), and the LFASTA program package (*see* Pearson and Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85(8): 2444-2448, 1988). The programs were operated using default parameters, and were used to assemble sequences and generate the consensus sequence of each isoform.

Secondary structure of the proteins was predicted using the TMHMM on-line server provided by the Center for Biological Sequence Analysis (CBS), Lyngby, Denmark.

Example 6

Localization of *ABCA13* to a Chromosome.

This example provides a description of how the human and mouse *ABCA13* genes were localized to their chromosomal locations.

The chromosomal localization of the human *ABCA13* gene on chromosome 7 at 7p12.3 was determined by PCR using primers E and F and T31 mouse/hamster radiation hybrid panel (Research Genetics) according to the manufacturer's protocol. Subsequently, the position of *ABCA13* on the draft human map was determined using chromosomal assignment using somatic cell hybrids (CASH) (*see* Ryu *et al.*, *Mol Cells* 10(5): 598-600, 2000).

Mouse *Abca13* was mapped to chromosome 11 between markers D11Mit226 and D11Mit259 using the T31 mouse/hamster radiation hybrid panel (Research Genetics, Huntsville, AL). PCR was performed using primers Ma13F2 5'-CAG CAC CTC ACA AAT TGC CAC C-3' (SEQ ID NO: 16) and Ma13R21 5'-GCA TGG GTA GGG TGC TGC CTG-3' (SEQ ID NO: 17) with Amplitaq Gold Taq polymerase. Samples were heated at 94°C for 10 minutes followed by 35 cycles at 94°C for 30 seconds, 65°C for 15 seconds, 72°C for 30 seconds, and a final extension at

-61-

72°C for 5 minutes. Reaction products were resolved on a 1.2% TAE agarose gel and data was submitted to online Jackson Laboratory Mouse Radiation Hybrid Database (Bar Harbor, Maine).

When the assembled mouse genome was analyzed, a sequence closely related to human *ABCA13* was found at mouse chromosome 11A1, a position syntenic with human chromosome 7p12.3. Further analysis revealed the presence of two large exons, similar to those found in the human *ABCA13*.

Example 7

Amino acid sequence comparison of ABC transporters.

This example provides a description of how the human *ABCA13* homology and identity was determined with respect to other known ABCA proteins.

Comparison of the *ABCA13* protein sequence to sequences of other known ABCA genes was performed by computer analysis searching. The results of the analysis are summarized in Table 14.

Table 14. Similarity / identity percentages between ABCA amino acid sequences.

	ABCA1	ABCA4	ABCA7	ABCA12	ABCA13
ABCA1	100/100	60.04/51.88	62.87/54.18	46.96/36.23	46.97/36.45
ABCA4		100/100	58.06/48.83	45.68/34.85	45.54/35.14
ABCA7			100/100	46.36/36.32	44.51/35.88
ABCA12				100/100	47.32/38.06
ABCA13					100/100

The overall similarities and identities of ABCA13 with ABCA1, ABCA4, ABCA7, and ABCA12 are 47/37%, 46/35%, 45/36% and 47/38%, respectively. These results indicate that ABCA13 shares the highest level of homology with ABCA1 and ABCA12 among the ABCA subfamily members. Comparison of all known members of the ABCA subfamily reveals that the highest levels of sequence homology are found in their ATP binding domains (NBD1 and NBD2).

Although the region of highest identity with other ABCA proteins starts around amino acid 3070 of ABCA13, there is a short, 90-amino acid region at the very N-terminus, which shows a high degree of identity with the N-terminal regions of other ABCA proteins (Figure 3). In addition, this N-terminal region contains a putative transmembrane helix found in all ABCA proteins. The enlarged size of ABCA13 is therefore due to a large insertion into the putative extracellular domain after the first N-terminal membrane-spanning helix.

Example 8

Phylogenetic analysis of the relationship between the ABCA13 to other ABCA transporters.

This example provides a description of how the evolutionary relationship between human *ABCA13* and other ABCA transporters was determined.

-62-

Phylogenetic analysis was performed using the ATP-binding domain, which is conserved across the species and in different subfamilies without large deletions or insertions (see Bodenmiller *et al.*, *DNA Seq.* 13(2): 77-83, 2002). Analysis was conducted using both ATP-binding domains, as all twelve human ABCA subfamily members known to date contain two such domains.

Phylogenetic analysis indicated that ABCA13 belongs to the subgroup inside the A subfamily, which contains the best-characterized members of this subfamily, ABCA1 (see Tanaka *et al.*, *Biochem. Biophys. Res. Commun.* 25; 283(5): 1019-1025, 2001) and ABCA4 (see Bungert *et al.*, *J. Biol. Chem.* 276(26): 23539-23546, 2001). The second subgroup includes five genes that form a tandem cluster at chromosome 17q24. When compared to other ABCA proteins, ABCA13 has a short (about 90 amino acids) region of similarity at the N-terminus, and larger region of similarity starting from amino acid 3070. This suggests that the large size of ABCA13 is due to the large insertion between these positions at the first extracellular domain. A graphic representation of the phylogenetic analysis is shown in Figure 2.

Example 9

Reverse Transcription Analysis of the Expression Pattern of ABCA13.

The profile of expression of *ABCA13* polynucleotides can be determined using art-known techniques, such as PCR-coupled reverse transcription, which has been described for instance by Sambrook *et al.* (*Molecular cloning: a laboratory manual*. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989). This example provides representative methods and results from various tissues.

By way of example, a pair of primers as described herein (*e.g.*, SEQ ID NO: 3 through 15) may be synthesized and used to amplify a human *ABCA13* cDNA (SEQ ID NO: 1).

The expression pattern of *ABCA13* was examined in a variety of human tissues using RT-PCR. The polymerase chain reaction (PCR) was carried out on commercially provided first-strand cDNA templates (CLONTECH) corresponding to retro-transcribed polyA⁺ mRNAs. Reverse transcription to cDNA was carried out with the enzyme SUPERScript II (GibcoBRL, Life Technologies) according to the conditions described by the manufacturer.

The polymerase chain reaction was carried out according to standard conditions, in 20 μ l of reaction mixture with 25 ng of cDNA preparation. The reaction mixture was composed of 400 μ M of each of the dNTPs, 2 units of *Thermus aquaticus* (Taq) DNA polymerase (Ampli Taq Gold; Perkin Elmer), 0.5 μ M of each primer, 2.5 mM MgCl₂, and PCR buffer. Thirty four PCR cycles [denaturing 30 seconds at 94°C, annealing of 30 seconds divided up as follows during the 34 cycles: 64°C (2 cycles), 61°C (2 cycles), 58°C (2 cycles), and 55°C (28 cycles), and an extension of one minute per kilobase at 72°C] were carried out after a first step of denaturing at 94°C for 10 minutes using a Perkin Elmer 9700 thermocycler. The PCR reactions were visualized on agarose gel by electrophoresis, and their relative level quantified.

-63-

Using RT-PCR, the amount of ABCA13 isoforms was measured in various tissues. *ABCA13* expression was found to be highest in human brain, lung, skeletal muscle and ovary. Thus like other ABCA subfamily genes, ABCA13 presents a restricted tissue distribution. The expression of the gene was also examined in a panel of tumor cell lines and was found to be expressed the

5 highest in the SR leukemia, SNB-19 CNS tumor and DU-145 prostate tumor cell lines.

Example 10

Real-Time PCR Analysis of the Expression Pattern of ABCA13

The profile of expression of *ABCA13* polynucleotides also can be determined using real-time

10 PCR analysis. This example provides representative methods and results from various tissues and cell lines.

Real-time PCR

Real-time PCR (ABI Prism 7700 Sequence Detection System, Applied Biosystems, Foster

15 City, CA) was used to determine the expression profiling of the different *ABCA13* transcripts, with β 2-microglobulin as a standard reference gene, in commercial total RNA samples extracted from various tissues (Figure 5). Measured RNA levels were normalized to β 2-microglobulin. One set of probe and primers was designed for each 3' end:

Transcripts 1 and 2

20 Fwd : SEQ ID NO: 57;
Rev : SEQ ID NO: 58;
Probe (MGB) : SEQ ID NO: 59;

Transcripts 3 and 4

25 Fwd : SEQ ID NO: 60;
Rev : SEQ ID NO: 61;
Probe (MGB) : SEQ ID NO: 62;

Transcripts 5 and 6

30 Fwd : SEQ ID NO: 63;
Rev : SEQ ID NO: 64;
Probe (TAMRA) : SEQ ID NO: 65).

The probes were labeled at the 5' position with 6-carboxyfluorescein (6-FAM) reporter. The 3' end was coupled to a quencher, either 6-carboxytetramethylrhodamine (TAMRA) or minor groove binder (MGB).

Reverse transcription of total RNAs was carried out in 5 to 10 different tubes after

35 denaturation of nucleic acid at 65°C for 10 minutes. Two micrograms of denatured total RNA were added into 100 μ l of reaction mix containing 1x Taqman Buffer, 5.5 mM MgCl₂, 500 μ M each dNTP, 2.5 μ M random hexamers, 2.5 μ M oligodT, 0.4 U/ μ l RNase inhibitor, and 1.25 U/ μ l

-64-

MultiScribeRTase. The thermal cycling conditions consisted of 25°C for 10 minutes, 42°C for 50 minutes, and 95°C for 5 minutes. The reaction products of the 5 to 10 tubes were pooled.

Real-time PCR was carried out under duplex conditions, using 10 µl of 1:10 diluted cDNA in a total volume of 35 µl. The PCR reaction master mix contains 1x Master Mix, 900 nM of each primer specific of selected *ABCA13* transcript, 50 nM of each primer specific of β 2-microglobulin transcript, and 200 nM of *ABCA13* and β 2-microglobulin probes. The thermal cycling conditions are 40 cycles of PCR amplification (UNG incubation: 50°C, 2 minutes; AmpliTaqGold activation: 95°C, 10 minutes; denaturation: 95°C, 15 seconds; annealing/extension: 60°C, 1 minute).

10 *Real-time PCR analysis of mouse tissue and tumor cell line RNA*

cDNA was quantitated for expression of the *ABCA13* gene in each of 60 tumor cell lines from the NCI tumor cell line screening panel (which includes the following cell lines: CCRF-CEM (leukemia); HL-60(TB) (leukemia); K-562 (leukemia); MOLT-4 (leukemia); RPMI-8226 (leukemia); SR (leukemia); A549/ATCC (non-small cell lung); EKVX (non-small cell lung); HOP-62 (non-small cell lung); HOP-92 (non-small cell lung); NCI-H226 (non-small cell lung); NCI-H23 (non-small cell lung); NCI-H322M (non-small cell lung); NCI-H460 (non-small cell lung); NCI-H522 (non-small cell lung); COLO 205 (colon); HCC-2998 (colon); HCT-116 (colon); HCT-15 (colon); HT29 (colon); KM12 (colon); SW-620 (colon); SF-268 (CNS); SF-295 (CNS); SF-539 (CNS); SNB-19 (CNS); SNB-75 (CNS); U251 (CNS); LOX IMVI (melanoma); MALME-3M (melanoma); M14 (melanoma); SK-MEL-2 (melanoma); SK-MEL-28 (melanoma); SK-MEL-5 (melanoma); UACC-257 (melanoma); UACC-62 (melanoma); IGR-OV1 (ovarian); OVCAR-3 (ovarian); OVCAR-4 (ovarian); OVCAR-5 (ovarian); OVCAR-8 (ovarian); SK-OV-3 (ovarian); 786-0 (renal); A498 (renal); ACHN (renal); CAKI-1 (renal); RXF 393 (renal); SN12C (renal); TK-10 (renal); UO-31 (renal); PC-3 (prostate); DU-145 (prostate); MCF7 (breast); NCI/ADR-RES (breast); MDA-MB-231/ATCC (breast); HS 578T (breast); MDA-MB-435 (breast); MDA-N (breast); BT-549 (breast); and T-47D (breast)).

Expression analyses of *ABCA13* cDNA in these tumor cell lines was performed using a real time PCR assay on the ABI 7000 apparatus. A multiplex reaction was performed along with primers and probe for β -actin provided by the manufacturer (Applied Biosystems) as the endogenous control. *ABCA13* mRNA was quantitated with primers *ABCA13*-TAQ-F (SEQ ID NO: 66), *ABCA13*-TAQ-R (SEQ ID NO: 67), and probe *ABCA13*-TAQ-P, fam-CTTGGGCAGCATCTTGGTCAATCTCTCT-tamra (SEQ ID NO: 68).

Expression of mouse *Abca13* was analyzed by PCR on the mouse multiple tissue cDNA panel (Clontech, Palo Alto, CA). Primers MmAbca13_F1 (SEQ ID NO: 69), and MmAbca13_R1 (SEQ ID NO: 70) were used to amplify a 255 bp fragment spanning exons one, two and three that are encoding the first transmembrane helix. Primers MmAbca13_F17 (SEQ ID NO: 71), and MmAbca13_R18 (SEQ ID NO: 72), were used to amplify a 381 bp fragment from exons 17 and 18. Thirty cycles of PCR, each cycle consisting of denaturation at 94°C for 30 s, annealing at 58°C for 30

seconds and extension at 68°C for 1 minute, were carried out using Platinum Taq polymerase (Invitrogen, Carlsbad, CA) and buffer conditions suggested by the manufacturer.

Results

5 The *ABCA13* expression pattern was examined using real-time-PCR. One set of probe and primers was designed for each 3' end. Transcripts 1 and 2 were found to be highest in human trachea, testis, and bone marrow (Figure 5B). The same tissue profiling has been observed with a second probe. However, under these conditions, the four other transcripts were not detected in any of the tested tissues (*i.e.*, no signal was detected before 35 cycles of real-time PCR). Although the
10 cloning was performed using transcripts derived from lung, *ABCA13* expression was not detected in lung. It is noted, however, that cloning was performed with cDNA derived from polyA⁺ RNA, whereas the real-time-PCR experiments used cDNA derived from total RNA.

Initial database mining did not reveal any published EST corresponding to the *ABCA13* gene. This analysis was repeated in 2002 and revealed a relatively low number of ESTs compared to
15 the expected length of the gene transcripts. Almost all of the ESTs match with the 3'UTR of the major transcript (called Transcript 1) and therefore could not be used to differentiate the expression pattern of the other transcripts. Like other *ABCA* subfamily genes, *ABCA13* presents a restricted tissue distribution.

The expression of the gene was also examined in a panel of tumor cell lines and was found
20 to have the highest expression in the SR leukemia, SNB-19 CNS tumor, and DU-145 prostate tumor cell lines (Table 15).

Table 15. Expression of *ABCA13* in tumor cell lines

CELL LINE	TUMOR	ABCA12	ABCA13
SR	Leukemia		87
SNB-19	CNS	7.8	18
U251	CNS		13
SF-539	CNS		9.7
MCF7	Breast	43	
T-47D	Breast	29	
COLO 205	Colon	13	
CAKI-1	Renal	6.3	
RXF 393	Renal	12	
TK-10	Renal	7.8	
DU-145	Prostate		28

25 The majority of the samples showed low expression level, and these samples were averaged to determine a background level of expression. Cell lines with at least 5-fold expression over background (an average of the sixty assayed) are shown. The expression of the closely related *ABCA12* gene is shown for comparison. *ABCA12* is also highly expressed in the SWB-19 cell line, but not in any of the others.

-66-

To detect the expression of mouse *Abca13*, PCR was performed on the panel of tissue-specific cDNAs with two pairs of primers (Figure 5B). One pair of primers was used to amplify the fragment from exons 1, 2 and 3 (which encode the first transmembrane segment of the N-terminal membrane-spanning domain). The second pair of primers was designed to amplify the fragment from large exons 17 and 18. Both reactions detected expression in kidney; using primers specific to exons 1 and 3, very weak signal was detected also in the skeletal muscle.

The public databases contain 12 ESTs matching predicted mouse *Abca13* cDNA. They derive only from two tissues: nine from kidney and three from retina. These data provide independent support to our observation that mouse *Abca13* is primarily expressed in kidney. Six ESTs are spliced and originate from regions that are different from those that were used in the research described herein. Six ESTs are unspliced and originate from the 3' end of the cDNA. Retina was not included in the current analysis.

The expression level of *ABCA13* is very low and could be detected by real-time-PCR in a small number of tissues and cell lines. In mouse, expression was detected only in kidney and in very low level also in skeletal muscle. Consistent with this is the observation that there are very few expressed sequence tags that represent portions of the *ABCA13* transcript. However it cannot currently be rule out that higher expression of one or more *ABCA13* transcripts is present in specific cell types that has not yet examined.

The expression pattern of *ABCA13* was also studied on a human Multiple Tissue Expression Array that contains polyA+ RNA from 76 different human normal and cancer tissues plus several control DNAs. Although the signal was clearly visible from human control DNA, no signal was detected from any RNAs on the array, even after prolonged exposure. No signal has been observed on human northern blots either, even after prolonged exposure. The tissues tested included kidney, skeletal muscle, lung, trachea, bone marrow, and testis. It is believed that these results are because blotting conditions do not permit good transfer of long transcripts.

Example 11

Northern Blot Analysis of Human *ABCA13*.

The profile of expression of *ABCA13* polynucleotides can be determined using art-known techniques, such as Northern blot analysis, which has been described for instance by Sambrook *et al.* (1989, Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). This example provides representative methods.

Preparation of probe

PCR products are gel-purified using Qiaquick® column (Qiagen). 10-20 ng of purified PCR product are radiolabeled with [γ^{32} P]dCTP (Amersham; 6000 Ci/mmol, 10 mCi/ml) by the random priming method (Rediprime kit; Amersham) according to the manufacturer's protocol. Unincorporated radioactive nucleotides are separated from the labeled probe by filtration on a G50

-67-

microcolumn (Pharmacia). The probe is competed with 50 µg of denatured human COT1 DNA for two hours at 65°C.

Northern blots (MTN blot, Clontech, Palo Alto, USA) containing 2 µg of poly(A⁺) mRNA per lane from various adult and/or fetal tissues of a subject are hybridized with randomly radiolabeled
5 with U17-L8 specific probe (2x10⁶ cpm/ml hybridization solution) in 50% formamide hybridization solution overnight at 65°C. The filters were washed in 2x SSC for 30 minutes at room temperature, twice in 2x SSC-0.1% SDS for 10 minutes at 65°C and twice in 1x SSC-0.1% SDS for 10 minutes at 65°C. The filters are analyzed using the StormTM blot imaging system (Molecular Dynamics, Sunnyvale, USA) by overnight exposure.

10

Example 12

Identification of a Causal Gene for a Disease Linked to ABCA13.

Among mutations identified in ABCA13 (for instance, mutations identified according to methods described herein), all those associated with the disease phenotype are capable of being
15 causal. These results are validated by sequencing the gene in the affected individuals and their relations (whose DNA is available).

Moreover, Northern blot or RT-PCR analysis, performed for instance according to the methods described herein, using RNA specific to affected or non-affected individuals makes it possible to detect notable variations in the level of expression of the gene studied, in particular the
20 absence of transcription of the gene.

Example 13

Validation of the Expression of Human ABCA13 cDNA.

Polyclonal antibodies specific for a human ABCA13 polypeptide may be prepared as
25 described herein in rabbits and chicks by injecting a synthetic polypeptide fragment derived from an ABCA13 protein, including all or part of an amino acid sequence as shown in SEQ ID NO: 2. These polyclonal antibodies are used to detect and/or quantify the expression of the ABCA13 gene in cells and animal models by immunoblotting and/or immunodetection.

30

Example 14

Expression *in vivo* of the ABCA13 Gene in Animal Models.

An appropriate volume (*e.g.*, 100 to 300 µl) of a medium containing the purified recombinant adenovirus (pABCA-AdV or pLucif-AdV) containing from 10⁸ to 10⁹ lysis plaque-forming units (pfu) are infused into the Saphenous vein of mice (C57BL/6, both control mice and
35 models of transgenic or knock-out mice) on day 0 of the experiment.

The evaluation of the physiological role of the ABCA13 protein in the transport of cholesterol or inflammatory lipid substances is carried out by determining the total quantity of cholesterol or appropriate inflammatory lipid substances before (day zero) and after (days 2, 4, 7, 10, 14) the administration of the adenovirus.

-68-

Kinetic studies with the aid of radioactively labeled products are carried out on day 5 after the administration of the vectors rLucif-AdV and rABCA-AdV in order to evaluate the effect of the expression or function of ABCA13 on the transport of cholesterol and inflammatory lipid substances.

Furthermore, transgenic mice and rabbits over-expressing gene may be produced, in
5 accordance with the teaching of Vaisman (*J. Biol. Chem.* 270(20): 12269-12275, 1995) and Hoeg (*Proc. Nat. Acad. Sci. USA.* 93(21):11448-11453, 1996) using constructs containing the human ABCA13 cDNA under the control of endogenous promoters such as ABCA13 or CMV or apoE.

The evaluation of the long-term effect of the expression of ABCA13 on the kinetics of the lipids involved in the mediation of the inflammation may be carried out as described above.

10

Embodiments of this disclosure provide several ABCA13 proteins and nucleic acid molecules, and methods of isolating, making, and using these molecules. Further embodiments provide methods for ameliorating, treating, detecting, prognosing and diagnosing diseases related to expression of ABCA13. It will be apparent that the precise details of the methods described may be
15 varied or modified without departing from the spirit of the described invention. We claim all such modifications and variations that fall within the scope and spirit of the claims below.

CLAIMS

1. A substantially purified human ABCA13 protein, comprising an amino acid sequence having at least 85% sequence identity to amino acid residues 2278 through 4942 of SEQ ID NO: 2.
2. The substantially purified human ABCA13 protein of claim 1, which protein has ABCA13 biological activity.
3. The substantially purified human ABCA13 protein of claim 1, comprising an amino acid sequence having at least 90% sequence identity to residues 2278 through 4942 of SEQ ID NO: 2.
4. The substantially purified human ABCA13 protein of claim 1, comprising an amino acid sequence having at least 95% sequence identity to residues 2278 through 4942 of SEQ ID NO: 2.
5. The substantially purified human ABCA13 protein of claim 1, comprising residues 2278 through 4942 of SEQ ID NO: 2, or a conservative substitution thereof.
6. The substantially purified human ABCA13 protein of claim 1, consisting essentially of the amino acid sequence as set forth in SEQ ID NO: 2, 19, 21, 23, 25, 27, or 29.
7. The substantially purified human ABCA13 protein of claim 1, having a sequence consisting essentially of the amino acid sequence as set forth in SEQ ID NO: 2, 19, 21, 23, 25, 27, or 29.
8. A substantially purified nucleic acid molecule that encodes the protein of claim 1.
9. The substantially purified nucleic acid molecule of claim 8, comprising a nucleotide sequence having at least 85% sequence identity to the nucleotide sequence as set forth as SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.
10. The substantially purified nucleic acid molecule of claim 8, comprising a nucleotide sequence having at least 90% sequence identity to the nucleotide sequence as set forth as SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.
11. The substantially purified nucleic acid molecule of claim 8, comprising a nucleotide sequence having at least 95% sequence identity to the nucleotide sequence as set forth as SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.

-70-

12. The substantially purified nucleic acid molecule of claim 8, consisting essentially of the nucleotide sequence as set forth as SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.

13. A recombinant polynucleotide encoding the protein of claim 1, wherein the
5 recombinant polynucleotide has a sequence comprising:
- (a) nucleotide residues 26 to 15202 of the nucleotide sequence as set forth as SEQ ID NO: 1;
 - (b) nucleotide residues 2209 to 10491 of the nucleotide sequence as set forth as SEQ ID NO: 18;
 - 10 (c) nucleotide residues 26 to 2179 of the nucleotide sequence as set forth as SEQ ID NO: 20;
 - (d) nucleotide residues 26 to 14902 of the nucleotide sequence as set forth as SEQ ID NO: 22;
 - (e) nucleotide residues 2209 to 10191 of the nucleotide sequence as set forth as SEQ ID NO: 24;
 - 15 (f) nucleotide residues 26 to 14884 of the nucleotide sequence as set forth as SEQ ID NO: 26;
 - (g) nucleotide residues 2209 to 10173 of the nucleotide sequence as set forth as SEQ ID NO: 28; or
 - (h) a nucleotide sequence having 85% sequence identity to any one of (a), (b), (c), (d), (e),
20 (f), or (g).

14. The recombinant polynucleotide of claim 13, wherein the recombinant polynucleotide has a sequence comprising the nucleotide sequence as set forth as SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.
25

15. A recombinant nucleic acid molecule comprising a promoter sequence operably linked to the nucleic acid molecule of claim 7.

16. A recombinant nucleic acid molecule according to claim 14, wherein the
30 recombinant polynucleotide is in antisense orientation relative to the promoter sequence.

17. A recombinant vector comprising the recombinant nucleic acid molecule of claim 14.
14.

- 35 18. A cell transformed with the recombinant nucleic acid molecule of claim 14.

19. A transgenic non-human animal, comprising the cell of claim 17.

20. A specific binding agent that specifically binds an epitope of the protein of claim 1.

-71-

21. The binding agent of claim 20, wherein the agent specifically binds an epitope of a protein having an amino acid sequence as set forth in SEQ ID NO: 2, 19, 21, 23, 25, 27, 29, or a conservative variation thereof.
- 5 22. The specific binding agent of claim 20, wherein the agent is an antibody.
23. A method of detecting a biological condition of a subject associated with altered expression of an *ABCA13* nucleic acid, comprising detecting in a biological sample:
- 10 a mutation in the *ABCA13* nucleic acid that is associated with disease;
altered expression of an *ABCA13* nucleic acid; or
expression of a mutant *ABCA13* nucleic acid,
wherein the mutation, altered expression, or expression of the mutant *ABCA13* nucleic acid indicates that the subject has the biological condition.
- 15 24. The method of claim 23, wherein the method comprises evaluating defective extra- or intra-cellular transport.
25. The method of claim 23, wherein the altered expression of the *ABCA13* nucleic acid comprises an alteration in a cellular level of *ABCA13* nucleic acid or *ABCA13* protein, in comparison to the level detected in a control sample.
- 20 26. The method of claim 23, wherein detecting comprises Southern blot analysis, quantitative polymerase chain reaction or semi-quantitative polymerase chain reaction.
- 25 27. The method of claim 23, wherein detecting comprises sequencing, chemical cleavage, denaturing gradient gel electrophoresis, or hybridization with allele specific oligonucleotides.
- 30 28. The method of claim 23, wherein the *ABCA13* nucleic acid is *in vitro* amplified using at least one oligonucleotide primer having a sequence identical to at least 10 contiguous nucleotides of the nucleic acid sequence as set forth in SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.
- 35 29. The method of claim 28, wherein the oligonucleotide primer comprises the nucleotide sequence as set forth in SEQ ID NOs: 3-15, SEQ ID NOs: 73-75, or a conservative substitution of any one thereof.
30. The method of claim 23, wherein the biological condition comprises hypercholesterolemia, drug resistance, retinal degeneration, or neurological disease.

31. The method of claim 23, wherein the biological condition comprises Shwachmann-Diamond syndrome, an inherited disorder of the pancreas, or T cell tumorigenesis.

5 32. The method of claim 23, wherein the method is used to detect a chemotherapy resistant cell.

33. The method of claim 23, wherein the altered expression of the *ABCA13* nucleic acid comprises an increased or decreased expression of *ABCA13* in a subject as compared to a
10 control.

34. The method of claim 23, comprising reacting at least one *ABCA13* molecule contained in a biological sample from the subject with a reagent comprising an *ABCA13* specific binding agent to form an *ABCA13*:agent complex.
15

35. The method of claim 34, wherein the *ABCA13* molecule is encoded by a nucleic acid having a nucleotide sequence as shown in SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28.

36. The method of claim 34, wherein the *ABCA13*:agent complex is detected by
20 nucleotide hybridization.

37. The method of claim 34, wherein the *ABCA13*:agent complex is detected by Western blot assay.

25 38. The method of claim 37, wherein the Western blot assay uses an antibody generated against a nucleotide binding or transmembrane domain of the protein encoded by the amino acid sequence as shown in SEQ ID NO: 2, 19, 21, 23, 25, 27, 29, or a conservative substitution thereof.

30 39. The method of claim 34, wherein the complexes are detected by ELISA.

40. The method of claim 38 or 39, wherein the antibody is a monoclonal antibody.

41. The antibody of claim 40, which recognizes an antigenic peptide comprising a
35 sequence within residues 2278 through 4942 of SEQ ID NO: 2.

42. The method of claim 41, wherein the sample comprises blood, a blood product, urine, saliva, a tissue biopsy, a surgical specimen, an amniocentesis sample, or autopsy material.

-73-

43. A method of screening for an agent that modulates ABCA13 transporter activity, the method comprising:
- 5 transfecting a cell with an expression vector, wherein the expression vector comprises a nucleic acid molecule having a sequence as set forth in SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28, or a conservative substitution thereof, operably linked to a promoter sequence;
- contacting the cell with a test agent; and
- detecting a change in the level of expression of the ABCA13 protein,
- 10 wherein a change in the level indicates that the test agent modulates the expression of the ABCA13 transporter protein.
44. A method of detecting an ABCA13 transporter protein in a biological sample, comprising:
- 15 amplifying in the biological sample nucleotide residues as set forth in SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28, or a conservative substitution thereof, with two or more oligonucleotide primers that specifically bind the nucleotide residues as set forth in SEQ ID NO: 1, 18, 20, 22, 24, 26, or 28; and
- detecting a level of an amplified product, thereby detecting the ABCA13
- 20 transporter protein.
45. An *in vitro* assay kit, comprising:
- a container comprising an ABCA13 protein specific binding agent; and
- instructions for using the kit, the instructions indicating steps for performing a
- 25 method to detect the level of expression of an ABCA13 nucleic acid in the sample; and analyzing data generated by the method,
- wherein the instructions indicate that altered expression of ABCA13 in the sample indicates that the individual has a biological condition associated with ABCA13 expression.
- 30 46. The kit of claim 45, wherein the agent is capable of specifically binding to an epitope within:
- (a) the amino acid sequence shown in SEQ ID NO: 2, 19; 21, 23, 25, 27, or 29;
- (b) the amino acid sequence shown in amino acid residues 2278 through 4942 of SEQ ID NO: 2; or
- 35 (c) a conservative variant of (a) or (b).
47. The kit of claim 46, wherein the altered ABCA13 expression results in altered extra- or intracellular transport.

-74-

48. The kit of claim 46, wherein the ABCA13 protein binding agent is an antibody.

49. An *in vitro* amplification assay kit, comprising:

a first container comprising an *in vitro* amplification primer that specifically

5 amplifies the nucleic acid that encodes ABCA13 protein;

a second container comprising a size marker, the size marker being the expected
size of amplified DNA if the nucleic acid that encodes ABCA13 protein is present in the sample; and

instructions for using the kit, wherein the instructions indicate steps for performing
a method to detect and/or quantify the nucleic acid that encodes ABCA13 protein in the sample; and

10 analyzing data generated by the method,

wherein the instructions indicate that an altered level of nucleic acid that encodes ABCA13 protein in
the sample indicates that the individual has a biological condition associated with ABCA13
expression.

FIGURE 1A

132286 - ABCA13

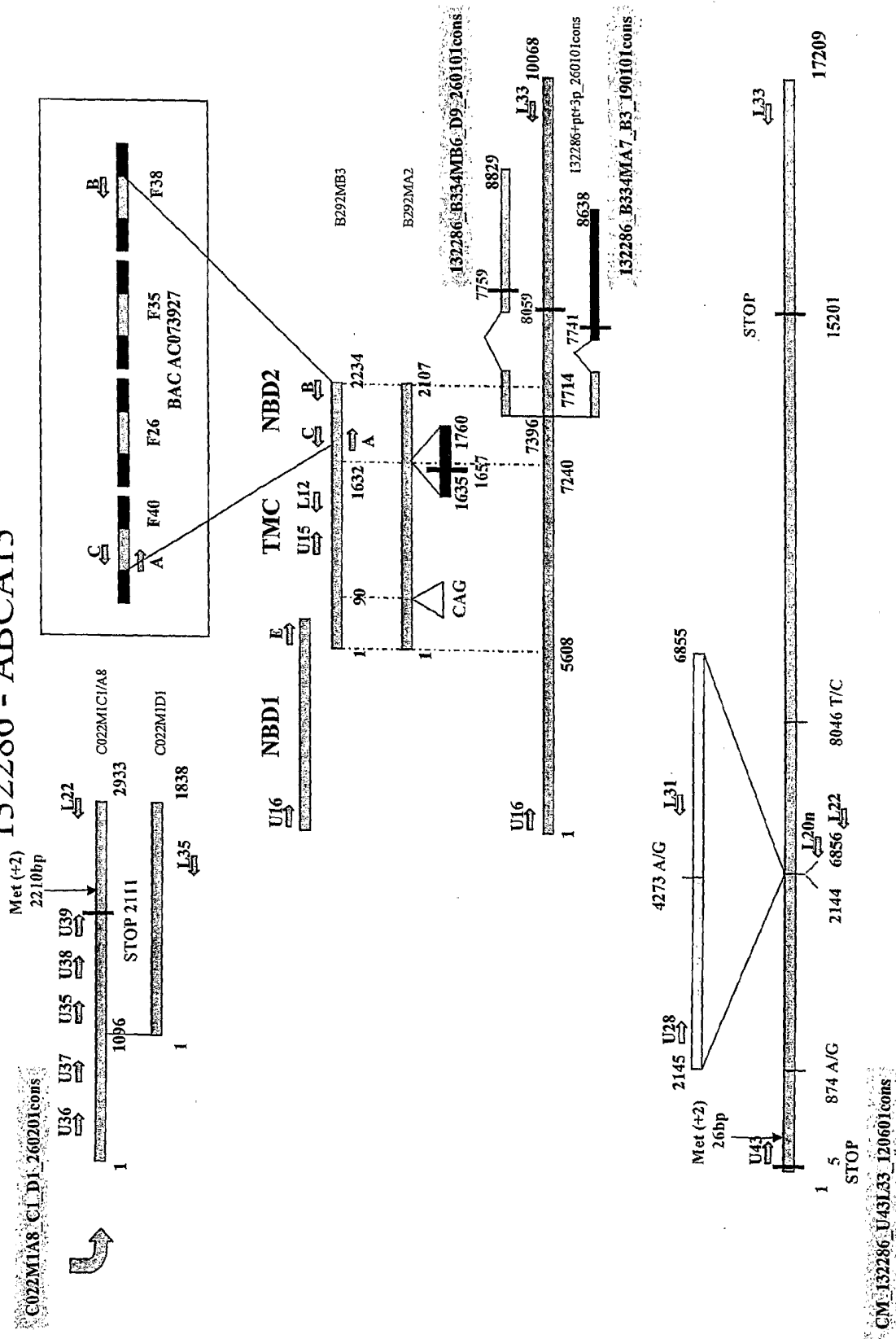


FIGURE 1C

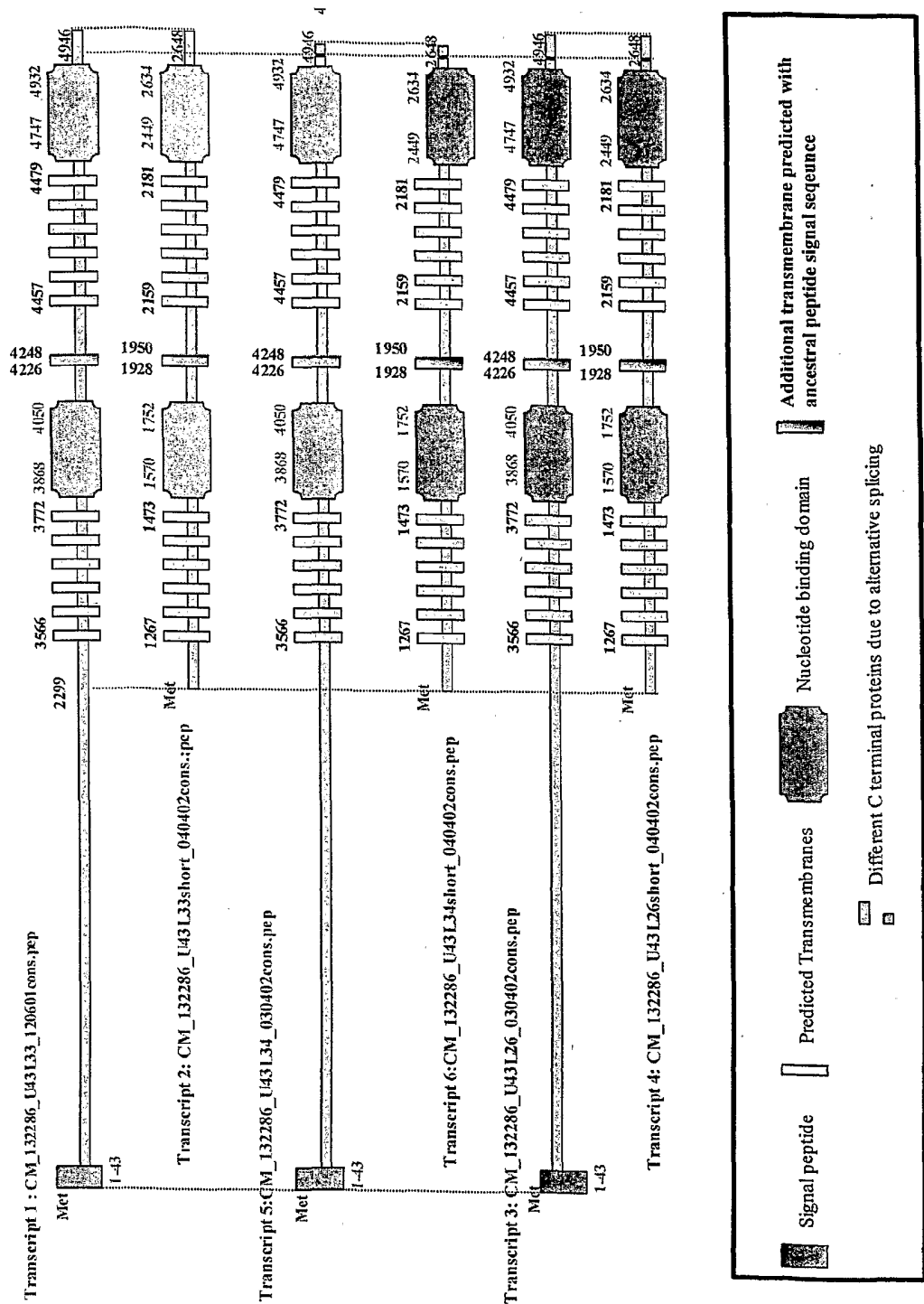


FIGURE 1D

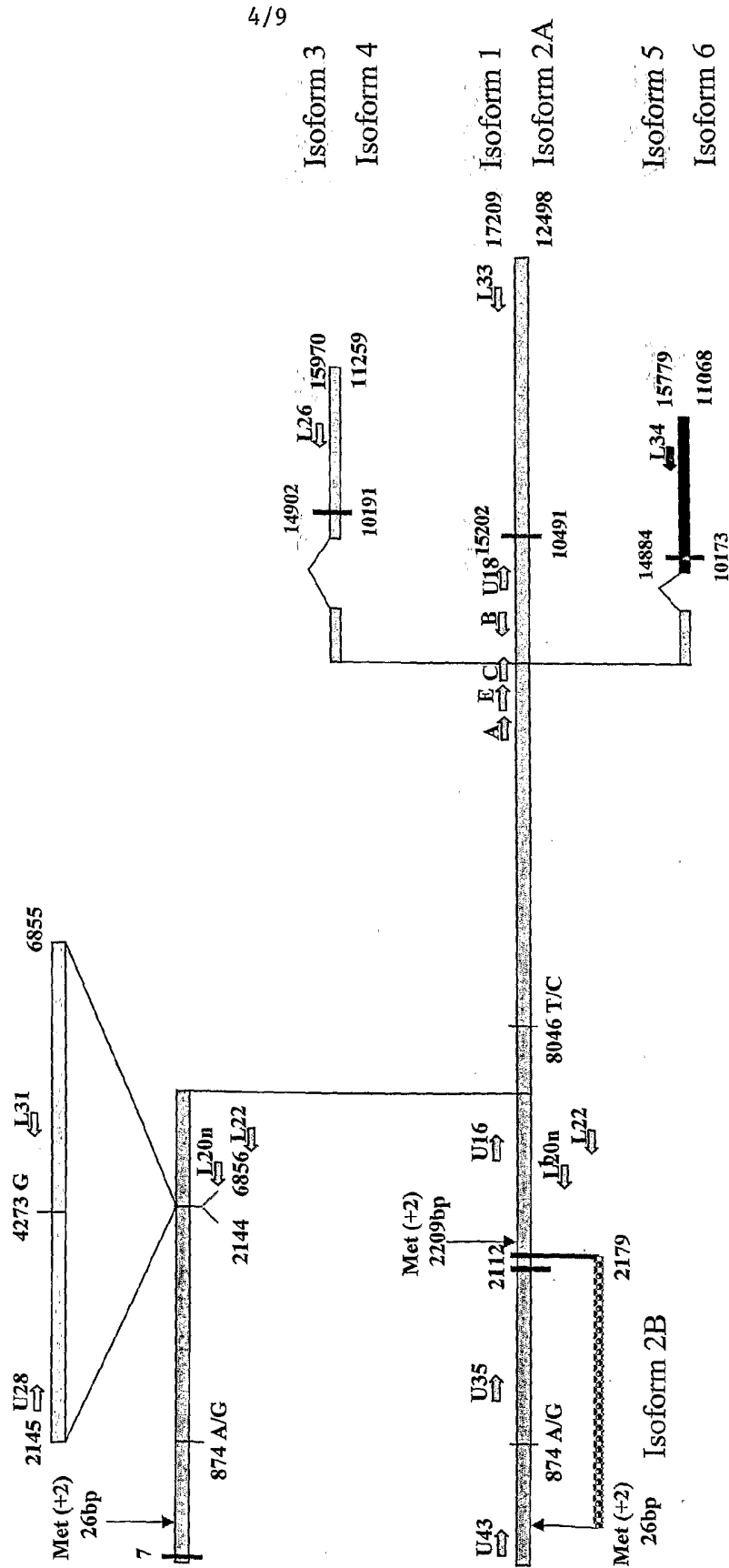


FIGURE 2

5/9

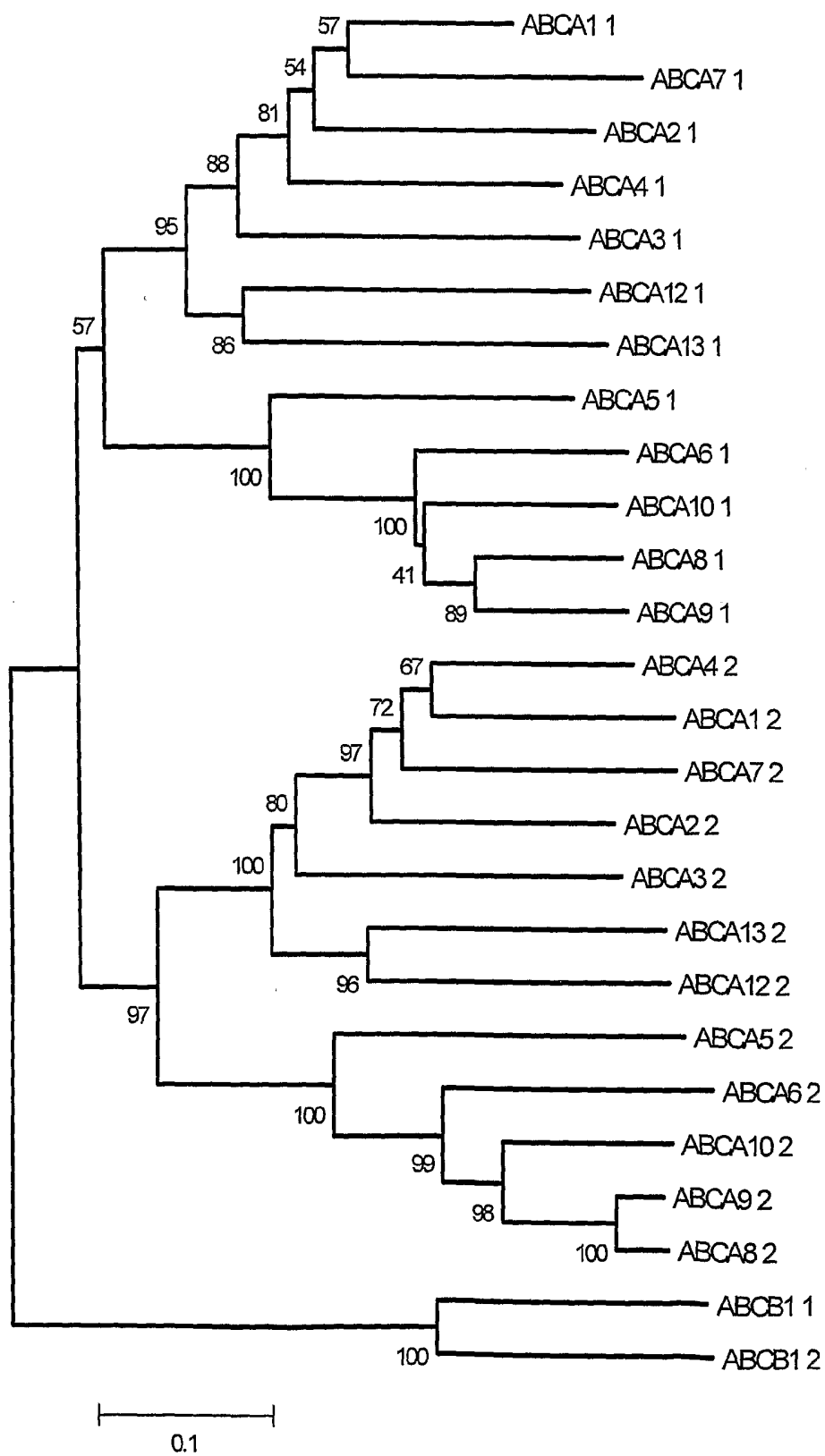


Figure 3

	10	20	30	40	50	60
1						
ABCA1	MA-----	CWPQLRL	LWKNTL	FRRRQT	CQLLLE	VAVPLF
ABCA2	MG-----	FLHQQL	LLWKNV	TLKRRS	PWVLA	FEFIF
ABCA3	MA-----	VLRLAL	LLWKNY	TLQKRK	VIVT	LELFL
ABCA4	MG-----	FVRQIQ	LLWKNW	TLRKQK	IRFV	VELV
ABCA5	MSTAI	REVG	VWRQ	TRTLL	KNYL	IKCR
ABCA6	MNMKQ	--	KSVY	QQTK	ALLC	KNFL
ABCA7	MA-----	FWTQL	MLLL	WKNF	MYRR	RRQP
ABCA8	MRKRK	--	ISVC	QQTW	ALLC	KNFL
ABCA9	MSKRR	--	MSVG	QQTW	ALLC	KNFL
ABCA12	MAS-----	LFHQQL	ILVW	KNWL	GVKR	QPLW
ABCA13	MGH-----	AGCQF	KALL	WKNW	LCRL	NPVL

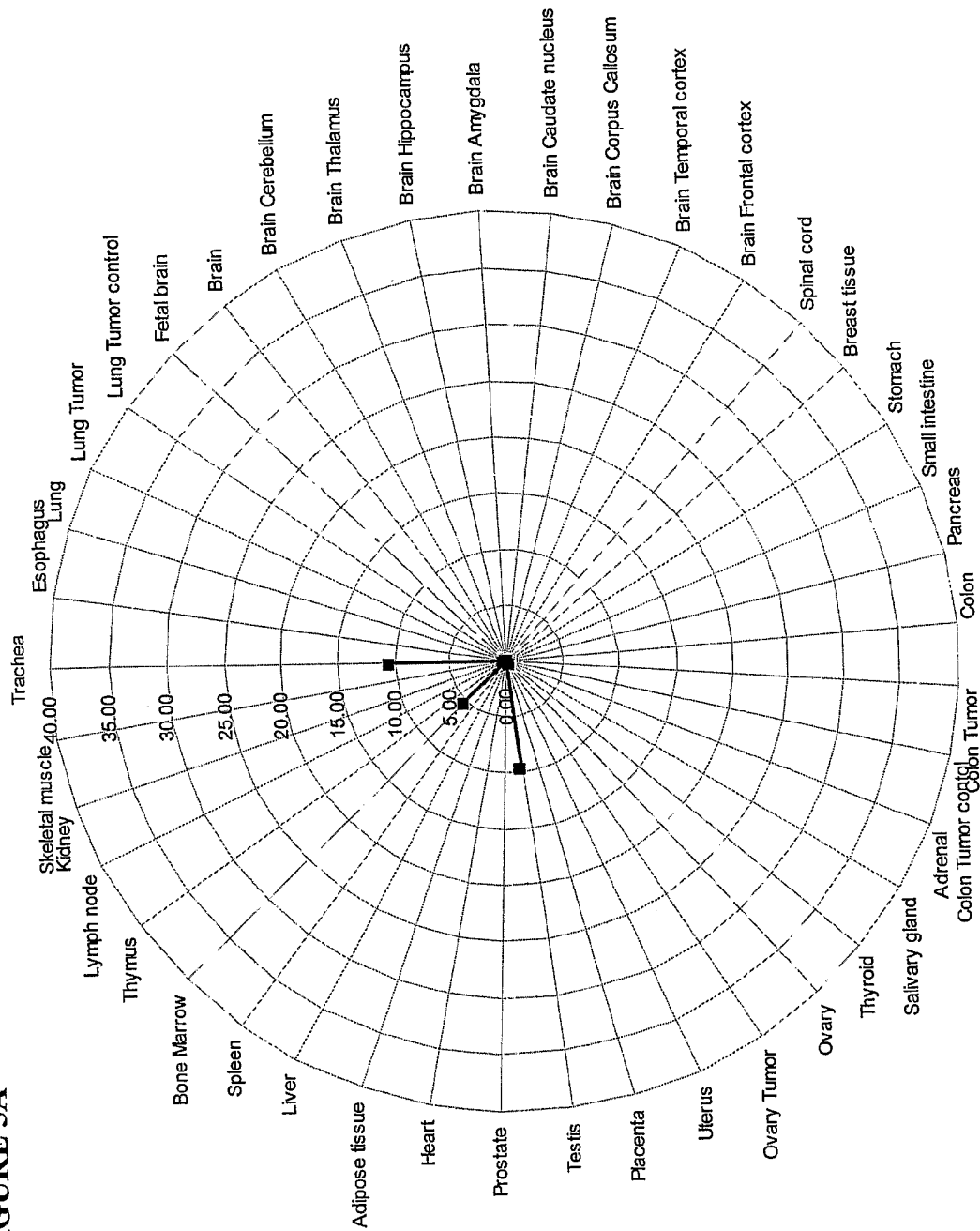
6/9

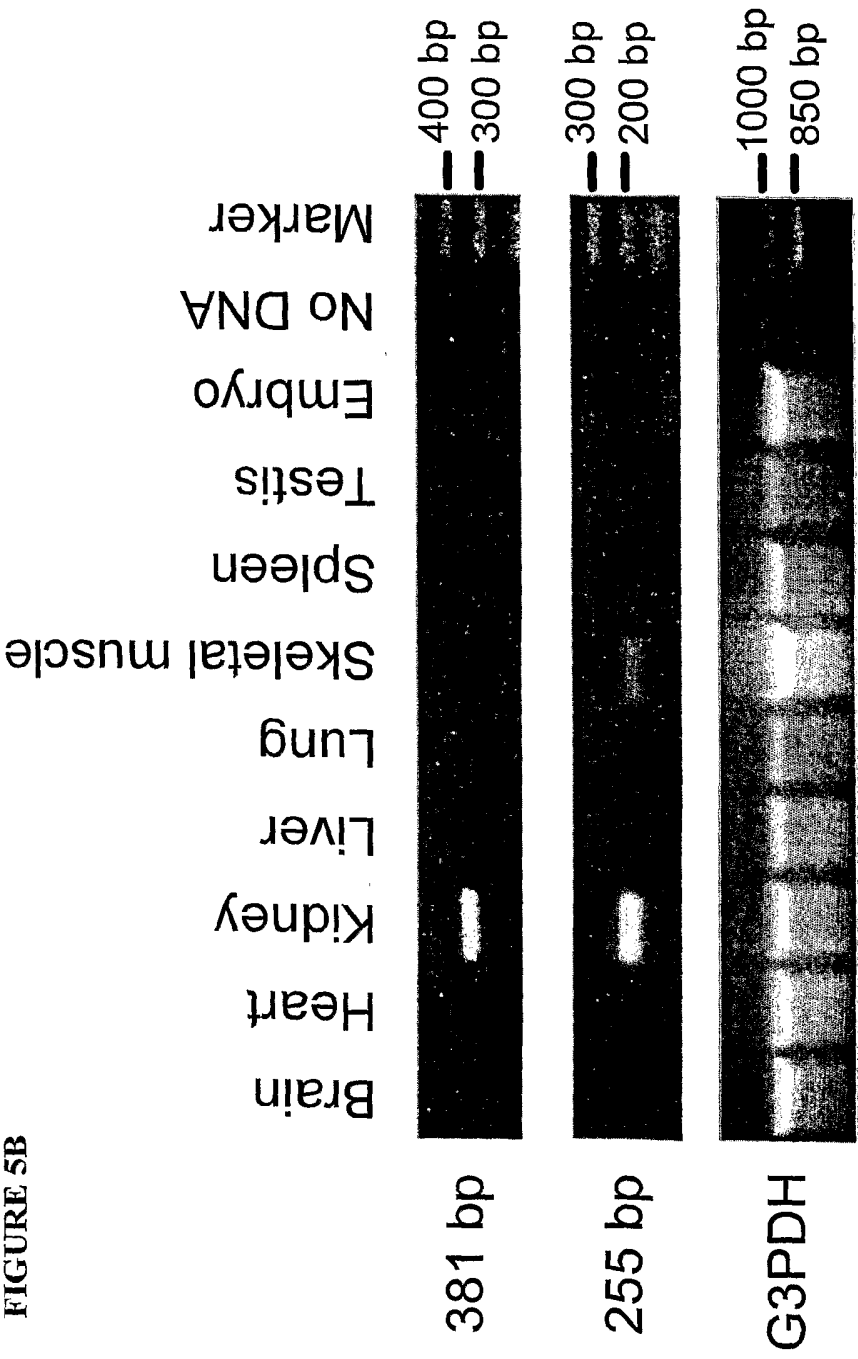
1320	1330	1340	1350	1360	1370	1380
ABCA1	SYQVKG	WKL	TQQQ	FVALL	WKRL	LIARR
ABCA2	SRKLDG	GWL	KVRQ	FHGL	LVKR	FHCARR
ABCA3	VKLNTG	LALH	CQQF	WAMF	LKKA	AYS
ABCA4	PQLNTG	TQLV	LQHV	QALL	VKR	FQHT
ABCA5	ASLVST	MSLW	KQOM	YTI	AKF	HEFT
ABCA6	QTAVSD	MGLW	RMQV	FAMAR	LRFL	KRQK
ABCA7	VGRVQG	WAL	TRQ	QALL	KRFL	LARR
ABCA8	RKTIGG	VAL	WRQ	QICAI	ARV	RLKL
ABCA9	RKTI	SG	VAL	WRQ	QVCAI	AKV
ABCA10	RKAV	SSA	L	WRQ	QIY	AVAT
ABCA12	GERLDG	FGL	LLK	IMAIL	IKR	FHH
ABCA13	PATVQ	GVQ	LLR	AQVAA	ILARR	LRRT

Figure 4

MGHAGCOPKALLKWNKWLRLERNEVFLAEEFWPCILFVILLTVLRFQBPBPRYRDI CYLQPRDLPSCGVIPFVQVLLCNGTSGRCRNFSEYEGSMHHEHFLSRFQTAAADPKKVN 110
 NLAFLKEIQDLAEIEHGMMDKAKNKLHLSVNTSPSSYSGSFFNDLNTKEEVILKLESHQQPHIWDIFLLPRALTHSHVDGMDVAVNLQTLINLSLIEDLD 220
 WPLNQTFQVSELVNTVITSTLTFQOHHGVAETEPYHLSMQNIWDPQKVQYDLKSGFQFDDLTHTQIILNSSAELKEIETDTISLEKMCVSVLSTSEDEAEKWHGVGG 330
 CHPKWSEARNYLVHAVSWLRVYQQVQWQGSLLQTLTGMSHIEALRNQFEESKPKWKEALHTALLNDLSADGPDNHTFPKILQHLWKLLQSLQNLQWPWA 440
 LKRFQLDQALRNAIAQNLHFQVEVLICLETSADEKWFELNQLKLEKDVFFWELKQMLAKNAVCPNGRSEKEVEFLPPGNSIWSGGLQGLICYNCSSETSVLNKLLGSV 550
 EDARILQAVITWKNMSVLIPEEYLDWQELMQLSASLSCTRLIFLLGADSPENDVSSDCHQLVSTVIFHTLEKTQFFELQAYYKAFKFRIRKTCVAVVNMQ 660
 ESFQNRLLAPPEESPCFENMDKMI SDNYFQFLNNLKSPTASISRALNFTKHLMEKKLHTLEDEOMNLFSPVEFFELKLLPNLFLSSVSPHSLPSTEDILNI 770
 SSLWTNHLKSLKRDPAQDKLLEFGNEVWKMQITIGSHWIRKEPNLLRFIEILFELINPKLLELWAYGISKGRKRAKLENNFLLNFSPENILISTSFNSQLFHSQ 880
 WPKSPAMNDFVRLSEAITSLHEFGELEQEI SEAINTVIAIRNASDLFSALSEQKQEVKIIITHILNFQDKDSALLQYSSFYRIYIILINISQSGSSLTFTQ 990
 ISKHILDIKQFNQNTSKAFALFKTAELVLCISNVSYCQQLLSFNFLELQAQSPMSTEGELEVHTTLGKQLLIIDDFRISLFQYMQFFNSVEDLDNKKCL 1100
 ISDNKHTSSVNSTSESSVFPLAQILFSLNLSANVVENKEMSIHCTVSWLQMWETIWIETISQLEKFDNMVFTSHHGFQTLDELEDDVKVSKCQGLPHTNVARIIL 1210
 NLFKNVTQANDFHNWEDFLDRDFVALGNALVSKKLNLEQVEKSLFTMEALHQLKTFPFNESTREFLNSELEVFIEFSSSEYIVRNLSINDFLSNLTNYGEKF 1320
 ENIITELREAI VFLRNVSHDRDLFCADIFQNVTECILEDGFLYVNTSQRMRLILDTLNSTFESSENTISLKGCIWLDVINHLVLLSNSSFSQGRLOLNLGNFEDIENK 1430
 MNSILKIVTWLNKKPLCSSNGSHNCVNIYLLKDVDFLNI VLTVEKEKPKRFEILLALNDSTKQVRMSINNLTTDFDASQSNWRYFTLILRPIEMSDIENQF 1540
 QNIWHLITLGKEFQKLVKGIYENILENNS SKTENLLNIFATSPKEKDVNSVGNISYHLASYLAFSLSHDLQNSPKIISPEIMKATGLGIQLIRDVFNLSMPVHHTS 1650
 PQNAGYMQALKKVTYSVRTKADIDLVDQLEQVSVNLMDFERNISVCGTGNLVNVLGLMEKFAADSHSNVNVHLLQSLRPPKDVDAVDVYVYVLPHAVKLLQGV 1760
 PGKNIETGLKDVYFTLLHGITISNITKEDFAIVIKILLDTIELVSDKPDIIISALACFPVVMWVWNHTNSGFRQNSKIDPCNVHCLMSSSFYKGVASILDHFLSPQGED 1870
 SPCSNESRMEITRKVVCIIHELVDWNSILLELSEVFNHVNISLVKTQKFWFKIIPFVPPSINOIRDSISELCPGSIKQVALQIIEKLNKVNFTKVTSGENILDKLSL 1980
 NKILININEDTETSQVNISSNLERTVQLISEDWLEKSTHNLISLMMLOQNAVNTGSSLEALSFTIEKSETPYNFEELWPKFQOIMKDLTQDFRIRHLLSEMKNKISIN 2090
 SMALQITLQFAHETLIDSPSLKTEIIEDEFLVTKNWLQEVANEDYSRMIEVFIPTVNESSTEDIALAKAIAATFWGSLKNSISRAGNEDVAFTHLLNQTLNFSV 2200
 VQLLFENILINILNLAGNSQEAANLNDTLQIMNFINILNMQSETSRKTVLSLSIVDFTEFLKTFSLFLEKSEKISLKLKFKHDAEMSFVPRDKILEI 2310
 LKLDQFTLMIQDRUMNIFSSLEKTIYHLMKSSFLDNGEYFDFHQGLKFMQDLNALLIRETSMKNKTENNIDFTTVSQLEFHVNKSEDLFKNLQDLSALHLVRES 2420
 TEMARLDTILHSPKNDFYALPYTQIEVILANLTDLLFFINNPSPLNRATLETKRLVGAISRASEESHVLPKLEMSGTVHMLLNDLADLRLATSMDSVTKLLKLVK 2530
 KVSCKMSTVEKTHFTSNTKDSVKKFDTLISIMQSVQNLVKEATIKKIDHFEFEKINDLLVFLDLAFEMICVEPYIISNDIFPSMSPSILSYMNOQSKDFDILEEIAE 2640
 FLTSVKMNLDMRGLAVANNETQTFMSDVNLREIEILGCLVFNNTIQTMDLVPNPISTHSGSPQDIKWEITHIEVILFLDKILSQNSTEISFSLKVMVICILEALWKNL 2750
 KKDNNVSVNLMTTHQPNNLKTIETVLEASGKISDIYEGDNKSLYFDTPLSQVVEICEVFQTVRSEAMQLQKVMVVRVLTIVAEPSWTKDLICATLSCQNGIRHLI 2860
 VTSKEERTEKEMIDFPYFSPKPFCELEKYLGLFVLTXYQOIPLTDQSVVEICEVFQTVRSEAMQLQKVMVVRVLTIVAEPSWTKDLICATLSCQNGIRHLI 2970
 LSAIQGVTLAQDHFQIEKINWSPNQNLNCESSKNSLSTLESPKSSLENATQDQTSQPRITVQOHLVMAKSLSEETWSSGNPIMTFLSNFTVTEDEKIDLMKNITKL 3080
 TEELKSSIQISNETHISILEANISHSKVLFSAITVALSGKQCEILHLLLTPEKGEKSWIAAEELCSLPGSKVYSILVLSRNLDRVAFIYKTLMPSEANGLNLSLDIV 3190
 SLSALLAKAQHVFEYLPFEFTFKITALLEITDFQQVSQNVQARSASFQSVKMKVCKDQASFLSDSNMFINLPVRKILEDDKEKFNIPEDSTPFCLIKLYOEILQL 3300
 PNGALVWTLKPLHKGILYVTENTPEINKVIQKANYTFYIVDKLKTSELEMSLFLQSGSGOMFNQLEALRNKFRVNFVENQHLIDVDKLTQYEGCLLDEMFN 3410
 HAGARFRFLGSIIVNLS SCVALNRFOALQSDVILETKAHELQONSFLASIFSNLSLDFNFRSESVKLPHVSYTIRTNVLYSVRTDVVKNPSWKPHQNLPAAGFKY 3520
 NYVAPLODMIERAILVQGOHALEPAAQOAPYCHTSDLFNNVGFPPILMMLTMMVSVASVVRKIVYEQEIQIERMYRMMGVHPVHFLAWPLENMAVLTISSA 3630
 TLAELVLTSGIFAHNSNTFIVFELLDLFGMSVMLSILLSAFTSQANTAALCHSLVYMISEFLYIVLLVHLNQLSEFVNQTFCLLSTTAFQGVFEFTFESQETGIQWNN 3740
 MTQALEGGMTGEGWVCMWLLPDSLYFLCGWYTHSNLIPETGARKPWFYFPPTASVYKSGEIVAEKROQFELSSLEFFENFENPDNKGSSLOQREGELEGASPGVTLVSVTE 3850
 YEGHKA VQDLSLTFYRQITALLGTNGAGKTIIISMLTGHHPPTSGTIIINGKNLQDLSNVRMELGVCPQDILLDNITVREHLLLEFASLKAPOWTKKELHQVNOQL 3960
 QVDLTDQHQRQTRALSGLKRLKLSIAFMSGMSRTVLDPTSGVPCSHSLWDILLKYREGRTIIFTHHLEDEALSDRVAVLQHLRCCGPPFCIKKAYGQGLR 4070
 LTLTRQPSVLEAHDLDKMACVTSILIKIYIPQAFKDSGSSILTYITPKDDIXACIKGLFQALDENLHQLHTGYIGISDTLEEVFLMLQD SNKKSHALGTESELOHNR 4180
 PTGHLGSGYCGSLARPATVQVQALLRAQVAATIAARRLRTTRAGKSLADLLLEVLVALAMLENVRPLATEYPLRLTPPGHYQRAETFFSSGDNLDLTRVLLRKFRD 4290
 QDLPCADLNQKNSSCWRTDPSHPEFQDSGCLCKPNRASAAPYLTNHLGHTLNLSCNMEEYLLASEPRLGWSFGLKIPSEACGNGNISKEPTLAKVWYNQK 4400
 GFHSLPSYLNHNNLNLWQHLPPTVDWROQYITLISHPYGCALLNEDKILSIRTCGVVALCIVGLFSISASIGSSVVDRVIGAKRLQHSIGLGYRMYFTNFLYDMFE 4510
 TLVSVCLCAVAVAFQLTAFRKNLAATAALLSLFGYATIPWMLMSRTSSSDVAFISVLSNFTFGCTMPTITIMPRLLAIIISKANLQNIYDVKWVFTIFFQFCL 4620
 GQGLVELCYNQIKYDILTHNEGLDSVSPFEMNELGWI FVQLASQCTVLLRLVLLHWDLLRPERGHSITQCTVKS KDDVEKEEKRYFEGRNGDILVNLKSHYRF 4730
 FQNIITAVQDLSGLPKSECEGLGVNGAGKSTFKMLNGVSVLTSCHAITRTPMGDAVDLISAGTAGVICYCQQQDAIDELLTGWEHIIYICSLRGPRQCIPEAGDL 4840
 TRLHLEAHADKPATVYSGGTRKRLSTALAVGKPDILLDEPSSGMDPCSKRYLWQTIMKEVREGCAAVLTSHSMECEALCTRLAINVNGSFKCLGSPQHIKNRFGDG 4950
 YTVKVLCKEANQCHTYSDBLKLYFPGLQKQKQHLNLELHYVPKRWGCLADLEKVIENNKTEFLNKHYSINTTLEQVFINFAEQOQTLQSTLDPSTDSHTHLPI 5058

FIGURE 5A





SEQUENCE LISTING

<110> The Government of the United States of America as
 represented by the Secretary of the Department of Health and
 Human Services
 Dean, Michael C.
 Shulenin, Sergey
 Rosier, Marie
 Arnould, Isabelle
 Prades, Catherine
 Denelfe, Patrice
 Anillo, Tarmo
 Triunfol, Marcia L.

<120> ABCA13 NUCLEIC ACIDS AND PROTEINS, AND USES THEREOF

<130> 4239-66326

<150> US 60/454,502
 <151> 2003-03-12

<150> US 60/405,006
 <151> 2002-03-20

<160> 75

<170> PatentIn version 3.2

<210> 1
 <211> 17209
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> CDS
 <222> (26)..(15202)

<400> 1
 ggactgagag cagggagcag caggc atg ggg cat gcc ggg tgc cag ttc aaa 52
 Met Gly His Ala Gly Cys Gln Phe Lys
 1 5

gcc ctg ctg tgg aag aat tgg ctc tgc aga ctc agg aac ccg gtc ctt 100
 Ala Leu Leu Trp Lys Asn Trp Leu Cys Arg Leu Arg Asn Pro Val Leu
 10 15 20 25

ttc ctt gct gaa ttc ttc tgg cct tgt atc ctg ttt gta att ctg aca 148
 Phe Leu Ala Glu Phe Phe Trp Pro Cys Ile Leu Phe Val Ile Leu Thr
 30 35 40

gtt ctt cgt ttt caa gaa cct ccc aga tac aga gac att tgt tat ttg 196
 Val Leu Arg Phe Gln Glu Pro Pro Arg Tyr Arg Asp Ile Cys Tyr Leu
 45 50 55

cag ccc cga gat cta ccc agc tgt ggt gtt atc ccc ttt gtt caa agc 244
 Gln Pro Arg Asp Leu Pro Ser Cys Gly Val Ile Pro Phe Val Gln Ser
 60 65 70

ctt ctt tgt aac act gga tca agg tgt agg aac ttc agc tat gaa ggg 292
 Leu Leu Cys Asn Thr Gly Ser Arg Cys Arg Asn Phe Ser Tyr Glu Gly

75	80	85	
tca atg gag cat cat ttt cgt ttg tct agg ttc caa act gca gct gac Ser Met Glu His His Phe Arg Leu Ser Arg Phe Gln Thr Ala Ala Asp 90 95 100 105			340
ccc aag aaa gtc aac aac ctg gcc ttt tta aaa gag ata caa gac ctg Pro Lys Lys Val Asn Asn Leu Ala Phe Leu Lys Glu Ile Gln Asp Leu 110 115 120			388
gca gag gaa att cat gga atg atg gac aag gca aaa aac tta aaa aga Ala Glu Glu Ile His Gly Met Met Asp Lys Ala Lys Asn Leu Lys Arg 125 130 135			436
ctt tgg gta gaa cga tcc aac act cca gat tct tct tat ggt tcc agt Leu Trp Val Glu Arg Ser Asn Thr Pro Asp Ser Ser Tyr Gly Ser Ser 140 145 150			484
ttt ttt aca atg gat ctc aat aag acc gag gag gta ata ttg aaa ctg Phe Phe Thr Met Asp Leu Asn Lys Thr Glu Glu Val Ile Leu Lys Leu 155 160 165			532
gaa agc ctc cat cag cag cct cat atc tgg gat ttt cta ctt tta ctg Glu Ser Leu His Gln Gln Pro His Ile Trp Asp Phe Leu Leu Leu Leu 170 175 180 185			580
ccg aga cta cac aca agc cat gat cat gtg gaa gat ggc atg gat gtt Pro Arg Leu His Thr Ser His Asp His Val Glu Asp Gly Met Asp Val 190 195 200			628
gca gtg aac ctt ctc cag acc att ttg aat tcc tta ata tcc cta gaa Ala Val Asn Leu Leu Gln Thr Ile Leu Asn Ser Leu Ile Ser Leu Glu 205 210 215			676
gat tta gat tgg ctt cca ctc aac caa act ttt tcc cag gtt tct gaa Asp Leu Asp Trp Leu Pro Leu Asn Gln Thr Phe Ser Gln Val Ser Glu 220 225 230			724
ctt gta ctg aat gtg acc att tcg aca ctg aca ttt ctg cag caa cat Leu Val Leu Asn Val Thr Ile Ser Thr Leu Thr Phe Leu Gln Gln His 235 240 245			772
gga gta gca gtc acc gag cca gtt tac cac ctg tcc atg cag aat ata Gly Val Ala Val Thr Glu Pro Val Tyr His Leu Ser Met Gln Asn Ile 250 255 260 265			820
gtg tgg gat cca cag aaa gtc cag tat gat ctc aaa tcc cag ttt ggc Val Trp Asp Pro Gln Lys Val Gln Tyr Asp Leu Lys Ser Gln Phe Gly 270 275 280			868
ttt grt gat ctt cac acg gaa cag atc ctg aac tct tca gct gaa ctg Phe Xaa Asp Leu His Thr Glu Gln Ile Leu Asn Ser Ser Ala Glu Leu 285 290 295			916
aag gag att ccc aca gac act tcc ttg gag aag atg gtg tgt tca gtc Lys Glu Ile Pro Thr Asp Thr Ser Leu Glu Lys Met Val Cys Ser Val 300 305 310			964
ttg tct agc aca tca gag gat gaa gct gag aaa tgg ggc cac gtt gga Leu Ser Ser Thr Ser Glu Asp Glu Ala Glu Lys Trp Gly His Val Gly 315 320 325			1012

ggc tgc cac cct aag tgg tca gaa gcc aaa aac tat ctt gtc cat gca	1060
Gly Cys His Pro Lys Trp Ser Glu Ala Lys Asn Tyr Leu Val His Ala	
330 335 340 345	
gtc agc tgg ctg cga gtc tac caa cag gtg ttt gtt cag tgg caa cag	1108
Val Ser Trp Leu Arg Val Tyr Gln Gln Val Phe Val Gln Trp Gln Gln	
350 355 360	
ggc agc ctg ctt cag aag aca ctc aca ggc atg ggc cat agt ctg gag	1156
Gly Ser Leu Leu Gln Lys Thr Leu Thr Gly Met Gly His Ser Leu Glu	
365 370 375	
gct ctc agg aat cag ttt gaa gaa gag agc aag ccc tgg aag gtg gtg	1204
Ala Leu Arg Asn Gln Phe Glu Glu Glu Ser Lys Pro Trp Lys Val Val	
380 385 390	
gaa gct ctg cac act gca ctg ctc ctg ctg aat gac agc ttg tca gca	1252
Glu Ala Leu His Thr Ala Leu Leu Leu Leu Asn Asp Ser Leu Ser Ala	
395 400 405	
gat ggc cca aaa gat aat cat aca ttt cca aag ata tta cag cat ctg	1300
Asp Gly Pro Lys Asp Asn His Thr Phe Pro Lys Ile Leu Gln His Leu	
410 415 420 425	
tgg aaa ttg caa agc ttg ctg caa aac ctg ccc cag tgg ccg gca ctg	1348
Trp Lys Leu Gln Ser Leu Leu Gln Asn Leu Pro Gln Trp Pro Ala Leu	
430 435 440	
aag aga ttt ctt cag ctt gat gga gct ctc aga aat gcg ata gct cag	1396
Lys Arg Phe Leu Gln Leu Asp Gly Ala Leu Arg Asn Ala Ile Ala Gln	
445 450 455	
aat tta cat ttt gtc caa gaa gtc ctc att tgc ctg gag aca tca gct	1444
Asn Leu His Phe Val Gln Glu Val Leu Ile Cys Leu Glu Thr Ser Ala	
460 465 470	
aat gat ttt aaa tgg ttt gaa ctt aac caa ttg aaa ctg gaa aag gat	1492
Asn Asp Phe Lys Trp Phe Glu Leu Asn Gln Leu Lys Leu Glu Lys Asp	
475 480 485	
gtg ttc ttt tgg gag ctg aaa cag atg ttg gcg aag aat gct gtc tgc	1540
Val Phe Phe Trp Glu Leu Lys Gln Met Leu Ala Lys Asn Ala Val Cys	
490 495 500 505	
ccg aat ggt cgt ttc tct gag aag gag gtc ttt ttg ccg cct gga aac	1588
Pro Asn Gly Arg Phe Ser Glu Lys Glu Val Phe Leu Pro Pro Gly Asn	
510 515 520	
tcc agc ata tgg ggt ggt ctc cag gga ctg ttg tgc tat tgt aac tcc	1636
Ser Ser Ile Trp Gly Gly Leu Gln Gly Leu Leu Cys Tyr Cys Asn Ser	
525 530 535	
tct gag acg agt gtt tta aac aag cta ctt ggt tca gta gag gat gct	1684
Ser Glu Thr Ser Val Leu Asn Lys Leu Leu Gly Ser Val Glu Asp Ala	
540 545 550	
gat cgt att ttg caa gag gtc att act tgg cac aaa aat atg tca gtt	1732
Asp Arg Ile Leu Gln Glu Val Ile Thr Trp His Lys Asn Met Ser Val	
555 560 565	

tta ata cct gaa gaa tat ttg gac tgg cag gaa ctt gag atg cag ctg Leu Ile Pro Glu Glu Tyr Leu Asp Trp Gln Glu Leu Glu Met Gln Leu 570 575 580 585	1780
tca gaa gca agc ctt tcc tgt act cgg ctc ttc ctg ctg ctg gga gct Ser Glu Ala Ser Leu Ser Cys Thr Arg Leu Phe Leu Leu Leu Gly Ala 590 595 600	1828
gat ccc tct cct gag aat gat gtc ttt tct agt gac tgt aag cac cag Asp Pro Ser Pro Glu Asn Asp Val Phe Ser Ser Asp Cys Lys His Gln 605 610 615	1876
ctt gtc tcc aca gtg ata ttt cat aca ctt gaa aaa aca caa ttt ttc Leu Val Ser Thr Val Ile Phe His Thr Leu Glu Lys Thr Gln Phe Phe 620 625 630	1924
ctg gaa caa gca tat tat tgg aaa gcc ttc aaa aag ttt atc agg aag Leu Glu Gln Ala Tyr Tyr Trp Lys Ala Phe Lys Lys Phe Ile Arg Lys 635 640 645	1972
act tgc gaa gtg gcc caa tat gta aat atg caa gag agt ttc cag aac Thr Cys Glu Val Ala Gln Tyr Val Asn Met Gln Glu Ser Phe Gln Asn 650 655 660 665	2020
aga cta ttg gct ttt cct gag gaa tct cct tgt ttt gaa gaa aac atg Arg Leu Leu Ala Phe Pro Glu Glu Ser Pro Cys Phe Glu Glu Asn Met 670 675 680	2068
gat tgg aaa atg atc agt gat aat tat ttt caa ttt ttg aat aac tta Asp Trp Lys Met Ile Ser Asp Asn Tyr Phe Gln Phe Leu Asn Asn Leu 685 690 695	2116
ctc aag tct cca aca gct tcc ata tcc agg gct tta aat ttc aca aag Leu Lys Ser Pro Thr Ala Ser Ile Ser Arg Ala Leu Asn Phe Thr Lys 700 705 710	2164
cac ctt cta atg atg gaa aag aag ttg cac acc ctt gag gat gaa caa His Leu Leu Met Met Glu Lys Lys Leu His Thr Leu Glu Asp Glu Gln 715 720 725	2212
atg aac ttt ctt tta tca ttt gtg gaa ttt ttt gag aaa tta ttg ttg Met Asn Phe Leu Leu Ser Phe Val Glu Phe Phe Glu Lys Leu Leu Leu 730 735 740 745	2260
cct aat ctt ttt gac tcc tcc att gtt ccc agt ttc cac agc ctc cca Pro Asn Leu Phe Asp Ser Ser Ile Val Pro Ser Phe His Ser Leu Pro 750 755 760	2308
tct ctc aca gag gat att ctg aat ata agt tct ctg tgg aca aat cat Ser Leu Thr Glu Asp Ile Leu Asn Ile Ser Ser Leu Trp Thr Asn His 765 770 775	2356
tta aaa agt tta aag aga gac cca tct gcc act gat gct cag aaa ctc Leu Lys Ser Leu Lys Arg Asp Pro Ser Ala Thr Asp Ala Gln Lys Leu 780 785 790	2404
ttg gaa ttt ggc aac gaa gtg att tgg aaa atg cag act ctc gga agt Leu Glu Phe Gly Asn Glu Val Ile Trp Lys Met Gln Thr Leu Gly Ser 795 800 805	2452
cac tgg ata agg aag gaa cca aaa aat ctt ttg aga ttc ata gaa tta	2500

His Trp Ile Arg Lys Glu Pro Lys Asn Leu Leu Arg Phe Ile Glu Leu	
810 815 820 825	
ata ctt ttt gaa att aat ccc aaa tta cta gaa tta tgg gcc tat ggc	2548
Ile Leu Phe Glu Ile Asn Pro Lys Leu Leu Glu Leu Trp Ala Tyr Gly	
830 835 840	
att tca aaa gga aaa aga gct aaa ttg gaa aac ttc ttt aca ctt tta	2596
Ile Ser Lys Gly Lys Arg Ala Lys Leu Glu Asn Phe Phe Thr Leu Leu	
845 850 855	
aat ttt tct gtt cca gaa aat gag att ctg agt aca agt ttt aac ttt	2644
Asn Phe Ser Val Pro Glu Asn Glu Ile Leu Ser Thr Ser Phe Asn Phe	
860 865 870	
tcc cag ttg ttc cat tca gat tgg cct aaa tca cca gct atg aac ata	2692
Ser Gln Leu Phe His Ser Asp Trp Pro Lys Ser Pro Ala Met Asn Ile	
875 880 885	
gat ttt gta cgt tta agt gag gct ata ata act agt ctc cat gaa ttt	2740
Asp Phe Val Arg Leu Ser Glu Ala Ile Ile Thr Ser Leu His Glu Phe	
890 895 900 905	
gga ttt ttg gag cag gaa cag atc tca gaa gct ctg aac aca gtc tac	2788
Gly Phe Leu Glu Gln Glu Gln Ile Ser Glu Ala Leu Asn Thr Val Tyr	
910 915 920	
gct atc agg aat gca tct gat ctt ttc tca gcc ctt tct gaa cca caa	2836
Ala Ile Arg Asn Ala Ser Asp Leu Phe Ser Ala Leu Ser Glu Pro Gln	
925 930 935	
aaa caa gaa gtt gat aaa att ttg act cac ata cac cta aat gtc ttc	2884
Lys Gln Glu Val Asp Lys Ile Leu Thr His Ile His Leu Asn Val Phe	
940 945 950	
cag gac aag gat tca gct tta ctt ctg caa att tat tct tca ttt tac	2932
Gln Asp Lys Asp Ser Ala Leu Leu Leu Gln Ile Tyr Ser Ser Phe Tyr	
955 960 965	
cga tat att tat gaa tta ttg aat att cag agt aga ggc tct tcg ttg	2980
Arg Tyr Ile Tyr Glu Leu Leu Asn Ile Gln Ser Arg Gly Ser Ser Leu	
970 975 980 985	
act ttc ctt aca caa atc tca aaa cac att ttg gat atc ata aaa caa	3028
Thr Phe Leu Thr Gln Ile Ser Lys His Ile Leu Asp Ile Ile Lys Gln	
990 995 1000	
ttt aat ttc caa aac atc agt aaa gca ttt gca ttt tta ttt aag	3073
Phe Asn Phe Gln Asn Ile Ser Lys Ala Phe Ala Phe Leu Phe Lys	
1005 1010 1015	
aca gca gag gtt ctt ggg gga att tct aat gta tct tac tgt cag	3118
Thr Ala Glu Val Leu Gly Gly Ile Ser Asn Val Ser Tyr Cys Gln	
1020 1025 1030	
caa ttg ctt tca att ttt aac ttt ttg gag ctt cag gcc caa tcc	3163
Gln Leu Leu Ser Ile Phe Asn Phe Leu Glu Leu Gln Ala Gln Ser	
1035 1040 1045	
ttc atg tct aca gag ggc caa gaa ctg gaa gtg atc cac act act	3208
Phe Met Ser Thr Glu Gly Gln Glu Leu Glu Val Ile His Thr Thr	

1050	1055	1060	
ttg aca ggc ctc Leu Thr Gly Leu 1065	aaa cag ctg ctc Lys Gln Leu Leu 1070	ata att gat gaa gat ttt Ile Ile Asp Glu Asp Phe 1075	3253 Arg
att tct tta ttt Ile Ser Leu Phe 1080	caa tat atg agc Gln Tyr Met Ser 1085	ttc ttc aac agt tca Phe Phe Asn Ser Ser 1090	3298 Val
gaa gac cta ttg Glu Asp Leu Leu 1095	gat aat aaa tgc Asp Asn Lys Cys 1100	att tcg gac aat aaa Ile Ser Asp Asn Lys 1105	3343 His
att tct tcc gta Ile Ser Ser Val 1110	aat tat tca aca Asn Tyr Ser Thr 1115	gag gag tct tca ttt Glu Glu Ser Ser Phe 1120	3388 Val
ttt cca ttg gca Phe Pro Leu Ala 1125	caa att ttt tca Gln Ile Phe Ser 1130	ctc tca gca aat gtc Leu Ser Ala Asn Val 1135	3433 Ser
gtg ttc aac aag Val Phe Asn Lys 1140	ttt atg tcc att Phe Met Ser Ile 1145	cac tgt acc gtt tca His Cys Thr Val Ser 1150	3478 Leu
caa atg tgg act Gln Met Trp Thr 1155	gaa atc tgg gaa Glu Ile Trp Glu 1160	acc ata tct caa tta Thr Ile Ser Gln Leu 1165	3523 Lys
ttt gac atg aat Phe Asp Met Asn 1170	gtt ttc aca tct Val Phe Thr Ser 1175	cat cat ggt ttc act His His Gly Phe Thr 1180	3568 Gln
ctt ttg gat gaa Leu Leu Asp Glu 1185	ttg gaa gat gat Leu Glu Asp Asp 1190	gtg aaa gtc tct aaa Val Lys Val Ser Lys 1195	3613 Cys
cag ggt ata ctt Gln Gly Ile Leu 1200	ccc acc cat aat Pro Thr His Asn 1205	gct aga ctc ata tta Ala Arg Leu Ile Leu 1210	3658 Asn
ttg ttt aaa aat Leu Phe Lys Asn 1215	gta act caa gcc Val Thr Gln Ala 1220	aat gac ttc cat aat Asn Asp Phe His Asn 1225	3703 Glu
gac ttc ctg gat Asp Phe Leu Asp 1230	ctc agg gat ttt Leu Arg Asp Phe 1235	ttg gta gct tta ggt Leu Val Ala Leu Gly 1240	3748 Ala
tta gtt tca gta Leu Val Ser Val 1245	aaa aaa ctt aac Lys Lys Leu Asn 1250	ttg gag caa gtg gag Leu Glu Gln Val Glu 1255	3793 Ser
ctt ttc acc atg Leu Phe Thr Met 1260	gaa gct gcc ctg Glu Ala Ala Leu 1265	cat cag ttg aag aca His Gln Leu Lys Thr 1270	3838 Pro
ttc aac gaa agt Phe Asn Glu Ser 1275	aca agc aga gag Thr Ser Arg Glu 1280	ttt tta aat tct ctg Phe Leu Asn Ser Leu 1285	3883 Glu

gtt ttc att gag ttt agc agt acc tca gaa tat ata gtc aga aat	3928
Val Phe Ile Glu Phe Ser Ser Thr Ser Glu Tyr Ile Val Arg Asn	
1290 1295 1300	
cta gat tca ata aat gac ttt ctt tca aat aat ctc aca aat tat	3973
Leu Asp Ser Ile Asn Asp Phe Leu Ser Asn Asn Leu Thr Asn Tyr	
1305 1310 1315	
gga gaa aaa ttt gaa aat atc atc act gag cta aga gaa gca ata	4018
Gly Glu Lys Phe Glu Asn Ile Ile Thr Glu Leu Arg Glu Ala Ile	
1320 1325 1330	
gta ttt ctt aga aat gta tca cat gat cga gat ttg ttt tcc tgt	4063
Val Phe Leu Arg Asn Val Ser His Asp Arg Asp Leu Phe Ser Cys	
1335 1340 1345	
gct gat att ttc caa aat gtt act gag tgt att tta gaa gat ggc	4108
Ala Asp Ile Phe Gln Asn Val Thr Glu Cys Ile Leu Glu Asp Gly	
1350 1355 1360	
ttt tta tat gta aat acc tca cag agg atg tta cgt att cta gac	4153
Phe Leu Tyr Val Asn Thr Ser Gln Arg Met Leu Arg Ile Leu Asp	
1365 1370 1375	
acg tta aat tcc aca ttt tcc tct gag aac aca att agc agt ctg	4198
Thr Leu Asn Ser Thr Phe Ser Ser Glu Asn Thr Ile Ser Ser Leu	
1380 1385 1390	
aaa gga tgc att gta tgg tta gat gtc ata aac cat ttg tat ttg	4243
Lys Gly Cys Ile Val Trp Leu Asp Val Ile Asn His Leu Tyr Leu	
1395 1400 1405	
ttg tct aac tcc agt ttt tca caa ggt cgt ctt caa aat att ttg	4288
Leu Ser Asn Ser Ser Phe Ser Gln Gly Arg Leu Gln Asn Ile Leu	
1410 1415 1420	
ggg aat ttc aga gat ata gaa aac aaa atg aac tct ata tta aaa	4333
Gly Asn Phe Arg Asp Ile Glu Asn Lys Met Asn Ser Ile Leu Lys	
1425 1430 1435	
att gta act tgg gtg tta aat ata aaa aaa cct ctt tgt tca tca	4378
Ile Val Thr Trp Val Leu Asn Ile Lys Lys Pro Leu Cys Ser Ser	
1440 1445 1450	
aat ggc tca cat ata aat tgt gtc aat att tac ttg aaa gat gta	4423
Asn Gly Ser His Ile Asn Cys Val Asn Ile Tyr Leu Lys Asp Val	
1455 1460 1465	
act gac ttt cta aat att gta ctt act aca gtc ttt gaa aaa gag	4468
Thr Asp Phe Leu Asn Ile Val Leu Thr Thr Val Phe Glu Lys Glu	
1470 1475 1480	
aag aaa cct aaa ttt gag att tta tta gct ctt tta aat gat tcc	4513
Lys Lys Pro Lys Phe Glu Ile Leu Leu Ala Leu Leu Asn Asp Ser	
1485 1490 1495	
aca aag caa gta agg atg agt atc aac aac tta aca aca gac ttt	4558
Thr Lys Gln Val Arg Met Ser Ile Asn Asn Leu Thr Thr Asp Phe	
1500 1505 1510	

gat ttt gca tct cag tcc aat tgg aga tat ttt act gaa tta att	4603
Asp Phe Ala Ser Gln Ser Asn Trp Arg Tyr Phe Thr Glu Leu Ile	
1515 1520 1525	
cta aga cca ata gaa atg tca gat gaa att cct aat cag ttt caa	4648
Leu Arg Pro Ile Glu Met Ser Asp Glu Ile Pro Asn Gln Phe Gln	
1530 1535 1540	
aat att tgg ctt cat tta ata aca ctg ggg aag gaa ttt cag aag	4693
Asn Ile Trp Leu His Leu Ile Thr Leu Gly Lys Glu Phe Gln Lys	
1545 1550 1555	
ctt gta aaa ggt att tat ttt aac atc ctg gaa aat aat tcc tct	4738
Leu Val Lys Gly Ile Tyr Phe Asn Ile Leu Glu Asn Asn Ser Ser	
1560 1565 1570	
tct aaa act gaa aac ttg tta aac ata ttt gcc acc agt cca aaa	4783
Ser Lys Thr Glu Asn Leu Leu Asn Ile Phe Ala Thr Ser Pro Lys	
1575 1580 1585	
gaa aag gat gta aac agt gta ggc aat tcc att tat cac tta gct	4828
Glu Lys Asp Val Asn Ser Val Gly Asn Ser Ile Tyr His Leu Ala	
1590 1595 1600	
agt tac ctt gcc ttc agc tta tct cat gac ctc caa aat tca cca	4873
Ser Tyr Leu Ala Phe Ser Leu Ser His Asp Leu Gln Asn Ser Pro	
1605 1610 1615	
aaa ata ata att tca cct gaa ata atg aaa gct aca ggt ctt ggt	4918
Lys Ile Ile Ile Ser Pro Glu Ile Met Lys Ala Thr Gly Leu Gly	
1620 1625 1630	
att caa ctg ata agg gat gtg ttc aac tcc tta atg cct gta gtt	4963
Ile Gln Leu Ile Arg Asp Val Phe Asn Ser Leu Met Pro Val Val	
1635 1640 1645	
cat cac act agt cca caa aat gca ggt tat atg caa gct ttg aag	5008
His His Thr Ser Pro Gln Asn Ala Gly Tyr Met Gln Ala Leu Lys	
1650 1655 1660	
aag gta act tct gtc atg cgt acc ctt aag aaa gca gac ata gac	5053
Lys Val Thr Ser Val Met Arg Thr Leu Lys Lys Ala Asp Ile Asp	
1665 1670 1675	
ctt tta gtg gat cag ctt gaa caa gtt agt gta aac cta atg gat	5098
Leu Leu Val Asp Gln Leu Glu Gln Val Ser Val Asn Leu Met Asp	
1680 1685 1690	
ttc ttt aag aat atc agt agt gtg gga act ggc aat tta gtg gtc	5143
Phe Phe Lys Asn Ile Ser Ser Val Gly Thr Gly Asn Leu Val Val	
1695 1700 1705	
aat ttg ctt gtt ggc ttg atg gaa aaa ttt gca gac agc tca cat	5188
Asn Leu Leu Val Gly Leu Met Glu Lys Phe Ala Asp Ser Ser His	
1710 1715 1720	
tct tgg aat gtt aat cat ctg ctg cag ctc tca cgc ctg ttt cct	5233
Ser Trp Asn Val Asn His Leu Leu Gln Leu Ser Arg Leu Phe Pro	
1725 1730 1735	
aaa gat gtt gtg gat gct gtg ata gat gtg tac tat gtg ctt cct	5278

Lys	Asp	Val	Val	Asp	Ala	Val	Ile	Asp	Val	Tyr	Tyr	Val	Leu	Pro	
			1740					1745					1750		
cat	gct	gta	agg	ctc	ctg	cag	gga	gta	cct	ggt	aaa	aac	atc	act	5323
His	Ala	Val	Arg	Leu	Leu	Gln	Gly	Val	Pro	Gly	Lys	Asn	Ile	Thr	
			1755					1760					1765		
gaa	ggc	ctc	aag	gat	gtc	tac	agc	ttc	aca	ctc	ctt	cat	ggc	ata	5368
Glu	Gly	Leu	Lys	Asp	Val	Tyr	Ser	Phe	Thr	Leu	Leu	His	Gly	Ile	
			1770					1775					1780		
acc	att	tca	aat	atc	acc	aag	gaa	gac	ttc	gca	att	gtg	ata	aaa	5413
Thr	Ile	Ser	Asn	Ile	Thr	Lys	Glu	Asp	Phe	Ala	Ile	Val	Ile	Lys	
			1785					1790					1795		
att	ctt	ttg	gat	aca	att	gaa	tta	gta	tca	gat	aag	cca	gat	att	5458
Ile	Leu	Leu	Asp	Thr	Ile	Glu	Leu	Val	Ser	Asp	Lys	Pro	Asp	Ile	
			1800					1805					1810		
att	tca	gag	gct	tta	gct	tgt	ttt	cct	gtg	gtt	tgg	tgc	tgg	aat	5503
Ile	Ser	Glu	Ala	Leu	Ala	Cys	Phe	Pro	Val	Val	Trp	Cys	Trp	Asn	
			1815					1820					1825		
cac	aca	aat	tct	gga	ttt	cgg	cag	aat	tca	aag	ata	gac	ccc	tgc	5548
His	Thr	Asn	Ser	Gly	Phe	Arg	Gln	Asn	Ser	Lys	Ile	Asp	Pro	Cys	
			1830					1835					1840		
aat	gtc	cat	ggg	ctc	atg	tct	tct	tcc	ttt	tat	ggc	aaa	gtg	gcc	5593
Asn	Val	His	Gly	Leu	Met	Ser	Ser	Ser	Phe	Tyr	Gly	Lys	Val	Ala	
			1845					1850					1855		
agt	ata	ctt	gat	cat	ttc	cac	ctg	tct	ccc	caa	ggt	gaa	gat	tca	5638
Ser	Ile	Leu	Asp	His	Phe	His	Leu	Ser	Pro	Gln	Gly	Glu	Asp	Ser	
			1860					1865					1870		
cca	tgt	tca	aat	gaa	agc	tcc	cga	atg	gaa	ata	act	agg	aaa	gtg	5683
Pro	Cys	Ser	Asn	Glu	Ser	Ser	Arg	Met	Glu	Ile	Thr	Arg	Lys	Val	
			1875					1880					1885		
gtc	tgc	ata	att	cat	gaa	tta	gtg	gac	tgg	aat	tct	att	ctt	ctg	5728
Val	Cys	Ile	Ile	His	Glu	Leu	Val	Asp	Trp	Asn	Ser	Ile	Leu	Leu	
			1890					1895					1900		
gag	ctc	tct	gaa	gtc	ttc	cat	gtt	aac	att	tct	ctt	gtg	aaa	act	5773
Glu	Leu	Ser	Glu	Val	Phe	His	Val	Asn	Ile	Ser	Leu	Val	Lys	Thr	
			1905					1910					1915		
gtg	cag	aaa	ttt	tgg	cat	aag	ata	tta	ccg	ttt	gtc	cca	cct	tca	5818
Val	Gln	Lys	Phe	Trp	His	Lys	Ile	Leu	Pro	Phe	Val	Pro	Pro	Ser	
			1920					1925					1930		
ata	aat	caa	act	agg	gat	agc	atc	tct	gaa	ctc	tgt	cct	agt	ggc	5863
Ile	Asn	Gln	Thr	Arg	Asp	Ser	Ile	Ser	Glu	Leu	Cys	Pro	Ser	Gly	
			1935					1940					1945		
tcc	ata	aag	caa	gtt	gct	ttg	caa	atc	ata	gaa	aaa	ctt	aaa	aat	5908
Ser	Ile	Lys	Gln	Val	Ala	Leu	Gln	Ile	Ile	Glu	Lys	Leu	Lys	Asn	
			1950					1955					1960		
gtc	aac	ttt	aca	aaa	gtt	aca	tca	ggc	gaa	aat	att	ctt	gac	aaa	5953
Val	Asn	Phe	Thr	Lys	Val	Thr	Ser	Gly	Glu	Asn	Ile	Leu	Asp	Lys	

1965				1970				1975							
cta	agt	agt	tta	aac	aag	atc	ctt	aac	att	aat	gaa	gac	aca	gag	5998
Leu	Ser	Ser	Leu	Asn	Lys	Ile	Leu	Asn	Ile	Asn	Glu	Asp	Thr	Glu	
			1980					1985					1990		
aca	tct	gtt	caa	aat	att	att	tcc	tca	aat	ttg	gaa	agg	aca	gta	6043
Thr	Ser	Val	Gln	Asn	Ile	Ile	Ser	Ser	Asn	Leu	Glu	Arg	Thr	Val	
			1995					2000					2005		
caa	ttg	att	tct	gaa	gac	tgg	agc	cta	gaa	aaa	agt	acg	cat	aat	6088
Gln	Leu	Ile	Ser	Glu	Asp	Trp	Ser	Leu	Glu	Lys	Ser	Thr	His	Asn	
			2010					2015					2020		
cta	ctc	tct	tta	ttc	atg	atg	ctc	cag	aat	gca	aat	gtc	aca	ggc	6133
Leu	Leu	Ser	Leu	Phe	Met	Met	Leu	Gln	Asn	Ala	Asn	Val	Thr	Gly	
			2025					2030					2035		
agc	agt	tta	gaa	gca	tta	tca	agt	ttt	att	gaa	aaa	agt	gaa	aca	6178
Ser	Ser	Leu	Glu	Ala	Leu	Ser	Ser	Phe	Ile	Glu	Lys	Ser	Glu	Thr	
			2040					2045					2050		
cct	tac	aac	ttt	gaa	gaa	cta	tgg	ccc	aag	ttt	caa	caa	atc	atg	6223
Pro	Tyr	Asn	Phe	Glu	Glu	Leu	Trp	Pro	Lys	Phe	Gln	Gln	Ile	Met	
			2055					2060					2065		
aaa	gac	cta	acc	caa	gat	ttt	aga	atc	aga	cac	ctg	ctt	tct	gaa	6268
Lys	Asp	Leu	Thr	Gln	Asp	Phe	Arg	Ile	Arg	His	Leu	Leu	Ser	Glu	
			2070					2075					2080		
atg	aac	aaa	gga	atc	aaa	agt	ata	aat	tca	atg	gct	ctt	caa	aag	6313
Met	Asn	Lys	Gly	Ile	Lys	Ser	Ile	Asn	Ser	Met	Ala	Leu	Gln	Lys	
			2085					2090					2095		
ata	act	ttg	cag	ttt	gcc	cat	ttc	ctg	gaa	atc	ctg	gat	tca	ccg	6358
Ile	Thr	Leu	Gln	Phe	Ala	His	Phe	Leu	Glu	Ile	Leu	Asp	Ser	Pro	
			2100					2105					2110		
tca	ttg	aag	aca	tta	gaa	att	att	gaa	gat	ttt	cta	ttg	gtc	aca	6403
Ser	Leu	Lys	Thr	Leu	Glu	Ile	Ile	Glu	Asp	Phe	Leu	Leu	Val	Thr	
			2115					2120					2125		
aaa	aac	tgg	ctt	cag	gaa	tat	gca	aat	gag	gat	tac	tcc	aga	atg	6448
Lys	Asn	Trp	Leu	Gln	Glu	Tyr	Ala	Asn	Glu	Asp	Tyr	Ser	Arg	Met	
			2130					2135					2140		
ata	gaa	aca	tta	ttc	att	cct	gtg	acc	aat	gag	agt	tca	act	gaa	6493
Ile	Glu	Thr	Leu	Phe	Ile	Pro	Val	Thr	Asn	Glu	Ser	Ser	Thr	Glu	
			2145					2150					2155		
gat	ata	gct	ttg	tta	gcc	aaa	gct	att	gct	act	ttt	tgg	ggc	tct	6538
Asp	Ile	Ala	Leu	Leu	Ala	Lys	Ala	Ile	Ala	Thr	Phe	Trp	Gly	Ser	
			2160					2165					2170		
tta	aaa	aat	ata	tct	aga	gca	ggc	aat	ttt	gat	gtt	gcc	ttt	ctt	6583
Leu	Lys	Asn	Ile	Ser	Arg	Ala	Gly	Asn	Phe	Asp	Val	Ala	Phe	Leu	
			2175					2180					2185		
acc	cat	ctg	cta	aat	caa	gaa	cag	ctg	act	aat	ttc	tca	gtt	gtt	6628
Thr	His	Leu	Leu	Asn	Gln	Glu	Gln	Leu	Thr	Asn	Phe	Ser	Val	Val	
			2190					2195					2200		

cag ctg ctt ttt	gaa aac atc cta att	aat ttg atc aat aac tta	6673
Gln Leu Leu Phe	Glu Asn Ile Leu Ile	Asn Leu Ile Asn Asn Leu	
2205	2210	2215	
gct ggg aat tct	cag gaa gca gct tgg	aac tta aat gat act gac	6718
Ala Gly Asn Ser	Gln Glu Ala Ala Trp	Asn Leu Asn Asp Thr Asp	
2220	2225	2230	
ctt caa ata atg	aat ttc att aac ctt	atc ttg aac cat atg cag	6763
Leu Gln Ile Met	Asn Phe Ile Asn Leu	Ile Leu Asn His Met Gln	
2235	2240	2245	
tca gaa act agt	agg aaa aca gtt ctc	tct ctg aga agc ata gta	6808
Ser Glu Thr Ser	Arg Lys Thr Val Leu	Ser Leu Arg Ser Ile Val	
2250	2255	2260	
gat ttc aca gaa	cag ttt ttg aaa aca	ttc ttc tcc ctt ttt cta	6853
Asp Phe Thr Glu	Gln Phe Leu Lys Thr	Phe Phe Ser Leu Phe Leu	
2265	2270	2275	
aag gaa gat tct	gag aac aaa ata tct	ctt ctg ctg aaa tat ttc	6898
Lys Glu Asp Ser	Glu Asn Lys Ile Ser	Leu Leu Leu Lys Tyr Phe	
2280	2285	2290	
cac aaa gat gtt	att gca gag atg agt	ttt gtc cca aaa gat aaa	6943
His Lys Asp Val	Ile Ala Glu Met Ser	Phe Val Pro Lys Asp Lys	
2295	2300	2305	
att cta gaa att	ctg aaa ctg gat caa	ttt ctt acc ctg atg ata	6988
Ile Leu Glu Ile	Leu Lys Leu Asp Gln	Phe Leu Thr Leu Met Ile	
2310	2315	2320	
caa gac aga ttg	atg aac att ttt tca	agt tta aag gag act ata	7033
Gln Asp Arg Leu	Met Asn Ile Phe Ser	Ser Leu Lys Glu Thr Ile	
2325	2330	2335	
tat cac cta atg	aaa agt tca ttt ata	tta gac aat gga gaa ttt	7078
Tyr His Leu Met	Lys Ser Ser Phe Ile	Leu Asp Asn Gly Glu Phe	
2340	2345	2350	
tat ttt gat act	cat caa gga ctg aag	ttc atg caa gat tta ttt	7123
Tyr Phe Asp Thr	His Gln Gly Leu Lys	Phe Met Gln Asp Leu Phe	
2355	2360	2365	
aat gcc ctt ctc	agg gaa act tca atg	aaa aat aag act gaa aat	7168
Asn Ala Leu Leu	Arg Glu Thr Ser Met	Lys Asn Lys Thr Glu Asn	
2370	2375	2380	
aat ata gac ttt	ttc aca gtg gtg agt	cag ttg ttt ttc cat gtg	7213
Asn Ile Asp Phe	Phe Thr Val Val Ser	Gln Leu Phe Phe His Val	
2385	2390	2395	
aat aag tct gag	gac ctc ttc aaa ctc	aat caa gat ctt ggg tca	7258
Asn Lys Ser Glu	Asp Leu Phe Lys Leu	Asn Gln Asp Leu Gly Ser	
2400	2405	2410	
gct ctt cac ctt	gta aga gaa tgt tca	aca gag atg gca aga ctt	7303
Ala Leu His Leu	Val Arg Glu Cys Ser	Thr Glu Met Ala Arg Leu	
2415	2420	2425	

ctg gat aca att tta cac tct cct aat aag gac ttc tat gct ttg	7348
Leu Asp Thr Ile Leu His Ser Pro Asn Lys Asp Phe Tyr Ala Leu	
2430 2435 2440	
tat cct acc ctg caa gaa gtt ata ctt gct aat cta acg gat ttg	7393
Tyr Pro Thr Leu Gln Glu Val Ile Leu Ala Asn Leu Thr Asp Leu	
2445 2450 2455	
ctt ttc ttt ata aat aat tca ttc cct cta aga aac aga gca aca	7438
Leu Phe Phe Ile Asn Asn Ser Phe Pro Leu Arg Asn Arg Ala Thr	
2460 2465 2470	
tta gaa att act aag aga tta gtt ggt gct att tca aga gca agt	7483
Leu Glu Ile Thr Lys Arg Leu Val Gly Ala Ile Ser Arg Ala Ser	
2475 2480 2485	
gaa gaa agt cac gtc ctg aaa ccc ctg tta gaa atg tct ggg act	7528
Glu Glu Ser His Val Leu Lys Pro Leu Leu Glu Met Ser Gly Thr	
2490 2495 2500	
ctg gtc atg ctg ttg aat gac agt gct gac ctg aga gat ctt gcc	7573
Leu Val Met Leu Leu Asn Asp Ser Ala Asp Leu Arg Asp Leu Ala	
2505 2510 2515	
aca tca atg gac tcc att gtg aaa ctt ctt aag ctg gtc aag aaa	7618
Thr Ser Met Asp Ser Ile Val Lys Leu Leu Lys Leu Val Lys Lys	
2520 2525 2530	
gtt tcg ggg aag atg tcc aca gtt ttt aaa act cat ttt atc tcc	7663
Val Ser Gly Lys Met Ser Thr Val Phe Lys Thr His Phe Ile Ser	
2535 2540 2545	
aat acc aag gac agt gtg aaa ttc ttt gac act ctg tat tcc atc	7708
Asn Thr Lys Asp Ser Val Lys Phe Phe Asp Thr Leu Tyr Ser Ile	
2550 2555 2560	
atg caa caa agt gtt caa aat ctt gtg aaa gaa ata gct act tta	7753
Met Gln Gln Ser Val Gln Asn Leu Val Lys Glu Ile Ala Thr Leu	
2565 2570 2575	
aaa aaa ata gat cat ttc aca ttt gaa aag ata aat gat ttg ttg	7798
Lys Lys Ile Asp His Phe Thr Phe Glu Lys Ile Asn Asp Leu Leu	
2580 2585 2590	
gtg cca ttt ctt gac ttg gcc ttt gaa atg att ggg gta gaa cct	7843
Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile Gly Val Glu Pro	
2595 2600 2605	
tat ata tca tca aac tct gat att ttc agt atg tca cct agc ata	7888
Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser Pro Ser Ile	
2610 2615 2620	
ctc tca tat atg aac caa tct aag gac ttt tct gat att ttg gaa	7933
Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile Leu Glu	
2625 2630 2635	
gaa att gct gaa ttt tta aca tct gtg aaa atg aac ttg gaa gat	7978
Glu Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu Asp	
2640 2645 2650	
atg agg agt ctt gcg gta gca ttt aac aat gag act caa aca ttt	8023

Met	Arg	Ser	Leu	Ala	Val	Ala	Phe	Asn	Asn	Glu	Thr	Gln	Thr	Phe	
			2655					2660					2665		
tct	atg	gat	tct	gtc	aac	tta	ygg	gaa	gaa	att	ctg	ggt	tgc	tta	8068
Ser	Met	Asp	Ser	Val	Asn	Leu	Xaa	Glu	Glu	Ile	Leu	Gly	Cys	Leu	
			2670					2675					2680		
gtt	cct	ata	aat	aac	atc	acc	aac	caa	atg	gac	ttc	tta	tac	cct	8113
Val	Pro	Ile	Asn	Asn	Ile	Thr	Asn	Gln	Met	Asp	Phe	Leu	Tyr	Pro	
			2685					2690					2695		
aat	cca	att	tcc	act	cat	agt	ggc	cct	caa	gat	ata	aaa	tgg	gaa	8158
Asn	Pro	Ile	Ser	Thr	His	Ser	Gly	Pro	Gln	Asp	Ile	Lys	Trp	Glu	
			2700					2705					2710		
ata	att	cat	gaa	gtg	atc	ctt	ttt	ttg	gat	aaa	ata	tta	tca	caa	8203
Ile	Ile	His	Glu	Val	Ile	Leu	Phe	Leu	Asp	Lys	Ile	Leu	Ser	Gln	
			2715					2720					2725		
aac	agc	aca	gaa	ata	gga	tct	ttc	ttg	aaa	atg	gtg	atc	tgt	ctc	8248
Asn	Ser	Thr	Glu	Ile	Gly	Ser	Phe	Leu	Lys	Met	Val	Ile	Cys	Leu	
			2730					2735					2740		
acc	tta	gaa	gct	ctt	tgg	aaa	aac	tta	aag	aaa	gat	aat	tgg	aat	8293
Thr	Leu	Glu	Ala	Leu	Trp	Lys	Asn	Leu	Lys	Lys	Asp	Asn	Trp	Asn	
			2745					2750					2755		
gtt	tct	aat	gtg	ttg	atg	acg	ttt	act	cag	cat	cca	aat	aac	ctt	8338
Val	Ser	Asn	Val	Leu	Met	Thr	Phe	Thr	Gln	His	Pro	Asn	Asn	Leu	
			2760					2765					2770		
ttg	aaa	acc	ata	gaa	aca	gtt	tta	gag	gcc	tcc	agt	gga	att	aaa	8383
Leu	Lys	Thr	Ile	Glu	Thr	Val	Leu	Glu	Ala	Ser	Ser	Gly	Ile	Lys	
			2775					2780					2785		
agt	gac	tat	gaa	ggc	gat	ttg	aat	aaa	agt	tta	tat	ttt	gac	aca	8428
Ser	Asp	Tyr	Glu	Gly	Asp	Leu	Asn	Lys	Ser	Leu	Tyr	Phe	Asp	Thr	
			2790					2795					2800		
cct	ttg	agt	cag	aat	ata	act	cat	cat	caa	ctt	gaa	aaa	gca	atc	8473
Pro	Leu	Ser	Gln	Asn	Ile	Thr	His	His	Gln	Leu	Glu	Lys	Ala	Ile	
			2805					2810					2815		
cat	aat	gtt	tta	agt	aga	ata	gct	ctc	tgg	agg	aaa	gga	ctt	cgt	8518
His	Asn	Val	Leu	Ser	Arg	Ile	Ala	Leu	Trp	Arg	Lys	Gly	Leu	Arg	
			2820					2825					2830		
ttt	aac	aac	tct	gaa	tgg	ata	act	tcc	aca	aga	act	ttg	ttt	cag	8563
Phe	Asn	Asn	Ser	Glu	Trp	Ile	Thr	Ser	Thr	Arg	Thr	Leu	Phe	Gln	
			2835					2840					2845		
cca	ctt	ttt	gag	att	ttc	att	aaa	gca	acc	acc	gga	aag	aat	gtc	8608
Pro	Leu	Phe	Glu	Ile	Phe	Ile	Lys	Ala	Thr	Thr	Gly	Lys	Asn	Val	
			2850					2855					2860		
aca	tca	gaa	aaa	gaa	gag	aga	acc	gag	aaa	gag	atg	att	gac	ttt	8653
Thr	Ser	Glu	Lys	Glu	Glu	Arg	Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	
			2865					2870					2875		
cct	tat	agt	ttc	aaa	cca	ttt	ttc	tgt	ttg	gag	aaa	tac	ctg	gga	8698
Pro	Tyr	Ser	Phe	Lys	Pro	Phe	Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	

2880								2885				2890				
gga tta ttt gta ttg act aaa tac tgg caa caa atc cca cta aca								8743								
Gly Leu Phe Val Leu Thr Lys Tyr Trp Gln Gln Ile Pro Leu Thr																
2895								2900				2905				
gat caa agt gtt gtt gag att tgt gaa gtt ttc cag cag act gtg								8788								
Asp Gln Ser Val Val Glu Ile Cys Glu Val Phe Gln Gln Thr Val																
2910								2915				2920				
aag ccc tca gaa gcc atg gag atg ctg cag aaa gtg aag atg atg								8833								
Lys Pro Ser Glu Ala Met Glu Met Leu Gln Lys Val Lys Met Met																
2925								2930				2935				
gtc gta cgt gtg ctc acc atc gtt gca gaa aac cct tcc tgg acc								8878								
Val Val Arg Val Leu Thr Ile Val Ala Glu Asn Pro Ser Trp Thr																
2940								2945				2950				
aag gac att ttg tgt gct act ctg agt tgc aag caa aat ggg ata								8923								
Lys Asp Ile Leu Cys Ala Thr Leu Ser Cys Lys Gln Asn Gly Ile																
2955								2960				2965				
agg cat ctc att tta tct gct ata caa ggg gtc act ttg gcg cag								8968								
Arg His Leu Ile Leu Ser Ala Ile Gln Gly Val Thr Leu Ala Gln																
2970								2975				2980				
gac cac ttc cag gaa att gaa aag ata tgg tcc tcg ccg aat cag								9013								
Asp His Phe Gln Glu Ile Glu Lys Ile Trp Ser Ser Pro Asn Gln																
2985								2990				2995				
cta aat tgt gaa agt ctt agc aag aat ctt tct agc acc ttg gag								9058								
Leu Asn Cys Glu Ser Leu Ser Lys Asn Leu Ser Ser Thr Leu Glu																
3000								3005				3010				
agc ttc aag agc agc ttg gaa aat gcc act ggc cag gac tgc aca								9103								
Ser Phe Lys Ser Ser Leu Glu Asn Ala Thr Gly Gln Asp Cys Thr																
3015								3020				3025				
agc cag ccg agg ctg gag acg gtg cag cag cac ttg tac atg ttg								9148								
Ser Gln Pro Arg Leu Glu Thr Val Gln Gln His Leu Tyr Met Leu																
3030								3035				3040				
gcc aaa agc ctc gag gaa act tgg tca tca ggg aat ccc atc atg								9193								
Ala Lys Ser Leu Glu Glu Thr Trp Ser Ser Gly Asn Pro Ile Met																
3045								3050				3055				
act ttt ctc agc aat ttc aca gta act gag gat gta aaa ata aaa								9238								
Thr Phe Leu Ser Asn Phe Thr Val Thr Glu Asp Val Lys Ile Lys																
3060								3065				3070				
gat ttg atg aag aat atc acc aag ttg act gag gag ctt cgc tct								9283								
Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu Glu Leu Arg Ser																
3075								3080				3085				
tcc atc caa atc tcg aat gag act atc cat agc att cta gaa gca								9328								
Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile Leu Glu Ala																
3090								3095				3100				
aat att tcc cac tcc aag gtt ctc ttc agt gcc ctc acc gta gct								9373								
Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr Val Ala																
3105								3110				3115				

ctg tct gga aag	tgt gat cag gaa atc	ott cat ctc ctg ctg	aca	9418
Leu Ser Gly Lys	Cys Asp Gln Glu Ile	Leu His Leu Leu Leu	Thr	
3120	3125	3130		
ttt ccc aaa ggg	gaa aaa tct tgg atc	gca gcg gag gaa ctc	tgt	9463
Phe Pro Lys Gly	Glu Lys Ser Trp Ile	Ala Ala Glu Glu Leu	Cys	
3135	3140	3145		
agc ctg cca ggg	tca aaa gtg tat tct	ctg att gtg ttg ctg	agt	9508
Ser Leu Pro Gly	Ser Lys Val Tyr Ser	Leu Ile Val Leu Leu	Ser	
3150	3155	3160		
cga aac ttg gat	gtg cga gct ttc att	tac aag act ctg atg	cct	9553
Arg Asn Leu Asp	Val Arg Ala Phe Ile	Tyr Lys Thr Leu Met	Pro	
3165	3170	3175		
tct gaa gca aat	ggc ttg ctc aac tcc	ttg ctg gat ata gtt	tcc	9598
Ser Glu Ala Asn	Gly Leu Leu Asn Ser	Leu Leu Asp Ile Val	Ser	
3180	3185	3190		
agc ctc agc gcc	ttg ctt gcc aaa gcc	cag cac gtc ttt gag	tat	9643
Ser Leu Ser Ala	Leu Leu Ala Lys Ala	Gln His Val Phe Glu	Tyr	
3195	3200	3205		
ctt cct gag ttt	ctt cac aca ttt aaa	atc act gcc ttg cta	gaa	9688
Leu Pro Glu Phe	Leu His Thr Phe Lys	Ile Thr Ala Leu Leu	Glu	
3210	3215	3220		
acc ctg gac ttt	caa cag gtt tca caa	aat gtc cag gcc aga	agt	9733
Thr Leu Asp Phe	Gln Gln Val Ser Gln	Asn Val Gln Ala Arg	Ser	
3225	3230	3235		
tca gct ttt ggt	tct ttc cag ttt gtg	atg aag atg gtt tgc	aag	9778
Ser Ala Phe Gly	Ser Phe Gln Phe Val	Met Lys Met Val Cys	Lys	
3240	3245	3250		
gac caa gca tca	ttc ctt agc gat tct	aat atg ttt att aat	ttg	9823
Asp Gln Ala Ser	Phe Leu Ser Asp Ser	Asn Met Phe Ile Asn	Leu	
3255	3260	3265		
ccc aga gtt aag	gaa ctc ttg gaa gat	gac aaa gaa aaa ttc	aac	9868
Pro Arg Val Lys	Glu Leu Leu Glu Asp	Asp Lys Glu Lys Phe	Asn	
3270	3275	3280		
att cct gaa gat	tca aca ccg ttt tgc	ttg aag ctt tat cag	gaa	9913
Ile Pro Glu Asp	Ser Thr Pro Phe Cys	Leu Lys Leu Tyr Gln	Glu	
3285	3290	3295		
att cta caa ttg	cca aat ggt gct ttg	gtg tgg acc ttc cta	aaa	9958
Ile Leu Gln Leu	Pro Asn Gly Ala Leu	Val Trp Thr Phe Leu	Lys	
3300	3305	3310		
ccc ata ttg cat	gga aaa ata cta tac	aca cca aac act cca	gaa	10003
Pro Ile Leu His	Gly Lys Ile Leu Tyr	Thr Pro Asn Thr Pro	Glu	
3315	3320	3325		
att aac aag gtc	att caa aag gct aat	tac acc ttt tat att	gtg	10048
Ile Asn Lys Val	Ile Gln Lys Ala Asn	Tyr Thr Phe Tyr Ile	Val	
3330	3335	3340		

gac aaa cta aaa act tta tca gaa aca ctg ctg gaa atg tcc agc	10093
Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu Leu Glu Met Ser Ser	
3345 3350 3355	
ctt ttc cag aga agt gga agt ggc cag atg ttc aac cag ctg cag	10138
Leu Phe Gln Arg Ser Gly Ser Gly Gln Met Phe Asn Gln Leu Gln	
3360 3365 3370	
gag gcc ctg aga aac aaa ttt gta aga aac ttt gta gaa aac cag	10183
Glu Ala Leu Arg Asn Lys Phe Val Arg Asn Phe Val Glu Asn Gln	
3375 3380 3385	
ttg cac att gat gta gac aaa ctt act gaa aaa ctc cag aca tac	10228
Leu His Ile Asp Val Asp Lys Leu Thr Glu Lys Leu Gln Thr Tyr	
3390 3395 3400	
gga ggg ctg ctg gat gag atg ttt aac cat gca ggc gct gga cgc	10273
Gly Gly Leu Leu Asp Glu Met Phe Asn His Ala Gly Ala Gly Arg	
3405 3410 3415	
ttc cgt ttc ttg ggc agc atc ttg gtc aat ctc tct tcc tgc gtg	10318
Phe Arg Phe Leu Gly Ser Ile Leu Val Asn Leu Ser Ser Cys Val	
3420 3425 3430	
gca ctg aac cgt ttc cag gct ctg cag tct gtc gac atc ctg gag	10363
Ala Leu Asn Arg Phe Gln Ala Leu Gln Ser Val Asp Ile Leu Glu	
3435 3440 3445	
act aaa gca cat gaa ctc ttg cag cag aac agc ttc ttg gcc agt	10408
Thr Lys Ala His Glu Leu Leu Gln Gln Asn Ser Phe Leu Ala Ser	
3450 3455 3460	
atc att ttc agc aat tcc tta ttc gac aag aac ttc aga tca gag	10453
Ile Ile Phe Ser Asn Ser Leu Phe Asp Lys Asn Phe Arg Ser Glu	
3465 3470 3475	
tct gtc aaa ctg cca ccc cat gtc tca tac aca atc cgg acc aat	10498
Ser Val Lys Leu Pro Pro His Val Ser Tyr Thr Ile Arg Thr Asn	
3480 3485 3490	
gtg tta tac agc gtg cga aca gat gtg gta aaa aac cct tct tgg	10543
Val Leu Tyr Ser Val Arg Thr Asp Val Val Lys Asn Pro Ser Trp	
3495 3500 3505	
aag ttc cac cct cag aat cta cca gct gat ggg ttc aaa tat aac	10588
Lys Phe His Pro Gln Asn Leu Pro Ala Asp Gly Phe Lys Tyr Asn	
3510 3515 3520	
tac gtc ttt gcc cca ctg caa gac atg atc gaa aga gcc atc att	10633
Tyr Val Phe Ala Pro Leu Gln Asp Met Ile Glu Arg Ala Ile Ile	
3525 3530 3535	
ttg gtg cag act ggg cag gaa gcc ctg gaa cca gca gca cag act	10678
Leu Val Gln Thr Gly Gln Glu Ala Leu Glu Pro Ala Ala Gln Thr	
3540 3545 3550	
cag gcg gcc cct tac ccc tgc cat acc agc gac cta ttc ctg aac	10723
Gln Ala Ala Pro Tyr Pro Cys His Thr Ser Asp Leu Phe Leu Asn	
3555 3560 3565	
aac gtt ggt ttc ttt ttt cca ctg ata atg atg ctg acg tgg atg	10768

Asn	Val	Gly	Phe	Phe	Phe	Pro	Leu	Ile	Met	Met	Leu	Thr	Trp	Met	
			3570					3575					3580		
gtg	tct	gtg	gcc	agc	atg	gtc	aga	aag	ttg	gtg	tat	gag	cag	gag	10813
Val	Ser	Val	Ala	Ser	Met	Val	Arg	Lys	Leu	Val	Tyr	Glu	Gln	Glu	
			3585					3590					3595		
ata	cag	ata	gaa	gag	tat	atg	cgg	atg	atg	gga	gtg	cat	cca	gtg	10858
Ile	Gln	Ile	Glu	Glu	Tyr	Met	Arg	Met	Met	Gly	Val	His	Pro	Val	
			3600					3605					3610		
atc	cat	ttc	ctg	gcc	tgg	ttc	ctg	gag	aac	atg	gct	gtg	ttg	acc	10903
Ile	His	Phe	Leu	Ala	Trp	Phe	Leu	Glu	Asn	Met	Ala	Val	Leu	Thr	
			3615					3620					3625		
ata	agc	agt	gct	act	ctg	gcc	atc	gtt	ctg	aaa	aca	agt	ggc	atc	10948
Ile	Ser	Ser	Ala	Thr	Leu	Ala	Ile	Val	Leu	Lys	Thr	Ser	Gly	Ile	
			3630					3635					3640		
ttt	gca	cac	agc	aat	acc	ttt	att	gtt	ttc	ctc	ttt	ctc	ttg	gat	10993
Phe	Ala	His	Ser	Asn	Thr	Phe	Ile	Val	Phe	Leu	Phe	Leu	Leu	Asp	
			3645					3650					3655		
ttt	ggg	atg	tca	gtc	gtc	atg	ctg	agc	tac	ctc	ttg	agt	gca	ttt	11038
Phe	Gly	Met	Ser	Val	Val	Met	Leu	Ser	Tyr	Leu	Leu	Ser	Ala	Phe	
			3660					3665					3670		
ttc	agc	caa	gct	aat	aca	gcg	gcc	ctt	tgt	acc	agc	ctg	gtg	tac	11083
Phe	Ser	Gln	Ala	Asn	Thr	Ala	Ala	Leu	Cys	Thr	Ser	Leu	Val	Tyr	
			3675					3680					3685		
atg	atc	agc	ttt	ctg	ccc	tac	ata	gtt	cta	ttg	gtt	cta	cat	aac	11128
Met	Ile	Ser	Phe	Leu	Pro	Tyr	Ile	Val	Leu	Leu	Val	Leu	His	Asn	
			3690					3695					3700		
caa	tta	agt	ttt	gtt	aat	cag	aca	ttt	ctg	tgc	ctt	ctt	tcg	aca	11173
Gln	Leu	Ser	Phe	Val	Asn	Gln	Thr	Phe	Leu	Cys	Leu	Leu	Ser	Thr	
			3705					3710					3715		
acc	gcc	ttt	gga	caa	ggg	gta	ttt	ttt	att	aca	ttc	ctg	gaa	gga	11218
Thr	Ala	Phe	Gly	Gln	Gly	Val	Phe	Phe	Ile	Thr	Phe	Leu	Glu	Gly	
			3720					3725					3730		
caa	gag	aca	ggg	att	caa	tgg	aat	aat	atg	tac	cag	gct	ctg	gaa	11263
Gln	Glu	Thr	Gly	Ile	Gln	Trp	Asn	Asn	Met	Tyr	Gln	Ala	Leu	Glu	
			3735					3740					3745		
caa	ggg	ggc	atg	aca	ttt	ggc	tgg	gtt	tgc	tgg	atg	att	ctt	ttt	11308
Gln	Gly	Gly	Met	Thr	Phe	Gly	Trp	Val	Cys	Trp	Met	Ile	Leu	Phe	
			3750					3755					3760		
gat	tca	agc	ctt	tat	ttt	ttg	tgt	gga	tgg	tac	ttg	agc	aac	ttg	11353
Asp	Ser	Ser	Leu	Tyr	Phe	Leu	Cys	Gly	Trp	Tyr	Leu	Ser	Asn	Leu	
			3765					3770					3775		
att	cct	gga	aca	ttt	ggg	tta	cgg	aaa	cca	tgg	tat	ttc	ccc	ttt	11398
Ile	Pro	Gly	Thr	Phe	Gly	Leu	Arg	Lys	Pro	Trp	Tyr	Phe	Pro	Phe	
			3780					3785					3790		
act	gcc	tca	tat	tgg	aag	agt	gtg	ggg	ttc	ttg	gtg	gag	aaa	agg	11443
Thr	Ala	Ser	Tyr	Trp	Lys	Ser	Val	Gly	Phe	Leu	Val	Glu	Lys	Arg	

3795	3800	3805	
caa tac ttt cta agt tct agt ctg ttc ttc ttc aat gag aac ttt Gln Tyr Phe Leu Ser Ser Ser Leu Phe Phe Phe Asn Glu Asn Phe 3810 3815 3820	11488		
gac aat aaa ggg tca tca ctg caa aac agg gaa gga gag ctt gaa Asp Asn Lys Gly Ser Ser Leu Gln Asn Arg Glu Gly Glu Leu Glu 3825 3830 3835	11533		
gga agt gcc ccg gga gtc acc ctg gtg tct gtg acc aag gaa tat Gly Ser Ala Pro Gly Val Thr Leu Val Ser Val Thr Lys Glu Tyr 3840 3845 3850	11578		
gag ggc cac aag gct gtg gtc caa gac ctc agc ctg acc ttc tac Glu Gly His Lys Ala Val Val Gln Asp Leu Ser Leu Thr Phe Tyr 3855 3860 3865	11623		
aga gac caa atc acc gcc ctg ctg ggg aca aac ggt gcc ggg aaa Arg Asp Gln Ile Thr Ala Leu Leu Gly Thr Asn Gly Ala Gly Lys 3870 3875 3880	11668		
acc act atc ata tcc atg ttg acg ggg ctc cac cct ccc act tct Thr Thr Ile Ile Ser Met Leu Thr Gly Leu His Pro Pro Thr Ser 3885 3890 3895	11713		
gga acc atc atc atc aat ggc aag aac cta cag aca gac ctg tcg Gly Thr Ile Ile Ile Asn Gly Lys Asn Leu Gln Thr Asp Leu Ser 3900 3905 3910	11758		
agg gtc aga atg gag ctt ggt gtg tgt ccg cag cag gac atc ctg Arg Val Arg Met Glu Leu Gly Val Cys Pro Gln Gln Asp Ile Leu 3915 3920 3925	11803		
ttg gac aac ctc acc gtc cgg gaa cat ttg ctg ctc ttt gct tcc Leu Asp Asn Leu Thr Val Arg Glu His Leu Leu Leu Phe Ala Ser 3930 3935 3940	11848		
ata aag gcg cct cag tgg acc aag aag gag ctg cat cag caa gtc Ile Lys Ala Pro Gln Trp Thr Lys Lys Glu Leu His Gln Gln Val 3945 3950 3955	11893		
aat caa act ctt cag gat gtg gac tta act cag cat cag cac aaa Asn Gln Thr Leu Gln Asp Val Asp Leu Thr Gln His Gln His Lys 3960 3965 3970	11938		
cag acc cga gct ctg tct gga ggc ctg aag agg aag ctc tcc ctt Gln Thr Arg Ala Leu Ser Gly Gly Leu Lys Arg Lys Leu Ser Leu 3975 3980 3985	11983		
ggc att gct ttc atg ggc atg tcg agg acc gtg gtt ctg gat gag Gly Ile Ala Phe Met Gly Met Ser Arg Thr Val Val Leu Asp Glu 3990 3995 4000	12028		
ccc acc agt ggg gtg gac cct tgc tcc cgg cat agc ctg tgg gac Pro Thr Ser Gly Val Asp Pro Cys Ser Arg His Ser Leu Trp Asp 4005 4010 4015	12073		
att ctg ctc aag tac cga gaa ggt cgt acg atc atc ttc aca acc Ile Leu Leu Lys Tyr Arg Glu Gly Arg Thr Ile Ile Phe Thr Thr 4020 4025 4030	12118		

cac	cac	ctg	gat	gaa	gcc	gaa	gcg	ctg	agt	gac	cgc	gtg	gcc	gtc	12163
His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser	Asp	Arg	Val	Ala	Val	
			4035					4040					4045		
ctc	cag	cat	ggg	agg	ctc	agg	tgc	tgc	ggg	cct	ccc	ttc	tgc	ctg	12208
Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly	Pro	Pro	Phe	Cys	Leu	
			4050					4055					4060		
aag	gag	gca	tat	ggc	cag	ggg	ctc	cgc	ctg	aca	ctc	acg	agg	cag	12253
Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu	Thr	Leu	Thr	Arg	Gln	
			4065					4070					4075		
cct	tct	gtt	ctg	gag	gcc	cat	gat	ctg	aaa	gac	atg	gct	tgt	gtt	12298
Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys	Asp	Met	Ala	Cys	Val	
			4080					4085					4090		
aca	tcc	ctg	ata	aag	atc	tat	att	cca	caa	gca	ttt	ctc	aaa	gac	12343
Thr	Ser	Leu	Ile	Lys	Ile	Tyr	Ile	Pro	Gln	Ala	Phe	Leu	Lys	Asp	
			4095					4100					4105		
agc	agt	gga	agt	gag	ctg	acc	tac	acc	att	cca	aag	gac	aca	gac	12388
Ser	Ser	Gly	Ser	Glu	Leu	Thr	Tyr	Thr	Ile	Pro	Lys	Asp	Thr	Asp	
			4110					4115					4120		
aag	gcc	tgc	ttg	aaa	ggg	ctc	ttc	cag	gcc	ctg	gat	gag	aac	ctg	12433
Lys	Ala	Cys	Leu	Lys	Gly	Leu	Phe	Gln	Ala	Leu	Asp	Glu	Asn	Leu	
			4125					4130					4135		
cat	cag	ctg	cac	ctg	acg	ggc	tat	ggg	atc	tca	gac	acc	acc	tta	12478
His	Gln	Leu	His	Leu	Thr	Gly	Tyr	Gly	Ile	Ser	Asp	Thr	Thr	Leu	
			4140					4145					4150		
gaa	gag	gtg	ttt	ttg	atg	ctt	ttg	caa	gat	tcc	aac	aag	aaa	tct	12523
Glu	Glu	Val	Phe	Leu	Met	Leu	Leu	Gln	Asp	Ser	Asn	Lys	Lys	Ser	
			4155					4160					4165		
cac	att	gcc	ctg	ggg	act	gag	tca	gag	ctg	cag	aac	cac	agg	cct	12568
His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu	Gln	Asn	His	Arg	Pro	
			4170					4175					4180		
aca	gga	cat	ctg	tct	ggc	tac	tgt	ggc	tcc	cta	gca	cgg	ccc	gca	12613
Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser	Leu	Ala	Arg	Pro	Ala	
			4185					4190					4195		
act	gtg	cag	ggc	gtc	cag	ctg	ctc	cgc	gca	caa	gtg	gcc	gcg	atc	12658
Thr	Val	Gln	Gly	Val	Gln	Leu	Leu	Arg	Ala	Gln	Val	Ala	Ala	Ile	
			4200					4205					4210		
ctg	gcc	cgg	agg	ctc	cgc	cgc	acg	ctg	cgc	gcc	ggg	aag	agc	acc	12703
Leu	Ala	Arg	Arg	Leu	Arg	Arg	Thr	Leu	Arg	Ala	Gly	Lys	Ser	Thr	
			4215					4220					4225		
ctc	gcc	gac	ctg	ctg	ctg	cca	gtc	ctc	ttc	gtg	gcc	ttg	gcc	atg	12748
Leu	Ala	Asp	Leu	Leu	Leu	Pro	Val	Leu	Phe	Val	Ala	Leu	Ala	Met	
			4230					4235					4240		
ggc	ttg	ttc	atg	gtg	aga	ccc	ctg	gcc	acc	gag	tac	cct	ccc	ctc	12793
Gly	Leu	Phe	Met	Val	Arg	Pro	Leu	Ala	Thr	Glu	Tyr	Pro	Pro	Leu	
			4245					4250					4255		

aga ctc aca cct	gga cat tac cag cgg	gcc gag acc tac ttt ttc	12838
Arg Leu Thr Pro	Gly His Tyr Gln Arg	Ala Glu Thr Tyr Phe Phe	
4260	4265	4270	
agc agt ggg ggc	gac aac ttg gac ctc	acc cgt gtg ctt ctg cgg	12883
Ser Ser Gly Gly	Asp Asn Leu Asp Leu	Thr Arg Val Leu Leu Arg	
4275	4280	4285	
aag ttt aga gat	caa gat ttg ccc tgt	gca gat tta aac cca cgc	12928
Lys Phe Arg Asp	Gln Asp Leu Pro Cys	Ala Asp Leu Asn Pro Arg	
4290	4295	4300	
cag aag aat tct	tca tgc tgg cgc aca	gat ccc ttt tct cac cca	12973
Gln Lys Asn Ser	Ser Cys Trp Arg Thr	Asp Pro Phe Ser His Pro	
4305	4310	4315	
gaa ttc cag gat	tca tgt ggc tgc ctg	aag tgt cca aat aga agt	13018
Glu Phe Gln Asp	Ser Cys Gly Cys Leu	Lys Cys Pro Asn Arg Ser	
4320	4325	4330	
gct agt gct ccc	tac ctg acc aac cac	ctg ggc cac aca ctg ttg	13063
Ala Ser Ala Pro	Tyr Leu Thr Asn His	Leu Gly His Thr Leu Leu	
4335	4340	4345	
aat ctc tca ggc	ttc aat atg gag gag	tac ttg ctg gca cca tct	13108
Asn Leu Ser Gly	Phe Asn Met Glu Glu	Tyr Leu Leu Ala Pro Ser	
4350	4355	4360	
gaa aaa cca agg	ctt gga ggt tgg tct	ttt gga tta aaa atc ccc	13153
Glu Lys Pro Arg	Leu Gly Gly Trp Ser	Phe Gly Leu Lys Ile Pro	
4365	4370	4375	
agt gaa gct gga	ggt gca aat gga aac	ata tca aaa ccc cca act	13198
Ser Glu Ala Gly	Gly Ala Asn Gly Asn	Ile Ser Lys Pro Pro Thr	
4380	4385	4390	
ctg gca aag gtg	tgg tat aat cag aag	ggt ttt cat tcc cta cct	13243
Leu Ala Lys Val	Trp Tyr Asn Gln Lys	Gly Phe His Ser Leu Pro	
4395	4400	4405	
tcc tac tta aat	cat cta aac aac ctt	att ttg tgg cag cac cta	13288
Ser Tyr Leu Asn	His Leu Asn Asn Leu	Ile Leu Trp Gln His Leu	
4410	4415	4420	
ccc cct act gtg	gac tgg aga caa tac	gga ata aca ctc tac agc	13333
Pro Pro Thr Val	Asp Trp Arg Gln Tyr	Gly Ile Thr Leu Tyr Ser	
4425	4430	4435	
cac cca tat gga	ggg gcc ttg ctg aac	gag gac aag atc ctg gag	13378
His Pro Tyr Gly	Gly Ala Leu Leu Asn	Glu Asp Lys Ile Leu Glu	
4440	4445	4450	
agc atc cgt cag	tgt gga gtg gcc ctc	tgc atc gtg ctg gga ttc	13423
Ser Ile Arg Gln	Cys Gly Val Ala Leu	Cys Ile Val Leu Gly Phe	
4455	4460	4465	
tcc atc ctg tct	gca tcc atc ggc agc	tct gtg gtg agg gac agg	13468
Ser Ile Leu Ser	Ala Ser Ile Gly Ser	Ser Val Val Arg Asp Arg	
4470	4475	4480	
gtg att gga gcc	aaa agg ttg cag cac	ata agt ggc ctt ggc tac	13513

Val Ile Gly Ala	Lys Arg Leu Gln His	Ile Ser Gly Leu Gly Tyr	
4485	4490	4495	
agg atg tac tgg	ttc aca aac ttc cta	tat gac atg ctc ttt tac	13558
Arg Met Tyr Trp	Phe Thr Asn Phe Leu	Tyr Asp Met Leu Phe Tyr	
4500	4505	4510	
ttg gtt tcc gtc	tgc ctg tgt gtt gcc	gtt att gtc gcc ttc cag	13603
Leu Val Ser Val	Cys Leu Cys Val Ala	Val Ile Val Ala Phe Gln	
4515	4520	4525	
tta aca gct ttt	act ttc cgc aag aac	ttg gca gcc acg gcc ctc	13648
Leu Thr Ala Phe	Thr Phe Arg Lys Asn	Leu Ala Ala Thr Ala Leu	
4530	4535	4540	
ctg ctg tca ctt	ttc gga tat gca act	ctt cca tgg atg tac ctg	13693
Leu Leu Ser Leu	Phe Gly Tyr Ala Thr	Leu Pro Trp Met Tyr Leu	
4545	4550	4555	
atg tcc aga atc	ttt tcc agt tcg gac	gtg gct ttc att tcc tat	13738
Met Ser Arg Ile	Phe Ser Ser Ser Asp	Val Ala Phe Ile Ser Tyr	
4560	4565	4570	
gtc tca cta aac	ttc atc ttt ggc ctt	tgt acc atg ccc ata acc	13783
Val Ser Leu Asn	Phe Ile Phe Gly Leu	Cys Thr Met Pro Ile Thr	
4575	4580	4585	
att atg ccc cgg	ttg cta gcc atc atc	tcc aaa gct aag aat tta	13828
Ile Met Pro Arg	Leu Leu Ala Ile Ile	Ser Lys Ala Lys Asn Leu	
4590	4595	4600	
cag aat atc tat	gat gtc ctc aag tgg	gtc ttt act att ttt cct	13873
Gln Asn Ile Tyr	Asp Val Leu Lys Trp	Val Phe Thr Ile Phe Pro	
4605	4610	4615	
caa ttc tgt ctt	ggc caa gga ctg gta	gaa ctc tgc tat aat cag	13918
Gln Phe Cys Leu	Gly Gln Gly Leu Val	Glu Leu Cys Tyr Asn Gln	
4620	4625	4630	
atc aaa tat gac	ctg acc cac aac ttc	ggc att gat tcc tat gtg	13963
Ile Lys Tyr Asp	Leu Thr His Asn Phe	Gly Ile Asp Ser Tyr Val	
4635	4640	4645	
agt ccc ttt gag	atg aac ttt ctg ggc	tgg atc ttc gtg caa ctg	14008
Ser Pro Phe Glu	Met Asn Phe Leu Gly	Trp Ile Phe Val Gln Leu	
4650	4655	4660	
gcc tcg cag ggc	aca gta ctt ctc ctc	ttg agg gtt ctg cta cac	14053
Ala Ser Gln Gly	Thr Val Leu Leu Leu	Leu Arg Val Leu Leu His	
4665	4670	4675	
tgg gac ctt ctg	cga tgg cca agg ggt	cat tct act ctc caa ggc	14098
Trp Asp Leu Leu	Arg Trp Pro Arg Gly	His Ser Thr Leu Gln Gly	
4680	4685	4690	
aca gtc aaa tct	tct aag gat aca gat	gtt gaa aaa gag gaa aag	14143
Thr Val Lys Ser	Ser Lys Asp Thr Asp	Val Glu Lys Glu Glu Lys	
4695	4700	4705	
aga gtg ttt gaa	gga agg acc aat gga	gac att ctt gtg tta tac	14188
Arg Val Phe Glu	Gly Arg Thr Asn Gly	Asp Ile Leu Val Leu Tyr	

4710	4715	4720	
aac ctt agt aaa	cat tat cga cgc ttt	ttc cag aat att att gct	14233
Asn Leu Ser Lys	His Tyr Arg Arg Phe	Phe Gln Asn Ile Ile Ala	
4725	4730	4735	
gtg caa gat att	agt ttg ggc ata cca	aaa gga gag tgc ttt gga	14278
Val Gln Asp Ile	Ser Leu Gly Ile Pro	Lys Gly Glu Cys Phe Gly	
4740	4745	4750	
ctt cta ggg gtg	aat gga gct ggg aag	agc acg act ttc aaa atg	14323
Leu Leu Gly Val	Asn Gly Ala Gly Lys	Ser Thr Thr Phe Lys Met	
4755	4760	4765	
ctg aat ggt gaa	gtt tct cta act tca	gga cat gct atc atc agg	14368
Leu Asn Gly Glu	Val Ser Leu Thr Ser	Gly His Ala Ile Ile Arg	
4770	4775	4780	
act ccc atg gga	gac gcc gtg gac ctg	tct tct gct ggc acg gca	14413
Thr Pro Met Gly	Asp Ala Val Asp Leu	Ser Ser Ala Gly Thr Ala	
4785	4790	4795	
ggc gtg ctc att	ggc tac tgt ccc cag	cag gat gcc ctg gac gag	14458
Gly Val Leu Ile	Gly Tyr Cys Pro Gln	Gln Asp Ala Leu Asp Glu	
4800	4805	4810	
ctt ctg act ggt	tgg gaa cat ctc tat	tat tac tgt agc tta cgc	14503
Leu Leu Thr Gly	Trp Glu His Leu Tyr	Tyr Tyr Cys Ser Leu Arg	
4815	4820	4825	
ggg att cca agg	cag tgc atc cct gag	gtt gct gga gac ctc atc	14548
Gly Ile Pro Arg	Gln Cys Ile Pro Glu	Val Ala Gly Asp Leu Ile	
4830	4835	4840	
agg cgc tta cac	ctc gaa gcc cac gcg	gac aaa cct gtg gcc acc	14593
Arg Arg Leu His	Leu Glu Ala His Ala	Asp Lys Pro Val Ala Thr	
4845	4850	4855	
tac agt ggg gga	acc aag cgg aaa ctc	tct aca gcc ctg gcc ctg	14638
Tyr Ser Gly Gly	Thr Lys Arg Lys Leu	Ser Thr Ala Leu Ala Leu	
4860	4865	4870	
gtg ggg aaa cct	gac att ctt tta ttg	gat gag ccc agc tct ggg	14683
Val Gly Lys Pro	Asp Ile Leu Leu Leu	Asp Glu Pro Ser Ser Gly	
4875	4880	4885	
atg gat ccc tgc	tct aag cgg tac ctg	tgg caa aca ata atg aag	14728
Met Asp Pro Cys	Ser Lys Arg Tyr Leu	Trp Gln Thr Ile Met Lys	
4890	4895	4900	
gag gtt cgg gaa	ggc tgt gct gcg gtg	ctg acc tcc cac agc atg	14773
Glu Val Arg Glu	Gly Cys Ala Ala Val	Leu Thr Ser His Ser Met	
4905	4910	4915	
gag gag tgt gag	gct ctt tgc aca aga	ctg gcc ata atg gtt aac	14818
Glu Glu Cys Glu	Ala Leu Cys Thr Arg	Leu Ala Ile Met Val Asn	
4920	4925	4930	
ggc agc ttc aaa	tgt ctt ggt tct cct	cag cac atc aaa aat agg	14863
Gly Ser Phe Lys	Cys Leu Gly Ser Pro	Gln His Ile Lys Asn Arg	
4935	4940	4945	

ttt ggt gat ggt	tat aca gtc aaa gtt	tgg ctc tgt aag gaa gca	14908
Phe Gly Asp Gly	Tyr Thr Val Lys Val	Trp Leu Cys Lys Glu Ala	
4950	4955	4960	
aat caa cat tgc	act gtt tct gac cac	ttg aag ctt tat ttt cca	14953
Asn Gln His Cys	Thr Val Ser Asp His	Leu Lys Leu Tyr Phe Pro	
4965	4970	4975	
gga att cag ttc	aag gga cag cac ctg	aat tta tta gaa tat cat	14998
Gly Ile Gln Phe	Lys Gly Gln His Leu	Asn Leu Leu Glu Tyr His	
4980	4985	4990	
gtg cca aaa aga	tgg gga tgc cta gct	gac ttg ttc aaa gtt ata	15043
Val Pro Lys Arg	Trp Gly Cys Leu Ala	Asp Leu Phe Lys Val Ile	
4995	5000	5005	
gag aac aat aaa	acc ttc ttg aat att	aag cat tat tcc att aac	15088
Glu Asn Asn Lys	Thr Phe Leu Asn Ile	Lys His Tyr Ser Ile Asn	
5010	5015	5020	
caa acc act ttg	gag cag gta ttt att	aat ttt gct tct gag cag	15133
Gln Thr Thr Leu	Glu Gln Val Phe Ile	Asn Phe Ala Ser Glu Gln	
5025	5030	5035	
cag caa act cta	caa tct act ctt gat	cca tcc act gac agt cac	15178
Gln Gln Thr Leu	Gln Ser Thr Leu Asp	Pro Ser Thr Asp Ser His	
5040	5045	5050	
cac aca cat cac	ttg ccc atc tga gcactaaaga	agtttccata aggaataaaa	15232
His Thr His His	Leu Pro Ile		
5055			
ccttgtcttc cattacaatt	aacagtcgaag gataaaacaa	gcacgcgcac aatcaaggag	15292
ctggaacaca ctctccaggc	cgtcaaatta ttctcttggt	cattttctat tttgaatctc	15352
cttgtagtt aataaccacc	aatggaaag gtcattcttt	ctgcagactt ttggggagct	15412
cctocaaaac atttgttctc	tttaccatgc cagatggaca	ccagcttctt tgtgacaaag	15472
gcatgaatga tttgacagtg	tccaaactga gacattctgg	agctggaaag cctgtcacac	15532
tagagtgtgt gtgacatgtc	cactctaaac atgtcacttt	tctgttaaga aaactgagcc	15592
ccctccccac aggttaaaaa	acttttagtaa cttgtttgta	tagaaaatag taacaaggac	15652
tattttctat tgttgtcatc	tatttactag atacatgttt	ttaatgattt taatgtaagc	15712
ttttattaat actgatgaca	ttatatggta tgatatgaaa	aaatcaccaa ttttttacat	15772
ataaaagata cctttttaaa	aaaatagggtt ttaagagctc	tttttagtata cacttttagca	15832
aaattaatta aattgaacta	gttactctgt atcaattaca	gtagttctac cagaatctcc	15892
caggttataa tttatgaggg	tagagaaata aaatgtagat	gcattttctt tttcttcatt	15952
tggatgaata attactgttt	tttgttatc taagtcagtg	tttttcaaag cgtagtggtc	16012
cccatattag ctgcattgcc	atcttctggg agcttgcaag	aaatgtacat tctcaggatc	16072

cactccagac ctattgaatc tcaaattctg gggcttaaac aagcacattc caagttaaga 16132
 accaatgacc taagggaatg tctggttacc tcctagttat acaagcaaaa tctgcatagt 16192
 atgtagtctt ttttatttat ttattttttt tttttgaggc ggagtctcgc tctgtcaccc 16252
 aggotggagt gcagtggcgc gatgtcggct cactgcaagc tccgcctcct gggttcacgc 16312
 cattctcctg cctcagcctc cccagcagct gggactacag gcacacatcg ccacacccgg 16372
 ctaattttta gtatttttag tagagacggg gtttcaccgt gttagccagg atggctctcta 16432
 tctcctgacc ttgtgatccg cccccctcca cctcccaaag tgctgggatt acaggcgtga 16492
 gccaccgtgt ccggccgtag tttattttta aatatatttt taaaagcttt gtaaaaatta 16552
 tgtcattctc agaattgttg tcttcaaagc attgtcagat gtagagtgtc cagatgtggc 16612
 tcttaaagac tatatacatc tgaatttttc atcctatagt tagtaagatg cataaaatca 16672
 atccactact gaaatagttt ccagtcagac atttctgagt tcagacattt ctcaacattc 16732
 ttttaacaac atttttctga atcctcaata gaaaatcaca ttaatcttat tttaaaattt 16792
 ggocctttttc aacactaacg ttgagtaccg gtagctttgt gatcaaaggc atatacttcc 16852
 ttatgagatt tctttactaa agcaagattt cattaaatct ctatttccta aatatcattc 16912
 tatacaaaag atatttttta aacggtaagg attaagacaa tcaactgatag ctttgttgtg 16972
 agcaattttg attcccatgt atcacatgaa ttacacttcc ttaaataaat tacagtttat 17032
 ggotgtatga tttattctct aattctaaca tagtctagtt gtcaaaagga aatatgtaat 17092
 ctttttatga ttgttgaatc aataaatacc aatttgtgaa acatgaatgt gtttaaaactg 17152
 cagtgaataa atgagatgtg ctttaattta acccaaaaaa aaaaaaaaaa aaaaaaa 17209

<210> 2
 <211> 5058
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> misc_feature
 <222> (283)..(283)
 <223> The 'Xaa' at location 283 stands for Gly, or Asp.

<220>
 <221> misc_feature
 <222> (2674)..(2674)
 <223> The 'Xaa' at location 2674 stands for Arg, or Trp.

<400> 2

Met Gly His Ala Gly Cys Gln Phe Lys Ala Leu Leu Trp Lys Asn Trp
 1 5 10 15

Leu Cys Arg Leu Arg Asn Pro Val Leu Phe Leu Ala Glu Phe Phe Trp

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

Gln Tyr Asp Leu Lys Ser Gln Phe Gly Phe Xaa Asp Leu His Thr Glu
 275 280 285

Gln Ile Leu Asn Ser Ser Ala Glu Leu Lys Glu Ile Pro Thr Asp Thr
 290 295 300

Ser Leu Glu Lys Met Val Cys Ser Val Leu Ser Ser Thr Ser Glu Asp
 305 310 315 320

Glu Ala Glu Lys Trp Gly His Val Gly Gly Cys His Pro Lys Trp Ser
 325 330 335

Glu Ala Lys Asn Tyr Leu Val His Ala Val Ser Trp Leu Arg Val Tyr
 340 345 350

Gln Gln Val Phe Val Gln Trp Gln Gln Gly Ser Leu Leu Gln Lys Thr
 355 360 365

Leu Thr Gly Met Gly His Ser Leu Glu Ala Leu Arg Asn Gln Phe Glu
 370 375 380

Glu Glu Ser Lys Pro Trp Lys Val Val Glu Ala Leu His Thr Ala Leu
 385 390 395 400

Leu Leu Leu Asn Asp Ser Leu Ser Ala Asp Gly Pro Lys Asp Asn His
 405 410 415

Thr Phe Pro Lys Ile Leu Gln His Leu Trp Lys Leu Gln Ser Leu Leu
 420 425 430

Gln Asn Leu Pro Gln Trp Pro Ala Leu Lys Arg Phe Leu Gln Leu Asp
 435 440 445

Gly Ala Leu Arg Asn Ala Ile Ala Gln Asn Leu His Phe Val Gln Glu
 450 455 460

Val Leu Ile Cys Leu Glu Thr Ser Ala Asn Asp Phe Lys Trp Phe Glu
 465 470 475 480

Leu Asn Gln Leu Lys Leu Glu Lys Asp Val Phe Phe Trp Glu Leu Lys
 485 490 495

Gln Met Leu Ala Lys Asn Ala Val Cys Pro Asn Gly Arg Phe Ser Glu
 500 505 510

Lys Glu Val Phe Leu Pro Pro Gly Asn Ser Ser Ile Trp Gly Gly Leu
 515 520 525

Gln Gly Leu Leu Cys Tyr Cys Asn Ser Ser Glu Thr Ser Val Leu Asn
 530 535 540

Lys Leu Leu Gly Ser Val Glu Asp Ala Asp Arg Ile Leu Gln Glu Val
 545 550 555 560

Ile Thr Trp His Lys Asn Met Ser Val Leu Ile Pro Glu Glu Tyr Leu
 565 570 575

Asp Trp Gln Glu Leu Glu Met Gln Leu Ser Glu Ala Ser Leu Ser Cys
 580 585 590

Thr Arg Leu Phe Leu Leu Leu Gly Ala Asp Pro Ser Pro Glu Asn Asp
 595 600 605

Val Phe Ser Ser Asp Cys Lys His Gln Leu Val Ser Thr Val Ile Phe
 610 615 620

His Thr Leu Glu Lys Thr Gln Phe Phe Leu Glu Gln Ala Tyr Tyr Trp
 625 630 635 640

Lys Ala Phe Lys Lys Phe Ile Arg Lys Thr Cys Glu Val Ala Gln Tyr
 645 650 655

Val Asn Met Gln Glu Ser Phe Gln Asn Arg Leu Leu Ala Phe Pro Glu
 660 665 670

Glu Ser Pro Cys Phe Glu Glu Asn Met Asp Trp Lys Met Ile Ser Asp
 675 680 685

Asn Tyr Phe Gln Phe Leu Asn Asn Leu Leu Lys Ser Pro Thr Ala Ser
 690 695 700

Ile Ser Arg Ala Leu Asn Phe Thr Lys His Leu Leu Met Met Glu Lys
 705 710 715 720

Lys Leu His Thr Leu Glu Asp Glu Gln Met Asn Phe Leu Leu Ser Phe
 725 730 735

Val Glu Phe Phe Glu Lys Leu Leu Leu Pro Asn Leu Phe Asp Ser Ser
 740 745 750

Ile Val Pro Ser Phe His Ser Leu Pro Ser Leu Thr Glu Asp Ile Leu
 755 760 765

Asn Ile Ser Ser Leu Trp Thr Asn His Leu Lys Ser Leu Lys Arg Asp
 770 775 780

Pro Ser Ala Thr Asp Ala Gln Lys Leu Leu Glu Phe Gly Asn Glu Val
 785 790 795 800

Ile Trp Lys Met Gln Thr Leu Gly Ser His Trp Ile Arg Lys Glu Pro
 805 810 815

Lys Asn Leu Leu Arg Phe Ile Glu Leu Ile Leu Phe Glu Ile Asn Pro
 820 825 830

Lys Leu Leu Glu Leu Trp Ala Tyr Gly Ile Ser Lys Gly Lys Arg Ala
 835 840 845

Lys Leu Glu Asn Phe Phe Thr Leu Leu Asn Phe Ser Val Pro Glu Asn
 850 855 860

Glu Ile Leu Ser Thr Ser Phe Asn Phe Ser Gln Leu Phe His Ser Asp
 865 870 875 880

Trp Pro Lys Ser Pro Ala Met Asn Ile Asp Phe Val Arg Leu Ser Glu
 885 890 895

Ala Ile Ile Thr Ser Leu His Glu Phe Gly Phe Leu Glu Gln Glu Gln
 900 905 910

Ile Ser Glu Ala Leu Asn Thr Val Tyr Ala Ile Arg Asn Ala Ser Asp
 915 920 925

Leu Phe Ser Ala Leu Ser Glu Pro Gln Lys Gln Glu Val Asp Lys Ile
 930 935 940

Leu Thr His Ile His Leu Asn Val Phe Gln Asp Lys Asp Ser Ala Leu
 945 950 955 960

Leu Leu Gln Ile Tyr Ser Ser Phe Tyr Arg Tyr Ile Tyr Glu Leu Leu
 965 970 975

Asn Ile Gln Ser Arg Gly Ser Ser Leu Thr Phe Leu Thr Gln Ile Ser
 980 985 990

Lys His Ile Leu Asp Ile Ile Lys Gln Phe Asn Phe Gln Asn Ile Ser
 995 1000 1005

Lys Ala Phe Ala Phe Leu Phe Lys Thr Ala Glu Val Leu Gly Gly
 1010 1015 1020

Ile Ser Asn Val Ser Tyr Cys Gln Gln Leu Leu Ser Ile Phe Asn
 1025 1030 1035

Phe Leu Glu Leu Gln Ala Gln Ser Phe Met Ser Thr Glu Gly Gln
 1040 1045 1050

Glu Leu Glu Val Ile His Thr Thr Leu Thr Gly Leu Lys Gln Leu
 1055 1060 1065

Leu Ile Ile Asp Glu Asp Phe Arg Ile Ser Leu Phe Gln Tyr Met
 1070 1075 1080

Ser Gln Phe Phe Asn Ser Ser Val Glu Asp Leu Leu Asp Asn Lys
 1085 1090 1095

Cys Leu Ile Ser Asp Asn Lys His Ile Ser Ser Val Asn Tyr Ser
 1100 1105 1110

Thr Ser Glu Glu Ser Ser Phe Val Phe Pro Leu Ala Gln Ile Phe
 1115 1120 1125

Ser Asn Leu Ser Ala Asn Val Ser Val Phe Asn Lys Phe Met Ser
 1130 1135 1140

Ile His Cys Thr Val Ser Trp Leu Gln Met Trp Thr Glu Ile Trp
 1145 1150 1155

Glu Thr Ile Ser Gln Leu Phe Lys Phe Asp Met Asn Val Phe Thr
 1160 1165 1170

Ser Leu His His Gly Phe Thr Gln Leu Leu Asp Glu Leu Glu Asp
 1175 1180 1185

Asp Val Lys Val Ser Lys Ser Cys Gln Gly Ile Leu Pro Thr His
 1190 1195 1200

Asn Val Ala Arg Leu Ile Leu Asn Leu Phe Lys Asn Val Thr Gln
 1205 1210 1215

Ala	Asn	Asp	Phe	His	Asn	Trp	Glu	Asp	Phe	Leu	Asp	Leu	Arg	Asp
1220						1225					1230			
Phe	Leu	Val	Ala	Leu	Gly	Asn	Ala	Leu	Val	Ser	Val	Lys	Lys	Leu
1235						1240					1245			
Asn	Leu	Glu	Gln	Val	Glu	Lys	Ser	Leu	Phe	Thr	Met	Glu	Ala	Ala
1250						1255					1260			
Leu	His	Gln	Leu	Lys	Thr	Phe	Pro	Phe	Asn	Glu	Ser	Thr	Ser	Arg
1265						1270					1275			
Glu	Phe	Leu	Asn	Ser	Leu	Leu	Glu	Val	Phe	Ile	Glu	Phe	Ser	Ser
1280						1285					1290			
Thr	Ser	Glu	Tyr	Ile	Val	Arg	Asn	Leu	Asp	Ser	Ile	Asn	Asp	Phe
1295						1300					1305			
Leu	Ser	Asn	Asn	Leu	Thr	Asn	Tyr	Gly	Glu	Lys	Phe	Glu	Asn	Ile
1310						1315					1320			
Ile	Thr	Glu	Leu	Arg	Glu	Ala	Ile	Val	Phe	Leu	Arg	Asn	Val	Ser
1325						1330					1335			
His	Asp	Arg	Asp	Leu	Phe	Ser	Cys	Ala	Asp	Ile	Phe	Gln	Asn	Val
1340						1345					1350			
Thr	Glu	Cys	Ile	Leu	Glu	Asp	Gly	Phe	Leu	Tyr	Val	Asn	Thr	Ser
1355						1360					1365			
Gln	Arg	Met	Leu	Arg	Ile	Leu	Asp	Thr	Leu	Asn	Ser	Thr	Phe	Ser
1370						1375					1380			
Ser	Glu	Asn	Thr	Ile	Ser	Ser	Leu	Lys	Gly	Cys	Ile	Val	Trp	Leu
1385						1390					1395			
Asp	Val	Ile	Asn	His	Leu	Tyr	Leu	Leu	Ser	Asn	Ser	Ser	Phe	Ser
1400						1405					1410			
Gln	Gly	Arg	Leu	Gln	Asn	Ile	Leu	Gly	Asn	Phe	Arg	Asp	Ile	Glu
1415						1420					1425			
Asn	Lys	Met	Asn	Ser	Ile	Leu	Lys	Ile	Val	Thr	Trp	Val	Leu	Asn
1430						1435					1440			
Ile	Lys	Lys	Pro	Leu	Cys	Ser	Ser	Asn	Gly	Ser	His	Ile	Asn	Cys

1445		1450		1455
Val Asn Ile Tyr Leu Lys Asp	Val Thr Asp Phe Leu Asn Ile Val			
1460	1465		1470	
Leu Thr Thr Val Phe Glu Lys	Glu Lys Lys Pro Lys Phe Glu Ile			
1475	1480		1485	
Leu Leu Ala Leu Leu Asn Asp	Ser Thr Lys Gln Val Arg Met Ser			
1490	1495		1500	
Ile Asn Asn Leu Thr Thr Asp	Phe Asp Phe Ala Ser Gln Ser Asn			
1505	1510		1515	
Trp Arg Tyr Phe Thr Glu Leu	Ile Leu Arg Pro Ile Glu Met Ser			
1520	1525		1530	
Asp Glu Ile Pro Asn Gln Phe	Gln Asn Ile Trp Leu His Leu Ile			
1535	1540		1545	
Thr Leu Gly Lys Glu Phe Gln	Lys Leu Val Lys Gly Ile Tyr Phe			
1550	1555		1560	
Asn Ile Leu Glu Asn Asn Ser	Ser Ser Lys Thr Glu Asn Leu Leu			
1565	1570		1575	
Asn Ile Phe Ala Thr Ser Pro	Lys Glu Lys Asp Val Asn Ser Val			
1580	1585		1590	
Gly Asn Ser Ile Tyr His Leu	Ala Ser Tyr Leu Ala Phe Ser Leu			
1595	1600		1605	
Ser His Asp Leu Gln Asn Ser	Pro Lys Ile Ile Ile Ser Pro Glu			
1610	1615		1620	
Ile Met Lys Ala Thr Gly Leu	Gly Ile Gln Leu Ile Arg Asp Val			
1625	1630		1635	
Phe Asn Ser Leu Met Pro Val	Val His His Thr Ser Pro Gln Asn			
1640	1645		1650	
Ala Gly Tyr Met Gln Ala Leu	Lys Lys Val Thr Ser Val Met Arg			
1655	1660		1665	
Thr Leu Lys Lys Ala Asp Ile	Asp Leu Leu Val Asp Gln Leu Glu			

1670		1675		1680
Gln Val Ser Val Asn Leu Met Asp Phe Phe Lys Asn Ile Ser Ser 1685 1690 1695				
Val Gly Thr Gly Asn Leu Val Val Asn Leu Leu Val Gly Leu Met 1700 1705 1710				
Glu Lys Phe Ala Asp Ser Ser His Ser Trp Asn Val Asn His Leu 1715 1720 1725				
Leu Gln Leu Ser Arg Leu Phe Pro Lys Asp Val Val Asp Ala Val 1730 1735 1740				
Ile Asp Val Tyr Tyr Val Leu Pro His Ala Val Arg Leu Leu Gln 1745 1750 1755				
Gly Val Pro Gly Lys Asn Ile Thr Glu Gly Leu Lys Asp Val Tyr 1760 1765 1770				
Ser Phe Thr Leu Leu His Gly Ile Thr Ile Ser Asn Ile Thr Lys 1775 1780 1785				
Glu Asp Phe Ala Ile Val Ile Lys Ile Leu Leu Asp Thr Ile Glu 1790 1795 1800				
Leu Val Ser Asp Lys Pro Asp Ile Ile Ser Glu Ala Leu Ala Cys 1805 1810 1815				
Phe Pro Val Val Trp Cys Trp Asn His Thr Asn Ser Gly Phe Arg 1820 1825 1830				
Gln Asn Ser Lys Ile Asp Pro Cys Asn Val His Gly Leu Met Ser 1835 1840 1845				
Ser Ser Phe Tyr Gly Lys Val Ala Ser Ile Leu Asp His Phe His 1850 1855 1860				
Leu Ser Pro Gln Gly Glu Asp Ser Pro Cys Ser Asn Glu Ser Ser 1865 1870 1875				
Arg Met Glu Ile Thr Arg Lys Val Val Cys Ile Ile His Glu Leu 1880 1885 1890				
Val Asp Trp Asn Ser Ile Leu Leu Glu Leu Ser Glu Val Phe His 1895 1900 1905				

Val	Asn	Ile	Ser	Leu	Val	Lys	Thr	Val	Gln	Lys	Phe	Trp	His	Lys
1910						1915					1920			
Ile	Leu	Pro	Phe	Val	Pro	Pro	Ser	Ile	Asn	Gln	Thr	Arg	Asp	Ser
1925						1930					1935			
Ile	Ser	Glu	Leu	Cys	Pro	Ser	Gly	Ser	Ile	Lys	Gln	Val	Ala	Leu
1940						1945					1950			
Gln	Ile	Ile	Glu	Lys	Leu	Lys	Asn	Val	Asn	Phe	Thr	Lys	Val	Thr
1955						1960					1965			
Ser	Gly	Glu	Asn	Ile	Leu	Asp	Lys	Leu	Ser	Ser	Leu	Asn	Lys	Ile
1970						1975					1980			
Leu	Asn	Ile	Asn	Glu	Asp	Thr	Glu	Thr	Ser	Val	Gln	Asn	Ile	Ile
1985						1990					1995			
Ser	Ser	Asn	Leu	Glu	Arg	Thr	Val	Gln	Leu	Ile	Ser	Glu	Asp	Trp
2000						2005					2010			
Ser	Leu	Glu	Lys	Ser	Thr	His	Asn	Leu	Leu	Ser	Leu	Phe	Met	Met
2015						2020					2025			
Leu	Gln	Asn	Ala	Asn	Val	Thr	Gly	Ser	Ser	Leu	Glu	Ala	Leu	Ser
2030						2035					2040			
Ser	Phe	Ile	Glu	Lys	Ser	Glu	Thr	Pro	Tyr	Asn	Phe	Glu	Glu	Leu
2045						2050					2055			
Trp	Pro	Lys	Phe	Gln	Gln	Ile	Met	Lys	Asp	Leu	Thr	Gln	Asp	Phe
2060						2065					2070			
Arg	Ile	Arg	His	Leu	Leu	Ser	Glu	Met	Asn	Lys	Gly	Ile	Lys	Ser
2075						2080					2085			
Ile	Asn	Ser	Met	Ala	Leu	Gln	Lys	Ile	Thr	Leu	Gln	Phe	Ala	His
2090						2095					2100			
Phe	Leu	Glu	Ile	Leu	Asp	Ser	Pro	Ser	Leu	Lys	Thr	Leu	Glu	Ile
2105						2110					2115			
Ile	Glu	Asp	Phe	Leu	Leu	Val	Thr	Lys	Asn	Trp	Leu	Gln	Glu	Tyr
2120						2125					2130			

Ala	Asn	Glu	Asp	Tyr	Ser	Arg	Met	Ile	Glu	Thr	Leu	Phe	Ile	Pro
2135							2140				2145			
Val	Thr	Asn	Glu	Ser	Ser	Thr	Glu	Asp	Ile	Ala	Leu	Leu	Ala	Lys
2150						2155					2160			
Ala	Ile	Ala	Thr	Phe	Trp	Gly	Ser	Leu	Lys	Asn	Ile	Ser	Arg	Ala
2165						2170					2175			
Gly	Asn	Phe	Asp	Val	Ala	Phe	Leu	Thr	His	Leu	Leu	Asn	Gln	Glu
2180						2185					2190			
Gln	Leu	Thr	Asn	Phe	Ser	Val	Val	Gln	Leu	Leu	Phe	Glu	Asn	Ile
2195						2200					2205			
Leu	Ile	Asn	Leu	Ile	Asn	Asn	Leu	Ala	Gly	Asn	Ser	Gln	Glu	Ala
2210						2215					2220			
Ala	Trp	Asn	Leu	Asn	Asp	Thr	Asp	Leu	Gln	Ile	Met	Asn	Phe	Ile
2225						2230					2235			
Asn	Leu	Ile	Leu	Asn	His	Met	Gln	Ser	Glu	Thr	Ser	Arg	Lys	Thr
2240						2245					2250			
Val	Leu	Ser	Leu	Arg	Ser	Ile	Val	Asp	Phe	Thr	Glu	Gln	Phe	Leu
2255						2260					2265			
Lys	Thr	Phe	Phe	Ser	Leu	Phe	Leu	Lys	Glu	Asp	Ser	Glu	Asn	Lys
2270						2275					2280			
Ile	Ser	Leu	Leu	Leu	Lys	Tyr	Phe	His	Lys	Asp	Val	Ile	Ala	Glu
2285						2290					2295			
Met	Ser	Phe	Val	Pro	Lys	Asp	Lys	Ile	Leu	Glu	Ile	Leu	Lys	Leu
2300						2305					2310			
Asp	Gln	Phe	Leu	Thr	Leu	Met	Ile	Gln	Asp	Arg	Leu	Met	Asn	Ile
2315						2320					2325			
Phe	Ser	Ser	Leu	Lys	Glu	Thr	Ile	Tyr	His	Leu	Met	Lys	Ser	Ser
2330						2335					2340			
Phe	Ile	Leu	Asp	Asn	Gly	Glu	Phe	Tyr	Phe	Asp	Thr	His	Gln	Gly
2345						2350					2355			

Leu Lys Phe Met Gln Asp Leu Phe Asn Ala Leu Leu Arg Glu Thr
 2360 2365 2370
 Ser Met Lys Asn Lys Thr Glu Asn Asn Ile Asp Phe Phe Thr Val
 2375 2380 2385
 Val Ser Gln Leu Phe Phe His Val Asn Lys Ser Glu Asp Leu Phe
 2390 2395 2400
 Lys Leu Asn Gln Asp Leu Gly Ser Ala Leu His Leu Val Arg Glu
 2405 2410 2415
 Cys Ser Thr Glu Met Ala Arg Leu Leu Asp Thr Ile Leu His Ser
 2420 2425 2430
 Pro Asn Lys Asp Phe Tyr Ala Leu Tyr Pro Thr Leu Gln Glu Val
 2435 2440 2445
 Ile Leu Ala Asn Leu Thr Asp Leu Leu Phe Phe Ile Asn Asn Ser
 2450 2455 2460
 Phe Pro Leu Arg Asn Arg Ala Thr Leu Glu Ile Thr Lys Arg Leu
 2465 2470 2475
 Val Gly Ala Ile Ser Arg Ala Ser Glu Glu Ser His Val Leu Lys
 2480 2485 2490
 Pro Leu Leu Glu Met Ser Gly Thr Leu Val Met Leu Leu Asn Asp
 2495 2500 2505
 Ser Ala Asp Leu Arg Asp Leu Ala Thr Ser Met Asp Ser Ile Val
 2510 2515 2520
 Lys Leu Leu Lys Leu Val Lys Lys Val Ser Gly Lys Met Ser Thr
 2525 2530 2535
 Val Phe Lys Thr His Phe Ile Ser Asn Thr Lys Asp Ser Val Lys
 2540 2545 2550
 Phe Phe Asp Thr Leu Tyr Ser Ile Met Gln Gln Ser Val Gln Asn
 2555 2560 2565
 Leu Val Lys Glu Ile Ala Thr Leu Lys Lys Ile Asp His Phe Thr
 2570 2575 2580

Phe Glu Lys Ile Asn Asp Leu Leu Val Pro Phe Leu Asp Leu Ala
 2585 2590 2595
 Phe Glu Met Ile Gly Val Glu Pro Tyr Ile Ser Ser Asn Ser Asp
 2600 2605 2610
 Ile Phe Ser Met Ser Pro Ser Ile Leu Ser Tyr Met Asn Gln Ser
 2615 2620 2625
 Lys Asp Phe Ser Asp Ile Leu Glu Glu Ile Ala Glu Phe Leu Thr
 2630 2635 2640
 Ser Val Lys Met Asn Leu Glu Asp Met Arg Ser Leu Ala Val Ala
 2645 2650 2655
 Phe Asn Asn Glu Thr Gln Thr Phe Ser Met Asp Ser Val Asn Leu
 2660 2665 2670
 Xaa Glu Glu Ile Leu Gly Cys Leu Val Pro Ile Asn Asn Ile Thr
 2675 2680 2685
 Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro Ile Ser Thr His Ser
 2690 2695 2700
 Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His Glu Val Ile Leu
 2705 2710 2715
 Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu Ile Gly Ser
 2720 2725 2730
 Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu Trp Lys
 2735 2740 2745
 Asn Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met Thr
 2750 2755 2760
 Phe Thr Gln His Pro Asn Asn Leu Leu Lys Thr Ile Glu Thr Val
 2765 2770 2775
 Leu Glu Ala Ser Ser Gly Ile Lys Ser Asp Tyr Glu Gly Asp Leu
 2780 2785 2790
 Asn Lys Ser Leu Tyr Phe Asp Thr Pro Leu Ser Gln Asn Ile Thr
 2795 2800 2805

His	His	Gln	Leu	Glu	Lys	Ala	Ile	His	Asn	Val	Leu	Ser	Arg	Ile
2810						2815					2820			
Ala	Leu	Trp	Arg	Lys	Gly	Leu	Arg	Phe	Asn	Asn	Ser	Glu	Trp	Ile
2825						2830					2835			
Thr	Ser	Thr	Arg	Thr	Leu	Phe	Gln	Pro	Leu	Phe	Glu	Ile	Phe	Ile
2840						2845					2850			
Lys	Ala	Thr	Thr	Gly	Lys	Asn	Val	Thr	Ser	Glu	Lys	Glu	Glu	Arg
2855						2860					2865			
Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	Pro	Tyr	Ser	Phe	Lys	Pro	Phe
2870						2875					2880			
Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	Gly	Leu	Phe	Val	Leu	Thr	Lys
2885						2890					2895			
Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	Asp	Gln	Ser	Val	Val	Glu	Ile
2900						2905					2910			
Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	Lys	Pro	Ser	Glu	Ala	Met	Glu
2915						2920					2925			
Met	Leu	Gln	Lys	Val	Lys	Met	Met	Val	Val	Arg	Val	Leu	Thr	Ile
2930						2935					2940			
Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	Lys	Asp	Ile	Leu	Cys	Ala	Thr
2945						2950					2955			
Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	Arg	His	Leu	Ile	Leu	Ser	Ala
2960						2965					2970			
Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	Asp	His	Phe	Gln	Glu	Ile	Glu
2975						2980					2985			
Lys	Ile	Trp	Ser	Ser	Pro	Asn	Gln	Leu	Asn	Cys	Glu	Ser	Leu	Ser
2990						2995					3000			
Lys	Asn	Leu	Ser	Ser	Thr	Leu	Glu	Ser	Phe	Lys	Ser	Ser	Leu	Glu
3005						3010					3015			
Asn	Ala	Thr	Gly	Gln	Asp	Cys	Thr	Ser	Gln	Pro	Arg	Leu	Glu	Thr
3020						3025					3030			
Val	Gln	Gln	His	Leu	Tyr	Met	Leu	Ala	Lys	Ser	Leu	Glu	Glu	Thr

3035		3040		3045	
Trp Ser Ser Gly Asn Pro Ile Met Thr Phe Leu Ser Asn Phe Thr	3050	3055		3060	
Val Thr Glu Asp Val Lys Ile Lys Asp Leu Met Lys Asn Ile Thr	3065	3070		3075	
Lys Leu Thr Glu Glu Leu Arg Ser Ser Ile Gln Ile Ser Asn Glu	3080	3085		3090	
Thr Ile His Ser Ile Leu Glu Ala Asn Ile Ser His Ser Lys Val	3095	3100		3105	
Leu Phe Ser Ala Leu Thr Val Ala Leu Ser Gly Lys Cys Asp Gln	3110	3115		3120	
Glu Ile Leu His Leu Leu Leu Thr Phe Pro Lys Gly Glu Lys Ser	3125	3130		3135	
Trp Ile Ala Ala Glu Glu Leu Cys Ser Leu Pro Gly Ser Lys Val	3140	3145		3150	
Tyr Ser Leu Ile Val Leu Leu Ser Arg Asn Leu Asp Val Arg Ala	3155	3160		3165	
Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu Ala Asn Gly Leu Leu	3170	3175		3180	
Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser Ala Leu Leu Ala	3185	3190		3195	
Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe Leu His Thr	3200	3205		3210	
Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln Gln Val	3215	3220		3225	
Ser Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe Gln	3230	3235		3240	
Phe Val Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser	3245	3250		3255	
Asp Ser Asn Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu	3260	3265		3270	

Glu 3275	Asp	Lys	Glu	Lys	Phe 3280	Asn	Ile	Pro	Glu	Asp 3285	Ser	Thr	Pro
Phe 3290	Cys	Leu	Lys	Leu	Tyr	Gln 3295	Glu	Ile	Leu	Gln 3300	Pro	Asn	Gly
Ala 3305	Leu	Val	Trp	Thr	Phe	Leu 3310	Lys	Pro	Ile	Leu 3315	Gly	Lys	Ile
Leu 3320	Tyr	Thr	Pro	Asn	Thr	Pro 3325	Glu	Ile	Asn	Lys 3330	Ile	Gln	Lys
Ala 3335	Asn	Tyr	Thr	Phe	Tyr	Ile 3340	Val	Asp	Lys	Leu 3345	Thr	Leu	Ser
Glu 3350	Thr	Leu	Leu	Glu	Met	Ser 3355	Ser	Leu	Phe	Gln 3360	Ser	Gly	Ser
Gly 3365	Gln	Met	Phe	Asn	Gln	Leu 3370	Gln	Glu	Ala	Leu 3375	Asn	Lys	Phe
Val 3380	Arg	Asn	Phe	Val	Glu	Asn 3385	Gln	Leu	His	Ile 3390	Val	Asp	Lys
Leu 3395	Thr	Glu	Lys	Leu	Gln	Thr 3400	Tyr	Gly	Gly	Leu 3405	Asp	Glu	Met
Phe 3410	Asn	His	Ala	Gly	Ala	Gly 3415	Arg	Phe	Arg	Phe 3420	Gly	Ser	Ile
Leu 3425	Val	Asn	Leu	Ser	Ser	Cys 3430	Val	Ala	Leu	Asn 3435	Phe	Gln	Ala
Leu 3440	Gln	Ser	Val	Asp	Ile	Leu 3445	Glu	Thr	Lys	Ala 3450	Glu	Leu	Leu
Gln 3455	Gln	Asn	Ser	Phe	Leu	Ala 3460	Ser	Ile	Ile	Phe 3465	Asn	Ser	Leu
Phe 3470	Asp	Lys	Asn	Phe	Arg	Ser 3475	Glu	Ser	Val	Lys 3480	Pro	Pro	His
Val 3485	Ser	Tyr	Thr	Ile	Arg	Thr 3490	Asn	Val	Leu	Tyr 3495	Val	Arg	Thr

Asp	Val	Val	Lys	Asn	Pro	Ser	Trp	Lys	Phe	His	Pro	Gln	Asn	Leu
3500						3505					3510			

Pro Ala Asp Gly Phe Lys Tyr Asn Tyr Val Phe Ala Pro Leu Gln
3515 3520 3525

Asp Met Ile Glu Arg Ala Ile Ile Leu Val Gln Thr Gly Gln Glu
3530 3535 3540

Ala Leu Glu Pro Ala Ala Gln Thr Gln Ala Ala Pro Tyr Pro Cys
3545 3550 3555

His Thr Ser Asp Leu Phe Leu Asn Asn Val Gly Phe Phe Phe Pro
3560 3565 3570

Leu Ile Met Met Leu Thr Trp Met Val Ser Val Ala Ser Met Val
3575 3580 3585

Arg Lys Leu Val Tyr Glu Gln Glu Ile Gln Ile Glu Glu Tyr Met
3590 3595 3600

Arg Met Met Gly Val His Pro Val Ile His Phe Leu Ala Trp Phe
3605 3610 3615

Leu Glu Asn Met Ala Val Leu Thr Ile Ser Ser Ala Thr Leu Ala
3620 3625 3630

Ile Val Leu Lys Thr Ser Gly Ile Phe Ala His Ser Asn Thr Phe
3635 3640 3645

```

Ile Val  Phe Leu Phe Leu Leu  Asp Phe Gly Met  Ser  Val Val Met
    3650                      3655                3660

```

Leu Ser Tyr Leu Leu Ser Ala Phe Phe Ser Gln Ala Asn Thr Ala
3665 3670 3675

Ala Leu Cys Thr Ser Leu Val Tyr Met Ile Ser Phe Leu Pro Tyr
3680 3685 3690

```

Ile Val  Leu Leu Val  Leu His  Asn Gln  Leu Ser  Phe  Val Asn Gln
    3695                3700                3705

```

Thr Phe Leu Cys Leu Leu Ser Thr Thr Ala Phe Gly Gln Gly Val
3710 3715 3720

Phe Phe Ile Thr Phe Leu Glu Gly Gln Glu Thr Gly Ile Gln Trp
 3725 3730 3735
 Asn Asn Met Tyr Gln Ala Leu Glu Gln Gly Gly Met Thr Phe Gly
 3740 3745 3750
 Trp Val Cys Trp Met Ile Leu Phe Asp Ser Ser Leu Tyr Phe Leu
 3755 3760 3765
 Cys Gly Trp Tyr Leu Ser Asn Leu Ile Pro Gly Thr Phe Gly Leu
 3770 3775 3780
 Arg Lys Pro Trp Tyr Phe Pro Phe Thr Ala Ser Tyr Trp Lys Ser
 3785 3790 3795
 Val Gly Phe Leu Val Glu Lys Arg Gln Tyr Phe Leu Ser Ser Ser
 3800 3805 3810
 Leu Phe Phe Phe Asn Glu Asn Phe Asp Asn Lys Gly Ser Ser Leu
 3815 3820 3825
 Gln Asn Arg Glu Gly Glu Leu Glu Gly Ser Ala Pro Gly Val Thr
 3830 3835 3840
 Leu Val Ser Val Thr Lys Glu Tyr Glu Gly His Lys Ala Val Val
 3845 3850 3855
 Gln Asp Leu Ser Leu Thr Phe Tyr Arg Asp Gln Ile Thr Ala Leu
 3860 3865 3870
 Leu Gly Thr Asn Gly Ala Gly Lys Thr Thr Ile Ile Ser Met Leu
 3875 3880 3885
 Thr Gly Leu His Pro Pro Thr Ser Gly Thr Ile Ile Ile Asn Gly
 3890 3895 3900
 Lys Asn Leu Gln Thr Asp Leu Ser Arg Val Arg Met Glu Leu Gly
 3905 3910 3915
 Val Cys Pro Gln Gln Asp Ile Leu Leu Asp Asn Leu Thr Val Arg
 3920 3925 3930
 Glu His Leu Leu Leu Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr
 3935 3940 3945

Lys Lys Glu Leu His Gln Gln Val Asn Gln Thr Leu Gln Asp Val
 3950 3955 3960
 Asp Leu Thr Gln His Gln His Lys Gln Thr Arg Ala Leu Ser Gly
 3965 3970 3975
 Gly Leu Lys Arg Lys Leu Ser Leu Gly Ile Ala Phe Met Gly Met
 3980 3985 3990
 Ser Arg Thr Val Val Leu Asp Glu Pro Thr Ser Gly Val Asp Pro
 3995 4000 4005
 Cys Ser Arg His Ser Leu Trp Asp Ile Leu Leu Lys Tyr Arg Glu
 4010 4015 4020
 Gly Arg Thr Ile Ile Phe Thr Thr His His Leu Asp Glu Ala Glu
 4025 4030 4035
 Ala Leu Ser Asp Arg Val Ala Val Leu Gln His Gly Arg Leu Arg
 4040 4045 4050
 Cys Cys Gly Pro Pro Phe Cys Leu Lys Glu Ala Tyr Gly Gln Gly
 4055 4060 4065
 Leu Arg Leu Thr Leu Thr Arg Gln Pro Ser Val Leu Glu Ala His
 4070 4075 4080
 Asp Leu Lys Asp Met Ala Cys Val Thr Ser Leu Ile Lys Ile Tyr
 4085 4090 4095
 Ile Pro Gln Ala Phe Leu Lys Asp Ser Ser Gly Ser Glu Leu Thr
 4100 4105 4110
 Tyr Thr Ile Pro Lys Asp Thr Asp Lys Ala Cys Leu Lys Gly Leu
 4115 4120 4125
 Phe Gln Ala Leu Asp Glu Asn Leu His Gln Leu His Leu Thr Gly
 4130 4135 4140
 Tyr Gly Ile Ser Asp Thr Thr Leu Glu Glu Val Phe Leu Met Leu
 4145 4150 4155
 Leu Gln Asp Ser Asn Lys Lys Ser His Ile Ala Leu Gly Thr Glu
 4160 4165 4170

Ser	Glu	Leu	Gln	Asn	His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr
4175						4180					4185			
Cys	Gly	Ser	Leu	Ala	Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu
4190						4195					4200			
Leu	Arg	Ala	Gln	Val	Ala	Ala	Ile	Leu	Ala	Arg	Arg	Leu	Arg	Arg
4205						4210					4215			
Thr	Leu	Arg	Ala	Gly	Lys	Ser	Thr	Leu	Ala	Asp	Leu	Leu	Leu	Pro
4220						4225					4230			
Val	Leu	Phe	Val	Ala	Leu	Ala	Met	Gly	Leu	Phe	Met	Val	Arg	Pro
4235						4240					4245			
Leu	Ala	Thr	Glu	Tyr	Pro	Pro	Leu	Arg	Leu	Thr	Pro	Gly	His	Tyr
4250						4255					4260			
Gln	Arg	Ala	Glu	Thr	Tyr	Phe	Phe	Ser	Ser	Gly	Gly	Asp	Asn	Leu
4265						4270					4275			
Asp	Leu	Thr	Arg	Val	Leu	Leu	Arg	Lys	Phe	Arg	Asp	Gln	Asp	Leu
4280						4285					4290			
Pro	Cys	Ala	Asp	Leu	Asn	Pro	Arg	Gln	Lys	Asn	Ser	Ser	Cys	Trp
4295						4300					4305			
Arg	Thr	Asp	Pro	Phe	Ser	His	Pro	Glu	Phe	Gln	Asp	Ser	Cys	Gly
4310						4315					4320			
Cys	Leu	Lys	Cys	Pro	Asn	Arg	Ser	Ala	Ser	Ala	Pro	Tyr	Leu	Thr
4325						4330					4335			
Asn	His	Leu	Gly	His	Thr	Leu	Leu	Asn	Leu	Ser	Gly	Phe	Asn	Met
4340						4345					4350			
Glu	Glu	Tyr	Leu	Leu	Ala	Pro	Ser	Glu	Lys	Pro	Arg	Leu	Gly	Gly
4355						4360					4365			
Trp	Ser	Phe	Gly	Leu	Lys	Ile	Pro	Ser	Glu	Ala	Gly	Gly	Ala	Asn
4370						4375					4380			
Gly	Asn	Ile	Ser	Lys	Pro	Pro	Thr	Leu	Ala	Lys	Val	Trp	Tyr	Asn
4385						4390					4395			

Gln	Lys	Gly	Phe	His	Ser	Leu	Pro	Ser	Tyr	Leu	Asn	His	Leu	Asn
4400						4405					4410			
Asn	Leu	Ile	Leu	Trp	Gln	His	Leu	Pro	Pro	Thr	Val	Asp	Trp	Arg
4415						4420					4425			
Gln	Tyr	Gly	Ile	Thr	Leu	Tyr	Ser	His	Pro	Tyr	Gly	Gly	Ala	Leu
4430						4435					4440			
Leu	Asn	Glu	Asp	Lys	Ile	Leu	Glu	Ser	Ile	Arg	Gln	Cys	Gly	Val
4445						4450					4455			
Ala	Leu	Cys	Ile	Val	Leu	Gly	Phe	Ser	Ile	Leu	Ser	Ala	Ser	Ile
4460						4465					4470			
Gly	Ser	Ser	Val	Val	Arg	Asp	Arg	Val	Ile	Gly	Ala	Lys	Arg	Leu
4475						4480					4485			
Gln	His	Ile	Ser	Gly	Leu	Gly	Tyr	Arg	Met	Tyr	Trp	Phe	Thr	Asn
4490						4495					4500			
Phe	Leu	Tyr	Asp	Met	Leu	Phe	Tyr	Leu	Val	Ser	Val	Cys	Leu	Cys
4505						4510					4515			
Val	Ala	Val	Ile	Val	Ala	Phe	Gln	Leu	Thr	Ala	Phe	Thr	Phe	Arg
4520						4525					4530			
Lys	Asn	Leu	Ala	Ala	Thr	Ala	Leu	Leu	Leu	Ser	Leu	Phe	Gly	Tyr
4535						4540					4545			
Ala	Thr	Leu	Pro	Trp	Met	Tyr	Leu	Met	Ser	Arg	Ile	Phe	Ser	Ser
4550						4555					4560			
Ser	Asp	Val	Ala	Phe	Ile	Ser	Tyr	Val	Ser	Leu	Asn	Phe	Ile	Phe
4565						4570					4575			
Gly	Leu	Cys	Thr	Met	Pro	Ile	Thr	Ile	Met	Pro	Arg	Leu	Leu	Ala
4580						4585					4590			
Ile	Ile	Ser	Lys	Ala	Lys	Asn	Leu	Gln	Asn	Ile	Tyr	Asp	Val	Leu
4595						4600					4605			
Lys	Trp	Val	Phe	Thr	Ile	Phe	Pro	Gln	Phe	Cys	Leu	Gly	Gln	Gly
4610						4615					4620			

Leu Val	Glu Leu Cys Tyr	Asn	Gln Ile Lys Tyr	Asp	Leu Thr His
4625		4630		4635	
Asn Phe	Gly Ile Asp Ser	Tyr	Val Ser Pro Phe	Glu	Met Asn Phe
4640		4645		4650	
Leu Gly	Trp Ile Phe Val	Gln	Leu Ala Ser Gln	Gly	Thr Val Leu
4655		4660		4665	
Leu Leu	Leu Arg Val Leu	Leu	His Trp Asp Leu	Leu	Arg Trp Pro
4670		4675		4680	
Arg Gly	His Ser Thr Leu	Gln	Gly Thr Val Lys	Ser	Ser Lys Asp
4685		4690		4695	
Thr Asp	Val Glu Lys Glu	Glu	Lys Arg Val Phe	Glu	Gly Arg Thr
4700		4705		4710	
Asn Gly	Asp Ile Leu Val	Leu	Tyr Asn Leu Ser	Lys	His Tyr Arg
4715		4720		4725	
Arg Phe	Phe Gln Asn Ile	Ile	Ala Val Gln Asp	Ile	Ser Leu Gly
4730		4735		4740	
Ile Pro	Lys Gly Glu Cys	Phe	Gly Leu Leu Gly	Val	Asn Gly Ala
4745		4750		4755	
Gly Lys	Ser Thr Thr Phe	Lys	Met Leu Asn Gly	Glu	Val Ser Leu
4760		4765		4770	
Thr Ser	Gly His Ala Ile	Ile	Arg Thr Pro Met	Gly	Asp Ala Val
4775		4780		4785	
Asp Leu	Ser Ser Ala Gly	Thr	Ala Gly Val Leu	Ile	Gly Tyr Cys
4790		4795		4800	
Pro Gln	Gln Asp Ala Leu	Asp	Glu Leu Leu Thr	Gly	Trp Glu His
4805		4810		4815	
Leu Tyr	Tyr Tyr Cys Ser	Leu	Arg Gly Ile Pro	Arg	Gln Cys Ile
4820		4825		4830	
Pro Glu	Val Ala Gly Asp	Leu	Ile Arg Arg Leu	His	Leu Glu Ala
4835		4840		4845	
His Ala	Asp Lys Pro Val	Ala	Thr Tyr Ser Gly	Gly	Thr Lys Arg

4850		4855		4860
Lys Leu Ser Thr Ala Leu Ala	Leu Val Gly Lys Pro	Asp Ile Leu		
4865	4870	4875		
Leu Leu Asp Glu Pro Ser Ser	Gly Met Asp Pro Cys	Ser Lys Arg		
4880	4885	4890		
Tyr Leu Trp Gln Thr Ile Met	Lys Glu Val Arg Glu	Gly Cys Ala		
4895	4900	4905		
Ala Val Leu Thr Ser His Ser	Met Glu Glu Cys Glu	Ala Leu Cys		
4910	4915	4920		
Thr Arg Leu Ala Ile Met Val	Asn Gly Ser Phe Lys	Cys Leu Gly		
4925	4930	4935		
Ser Pro Gln His Ile Lys Asn	Arg Phe Gly Asp Gly	Tyr Thr Val		
4940	4945	4950		
Lys Val Trp Leu Cys Lys Glu	Ala Asn Gln His Cys	Thr Val Ser		
4955	4960	4965		
Asp His Leu Lys Leu Tyr Phe	Pro Gly Ile Gln Phe	Lys Gly Gln		
4970	4975	4980		
His Leu Asn Leu Leu Glu Tyr	His Val Pro Lys Arg	Trp Gly Cys		
4985	4990	4995		
Leu Ala Asp Leu Phe Lys Val	Ile Glu Asn Asn Lys	Thr Phe Leu		
5000	5005	5010		
Asn Ile Lys His Tyr Ser Ile	Asn Gln Thr Thr Leu	Glu Gln Val		
5015	5020	5025		
Phe Ile Asn Phe Ala Ser Glu	Gln Gln Gln Thr Leu	Gln Ser Thr		
5030	5035	5040		
Leu Asp Pro Ser Thr Asp Ser	His His Thr His His	Leu Pro Ile		
5045	5050	5055		

<210> 3
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 3

ctaagagggg tttcaggacg tgac

24

<210> 4

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 4

aaactgtgga catcttcccc gaaac

25

<210> 5

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 5

gttgatactc atccttactt gc

22

<210> 6

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 6

acatatttcc ttttgacaac tagac

25

<210> 7

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 7

ccctctaaga aacagagcaa c

21

<210> 8

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 8 22
ggagacctca tcaggcgctt ac

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 9 20
ttgactcctc cattgttccc

<210> 10
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 10 21
ctcagaaatg cgatagctca g

<210> 11
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 11 18
gactcaggaa cccggtcc

<210> 12
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 12 18
ctgccagtcc tcttcgtg

<210> 13
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 13
gatgtgctga ggagaaccaa g 21

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 14
taccctcccc tcagactcac 20

<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 15
gggcttggttc atggtgag 18

<210> 16
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 16
cagcacctca caaattgccca cc 22

<210> 17
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 17
gcatgggtag ggtgctgcct g 21

<210> 18
<211> 12498
<212> DNA
<213> Homo Sapiens

<220>
<221> misc_feature

<222> (873)..(873)
 <223> Xaa = any amino acid

<220>
 <221> CDS
 <222> (2209)..(10491)

<220>
 <221> misc_feature
 <222> (3334)..(3334)
 <223> N = any nucleic acid residue

<400> 18
 ggactgagag cagggagcag caggcatggg gcatgccggg tgccagttca aagccctgct 60
 gtggaagaat tggctctgca gactcaggaa cccggtcctt ttccttgctg aattcttctg 120
 gccttgtatc ctgtttgtaa ttctgacagt tcttcgtttt caagaacctc ccagatacag 180
 agacatttgt tatttgcagc cccgagatct acccagctgt ggtgttatcc cctttgttca 240
 aagccttctt tgtaacctg gatcaagggt taggaacttc agctatgaag ggtcaatgga 300
 gcatcatttt cgtttgtcta gggtccaaac tgcagctgac cccaagaaag tcaacaacct 360
 ggccttttta aaagagatac aagacctggc agaggaaatt catggaatga tggacaaggc 420
 aaaaaactta aaaagacttt gggtagaacg atccaacact ccagattctt cttatggttc 480
 cagttttttt acaatggatc tcaataagac cgaggaggta atattgaaac tggaaagcct 540
 ccatcagcag cctcatatct gggattttct acttttactg ccgagactac acacaagcca 600
 tgatcatgtg gaagatggca tggatgttgc agtgaacctt ctccagacca ttttgaattc 660
 cttaatatcc ctagaagatt tagattggct tccactcaac caaacttttt cccaggtttc 720
 tgaacttgta ctgaatgtga ccatttcgac actgacattt ctgcagcaac atggagtagc 780
 agtcaccgag ccagtttacc acctgtccat gcagaatata gtgtgggatc cacagaaagt 840
 ccagtatgat ctcaaattcc agtttggcct tgntgatctt cacacggaac agatcctgaa 900
 ctcttcagct gaactgaagg agattccac agacacttcc ttggagaaga tgggtgtgtc 960
 agtcttgtct agcacatcag aggatgaagc tgagaaatgg ggccacgttg gaggctgcca 1020
 ccctaagtgg tcagaagcca aaaactatct tgtccatgca gtcagctggc tgcgagtcta 1080
 ccaacagggtg tttgttcagt ggcaacaggg tagcctgctt cagaagacac tcacaggcat 1140
 gggccatagt ctggaggctc tcaggaatca gtttgaagaa gagagcaagc cctggaagggt 1200
 ggtggaagct ctgcacactg cactgctcct gctgaatgac agcttgtcag cagatggccc 1260
 aaaagataat catacatttc caaagatatt acagcatctg tggaaattgc aaagcttgct 1320
 gcaaaacctg cccagtggtc cggcactgaa gagatttctt cagcttgatg gagctotcag 1380
 aatgcgata gctcagaatt tacattttgt ccaagaagtc ctcatttgcc tggagacatc 1440

agctaatagat	tttaaatggt	ttgaacttaa	ccaattgaaa	ctggaaaagg	atgtgttctt	1500
ttgggagctg	aaacagatgt	tggcgaagaa	tgctgtctgc	ccgaatggtc	gtttctctga	1560
gaaggaggtc	tttttgccgc	ctggaaactc	cagcatatgg	ggtgggtctcc	agggaactgtt	1620
gtgctattgt	aactcctctg	agacgagtgt	tttaacaacag	ctacttggtt	cagtagagga	1680
tgctgatcgt	attttgcaag	aggtcattac	ttggcacaaa	aatatgtcag	ttttaataacc	1740
tgaagaatat	ttggactggc	aggaacttga	gatgcagctg	tcagaagcaa	gcctttcctg	1800
tactcggctc	ttcctgctgc	tgggagctga	tcctctctct	gagaatgatg	tcttttctag	1860
tgactgtaag	caccagcttg	tctccacagt	gatatttcat	acacttgaaa	aaacacaatt	1920
tttcttgcaa	caagcatatt	attggaaaagc	cttcaaaaag	tttatcagga	agacttgcca	1980
agtggcccaa	tatgtaaata	tgcaagagag	tttcagaaac	agactattgg	cttttctctga	2040
ggaatctcct	tgttttgaag	aaaacatgga	ttggaaaatg	atcagtgata	attattttca	2100
atttttgaat	aacttactca	agtctccaac	agcttccata	tcaggaaga	ttctgagaac	2160
aaaatatctc	ttctgctgaa	atattttccac	aaagatgtta	ttgcagag	atg agt ttt	2217
				Met Ser Phe		

gtc cca aaa gat aaa att cta gaa att ctg aaa ctg gat caa ttt ctt 2265
Val Pro Lys Asp Lys Ile Leu Glu Ile Leu Lys Leu Asp Gln Phe Leu
5 10 15

acc ctg atg ata caa gac aga ttg atg aac att ttt tca agt tta aag 2313
Thr Leu Met Ile Gln Asp Arg Leu Met Asn Ile Phe Ser Ser Leu Lys
20 25 30 35

gag act ata tat cac cta atg aaa agt tca ttt ata tta gac aat gga 2361
Glu Thr Ile Tyr His Leu Met Lys Ser Ser Phe Ile Leu Asp Asn Gly
40 45 50

gaa ttt tat ttt gat act cat caa gga ctg aag ttc atg caa gat tta 2409
Glu Phe Tyr Phe Asp Thr His Gln Gly Leu Lys Phe Met Gln Asp Leu
55 60 65

ttt aat gcc ctt ctg agg gaa act tca atg aaa aat aag act gaa aat 2457
Phe Asn Ala Leu Leu Arg Glu Thr Ser Met Lys Asn Lys Thr Glu Asn
70 75 80

aat ata gac ttt ttc aca gtg gtg agt cag ttg ttt ttc cat gtg aat 2505
Asn Ile Asp Phe Phe Thr Val Val Ser Gln Leu Phe Phe His Val Asn
85 90 95

aag tct gag gac ctc ttc aaa ctc aat caa gat ctt ggg tca gct ctt 2553
Lys Ser Glu Asp Leu Phe Lys Leu Asn Gln Asp Leu Gly Ser Ala Leu
100 105 110 115

cac ctt gta aga gaa tgt tca aca gag atg gca aga ctt ctg gat aca 2601
 His Leu Val Arg Glu Cys Ser Thr Glu Met Ala Arg Leu Leu Asp Thr
 120 125 130

att tta cac tct cct aat aag gac ttc tat gct ttg tat cct acc ctc Ile Leu His Ser Pro Asn Lys Asp Phe Tyr Ala Leu Tyr Pro Thr Leu 135 140 145	2649
caa gaa gtt ata ctt gct aat cta acg gat ttg ctt ttc ttt ata aat Gln Glu Val Ile Leu Ala Asn Leu Thr Asp Leu Leu Phe Phe Ile Asn 150 155 160	2697
aat tca ttc cct cta aga aac aga gca aca tta gaa att act aag aga Asn Ser Phe Pro Leu Arg Asn Arg Ala Thr Leu Glu Ile Thr Lys Arg 165 170 175	2745
tta gtt ggt gct att tca aga gca agt gaa gaa agt cac gtc ctg aaa Leu Val Gly Ala Ile Ser Arg Ala Ser Glu Glu Ser His Val Leu Lys 180 185 190 195	2793
ccc ctc tta gaa atg tct ggg act ctg gtc atg ctg ttg aat gac agt Pro Leu Leu Glu Met Ser Gly Thr Leu Val Met Leu Leu Asn Asp Ser 200 205 210	2841
gct gac ctg aga gat ctt gcc aca tca atg gac tcc att gtg aaa ctt Ala Asp Leu Arg Asp Leu Ala Thr Ser Met Asp Ser Ile Val Lys Leu 215 220 225	2889
ctt aag ctg gtc aag aaa gtt tgc ggg aag atg tcc aca gtt ttt aaa Leu Lys Leu Val Lys Lys Val Ser Gly Lys Met Ser Thr Val Phe Lys 230 235 240	2937
act cat ttt atc tcc aat acc aag gac agt gtg aaa ttc ttt gac act Thr His Phe Ile Ser Asn Thr Lys Asp Ser Val Lys Phe Phe Asp Thr 245 250 255	2985
ctg tat tcc atc atg caa caa agt gtt caa aat ctt gtg aaa gaa ata Leu Tyr Ser Ile Met Gln Gln Ser Val Gln Asn Leu Val Lys Glu Ile 260 265 270 275	3033
gct act tta aaa aaa ata gat cat ttc aca ttt gaa aag ata aat gat Ala Thr Leu Lys Lys Ile Asp His Phe Thr Phe Glu Lys Ile Asn Asp 280 285 290	3081
ttg ttg gtg cca ttt ctt gac ttg gcc ttt gaa atg att ggg gta gaa Leu Leu Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile Gly Val Glu 295 300 305	3129
cct tat ata tca tca aac tct gat att ttc agt atg tca cct agc ata Pro Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser Pro Ser Ile 310 315 320	3177
ctc tca tat atg aac caa tct aag gac ttt tct gat att ttg gaa gaa Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile Leu Glu Glu 325 330 335	3225
att gct gaa ttt tta aca tct gtg aaa atg aac ttg gaa gat atg agg Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu Asp Met Arg 340 345 350 355	3273
agt ctt gcg gta gca ttt aac aat gag act caa aca ttt tct atg gat Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe Ser Met Asp 3321	

360	365	370	
tct gtc aac tta ngg gaa gaa att ctg ggt tgc tta gtt cct ata aat			3369
Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu Val Pro Ile Asn			
375	380	385	
aac atc acc aac caa atg gac ttc tta tac cct aat cca att tcc act			3417
Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro Ile Ser Thr			
390	395	400	
cat agt ggc cct caa gat ata aaa tgg gaa ata att cat gaa gtg atc			3465
His Ser Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His Glu Val Ile			
405	410	415	
ctt ttt ttg gat aaa ata tta tca caa aac agc aca gaa ata gga tct			3513
Leu Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu Ile Gly Ser			
420	425	430	435
ttc ttg aaa atg gtg atc tgt ctc acc tta gaa gct ctt tgg aaa aac			3561
Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu Trp Lys Asn			
440	445	450	
tta aag aaa gat aat tgg aat gtt tct aat gtg ttg atg acg ttt act			3609
Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met Thr Phe Thr			
455	460	465	
cag cat cca aat aac ctt ttg aaa acc ata gaa aca gtt tta gag gcc			3657
Gln His Pro Asn Asn Leu Leu Lys Thr Ile Glu Thr Val Leu Glu Ala			
470	475	480	
tcc agt gga att aaa agt gac tat gaa ggt gat ttg aat aaa agt tta			3705
Ser Ser Gly Ile Lys Ser Asp Tyr Glu Gly Asp Leu Asn Lys Ser Leu			
485	490	495	
tat ttt gac aca cct ttg agt cag aat ata act cat cat caa ctt gaa			3753
Tyr Phe Asp Thr Pro Leu Ser Gln Asn Ile Thr His His Gln Leu Glu			
500	505	510	515
aaa gca atc cat aat gtt tta agt aga ata gct ctc tgg agg aaa gga			3801
Lys Ala Ile His Asn Val Leu Ser Arg Ile Ala Leu Trp Arg Lys Gly			
520	525	530	
ctt cgt ttt aac aac tct gaa tgg ata act tcc aca aga act ttg ttt			3849
Leu Arg Phe Asn Asn Ser Glu Trp Ile Thr Ser Thr Arg Thr Leu Phe			
535	540	545	
cag cca ctt ttt gag att ttc att aaa gca acc acc gga aag aat gtc			3897
Gln Pro Leu Phe Glu Ile Phe Ile Lys Ala Thr Thr Gly Lys Asn Val			
550	555	560	
aca tca gaa aaa gaa gag aga acc gag aaa gag atg att gac ttt cct			3945
Thr Ser Glu Lys Glu Glu Arg Thr Glu Lys Glu Met Ile Asp Phe Pro			
565	570	575	
tat agt ttc aaa cca ttt ttc tgt ttg gag aaa tac ctg gga gga tta			3993
Tyr Ser Phe Lys Pro Phe Phe Cys Leu Glu Lys Tyr Leu Gly Gly Leu			
580	585	590	595
ttt gta ttg act aaa tac tgg caa caa atc cca cta aca gat caa agt			4041

Phe	Val	Leu	Thr	Lys	Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	Asp	Gln	Ser	
				600					605					610		
gtt	gtt	gag	att	tgt	gaa	gtt	ttc	cag	cag	act	gtg	aag	ccc	tca	gaa	4089
Val	Val	Glu	Ile	Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	Lys	Pro	Ser	Glu	
			615					620					625			
gcc	atg	gag	atg	ctg	cag	aaa	gtg	aag	atg	atg	gtc	gta	cgt	gtg	ctc	4137
Ala	Met	Glu	Met	Leu	Gln	Lys	Val	Lys	Met	Met	Val	Val	Arg	Val	Leu	
			630					635					640			
acc	atc	gtt	gca	gaa	aac	cct	tcc	tgg	acc	aag	gac	att	ttg	tgt	gct	4185
Thr	Ile	Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	Lys	Asp	Ile	Leu	Cys	Ala	
			645					650					655			
act	ctg	agt	tgc	aag	caa	aat	ggg	ata	agg	cat	ctc	att	tta	tct	gct	4233
Thr	Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	Arg	His	Leu	Ile	Leu	Ser	Ala	
															675	
ata	caa	ggg	gtc	act	ttg	gcg	cag	gac	cac	ttc	cag	gaa	att	gaa	aag	4281
Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	Asp	His	Phe	Gln	Glu	Ile	Glu	Lys	
															690	
ata	tgg	tcc	tcg	ccg	aat	cag	cta	aat	tgt	gaa	agt	ctt	agc	aag	aat	4329
Ile	Trp	Ser	Ser	Pro	Asn	Gln	Leu	Asn	Cys	Glu	Ser	Leu	Ser	Lys	Asn	
ctt	tct	agc	acc	ttg	gag	agc	ttc	aag	agc	agc	ttg	gaa	aat	gcc	act	4377
Leu	Ser	Ser	Thr	Leu	Glu	Ser	Phe	Lys	Ser	Ser	Leu	Glu	Asn	Ala	Thr	
ggc	cag	gac	tgc	aca	agc	cag	ccg	agg	ctg	gag	acg	gtg	cag	cag	cac	4425
Gly	Gln	Asp	Cys	Thr	Ser	Gln	Pro	Arg	Leu	Glu	Thr	Val	Gln	Gln	His	
ttg	tac	atg	ttg	gcc	aaa	agc	ctc	gag	gaa	act	tgg	tca	tca	ggg	aat	4473
Leu	Tyr	Met	Leu	Ala	Lys	Ser	Leu	Glu	Glu	Thr	Trp	Ser	Ser	Gly	Asn	
ccg	atc	atg	act	ttt	ctc	agc	aat	ttc	aca	gta	act	gag	gat	gta	aaa	4521
Pro	Ile	Met	Thr	Phe	Leu	Ser	Asn	Phe	Thr	Val	Thr	Glu	Asp	Val	Lys	
ata	aaa	gat	ttg	atg	aag	aat	atc	acc	aag	ttg	act	gag	gag	ctt	cgc	4569
Ile	Lys	Asp	Leu	Met	Lys	Asn	Ile	Thr	Lys	Leu	Thr	Glu	Glu	Leu	Arg	
tct	tcc	atc	caa	atc	tcg	aat	gag	act	atc	cat	agc	att	cta	gaa	gca	4617
Ser	Ser	Ile	Gln	Ile	Ser	Asn	Glu	Thr	Ile	His	Ser	Ile	Leu	Glu	Ala	
aat	att	tcc	cac	tcc	aag	gtt	ctc	ttc	agt	gcc	ctc	acc	gta	gct	ctg	4665
Asn	Ile	Ser	His	Ser	Lys	Val	Leu	Phe	Ser	Ala	Leu	Thr	Val	Ala	Leu	
tct	gga	aag	tgt	gat	cag	gaa	atc	ctt	cat	ctc	ctg	ctg	aca	ttt	ccc	4713
Ser	Gly	Lys	Cys	Asp	Gln	Glu	Ile	Leu	His	Leu	Leu	Leu	Thr	Phe	Pro	

aaa ggg gaa aaa tct tgg atc gca gcg gag gaa ctc tgt agc ctg cca	4761
Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys Ser Leu Pro	
840 845 850	
ggg tca aaa gtg tat tct ctg att gtg ttg ctg agt cga aac ttg gat	4809
Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg Asn Leu Asp	
855 860 865	
gtg cga gct ttc att tac aag act ctg atg cct tct gaa gca aat ggc	4857
Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu Ala Asn Gly	
870 875 880	
ttg ctc aac tcc ttg ctg gat ata gtt tcc agc ctc agc gcc ttg ctt	4905
Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser Ala Leu Leu	
885 890 895	
gcc aaa gcc cag cac gtc ttt gag tat ctt cct gag ttt ctt cac aca	4953
Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe Leu His Thr	
900 905 910 915	
ttt aaa atc act gcc ttg cta gaa acc ctg gac ttt caa cag gtt tca	5001
Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln Gln Val Ser	
920 925 930	
caa aat gtc cag gcc aga agt tca gct ttt ggt tct ttc cag ttt gtg	5049
Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe Gln Phe Val	
935 940 945	
atg aag atg gtt tgc aag gac caa gca tca ttc ctt agc gat tct aat	5097
Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser Asp Ser Asn	
950 955 960	
atg ttt att aat ttg ccc aga gtt aag gaa ctc ttg gaa gat gac aaa	5145
Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Glu Asp Asp Lys	
965 970 975	
gaa aaa ttc aac att cct gaa gat tca aca ccg ttt tgc ttg aag ctt	5193
Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys Leu Lys Leu	
980 985 990 995	
tat cag gaa att cta caa ttg cca aat ggt gct ttg gtg tgg acc	5238
Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val Trp Thr	
1000 1005 1010	
ttc cta aaa ccc ata ttg cat gga aaa ata cta tac aca cca aac	5283
Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr Pro Asn	
1015 1020 1025	
act cca gaa att aac aag gtc att caa aag gct aat tac acc ttt	5328
Thr Pro Glu Ile Asn Lys Val Ile Gln Lys Ala Asn Tyr Thr Phe	
1030 1035 1040	
tat att gtg gac aaa cta aaa act tta tca gaa aca ctg ctg gaa	5373
Tyr Ile Val Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu Leu Glu	
1045 1050 1055	
atg tcc agc ctt ttc cag aga agt gga agt ggc cag atg ttc aac	5418
Met Ser Ser Leu Phe Gln Arg Ser Gly Ser Gly Gln Met Phe Asn	
1060 1065 1070	

cag ctg cag gag gcc	ctg aga aac aaa ttt	gta aga aac ttt gta	5463
Gln Leu Gln Glu Ala	Leu Arg Asn Lys Phe	Val Arg Asn Phe Val	
1075	1080	1085	
gaa aac cag ttg cac	att gat gta gac aaa	ctt act gaa aaa ctc	5508
Glu Asn Gln Leu His	Ile Asp Val Asp Lys	Leu Thr Glu Lys Leu	
1090	1095	1100	
cag aca tac gga ggg	ctg ctg gat gag atg	ttt aac cat gca ggc	5553
Gln Thr Tyr Gly Gly	Leu Leu Asp Glu Met	Phe Asn His Ala Gly	
1105	1110	1115	
gct gga cgc ttc cgt	ttc ttg ggc agc atc	ttg gtc aat ctc tct	5598
Ala Gly Arg Phe Arg	Phe Leu Gly Ser Ile	Leu Val Asn Leu Ser	
1120	1125	1130	
tcc tgc gtg gca ctg	aac cgt ttc cag gct	ctg cag tct gtc gac	5643
Ser Cys Val Ala Leu	Asn Arg Phe Gln Ala	Leu Gln Ser Val Asp	
1135	1140	1145	
atc ctg gag act aaa	gca cat gaa ctc ttg	cag cag aac agc ttc	5688
Ile Leu Glu Thr Lys	Ala His Glu Leu Leu	Gln Gln Asn Ser Phe	
1150	1155	1160	
ttg gcc agt atc att	ttc agc aat tcc tta	ttc gac aag aac ttc	5733
Leu Ala Ser Ile Ile	Phe Ser Asn Ser Leu	Phe Asp Lys Asn Phe	
1165	1170	1175	
aga tca gag tct gtc	aaa ctg cca ccc cat	gtc tca tac aca atc	5778
Arg Ser Glu Ser Val	Lys Leu Pro Pro His	Val Ser Tyr Thr Ile	
1180	1185	1190	
cgg acc aat gtg tta	tac agc gtg cga aca	gat gtg gta aaa aac	5823
Arg Thr Asn Val Leu	Tyr Ser Val Arg Thr	Asp Val Val Lys Asn	
1195	1200	1205	
cct tct tgg aag ttc	cac cct cag aat cta	cca gct gat ggg ttc	5868
Pro Ser Trp Lys Phe	His Pro Gln Asn Leu	Pro Ala Asp Gly Phe	
1210	1215	1220	
aaa tat aac tac gtc	ttt gcc cca ctg caa	gac atg atc gaa aga	5913
Lys Tyr Asn Tyr Val	Phe Ala Pro Leu Gln	Asp Met Ile Glu Arg	
1225	1230	1235	
gcc atc att ttg gtg	cag act ggg cag gaa	gcc ctg gaa cca gca	5958
Ala Ile Ile Leu Val	Gln Thr Gly Gln Glu	Ala Leu Glu Pro Ala	
1240	1245	1250	
gca cag act cag gcg	gcc cct tac ccc tgc	cat acc agc gac cta	6003
Ala Gln Thr Gln Ala	Ala Pro Tyr Pro Cys	His Thr Ser Asp Leu	
1255	1260	1265	
ttc ctg aac aac gtt	ggt ttc ttt ttt cca	ctg ata atg atg ctg	6048
Phe Leu Asn Asn Val	Gly Phe Phe Phe Pro	Leu Ile Met Met Leu	
1270	1275	1280	
acg tgg atg gtg tct	gtg gcc agc atg gtc	aga aag ttg gtg tat	6093
Thr Trp Met Val Ser	Val Ala Ser Met Val	Arg Lys Leu Val Tyr	

1285					1290					1295					
gag	cag	gag	ata	cag	ata	gaa	gag	tat	atg	cgg	atg	atg	gga	gtg	6138
Glu	Gln	Glu	Ile	Gln	Ile	Glu	Glu	Tyr	Met	Arg	Met	Met	Gly	Val	
				1300					1305					1310	
cat	cca	gtg	atc	cat	ttc	ctg	gcc	tgg	ttc	ctg	gag	aac	atg	gct	6183
His	Pro	Val	Ile	His	Phe	Leu	Ala	Trp	Phe	Leu	Glu	Asn	Met	Ala	
				1315					1320					1325	
gtg	ttg	acc	ata	agc	agt	gct	act	ctg	gcc	atc	gtt	ctg	aaa	aca	6228
Val	Leu	Thr	Ile	Ser	Ser	Ala	Thr	Leu	Ala	Ile	Val	Leu	Lys	Thr	
				1330					1335					1340	
agt	ggc	atc	ttt	gca	cac	agc	aat	acc	ttt	att	gtt	ttc	ctc	ttt	6273
Ser	Gly	Ile	Phe	Ala	His	Ser	Asn	Thr	Phe	Ile	Val	Phe	Leu	Phe	
				1345					1350					1355	
ctc	ttg	gat	ttt	ggg	atg	tca	gtc	gtc	atg	ctg	agc	tac	ctc	ttg	6318
Leu	Leu	Asp	Phe	Gly	Met	Ser	Val	Val	Met	Leu	Ser	Tyr	Leu	Leu	
				1360					1365					1370	
agt	gca	ttt	ttc	agc	caa	gct	aat	aca	gcg	gcc	ctt	tgt	acc	agc	6363
Ser	Ala	Phe	Phe	Ser	Gln	Ala	Asn	Thr	Ala	Ala	Leu	Cys	Thr	Ser	
				1375					1380					1385	
ctg	gtg	tac	atg	atc	agc	ttt	ctg	ccc	tac	ata	gtt	cta	ttg	gtt	6408
Leu	Val	Tyr	Met	Ile	Ser	Phe	Leu	Pro	Tyr	Ile	Val	Leu	Leu	Val	
				1390					1395					1400	
cta	cat	aac	caa	tta	agt	ttt	gtt	aat	cag	aca	ttt	ctg	tgc	ctt	6453
Leu	His	Asn	Gln	Leu	Ser	Phe	Val	Asn	Gln	Thr	Phe	Leu	Cys	Leu	
				1405					1410					1415	
ctt	tcg	aca	acc	gcc	ttt	gga	caa	ggg	gta	ttt	ttt	att	aca	ttc	6498
Leu	Ser	Thr	Thr	Ala	Phe	Gly	Gln	Gly	Val	Phe	Phe	Ile	Thr	Phe	
				1420					1425					1430	
ctg	gaa	gga	caa	gag	aca	ggg	att	caa	tgg	aat	aat	atg	tac	cag	6543
Leu	Glu	Gly	Gln	Glu	Thr	Gly	Ile	Gln	Trp	Asn	Asn	Met	Tyr	Gln	
				1435					1440					1445	
gct	ctg	gaa	caa	ggg	ggc	atg	aca	ttt	ggc	tgg	gtt	tgc	tgg	atg	6588
Ala	Leu	Glu	Gln	Gly	Gly	Met	Thr	Phe	Gly	Trp	Val	Cys	Trp	Met	
				1450					1455					1460	
att	ctt	ttt	gat	tca	agc	ctt	tat	ttt	ttg	tgt	gga	tgg	tac	ttg	6633
Ile	Leu	Phe	Asp	Ser	Ser	Leu	Tyr	Phe	Leu	Cys	Gly	Trp	Tyr	Leu	
				1465					1470					1475	
agc	aac	ttg	att	cct	gga	aca	ttt	ggt	tta	cgg	aaa	cca	tgg	tat	6678
Ser	Asn	Leu	Ile	Pro	Gly	Thr	Phe	Gly	Leu	Arg	Lys	Pro	Trp	Tyr	
				1480					1485					1490	
ttc	ccc	ttt	act	gcc	tca	tat	tgg	aag	agt	gtg	ggt	ttc	ttg	gtg	6723
Phe	Pro	Phe	Thr	Ala	Ser	Tyr	Trp	Lys	Ser	Val	Gly	Phe	Leu	Val	
				1495					1500					1505	
gag	aaa	agg	caa	tac	ttt	cta	agt	tct	agt	ctg	ttc	ttc	ttc	aat	6768

Glu Lys Arg Gln Tyr Phe Leu Ser Ser Ser Leu Phe Phe Phe Asn	
1510 1515 1520	
gag aac ttt gac aat aaa ggg tca tca ctg caa aac agg gaa gga	6813
Glu Asn Phe Asp Asn Lys Gly Ser Ser Leu Gln Asn Arg Glu Gly	
1525 1530 1535	
gag ctt gaa gga agt gcc ccg gga gtc acc ctg gtg tct gtg acc	6858
Glu Leu Glu Gly Ser Ala Pro Gly Val Thr Leu Val Ser Val Thr	
1540 1545 1550	
aag gaa tat gag ggc cac aag gct gtg gtc caa gac ctc agc ctg	6903
Lys Glu Tyr Glu Gly His Lys Ala Val Val Gln Asp Leu Ser Leu	
1555 1560 1565	
acc ttc tac aga gac caa atc acc gcc ctg ctg ggg aca aac ggt	6948
Thr Phe Tyr Arg Asp Gln Ile Thr Ala Leu Leu Gly Thr Asn Gly	
1570 1575 1580	
gcc ggg aaa acc act atc ata tcc atg ttg acg ggg ctc cac cct	6993
Ala Gly Lys Thr Thr Ile Ile Ser Met Leu Thr Gly Leu His Pro	
1585 1590 1595	
ccc act tct gga acc atc atc atc aat ggc aag aac cta cag aca	7038
Pro Thr Ser Gly Thr Ile Ile Ile Asn Gly Lys Asn Leu Gln Thr	
1600 1605 1610	
gac ctg tcg agg gtc aga atg gag ctt ggt gtg tgt ccg cag cag	7083
Asp Leu Ser Arg Val Arg Met Glu Leu Gly Val Cys Pro Gln Gln	
1615 1620 1625	
gac atc ctg ttg gac aac ctc acc gtc cgg gaa cat ttg ctg ctc	7128
Asp Ile Leu Leu Asp Asn Leu Thr Val Arg Glu His Leu Leu Leu	
1630 1635 1640	
ttt gct tcc ata aag gcg cct cag tgg acc aag aag gag ctg cat	7173
Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr Lys Lys Glu Leu His	
1645 1650 1655	
cag caa gtc aat caa act ctt cag gat gtg gac tta act cag cat	7218
Gln Gln Val Asn Gln Thr Leu Gln Asp Val Asp Leu Thr Gln His	
1660 1665 1670	
cag cac aaa cag acc cga gct ctg tct gga ggc ctg aag agg aag	7263
Gln His Lys Gln Thr Arg Ala Leu Ser Gly Gly Leu Lys Arg Lys	
1675 1680 1685	
ctc tcc ctt ggc att gct ttc atg ggc atg tcg agg acc gtg gtt	7308
Leu Ser Leu Gly Ile Ala Phe Met Gly Met Ser Arg Thr Val Val	
1690 1695 1700	
ctg gat gag ccc acc agt ggg gtg gac cct tgc tcc cgg cat agc	7353
Leu Asp Glu Pro Thr Ser Gly Val Asp Pro Cys Ser Arg His Ser	
1705 1710 1715	
ctg tgg gac att ctg ctc aag tac cga gaa ggt cgt acg atc atc	7398
Leu Trp Asp Ile Leu Leu Lys Tyr Arg Glu Gly Arg Thr Ile Ile	
1720 1725 1730	
ttc aca acc cac cac ctg gat gaa gcc gaa gcg ctg agt gac cgc	7443

Phe	Thr	Thr	His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser	Asp	Arg	
				1735					1740					1745	
gtg	gcc	gtc	ctc	cag	cat	ggg	agg	ctc	agg	tgc	tgc	ggt	cct	ccc	7488
Val	Ala	Val	Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly	Pro	Pro	
				1750					1755					1760	
ttc	tgc	ctg	aag	gag	gca	tat	ggc	cag	ggg	ctc	cgc	ctg	aca	ctc	7533
Phe	Cys	Leu	Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu	Thr	Leu	
				1765					1770					1775	
acg	agg	cag	cct	tct	gtt	ctg	gag	gcc	cat	gat	ctg	aaa	gac	atg	7578
Thr	Arg	Gln	Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys	Asp	Met	
				1780					1785					1790	
gct	tgt	gtt	aca	tcc	ctg	ata	aag	atc	tat	att	cca	caa	gca	ttt	7623
Ala	Cys	Val	Thr	Ser	Leu	Ile	Lys	Ile	Tyr	Ile	Pro	Gln	Ala	Phe	
				1795					1800					1805	
ctc	aaa	gac	agc	agt	gga	agt	gag	ctg	acc	tac	acc	att	cca	aag	7668
Leu	Lys	Asp	Ser	Ser	Gly	Ser	Glu	Leu	Thr	Tyr	Thr	Ile	Pro	Lys	
				1810					1815					1820	
gac	aca	gac	aag	gcc	tgc	ttg	aaa	ggg	ctc	ttc	cag	gcc	ctg	gat	7713
Asp	Thr	Asp	Lys	Ala	Cys	Leu	Lys	Gly	Leu	Phe	Gln	Ala	Leu	Asp	
				1825					1830					1835	
gag	aac	ctg	cat	cag	ctg	cac	ctg	acg	ggc	tat	ggg	atc	tca	gac	7758
Glu	Asn	Leu	His	Gln	Leu	His	Leu	Thr	Gly	Tyr	Gly	Ile	Ser	Asp	
				1840					1845					1850	
acc	acc	tta	gaa	gag	gtg	ttt	ttg	atg	ctt	ttg	caa	gat	tcc	aac	7803
Thr	Thr	Leu	Glu	Glu	Val	Phe	Leu	Met	Leu	Leu	Gln	Asp	Ser	Asn	
				1855					1860					1865	
aag	aaa	tct	cac	att	gcc	ctg	ggg	act	gag	tca	gag	ctg	cag	aac	7848
Lys	Lys	Ser	His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu	Gln	Asn	
				1870					1875					1880	
cac	agg	cct	aca	gga	cat	ctg	tct	ggc	tac	tgt	ggc	tcc	cta	gca	7893
His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser	Leu	Ala	
				1885					1890					1895	
cgg	ccc	gca	act	gtg	cag	ggc	gtc	cag	ctg	ctc	cgc	gca	caa	gtg	7938
Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu	Leu	Arg	Ala	Gln	Val	
				1900					1905					1910	
gcc	gcg	atc	ctg	gcc	cgg	agg	ctc	cgc	cgc	acg	ctg	cgc	gcc	ggg	7983
Ala	Ala	Ile	Leu	Ala	Arg	Arg	Leu	Arg	Arg	Thr	Leu	Arg	Ala	Gly	
				1915					1920					1925	
aag	agc	acc	ctc	gcc	gac	ctg	ctg	ctg	cca	gtc	ctc	ttc	gtg	gcc	8028
Lys	Ser	Thr	Leu	Ala	Asp	Leu	Leu	Leu	Pro	Val	Leu	Phe	Val	Ala	
				1930					1935					1940	
ttg	gcc	atg	ggc	ttg	ttc	atg	gtg	aga	ccc	ctg	gcc	acc	gag	tac	8073
Leu	Ala	Met	Gly	Leu	Phe	Met	Val	Arg	Pro	Leu	Ala	Thr	Glu	Tyr	
				1945					1950					1955	

cct ccc ctc aga ctc	aca cct gga cat tac	cag cgg gcc gag acc	8118
Pro Pro Leu Arg Leu	Thr Pro Gly His Tyr	Gln Arg Ala Glu Thr	
1960	1965	1970	
tac ttt ttc agc agt	ggg ggc gac aac ttg	gac ctc acc cgt gtg	8163
Tyr Phe Phe Ser Ser	Gly Gly Asp Asn Leu	Asp Leu Thr Arg Val	
1975	1980	1985	
ctt ctg cgg aag ttt	aga gat caa gat ttg	ccc tgt gca gat tta	8208
Leu Leu Arg Lys Phe	Arg Asp Gln Asp Leu	Pro Cys Ala Asp Leu	
1990	1995	2000	
aac cca cgc cag aag	aat tct tca tgc tgg	cgc aca gat ccc ttt	8253
Asn Pro Arg Gln Lys	Asn Ser Ser Cys Trp	Arg Thr Asp Pro Phe	
2005	2010	2015	
tct cac cca gaa ttc	cag gat tca tgt ggc	tgc ctg aag tgt cca	8298
Ser His Pro Glu Phe	Gln Asp Ser Cys Gly	Cys Leu Lys Cys Pro	
2020	2025	2030	
aat aga agt gct agt	gct ccc tac ctg acc	aac cac ctg ggc cac	8343
Asn Arg Ser Ala Ser	Ala Pro Tyr Leu Thr	Asn His Leu Gly His	
2035	2040	2045	
aca ctg ttg aat ctc	tca ggc ttc aat atg	gag gag tac ttg ctg	8388
Thr Leu Leu Asn Leu	Ser Gly Phe Asn Met	Glu Glu Tyr Leu Leu	
2050	2055	2060	
gca cca tct gaa aaa	cca agg ctt gga ggt	tgg tct ttt gga tta	8433
Ala Pro Ser Glu Lys	Pro Arg Leu Gly Gly	Trp Ser Phe Gly Leu	
2065	2070	2075	
aaa atc ccc agt gaa	gct gga ggt gca aat	gga aac ata tca aaa	8478
Lys Ile Pro Ser Glu	Ala Gly Gly Ala Asn	Gly Asn Ile Ser Lys	
2080	2085	2090	
ccc cca act ctg gca	aag gtg tgg tat aat	cag aag ggt ttt cat	8523
Pro Pro Thr Leu Ala	Lys Val Trp Tyr Asn	Gln Lys Gly Phe His	
2095	2100	2105	
tcc cta cct tcc tac	tta aat cat cta aac	aac ctt att ttg tgg	8568
Ser Leu Pro Ser Tyr	Leu Asn His Leu Asn	Asn Leu Ile Leu Trp	
2110	2115	2120	
cag cac cta ccc cct	act gtg gac tgg aga	caa tac gga ata aca	8613
Gln His Leu Pro Pro	Thr Val Asp Trp Arg	Gln Tyr Gly Ile Thr	
2125	2130	2135	
ctc tac agc cac cca	tat gga ggg gcc ttg	ctg aac gag gac aag	8658
Leu Tyr Ser His Pro	Tyr Gly Gly Ala Leu	Leu Asn Glu Asp Lys	
2140	2145	2150	
atc ctg gag agc atc	cgt cag tgt gga gtg	gcc ctc tgc atc gtg	8703
Ile Leu Glu Ser Ile	Arg Gln Cys Gly Val	Ala Leu Cys Ile Val	
2155	2160	2165	
ctg gga ttc tcc atc	ctg tct gca tcc atc	ggc agc tct gtg gtg	8748
Leu Gly Phe Ser Ile	Leu Ser Ala Ser Ile	Gly Ser Ser Val Val	
2170	2175	2180	

agg gac agg gtg att	gga gcc aaa agg ttg	cag cac ata agt ggc	8793
Arg Asp Arg Val Ile	Gly Ala Lys Arg Leu	Gln His Ile Ser Gly	
2185	2190	2195	
ctt ggc tac agg atg	tac tgg ttc aca aac	ttc cta tat gac atg	8838
Leu Gly Tyr Arg Met	Tyr Trp Phe Thr Asn	Phe Leu Tyr Asp Met	
2200	2205	2210	
ctc ttt tac ttg gtt	tcc gtc tgc ctg tgt	gtt gcc gtt att gtc	8883
Leu Phe Tyr Leu Val	Ser Val Cys Leu Cys	Val Ala Val Ile Val	
2215	2220	2225	
gcc ttc cag tta aca	gct ttt act ttc cgc	aag aac ttg gca gcc	8928
Ala Phe Gln Leu Thr	Ala Phe Thr Phe Arg	Lys Asn Leu Ala Ala	
2230	2235	2240	
acg gcc ctc ctg ctg	tca ctt ttc gga tat	gca act ctt cca tgg	8973
Thr Ala Leu Leu Leu	Ser Leu Phe Gly Tyr	Ala Thr Leu Pro Trp	
2245	2250	2255	
atg tac ctg atg tcc	aga atc ttt tcc agt	tcg gac gtg gct ttc	9018
Met Tyr Leu Met Ser	Arg Ile Phe Ser Ser	Ser Asp Val Ala Phe	
2260	2265	2270	
att tcc tat gtc tca	cta aac ttc atc ttt	ggc ctt tgt acc atg	9063
Ile Ser Tyr Val Ser	Leu Asn Phe Ile Phe	Gly Leu Cys Thr Met	
2275	2280	2285	
ccc ata acc att atg	ccc cgg ttg cta gcc	atc atc tcc aaa gct	9108
Pro Ile Thr Ile Met	Pro Arg Leu Leu Ala	Ile Ile Ser Lys Ala	
2290	2295	2300	
aag aat tta cag aat	atc tat gat gtc ctc	aag tgg gtc ttt act	9153
Lys Asn Leu Gln Asn	Ile Tyr Asp Val Leu	Lys Trp Val Phe Thr	
2305	2310	2315	
att ttt cct caa ttc	tgt ctt ggt caa gga	ctg gta gaa ctc tgc	9198
Ile Phe Pro Gln Phe	Cys Leu Gly Gln Gly	Leu Val Glu Leu Cys	
2320	2325	2330	
tat aat cag atc aaa	tat gac ctg acc cac	aac ttc ggc att gat	9243
Tyr Asn Gln Ile Lys	Tyr Asp Leu Thr His	Asn Phe Gly Ile Asp	
2335	2340	2345	
tcc tat gtg agt ccc	ttt gag atg aac ttt	ctg ggc tgg atc ttc	9288
Ser Tyr Val Ser Pro	Phe Glu Met Asn Phe	Leu Gly Trp Ile Phe	
2350	2355	2360	
gtg caa ctg gcc tcg	cag ggc aca gta ctt	ctc ctc ttg agg gtt	9333
Val Gln Leu Ala Ser	Gln Gly Thr Val Leu	Leu Leu Leu Arg Val	
2365	2370	2375	
ctg cta cac tgg gac	ctt ctg cga tgg cca	agg ggt cat tct act	9378
Leu Leu His Trp Asp	Leu Leu Arg Trp Pro	Arg Gly His Ser Thr	
2380	2385	2390	
ctc caa ggc aca gtc	aaa tct tct aag gat	aca gat gtt gaa aaa	9423
Leu Gln Gly Thr Val	Lys Ser Ser Lys Asp	Thr Asp Val Glu Lys	

2395					2400					2405					
gag	gaa	aag	aga	gtg	ttt	gaa	gga	agg	acc	aat	gga	gac	att	ctt	9468
Glu	Glu	Lys	Arg	Val	Phe	Glu	Gly	Arg	Thr	Asn	Gly	Asp	Ile	Leu	
2410					2415					2420					
gtg	tta	tac	aac	ctt	agt	aaa	cat	tat	cga	cgc	ttt	ttc	cag	aat	9513
Val	Leu	Tyr	Asn	Leu	Ser	Lys	His	Tyr	Arg	Arg	Phe	Phe	Gln	Asn	
2425					2430					2435					
att	att	gct	gtg	caa	gat	att	agt	ttg	ggc	ata	cca	aaa	gga	gag	9558
Ile	Ile	Ala	Val	Gln	Asp	Ile	Ser	Leu	Gly	Ile	Pro	Lys	Gly	Glu	
2440					2445					2450					
tgc	ttt	gga	ctt	cta	ggg	gtg	aat	gga	gct	ggg	aag	agc	acg	act	9603
Cys	Phe	Gly	Leu	Leu	Gly	Val	Asn	Gly	Ala	Gly	Lys	Ser	Thr	Thr	
2455					2460					2465					
ttc	aaa	atg	ctg	aat	ggg	gaa	gtt	tct	cta	act	tca	gga	cat	gct	9648
Phe	Lys	Met	Leu	Asn	Gly	Glu	Val	Ser	Leu	Thr	Ser	Gly	His	Ala	
2470					2475					2480					
atc	atc	agg	act	ccc	atg	gga	gac	gcc	gtg	gac	ctg	tct	tct	gct	9693
Ile	Ile	Arg	Thr	Pro	Met	Gly	Asp	Ala	Val	Asp	Leu	Ser	Ser	Ala	
2485					2490					2495					
ggc	acg	gca	ggc	gtg	ctc	att	ggc	tac	tgt	ccc	cag	cag	gat	gcc	9738
Gly	Thr	Ala	Gly	Val	Leu	Ile	Gly	Tyr	Cys	Pro	Gln	Gln	Asp	Ala	
2500					2505					2510					
ctg	gac	gag	ctt	ctg	act	ggg	tgg	gaa	cat	ctc	tat	tat	tac	tgt	9783
Leu	Asp	Glu	Leu	Leu	Thr	Gly	Trp	Glu	His	Leu	Tyr	Tyr	Tyr	Cys	
2515					2520					2525					
agc	tta	cgc	ggg	att	cca	agg	cag	tgc	atc	cct	gag	gtt	gct	gga	9828
Ser	Leu	Arg	Gly	Ile	Pro	Arg	Gln	Cys	Ile	Pro	Glu	Val	Ala	Gly	
2530					2535					2540					
gac	ctc	atc	agg	cgc	tta	cac	ctc	gaa	gcc	cac	gcg	gac	aaa	cct	9873
Asp	Leu	Ile	Arg	Arg	Leu	His	Leu	Glu	Ala	His	Ala	Asp	Lys	Pro	
2545					2550					2555					
gtg	gcc	acc	tac	agt	ggg	gga	acc	aag	cgg	aaa	ctc	tct	aca	gcc	9918
Val	Ala	Thr	Tyr	Ser	Gly	Gly	Thr	Lys	Arg	Lys	Leu	Ser	Thr	Ala	
2560					2565					2570					
ctg	gcc	ctg	gtg	ggg	aaa	cct	gac	att	ctt	tta	ttg	gat	gag	ccc	9963
Leu	Ala	Leu	Val	Gly	Lys	Pro	Asp	Ile	Leu	Leu	Leu	Asp	Glu	Pro	
2575					2580					2585					
agc	tct	ggg	atg	gat	ccc	tgc	tct	aag	cgg	tac	ctg	tgg	caa	aca	10008
Ser	Ser	Gly	Met	Asp	Pro	Cys	Ser	Lys	Arg	Tyr	Leu	Trp	Gln	Thr	
2590					2595					2600					
ata	atg	aag	gag	gtt	cgg	gaa	ggc	tgt	gct	gcg	gtg	ctg	acc	tcc	10053
Ile	Met	Lys	Glu	Val	Arg	Glu	Gly	Cys	Ala	Ala	Val	Leu	Thr	Ser	
2605					2610					2615					
cac	agc	atg	gag	gag	tgt	gag	gct	ctt	tgc	aca	aga	ctg	gcc	ata	10098

His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile	
2620 2625 2630	
atg gtt aac ggc agc ttc aaa tgt ctt ggt tct cct cag cac atc	10143
Met Val Asn Gly Ser Phe Lys Cys Leu Gly Ser Pro Gln His Ile	
2635 2640 2645	
aaa aat agg ttt ggt gat ggt tat aca gtc aaa gtt tgg ctc tgt	10188
Lys Asn Arg Phe Gly Asp Gly Tyr Thr Val Lys Val Trp Leu Cys	
2650 2655 2660	
aag gaa gca aat caa cat tgc act gtt tct gac cac ttg aag ctt	10233
Lys Glu Ala Asn Gln His Cys Thr Val Ser Asp His Leu Lys Leu	
2665 2670 2675	
tat ttt cca gga att cag ttc aag gga cag cac ctg aat tta tta	10278
Tyr Phe Pro Gly Ile Gln Phe Lys Gly Gln His Leu Asn Leu Leu	
2680 2685 2690	
gaa tat cat gtg cca aaa aga tgg gga tgc cta gct gac ttg ttc	10323
Glu Tyr His Val Pro Lys Arg Trp Gly Cys Leu Ala Asp Leu Phe	
2695 2700 2705	
aaa gtt ata gag aac aat aaa acc ttc ttg aat att aag cat tat	10368
Lys Val Ile Glu Asn Asn Lys Thr Phe Leu Asn Ile Lys His Tyr	
2710 2715 2720	
tcc att aac caa acc act ttg gag cag gta ttt att aat ttt gct	10413
Ser Ile Asn Gln Thr Thr Leu Glu Gln Val Phe Ile Asn Phe Ala	
2725 2730 2735	
tct gag cag cag caa act cta caa tct act ctt gat cca tcc act	10458
Ser Glu Gln Gln Gln Thr Leu Gln Ser Thr Leu Asp Pro Ser Thr	
2740 2745 2750	
gac agt cac cac aca cat cac ttg ccc atc tga gcactaaaga	10501
Asp Ser His His Thr His His Leu Pro Ile	
2755 2760	
agttttccata aggaataaaaa ccttgtcttc cattacaatt aacagtcaag gataaaacaa	10561
gcacgcgcac aatcaaggag ctggaacaca ctctccaggc cgtcaaatta ttctcttggt	10621
catttttctat tttgaatctc cttgttagtt aataaccacc aaatggaaag gtcattcttt	10681
ctgcagactt ttggggagct cctccaaaac atttgttctc tttaccatgc cagatggaca	10741
ccagctttctt tgtgacaaag gcatgaatga tttgacagtg tccaaactga gacattctgg	10801
agctggaaag cctgtcacac tagagtgtgt gtgacatgtc cactctaaac atgtcacttt	10861
tctgttaaga aaactgagcc cctccccac aggttaaaaa acttttagtaa cttgttttga	10921
tagaaaatag taacaaggac tattttctat tgttgtcatc tatttactag atacatgttt	10981
ttaatgattt taatgtaagc ttttattaat actgatgaca ttatatggta tgatatgaaa	11041
aaatcaccaa ttttttacat ataaaagata cttttttaaa aaaatagggtt ttaagagctc	11101
tttttagtata cacttttagca aaattaatta aattgaacta gttactctgt atcaattaca	11161


```

gtagttctac cagaatctcc caggttataa tttatgaggg tagagaaata aaatgtagat 11221
gcattttctt tttcttcatt tggatgaata attactgttt tttgttattc taagtcaagt 11281
tttttcaaag cgtagtggtc cccatattag ctgcattgcc atcttctggg agcttgcaag 11341
aaatgtacat tctcaggatc cactocagac ctattgaatc tcaaattctg gggcttaaac 11401
aagcacattc caagttaaga accaatgacc taagggaatg tctggttacc tcctagttat 11461
acaagcaaaa tctgcatagt atgtagtctt ttttatttat ttattttttt tttttgaggc 11521
ggagtctcgc tctgtcaccc aggctggagt gcagtggcgc gatgtcggct cactgcaagc 11581
tccgcctcct gggttcacgc cattctcctg cctcagcctc cccagcagct gggactacag 11641
gcacacatcg ccacaccgg ctaattttta gtatttttag tagagacggg gtttcaccgt 11701
gttagccagg atggtctcta tctcctgacc ttgtgatccg cccccctcca cctcccaaag 11761
tgctgggatt acaggcgtga gccaccgtgt ccggccgtag tttattttta aatatatttt 11821
taaaagcttt gtaaaaatta tgtcattctc agaattgttg tcttcaaagc attgtcagat 11881
gtagagtgc cagatgtggc tcttaaagac tatatacatc tgaatttttc atcctatagt 11941
tagtaagatg cataaaatca atccactact gaaatagttt ccagtcagac atttctgagt 12001
tcagacattt ctcaacattc ttttaacaac atttttctga atcctcaata gaaaatcaca 12061
ttaatcttat tttaaaattt ggcctttttc aacactaacg ttgagtaccg gtagctttgt 12121
gatcaaaggc atatacttcc ttatgagatt tctttactaa agcaagattt cattaaatct 12181
ctatttctta aatatcattc tatacaaaag atatttttta aacggtaagg attaagacaa 12241
tcaactgatag ctttgttgtg agcaattttg attcccatgt atcacatgaa ttacacttcc 12301
ttaaataaat tacagtttat ggctgtatga tttattctct aattctaaca tagtctagtt 12361
gtcaaaagga aatatgtaat ctttttatga ttgttgaatc aataaatacc aatttgtgaa 12421

acatgaatgt gtttaaactg cagtgaataa atgagatgtg ctttaattta acccaaaaaa 12481
aaaaaaaaaa aaaaaaaa 12498

```

<210> 19
 <211> 2760
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> misc_feature
 <222> (376)..(376)
 <223> The 'Xaa' at location 376 stands for Arg, Gly, or Trp.
 <400> 19

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

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

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

Gln Gln His Leu Tyr Met Leu Ala Lys Ser Leu Glu Glu Thr Trp Ser
 740 745 750
 Ser Gly Asn Pro Ile Met Thr Phe Leu Ser Asn Phe Thr Val Thr Glu
 755 760 765
 Asp Val Lys Ile Lys Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu
 770 775 780
 Glu Leu Arg Ser Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile
 785 790 795 800
 Leu Glu Ala Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr
 805 810 815
 Val Ala Leu Ser Gly Lys Cys Asp Gln Glu Ile Leu His Leu Leu Leu
 820 825 830
 Thr Phe Pro Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys
 835 840 845
 Ser Leu Pro Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg
 850 855 860
 Asn Leu Asp Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu
 865 870 875 880
 Ala Asn Gly Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser
 885 890 895
 Ala Leu Leu Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe
 900 905 910
 Leu His Thr Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln
 915 920 925
 Gln Val Ser Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe
 930 935 940
 Gln Phe Val Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser
 945 950 955 960
 Asp Ser Asn Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu Glu
 965 970 975

Asp Asp Lys Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys
 980 985 990

Leu Lys Leu Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val
 995 1000 1005

Trp Thr Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr
 1010 1015 1020

Pro Asn Thr Pro Glu Ile Asn Lys Val Ile Gln Lys Ala Asn Tyr
 1025 1030 1035

Thr Phe Tyr Ile Val Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu
 1040 1045 1050

Leu Glu Met Ser Ser Leu Phe Gln Arg Ser Gly Ser Gly Gln Met
 1055 1060 1065

Phe Asn Gln Leu Gln Glu Ala Leu Arg Asn Lys Phe Val Arg Asn
 1070 1075 1080

Phe Val Glu Asn Gln Leu His Ile Asp Val Asp Lys Leu Thr Glu
 1085 1090 1095

Lys Leu Gln Thr Tyr Gly Gly Leu Leu Asp Glu Met Phe Asn His
 1100 1105 1110

Ala Gly Ala Gly Arg Phe Arg Phe Leu Gly Ser Ile Leu Val Asn
 1115 1120 1125

Leu Ser Ser Cys Val Ala Leu Asn Arg Phe Gln Ala Leu Gln Ser
 1130 1135 1140

Val Asp Ile Leu Glu Thr Lys Ala His Glu Leu Leu Gln Gln Asn
 1145 1150 1155

Ser Phe Leu Ala Ser Ile Ile Phe Ser Asn Ser Leu Phe Asp Lys
 1160 1165 1170

Asn Phe Arg Ser Glu Ser Val Lys Leu Pro Pro His Val Ser Tyr
 1175 1180 1185

Thr Ile Arg Thr Asn Val Leu Tyr Ser Val Arg Thr Asp Val Val
 1190 1195 1200

Lys	Asn	Pro	Ser	Trp	Lys	Phe	His	Pro	Gln	Asn	Leu	Pro	Ala	Asp
1205						1210					1215			
Gly	Phe	Lys	Tyr	Asn	Tyr	Val	Phe	Ala	Pro	Leu	Gln	Asp	Met	Ile
1220						1225					1230			
Glu	Arg	Ala	Ile	Ile	Leu	Val	Gln	Thr	Gly	Gln	Glu	Ala	Leu	Glu
1235						1240					1245			
Pro	Ala	Ala	Gln	Thr	Gln	Ala	Ala	Pro	Tyr	Pro	Cys	His	Thr	Ser
1250						1255					1260			
Asp	Leu	Phe	Leu	Asn	Asn	Val	Gly	Phe	Phe	Phe	Pro	Leu	Ile	Met
1265						1270					1275			
Met	Leu	Thr	Trp	Met	Val	Ser	Val	Ala	Ser	Met	Val	Arg	Lys	Leu
1280						1285					1290			
Val	Tyr	Glu	Gln	Glu	Ile	Gln	Ile	Glu	Glu	Tyr	Met	Arg	Met	Met
1295						1300					1305			
Gly	Val	His	Pro	Val	Ile	His	Phe	Leu	Ala	Trp	Phe	Leu	Glu	Asn
1310						1315					1320			
Met	Ala	Val	Leu	Thr	Ile	Ser	Ser	Ala	Thr	Leu	Ala	Ile	Val	Leu
1325						1330					1335			
Lys	Thr	Ser	Gly	Ile	Phe	Ala	His	Ser	Asn	Thr	Phe	Ile	Val	Phe
1340						1345					1350			
Leu	Phe	Leu	Leu	Asp	Phe	Gly	Met	Ser	Val	Val	Met	Leu	Ser	Tyr
1355						1360					1365			
Leu	Leu	Ser	Ala	Phe	Phe	Ser	Gln	Ala	Asn	Thr	Ala	Ala	Leu	Cys
1370						1375					1380			
Thr	Ser	Leu	Val	Tyr	Met	Ile	Ser	Phe	Leu	Pro	Tyr	Ile	Val	Leu
1385						1390					1395			
Leu	Val	Leu	His	Asn	Gln	Leu	Ser	Phe	Val	Asn	Gln	Thr	Phe	Leu
1400						1405					1410			
Cys	Leu	Leu	Ser	Thr	Thr	Ala	Phe	Gly	Gln	Gly	Val	Phe	Phe	Ile
1415						1420					1425			

Thr	Phe	Leu	Glu	Gly	Gln	Glu	Thr	Gly	Ile	Gln	Trp	Asn	Asn	Met
1430						1435					1440			
Tyr	Gln	Ala	Leu	Glu	Gln	Gly	Gly	Met	Thr	Phe	Gly	Trp	Val	Cys
1445						1450					1455			
Trp	Met	Ile	Leu	Phe	Asp	Ser	Ser	Leu	Tyr	Phe	Leu	Cys	Gly	Trp
1460						1465					1470			
Tyr	Leu	Ser	Asn	Leu	Ile	Pro	Gly	Thr	Phe	Gly	Leu	Arg	Lys	Pro
1475						1480					1485			
Trp	Tyr	Phe	Pro	Phe	Thr	Ala	Ser	Tyr	Trp	Lys	Ser	Val	Gly	Phe
1490						1495					1500			
Leu	Val	Glu	Lys	Arg	Gln	Tyr	Phe	Leu	Ser	Ser	Ser	Leu	Phe	Phe
1505						1510					1515			
Phe	Asn	Glu	Asn	Phe	Asp	Asn	Lys	Gly	Ser	Ser	Leu	Gln	Asn	Arg
1520						1525					1530			
Glu	Gly	Glu	Leu	Glu	Gly	Ser	Ala	Pro	Gly	Val	Thr	Leu	Val	Ser
1535						1540					1545			
Val	Thr	Lys	Glu	Tyr	Glu	Gly	His	Lys	Ala	Val	Val	Gln	Asp	Leu
1550						1555					1560			
Ser	Leu	Thr	Phe	Tyr	Arg	Asp	Gln	Ile	Thr	Ala	Leu	Leu	Gly	Thr
1565						1570					1575			
Asn	Gly	Ala	Gly	Lys	Thr	Thr	Ile	Ile	Ser	Met	Leu	Thr	Gly	Leu
1580						1585					1590			
His	Pro	Pro	Thr	Ser	Gly	Thr	Ile	Ile	Ile	Asn	Gly	Lys	Asn	Leu
1595						1600					1605			
Gln	Thr	Asp	Leu	Ser	Arg	Val	Arg	Met	Glu	Leu	Gly	Val	Cys	Pro
1610						1615					1620			
Gln	Gln	Asp	Ile	Leu	Leu	Asp	Asn	Leu	Thr	Val	Arg	Glu	His	Leu
1625						1630					1635			
Leu	Leu	Phe	Ala	Ser	Ile	Lys	Ala	Pro	Gln	Trp	Thr	Lys	Lys	Glu
1640						1645					1650			

Leu	His	Gln	Gln	Val	Asn	Gln	Thr	Leu	Gln	Asp	Val	Asp	Leu	Thr
1655						1660					1665			
Gln	His	Gln	His	Lys	Gln	Thr	Arg	Ala	Leu	Ser	Gly	Gly	Leu	Lys
1670						1675					1680			
Arg	Lys	Leu	Ser	Leu	Gly	Ile	Ala	Phe	Met	Gly	Met	Ser	Arg	Thr
1685						1690					1695			
Val	Val	Leu	Asp	Glu	Pro	Thr	Ser	Gly	Val	Asp	Pro	Cys	Ser	Arg
1700						1705					1710			
His	Ser	Leu	Trp	Asp	Ile	Leu	Leu	Lys	Tyr	Arg	Glu	Gly	Arg	Thr
1715						1720					1725			
Ile	Ile	Phe	Thr	Thr	His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser
1730						1735					1740			
Asp	Arg	Val	Ala	Val	Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly
1745						1750					1755			
Pro	Pro	Phe	Cys	Leu	Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu
1760						1765					1770			
Thr	Leu	Thr	Arg	Gln	Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys
1775						1780					1785			
Asp	Met	Ala	Cys	Val	Thr	Ser	Leu	Ile	Lys	Ile	Tyr	Ile	Pro	Gln
1790						1795					1800			
Ala	Phe	Leu	Lys	Asp	Ser	Ser	Gly	Ser	Glu	Leu	Thr	Tyr	Thr	Ile
1805						1810					1815			
Pro	Lys	Asp	Thr	Asp	Lys	Ala	Cys	Leu	Lys	Gly	Leu	Phe	Gln	Ala
1820						1825					1830			
Leu	Asp	Glu	Asn	Leu	His	Gln	Leu	His	Leu	Thr	Gly	Tyr	Gly	Ile
1835						1840					1845			
Ser	Asp	Thr	Thr	Leu	Glu	Glu	Val	Phe	Leu	Met	Leu	Leu	Gln	Asp
1850						1855					1860			
Ser	Asn	Lys	Lys	Ser	His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu
1865						1870					1875			
Gln	Asn	His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser

1880		1885		1890
Leu Ala Arg Pro Ala Thr Val Gln Gly Val Gln Leu Leu Arg Ala				
1895		1900		1905
Gln Val Ala Ala Ile Leu Ala Arg Arg Leu Arg Arg Thr Leu Arg				
1910		1915		1920
Ala Gly Lys Ser Thr Leu Ala Asp Leu Leu Leu Pro Val Leu Phe				
1925		1930		1935
Val Ala Leu Ala Met Gly Leu Phe Met Val Arg Pro Leu Ala Thr				
1940		1945		1950
Glu Tyr Pro Pro Leu Arg Leu Thr Pro Gly His Tyr Gln Arg Ala				
1955		1960		1965
Glu Thr Tyr Phe Phe Ser Ser Gly Gly Asp Asn Leu Asp Leu Thr				
1970		1975		1980
Arg Val Leu Leu Arg Lys Phe Arg Asp Gln Asp Leu Pro Cys Ala				
1985		1990		1995
Asp Leu Asn Pro Arg Gln Lys Asn Ser Ser Cys Trp Arg Thr Asp				
2000		2005		2010
Pro Phe Ser His Pro Glu Phe Gln Asp Ser Cys Gly Cys Leu Lys				
2015		2020		2025
Cys Pro Asn Arg Ser Ala Ser Ala Pro Tyr Leu Thr Asn His Leu				
2030		2035		2040
Gly His Thr Leu Leu Asn Leu Ser Gly Phe Asn Met Glu Glu Tyr				
2045		2050		2055
Leu Leu Ala Pro Ser Glu Lys Pro Arg Leu Gly Gly Trp Ser Phe				
2060		2065		2070
Gly Leu Lys Ile Pro Ser Glu Ala Gly Gly Ala Asn Gly Asn Ile				
2075		2080		2085
Ser Lys Pro Pro Thr Leu Ala Lys Val Trp Tyr Asn Gln Lys Gly				
2090		2095		2100
Phe His Ser Leu Pro Ser Tyr Leu Asn His Leu Asn Asn Leu Ile				
2105		2110		2115

Leu	Trp	Gln	His	Leu	Pro	Pro	Thr	Val	Asp	Trp	Arg	Gln	Tyr	Gly
2120						2125					2130			
Ile	Thr	Leu	Tyr	Ser	His	Pro	Tyr	Gly	Gly	Ala	Leu	Leu	Asn	Glu
2135						2140					2145			
Asp	Lys	Ile	Leu	Glu	Ser	Ile	Arg	Gln	Cys	Gly	Val	Ala	Leu	Cys
2150						2155					2160			
Ile	Val	Leu	Gly	Phe	Ser	Ile	Leu	Ser	Ala	Ser	Ile	Gly	Ser	Ser
2165						2170					2175			
Val	Val	Arg	Asp	Arg	Val	Ile	Gly	Ala	Lys	Arg	Leu	Gln	His	Ile
2180						2185					2190			
Ser	Gly	Leu	Gly	Tyr	Arg	Met	Tyr	Trp	Phe	Thr	Asn	Phe	Leu	Tyr
2195						2200					2205			
Asp	Met	Leu	Phe	Tyr	Leu	Val	Ser	Val	Cys	Leu	Cys	Val	Ala	Val
2210						2215					2220			
Ile	Val	Ala	Phe	Gln	Leu	Thr	Ala	Phe	Thr	Phe	Arg	Lys	Asn	Leu
2225						2230					2235			
Ala	Ala	Thr	Ala	Leu	Leu	Leu	Ser	Leu	Phe	Gly	Tyr	Ala	Thr	Leu
2240						2245					2250			
Pro	Trp	Met	Tyr	Leu	Met	Ser	Arg	Ile	Phe	Ser	Ser	Ser	Asp	Val
2255						2260					2265			
Ala	Phe	Ile	Ser	Tyr	Val	Ser	Leu	Asn	Phe	Ile	Phe	Gly	Leu	Cys
2270						2275					2280			
Thr	Met	Pro	Ile	Thr	Ile	Met	Pro	Arg	Leu	Leu	Ala	Ile	Ile	Ser
2285						2290					2295			
Lys	Ala	Lys	Asn	Leu	Gln	Asn	Ile	Tyr	Asp	Val	Leu	Lys	Trp	Val
2300						2305					2310			
Phe	Thr	Ile	Phe	Pro	Gln	Phe	Cys	Leu	Gly	Gln	Gly	Leu	Val	Glu
2315						2320					2325			
Leu	Cys	Tyr	Asn	Gln	Ile	Lys	Tyr	Asp	Leu	Thr	His	Asn	Phe	Gly
2330						2335					2340			

Ile	Asp	Ser	Tyr	Val	Ser	Pro	Phe	Glu	Met	Asn	Phe	Leu	Gly	Trp
2345						2350					2355			
Ile	Phe	Val	Gln	Leu	Ala	Ser	Gln	Gly	Thr	Val	Leu	Leu	Leu	Leu
2360						2365					2370			
Arg	Val	Leu	Leu	His	Trp	Asp	Leu	Leu	Arg	Trp	Pro	Arg	Gly	His
2375						2380					2385			
Ser	Thr	Leu	Gln	Gly	Thr	Val	Lys	Ser	Ser	Lys	Asp	Thr	Asp	Val
2390						2395					2400			
Glu	Lys	Glu	Glu	Lys	Arg	Val	Phe	Glu	Gly	Arg	Thr	Asn	Gly	Asp
2405						2410					2415			
Ile	Leu	Val	Leu	Tyr	Asn	Leu	Ser	Lys	His	Tyr	Arg	Arg	Phe	Phe
2420						2425					2430			
Gln	Asn	Ile	Ile	Ala	Val	Gln	Asp	Ile	Ser	Leu	Gly	Ile	Pro	Lys
2435						2440					2445			
Gly	Glu	Cys	Phe	Gly	Leu	Leu	Gly	Val	Asn	Gly	Ala	Gly	Lys	Ser
2450						2455					2460			
Thr	Thr	Phe	Lys	Met	Leu	Asn	Gly	Glu	Val	Ser	Leu	Thr	Ser	Gly
2465						2470					2475			
His	Ala	Ile	Ile	Arg	Thr	Pro	Met	Gly	Asp	Ala	Val	Asp	Leu	Ser
2480						2485					2490			
Ser	Ala	Gly	Thr	Ala	Gly	Val	Leu	Ile	Gly	Tyr	Cys	Pro	Gln	Gln
2495						2500					2505			
Asp	Ala	Leu	Asp	Glu	Leu	Leu	Thr	Gly	Trp	Glu	His	Leu	Tyr	Tyr
2510						2515					2520			
Tyr	Cys	Ser	Leu	Arg	Gly	Ile	Pro	Arg	Gln	Cys	Ile	Pro	Glu	Val
2525						2530					2535			
Ala	Gly	Asp	Leu	Ile	Arg	Arg	Leu	His	Leu	Glu	Ala	His	Ala	Asp
2540						2545					2550			
Lys	Pro	Val	Ala	Thr	Tyr	Ser	Gly	Gly	Thr	Lys	Arg	Lys	Leu	Ser
2555						2560					2565			

Thr Ala Leu Ala Leu Val Gly Lys Pro Asp Ile Leu Leu Leu Asp
 2570 2575 2580
 Glu Pro Ser Ser Gly Met Asp Pro Cys Ser Lys Arg Tyr Leu Trp
 2585 2590 2595
 Gln Thr Ile Met Lys Glu Val Arg Glu Gly Cys Ala Ala Val Leu
 2600 2605 2610
 Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu
 2615 2620 2625
 Ala Ile Met Val Asn Gly Ser Phe Lys Cys Leu Gly Ser Pro Gln
 2630 2635 2640
 His Ile Lys Asn Arg Phe Gly Asp Gly Tyr Thr Val Lys Val Trp
 2645 2650 2655
 Leu Cys Lys Glu Ala Asn Gln His Cys Thr Val Ser Asp His Leu
 2660 2665 2670
 Lys Leu Tyr Phe Pro Gly Ile Gln Phe Lys Gly Gln His Leu Asn
 2675 2680 2685
 Leu Leu Glu Tyr His Val Pro Lys Arg Trp Gly Cys Leu Ala Asp
 2690 2695 2700
 Leu Phe Lys Val Ile Glu Asn Asn Lys Thr Phe Leu Asn Ile Lys
 2705 2710 2715
 His Tyr Ser Ile Asn Gln Thr Thr Leu Glu Gln Val Phe Ile Asn
 2720 2725 2730
 Phe Ala Ser Glu Gln Gln Gln Thr Leu Gln Ser Thr Leu Asp Pro
 2735 2740 2745
 Ser Thr Asp Ser His His Thr His His Leu Pro Ile
 2750 2755 2760

<210> 20
 <211> 12498
 <212> DNA
 <213> Homo Sapiens

<220>

```

<221> CDS
<222> (26)..(2179)

<220>
<221> misc_feature
<222> (873)..(873)
<223> n = any nucleic acid residue

<220>
<221> misc_feature
<222> (3334)..(3334)
<223> Xaa = any amino acid

<400> 20
ggactgagag cagggagcag caggc atg ggg cat gcc ggg tgc cag ttc aaa      52
                               Met Gly His Ala Gly Cys Gln Phe Lys
                               1                               5

gcc ctg ctg tgg aag aat tgg ctc tgc aga ctc agg aac ccg gtc ctt      100
Ala Leu Leu Trp Lys Asn Trp Leu Cys Arg Leu Arg Asn Pro Val Leu
10                               15                               20                               25

ttc ctt gct gaa ttc ttc tgg cct tgt atc ctg ttt gta att ctg aca      148
Phe Leu Ala Glu Phe Phe Trp Pro Cys Ile Leu Phe Val Ile Leu Thr
                               30                               35                               40

gtt ctt cgt ttt caa gaa cct ccc aga tac aga gac att tgt tat ttg      196
Val Leu Arg Phe Gln Glu Pro Pro Arg Tyr Arg Asp Ile Cys Tyr Leu
                               45                               50                               55

cag ccc cga gat cta ccc agc tgt ggt gtt atc ccc ttt gtt caa agc      244
Gln Pro Arg Asp Leu Pro Ser Cys Gly Val Ile Pro Phe Val Gln Ser
                               60                               65                               70

ctt ctt tgt aac act gga tca agg tgt agg aac ttc agc tat gaa ggg      292
Leu Leu Cys Asn Thr Gly Ser Arg Cys Arg Asn Phe Ser Tyr Glu Gly
75                               80                               85

tca atg gag cat cat ttt cgt ttg tct agg ttc caa act gca gct gac      340
Ser Met Glu His His Phe Arg Leu Ser Arg Phe Gln Thr Ala Ala Asp
90                               95                               100                               105

ccc aag aaa gtc aac aac ctg gcc ttt tta aaa gag ata caa gac ctg      388
Pro Lys Lys Val Asn Asn Leu Ala Phe Leu Lys Glu Ile Gln Asp Leu
                               110                               115                               120

gca gag gaa att cat gga atg atg gac aag gca aaa aac tta aaa aga      436
Ala Glu Glu Ile His Gly Met Met Asp Lys Ala Lys Asn Leu Lys Arg
125                               130                               135

ctt tgg gta gaa cga tcc aac act cca gat tct tct tat ggt tcc agt      484
Leu Trp Val Glu Arg Ser Asn Thr Pro Asp Ser Ser Tyr Gly Ser Ser
140                               145                               150

ttt ttt aca atg gat ctc aat aag acc gag gag gta ata ttg aaa ctg      532
Phe Phe Thr Met Asp Leu Asn Lys Thr Glu Glu Val Ile Leu Lys Leu
155                               160                               165

gaa agc ctc cat cag cag cct cat atc tgg gat ttt cta ctt tta ctg      580
Glu Ser Leu His Gln Gln Pro His Ile Trp Asp Phe Leu Leu Leu Leu
170                               175                               180                               185

```

ccg aga cta cac aca agc cat gat cat gtg gaa gat ggc atg gat gtt	628
Pro Arg Leu His Thr Ser His Asp His Val Glu Asp Gly Met Asp Val	
190 195 200	
gca gtg aac ctt ctc cag acc att ttg aat tcc tta ata tcc cta gaa	676
Ala Val Asn Leu Leu Gln Thr Ile Leu Asn Ser Leu Ile Ser Leu Glu	
205 210 215	
gat tta gat tgg ctt cca ctc aac caa act ttt tcc cag gtt tct gaa	724
Asp Leu Asp Trp Leu Pro Leu Asn Gln Thr Phe Ser Gln Val Ser Glu	
220 225 230	
ctt gta ctg aat gtg acc att tcg aca ctg aca ttt ctg cag caa cat	772
Leu Val Leu Asn Val Thr Ile Ser Thr Leu Thr Phe Leu Gln Gln His	
235 240 245	
gga gta gca gtc acc gag cca gtt tac cac ctg tcc atg cag aat ata	820
Gly Val Ala Val Thr Glu Pro Val Tyr His Leu Ser Met Gln Asn Ile	
250 255 260 265	
gtg tgg gat cca cag aaa gtc cag tat gat ctc aaa tcc cag ttt ggc	868
Val Trp Asp Pro Gln Lys Val Gln Tyr Asp Leu Lys Ser Gln Phe Gly	
270 275 280	
ttt gnt gat ctt cac acg gaa cag atc ctg aac tct tca gct gaa ctg	916
Phe Xaa Asp Leu His Thr Glu Gln Ile Leu Asn Ser Ser Ala Glu Leu	
285 290 295	
aag gag att ccc aca gac act tcc ttg gag aag atg gtg tgt tca gtc	964
Lys Glu Ile Pro Thr Asp Thr Ser Leu Glu Lys Met Val Cys Ser Val	
300 305 310	
ttg tct agc aca tca gag gat gaa gct gag aaa tgg ggc cac gtt gga	1012
Leu Ser Ser Thr Ser Glu Asp Glu Ala Glu Lys Trp Gly His Val Gly	
315 320 325	
ggc tgc cac cct aag tgg tca gaa gcc aaa aac tat ctt gtc cat gca	1060
Gly Cys His Pro Lys Trp Ser Glu Ala Lys Asn Tyr Leu Val His Ala	
330 335 340 345	
gtc agc tgg ctg cga gtc tac caa cag gtg ttt gtt cag tgg caa cag	1108
Val Ser Trp Leu Arg Val Tyr Gln Gln Val Phe Val Gln Trp Gln Gln	
350 355 360	
ggt agc ctg ctt cag aag aca ctc aca ggc atg ggc cat agt ctg gag	1156
Gly Ser Leu Leu Gln Lys Thr Leu Thr Gly Met Gly His Ser Leu Glu	
365 370 375	
gct ctc agg aat cag ttt gaa gaa gag agc aag ccc tgg aag gtg gtg	1204
Ala Leu Arg Asn Gln Phe Glu Glu Glu Ser Lys Pro Trp Lys Val Val	
380 385 390	
gaa gct ctg cac act gca ctg ctc ctg ctg aat gac agc ttg tca gca	1252
Glu Ala Leu His Thr Ala Leu Leu Leu Leu Asn Asp Ser Leu Ser Ala	
395 400 405	
gat ggc cca aaa gat aat cat aca ttt cca aag ata tta cag cat ctg	1300
Asp Gly Pro Lys Asp Asn His Thr Phe Pro Lys Ile Leu Gln His Leu	
410 415 420 425	

tgg aaa ttg caa agc ttg ctg caa aac ctg ccc cag tgg ccg gca ctg	1348
Trp Lys Leu Gln Ser Leu Leu Gln Asn Leu Pro Gln Trp Pro Ala Leu	
430 435 440	
aag aga ttt ctt cag ctt gat gga gct ctc aga aat gcg ata gct cag	1396
Lys Arg Phe Leu Gln Leu Asp Gly Ala Leu Arg Asn Ala Ile Ala Gln	
445 450 455	
aat tta cat ttt gtc caa gaa gtc ctc att tgc ctg gag aca tca gct	1444
Asn Leu His Phe Val Gln Glu Val Leu Ile Cys Leu Glu Thr Ser Ala	
460 465 470	
aat gat ttt aaa tgg ttt gaa ctt aac caa ttg aaa ctg gaa aag gat	1492
Asn Asp Phe Lys Trp Phe Glu Leu Asn Gln Leu Lys Leu Glu Lys Asp	
475 480 485	
gtg ttc ttt tgg gag ctg aaa cag atg ttg gcg aag aat gct gtc tgc	1540
Val Phe Phe Trp Glu Leu Lys Gln Met Leu Ala Lys Asn Ala Val Cys	
490 495 500 505	
ccg aat ggt cgt ttc tct gag aag gag gtc ttt ttg ccg cct gga aac	1588
Pro Asn Gly Arg Phe Ser Glu Lys Glu Val Phe Leu Pro Pro Gly Asn	
510 515 520	
tcc agc ata tgg ggt ggt ctc cag gga ctg ttg tgc tat tgt aac tcc	1636
Ser Ser Ile Trp Gly Gly Leu Gln Gly Leu Leu Cys Tyr Cys Asn Ser	
525 530 535	
tct gag acg agt gtt tta aac aag cta ctt ggt tca gta gag gat gct	1684
Ser Glu Thr Ser Val Leu Asn Lys Leu Leu Gly Ser Val Glu Asp Ala	
540 545 550	
gat cgt att ttg caa gag gtc att act tgg cac aaa aat atg tca gtt	1732
Asp Arg Ile Leu Gln Glu Val Ile Thr Trp His Lys Asn Met Ser Val	
555 560 565	
tta ata cct gaa gaa tat ttg gac tgg cag gaa ctt gag atg cag ctg	1780
Leu Ile Pro Glu Glu Tyr Leu Asp Trp Gln Glu Leu Glu Met Gln Leu	
570 575 580 585	
tca gaa gca agc ctt tcc tgt act cgg ctc ttc ctg ctg ctg gga gct	1828
Ser Glu Ala Ser Leu Ser Cys Thr Arg Leu Phe Leu Leu Leu Gly Ala	
590 595 600	
gat ccc tct cct gag aat gat gtc ttt tct agt gac tgt aag cac cag	1876
Asp Pro Ser Pro Glu Asn Asp Val Phe Ser Ser Asp Cys Lys His Gln	
605 610 615	
ctt gtc tcc aca gtg ata ttt cat aca ctt gaa aaa aca caa ttt ttc	1924
Leu Val Ser Thr Val Ile Phe His Thr Leu Glu Lys Thr Gln Phe Phe	
620 625 630	
ctg gaa caa gca tat tat tgg aaa gcc ttc aaa aag ttt atc agg aag	1972
Leu Glu Gln Ala Tyr Tyr Trp Lys Ala Phe Lys Lys Phe Ile Arg Lys	
635 640 645	
act tgc gaa gtg gcc caa tat gta aat atg caa gag agt ttc cag aac	2020
Thr Cys Glu Val Ala Gln Tyr Val Asn Met Gln Glu Ser Phe Gln Asn	
650 655 660 665	
aga cta ttg gct ttt cct gag gaa tct cct tgt ttt gaa gaa aac atg	2068

Arg Leu Leu Ala Phe Pro Glu Glu Ser Pro Cys Phe Glu Glu Asn Met	
670 675 680	
gat tgg aaa atg atc agt gat aat tat ttt caa ttt ttg aat aac tta	2116
Asp Trp Lys Met Ile Ser Asp Asn Tyr Phe Gln Phe Leu Asn Asn Leu	
685 690 695	
ctc aag tct cca aca gct tcc ata tcc agg aag att ctg aga aca aaa	2164
Leu Lys Ser Pro Thr Ala Ser Ile Ser Arg Lys Ile Leu Arg Thr Lys	
700 705 710	
tat ctc ttc tgc tga aatatttcca caaagatggt attgcagaga tgagttttgt	2219
Tyr Leu Phe Cys	
715	
cccaaaagat aaaattctag aaattctgaa actggatcaa tttcttaccg tgatgataca	2279
agacagattg atgaacattt tttcaagttt aaaggagact atatatcacc taatgaaaag	2339
ttcattttata ttagacaatg gagaattttta ttttgatact catcaaggac tgaagttcat	2399
gcaagatttta tttaatgccc ttctcaggga aacttcaatg aaaaataaga ctgaaaataa	2459
tatagactttt ttcacagtgg tgagtcagtt gtttttccat gtgaataagt ctgaggacct	2519
cttcaaactc aatcaagatc ttgggtcagc tcttcacctt gtaagagaat gttcaacaga	2579
gatggcaaga cttctggata caatttttaca ctctcctaata aaggacttct atgctttgta	2639
tcctaccctc caagaagtta tacttgctaa tctaacggat ttgcttttct ttataaataa	2699
ttcattccct ctaagaaaca gagcaacatt agaaattact aagagattag ttggtgctat	2759
ttcaagagca agtgaagaaa gtcacgtcct gaaacccctc ttagaaatgt ctgggactct	2819
ggtcatgctg ttgaatgaca gtgctgacct gagagatctt gccacatcaa tggactccat	2879
tgtgaaactt ctttaagctgg tcaagaaagt ttoggggaag atgtccacag tttttaaaac	2939
tcattttatc tccaatacca aggacagtgt gaaattcttt gacactctgt attccatcat	2999
gcaacaaagt gttcaaaaac ttgtgaaaga aatagctact ttaaaaaaaaa tagatcattt	3059
cacatttgaa aagataaatg atttggttgg gccatttctt gacttggcct ttgaaatgat	3119
tggggtagaa cottatatat catcaaactc tgatattttc agtatgtcac ctagcatact	3179
ctcatatatg aaccaatcta aggacttttc tgatattttg gaagaaattg ctgaattttt	3239
aacatctgtg aaaatgaact tggaagatat gaggagtctt gcggtagcat ttaacaatga	3299
gactcaaaca ttttctatgg attctgtcaa cttangggaa gaaattcttg gttgottagt	3359
tcctataaat aacatcacca accaaatgga cttcttatac cctaatacaa tttccactca	3419
tagtggccct caagatataa aatgggaaat aattcatgaa gtgatccttt ttttgataa	3479
aatattatca caaaacagca cagaaatagg atctttcttg aaaatggtga tctgtctcac	3539
cttagaagct ctttggaaaa acttaagaa agataattgg aatgtttcta atgtgttgat	3599

gacgtttact	cagcatccaa	ataacotttt	gaaaaccata	gaaacagttt	tagaggcctc	3659
cagtgggaatt	aaaagtgact	atgaaggtga	tttgaataaa	agtttatatt	ttgacacacc	3719
tttgagtcag	aataataactc	atcatcaact	tgaaaaagca	atccataatg	ttttaagtag	3779
aatagctctc	tggaggaaaag	gacttcgttt	taacaactct	gaatggataa	cttcacaaag	3839
aactttgttt	cagccacttt	ttgagatttt	cattaaagca	accaccggaa	agaatgtcac	3899
atcagaaaaa	gaagagagaa	ccgagaaaaga	gatgattgac	tttccttata	gtttcaaacc	3959
atttttctgt	ttggagaaat	acctgggagg	attatttgta	ttgactaaat	actggcaaca	4019
aatcccacta	acagatcaaa	gtgttggtga	gattttgtgaa	gttttccagc	agactgtgaa	4079
gccctcagaa	gccatggaga	tgctgcagaa	agtgaagatg	atggctcgtac	gtgtgctcac	4139
catcgttgca	gaaaaccott	cctggaccaa	ggacattttg	tgtgctactc	tgagttgcaa	4199
gcaaaatggg	ataaggcatc	tcattttatc	tgctatacaa	ggggtcactt	tggcgcagga	4259
ccacttccag	gaaattgaaa	agatatggtc	ctcgccgaat	cagctaaatt	gtgaaagtct	4319
tagcaagaat	ctttctagca	ccttggagag	cttcaagagc	agcttggaat	atgccactgg	4379
ccaggactgc	acaagccagc	cgaggctgga	gacggtgcag	cagcacttgt	acatgtttggc	4439
caaaagcctc	gaggaaactt	ggatcatcagg	gaatcccatc	atgacttttc	tcagcaattt	4499
cacagtaact	gaggatgtaa	aaataaaaga	tttgatgaag	aatatcacca	agttgactga	4559
ggagcttcgc	tcttccatcc	aaatctogaa	tgagactatc	catagcattc	tagaagcaaa	4619
tatttccac	tccaagggtc	tcttcagtgc	cctcacgta	gctctgtctg	gaaagtgtga	4679
tcaggaaatc	cttcatctcc	tgctgacatt	tcccaaaggg	gaaaaatctt	ggatcgcagc	4739
ggaggaaactc	tgtagcctgc	cagggtcaaa	agtgtattct	ctgatttgtgt	tgctgagtcg	4799
aaacttggat	gtgcgagctt	tcattttaca	gactctgatg	ccttctgaag	caaattggctt	4859
gctcaactcc	ttgctggata	tagtttccag	cctcagcgcc	ttgcttgcca	aagcccagca	4919
cgtctttgag	tatcttcctg	agtttcttca	cacatttaaa	atcactgcct	tgctagaaac	4979
cctggacttt	caacagggtt	cacaaaatgt	ccaggccaga	agttcagctt	ttggttcttt	5039
ccagtttgtg	atgaagatgg	tttgcaagga	ccaagcatca	ttccttagcg	attctaatat	5099
gtttattaat	ttgccagag	ttaaggaact	cttggaagat	gacaaagaaa	aattcaacat	5159
tcctgaagat	tcaacaccgt	tttgcttgaa	gctttatcag	gaaattctac	aattgccaaa	5219
tggtgctttg	gtgtggacct	tcctaaaacc	catattgcat	ggaaaaatac	tatacacacc	5279
aaacactcca	gaaattaaca	aggtcattca	aaaggcta	tacacctttt	atattgtgga	5339
caaaactaaa	actttatcag	aaacactgct	ggaaatgtcc	agccttttcc	agagaagtgg	5399
aagtggccag	atgttcaacc	agctgcagga	ggccctgaga	aacaaatttg	taagaaactt	5459

tgtagaaaac cagttgcaca ttgatgtaga caaacttact gaaaaactcc agacatacgg	5519
agggctgctg gatgagatgt ttaacatgc aggcgctgga cgcttccgtt tcttgggcag	5579
catcttggtc aatctctctt cctgcgtggc actgaaccgt ttccaggctc tgcagtctgt	5639
cgacatcctg gagactaaag cacatgaact cttgcagcag aacagcttct tggccagtat	5699
cattttcagc aattccttat togacaagaa cttcagatca gagtctgtca aactgccacc	5759
ccatgtctca tacacaatcc ggaaccaatgt gttatacagc gtgcgaacag atgtggtaaa	5819
aaacccttct tggaagtcc accctcagaa tctaccagct gatgggttca aatataacta	5879
cgtctttgcc ccaactgcaag acatgatcga aagagccatc attttgggtgc agactgggca	5939
ggaagccctg gaaccagcag cacagactca ggcggcccct taccctgcc ataccagcga	5999
cctattcctg aacaacgttg gtttcttttt tccactgata atgatgctga cgtggatgg	6059
gtctgtggcc agcatgggtca gaaagtgggt gtatgagcag gagatacaga tagaagagta	6119
tatgcggatg atgggagtgc atccagtgat ccatttcctg gcctgggttc tggagaacat	6179
ggctgtgttg accataagca gtgctactct ggccatcgtt ctgaaaacaa gtggcatctt	6239
tgcacacagc aataccttta ttgttttctt ctttctcttg gattttggga tgtcagtcgt	6299
catgctgagc tacctottga gtgcattttt cagccaagct aatacagcgg ccctttgtac	6359
cagcctggtg tacatgatca gctttctgcc ctacatagtt ctattgggtc tacataacca	6419
attaagtttt gttaatcaga catttctgtg ccttctttcg acaaccgcct ttggacaagg	6479
ggtatttttt attacattcc tggaaggaca agagacaggg attcaatgga ataatatgta	6539
ccaggctctg gaacaagggg gcatgacatt tggctgggtt tgctggatga ttctttttga	6599
ttcaagcctt tattttttgt gtggatggta cttgagcaac ttgattcctg gaacatttgg	6659
tttacggaaa ccatggtatt tcccccttac tgccatcatat tggaagagtg tgggtttctt	6719
ggtggagaaa aggcaatact ttctaagttc tagtctgttc ttcttcaatg agaactttga	6779
caataaaggg tcatcactgc aaaacaggga aggagagctt gaaggaagtg ccccgaggat	6839
caccctggtg tctgtgacca aggaatatga gggccacaag gctgtggtcc aagacctcag	6899
cctgaccttc tacagagacc aaatcacccg cctgctgggg acaaacgggt cgggaaaaac	6959
cactatcata tccatgttga cggggctcca cctccact tctggaacca tcatcatcaa	7019
tggcaagaac ctacagacag acctgtcgag ggtcagaatg gagcttgggtg tgtgtccgca	7079
gcaggacatc ctgttggaca acctcacccg ccgggaacat ttgctgctct ttgcttccat	7139
aaaggcgcct cagtggacca agaaggagct gcatcagcaa gtcaatcaaa ctcttcagga	7199
tgtggactta actcagcatc agcacaacaa gaccogagct ctgtctggag gcctgaagag	7259

gaagctctcc cttggcattg ctttcatggg catgtcgagg accgtgggtc tggatgagcc	7319
caccagtggg gtggaccctt gctcccgga tagcctgtgg gacattctgc tcaagtaccg	7379
agaaggctgt acgatcatct tcacaacca ccacctggat gaagccgaag cgctgagtga	7439
ccgcgtggcc gtccctccagc atgggaggct cagggtgctgc ggtcctccct tctgcctgaa	7499
ggaggcatat ggccaggggc tccgcctgac actcacgagg cagccttctg ttctggaggc	7559
ccatgatctg aaagacatgg cttgtgttac atccctgata aagatctata ttccacaagc	7619
atttctcaaa gacagcagtg gaagtgaagt gacctacacc attccaaagg acacagacaa	7679
ggcctgcttg aaagggctct tccaggccct ggatgagaac ctgcatcagc tgcacctgac	7739
gggctatggg atctcagaca ccacctaga agaggtgttt ttgatgcttt tgcaagattc	7799
caacaagaaa tctcacattg ccctggggac tgagtcagag ctgcagaacc acaggcctac	7859
aggacatctg tctggctact gtggctccct agcacggccc gcaactgtgc agggcgcca	7919
gctgtccgc gcacaagtgg ccgcgacct ggcccgagg ctccgcgca cgctgcgcgc	7979
cggaagagc accctcgccg acctgctgct gccagtctc ttcgtggcct tggccatggg	8039
cttgttcatg gtgagacccc tggccacoga gtacctccc ctgagactca cacctggaca	8099
ttaccagcgg gccgagacct actttttcag cagtgggggc gacaacttgg acctcaccg	8159
tgtgcttctg cggaagtta gagatcaaga tttgccctgt gcagatttaa acccacgcca	8219
gaagaattct tcatgctggc gcacagatcc cttttctcac ccagaattcc aggatctatg	8279
tggctgcctg aagtgtccaa atagaagtgc tagtgctccc tacctgacca accacctggg	8339
ccacacactg ttgaatctct caggcttcaa tatggaggag tacttgctgg caccatctga	8399
aaaaccaagg cttggagggtt ggtcttttgg attaaaaatc ccagtgagc ctggagggtgc	8459
aatggaaac atatcaaaac cccaactct ggcaaagggtg tggataatc agaagggttt	8519
tcattcccta ctttctact taaatcatct aaacaacctt attttgtggc agcacctacc	8579
ccctactgtg gactggagac aatacgaat aacactctac agccacccat atggaggggc	8639
cttgctgaac gaggacaaga tcctggagag catccgtcag tgtggagtgg ccctctgcat	8699
cgtgctggga ttctccatcc tgtctgcac catcggcagc tctgtggtga gggacagggt	8759
gattggagcc aaaagggttc agcacataag tggccttggc tacaggatgt actggttcac	8819
aaacttcta tatgacatgc tcttttactt ggtttccgtc tgccctgtgtg ttgcgttat	8879
tgtcgcttc cagttaacag cttttacttt ccgcaagaac ttggcagcca cgccctcct	8939
gctgtcactt ttcggatatg caactcttcc atggatgtac ctgatgtcca gaatcttttc	8999
cagttcggac gtggctttca tttctatgt ctcaactaaac ttcattcttg gcctttgtac	9059
catgcccata accattatgc cccggttgc agccatcatc tccaaagcta agaatttaca	9119

gaatatctat gatgtcctca agtgggtcctt tactatTTTT cctcaattct gtcttggtca	9179
aggactggta gaactctgct ataatcagat caaatatgac ctgaccaca acttcggcat	9239
tgattcctat gtgagtcctt ttgagatgaa ctttctgggc tggatcttcg tgcaactggc	9299
ctcgcagggc acagtacttc tcctcttgag ggttctgcta cactgggacc ttctgcgatg	9359
gccaaggggt cattctactc tccaaggcac agtcaaactc tctaaggata cagatgttga	9419
aaaagaggaa aagagagtgt ttgaaggag gaccaatgga gacattcttg tgttatacaa	9479
ccttagtaaa cattatcgac gctttttoca gaatattatt gctgtgcaag atattagttt	9539
gggcatacca aaaggagagt gctttggact tctaggggtg aatggagctg ggaagagcac	9599
gactttcaaa atgctgaatg gtgaagtctc tctaacttca ggacatgcta tcatcaggac	9659
tcccatggga gacgccgtgg acctgtcttc tgctggcacg gcaggcgtgc tcattggcta	9719
ctgtccccag caggatgccc tggacgagct totgactggg tgggaacatc tctattatta	9779
ctgtagctta cgcgggattc caaggcagtg catccctgag gttgctggag acctcatcag	9839
gcgcttacac ctgcaagccc acgcggacaa acctgtggcc acctacagtg ggggaaccaa	9899
gcggaaactc tctacagccc tggccctggg ggggaaacct gacattcttt tattggatga	9959
gccagctctt gggatggatc cctgctctaa gcggtacctg tggcaaacaa taatgaagga	10019
ggttcgggaa ggctgtgctg cggtgctgac ctcccacagc atggaggagt gtgaggctct	10079
ttgcacaaga ctggccataa tggttaacgg cagcttcaaa tgtcttggtt ctccctcagca	10139
catcaaaaat aggtttgggt atggttatac agtcaaagtt tggctctgta aggaagcaaa	10199
tcaacattgc actgtttctg accacttgaa gctttatTTT ccaggaattc agttcaaggg	10259
acagcacctg aatttattag aatatcatgt gccaaaaaga tggggatgcc tagctgactt	10319
gttcaaagtt atagagaaca ataaaacctt cttgaatatt aagcattatt ccattaacca	10379
aaccactttg gagcaggtat ttattaattt tgcttctgag cagcagcaaa ctctacaatc	10439
tactcttgat ccatccactg acagtcacca cacacatcac ttgcccactc gagcactaaa	10499
gaagtttcca taaggaataa aaccttgtct tcattacaa ttaacagtca aggataaaac	10559
aagcacgcg acaatcaagg agctggaaca cactctccag gccgtcaaat tattctcttg	10619
ttcattttct attttgaatc tccttgtag ttaataacca ccaaatggaa aggtcattct	10679
ttctgcagac ttttggggag ctccctcaaaa acatttgTtc tctttaccat gccagatgga	10739
caccagcttc tttgtgacaa aggcatgaat gatttgacag tgtccaaact gagacattct	10799
ggagctggaa agcctgtcac actagagtgt gtgtgacatg tccactctaa acatgtcact	10859
tttctgttaa gaaaactgag cccctcccc acagggttaa aaactttagt aacttgtttg	10919

tatagaaaat agtaacaagg actatTTTTct attgTtGtca tctatTTtact agatacatgt	10979
TTTTaatgat TTTaatgtaa gCTTTTTatta atactgatga cattatatgg tatgatatga	11039
aaaaatcacc aATTTTTttac atataaaaga tacCTTTTTta aaaaaaatagg TTTtaagagc	11099
tCTTTtagta tacactTTtag caaaattaat taaattgaac tagTTactct gtatcaatta	11159
cagtagTTct accagaatct cccaggTTat aATTTatgag ggtagagaaa taaaatgtag	11219
atgcATTTTc TTTTTcttca TTTggatgaa taattactgt TTTTgttat tctaagtcag	11279
TgtTTTTcaa agcgtagtgg TCCCCatatt agctgcattg ccatCTTctg ggagCTtgca	11339
agaaatgtac attctcagga tccactccag acctattgaa tctcaaattc Tggggcttaa	11399
acaagcacat tccaagTtaa gaaccaatga cctaagggaa Tgtctggtta cctcctagtt	11459
atacaagcaa aatctgcata gtatgtagtc TTTTTatTTt atttattTTTt TTTTTtgag	11519
gcggagtctc gctctgtcac ccaggctgga gtgcagtggc gcgatgtcgg ctactgcaa	11579
gctccgcctc ctgggttcac gccattctcc Tgcctcagcc TCCCCagcag ctgggactac	11639
aggcacacat cgccacaccc ggctaATTTt tagtattTTTt agtagagacg gggTTTcacc	11699
gtgTtagcca ggatggTctc tatctcctga cTTgtgatc cGCCCCctc cacctcccaa	11759
agtgtctggga ttacaggcgt gagccaccgt gtccggccgt agTTtattTTt aaaatatatt	11819
TTTaaaagct ttgtaaaaat tatgtcattc tcagaattgt Tgtcttcaaa gcattgtcag	11879
atgtagagtg ctcagatgtg gctcttaaag actatataca tctgaATTTt tcatctata	11939
gttagtaaga Tgcataaaat caatccacta ctgaaatagt ttccagtcag acatttctga	11999
gttcagacat ttctcaacat tCTTTtaaca acattTTTct gaatcctcaa tagaaaaatca	12059
cattaatctt atTTTaaaat ttggcCTTTt tcaacactaa cgttgagtac cggtagcttt	12119
gtgatcaaag gcatatactt cTTtatgaga tttctTTtact aaagcaagat ttcattaaat	12179
ctctattTcc taaatatcat tctatacaaa agatattTTTt taaacggtaa ggattaagac	12239
aatcactgat agctTtgttg tgagcaattt tgattcccat gtatcacatg aattacactt	12299
ccttaaataa attacagTTt atggctgtat gatttattct ctaattctaa catagtctag	12359
ttgtcaaaag gaaatatgta atCTTTTtat gattgttgaa tcaataaata ccaatttgtg	12419
aaacatgaat gtgtTTaaac Tgcagtgaat aaatgagatg TgctTTaatt taacccaaaa	12479
aaaaaaaaa aaaaaaaaaa	12498

<210> 21
 <211> 717
 <212> PRT
 <213> Homo Sapiens

<220>

<221> misc_feature
 <222> (283)..(283)
 <223> The 'Xaa' at location 283 stands for Asp, Gly, Ala, or Val.
 <400> 21

Met Gly His Ala Gly Cys Gln Phe Lys Ala Leu Leu Trp Lys Asn Trp
 1 5 10 15

Leu Cys Arg Leu Arg Asn Pro Val Leu Phe Leu Ala Glu Phe Phe Trp
 20 25 30

Pro Cys Ile Leu Phe Val Ile Leu Thr Val Leu Arg Phe Gln Glu Pro
 35 40 45

Pro Arg Tyr Arg Asp Ile Cys Tyr Leu Gln Pro Arg Asp Leu Pro Ser
 50 55 60

Cys Gly Val Ile Pro Phe Val Gln Ser Leu Leu Cys Asn Thr Gly Ser
 65 70 75 80

Arg Cys Arg Asn Phe Ser Tyr Glu Gly Ser Met Glu His His Phe Arg
 85 90 95

Leu Ser Arg Phe Gln Thr Ala Ala Asp Pro Lys Lys Val Asn Asn Leu
 100 105 110

Ala Phe Leu Lys Glu Ile Gln Asp Leu Ala Glu Glu Ile His Gly Met
 115 120 125

Met Asp Lys Ala Lys Asn Leu Lys Arg Leu Trp Val Glu Arg Ser Asn
 130 135 140

Thr Pro Asp Ser Ser Tyr Gly Ser Ser Phe Phe Thr Met Asp Leu Asn
 145 150 155 160

Lys Thr Glu Glu Val Ile Leu Lys Leu Glu Ser Leu His Gln Gln Pro
 165 170 175

His Ile Trp Asp Phe Leu Leu Leu Leu Pro Arg Leu His Thr Ser His
 180 185 190

Asp His Val Glu Asp Gly Met Asp Val Ala Val Asn Leu Leu Gln Thr
 195 200 205

Ile Leu Asn Ser Leu Ile Ser Leu Glu Asp Leu Asp Trp Leu Pro Leu
 210 215 220

Asn Gln Thr Phe Ser Gln Val Ser Glu Leu Val Leu Asn Val Thr Ile
 225 230 235 240

Ser Thr Leu Thr Phe Leu Gln Gln His Gly Val Ala Val Thr Glu Pro
 245 250 255

Val Tyr His Leu Ser Met Gln Asn Ile Val Trp Asp Pro Gln Lys Val
 260 265 270

Gln Tyr Asp Leu Lys Ser Gln Phe Gly Phe Xaa Asp Leu His Thr Glu
 275 280 285

Gln Ile Leu Asn Ser Ser Ala Glu Leu Lys Glu Ile Pro Thr Asp Thr
 290 295 300

Ser Leu Glu Lys Met Val Cys Ser Val Leu Ser Ser Thr Ser Glu Asp
 305 310 315 320

Glu Ala Glu Lys Trp Gly His Val Gly Gly Cys His Pro Lys Trp Ser
 325 330 335

Glu Ala Lys Asn Tyr Leu Val His Ala Val Ser Trp Leu Arg Val Tyr
 340 345 350

Gln Gln Val Phe Val Gln Trp Gln Gln Gly Ser Leu Leu Gln Lys Thr
 355 360 365

Leu Thr Gly Met Gly His Ser Leu Glu Ala Leu Arg Asn Gln Phe Glu
 370 375 380

Glu Glu Ser Lys Pro Trp Lys Val Val Glu Ala Leu His Thr Ala Leu
 385 390 395 400

Leu Leu Leu Asn Asp Ser Leu Ser Ala Asp Gly Pro Lys Asp Asn His
 405 410 415

Thr Phe Pro Lys Ile Leu Gln His Leu Trp Lys Leu Gln Ser Leu Leu
 420 425 430

Gln Asn Leu Pro Gln Trp Pro Ala Leu Lys Arg Phe Leu Gln Leu Asp
 435 440 445

Gly Ala Leu Arg Asn Ala Ile Ala Gln Asn Leu His Phe Val Gln Glu
 450 455 460

Val Leu Ile Cys Leu Glu Thr Ser Ala Asn Asp Phe Lys Trp Phe Glu
465 470 475 480

Leu Asn Gln Leu Lys Leu Glu Lys Asp Val Phe Phe Trp Glu Leu Lys
485 490 495

Gln Met Leu Ala Lys Asn Ala Val Cys Pro Asn Gly Arg Phe Ser Glu
500 505 510

Lys Glu Val Phe Leu Pro Pro Gly Asn Ser Ser Ile Trp Gly Gly Leu
515 520 525

Gln Gly Leu Leu Cys Tyr Cys Asn Ser Ser Glu Thr Ser Val Leu Asn
530 535 540

Lys Leu Leu Gly Ser Val Glu Asp Ala Asp Arg Ile Leu Gln Glu Val
545 550 555 560

Ile Thr Trp His Lys Asn Met Ser Val Leu Ile Pro Glu Glu Tyr Leu
565 570 575

Asp Trp Gln Glu Leu Glu Met Gln Leu Ser Glu Ala Ser Leu Ser Cys
580 585 590

Thr Arg Leu Phe Leu Leu Leu Gly Ala Asp Pro Ser Pro Glu Asn Asp
595 600 605

Val Phe Ser Ser Asp Cys Lys His Gln Leu Val Ser Thr Val Ile Phe
610 615 620

His Thr Leu Glu Lys Thr Gln Phe Phe Leu Glu Gln Ala Tyr Tyr Trp
625 630 635 640

Lys Ala Phe Lys Lys Phe Ile Arg Lys Thr Cys Glu Val Ala Gln Tyr
645 650 655

Val Asn Met Gln Glu Ser Phe Gln Asn Arg Leu Leu Ala Phe Pro Glu
660 665 670

Glu Ser Pro Cys Phe Glu Glu Asn Met Asp Trp Lys Met Ile Ser Asp
675 680 685

Asn Tyr Phe Gln Phe Leu Asn Asn Leu Leu Lys Ser Pro Thr Ala Ser
690 695 700

Ile Ser Arg Lys Ile Leu Arg Thr Lys Tyr Leu Phe Cys

Leu	Trp	Val	Glu	Arg	Ser	Asn	Thr	Pro	Asp	Ser	Ser	Tyr	Gly	Ser	Ser		
		140					145					150					
ttt	ttt	aca	atg	gat	ctc	aat	aag	acc	gag	gag	gta	ata	ttg	aaa	ctg	532	
Phe	Phe	Thr	Met	Asp	Leu	Asn	Lys	Thr	Glu	Glu	Val	Ile	Leu	Lys	Leu		
	155					160					165						
gaa	agc	ctc	cat	cag	cag	cct	cat	atc	tgg	gat	ttt	cta	ctt	tta	ctg	580	
Glu	Ser	Leu	His	Gln	Gln	Pro	His	Ile	Trp	Asp	Phe	Leu	Leu	Leu	Leu		
170					175					180					185		
ccg	aga	cta	cac	aca	agc	cat	gat	cat	gtg	gaa	gat	ggc	atg	gat	gtt	628	
Pro	Arg	Leu	His	Thr	Ser	His	Asp	His	Val	Glu	Asp	Gly	Met	Asp	Val		
				190					195					200			
gca	gtg	aac	ctt	ctc	cag	acc	att	ttg	aat	tcc	tta	ata	tcc	cta	gaa	676	
Ala	Val	Asn	Leu	Leu	Gln	Thr	Ile	Leu	Asn	Ser	Leu	Ile	Ser	Leu	Glu		
		205					210						215				
gat	tta	gat	tgg	ctt	cca	ctc	aac	caa	act	ttt	tcc	cag	gtt	tct	gaa	724	
Asp	Leu	Asp	Trp	Leu	Pro	Leu	Asn	Gln	Thr	Phe	Ser	Gln	Val	Ser	Glu		
	220					225						230					
ctt	gta	ctg	aat	gtg	acc	att	tcg	aca	ctg	aca	ttt	ctg	cag	caa	cat	772	
Leu	Val	Leu	Asn	Val	Thr	Ile	Ser	Thr	Leu	Thr	Phe	Leu	Gln	Gln	His		
	235					240					245						
gga	gta	gca	gtc	acc	gag	cca	gtt	tac	cac	ctg	tcc	atg	cag	aat	ata	820	
Gly	Val	Ala	Val	Thr	Glu	Pro	Val	Tyr	His	Leu	Ser	Met	Gln	Asn	Ile		
250					255					260					265		
gtg	tgg	gat	cca	cag	aaa	gtc	cag	tat	gat	ctc	aaa	tcc	cag	ttt	ggc	868	
Val	Trp	Asp	Pro	Gln	Lys	Val	Gln	Tyr	Asp	Leu	Lys	Ser	Gln	Phe	Gly		
				270					275					280			
ttt	gnt	gat	ctt	cac	acg	gaa	cag	atc	ctg	aac	tct	tca	gct	gaa	ctg	916	
Phe	Xaa	Asp	Leu	His	Thr	Glu	Gln	Ile	Leu	Asn	Ser	Ser	Ala	Glu	Leu		
			285					290					295				
aag	gag	att	ccc	aca	gac	act	tcc	ttg	gag	aag	atg	gtg	tgt	tca	gtc	964	
Lys	Glu	Ile	Pro	Thr	Asp	Thr	Ser	Leu	Glu	Lys	Met	Val	Cys	Ser	Val		
	300						305					310					
ttg	tct	agc	aca	tca	gag	gat	gaa	gct	gag	aaa	tgg	ggc	cac	gtt	gga	1012	
Leu	Ser	Ser	Thr	Ser	Glu	Asp	Glu	Ala	Glu	Lys	Trp	Gly	His	Val	Gly		
	315					320					325						
ggc	tgc	cac	cct	aag	tgg	tca	gaa	gcc	aaa	aac	tat	ctt	gtc	cat	gca	1060	
Gly	Cys	His	Pro	Lys	Trp	Ser	Glu	Ala	Lys	Asn	Tyr	Leu	Val	His	Ala		
330					335					340					345		
gtc	agc	tgg	ctg	cga	gtc	tac	caa	cag	gtg	ttt	gtt	cag	tgg	caa	cag	1108	
Val	Ser	Trp	Leu	Arg	Val	Tyr	Gln	Gln	Val	Phe	Val	Gln	Trp	Gln	Gln		
				350					355					360			
ggt	agc	ctg	ctt	cag	aag	aca	ctc	aca	ggc	atg	ggc	cat	agt	ctg	gag	1156	
Gly	Ser	Leu	Leu	Gln	Lys	Thr	Leu	Thr	Gly	Met	Gly	His	Ser	Leu	Glu		
			365					370					375				
gct	ctc	agg	aat	cag	ttt	gaa	gaa	gag	agc	aag	ccc	tgg	aag	gtg	gtg	1204	
Ala	Leu	Arg	Asn	Gln	Phe	Glu	Glu	Glu	Ser	Lys	Pro	Trp	Lys	Val	Val		

380	385	390	
gaa gct ctg cac act gca ctg ctc ctg ctg aat gac agc ttg tca gca Glu Ala Leu His Thr Ala Leu Leu Leu Leu Asn Asp Ser Leu Ser Ala 395 400 405			1252
gat ggc cca aaa gat aat cat aca ttt cca aag ata tta cag cat ctg Asp Gly Pro Lys Asp Asn His Thr Phe Pro Lys Ile Leu Gln His Leu 410 415 420 425			1300
tgg aaa ttg caa agc ttg ctg caa aac ctg ccc cag tgg ccg gca ctg Trp Lys Leu Gln Ser Leu Leu Gln Asn Leu Pro Gln Trp Pro Ala Leu 430 435 440			1348
aag aga ttt ctt cag ctt gat gga gct ctc aga aat gcg ata gct cag Lys Arg Phe Leu Gln Leu Asp Gly Ala Leu Arg Asn Ala Ile Ala Gln 445 450 455			1396
aat tta cat ttt gtc caa gaa gtc ctc att tgc ctg gag aca tca gct Asn Leu His Phe Val Gln Glu Val Leu Ile Cys Leu Glu Thr Ser Ala 460 465 470			1444
aat gat ttt aaa tgg ttt gaa ctt aac caa ttg aaa ctg gaa aag gat Asn Asp Phe Lys Trp Phe Glu Leu Asn Gln Leu Lys Leu Glu Lys Asp 475 480 485			1492
gtg ttc ttt tgg gag ctg aaa cag atg ttg gcg aag aat gct gtc tgc Val Phe Phe Trp Glu Leu Lys Gln Met Leu Ala Lys Asn Ala Val Cys 490 495 500 505			1540
ccg aat ggt cgt ttc tct gag aag gag gtc ttt ttg ccg cct gga aac Pro Asn Gly Arg Phe Ser Glu Lys Glu Val Phe Leu Pro Pro Gly Asn 510 515 520			1588
tcc agc ata tgg ggt ggt ctc cag gga ctg ttg tgc tat tgt aac tcc Ser Ser Ile Trp Gly Gly Leu Gln Gly Leu Leu Cys Tyr Cys Asn Ser 525 530 535			1636
tct gag acg agt gtt tta aac aag cta ctt ggt tca gta gag gat gct Ser Glu Thr Ser Val Leu Asn Lys Leu Leu Gly Ser Val Glu Asp Ala 540 545 550			1684
gat cgt att ttg caa gag gtc att act tgg cac aaa aat atg tca gtt Asp Arg Ile Leu Gln Glu Val Ile Thr Trp His Lys Asn Met Ser Val 555 560 565			1732
tta ata cct gaa gaa tat ttg gac tgg cag gaa ctt gag atg cag ctg Leu Ile Pro Glu Glu Tyr Leu Asp Trp Gln Glu Leu Glu Met Gln Leu 570 575 580 585			1780
tca gaa gca agc ctt tcc tgt act cgg ctc ttc ctg ctg ctg gga gct Ser Glu Ala Ser Leu Ser Cys Thr Arg Leu Phe Leu Leu Leu Gly Ala 590 595 600			1828
gat ccc tct cct gag aat gat gtc ttt tct agt gac tgt aag cac cag Asp Pro Ser Pro Glu Asn Asp Val Phe Ser Ser Asp Cys Lys His Gln 605 610 615			1876
ctt gtc tcc aca gtg ata ttt cat aca ctt gaa aaa aca caa ttt ttc Leu Val Ser Thr Val Ile Phe His Thr Leu Glu Lys Thr Gln Phe Phe 620 625 630			1924

ctg gaa caa gca tat tat tgg aaa gcc ttc aaa aag ttt atc agg aag	1972
Leu Glu Gln Ala Tyr Tyr Trp Lys Ala Phe Lys Lys Phe Ile Arg Lys	
635 640 645	
act tgc gaa gtg gcc caa tat gta aat atg caa gag agt ttc cag aac	2020
Thr Cys Glu Val Ala Gln Tyr Val Asn Met Gln Glu Ser Phe Gln Asn	
650 655 660 665	
aga cta ttg gct ttt cct gag gaa tct cct tgt ttt gaa gaa aac atg	2068
Arg Leu Leu Ala Phe Pro Glu Glu Ser Pro Cys Phe Glu Glu Asn Met	
670 675 680	
gat tgg aaa atg atc agt gat aat tat ttt caa ttt ttg aat aac tta	2116
Asp Trp Lys Met Ile Ser Asp Asn Tyr Phe Gln Phe Leu Asn Asn Leu	
685 690 695	
ctc aag tct cca aca gct tcc ata tcc agg gct tta aat ttc aca aag	2164
Leu Lys Ser Pro Thr Ala Ser Ile Ser Arg Ala Leu Asn Phe Thr Lys	
700 705 710	
cac ctt cta atg atg gaa aag aag ttg cac acc ctt gag gat gaa caa	2212
His Leu Leu Met Met Glu Lys Lys Leu His Thr Leu Glu Asp Glu Gln	
715 720 725	
atg aac ttt ctt tta tca ttt gtg gaa ttt ttt gag aaa tta ttg ttg	2260
Met Asn Phe Leu Leu Ser Phe Val Glu Phe Phe Glu Lys Leu Leu Leu	
730 735 740 745	
cct aat ctt ttt gac tcc tcc att gtt ccc agt ttc cac agc ctc cca	2308
Pro Asn Leu Phe Asp Ser Ser Ile Val Pro Ser Phe His Ser Leu Pro	
750 755 760	
tct ctc aca gag gat att ctg aat ata agt tct ctg tgg aca aat cat	2356
Ser Leu Thr Glu Asp Ile Leu Asn Ile Ser Ser Leu Trp Thr Asn His	
765 770 775	
tta aaa agt tta aag aga gac cca tct gcc act gat gct cag aaa ctc	2404
Leu Lys Ser Leu Lys Arg Asp Pro Ser Ala Thr Asp Ala Gln Lys Leu	
780 785 790	
ttg gaa ttt ggc aac gaa gtg att tgg aaa atg cag act ctc gga agt	2452
Leu Glu Phe Gly Asn Glu Val Ile Trp Lys Met Gln Thr Leu Gly Ser	
795 800 805	
cac tgg ata agg aag gaa cca aaa aat ctt ttg aga ttc ata gaa tta	2500
His Trp Ile Arg Lys Glu Pro Lys Asn Leu Leu Arg Phe Ile Glu Leu	
810 815 820 825	
ata ctt ttt gaa att aat ccc aaa tta cta gaa tta tgg gcc tat ggc	2548
Ile Leu Phe Glu Ile Asn Pro Lys Leu Leu Glu Leu Trp Ala Tyr Gly	
830 835 840	
att tca aaa gga aaa aga gct aaa ttg gaa aac ttc ttt aca ctt tta	2596
Ile Ser Lys Gly Lys Arg Ala Lys Leu Glu Asn Phe Phe Thr Leu Leu	
845 850 855	
aat ttt tct gtt cca gaa aat gag att ctg agt aca agt ttt aac ttt	2644
Asn Phe Ser Val Pro Glu Asn Glu Ile Leu Ser Thr Ser Phe Asn Phe	
860 865 870	

tcc cag ttg ttc cat tca gat tgg cct aaa tca cca gct atg aac ata Ser Gln Leu Phe His Ser Asp Trp Pro Lys Ser Pro Ala Met Asn Ile 875 880 885	2692
gat ttt gta cgt tta agt gag gct ata ata act agt ctc cat gaa ttt Asp Phe Val Arg Leu Ser Glu Ala Ile Ile Thr Ser Leu His Glu Phe 890 895 900 905	2740
gga ttt ttg gag cag gaa cag atc tca gaa gct ctg aac aca gtc tac Gly Phe Leu Glu Gln Glu Gln Ile Ser Glu Ala Leu Asn Thr Val Tyr 910 915 920	2788
gct atc agg aat gca tct gat ctt ttc tca gcc ctt tct gaa cca caa Ala Ile Arg Asn Ala Ser Asp Leu Phe Ser Ala Leu Ser Glu Pro Gln 925 930 935	2836
aaa caa gaa gtt gat aaa att ttg act cac ata cac cta aat gtc ttc Lys Gln Glu Val Asp Lys Ile Leu Thr His Ile His Leu Asn Val Phe 940 945 950	2884
cag gac aag gat tca gct tta ctt ctg caa att tat tct tca ttt tac Gln Asp Lys Asp Ser Ala Leu Leu Leu Gln Ile Tyr Ser Ser Phe Tyr 955 960 965	2932
cga tat att tat gaa tta ttg aat att cag agt aga ggc tct tcg ttg Arg Tyr Ile Tyr Glu Leu Leu Asn Ile Gln Ser Arg Gly Ser Ser Leu 970 975 980 985	2980
act ttc ctt aca caa atc tca aaa cac att ttg gat atc ata aaa caa Thr Phe Leu Thr Gln Ile Ser Lys His Ile Leu Asp Ile Ile Lys Gln 990 995 1000	3028
ttt aat ttc caa aac atc agt aaa gca ttt gca ttt tta ttt aag Phe Asn Phe Gln Asn Ile Ser Lys Ala Phe Ala Phe Leu Phe Lys 1005 1010 1015	3073
aca gca gag gtt ctt ggg gga att tct aat gta tct tac tgt cag Thr Ala Glu Val Leu Gly Gly Ile Ser Asn Val Ser Tyr Cys Gln 1020 1025 1030	3118
caa ttg ctt tca att ttt aac ttt ttg gag ctt cag gcc caa tcc Gln Leu Leu Ser Ile Phe Asn Phe Leu Glu Leu Gln Ala Gln Ser 1035 1040 1045	3163
ttc atg tct aca gag ggc caa gaa ctg gaa gtg atc cac act act Phe Met Ser Thr Glu Gly Gln Glu Leu Glu Val Ile His Thr Thr 1050 1055 1060	3208
ttg aca ggc ctc aaa cag ctg ctc ata att gat gaa gat ttt cgt Leu Thr Gly Leu Lys Gln Leu Leu Ile Ile Asp Glu Asp Phe Arg 1065 1070 1075	3253
att tct tta ttt caa tat atg agc caa ttc ttc aac agt tca gta Ile Ser Leu Phe Gln Tyr Met Ser Gln Phe Phe Asn Ser Ser Val 1080 1085 1090	3298
gaa gac cta ttg gat aat aaa tgc ttg att tcg gac aat aaa cac Glu Asp Leu Leu Asp Asn Lys Cys Leu Ile Ser Asp Asn Lys His 1095 1100 1105	3343
att tct tcc gta aat tat tca aca agt gag gag tct tca ttt gtt	3388

Ile Ser Ser Val	Asn Tyr Ser Thr Ser	Glu Glu Ser Ser Phe Val	
1110	1115	1120	
ttt cca ttg gca	caa att ttt tca aac	ctc tca gca aat gtc agt	3433
Phe Pro Leu Ala	Gln Ile Phe Ser Asn	Leu Ser Ala Asn Val Ser	
1125	1130	1135	
gtg ttc aac aag	ttt atg tcc att cac	tgt acc gtt tca tgg ctt	3478
Val Phe Asn Lys	Phe Met Ser Ile His	Cys Thr Val Ser Trp Leu	
1140	1145	1150	
caa atg tgg act	gaa atc tgg gaa acc	ata tct caa tta ttt aag	3523
Gln Met Trp Thr	Glu Ile Trp Glu Thr	Ile Ser Gln Leu Phe Lys	
1155	1160	1165	
ttt gac atg aat	gtt ttc aca tct ctt	cat cat ggt ttc act cag	3568
Phe Asp Met Asn	Val Phe Thr Ser Leu	His His Gly Phe Thr Gln	
1170	1175	1180	
ctt ttg gat gaa	ttg gaa gat gat gtg	aaa gtc tct aaa agc tgc	3613
Leu Leu Asp Glu	Leu Glu Asp Asp Val	Lys Val Ser Lys Ser Cys	
1185	1190	1195	
cag ggt ata ctt	ccc acc cat aat gtt	gct aga ctc ata tta aat	3658
Gln Gly Ile Leu	Pro Thr His Asn Val	Ala Arg Leu Ile Leu Asn	
1200	1205	1210	
ttg ttt aaa aat	gta act caa gcc aat	gac ttc cat aat tgg gag	3703
Leu Phe Lys Asn	Val Thr Gln Ala Asn	Asp Phe His Asn Trp Glu	
1215	1220	1225	
gac ttc ctg gat	ctc agg gat ttt ttg	gta gct tta ggt aat gca	3748
Asp Phe Leu Asp	Leu Arg Asp Phe Leu	Val Ala Leu Gly Asn Ala	
1230	1235	1240	
tta gtt tca gta	aaa aaa ctt aac ttg	gag caa gtg gag aaa tcc	3793
Leu Val Ser Val	Lys Lys Leu Asn Leu	Glu Gln Val Glu Lys Ser	
1245	1250	1255	
ctt ttc acc atg	gaa gct gcc ctg cat	cag ttg aag aca ttt cca	3838
Leu Phe Thr Met	Glu Ala Ala Leu His	Gln Leu Lys Thr Phe Pro	
1260	1265	1270	
ttc aac gaa agt	aca agc aga gag ttt	tta aat tct ctg ctt gaa	3883
Phe Asn Glu Ser	Thr Ser Arg Glu Phe	Leu Asn Ser Leu Leu Glu	
1275	1280	1285	
gtt ttc att gag	ttt agc agt acc tca	gaa tat ata gtc aga aat	3928
Val Phe Ile Glu	Phe Ser Ser Thr Ser	Glu Tyr Ile Val Arg Asn	
1290	1295	1300	
cta gat tca ata	aat gac ttt ctt tca	aat aat ctc aca aat tat	3973
Leu Asp Ser Ile	Asn Asp Phe Leu Ser	Asn Asn Leu Thr Asn Tyr	
1305	1310	1315	
gga gaa aaa ttt	gaa aat atc atc act	gag cta aga gaa gca ata	4018
Gly Glu Lys Phe	Glu Asn Ile Ile Thr	Glu Leu Arg Glu Ala Ile	
1320	1325	1330	
gta ttt ctt aga	aat gta tca cat gat	cga gat ttg ttt tcc tgt	4063
Val Phe Leu Arg	Asn Val Ser His Asp	Arg Asp Leu Phe Ser Cys	

1335	1340	1345	
gct gat att ttc Ala Asp Ile Phe 1350	caa aat gtt act gag Gln Asn Val Thr 1355	tgt att tta gaa gat ggc Cys Ile Leu Glu Asp Gly 1360	4108
ttt tta tat gta Phe Leu Tyr Val 1365	aat acc tca cag agg Asn Thr Ser Gln Arg 1370	atg tta cgt att cta gac Met Leu Arg Ile Leu Asp 1375	4153
acg tta aat tcc Thr Leu Asn Ser 1380	aca ttt tcc tct gag Thr Phe Ser Ser Glu 1385	aac aca att agc agt ctg Asn Thr Ile Ser Ser Leu 1390	4198
aaa gga tgc att Lys Gly Cys Ile 1395	gta tgg tta gat gtc Val Trp Leu Asp Val 1400	ata aac cat ttg tat ttg Ile Asn His Leu Tyr Leu 1405	4243
ttg tct aac tcc Leu Ser Asn Ser 1410	agt ttt tca caa ggt Ser Phe Ser Gln Gly 1415	cgt ctt caa aat att ttg Arg Leu Gln Asn Ile Leu 1420	4288
ggg aat ttc aga Gly Asn Phe Arg 1425	gat ata gaa aac aaa Asp Ile Glu Asn Lys 1430	atg aac tct ata tta aaa Met Asn Ser Ile Leu Lys 1435	4333
att gta act tgg Ile Val Thr Trp 1440	gtg tta aat ata aaa Val Leu Asn Ile Lys 1445	aaa cct ctt tgt tca tca Lys Pro Leu Cys Ser Ser 1450	4378
aat ggc tca cat Asn Gly Ser His 1455	ata aat tgt gtc aat Ile Asn Cys Val Asn 1460	att tac ttg aaa gat gta Ile Tyr Leu Lys Asp Val 1465	4423
act gac ttt cta Thr Asp Phe Leu 1470	aat att gta ctt act Asn Ile Val Leu Thr 1475	aca gtc ttt gaa aaa gag Thr Val Phe Glu Lys Glu 1480	4468
aag aaa cct aaa Lys Lys Pro Lys 1485	ttt gag att tta tta Phe Glu Ile Leu Leu 1490	gct ctt tta aat gat tcc Ala Leu Leu Asn Asp Ser 1495	4513
aca aag caa gta Thr Lys Gln Val 1500	agg atg agt atc aac Arg Met Ser Ile Asn 1505	aac tta aca aca gac ttt Asn Leu Thr Thr Asp Phe 1510	4558
gat ttt gca tct Asp Phe Ala Ser 1515	cag tcc aat tgg aga Gln Ser Asn Trp Arg 1520	tat ttt act gaa tta att Tyr Phe Thr Glu Leu Ile 1525	4603
cta aga cca ata Leu Arg Pro Ile 1530	gaa atg tca gat gaa Glu Met Ser Asp Glu 1535	att cct aat cag ttt caa Ile Pro Asn Gln Phe Gln 1540	4648
aat att tgg ctt Asn Ile Trp Leu 1545	cat tta ata aca ctg His Leu Ile Thr Leu 1550	ggg aag gaa ttt cag aag Gly Lys Glu Phe Gln Lys 1555	4693
ctt gta aaa ggt Leu Val Lys Gly 1560	att tat ttt aac atc Ile Tyr Phe Asn Ile 1565	ctg gaa aat aat tcc tct Leu Glu Asn Asn Ser Ser 1570	4738

tct	aaa	act	gaa	aac	ttg	tta	aac	ata	ttt	gcc	acc	agt	cca	aaa	4783
Ser	Lys	Thr	Glu	Asn	Leu	Leu	Asn	Ile	Phe	Ala	Thr	Ser	Pro	Lys	
			1575					1580					1585		
gaa	aag	gat	gta	aac	agt	gta	ggc	aat	tcc	att	tat	cac	tta	gct	4828
Glu	Lys	Asp	Val	Asn	Ser	Val	Gly	Asn	Ser	Ile	Tyr	His	Leu	Ala	
			1590					1595					1600		
agt	tac	ctt	gcc	ttc	agc	tta	tct	cat	gac	ctc	caa	aat	tca	cca	4873
Ser	Tyr	Leu	Ala	Phe	Ser	Leu	Ser	His	Asp	Leu	Gln	Asn	Ser	Pro	
			1605					1610					1615		
aaa	ata	ata	att	tca	cct	gaa	ata	atg	aaa	gct	aca	ggc	ctt	ggc	4918
Lys	Ile	Ile	Ile	Ser	Pro	Glu	Ile	Met	Lys	Ala	Thr	Gly	Leu	Gly	
			1620					1625					1630		
att	caa	ctg	ata	agg	gat	gtg	ttc	aac	tcc	tta	atg	cct	gta	gtt	4963
Ile	Gln	Leu	Ile	Arg	Asp	Val	Phe	Asn	Ser	Leu	Met	Pro	Val	Val	
			1635					1640					1645		
cat	cac	act	agt	cca	caa	aat	gca	ggc	tat	atg	caa	gct	ttg	aag	5008
His	His	Thr	Ser	Pro	Gln	Asn	Ala	Gly	Tyr	Met	Gln	Ala	Leu	Lys	
			1650					1655					1660		
aag	gta	act	tct	gtc	atg	cgt	acc	ctt	aag	aaa	gca	gac	ata	gac	5053
Lys	Val	Thr	Ser	Val	Met	Arg	Thr	Leu	Lys	Lys	Ala	Asp	Ile	Asp	
			1665					1670					1675		
ctt	tta	gtg	gat	cag	ctt	gaa	caa	gtt	agt	gta	aac	cta	atg	gat	5098
Leu	Leu	Val	Asp	Gln	Leu	Glu	Gln	Val	Ser	Val	Asn	Leu	Met	Asp	
			1680					1685					1690		
ttc	ttt	aag	aat	atc	agt	agt	gtg	gga	act	ggc	aat	tta	gtg	gtc	5143
Phe	Phe	Lys	Asn	Ile	Ser	Ser	Val	Gly	Thr	Gly	Asn	Leu	Val	Val	
			1695					1700					1705		
aat	ttg	ctt	gtt	ggc	ttg	atg	gaa	aaa	ttt	gca	gac	agc	tca	cat	5188
Asn	Leu	Leu	Val	Gly	Leu	Met	Glu	Lys	Phe	Ala	Asp	Ser	Ser	His	
			1710					1715					1720		
tct	tgg	aat	gtt	aat	cat	ctg	ctg	cag	ctc	tca	cgc	ctg	ttt	cct	5233
Ser	Trp	Asn	Val	Asn	His	Leu	Leu	Gln	Leu	Ser	Arg	Leu	Phe	Pro	
			1725					1730					1735		
aaa	gat	gtt	gtg	gat	gct	gtg	ata	gat	gtg	tac	tat	gtg	ctt	cct	5278
Lys	Asp	Val	Val	Asp	Ala	Val	Ile	Asp	Val	Tyr	Tyr	Val	Leu	Pro	
			1740					1745					1750		
cat	gct	gta	agg	ctc	ctg	cag	gga	gta	cct	ggc	aaa	aac	atc	act	5323
His	Ala	Val	Arg	Leu	Leu	Gln	Gly	Val	Pro	Gly	Lys	Asn	Ile	Thr	
			1755					1760					1765		
gaa	ggc	ctc	aag	gat	gtc	tac	agc	ttc	aca	ctc	ctt	cat	ggc	ata	5368
Glu	Gly	Leu	Lys	Asp	Val	Tyr	Ser	Phe	Thr	Leu	Leu	His	Gly	Ile	
			1770					1775					1780		
acc	att	tca	aat	atc	acc	aag	gaa	gac	ttc	gca	att	gtg	ata	aaa	5413
Thr	Ile	Ser	Asn	Ile	Thr	Lys	Glu	Asp	Phe	Ala	Ile	Val	Ile	Lys	
			1785					1790					1795		

att ctt ttg gat	aca att gaa tta gta	tca gat aag cca gat	att	5458
Ile Leu Leu Asp	Thr Ile Glu Leu Val	Ser Asp Lys Pro Asp	Ile	
1800	1805	1810		
att tca gag gct	tta gct tgt ttt cct	gtg gtt tgg tgc tgg	aat	5503
Ile Ser Glu Ala	Leu Ala Cys Phe Pro	Val Val Trp Cys Trp	Asn	
1815	1820	1825		
cac aca aat tct	gga ttt cgg cag aat	tca aag ata gac ccc	tgc	5548
His Thr Asn Ser	Gly Phe Arg Gln Asn	Ser Lys Ile Asp Pro	Cys	
1830	1835	1840		
aat gtc cat ggg	ctc atg tct tct tcc	ttt tat ggc aaa gtg	gcc	5593
Asn Val His Gly	Leu Met Ser Ser Ser	Phe Tyr Gly Lys Val	Ala	
1845	1850	1855		
agt ata ctt gat	cat ttc cac ctg tct	ccc caa ggt gaa gat	tca	5638
Ser Ile Leu Asp	His Phe His Leu Ser	Pro Gln Gly Glu Asp	Ser	
1860	1865	1870		
cca tgt tca aat	gaa agc tcc cga atg	gaa ata act agg aaa	gtg	5683
Pro Cys Ser Asn	Glu Ser Ser Arg Met	Glu Ile Thr Arg Lys	Val	
1875	1880	1885		
gtc tgc ata att	cat gaa tta gtg gac	tgg aat tct att ctt	ctg	5728
Val Cys Ile Ile	His Glu Leu Val Asp	Trp Asn Ser Ile Leu	Leu	
1890	1895	1900		
gag ctc tct gaa	gtc ttc cat gtt aac	att tct ctt gtg aaa	act	5773
Glu Leu Ser Glu	Val Phe His Val Asn	Ile Ser Leu Val Lys	Thr	
1905	1910	1915		
gtg cag aaa ttt	tgg cat aag ata tta	ccg ttt gtc cca cct	tca	5818
Val Gln Lys Phe	Trp His Lys Ile Leu	Pro Phe Val Pro Pro	Ser	
1920	1925	1930		
ata aat caa act	agg gat agc atc tct	gaa ctc tgt cct agt	ggc	5863
Ile Asn Gln Thr	Arg Asp Ser Ile Ser	Glu Leu Cys Pro Ser	Gly	
1935	1940	1945		
tcc ata aag caa	gtt gct ttg caa atc	ata gaa aaa ctt aaa	aat	5908
Ser Ile Lys Gln	Val Ala Leu Gln Ile	Ile Glu Lys Leu Lys	Asn	
1950	1955	1960		
gtc aac ttt aca	aaa gtt aca tca ggt	gaa aat att ctt gac	aaa	5953
Val Asn Phe Thr	Lys Val Thr Ser Gly	Glu Asn Ile Leu Asp	Lys	
1965	1970	1975		
cta agt agt tta	aac aag atc ctt aac	att aat gaa gac aca	gag	5998
Leu Ser Ser Leu	Asn Lys Ile Leu Asn	Ile Asn Glu Asp Thr	Glu	
1980	1985	1990		
aca tct gtt caa	aat att att tcc tca	aat ttg gaa agg aca	gta	6043
Thr Ser Val Gln	Asn Ile Ile Ser Ser	Asn Leu Glu Arg Thr	Val	
1995	2000	2005		
caa ttg att tct	gaa gac tgg agc cta	gaa aaa agt acg cat	aat	6088
Gln Leu Ile Ser	Glu Asp Trp Ser Leu	Glu Lys Ser Thr His	Asn	
2010	2015	2020		
cta ctc tct tta	ttc atg atg ctc cag	aat gca aat gtc aca	ggc	6133

Leu	Leu	Ser	Leu	Phe	Met	Met	Leu	Gln	Asn	Ala	Asn	Val	Thr	Gly	
			2025					2030					2035		
agc	agt	tta	gaa	gca	tta	tca	agt	ttt	att	gaa	aaa	agt	gaa	aca	6178
Ser	Ser	Leu	Glu	Ala	Leu	Ser	Ser	Phe	Ile	Glu	Lys	Ser	Glu	Thr	
			2040					2045					2050		
cct	tac	aac	ttt	gaa	gaa	cta	tgg	ccc	aag	ttt	caa	caa	atc	atg	6223
Pro	Tyr	Asn	Phe	Glu	Glu	Leu	Trp	Pro	Lys	Phe	Gln	Gln	Ile	Met	
			2055					2060					2065		
aaa	gac	cta	acc	caa	gat	ttt	aga	atc	aga	cac	ctg	ctt	tct	gaa	6268
Lys	Asp	Leu	Thr	Gln	Asp	Phe	Arg	Ile	Arg	His	Leu	Leu	Ser	Glu	
			2070					2075					2080		
atg	aac	aaa	gga	atc	aaa	agt	ata	aat	tca	atg	gct	ctt	caa	aag	6313
Met	Asn	Lys	Gly	Ile	Lys	Ser	Ile	Asn	Ser	Met	Ala	Leu	Gln	Lys	
			2085					2090					2095		
ata	act	ttg	cag	ttt	gcc	cat	ttc	ctg	gaa	atc	ctg	gat	tca	ccg	6358
Ile	Thr	Leu	Gln	Phe	Ala	His	Phe	Leu	Glu	Ile	Leu	Asp	Ser	Pro	
			2100					2105					2110		
tca	ttg	aag	aca	tta	gaa	att	att	gaa	gat	ttt	cta	ttg	gtc	aca	6403
Ser	Leu	Lys	Thr	Leu	Glu	Ile	Ile	Glu	Asp	Phe	Leu	Leu	Val	Thr	
			2115					2120					2125		
aaa	aac	tgg	ctt	cag	gaa	tat	gca	aat	gag	gat	tac	tcc	aga	atg	6448
Lys	Asn	Trp	Leu	Gln	Glu	Tyr	Ala	Asn	Glu	Asp	Tyr	Ser	Arg	Met	
			2130					2135					2140		
ata	gaa	aca	tta	ttc	att	cct	gtg	acc	aat	gag	agt	tca	act	gaa	6493
Ile	Glu	Thr	Leu	Phe	Ile	Pro	Val	Thr	Asn	Glu	Ser	Ser	Thr	Glu	
			2145					2150					2155		
gat	ata	gct	ttg	tta	gcc	aaa	gct	att	gct	act	ttt	tgg	ggc	tct	6538
Asp	Ile	Ala	Leu	Leu	Ala	Lys	Ala	Ile	Ala	Thr	Phe	Trp	Gly	Ser	
			2160					2165					2170		
tta	aaa	aat	ata	tct	aga	gca	ggc	aat	ttt	gat	gtt	gcc	ttt	ctt	6583
Leu	Lys	Asn	Ile	Ser	Arg	Ala	Gly	Asn	Phe	Asp	Val	Ala	Phe	Leu	
			2175					2180					2185		
acc	cat	ctg	cta	aat	caa	gaa	cag	ctg	act	aat	ttc	tca	gtt	gtt	6628
Thr	His	Leu	Leu	Asn	Gln	Glu	Gln	Leu	Thr	Asn	Phe	Ser	Val	Val	
			2190					2195					2200		
cag	ctg	ctt	ttt	gaa	aac	atc	cta	att	aat	ttg	atc	aat	aac	tta	6673
Gln	Leu	Leu	Phe	Glu	Asn	Ile	Leu	Ile	Asn	Leu	Ile	Asn	Asn	Leu	
			2205					2210					2215		
gct	ggg	aat	tct	cag	gaa	gca	gct	tgg	aac	tta	aat	gat	act	gac	6718
Ala	Gly	Asn	Ser	Gln	Glu	Ala	Ala	Trp	Asn	Leu	Asn	Asp	Thr	Asp	
			2220					2225					2230		
ctt	caa	ata	atg	aat	ttc	att	aac	ctt	atc	ttg	aac	cat	atg	cag	6763
Leu	Gln	Ile	Met	Asn	Phe	Ile	Asn	Leu	Ile	Leu	Asn	His	Met	Gln	
			2235					2240					2245		
tca	gaa	act	agt	agg	aaa	aca	gtt	ctc	tct	ctg	aga	agc	ata	gta	6808
Ser	Glu	Thr	Ser	Arg	Lys	Thr	Val	Leu	Ser	Leu	Arg	Ser	Ile	Val	

2250	2255	2260	
gat ttc aca gaa	cag ttt ttg aaa	aca ttc ttc tcc ctt ttt cta	6853
Asp Phe Thr Glu	Gln Phe Leu Lys Thr	Phe Phe Ser Leu Phe Leu	
2265	2270	2275	
aag gaa gat tct	gag aac aaa ata tct	ctt ctg ctg aaa tat ttc	6898
Lys Glu Asp Ser	Glu Asn Lys Ile Ser	Leu Leu Leu Lys Tyr Phe	
2280	2285	2290	
cac aaa gat gtt	att gca gag atg agt	ttt gtc cca aaa gat aaa	6943
His Lys Asp Val	Ile Ala Glu Met Ser	Phe Val Pro Lys Asp Lys	
2295	2300	2305	
att cta gaa att	ctg aaa ctg gat caa	ttt ctt acc ctg atg ata	6988
Ile Leu Glu Ile	Leu Lys Leu Asp Gln	Phe Leu Thr Leu Met Ile	
2310	2315	2320	
caa gac aga ttg	atg aac att ttt tca	agt tta aag gag act ata	7033
Gln Asp Arg Leu	Met Asn Ile Phe Ser	Ser Leu Lys Glu Thr Ile	
2325	2330	2335	
tat cac cta atg	aaa agt tca ttt ata	tta gac aat gga gaa ttt	7078
Tyr His Leu Met	Lys Ser Ser Phe Ile	Leu Asp Asn Gly Glu Phe	
2340	2345	2350	
tat ttt gat act	cat caa gga ctg aag	ttc atg caa gat tta ttt	7123
Tyr Phe Asp Thr	His Gln Gly Leu Lys	Phe Met Gln Asp Leu Phe	
2355	2360	2365	
aat gcc ctt ctc	agg gaa act tca atg	aaa aat aag act gaa aat	7168
Asn Ala Leu Leu	Arg Glu Thr Ser Met	Lys Asn Lys Thr Glu Asn	
2370	2375	2380	
aat ata gac ttt	ttc aca gtg gtg agt	cag ttg ttt ttc cat gtg	7213
Asn Ile Asp Phe	Phe Thr Val Val Ser	Gln Leu Phe Phe His Val	
2385	2390	2395	
aat aag tct gag	gac ctc ttc aaa ctc	aat caa gat ctt ggg tca	7258
Asn Lys Ser Glu	Asp Leu Phe Lys Leu	Asn Gln Asp Leu Gly Ser	
2400	2405	2410	
gct ctt cac ctt	gta aga gaa tgt tca	aca gag atg gca aga ctt	7303
Ala Leu His Leu	Val Arg Glu Cys Ser	Thr Glu Met Ala Arg Leu	
2415	2420	2425	
ctg gat aca att	tta cac tct cct aat	aag gac ttc tat gct ttg	7348
Leu Asp Thr Ile	Leu His Ser Pro Asn	Lys Asp Phe Tyr Ala Leu	
2430	2435	2440	
tat cct acc ctc	caa gaa gtt ata ctt	gct aat cta acg gat ttg	7393
Tyr Pro Thr Leu	Gln Glu Val Ile Leu	Ala Asn Leu Thr Asp Leu	
2445	2450	2455	
ctt ttc ttt ata	aat aat tca ttc cct	cta aga aac aga gca aca	7438
Leu Phe Phe Ile	Asn Asn Ser Phe Pro	Leu Arg Asn Arg Ala Thr	
2460	2465	2470	
tta gaa att act	aag aga tta gtt ggt	gct att tca aga gca agt	7483
Leu Glu Ile Thr	Lys Arg Leu Val Gly	Ala Ile Ser Arg Ala Ser	
2475	2480	2485	

gaa gaa agt cac gtc ctg aaa ccc ctc tta gaa atg tct ggg act	7528
Glu Glu Ser His Val Leu Lys Pro Leu Leu Glu Met Ser Gly Thr	
2490 2495 2500	
ctg gtc atg ctg ttg aat gac agt gct gac ctg aga gat ctt gcc	7573
Leu Val Met Leu Leu Asn Asp Ser Ala Asp Leu Arg Asp Leu Ala	
2505 2510 2515	
aca tca atg gac tcc att gtg aaa ctt ctt aag ctg gtc aag aaa	7618
Thr Ser Met Asp Ser Ile Val Lys Leu Leu Lys Leu Val Lys Lys	
2520 2525 2530	
gtt tgc ggg aag atg tcc aca gtt ttt aaa act cat ttt atc tcc	7663
Val Ser Gly Lys Met Ser Thr Val Phe Lys Thr His Phe Ile Ser	
2535 2540 2545	
aat acc aag gac agt gtg aaa ttc ttt gac act ctg tat tcc atc	7708
Asn Thr Lys Asp Ser Val Lys Phe Phe Asp Thr Leu Tyr Ser Ile	
2550 2555 2560	
atg caa caa agt gtt caa aat ctt gtg aaa gaa ata gct act tta	7753
Met Gln Gln Ser Val Gln Asn Leu Val Lys Glu Ile Ala Thr Leu	
2565 2570 2575	
aaa aaa ata gat cat ttc aca ttt gaa aag ata aat gat ttg ttg	7798
Lys Lys Ile Asp His Phe Thr Phe Glu Lys Ile Asn Asp Leu Leu	
2580 2585 2590	
gtg cca ttt ctt gac ttg gcc ttt gaa atg att ggg gta gaa cct	7843
Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile Gly Val Glu Pro	
2595 2600 2605	
tat ata tca tca aac tct gat att ttc agt atg tca cct agc ata	7888
Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser Pro Ser Ile	
2610 2615 2620	
ctc tca tat atg aac caa tct aag gac ttt tct gat att ttg gaa	7933
Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile Leu Glu	
2625 2630 2635	
gaa att gct gaa ttt tta aca tct gtg aaa atg aac ttg gaa gat	7978
Glu Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu Asp	
2640 2645 2650	
atg agg agt ctt gcg gta gca ttt aac aat gag act caa aca ttt	8023
Met Arg Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe	
2655 2660 2665	
tct atg gat tct gtc aac tta ngg gaa gaa att ctg ggt tgc tta	8068
Ser Met Asp Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu	
2670 2675 2680	
gtt cct ata aat aac atc acc aac caa atg gac ttc tta tac cct	8113
Val Pro Ile Asn Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro	
2685 2690 2695	
aat cca att tcc act cat agt ggc cct caa gat ata aaa tgg gaa	8158
Asn Pro Ile Ser Thr His Ser Gly Pro Gln Asp Ile Lys Trp Glu	
2700 2705 2710	

ata att cat gaa gtg atc ctt ttt ttg gat aaa ata tta tca caa	8203
Ile Ile His Glu Val Ile Leu Phe Leu Asp Lys Ile Leu Ser Gln	
2715 2720 2725	
aac agc aca gaa ata gga tct ttc ttg aaa atg gtg atc tgt ctc	8248
Asn Ser Thr Glu Ile Gly Ser Phe Leu Lys Met Val Ile Cys Leu	
2730 2735 2740	
acc tta gaa gct ctt tgg aaa aac tta aag aaa gat aat tgg aat	8293
Thr Leu Glu Ala Leu Trp Lys Asn Leu Lys Lys Asp Asn Trp Asn	
2745 2750 2755	
gtt tct aat gtg ttg atg acg ttt act cag cat cca aat aac ctt	8338
Val Ser Asn Val Leu Met Thr Phe Thr Gln His Pro Asn Asn Leu	
2760 2765 2770	
ttg aaa acc ata gaa aca gtt tta gag gcc tcc agt gga att aaa	8383
Leu Lys Thr Ile Glu Thr Val Leu Glu Ala Ser Ser Gly Ile Lys	
2775 2780 2785	
agt gac tat gaa ggt gat ttg aat aaa agt tta tat ttt gac aca	8428
Ser Asp Tyr Glu Gly Asp Leu Asn Lys Ser Leu Tyr Phe Asp Thr	
2790 2795 2800	
cct ttg agt cag aat ata act cat cat caa ctt gaa aaa gca atc	8473
Pro Leu Ser Gln Asn Ile Thr His His Gln Leu Glu Lys Ala Ile	
2805 2810 2815	
cat aat gtt tta agt aga ata gct ctc tgg agg aaa gga ctt cgt	8518
His Asn Val Leu Ser Arg Ile Ala Leu Trp Arg Lys Gly Leu Arg	
2820 2825 2830	
ttt aac aac tct gaa tgg ata act tcc aca aga act ttg ttt cag	8563
Phe Asn Asn Ser Glu Trp Ile Thr Ser Thr Arg Thr Leu Phe Gln	
2835 2840 2845	
cca ctt ttt gag att ttc att aaa gca acc acc gga aag aat gtc	8608
Pro Leu Phe Glu Ile Phe Ile Lys Ala Thr Thr Gly Lys Asn Val	
2850 2855 2860	
aca tca gaa aaa gaa gag aga acc gag aaa gag atg att gac ttt	8653
Thr Ser Glu Lys Glu Glu Arg Thr Glu Lys Glu Met Ile Asp Phe	
2865 2870 2875	
cct tat agt ttc aaa cca ttt ttc tgt ttg gag aaa tac ctg gga	8698
Pro Tyr Ser Phe Lys Pro Phe Phe Cys Leu Glu Lys Tyr Leu Gly	
2880 2885 2890	
gga tta ttt gta ttg act aaa tac tgg caa caa atc cca cta aca	8743
Gly Leu Phe Val Leu Thr Lys Tyr Trp Gln Gln Ile Pro Leu Thr	
2895 2900 2905	
gat caa agt gtt gtt gag att tgt gaa gtt ttc cag cag act gtg	8788
Asp Gln Ser Val Val Glu Ile Cys Glu Val Phe Gln Gln Thr Val	
2910 2915 2920	
aag ccc tca gaa gcc atg gag atg ctg cag aaa gtg aag atg atg	8833
Lys Pro Ser Glu Ala Met Glu Met Leu Gln Lys Val Lys Met Met	
2925 2930 2935	
gtc gta cgt gtg ctc acc atc gtt gca gaa aac cct tcc tgg acc	8878

Val	Val	Arg	Val	Leu	Thr	Ile	Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	
			2940					2945					2950		
aag	gac	att	ttg	tgt	gct	act	ctg	agt	tgc	aag	caa	aat	ggg	ata	8923
Lys	Asp	Ile	Leu	Cys	Ala	Thr	Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	
			2955					2960					2965		
agg	cat	ctc	att	tta	tct	gct	ata	caa	ggg	gtc	act	ttg	gcg	cag	8968
Arg	His	Leu	Ile	Leu	Ser	Ala	Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	
			2970					2975					2980		
gac	cac	ttc	cag	gaa	att	gaa	aag	ata	tgg	tcc	tcg	ccg	aat	cag	9013
Asp	His	Phe	Gln	Glu	Ile	Glu	Lys	Ile	Trp	Ser	Ser	Pro	Asn	Gln	
			2985					2990					2995		
cta	aat	tgt	gaa	agt	ctt	agc	aag	aat	ctt	tct	agc	acc	ttg	gag	9058
Leu	Asn	Cys	Glu	Ser	Leu	Ser	Lys	Asn	Leu	Ser	Ser	Thr	Leu	Glu	
			3000					3005					3010		
agc	ttc	aag	agc	agc	ttg	gaa	aat	gcc	act	ggc	cag	gac	tgc	aca	9103
Ser	Phe	Lys	Ser	Ser	Leu	Glu	Asn	Ala	Thr	Gly	Gln	Asp	Cys	Thr	
			3015					3020					3025		
agc	cag	ccg	agg	ctg	gag	acg	gtg	cag	cag	cac	ttg	tac	atg	ttg	9148
Ser	Gln	Pro	Arg	Leu	Glu	Thr	Val	Gln	Gln	His	Leu	Tyr	Met	Leu	
			3030					3035					3040		
gcc	aaa	agc	ctc	gag	gaa	act	tgg	tca	tca	ggg	aat	ccc	atc	atg	9193
Ala	Lys	Ser	Leu	Glu	Glu	Thr	Trp	Ser	Ser	Gly	Ash	Pro	Ile	Met	
			3045					3050					3055		
act	ttt	ctc	agc	aat	ttc	aca	gta	act	gag	gat	gta	aaa	ata	aaa	9238
Thr	Phe	Leu	Ser	Asn	Phe	Thr	Val	Thr	Glu	Asp	Val	Lys	Ile	Lys	
			3060					3065					3070		
gat	ttg	atg	aag	aat	atc	acc	aag	ttg	act	gag	gag	ctt	cgc	tct	9283
Asp	Leu	Met	Lys	Asn	Ile	Thr	Lys	Leu	Thr	Glu	Glu	Leu	Arg	Ser	
			3075					3080					3085		
tcc	atc	caa	atc	tcg	aat	gag	act	atc	cat	agc	att	cta	gaa	gca	9328
Ser	Ile	Gln	Ile	Ser	Asn	Glu	Thr	Ile	His	Ser	Ile	Leu	Glu	Ala	
			3090					3095					3100		
aat	att	tcc	cac	tcc	aag	gtt	ctc	ttc	agt	gcc	ctc	acc	gta	gct	9373
Asn	Ile	Ser	His	Ser	Lys	Val	Leu	Phe	Ser	Ala	Leu	Thr	Val	Ala	
			3105					3110					3115		
ctg	tct	gga	aag	tgt	gat	cag	gaa	atc	ctt	cat	ctc	ctg	ctg	aca	9418
Leu	Ser	Gly	Lys	Cys	Asp	Gln	Glu	Ile	Leu	His	Leu	Leu	Leu	Thr	
			3120					3125					3130		
ttt	ccc	aaa	ggg	gaa	aaa	tct	tgg	atc	gca	gcg	gag	gaa	ctc	tgt	9463
Phe	Pro	Lys	Gly	Glu	Lys	Ser	Trp	Ile	Ala	Ala	Glu	Glu	Leu	Cys	
			3135					3140					3145		
agc	ctg	cca	ggg	tca	aaa	gtg	tat	tct	ctg	att	gtg	ttg	ctg	agt	9508
Ser	Leu	Pro	Gly	Ser	Lys	Val	Tyr	Ser	Leu	Ile	Val	Leu	Leu	Ser	
			3150					3155					3160		
cga	aac	ttg	gat	gtg	cga	gct	ttc	att	tac	aag	act	ctg	atg	cct	9553
Arg	Asn	Leu	Asp	Val	Arg	Ala	Phe	Ile	Tyr	Lys	Thr	Leu	Met	Pro	

3165	3170	3175	
tct gaa gca aat	ggc ttg ctc aac tcc	ttg ctg gat ata gtt tcc	9598
Ser Glu Ala Asn	Gly Leu Leu Asn Ser	Leu Leu Asp Ile Val Ser	
3180	3185	3190	
agc ctc agc gcc	ttg ctt gcc aaa gcc	cag cac gtc ttt gag tat	9643
Ser Leu Ser Ala	Leu Leu Ala Lys Ala	Gln His Val Phe Glu Tyr	
3195	3200	3205	
ctt cct gag ttt	ctt cac aca ttt aaa	atc act gcc ttg cta gaa	9688
Leu Pro Glu Phe	Leu His Thr Phe Lys	Ile Thr Ala Leu Leu Glu	
3210	3215	3220	
acc ctg gac ttt	caa cag gtt tca caa	aat gtc cag gcc aga agt	9733
Thr Leu Asp Phe	Gln Gln Val Ser Gln	Asn Val Gln Ala Arg Ser	
3225	3230	3235	
tca gct ttt ggt	tct ttc cag ttt gtg	atg aag atg gtt tgc aag	9778
Ser Ala Phe Gly	Ser Phe Gln Phe Val	Met Lys Met Val Cys Lys	
3240	3245	3250	
gac caa gca tca	ttc ctt agc gat tct	aat atg ttt att aat ttg	9823
Asp Gln Ala Ser	Phe Leu Ser Asp Ser	Asn Met Phe Ile Asn Leu	
3255	3260	3265	
ccc aga gtt aag	gaa ctc ttg gaa gat	gac aaa gaa aaa ttc aac	9868
Pro Arg Val Lys	Glu Leu Leu Glu Asp	Asp Lys Glu Lys Phe Asn	
3270	3275	3280	
att cct gaa gat	tca aca ccg ttt tgc	ttg aag ctt tat cag gaa	9913
Ile Pro Glu Asp	Ser Thr Pro Phe Cys	Leu Lys Leu Tyr Gln Glu	
3285	3290	3295	
att cta caa ttg	cca aat ggt gct ttg	gtg tgg acc ttc cta aaa	9958
Ile Leu Gln Leu	Pro Asn Gly Ala Leu	Val Trp Thr Phe Leu Lys	
3300	3305	3310	
ccc ata ttg cat	gga aaa ata cta tac	aca cca aac act cca gaa	10003
Pro Ile Leu His	Gly Lys Ile Leu Tyr	Thr Pro Asn Thr Pro Glu	
3315	3320	3325	
att aac aag gtc	att caa aag gct aat	tac acc ttt tat att gtg	10048
Ile Asn Lys Val	Ile Gln Lys Ala Asn	Tyr Thr Phe Tyr Ile Val	
3330	3335	3340	
gac aaa cta aaa	act tta tca gaa aca	ctg ctg gaa atg tcc agc	10093
Asp Lys Leu Lys	Thr Leu Ser Glu Thr	Leu Leu Glu Met Ser Ser	
3345	3350	3355	
ctt ttc cag aga	agt gga agt ggc cag	atg ttc aac cag ctg cag	10138
Leu Phe Gln Arg	Ser Gly Ser Gly Gln	Met Phe Asn Gln Leu Gln	
3360	3365	3370	
gag gcc ctg aga	aac aaa ttt gta aga	aac ttt gta gaa aac cag	10183
Glu Ala Leu Arg	Asn Lys Phe Val Arg	Asn Phe Val Glu Asn Gln	
3375	3380	3385	
ttg cac att gat	gta gac aaa ctt act	gaa aaa ctc cag aca tac	10228
Leu His Ile Asp	Val Asp Lys Leu Thr	Glu Lys Leu Gln Thr Tyr	
3390	3395	3400	

gga ggg ctg ctg	gat gag atg ttt aac	cat gca ggc gct gga cgc	10273
Gly Gly Leu Leu	Asp Glu Met Phe Asn	His Ala Gly Ala Gly Arg	
3405	3410	3415	
ttc cgt ttc ttg	ggc agc atc ttg gtc	aat ctc tct tcc tgc gtg	10318
Phe Arg Phe Leu	Gly Ser Ile Leu Val	Asn Leu Ser Ser Cys Val	
3420	3425	3430	
gca ctg aac cgt	ttc cag gct ctg cag	tct gtc gac atc ctg gag	10363
Ala Leu Asn Arg	Phe Gln Ala Leu Gln	Ser Val Asp Ile Leu Glu	
3435	3440	3445	
act aaa gca cat	gaa ctc ttg cag cag	aac agc ttc ttg gcc agt	10408
Thr Lys Ala His	Glu Leu Leu Gln Gln	Asn Ser Phe Leu Ala Ser	
3450	3455	3460	
atc att ttc agc	aat tcc tta ttc gac	aag aac ttc aga tca gag	10453
Ile Ile Phe Ser	Asn Ser Leu Phe Asp	Lys Asn Phe Arg Ser Glu	
3465	3470	3475	
tct gtc aaa ctg	cca ccc cat gtc tca	tac aca atc cgg acc aat	10498
Ser Val Lys Leu	Pro Pro His Val Ser	Tyr Thr Ile Arg Thr Asn	
3480	3485	3490	
gtg tta tac agc	gtg cga aca gat gtg	gta aaa aac cct tct tgg	10543
Val Leu Tyr Ser	Val Arg Thr Asp Val	Val Lys Asn Pro Ser Trp	
3495	3500	3505	
aag ttc cac cct	cag aat cta cca gct	gat ggg ttc aaa tat aac	10588
Lys Phe His Pro	Gln Asn Leu Pro Ala	Asp Gly Phe Lys Tyr Asn	
3510	3515	3520	
tac gtc ttt gcc	cca ctg caa gac atg	atc gaa aga gcc atc att	10633
Tyr Val Phe Ala	Pro Leu Gln Asp Met	Ile Glu Arg Ala Ile Ile	
3525	3530	3535	
ttg gtg cag act	ggg cag gaa gcc ctg	gaa cca gca gca cag act	10678
Leu Val Gln Thr	Gly Gln Glu Ala Leu	Glu Pro Ala Ala Gln Thr	
3540	3545	3550	
cag gcg gcc cct	tac ccc tgc cat acc	agc gac cta ttc ctg aac	10723
Gln Ala Ala Pro	Tyr Pro Cys His Thr	Ser Asp Leu Phe Leu Asn	
3555	3560	3565	
aac gtt ggt ttc	ttt ttt cca ctg ata	atg atg ctg acg tgg atg	10768
Asn Val Gly Phe	Phe Phe Pro Leu Ile	Met Met Leu Thr Trp Met	
3570	3575	3580	
gtg tct gtg gcc	agc atg gtc aga aag	ttg gtg tat gag cag gag	10813
Val Ser Val Ala	Ser Met Val Arg Lys	Leu Val Tyr Glu Gln Glu	
3585	3590	3595	
ata cag ata gaa	gag tat atg cgg atg	atg gga gtg cat cca gtg	10858
Ile Gln Ile Glu	Glu Tyr Met Arg Met	Met Gly Val His Pro Val	
3600	3605	3610	
atc cat ttc ctg	gcc tgg ttc ctg gag	aac atg gct gtg ttg acc	10903
Ile His Phe Leu	Ala Trp Phe Leu Glu	Asn Met Ala Val Leu Thr	
3615	3620	3625	

ata Ile	agc Ser	agt Ser	gct Ala 3630	act Thr	ctg Leu	gcc Ala	atc Ile	gtt Val 3635	ctg Leu	aaa Lys	aca Thr	agt Ser	ggc Gly 3640	atc Ile	10948
ttt Phe	gca Ala	cac His	agc Ser 3645	aat Asn	acc Thr	ttt Phe	att Ile	gtt Val 3650	ttc Phe	ctc Leu	ttt Phe	ctc Leu	ttg Leu 3655	gat Asp	10993
ttt Phe	ggg Gly	atg Met	tca Ser 3660	gtc Val	gtc Val	atg Met	ctg Leu	agc Ser 3665	tac Tyr	ctc Leu	ttg Leu	agt Ser	gca Ala 3670	ttt Phe	11038
ttc Phe	agc Ser	caa Gln	gct Ala 3675	aat Asn	aca Thr	gcg Ala	gcc Ala	ctt Leu 3680	tgt Cys	acc Thr	agc Ser	ctg Leu	gtg Val 3685	tac Tyr	11083
atg Met	atc Ile	agc Ser	ttt Phe 3690	ctg Leu	ccc Pro	tac Tyr	ata Ile	gtt Val 3695	cta Leu	ttg Leu	gtt Val	cta Leu	cat His 3700	aac Asn	11128
caa Gln	tta Leu	agt Ser	ttt Phe 3705	gtt Val	aat Asn	cag Gln	aca Thr	ttt Phe 3710	ctg Leu	tgc Cys	ctt Leu	ctt Leu	tcg Ser 3715	aca Thr	11173
acc Thr	gcc Ala	ttt Phe	gga Gly 3720	caa Gln	ggg Gly	gta Val	ttt Phe	ttt Phe 3725	att Ile	aca Thr	ttc Phe	ctg Leu	gaa Glu 3730	gga Gly	11218
caa Gln	gag Glu	aca Thr	ggg Gly 3735	att Ile	caa Gln	tgg Trp	aat Asn	aat Asn 3740	atg Met	tac Tyr	cag Gln	gct Ala	ctg Leu 3745	gaa Glu	11263
caa Gln	ggg Gly	ggc Gly	atg Met 3750	aca Thr	ttt Phe	ggc Gly	tgg Trp	gtt Val 3755	tgc Cys	tgg Trp	atg Met	att Ile	ctt Leu 3760	ttt Phe	11308
gat Asp	tca Ser	agc Ser	ctt Leu 3765	tat Tyr	ttt Phe	ttg Leu	tgt Cys	gga Gly 3770	tgg Trp	tac Tyr	ttg Leu	agc Ser	aac Asn 3775	ttg Leu	11353
att Ile	cct Pro	gga Gly	aca Thr 3780	ttt Phe	ggt Gly	tta Leu	cgg Arg	aaa Lys 3785	cca Pro	tgg Trp	tat Tyr	ttc Phe	ccc Pro 3790	ttt Phe	11398
act Thr	gcc Ala	tca Ser	tat Tyr 3795	tgg Trp	aag Lys	agt Ser	gtg Val	ggt Gly 3800	ttc Phe	ttg Leu	gtg Val	gag Glu	aaa Lys 3805	agg Arg	11443
caa Gln	tac Tyr	ttt Phe	cta Leu 3810	agt Ser	tct Ser	agt Ser	ctg Leu	ttc Phe 3815	ttc Phe	ttc Phe	aat Asn	gag Glu	aac Asn 3820	ttt Phe	11488
gac Asp	aat Asn	aaa Lys	ggg Gly 3825	tca Ser	tca Ser	ctg Leu	caa Gln	aac Asn 3830	agg Arg	gaa Glu	gga Gly	gag Glu	ctt Leu 3835	gaa Glu	11533
gga Gly	agt Ser	gcc Ala	ccg Pro 3840	gga Gly	gtc Val	acc Thr	ctg Leu	gtg Val 3845	tct Ser	gtg Val	acc Thr	aag Lys	gaa Glu 3850	tat Tyr	11578
gag	ggc	cac	aag	gct	gtg	gtc	caa	gac	ctc	agc	ctg	acc	ttc	tac	11623

Glu	Gly	His	Lys	Ala	Val	Val	Gln	Asp	Leu	Ser	Leu	Thr	Phe	Tyr	
			3855					3860					3865		
aga	gac	caa	atc	acc	gcc	ctg	ctg	ggg	aca	aac	ggt	gcc	ggg	aaa	11668
Arg	Asp	Gln	Ile	Thr	Ala	Leu	Leu	Gly	Thr	Asn	Gly	Ala	Gly	Lys	
			3870					3875					3880		
acc	act	atc	ata	tcc	atg	ttg	acg	ggg	ctc	cac	cct	ccc	act	tct	11713
Thr	Thr	Ile	Ile	Ser	Met	Leu	Thr	Gly	Leu	His	Pro	Pro	Thr	Ser	
			3885					3890					3895		
gga	acc	atc	atc	atc	aat	ggc	aag	aac	cta	cag	aca	gac	ctg	tcg	11758
Gly	Thr	Ile	Ile	Ile	Asn	Gly	Lys	Asn	Leu	Gln	Thr	Asp	Leu	Ser	
			3900					3905					3910		
agg	gtc	aga	atg	gag	ctt	ggt	gtg	tgt	ccg	cag	cag	gac	atc	ctg	11803
Arg	Val	Arg	Met	Glu	Leu	Gly	Val	Cys	Pro	Gln	Gln	Asp	Ile	Leu	
			3915					3920					3925		
ttg	gac	aac	ctc	acc	gtc	cgg	gaa	cat	ttg	ctg	ctc	ttt	gct	tcc	11848
Leu	Asp	Asn	Leu	Thr	Val	Arg	Glu	His	Leu	Leu	Leu	Phe	Ala	Ser	
			3930					3935					3940		
ata	aag	gcg	cct	cag	tgg	acc	aag	aag	gag	ctg	cat	cag	caa	gtc	11893
Ile	Lys	Ala	Pro	Gln	Trp	Thr	Lys	Lys	Glu	Leu	His	Gln	Gln	Val	
			3945					3950					3955		
aat	caa	act	ctt	cag	gat	gtg	gac	tta	act	cag	cat	cag	cac	aaa	11938
Asn	Gln	Thr	Leu	Gln	Asp	Val	Asp	Leu	Thr	Gln	His	Gln	His	Lys	
			3960					3965					3970		
cag	acc	cga	gct	ctg	tct	gga	ggc	ctg	aag	agg	aag	ctc	tcc	ctt	11983
Gln	Thr	Arg	Ala	Leu	Ser	Gly	Gly	Leu	Lys	Arg	Lys	Leu	Ser	Leu	
			3975					3980					3985		
ggc	att	gct	ttc	atg	ggc	atg	tcg	agg	acc	gtg	ggt	ctg	gat	gag	12028
Gly	Ile	Ala	Phe	Met	Gly	Met	Ser	Arg	Thr	Val	Val	Leu	Asp	Glu	
			3990					3995					4000		
ccc	acc	agt	ggg	gtg	gac	cct	tgc	tcc	cgg	cat	agc	ctg	tgg	gac	12073
Pro	Thr	Ser	Gly	Val	Asp	Pro	Cys	Ser	Arg	His	Ser	Leu	Trp	Asp	
			4005					4010					4015		
att	ctg	ctc	aag	tac	cga	gaa	ggt	cgt	acg	atc	atc	ttc	aca	acc	12118
Ile	Leu	Leu	Lys	Tyr	Arg	Glu	Gly	Arg	Thr	Ile	Ile	Phe	Thr	Thr	
			4020					4025					4030		
cac	cac	ctg	gat	gaa	gcc	gaa	gcg	ctg	agt	gac	cgc	gtg	gcc	gtc	12163
His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser	Asp	Arg	Val	Ala	Val	
			4035					4040					4045		
ctc	cag	cat	ggg	agg	ctc	agg	tgc	tgc	ggt	cct	ccc	ttc	tgc	ctg	12208
Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly	Pro	Pro	Phe	Cys	Leu	
			4050					4055					4060		
aag	gag	gca	tat	ggc	cag	ggg	ctc	cgc	ctg	aca	ctc	acg	agg	cag	12253
Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu	Thr	Leu	Thr	Arg	Gln	
			4065					4070					4075		
cct	tct	ggt	ctg	gag	gcc	cat	gat	ctg	aaa	gac	atg	gct	tgt	gtt	12298
Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys	Asp	Met	Ala	Cys	Val	

4080	4085	4090	
aca tcc ctg ata aag atc tat att cca caa gca ttt ctc aaa gac			12343
Thr Ser Leu Ile Lys Ile Tyr Ile Pro Gln Ala Phe Leu Lys Asp			
4095	4100	4105	
agc agt gga agt gag ctg acc tac acc att cca aag gac aca gac			12388
Ser Ser Gly Ser Glu Leu Thr Tyr Thr Ile Pro Lys Asp Thr Asp			
4110	4115	4120	
aag gcc tgc ttg aaa ggg ctc ttc cag gcc ctg gat gag aac ctg			12433
Lys Ala Cys Leu Lys Gly Leu Phe Gln Ala Leu Asp Glu Asn Leu			
4125	4130	4135	
cat cag ctg cac ctg acg ggc tat ggg atc tca gac acc acc tta			12478
His Gln Leu His Leu Thr Gly Tyr Gly Ile Ser Asp Thr Thr Leu			
4140	4145	4150	
gaa gag gtg ttt ttg atg ott ttg caa gat tcc aac aag aaa tct			12523
Glu Glu Val Phe Leu Met Leu Leu Gln Asp Ser Asn Lys Lys Ser			
4155	4160	4165	
cac att gcc ctg ggg act gag tca gag ctg cag aac cac agg cct			12568
His Ile Ala Leu Gly Thr Glu Ser Glu Leu Gln Asn His Arg Pro			
4170	4175	4180	
aca gga cat ctg tct ggc tac tgt ggc tcc cta gca cgg ccc gca			12613
Thr Gly His Leu Ser Gly Tyr Cys Gly Ser Leu Ala Arg Pro Ala			
4185	4190	4195	
act gtg cag ggc gtc cag ctg ctc cgc gca caa gtg gcc gcg atc			12658
Thr Val Gln Gly Val Gln Leu Leu Arg Ala Gln Val Ala Ala Ile			
4200	4205	4210	
ctg gcc cgg agg ctc cgc cgc acg ctg cgc gcc ggg aag agc acc			12703
Leu Ala Arg Arg Leu Arg Arg Thr Leu Arg Ala Gly Lys Ser Thr			
4215	4220	4225	
ctc gcc gac ctg ctg ctg cca gtc ctc ttc gtg gcc ttg gcc atg			12748
Leu Ala Asp Leu Leu Leu Pro Val Leu Phe Val Ala Leu Ala Met			
4230	4235	4240	
ggc ttg ttc atg gtg aga ccc ctg gcc acc gag tac cct ccc ctc			12793
Gly Leu Phe Met Val Arg Pro Leu Ala Thr Glu Tyr Pro Pro Leu			
4245	4250	4255	
aga ctc aca cct gga cat tac cag cgg gcc gag acc tac ttt ttc			12838
Arg Leu Thr Pro Gly His Tyr Gln Arg Ala Glu Thr Tyr Phe Phe			
4260	4265	4270	
agc agt ggg ggc gac aac ttg gac ctc acc cgt gtg ctt ctg cgg			12883
Ser Ser Gly Gly Asp Asn Leu Asp Leu Thr Arg Val Leu Leu Arg			
4275	4280	4285	
aag ttt aga gat caa gat ttg ccc tgt gca gat tta aac cca cgc			12928
Lys Phe Arg Asp Gln Asp Leu Pro Cys Ala Asp Leu Asn Pro Arg			
4290	4295	4300	
cag aag aat tct tca tgc tgg cgc aca gat ccc ttt tct cac cca			12973
Gln Lys Asn Ser Ser Cys Trp Arg Thr Asp Pro Phe Ser His Pro			
4305	4310	4315	

gaa ttc cag gat	tca tgt ggc tgc ctg	aag tgt cca aat aga	agt	13018
Glu Phe Gln Asp	Ser Cys Gly Cys Leu	Lys Cys Pro Asn Arg	Ser	
4320	4325	4330		
gct agt gct ccc	tac ctg acc aac cac	ctg ggc cac aca ctg	ttg	13063
Ala Ser Ala Pro	Tyr Leu Thr Asn His	Leu Gly His Thr Leu	Leu	
4335	4340	4345		
aat ctc tca ggc	ttc aat atg gag gag	tac ttg ctg gca cca	tct	13108
Asn Leu Ser Gly	Phe Asn Met Glu Glu	Tyr Leu Leu Ala Pro	Ser	
4350	4355	4360		
gaa aaa cca agg	ctt gga ggt tgg tct	ttt gga tta aaa atc	ccc	13153
Glu Lys Pro Arg	Leu Gly Gly Trp Ser	Phe Gly Leu Lys Ile	Pro	
4365	4370	4375		
agt gaa gct gga	ggt gca aat gga aac	ata tca aaa ccc cca	act	13198
Ser Glu Ala Gly	Gly Ala Asn Gly Asn	Ile Ser Lys Pro Pro	Thr	
4380	4385	4390		
ctg gca aag gtg	tgg tat aat cag aag	ggt ttt cat tcc cta	cct	13243
Leu Ala Lys Val	Trp Tyr Asn Gln Lys	Gly Phe His Ser Leu	Pro	
4395	4400	4405		
tcc tac tta aat	cat cta aac aac ctt	att ttg tgg cag cac	cta	13288
Ser Tyr Leu Asn	His Leu Asn Asn Leu	Ile Leu Trp Gln His	Leu	
4410	4415	4420		
ccc cct act gtg	gac tgg aga caa tac	gga ata aca ctc tac	agc	13333
Pro Pro Thr Val	Asp Trp Arg Gln Tyr	Gly Ile Thr Leu Tyr	Ser	
4425	4430	4435		
cac cca tat gga	ggg gcc ttg ctg aac	gag gac aag atc ctg	gag	13378
His Pro Tyr Gly	Gly Ala Leu Leu Asn	Glu Asp Lys Ile Leu	Glu	
4440	4445	4450		
agc atc cgt cag	tgt gga gtg gcc ctc	tgc atc gtg ctg gga	ttc	13423
Ser Ile Arg Gln	Cys Gly Val Ala Leu	Cys Ile Val Leu Gly	Phe	
4455	4460	4465		
tcc atc ctg tct	gca tcc atc ggc agc	tct gtg gtg agg gac	agg	13468
Ser Ile Leu Ser	Ala Ser Ile Gly Ser	Ser Val Val Arg Asp	Arg	
4470	4475	4480		
gtg att gga gcc	aaa agg ttg cag cac	ata agt ggc ctt ggc	tac	13513
Val Ile Gly Ala	Lys Arg Leu Gln His	Ile Ser Gly Leu Gly	Tyr	
4485	4490	4495		
agg atg tac tgg	ttc aca aac ttc cta	tat gac atg ctc ttt	tac	13558
Arg Met Tyr Trp	Phe Thr Asn Phe Leu	Tyr Asp Met Leu Phe	Tyr	
4500	4505	4510		
ttg gtt tcc gtc	tgc ctg tgt gtt gcc	gtt att gtc gcc ttc	cag	13603
Leu Val Ser Val	Cys Leu Cys Val Ala	Val Ile Val Ala Phe	Gln	
4515	4520	4525		
tta aca gct ttt	act ttc cgc aag aac	ttg gca gcc acg gcc	ctc	13648
Leu Thr Ala Phe	Thr Phe Arg Lys Asn	Leu Ala Ala Thr Ala	Leu	
4530	4535	4540		

ctg ctg tca ctt ttc gga tat gca act ctt cca tgg atg tac ctg	13693
Leu Leu Ser Leu Phe Gly Tyr Ala Thr Leu Pro Trp Met Tyr Leu	
4545 4550 4555	
atg tcc aga atc ttt tcc agt tcg gac gtg gct ttc att tcc tat	13738
Met Ser Arg Ile Phe Ser Ser Ser Asp Val Ala Phe Ile Ser Tyr	
4560 4565 4570	
gtc tca cta aac ttc atc ttt ggc ctt tgt acc atg ccc ata acc	13783
Val Ser Leu Asn Phe Ile Phe Gly Leu Cys Thr Met Pro Ile Thr	
4575 4580 4585	
att atg ccc cgg ttg cta gcc atc atc tcc aaa gct aag aat tta	13828
Ile Met Pro Arg Leu Leu Ala Ile Ile Ser Lys Ala Lys Asn Leu	
4590 4595 4600	
cag aat atc tat gat gtc ctc aag tgg gtc ttt act att ttt cct	13873
Gln Asn Ile Tyr Asp Val Leu Lys Trp Val Phe Thr Ile Phe Pro	
4605 4610 4615	
caa ttc tgt ctt ggt caa gga ctg gta gaa ctc tgc tat aat cag	13918
Gln Phe Cys Leu Gly Gln Gly Leu Val Glu Leu Cys Tyr Asn Gln	
4620 4625 4630	
atc aaa tat gac ctg acc cac aac ttc ggc att gat tcc tat gtg	13963
Ile Lys Tyr Asp Leu Thr His Asn Phe Gly Ile Asp Ser Tyr Val	
4635 4640 4645	
agt ccc ttt gag atg aac ttt ctg ggc tgg atc ttc gtg caa ctg	14008
Ser Pro Phe Glu Met Asn Phe Leu Gly Trp Ile Phe Val Gln Leu	
4650 4655 4660	
gcc tcg cag ggc aca gta ctt ctc ctc ttg agg gtt ctg cta cac	14053
Ala Ser Gln Gly Thr Val Leu Leu Leu Leu Arg Val Leu Leu His	
4665 4670 4675	
tgg gac ctt ctg cga tgg cca agg ggt cat tct act ctc caa ggc	14098
Trp Asp Leu Leu Arg Trp Pro Arg Gly His Ser Thr Leu Gln Gly	
4680 4685 4690	
aca gtc aaa tct tct aag gat aca gat gtt gaa aaa gag gaa aag	14143
Thr Val Lys Ser Ser Lys Asp Thr Asp Val Glu Lys Glu Glu Lys	
4695 4700 4705	
aga gtg ttt gaa gga agg acc aat gga gac att ctt gtg tta tac	14188
Arg Val Phe Glu Gly Arg Thr Asn Gly Asp Ile Leu Val Leu Tyr	
4710 4715 4720	
aac ctt agt aaa cat tat cga cgc ttt ttc cag aat att att gct	14233
Asn Leu Ser Lys His Tyr Arg Arg Phe Phe Gln Asn Ile Ile Ala	
4725 4730 4735	
gtg caa gat att agt ttg ggc ata cca aaa gga gag tgc ttt gga	14278
Val Gln Asp Ile Ser Leu Gly Ile Pro Lys Gly Glu Cys Phe Gly	
4740 4745 4750	
ctt cta ggg gtg aat gga gct ggg aag agc acg act ttc aaa atg	14323
Leu Leu Gly Val Asn Gly Ala Gly Lys Ser Thr Thr Phe Lys Met	
4755 4760 4765	
ctg aat ggt gaa gtt tct cta act tca gga cat gct atc atc agg	14368

Leu Asn Gly Glu Val Ser Leu Thr Ser Gly His Ala Ile Ile Arg	
4770 4775 4780	
act ccc atg gga gac gcc gtg gac ctg tct tct gct ggc acg gca	14413
Thr Pro Met Gly Asp Ala Val Asp Leu Ser Ser Ala Gly Thr Ala	
4785 4790 4795	
ggc gtg ctc att ggc tac tgt ccc cag cag gat gcc ctg gac gag	14458
Gly Val Leu Ile Gly Tyr Cys Pro Gln Gln Asp Ala Leu Asp Glu	
4800 4805 4810	
ctt ctg act ggt tgg gaa cat ctc tat tat tac tgt agc tta cgc	14503
Leu Leu Thr Gly Trp Glu His Leu Tyr Tyr Tyr Cys Ser Leu Arg	
4815 4820 4825	
ggg att cca agg cag tgc atc cct gag gtt gct gga gac ctc atc	14548
Gly Ile Pro Arg Gln Cys Ile Pro Glu Val Ala Gly Asp Leu Ile	
4830 4835 4840	
agg cgc tta cac ctc gaa gcc cac gcg gac aaa cct gtg gcc acc	14593
Arg Arg Leu His Leu Glu Ala His Ala Asp Lys Pro Val Ala Thr	
4845 4850 4855	
tac agt ggg gga acc aag cgg aaa ctc tct aca gcc ctg gcc ctg	14638
Tyr Ser Gly Gly Thr Lys Arg Lys Leu Ser Thr Ala Leu Ala Leu	
4860 4865 4870	
gtg ggg aaa cct gac att ctt tta ttg gat gag ccc agc tct ggg	14683
Val Gly Lys Pro Asp Ile Leu Leu Leu Asp Glu Pro Ser Ser Gly	
4875 4880 4885	
atg gat ccc tgc tct aag cgg tac ctg tgg caa aca ata atg aag	14728
Met Asp Pro Cys Ser Lys Arg Tyr Leu Trp Gln Thr Ile Met Lys	
4890 4895 4900	
gag gtt cgg gaa ggc tgt gct gcg gtg ctg acc tcc cac agc atg	14773
Glu Val Arg Glu Gly Cys Ala Ala Val Leu Thr Ser His Ser Met	
4905 4910 4915	
gag gag tgt gag gct ctt tgc aca aga ctg gcc ata atg gtt aac	14818
Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Asn	
4920 4925 4930	
ggc agc ttc aaa tgt ctt ggt tct cct cag cac atc aaa aat agg	14863
Gly Ser Phe Lys Cys Leu Gly Ser Pro Gln His Ile Lys Asn Arg	
4935 4940 4945	
tgc gtt gaa gtt ctc ctt ttg ctt aga tat tta cta tga aatcaatagc	14912
Cys Val Glu Val Leu Leu Leu Leu Arg Tyr Leu Leu	
4950 4955	
atgatgcagg gttatgactg gagatgcttt agaggcaagt atgttccaat tacctgctgg	14972
cataaaatag gagatgaagg ctctgagggg ataaaataact gtggctagta agtcaaacac	15032
atagcaactc ccaaacacag tctagtgttc tccctgtaag ggagacattt tgctttttct	15092
cttgccaggt agacattgct ccttccttgt agtgacattg gtaggggtga gaataagatt	15152
gcagagtcaa gatgatgaca ggacacaaaa tgtcctcttt acacacaaaa tgccttatga	15212

```

aaaggctgtg ctgatatgtc cctcatctgc agtgaatcaa agttccatt ttatcctacc 15272
aaacactatc cttttgaaaa agaccatcaa ttataaggg taggagaggg ttgttatggg 15332
tgctttccat ttggattatg aaagactcaa gacttctttc ctgtgctcat tgatccttta 15392
tatttcttcc tggttaattc tcttcatgta ttttctttat ttcttttttg gaatcttaat 15452
aattttgttg tcaatttggt tgagcattgt tacatcctgg tataacatta aaaaaattaa 15512
agttgcctaa atacacgtcc aattttcctt tgatttgttt aacctgtcat tcttatgtag 15572
aataatttat tcatattatg tgataaatct cttattttca tttattcagt gtttttaagg 15632
ttaggattct caagcagaga ttggtaaagt agtcatatgt aaattttaaa tacttttaat 15692
ttatggagaa aatagctgat gctcaggatt gagatgacat agtagggagt gagtgtccat 15752
ttttagagtt aggcaactgt gggatgggtc gcctcacagt tgcttgctgc ttcggggtcaa 15812
ccactgggtc gatgcaccta ctggaaacta ttctcggtat gtactctgcc catctttaat 15872
aagtatcatc aaaagaaaaa taatatttca tgggtctcaa cctttatgaa tgcctacata 15932
aagttttata acccaaaaaa aaaaaaaaaa aaaaaaaaaa 15970

```

<210> 23
 <211> 4958
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> misc_feature
 <222> (283)..(283)
 <223> The 'Xaa' at location 283 stands for Asp, Gly, Ala, or Val.

<220>
 <221> misc_feature
 <222> (2674)..(2674)
 <223> The 'Xaa' at location 2674 stands for Arg, Gly, or Trp.

<400> 23

```

Met Gly His Ala Gly Cys Gln Phe Lys Ala Leu Leu Trp Lys Asn Trp
1           5           10           15

```

```

Leu Cys Arg Leu Arg Asn Pro Val Leu Phe Leu Ala Glu Phe Phe Trp
20           25           30

```

```

Pro Cys Ile Leu Phe Val Ile Leu Thr Val Leu Arg Phe Gln Glu Pro
35           40           45

```

```

Pro Arg Tyr Arg Asp Ile Cys Tyr Leu Gln Pro Arg Asp Leu Pro Ser
50           55           60

```

```

Cys Gly Val Ile Pro Phe Val Gln Ser Leu Leu Cys Asn Thr Gly Ser

```


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

Glu Ala Glu Lys Trp Gly His Val Gly Gly Cys His Pro Lys Trp Ser
 325 330 335
 Glu Ala Lys Asn Tyr Leu Val His Ala Val Ser Trp Leu Arg Val Tyr
 340 345 350
 Gln Gln Val Phe Val Gln Trp Gln Gln Gly Ser Leu Leu Gln Lys Thr
 355 360 365
 Leu Thr Gly Met Gly His Ser Leu Glu Ala Leu Arg Asn Gln Phe Glu
 370 375 380
 Glu Glu Ser Lys Pro Trp Lys Val Val Glu Ala Leu His Thr Ala Leu
 385 390 395 400
 Leu Leu Leu Asn Asp Ser Leu Ser Ala Asp Gly Pro Lys Asp Asn His
 405 410 415
 Thr Phe Pro Lys Ile Leu Gln His Leu Trp Lys Leu Gln Ser Leu Leu
 420 425 430
 Gln Asn Leu Pro Gln Trp Pro Ala Leu Lys Arg Phe Leu Gln Leu Asp
 435 440 445
 Gly Ala Leu Arg Asn Ala Ile Ala Gln Asn Leu His Phe Val Gln Glu
 450 455 460
 Val Leu Ile Cys Leu Glu Thr Ser Ala Asn Asp Phe Lys Trp Phe Glu
 465 470 475 480
 Leu Asn Gln Leu Lys Leu Glu Lys Asp Val Phe Phe Trp Glu Leu Lys
 485 490 495
 Gln Met Leu Ala Lys Asn Ala Val Cys Pro Asn Gly Arg Phe Ser Glu
 500 505 510
 Lys Glu Val Phe Leu Pro Pro Gly Asn Ser Ser Ile Trp Gly Gly Leu
 515 520 525
 Gln Gly Leu Leu Cys Tyr Cys Asn Ser Ser Glu Thr Ser Val Leu Asn
 530 535 540
 Lys Leu Leu Gly Ser Val Glu Asp Ala Asp Arg Ile Leu Gln Glu Val
 545 550 555 560

Ile Thr Trp His Lys Asn Met Ser Val Leu Ile Pro Glu Glu Tyr Leu
565 570 575

Asp Trp Gln Glu Leu Glu Met Gln Leu Ser Glu Ala Ser Leu Ser Cys
580 585 590

Thr Arg Leu Phe Leu Leu Leu Gly Ala Asp Pro Ser Pro Glu Asn Asp
595 600 605

Val Phe Ser Ser Asp Cys Lys His Gln Leu Val Ser Thr Val Ile Phe
610 615 620

His Thr Leu Glu Lys Thr Gln Phe Phe Leu Glu Gln Ala Tyr Tyr Trp
625 630 635 640

Lys Ala Phe Lys Lys Phe Ile Arg Lys Thr Cys Glu Val Ala Gln Tyr
645 650 655

Val Asn Met Gln Glu Ser Phe Gln Asn Arg Leu Leu Ala Phe Pro Glu
660 665 670

Glu Ser Pro Cys Phe Glu Glu Asn Met Asp Trp Lys Met Ile Ser Asp
675 680 685

Asn Tyr Phe Gln Phe Leu Asn Asn Leu Leu Lys Ser Pro Thr Ala Ser
690 695 700

Ile Ser Arg Ala Leu Asn Phe Thr Lys His Leu Leu Met Met Glu Lys
705 710 715 720

Lys Leu His Thr Leu Glu Asp Glu Gln Met Asn Phe Leu Leu Ser Phe
725 730 735

Val Glu Phe Phe Glu Lys Leu Leu Leu Pro Asn Leu Phe Asp Ser Ser
740 745 750

Ile Val Pro Ser Phe His Ser Leu Pro Ser Leu Thr Glu Asp Ile Leu
755 760 765

Asn Ile Ser Ser Leu Trp Thr Asn His Leu Lys Ser Leu Lys Arg Asp
770 775 780

Pro Ser Ala Thr Asp Ala Gln Lys Leu Leu Glu Phe Gly Asn Glu Val
785 790 795 800

Ile Trp Lys Met Gln Thr Leu Gly Ser His Trp Ile Arg Lys Glu Pro
805 810 815

Lys Asn Leu Leu Arg Phe Ile Glu Leu Ile Leu Phe Glu Ile Asn Pro
820 825 830

Lys Leu Leu Glu Leu Trp Ala Tyr Gly Ile Ser Lys Gly Lys Arg Ala
835 840 845

Lys Leu Glu Asn Phe Phe Thr Leu Leu Asn Phe Ser Val Pro Glu Asn
850 855 860

Glu Ile Leu Ser Thr Ser Phe Asn Phe Ser Gln Leu Phe His Ser Asp
865 870 875 880

Trp Pro Lys Ser Pro Ala Met Asn Ile Asp Phe Val Arg Leu Ser Glu
885 890 895

Ala Ile Ile Thr Ser Leu His Glu Phe Gly Phe Leu Glu Gln Glu Gln
900 905 910

Ile Ser Glu Ala Leu Asn Thr Val Tyr Ala Ile Arg Asn Ala Ser Asp
915 920 925

Leu Phe Ser Ala Leu Ser Glu Pro Gln Lys Gln Glu Val Asp Lys Ile
930 935 940

Leu Thr His Ile His Leu Asn Val Phe Gln Asp Lys Asp Ser Ala Leu
945 950 955 960

Leu Leu Gln Ile Tyr Ser Ser Phe Tyr Arg Tyr Ile Tyr Glu Leu Leu
965 970 975

Asn Ile Gln Ser Arg Gly Ser Ser Leu Thr Phe Leu Thr Gln Ile Ser
980 985 990

Lys His Ile Leu Asp Ile Ile Lys Gln Phe Asn Phe Gln Asn Ile Ser
995 1000 1005

Lys Ala Phe Ala Phe Leu Phe Lys Thr Ala Glu Val Leu Gly Gly
1010 1015 1020

Ile Ser Asn Val Ser Tyr Cys Gln Gln Leu Leu Ser Ile Phe Asn
1025 1030 1035

Phe Leu Glu Leu Gln Ala Gln Ser Phe Met Ser Thr Glu Gly Gln

1040		1045		1050
Glu Leu Glu Val Ile His Thr Thr Leu Thr Gly Leu Lys Gln Leu				
1055		1060		1065
Leu Ile Ile Asp Glu Asp Phe Arg Ile Ser Leu Phe Gln Tyr Met				
1070		1075		1080
Ser Gln Phe Phe Asn Ser Ser Val Glu Asp Leu Leu Asp Asn Lys				
1085		1090		1095
Cys Leu Ile Ser Asp Asn Lys His Ile Ser Ser Val Asn Tyr Ser				
1100		1105		1110
Thr Ser Glu Glu Ser Ser Phe Val Phe Pro Leu Ala Gln Ile Phe				
1115		1120		1125
Ser Asn Leu Ser Ala Asn Val Ser Val Phe Asn Lys Phe Met Ser				
1130		1135		1140
Ile His Cys Thr Val Ser Trp Leu Gln Met Trp Thr Glu Ile Trp				
1145		1150		1155
Glu Thr Ile Ser Gln Leu Phe Lys Phe Asp Met Asn Val Phe Thr				
1160		1165		1170
Ser Leu His His Gly Phe Thr Gln Leu Leu Asp Glu Leu Glu Asp				
1175		1180		1185
Asp Val Lys Val Ser Lys Ser Cys Gln Gly Ile Leu Pro Thr His				
1190		1195		1200
Asn Val Ala Arg Leu Ile Leu Asn Leu Phe Lys Asn Val Thr Gln				
1205		1210		1215
Ala Asn Asp Phe His Asn Trp Glu Asp Phe Leu Asp Leu Arg Asp				
1220		1225		1230
Phe Leu Val Ala Leu Gly Asn Ala Leu Val Ser Val Lys Lys Leu				
1235		1240		1245
Asn Leu Glu Gln Val Glu Lys Ser Leu Phe Thr Met Glu Ala Ala				
1250		1255		1260
Leu His Gln Leu Lys Thr Phe Pro Phe Asn Glu Ser Thr Ser Arg				
1265		1270		1275

Glu Phe Leu Asn Ser Leu Leu Glu Val Phe Ile Glu Phe Ser Ser	1280	1285	1290
Thr Ser Glu Tyr Ile Val Arg Asn Leu Asp Ser Ile Asn Asp Phe	1295	1300	1305
Leu Ser Asn Asn Leu Thr Asn Tyr Gly Glu Lys Phe Glu Asn Ile	1310	1315	1320
Ile Thr Glu Leu Arg Glu Ala Ile Val Phe Leu Arg Asn Val Ser	1325	1330	1335
His Asp Arg Asp Leu Phe Ser Cys Ala Asp Ile Phe Gln Asn Val	1340	1345	1350
Thr Glu Cys Ile Leu Glu Asp Gly Phe Leu Tyr Val Asn Thr Ser	1355	1360	1365
Gln Arg Met Leu Arg Ile Leu Asp Thr Leu Asn Ser Thr Phe Ser	1370	1375	1380
Ser Glu Asn Thr Ile Ser Ser Leu Lys Gly Cys Ile Val Trp Leu	1385	1390	1395
Asp Val Ile Asn His Leu Tyr Leu Leu Ser Asn Ser Ser Phe Ser	1400	1405	1410
Gln Gly Arg Leu Gln Asn Ile Leu Gly Asn Phe Arg Asp Ile Glu	1415	1420	1425
Asn Lys Met Asn Ser Ile Leu Lys Ile Val Thr Trp Val Leu Asn	1430	1435	1440
Ile Lys Lys Pro Leu Cys Ser Ser Asn Gly Ser His Ile Asn Cys	1445	1450	1455
Val Asn Ile Tyr Leu Lys Asp Val Thr Asp Phe Leu Asn Ile Val	1460	1465	1470
Leu Thr Thr Val Phe Glu Lys Glu Lys Lys Pro Lys Phe Glu Ile	1475	1480	1485
Leu Leu Ala Leu Leu Asn Asp Ser Thr Lys Gln Val Arg Met Ser	1490	1495	1500

Ile Asn Asn Leu Thr Thr Asp Phe Asp Phe Ala Ser Gln Ser Asn	
1505	1510 1515
Trp Arg Tyr Phe Thr Glu Leu Ile Leu Arg Pro Ile Glu Met Ser	
1520	1525 1530
Asp Glu Ile Pro Asn Gln Phe Gln Asn Ile Trp Leu His Leu Ile	
1535	1540 1545
Thr Leu Gly Lys Glu Phe Gln Lys Leu Val Lys Gly Ile Tyr Phe	
1550	1555 1560
Asn Ile Leu Glu Asn Asn Ser Ser Ser Lys Thr Glu Asn Leu Leu	
1565	1570 1575
Asn Ile Phe Ala Thr Ser Pro Lys Glu Lys Asp Val Asn Ser Val	
1580	1585 1590
Gly Asn Ser Ile Tyr His Leu Ala Ser Tyr Leu Ala Phe Ser Leu	
1595	1600 1605
Ser His Asp Leu Gln Asn Ser Pro Lys Ile Ile Ile Ser Pro Glu	
1610	1615 1620
Ile Met Lys Ala Thr Gly Leu Gly Ile Gln Leu Ile Arg Asp Val	
1625	1630 1635
Phe Asn Ser Leu Met Pro Val Val His His Thr Ser Pro Gln Asn	
1640	1645 1650
Ala Gly Tyr Met Gln Ala Leu Lys Lys Val Thr Ser Val Met Arg	
1655	1660 1665
Thr Leu Lys Lys Ala Asp Ile Asp Leu Leu Val Asp Gln Leu Glu	
1670	1675 1680
Gln Val Ser Val Asn Leu Met Asp Phe Phe Lys Asn Ile Ser Ser	
1685	1690 1695
Val Gly Thr Gly Asn Leu Val Val Asn Leu Leu Val Gly Leu Met	
1700	1705 1710
Glu Lys Phe Ala Asp Ser Ser His Ser Trp Asn Val Asn His Leu	
1715	1720 1725

Leu Gln Leu Ser Arg Leu Phe Pro Lys Asp Val Val Asp Ala Val
1730 1735 1740

Ile Asp Val Tyr Tyr Val Leu Pro His Ala Val Arg Leu Leu Gln
1745 1750 1755

Gly Val Pro Gly Lys Asn Ile Thr Glu Gly Leu Lys Asp Val Tyr
1760 1765 1770

Ser Phe Thr Leu Leu His Gly Ile Thr Ile Ser Asn Ile Thr Lys
1775 1780 1785

Glu Asp Phe Ala Ile Val Ile Lys Ile Leu Leu Asp Thr Ile Glu
1790 1795 1800

Leu Val Ser Asp Lys Pro Asp Ile Ile Ser Glu Ala Leu Ala Cys
1805 1810 1815

Phe Pro Val Val Trp Cys Trp Asn His Thr Asn Ser Gly Phe Arg
1820 1825 1830

Gln Asn Ser Lys Ile Asp Pro Cys Asn Val His Gly Leu Met Ser
1835 1840 1845

Ser Ser Phe Tyr Gly Lys Val Ala Ser Ile Leu Asp His Phe His
1850 1855 1860

Leu Ser Pro Gln Gly Glu Asp Ser Pro Cys Ser Asn Glu Ser Ser
1865 1870 1875

Arg Met Glu Ile Thr Arg Lys Val Val Cys Ile Ile His Glu Leu
1880 1885 1890

Val	Asp	Trp	Asn	Ser	Ile	Leu	Leu	Glu	Leu	Ser	Glu	Val	Phe	His
1895						1900					1905			

Val	Asn	Ile	Ser	Leu	Val	Lys	Thr	Val	Gln	Lys	Phe	Trp	His	Lys
1910						1915					1920			

Ile Leu Pro Phe Val Pro Pro Ser Ile Asn Gln Thr Arg Asp Ser
1925 1930 1935

Ile Ser Glu Leu Cys Pro Ser Gly Ser Ile Lys Gln Val Ala Leu
1940 1945 1950

Gln Ile Ile Glu Lys Leu Lys Asn Val Asn Phe Thr Lys Val Thr

1955	1960	1965
Ser Gly Glu Asn Ile Leu Asp 1970	Lys Leu Ser Ser 1975	Leu Asn Lys Ile 1980
Leu Asn Ile Asn Glu Asp Thr 1985	Glu Thr Ser Val 1990	Gln Asn Ile Ile 1995
Ser Ser Asn Leu Glu Arg Thr 2000	Val Gln Leu Ile 2005	Ser Glu Asp Trp 2010
Ser Leu Glu Lys Ser Thr His 2015	Asn Leu Leu Ser 2020	Leu Phe Met Met 2025
Leu Gln Asn Ala Asn Val Thr 2030	Gly Ser Ser Leu 2035	Glu Ala Leu Ser 2040
Ser Phe Ile Glu Lys Ser Glu 2045	Thr Pro Tyr Asn 2050	Phe Glu Glu Leu 2055
Trp Pro Lys Phe Gln Gln Ile 2060	Met Lys Asp Leu 2065	Thr Gln Asp Phe 2070
Arg Ile Arg His Leu Leu Ser 2075	Glu Met Asn Lys 2080	Gly Ile Lys Ser 2085
Ile Asn Ser Met Ala Leu Gln 2090	Lys Ile Thr Leu 2095	Gln Phe Ala His 2100
Phe Leu Glu Ile Leu Asp Ser 2105	Pro Ser Leu Lys 2110	Thr Leu Glu Ile 2115
Ile Glu Asp Phe Leu Leu Val 2120	Thr Lys Asn Trp 2125	Leu Gln Glu Tyr 2130
Ala Asn Glu Asp Tyr Ser Arg 2135	Met Ile Glu Thr 2140	Leu Phe Ile Pro 2145
Val Thr Asn Glu Ser Ser Thr 2150	Glu Asp Ile Ala 2155	Leu Leu Ala Lys 2160
Ala Ile Ala Thr Phe Trp Gly 2165	Ser Leu Lys Asn 2170	Ile Ser Arg Ala 2175
Gly Asn Phe Asp Val Ala Phe 2180	Leu Thr His Leu 2185	Leu Asn Gln Glu 2190

Gln Leu Thr Asn Phe Ser Val	Val Gln Leu Leu Phe	Glu Asn Ile
2195	2200	2205
Leu Ile Asn Leu Ile Asn Asn	Leu Ala Gly Asn Ser	Gln Glu Ala
2210	2215	2220
Ala Trp Asn Leu Asn Asp Thr	Asp Leu Gln Ile Met	Asn Phe Ile
2225	2230	2235
Asn Leu Ile Leu Asn His Met	Gln Ser Glu Thr Ser	Arg Lys Thr
2240	2245	2250
Val Leu Ser Leu Arg Ser Ile	Val Asp Phe Thr Glu	Gln Phe Leu
2255	2260	2265
Lys Thr Phe Phe Ser Leu Phe	Leu Lys Glu Asp Ser	Glu Asn Lys
2270	2275	2280
Ile Ser Leu Leu Leu Lys Tyr	Phe His Lys Asp Val	Ile Ala Glu
2285	2290	2295
Met Ser Phe Val Pro Lys Asp	Lys Ile Leu Glu Ile	Leu Lys Leu
2300	2305	2310
Asp Gln Phe Leu Thr Leu Met	Ile Gln Asp Arg Leu	Met Asn Ile
2315	2320	2325
Phe Ser Ser Leu Lys Glu Thr	Ile Tyr His Leu Met	Lys Ser Ser
2330	2335	2340
Phe Ile Leu Asp Asn Gly Glu	Phe Tyr Phe Asp Thr	His Gln Gly
2345	2350	2355
Leu Lys Phe Met Gln Asp Leu	Phe Asn Ala Leu Leu	Arg Glu Thr
2360	2365	2370
Ser Met Lys Asn Lys Thr Glu	Asn Asn Ile Asp Phe	Phe Thr Val
2375	2380	2385
Val Ser Gln Leu Phe Phe His	Val Asn Lys Ser Glu	Asp Leu Phe
2390	2395	2400
Lys Leu Asn Gln Asp Leu Gly	Ser Ala Leu His Leu	Val Arg Glu
2405	2410	2415

Cys Ser Thr Glu Met Ala Arg	Leu Leu Asp Thr Ile	Leu His Ser
2420	2425	2430
Pro Asn Lys Asp Phe Tyr Ala	Leu Tyr Pro Thr Leu	Gln Glu Val
2435	2440	2445
Ile Leu Ala Asn Leu Thr Asp	Leu Leu Phe Phe Ile	Asn Asn Ser
2450	2455	2460
Phe Pro Leu Arg Asn Arg Ala	Thr Leu Glu Ile Thr	Lys Arg Leu
2465	2470	2475
Val Gly Ala Ile Ser Arg Ala	Ser Glu Glu Ser His	Val Leu Lys
2480	2485	2490
Pro Leu Leu Glu Met Ser Gly	Thr Leu Val Met Leu	Leu Asn Asp
2495	2500	2505
Ser Ala Asp Leu Arg Asp Leu	Ala Thr Ser Met Asp	Ser Ile Val
2510	2515	2520
Lys Leu Leu Lys Leu Val Lys	Lys Val Ser Gly Lys	Met Ser Thr
2525	2530	2535
Val Phe Lys Thr His Phe Ile	Ser Asn Thr Lys Asp	Ser Val Lys
2540	2545	2550
Phe Phe Asp Thr Leu Tyr Ser	Ile Met Gln Gln Ser	Val Gln Asn
2555	2560	2565
Leu Val Lys Glu Ile Ala Thr	Leu Lys Lys Ile Asp	His Phe Thr
2570	2575	2580
Phe Glu Lys Ile Asn Asp Leu	Leu Val Pro Phe Leu	Asp Leu Ala
2585	2590	2595
Phe Glu Met Ile Gly Val Glu	Pro Tyr Ile Ser Ser	Asn Ser Asp
2600	2605	2610
Ile Phe Ser Met Ser Pro Ser	Ile Leu Ser Tyr Met	Asn Gln Ser
2615	2620	2625
Lys Asp Phe Ser Asp Ile Leu	Glu Glu Ile Ala Glu	Phe Leu Thr
2630	2635	2640

Ser Val 2645	Lys Met Asn Leu Glu 2650	Asp Met Arg Ser Leu 2655	Ala Val Ala
Phe Asn 2660	Asn Glu Thr Gln Thr 2665	Phe Ser Met Asp Ser 2670	Val Asn Leu
Xaa Glu 2675	Glu Ile Leu Gly Cys 2680	Leu Val Pro Ile Asn 2685	Asn Ile Thr
Asn Gln 2690	Met Asp Phe Leu Tyr 2695	Pro Asn Pro Ile Ser 2700	Thr His Ser
Gly Pro 2705	Gln Asp Ile Lys Trp 2710	Glu Ile Ile His Glu 2715	Val Ile Leu
Phe Leu 2720	Asp Lys Ile Leu Ser 2725	Gln Asn Ser Thr Glu 2730	Ile Gly Ser
Phe Leu 2735	Lys Met Val Ile Cys 2740	Leu Thr Leu Glu Ala 2745	Leu Trp Lys
Asn Leu 2750	Lys Lys Asp Asn Trp 2755	Asn Val Ser Asn Val 2760	Leu Met Thr
Phe Thr 2765	Gln His Pro Asn Asn 2770	Leu Leu Lys Thr Ile 2775	Glu Thr Val
Leu Glu 2780	Ala Ser Ser Gly Ile 2785	Lys Ser Asp Tyr Glu 2790	Gly Asp Leu
Asn Lys 2795	Ser Leu Tyr Phe Asp 2800	Thr Pro Leu Ser Gln 2805	Asn Ile Thr
His His 2810	Gln Leu Glu Lys Ala 2815	Ile His Asn Val Leu 2820	Ser Arg Ile
Ala Leu 2825	Trp Arg Lys Gly Leu 2830	Arg Phe Asn Asn Ser 2835	Glu Trp Ile
Thr Ser 2840	Thr Arg Thr Leu Phe 2845	Gln Pro Leu Phe Glu 2850	Ile Phe Ile
Lys Ala 2855	Thr Thr Gly Lys Asn 2860	Val Thr Ser Glu Lys 2865	Glu Glu Arg

Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	Pro	Tyr	Ser	Phe	Lys	Pro	Phe
2870						2875					2880			
Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	Gly	Leu	Phe	Val	Leu	Thr	Lys
2885						2890					2895			
Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	Asp	Gln	Ser	Val	Val	Glu	Ile
2900						2905					2910			
Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	Lys	Pro	Ser	Glu	Ala	Met	Glu
2915						2920					2925			
Met	Leu	Gln	Lys	Val	Lys	Met	Met	Val	Val	Arg	Val	Leu	Thr	Ile
2930						2935					2940			
Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	Lys	Asp	Ile	Leu	Cys	Ala	Thr
2945						2950					2955			
Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	Arg	His	Leu	Ile	Leu	Ser	Ala
2960						2965					2970			
Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	Asp	His	Phe	Gln	Glu	Ile	Glu
2975						2980					2985			
Lys	Ile	Trp	Ser	Ser	Pro	Asn	Gln	Leu	Asn	Cys	Glu	Ser	Leu	Ser
2990						2995					3000			
Lys	Asn	Leu	Ser	Ser	Thr	Leu	Glu	Ser	Phe	Lys	Ser	Ser	Leu	Glu
3005						3010					3015			
Asn	Ala	Thr	Gly	Gln	Asp	Cys	Thr	Ser	Gln	Pro	Arg	Leu	Glu	Thr
3020						3025					3030			
Val	Gln	Gln	His	Leu	Tyr	Met	Leu	Ala	Lys	Ser	Leu	Glu	Glu	Thr
3035						3040					3045			
Trp	Ser	Ser	Gly	Asn	Pro	Ile	Met	Thr	Phe	Leu	Ser	Asn	Phe	Thr
3050						3055					3060			
Val	Thr	Glu	Asp	Val	Lys	Ile	Lys	Asp	Leu	Met	Lys	Asn	Ile	Thr
3065						3070					3075			
Lys	Leu	Thr	Glu	Glu	Leu	Arg	Ser	Ser	Ile	Gln	Ile	Ser	Asn	Glu
3080						3085					3090			
Thr	Ile	His	Ser	Ile	Leu	Glu	Ala	Asn	Ile	Ser	His	Ser	Lys	Val

3095		3100		3105
Leu Phe Ser Ala Leu Thr Val Ala Leu Ser Gly Lys Cys Asp Gln				
3110		3115		3120
Glu Ile Leu His Leu Leu Leu Thr Phe Pro Lys Gly Glu Lys Ser				
3125		3130		3135
Trp Ile Ala Ala Glu Glu Leu Cys Ser Leu Pro Gly Ser Lys Val				
3140		3145		3150
Tyr Ser Leu Ile Val Leu Leu Ser Arg Asn Leu Asp Val Arg Ala				
3155		3160		3165
Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu Ala Asn Gly Leu Leu				
3170		3175		3180
Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser Ala Leu Leu Ala				
3185		3190		3195
Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe Leu His Thr				
3200		3205		3210
Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln Gln Val				
3215		3220		3225
Ser Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe Gln				
3230		3235		3240
Phe Val Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser				
3245		3250		3255
Asp Ser Asn Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu				
3260		3265		3270
Glu Asp Asp Lys Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro				
3275		3280		3285
Phe Cys Leu Lys Leu Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly				
3290		3295		3300
Ala Leu Val Trp Thr Phe Leu Lys Pro Ile Leu His Gly Lys Ile				
3305		3310		3315
Leu Tyr Thr Pro Asn Thr Pro Glu Ile Asn Lys Val Ile Gln Lys				
3320		3325		3330

Ala Asn Tyr Thr Phe Tyr Ile	Val Asp Lys Leu Lys	Thr Leu Ser
3335	3340	3345
Glu Thr Leu Leu Glu Met Ser	Ser Leu Phe Gln Arg	Ser Gly Ser
3350	3355	3360
Gly Gln Met Phe Asn Gln Leu	Gln Glu Ala Leu Arg	Asn Lys Phe
3365	3370	3375
Val Arg Asn Phe Val Glu Asn	Gln Leu His Ile Asp	Val Asp Lys
3380	3385	3390
Leu Thr Glu Lys Leu Gln Thr	Tyr Gly Gly Leu Leu	Asp Glu Met
3395	3400	3405
Phe Asn His Ala Gly Ala Gly	Arg Phe Arg Phe Leu	Gly Ser Ile
3410	3415	3420
Leu Val Asn Leu Ser Ser Cys	Val Ala Leu Asn Arg	Phe Gln Ala
3425	3430	3435
Leu Gln Ser Val Asp Ile Leu	Glu Thr Lys Ala His	Glu Leu Leu
3440	3445	3450
Gln Gln Asn Ser Phe Leu Ala	Ser Ile Ile Phe Ser	Asn Ser Leu
3455	3460	3465
Phe Asp Lys Asn Phe Arg Ser	Glu Ser Val Lys Leu	Pro Pro His
3470	3475	3480
Val Ser Tyr Thr Ile Arg Thr	Asn Val Leu Tyr Ser	Val Arg Thr
3485	3490	3495
Asp Val Val Lys Asn Pro Ser	Trp Lys Phe His Pro	Gln Asn Leu
3500	3505	3510
Pro Ala Asp Gly Phe Lys Tyr	Asn Tyr Val Phe Ala	Pro Leu Gln
3515	3520	3525
Asp Met Ile Glu Arg Ala Ile	Ile Leu Val Gln Thr	Gly Gln Glu
3530	3535	3540
Ala Leu Glu Pro Ala Ala Gln	Thr Gln Ala Ala Pro	Tyr Pro Cys
3545	3550	3555

His Thr Ser Asp Leu Phe Leu Asn Asn Val Gly Phe Phe Phe Pro
 3560 3565 3570

Leu Ile Met Met Leu Thr Trp Met Val Ser Val Ala Ser Met Val
 3575 3580 3585

Arg Lys Leu Val Tyr Glu Gln Glu Ile Gln Ile Glu Glu Tyr Met
 3590 3595 3600

Arg Met Met Gly Val His Pro Val Ile His Phe Leu Ala Trp Phe
 3605 3610 3615

Leu Glu Asn Met Ala Val Leu Thr Ile Ser Ser Ala Thr Leu Ala
 3620 3625 3630

Ile Val Leu Lys Thr Ser Gly Ile Phe Ala His Ser Asn Thr Phe
 3635 3640 3645

Ile Val Phe Leu Phe Leu Leu Asp Phe Gly Met Ser Val Val Met
 3650 3655 3660

Leu Ser Tyr Leu Leu Ser Ala Phe Phe Ser Gln Ala Asn Thr Ala
 3665 3670 3675

Ala Leu Cys Thr Ser Leu Val Tyr Met Ile Ser Phe Leu Pro Tyr
 3680 3685 3690

Ile Val Leu Leu Val Leu His Asn Gln Leu Ser Phe Val Asn Gln
 3695 3700 3705

Thr Phe Leu Cys Leu Leu Ser Thr Thr Ala Phe Gly Gln Gly Val
 3710 3715 3720

Phe Phe Ile Thr Phe Leu Glu Gly Gln Glu Thr Gly Ile Gln Trp
 3725 3730 3735

Asn Asn Met Tyr Gln Ala Leu Glu Gln Gly Gly Met Thr Phe Gly
 3740 3745 3750

Trp Val Cys Trp Met Ile Leu Phe Asp Ser Ser Leu Tyr Phe Leu
 3755 3760 3765

Cys Gly Trp Tyr Leu Ser Asn Leu Ile Pro Gly Thr Phe Gly Leu
 3770 3775 3780

Arg Lys	Pro Trp Tyr Phe	Pro	Phe Thr Ala Ser Tyr	Trp Lys Ser
3785		3790		3795
Val Gly	Phe Leu Val Glu Lys	Arg Gln Tyr Phe	Leu	Ser Ser Ser
3800		3805		3810
Leu Phe	Phe Phe Asn Glu Asn	Phe Asp Asn Lys	Gly	Ser Ser Leu
3815		3820		3825
Gln Asn	Arg Glu Gly Glu Leu	Glu Gly Ser Ala	Pro	Gly Val Thr
3830		3835		3840
Leu Val	Ser Val Thr Lys Glu	Tyr Glu Gly His	Lys	Ala Val Val
3845		3850		3855
Gln Asp	Leu Ser Leu Thr Phe	Tyr Arg Asp Gln	Ile	Thr Ala Leu
3860		3865		3870
Leu Gly	Thr Asn Gly Ala Gly	Lys Thr Thr Ile	Ile	Ser Met Leu
3875		3880		3885
Thr Gly	Leu His Pro Pro Thr	Ser Gly Thr Ile	Ile	Ile Asn Gly
3890		3895		3900
Lys Asn	Leu Gln Thr Asp Leu	Ser Arg Val Arg	Met	Glu Leu Gly
3905		3910		3915
Val Cys	Pro Gln Gln Asp Ile	Leu Leu Asp Asn	Leu	Thr Val Arg
3920		3925		3930
Glu His	Leu Leu Leu Phe Ala	Ser Ile Lys Ala	Pro	Gln Trp Thr
3935		3940		3945
Lys Lys	Glu Leu His Gln Gln	Val Asn Gln Thr	Leu	Gln Asp Val
3950		3955		3960
Asp Leu	Thr Gln His Gln His	Lys Gln Thr Arg	Ala	Leu Ser Gly
3965		3970		3975
Gly Leu	Lys Arg Lys Leu Ser	Leu Gly Ile Ala	Phe	Met Gly Met
3980		3985		3990
Ser Arg	Thr Val Val Leu Asp	Glu Pro Thr Ser	Gly	Val Asp Pro
3995		4000		4005
Cys Ser	Arg His Ser Leu Trp	Asp Ile Leu Leu Lys	Tyr	Arg Glu

4010		4015		4020
Gly Arg Thr Ile Ile Phe Thr Thr His His Leu Asp Glu Ala Glu				
4025		4030		4035
Ala Leu Ser Asp Arg Val Ala Val Leu Gln His Gly Arg Leu Arg				
4040		4045		4050
Cys Cys Gly Pro Pro Phe Cys Leu Lys Glu Ala Tyr Gly Gln Gly				
4055		4060		4065
Leu Arg Leu Thr Leu Thr Arg Gln Pro Ser Val Leu Glu Ala His				
4070		4075		4080
Asp Leu Lys Asp Met Ala Cys Val Thr Ser Leu Ile Lys Ile Tyr				
4085		4090		4095
Ile Pro Gln Ala Phe Leu Lys Asp Ser Ser Gly Ser Glu Leu Thr				
4100		4105		4110
Tyr Thr Ile Pro Lys Asp Thr Asp Lys Ala Cys Leu Lys Gly Leu				
4115		4120		4125
Phe Gln Ala Leu Asp Glu Asn Leu His Gln Leu His Leu Thr Gly				
4130		4135		4140
Tyr Gly Ile Ser Asp Thr Thr Leu Glu Glu Val Phe Leu Met Leu				
4145		4150		4155
Leu Gln Asp Ser Asn Lys Lys Ser His Ile Ala Leu Gly Thr Glu				
4160		4165		4170
Ser Glu Leu Gln Asn His Arg Pro Thr Gly His Leu Ser Gly Tyr				
4175		4180		4185
Cys Gly Ser Leu Ala Arg Pro Ala Thr Val Gln Gly Val Gln Leu				
4190		4195		4200
Leu Arg Ala Gln Val Ala Ala Ile Leu Ala Arg Arg Leu Arg Arg				
4205		4210		4215
Thr Leu Arg Ala Gly Lys Ser Thr Leu Ala Asp Leu Leu Leu Pro				
4220		4225		4230
Val Leu Phe Val Ala Leu Ala Met Gly Leu Phe Met Val Arg Pro				
4235		4240		4245

Leu Ala	Thr Glu Tyr Pro Pro	Leu Arg Leu Thr Pro	Gly His Tyr
4250	4255	4260	
Gln Arg	Ala Glu Thr Tyr Phe	Phe Ser Ser Gly Gly	Asp Asn Leu
4265	4270	4275	
Asp Leu	Thr Arg Val Leu Leu	Arg Lys Phe Arg Asp	Gln Asp Leu
4280	4285	4290	
Pro Cys	Ala Asp Leu Asn Pro	Arg Gln Lys Asn Ser	Ser Cys Trp
4295	4300	4305	
Arg Thr	Asp Pro Phe Ser His	Pro Glu Phe Gln Asp	Ser Cys Gly
4310	4315	4320	
Cys Leu	Lys Cys Pro Asn Arg	Ser Ala Ser Ala Pro	Tyr Leu Thr
4325	4330	4335	
Asn His	Leu Gly His Thr Leu	Leu Asn Leu Ser Gly	Phe Asn Met
4340	4345	4350	
Glu Glu	Tyr Leu Leu Ala Pro	Ser Glu Lys Pro Arg	Leu Gly Gly
4355	4360	4365	
Trp Ser	Phe Gly Leu Lys Ile	Pro Ser Glu Ala Gly	Gly Ala Asn
4370	4375	4380	
Gly Asn	Ile Ser Lys Pro Pro	Thr Leu Ala Lys Val	Trp Tyr Asn
4385	4390	4395	
Gln Lys	Gly Phe His Ser Leu	Pro Ser Tyr Leu Asn	His Leu Asn
4400	4405	4410	
Asn Leu	Ile Leu Trp Gln His	Leu Pro Pro Thr Val	Asp Trp Arg
4415	4420	4425	
Gln Tyr	Gly Ile Thr Leu Tyr	Ser His Pro Tyr Gly	Gly Ala Leu
4430	4435	4440	
Leu Asn	Glu Asp Lys Ile Leu	Glu Ser Ile Arg Gln	Cys Gly Val
4445	4450	4455	
Ala Leu	Cys Ile Val Leu Gly	Phe Ser Ile Leu Ser	Ala Ser Ile
4460	4465	4470	

Gly Ser Ser Val Val Arg Asp Arg Val Ile Gly Ala Lys Arg Leu
 4475 4480 4485

 Gln His Ile Ser Gly Leu Gly Tyr Arg Met Tyr Trp Phe Thr Asn
 4490 4495 4500

 Phe Leu Tyr Asp Met Leu Phe Tyr Leu Val Ser Val Cys Leu Cys
 4505 4510 4515

 Val Ala Val Ile Val Ala Phe Gln Leu Thr Ala Phe Thr Phe Arg
 4520 4525 4530

 Lys Asn Leu Ala Ala Thr Ala Leu Leu Leu Ser Leu Phe Gly Tyr
 4535 4540 4545

 Ala Thr Leu Pro Trp Met Tyr Leu Met Ser Arg Ile Phe Ser Ser
 4550 4555 4560

 Ser Asp Val Ala Phe Ile Ser Tyr Val Ser Leu Asn Phe Ile Phe
 4565 4570 4575

 Gly Leu Cys Thr Met Pro Ile Thr Ile Met Pro Arg Leu Leu Ala
 4580 4585 4590

 Ile Ile Ser Lys Ala Lys Asn Leu Gln Asn Ile Tyr Asp Val Leu
 4595 4600 4605

 Lys Trp Val Phe Thr Ile Phe Pro Gln Phe Cys Leu Gly Gln Gly
 4610 4615 4620

 Leu Val Glu Leu Cys Tyr Asn Gln Ile Lys Tyr Asp Leu Thr His
 4625 4630 4635

 Asn Phe Gly Ile Asp Ser Tyr Val Ser Pro Phe Glu Met Asn Phe
 4640 4645 4650

 Leu Gly Trp Ile Phe Val Gln Leu Ala Ser Gln Gly Thr Val Leu
 4655 4660 4665

 Leu Leu Leu Arg Val Leu Leu His Trp Asp Leu Leu Arg Trp Pro
 4670 4675 4680

 Arg Gly His Ser Thr Leu Gln Gly Thr Val Lys Ser Ser Lys Asp
 4685 4690 4695

Thr	Asp	Val	Glu	Lys	Glu	Glu	Lys	Arg	Val	Phe	Glu	Gly	Arg	Thr
4700						4705					4710			
Asn	Gly	Asp	Ile	Leu	Val	Leu	Tyr	Asn	Leu	Ser	Lys	His	Tyr	Arg
4715						4720					4725			
Arg	Phe	Phe	Gln	Asn	Ile	Ile	Ala	Val	Gln	Asp	Ile	Ser	Leu	Gly
4730						4735					4740			
Ile	Pro	Lys	Gly	Glu	Cys	Phe	Gly	Leu	Leu	Gly	Val	Asn	Gly	Ala
4745						4750					4755			
Gly	Lys	Ser	Thr	Thr	Phe	Lys	Met	Leu	Asn	Gly	Glu	Val	Ser	Leu
4760						4765					4770			
Thr	Ser	Gly	His	Ala	Ile	Ile	Arg	Thr	Pro	Met	Gly	Asp	Ala	Val
4775						4780					4785			
Asp	Leu	Ser	Ser	Ala	Gly	Thr	Ala	Gly	Val	Leu	Ile	Gly	Tyr	Cys
4790						4795					4800			
Pro	Gln	Gln	Asp	Ala	Leu	Asp	Glu	Leu	Leu	Thr	Gly	Trp	Glu	His
4805						4810					4815			
Leu	Tyr	Tyr	Tyr	Cys	Ser	Leu	Arg	Gly	Ile	Pro	Arg	Gln	Cys	Ile
4820						4825					4830			
Pro	Glu	Val	Ala	Gly	Asp	Leu	Ile	Arg	Arg	Leu	His	Leu	Glu	Ala
4835						4840					4845			
His	Ala	Asp	Lys	Pro	Val	Ala	Thr	Tyr	Ser	Gly	Gly	Thr	Lys	Arg
4850						4855					4860			
Lys	Leu	Ser	Thr	Ala	Leu	Ala	Leu	Val	Gly	Lys	Pro	Asp	Ile	Leu
4865						4870					4875			
Leu	Leu	Asp	Glu	Pro	Ser	Ser	Gly	Met	Asp	Pro	Cys	Ser	Lys	Arg
4880						4885					4890			
Tyr	Leu	Trp	Gln	Thr	Ile	Met	Lys	Glu	Val	Arg	Glu	Gly	Cys	Ala
4895						4900					4905			
Ala	Val	Leu	Thr	Ser	His	Ser	Met	Glu	Glu	Cys	Glu	Ala	Leu	Cys
4910						4915					4920			

Thr Arg Leu Ala Ile Met Val Asn Gly Ser Phe Lys Cys Leu Gly
 4925 4930 4935

Ser Pro Gln His Ile Lys Asn Arg Cys Val Glu Val Leu Leu Leu
 4940 4945 4950

Leu Arg Tyr Leu Leu
 4955

<210> 24
 <211> 11259
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> misc_feature
 <222> (873)..(873)
 <223> Xaa = any amino acid

<220>
 <221> CDS
 <222> (2209)..(10191)

<220>
 <221> misc_feature
 <222> (3334)..(3334)
 <223> n = any nucleic acid residue

<400> 24
 ggactgagag cagggagcag caggcatggg gcatgccggg tgccagttca aagccctgct 60
 gtggaagaat tggctctgca gactcaggaa cccggtcctt ttccttgctg aattcttctg 120
 gccttgtatc ctgtttgtaa ttctgacagt tcttggtttt caagaacctc ccagatacag 180
 agacatttgt tatttgcagc cccgagatct acccagctgt ggtgttatcc ctttgtttca 240
 aagccttctt tgtaacactg gatcaagggtg taggaacttc agctatgaag ggtcaatgga 300
 gcatcatttt cgtttgtcta gggtccaaac tgcagctgac cccaagaaag tcaacaacct 360
 ggccttttta aaagagatac aagacctggc agaggaaatt catggaatga tggacaaggc 420
 aaaaaactta aaaagacttt gggtagaacg atccaacact ccagattott cttatggttc 480
 cagttttttt acaatggatc tcaataagac cgaggaggta atattgaaac tggaaagcct 540
 ccatcagcag cctcatatct gggattttct acttttactg ccgagactac acacaagcca 600
 tgatcatgtg gaagatggca tggatgttgc agtgaacctt ctccagacca ttttgaattc 660
 cttaatatcc ctagaagatt tagattggct tccactcaac caaacttttt ccaggttttc 720
 tgaacttgta ctgaatgtga ccatttcgac actgacattt ctgcagcaac atggagtagc 780
 agtcaccgag ccagttttacc acctgtccat gcagaatata gtgtgggata cacagaaagt 840

ccagtatgat ctcaaatccc agtttggctt tgntgatctt cacacggaac agatcctgaa	900
ctcttcagct gaactgaagg agattcccac agacacttcc ttggagaaga tgggtgtgttc	960
agtcttgtct agcacatcag aggatgaagc tgagaaatgg ggccacgttg gaggctgcca	1020
ccctaagtgg tcagaagcca aaaactatct tgtccatgca gtcagctggc tgcgagtcta	1080
ccaacagggtg tttgttcagt ggcaacaggg tagcctgctt cagaagacac tcacaggcat	1140
gggccatagt ctggaggctc tcaggaatca gtttgaagaa gagagcaagc cctggaaggt	1200
ggtggaagct ctgcacactg cactgctcct gctgaatgac agcttgtcag cagatggccc	1260
aaaagataat catacatttc caaagatatt acagcatctg tggaaattgc aaagcttget	1320
gcaaaacctg cccagtggc cggcactgaa gagatttctt cagcttgatg gagctctcag	1380
aaatgcgata gctcagaatt tacattttgt ccaagaagtc ctcatctgcc tggagacatc	1440
agctaattgat tttaaatggt ttgaacttaa ccaattgaaa ctggaaaagg atgtgttctt	1500
ttgggagctg aaacagatgt tggcgaagaa tgctgtctgc ccgaatggtc gtttctctga	1560
gaaggaggctc tttttgccgc ctggaaactc cagcatatgg ggtggtctcc agggactgtt	1620
gtgctattgt aactcctctg agacgagtgt tttaaacaag ctacttggtt cagtagagga	1680
tgctgatcgt attttgcaag aggtcattac ttggcacaaa aatatgtcag ttttaatacc	1740
tgaagaatat ttggactggc aggaacttga gatgcagctg tcagaagcaa gcctttcctg	1800
tactcggctc ttctgtctgc tgggagctga tccctctcct gagaatgatg tcttttctag	1860
tgactgtaag caccagcttg tctccacagt gatatttcat acacttgaaa aaacacaatt	1920
tttcttgcaa caagcatatt attggaaagc cttcaaaaag tttatcagga agacttgcca	1980
agtggcccaa tatgtaaata tgcaagagag tttccagaac agactattgg cttttcctga	2040
ggaatctcct tgttttgaag aaaacatgga ttggaaaatg atcagtgata attattttca	2100
atttttgaat aacttactca agtctccaac agcttcata tccaggaaga ttctgagaac	2160
aaaatatctc ttctgctgaa atatttccac aaagatgtta ttgcagag atg agt ttt	2217
	Met Ser Phe
	1
gtc cca aaa gat aaa att cta gaa att ctg aaa ctg gat caa ttt ctt	2265
Val Pro Lys Asp Lys Ile Leu Glu Ile Leu Lys Leu Asp Gln Phe Leu	
5 10 15	
acc ctg atg ata caa gac aga ttg atg aac att ttt tca agt tta aag	2313
Thr Leu Met Ile Gln Asp Arg Leu Met Asn Ile Phe Ser Ser Leu Lys	
20 25 30 35	
gag act ata tat cac cta atg aaa agt tca ttt ata tta gac aat gga	2361
Glu Thr Ile Tyr His Leu Met Lys Ser Phe Ile Leu Asp Asn Gly	
40 45 50	
gaa ttt tat ttt gat act cat caa gga ctg aag ttc atg caa gat tta	2409

Glu	Phe	Tyr	Phe	Asp	Thr	His	Gln	Gly	Leu	Lys	Phe	Met	Gln	Asp	Leu	
			55					60					65			
ttt	aat	gcc	ctt	ctc	agg	gaa	act	tca	atg	aaa	aat	aag	act	gaa	aat	2457
Phe	Asn	Ala	Leu	Leu	Arg	Glu	Thr	Ser	Met	Lys	Asn	Lys	Thr	Glu	Asn	
		70					75					80				
aat	ata	gac	ttt	ttc	aca	gtg	gtg	agt	cag	ttg	ttt	ttc	cat	gtg	aat	2505
Asn	Ile	Asp	Phe	Phe	Thr	Val	Val	Ser	Gln	Leu	Phe	Phe	His	Val	Asn	
	85					90					95					
aag	tct	gag	gac	ctc	ttc	aaa	ctc	aat	caa	gat	ctt	ggg	tca	gct	ctt	2553
Lys	Ser	Glu	Asp	Leu	Phe	Lys	Leu	Asn	Gln	Asp	Leu	Gly	Ser	Ala	Leu	
100					105					110					115	
cac	ctt	gta	aga	gaa	tgt	tca	aca	gag	atg	gca	aga	ctt	ctg	gat	aca	2601
His	Leu	Val	Arg	Glu	Cys	Ser	Thr	Glu	Met	Ala	Arg	Leu	Leu	Asp	Thr	
				120					125					130		
att	tta	cac	tct	cct	aat	aag	gac	ttc	tat	gct	ttg	tat	cct	acc	ctc	2649
Ile	Leu	His	Ser	Pro	Asn	Lys	Asp	Phe	Tyr	Ala	Leu	Tyr	Pro	Thr	Leu	
			135					140					145			
caa	gaa	gtt	ata	ctt	gct	aat	cta	acg	gat	ttg	ctt	ttc	ttt	ata	aat	2697
Gln	Glu	Val	Ile	Leu	Ala	Asn	Leu	Thr	Asp	Leu	Leu	Phe	Phe	Ile	Asn	
		150					155					160				
aat	tca	ttc	cct	cta	aga	aac	aga	gca	aca	tta	gaa	att	act	aag	aga	2745
Asn	Ser	Phe	Pro	Leu	Arg	Asn	Arg	Ala	Thr	Leu	Glu	Ile	Thr	Lys	Arg	
	165					170					175					
tta	gtt	ggt	gct	att	tca	aga	gca	agt	gaa	gaa	agt	cac	gtc	ctg	aaa	2793
Leu	Val	Gly	Ala	Ile	Ser	Arg	Ala	Ser	Glu	Glu	Ser	His	Val	Leu	Lys	
180					185					190					195	
ccc	ctc	tta	gaa	atg	tct	ggg	act	ctg	gtc	atg	ctg	ttg	aat	gac	agt	2841
Pro	Leu	Leu	Glu	Met	Ser	Gly	Thr	Leu	Val	Met	Leu	Leu	Asn	Asp	Ser	
				200					205					210		
gct	gac	ctg	aga	gat	ctt	gcc	aca	tca	atg	gac	tcc	att	gtg	aaa	ctt	2889
Ala	Asp	Leu	Arg	Asp	Leu	Ala	Thr	Ser	Met	Asp	Ser	Ile	Val	Lys	Leu	
			215					220					225			
ctt	aag	ctg	gtc	aag	aaa	gtt	tcg	ggg	aag	atg	tcc	aca	gtt	ttt	aaa	2937
Leu	Lys	Leu	Val	Lys	Lys	Val	Ser	Gly	Lys	Met	Ser	Thr	Val	Phe	Lys	
		230					235					240				
act	cat	ttt	atc	tcc	aat	acc	aag	gac	agt	gtg	aaa	ttc	ttt	gac	act	2985
Thr	His	Phe	Ile	Ser	Asn	Thr	Lys	Asp	Ser	Val	Lys	Phe	Phe	Asp	Thr	
	245					250					255					
ctg	tat	tcc	atc	atg	caa	caa	agt	gtt	caa	aat	ctt	gtg	aaa	gaa	ata	3033
Leu	Tyr	Ser	Ile	Met	Gln	Gln	Ser	Val	Gln	Asn	Leu	Val	Lys	Glu	Ile	
260					265					270					275	
gct	act	tta	aaa	aaa	ata	gat	cat	ttc	aca	ttt	gaa	aag	ata	aat	gat	3081
Ala	Thr	Leu	Lys	Lys	Ile	Asp	His	Phe	Thr	Phe	Glu	Lys	Ile	Asn	Asp	
				280				285						290		
ttg	ttg	gtg	cca	ttt	ctt	gac	ttg	gcc	ttt	gaa	atg	att	ggg	gta	gaa	3129
Leu	Leu	Val	Pro	Phe	Leu	Asp	Leu	Ala	Phe	Glu	Met	Ile	Gly	Val	Glu	

295						300						305						
cct Pro	tat Tyr	ata Ile 310	tca Ser	tca Ser	aac Asn	tct Ser	gat Asp 315	att Ile	ttc Phe	agt Ser	atg Met	tca Ser 320	cct Pro	agc Ser	ata Ile	3177		
ctc Leu	tca Ser 325	tat Tyr	atg Met	aac Asn	caa Gln	tct Ser 330	aag Lys	gac Asp	ttt Phe	tct Ser	gat Asp 335	att Ile	ttg Leu	gaa Glu	gaa Glu	3225		
att Ile 340	gct Ala	gaa Glu	ttt Phe	tta Leu	aca Thr 345	tct Ser	gtg Val	aaa Lys	atg Met	aac Asn 350	ttg Leu	gaa Glu	gat Asp	atg Met	agg Arg 355	3273		
agt Ser	ctt Leu	gcg Ala	gta Val	gca Ala 360	ttt Phe	aac Asn	aat Asn	gag Glu	act Thr 365	caa Gln	aca Thr	ttt Phe	tct Ser	atg Met	gat Asp	3321		
tct Ser	gtc Val	aac Asn	tta Leu 375	ngg Xaa	gaa Glu	gaa Glu	att Ile	ctg Leu 380	ggg Gly	tgc Cys	tta Leu	gtt Val	cct Pro 385	ata Ile	aat Asn	3369		
aac Asn	atc Ile	acc Thr 390	aac Asn	caa Gln	atg Met	gac Asp	ttc Phe 395	tta Leu	tac Tyr	cct Pro	aat Asn	cca Pro 400	att Ile	tcc Ser	act Thr	3417		
cat His	agt Ser 405	ggc Gly	cct Pro	caa Gln	gat Asp	ata Ile 410	aaa Lys	tgg Trp	gaa Glu	ata Ile	att Ile 415	cat His	gaa Glu	gtg Val	atc Ile	3465		
ctt Leu 420	ttt Phe	ttg Leu	gat Asp	aaa Lys	ata Ile 425	tta Leu	tca Ser	caa Gln	aac Asn	agc Ser 430	aca Thr	gaa Glu	ata Ile	gga Gly	tct Ser 435	3513		
ttc Phe	ttg Leu	aaa Lys	atg Met	gtg Val 440	atc Ile	tgt Cys	ctc Leu	acc Thr	tta Leu 445	gaa Glu	gct Ala	ctt Leu	tgg Trp	aaa Lys 450	aac Asn	3561		
tta Leu	aag Lys	aaa Lys	gat Asp 455	aat Asn	tgg Trp	aat Asn	gtt Val	tct Ser 460	aat Asn	gtg Val	ttg Leu	atg Met	acg Thr 465	ttt Phe	act Thr	3609		
cag Gln	cat His	cca Pro 470	aat Asn	aac Asn	ctt Leu	ttg Leu	aaa Lys 475	acc Thr	ata Ile	gaa Glu	aca Thr	gtt Val 480	tta Leu	gag Glu	gcc Ala	3657		
tcc Ser	agt Ser 485	gga Gly	att Ile	aaa Lys	agt Ser	gac Asp 490	tat Tyr	gaa Glu	ggg Gly	gat Asp	ttg Leu 495	aat Asn	aaa Lys	agt Ser	tta Leu	3705		
tat Tyr 500	ttt Phe	gac Asp	aca Thr	cct Pro	ttg Leu 505	agt Ser	cag Gln	aat Asn	ata Ile	act Thr 510	cat His	cat His	caa Gln	ctt Leu	gaa Glu 515	3753		
aaa Lys	gca Ala	atc Ile	cat His	aat Asn 520	gtt Val	tta Leu	agt Ser	aga Arg	ata Ile 525	gct Ala	ctc Leu	tgg Trp	agg Arg	aaa Lys 530	gga Gly	3801		
ctt Leu	cgt Arg	ttt Phe	aac Asn 535	aac Asn	tct Ser	gaa Glu	tgg Trp	ata Ile 540	act Thr	tcc Ser	aca Thr	aga Arg	act Thr 545	ttg Leu	ttt Phe	3849		

cag cca ctt ttt gag att ttc att aaa gca acc acc gga aag aat gtc	3897
Gln Pro Leu Phe Glu Ile Phe Ile Lys Ala Thr Thr Gly Lys Asn Val	
550 555 560	
aca tca gaa aaa gaa gag aga acc gag aaa gag atg att gac ttt cct	3945
Thr Ser Glu Lys Glu Glu Arg Thr Glu Lys Glu Met Ile Asp Phe Pro	
565 570 575	
tat agt ttc aaa cca ttt ttc tgt ttg gag aaa tac ctg gga gga tta	3993
Tyr Ser Phe Lys Pro Phe Phe Cys Leu Glu Lys Tyr Leu Gly Gly Leu	
580 585 590 595	
ttt gta ttg act aaa tac tgg caa caa atc cca cta aca gat caa agt	4041
Phe Val Leu Thr Lys Tyr Trp Gln Gln Ile Pro Leu Thr Asp Gln Ser	
600 605 610	
gtt gtt gag att tgt gaa gtt ttc cag cag act gtg aag ccc tca gaa	4089
Val Val Glu Ile Cys Glu Val Phe Gln Gln Thr Val Lys Pro Ser Glu	
615 620 625	
gcc atg gag atg ctg cag aaa gtg aag atg atg gtc gta cgt gtg ctc	4137
Ala Met Glu Met Leu Gln Lys Val Lys Met Met Val Val Arg Val Leu	
630 635 640	
acc atc gtt gca gaa aac cct tcc tgg acc aag gac att ttg tgt gct	4185
Thr Ile Val Ala Glu Asn Pro Ser Trp Thr Lys Asp Ile Leu Cys Ala	
645 650 655	
act ctg agt tgc aag caa aat ggg ata agg cat ctc att tta tct gct	4233
Thr Leu Ser Cys Lys Gln Asn Gly Ile Arg His Leu Ile Leu Ser Ala	
660 665 670 675	
ata caa ggg gtc act ttg gcg cag gac cac ttc cag gaa att gaa aag	4281
Ile Gln Gly Val Thr Leu Ala Gln Asp His Phe Gln Glu Ile Glu Lys	
680 685 690	
ata tgg tcc tcg ccg aat cag cta aat tgt gaa agt ctt agc aag aat	4329
Ile Trp Ser Ser Pro Asn Gln Leu Asn Cys Glu Ser Leu Ser Lys Asn	
695 700 705	
ctt tct agc acc ttg gag agc ttc aag agc agc ttg gaa aat gcc act	4377
Leu Ser Ser Thr Leu Glu Ser Phe Lys Ser Ser Leu Glu Asn Ala Thr	
710 715 720	
ggc cag gac tgc aca agc cag ccg agg ctg gag acg gtg cag cag cac	4425
Gly Gln Asp Cys Thr Ser Gln Pro Arg Leu Glu Thr Val Gln Gln His	
725 730 735	
ttg tac atg ttg gcc aaa agc ctc gag gaa act tgg tca tca ggg aat	4473
Leu Tyr Met Leu Ala Lys Ser Leu Glu Glu Thr Trp Ser Ser Gly Asn	
740 745 750 755	
ccc atc atg act ttt ctc agc aat ttc aca gta act gag gat gta aaa	4521
Pro Ile Met Thr Phe Leu Ser Asn Phe Thr Val Thr Glu Asp Val Lys	
760 765 770	
ata aaa gat ttg atg aag aat atc acc aag ttg act gag gag ctt cgc	4569
Ile Lys Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu Glu Leu Arg	
775 780 785	

tct tcc atc caa atc tcg aat gag act atc cat agc att cta gaa gca	4617
Ser Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile Leu Glu Ala	
790 795 800	
aat att tcc cac tcc aag gtt ctg ttc agt gcc ctg acc gta gct ctg	4665
Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr Val Ala Leu	
805 810 815	
tct gga aag tgt gat cag gaa atc ctt cat ctg ctg ctg aca ttt ccc	4713
Ser Gly Lys Cys Asp Gln Glu Ile Leu His Leu Leu Leu Thr Phe Pro	
820 825 830 835	
aaa ggg gaa aaa tct tgg atc gca gcg gag gaa ctg tgt agc ctg cca	4761
Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys Ser Leu Pro	
840 845 850	
ggg tca aaa gtg tat tct ctg att gtg ttg ctg agt cga aac ttg gat	4809
Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg Asn Leu Asp	
855 860 865	
gtg cga gct ttc att tac aag act ctg atg cct tct gaa gca aat ggc	4857
Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu Ala Asn Gly	
870 875 880	
ttg ctg aac tcc ttg ctg gat ata gtt tcc agc ctg agc gcc ttg ctt	4905
Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser Ala Leu Leu	
885 890 895	
gcc aaa gcc cag cac gtc ttt gag tat ctt cct gag ttt ctt cac aca	4953
Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe Leu His Thr	
900 905 910 915	
ttt aaa atc act gcc ttg cta gaa acc ctg gac ttt caa cag gtt tca	5001
Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln Gln Val Ser	
920 925 930	
caa aat gtc cag gcc aga agt tca gct ttt ggt tct ttc cag ttt gtg	5049
Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe Gln Phe Val	
935 940 945	
atg aag atg gtt tgc aag gac caa gca tca ttc ctt agc gat tct aat	5097
Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser Asp Ser Asn	
950 955 960	
atg ttt att aat ttg ccc aga gtt aag gaa ctg ttg gaa gat gac aaa	5145
Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Glu Asp Asp Lys	
965 970 975	
gaa aaa ttc aac att cct gaa gat tca aca ccg ttt tgc ttg aag ctt	5193
Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys Leu Lys Leu	
980 985 990 995	
tat cag gaa att cta caa ttg cca aat ggt gct ttg gtg tgg acc	5238
Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val Trp Thr	
1000 1005 1010	
ttc cta aaa ccc ata ttg cat gga aaa ata cta tac aca cca aac	5283
Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr Pro Asn	
1015 1020 1025	
act cca gaa att aac aag gtc att caa aag gct aat tac acc ttt	5328

Thr	Pro	Glu	Ile	Asn	Lys	Val	Ile	Gln	Lys	Ala	Asn	Tyr	Thr	Phe		
				1030					1035					1040		
tat	att	gtg	gac	aaa	cta	aaa	act	tta	tca	gaa	aca	ctg	ctg	gaa		5373
Tyr	Ile	Val	Asp	Lys	Leu	Lys	Thr	Leu	Ser	Glu	Thr	Leu	Leu	Glu		
				1045					1050					1055		
atg	tcc	agc	ctt	ttc	cag	aga	agt	gga	agt	ggc	cag	atg	ttc	aac		5418
Met	Ser	Ser	Leu	Phe	Gln	Arg	Ser	Gly	Ser	Gly	Gln	Met	Phe	Asn		
				1060					1065					1070		
cag	ctg	cag	gag	gcc	ctg	aga	aac	aaa	ttt	gta	aga	aac	ttt	gta		5463
Gln	Leu	Gln	Glu	Ala	Leu	Arg	Asn	Lys	Phe	Val	Arg	Asn	Phe	Val		
				1075					1080					1085		
gaa	aac	cag	ttg	cac	att	gat	gta	gac	aaa	ctt	act	gaa	aaa	ctc		5508
Glu	Asn	Gln	Leu	His	Ile	Asp	Val	Asp	Lys	Leu	Thr	Glu	Lys	Leu		
				1090					1095					1100		
cag	aca	tac	gga	ggg	ctg	ctg	gat	gag	atg	ttt	aac	cat	gca	ggc		5553
Gln	Thr	Tyr	Gly	Gly	Leu	Leu	Asp	Glu	Met	Phe	Asn	His	Ala	Gly		
				1105					1110					1115		
gct	gga	cgc	ttc	cgt	ttc	ttg	ggc	agc	atc	ttg	gtc	aat	ctc	tct		5598
Ala	Gly	Arg	Phe	Arg	Phe	Leu	Gly	Ser	Ile	Leu	Val	Asn	Leu	Ser		
				1120					1125					1130		
tcc	tgc	gtg	gca	ctg	aac	cgt	ttc	cag	gct	ctg	cag	tct	gtc	gac		5643
Ser	Cys	Val	Ala	Leu	Asn	Arg	Phe	Gln	Ala	Leu	Gln	Ser	Val	Asp		
				1135					1140					1145		
atc	ctg	gag	act	aaa	gca	cat	gaa	ctc	ttg	cag	cag	aac	agc	ttc		5688
Ile	Leu	Glu	Thr	Lys	Ala	His	Glu	Leu	Leu	Gln	Gln	Asn	Ser	Phe		
				1150					1155					1160		
ttg	gcc	agt	atc	att	ttc	agc	aat	tcc	tta	ttc	gac	aag	aac	ttc		5733
Leu	Ala	Ser	Ile	Ile	Phe	Ser	Asn	Ser	Leu	Phe	Asp	Lys	Asn	Phe		
				1165					1170					1175		
aga	tca	gag	tct	gtc	aaa	ctg	cca	ccc	cat	gtc	tca	tac	aca	atc		5778
Arg	Ser	Glu	Ser	Val	Lys	Leu	Pro	Pro	His	Val	Ser	Tyr	Thr	Ile		
				1180					1185					1190		
cgg	acc	aat	gtg	tta	tac	agc	gtg	cga	aca	gat	gtg	gta	aaa	aac		5823
Arg	Thr	Asn	Val	Leu	Tyr	Ser	Val	Arg	Thr	Asp	Val	Val	Lys	Asn		
				1195					1200					1205		
cct	tct	tgg	aag	ttc	cac	cct	cag	aat	cta	cca	gct	gat	ggg	ttc		5868
Pro	Ser	Trp	Lys	Phe	His	Pro	Gln	Asn	Leu	Pro	Ala	Asp	Gly	Phe		
				1210					1215					1220		
aaa	tat	aac	tac	gtc	ttt	gcc	cca	ctg	caa	gac	atg	atc	gaa	aga		5913
Lys	Tyr	Asn	Tyr	Val	Phe	Ala	Pro	Leu	Gln	Asp	Met	Ile	Glu	Arg		
				1225					1230					1235		
gcc	atc	att	ttg	gtg	cag	act	ggg	cag	gaa	gcc	ctg	gaa	cca	gca		5958
Ala	Ile	Ile	Leu	Val	Gln	Thr	Gly	Gln	Glu	Ala	Leu	Glu	Pro	Ala		
				1240					1245					1250		
gca	cag	act	cag	gcg	gcc	cct	tac	ccc	tgc	cat	acc	agc	gac	cta		6003
Ala	Gln	Thr	Gln	Ala	Ala	Pro	Tyr	Pro	Cys	His	Thr	Ser	Asp	Leu		

1255	1260	1265	
ttc ctg aac aac gtt ggt ttc ttt ttt cca ctg ata atg atg ctg Phe Leu Asn Asn Val Gly Phe Phe Phe Pro Leu Ile Met Met Leu 1270 1275 1280			6048
acg tgg atg gtg tct gtg gcc agc atg gtc aga aag ttg gtg tat Thr Trp Met Val Ser Val Ala Ser Met Val Arg Lys Leu Val Tyr 1285 1290 1295			6093
gag cag gag ata cag ata gaa gag tat atg cgg atg atg gga gtg Glu Gln Glu Ile Gln Ile Glu Glu Tyr Met Arg Met Met Gly Val 1300 1305 1310			6138
cat cca gtg atc cat ttc ctg gcc tgg ttc ctg gag aac atg gct His Pro Val Ile His Phe Leu Ala Trp Phe Leu Glu Asn Met Ala 1315 1320 1325			6183
gtg ttg acc ata agc agt gct act ctg gcc atc gtt ctg aaa aca Val Leu Thr Ile Ser Ser Ala Thr Leu Ala Ile Val Leu Lys Thr 1330 1335 1340			6228
agt ggc atc ttt gca cac agc aat acc ttt att gtt ttc ctc ttt Ser Gly Ile Phe Ala His Ser Asn Thr Phe Ile Val Phe Leu Phe 1345 1350 1355			6273
ctc ttg gat ttt ggg atg tca gtc gtc atg ctg agc tac ctc ttg Leu Leu Asp Phe Gly Met Ser Val Val Met Leu Ser Tyr Leu Leu 1360 1365 1370			6318
agt gca ttt ttc agc caa gct aat aca gcg gcc ctt tgt acc agc Ser Ala Phe Phe Ser Gln Ala Asn Thr Ala Ala Leu Cys Thr Ser 1375 1380 1385			6363
ctg gtg tac atg atc agc ttt ctg ccc tac ata gtt cta ttg gtt Leu Val Tyr Met Ile Ser Phe Leu Pro Tyr Ile Val Leu Leu Val 1390 1395 1400			6408
cta cat aac caa tta agt ttt gtt aat cag aca ttt ctg tgc ctt Leu His Asn Gln Leu Ser Phe Val Asn Gln Thr Phe Leu Cys Leu 1405 1410 1415			6453
ctt tcg aca acc gcc ttt gga caa ggg gta ttt ttt att aca ttc Leu Ser Thr Thr Ala Phe Gly Gln Gly Val Phe Phe Ile Thr Phe 1420 1425 1430			6498
ctg gaa gga caa gag aca ggg att caa tgg aat aat atg tac cag Leu Glu Gly Gln Glu Thr Gly Ile Gln Trp Asn Asn Met Tyr Gln 1435 1440 1445			6543
gct ctg gaa caa ggg ggc atg aca ttt ggc tgg gtt tgc tgg atg Ala Leu Glu Gln Gly Gly Met Thr Phe Gly Trp Val Cys Trp Met 1450 1455 1460			6588
att ctt ttt gat tca agc ctt tat ttt ttg tgt gga tgg tac ttg Ile Leu Phe Asp Ser Ser Leu Tyr Phe Leu Cys Gly Trp Tyr Leu 1465 1470 1475			6633
agc aac ttg att cct gga aca ttt ggt tta cgg aaa cca tgg tat Ser Asn Leu Ile Pro Gly Thr Phe Gly Leu Arg Lys Pro Trp Tyr 1480 1485 1490			6678

ttc ccc ttt act gcc tca tat tgg aag agt gtg ggt ttc ttg gtg	6723
Phe Pro Phe Thr Ala Ser Tyr Trp Lys Ser Val Gly Phe Leu Val	
1495 1500 1505	
gag aaa agg caa tac ttt cta agt tct agt ctg ttc ttc ttc aat	6768
Glu Lys Arg Gln Tyr Phe Leu Ser Ser Ser Leu Phe Phe Phe Asn	
1510 1515 1520	
gag aac ttt gac aat aaa ggg tca tca ctg caa aac agg gaa gga	6813
Glu Asn Phe Asp Asn Lys Gly Ser Ser Leu Gln Asn Arg Glu Gly	
1525 1530 1535	
gag ctt gaa gga agt gcc ccg gga gtc acc ctg gtg tct gtg acc	6858
Glu Leu Glu Gly Ser Ala Pro Gly Val Thr Leu Val Ser Val Thr	
1540 1545 1550	
aag gaa tat gag ggc cac aag gct gtg gtc caa gac ctc agc ctg	6903
Lys Glu Tyr Glu Gly His Lys Ala Val Val Gln Asp Leu Ser Leu	
1555 1560 1565	
acc ttc tac aga gac caa atc acc gcc ctg ctg ggg aca aac ggt	6948
Thr Phe Tyr Arg Asp Gln Ile Thr Ala Leu Leu Gly Thr Asn Gly	
1570 1575 1580	
gcc ggg aaa acc act atc ata tcc atg ttg acg ggg ctc cac cct	6993
Ala Gly Lys Thr Thr Ile Ile Ser Met Leu Thr Gly Leu His Pro	
1585 1590 1595	
ccc act tct gga acc atc atc atc aat ggc aag aac cta cag aca	7038
Pro Thr Ser Gly Thr Ile Ile Ile Asn Gly Lys Asn Leu Gln Thr	
1600 1605 1610	
gac ctg tcg agg gtc aga atg gag ctt ggt gtg tgt ccg cag cag	7083
Asp Leu Ser Arg Val Arg Met Glu Leu Gly Val Cys Pro Gln Gln	
1615 1620 1625	
gac atc ctg ttg gac aac ctc acc gtc cgg gaa cat ttg ctg ctc	7128
Asp Ile Leu Leu Asp Asn Leu Thr Val Arg Glu His Leu Leu Leu	
1630 1635 1640	
ttt gct tcc ata aag gcg cct cag tgg acc aag aag gag ctg cat	7173
Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr Lys Lys Glu Leu His	
1645 1650 1655	
cag caa gtc aat caa act ctt cag gat gtg gac tta act cag cat	7218
Gln Gln Val Asn Gln Thr Leu Gln Asp Val Asp Leu Thr Gln His	
1660 1665 1670	
cag cac aaa cag acc cga gct ctg tct gga ggc ctg aag agg aag	7263
Gln His Lys Gln Thr Arg Ala Leu Ser Gly Gly Leu Lys Arg Lys	
1675 1680 1685	
ctc tcc ctt ggc att gct ttc atg ggc atg tcg agg acc gtg gtt	7308
Leu Ser Leu Gly Ile Ala Phe Met Gly Met Ser Arg Thr Val Val	
1690 1695 1700	
ctg gat gag ccc acc agt ggg gtg gac cct tgc tcc cgg cat agc	7353
Leu Asp Glu Pro Thr Ser Gly Val Asp Pro Cys Ser Arg His Ser	
1705 1710 1715	

ctg tgg gac att ctg	ctc aag tac cga gaa ggt cgt acg atc atc	7398
Leu Trp Asp Ile Leu	Leu Lys Tyr Arg Glu Gly Arg Thr Ile Ile	
1720	1725	1730
ttc aca acc cac cac	ctg gat gaa gcc gaa gcg ctg agt gac cgc	7443
Phe Thr Thr His His	Leu Asp Glu Ala Glu Ala Leu Ser Asp Arg	
1735	1740	1745
gtg gcc gtc ctc cag	cat ggg agg ctc agg tgc tgc ggt cct ccc	7488
Val Ala Val Leu Gln	His Gly Arg Leu Arg Cys Cys Gly Pro Pro	
1750	1755	1760
ttc tgc ctg aag gag	gca tat ggc cag ggg ctc cgc ctg aca ctc	7533
Phe Cys Leu Lys Glu	Ala Tyr Gly Gln Gly Leu Arg Leu Thr Leu	
1765	1770	1775
acg agg cag cct tct	gtt ctg gag gcc cat gat ctg aaa gac atg	7578
Thr Arg Gln Pro Ser	Val Leu Glu Ala His Asp Leu Lys Asp Met	
1780	1785	1790
gct tgt gtt aca tcc	ctg ata aag atc tat att cca caa gca ttt	7623
Ala Cys Val Thr Ser	Leu Ile Lys Ile Tyr Ile Pro Gln Ala Phe	
1795	1800	1805
ctc aaa gac agc agt	gga agt gag ctg acc tac acc att cca aag	7668
Leu Lys Asp Ser Ser	Gly Ser Glu Leu Thr Tyr Thr Ile Pro Lys	
1810	1815	1820
gac aca gac aag gcc	tgc ttg aaa ggc ctc ttc cag gcc ctg gat	7713
Asp Thr Asp Lys Ala	Cys Leu Lys Gly Leu Phe Gln Ala Leu Asp	
1825	1830	1835
gag aac ctg cat cag	ctg cac ctg acg ggc tat ggg atc tca gac	7758
Glu Asn Leu His Gln	Leu His Leu Thr Gly Tyr Gly Ile Ser Asp	
1840	1845	1850
acc acc tta gaa gag	gtg ttt ttg atg ctt ttg caa gat tcc aac	7803
Thr Thr Leu Glu Glu	Val Phe Leu Met Leu Leu Gln Asp Ser Asn	
1855	1860	1865
aag aaa tct cac att	gcc ctg ggg act gag tca gag ctg cag aac	7848
Lys Lys Ser His Ile	Ala Leu Gly Thr Glu Ser Glu Leu Gln Asn	
1870	1875	1880
cac agg cct aca gga	cat ctg tct ggc tac tgt ggc tcc cta gca	7893
His Arg Pro Thr Gly	His Leu Ser Gly Tyr Cys Gly Ser Leu Ala	
1885	1890	1895
cgg ccc gca act gtg	cag ggc gtc cag ctg ctc cgc gca caa gtg	7938
Arg Pro Ala Thr Val	Gln Gly Val Gln Leu Leu Arg Ala Gln Val	
1900	1905	1910
gcc gcg atc ctg gcc	cgg agg ctc cgc cgc acg ctg cgc gcc ggg	7983
Ala Ala Ile Leu Ala	Arg Arg Leu Arg Arg Thr Leu Arg Ala Gly	
1915	1920	1925
aag agc acc ctc gcc	gac ctg ctg ctg cca gtc ctc ttc gtg gcc	8028
Lys Ser Thr Leu Ala	Asp Leu Leu Leu Pro Val Leu Phe Val Ala	
1930	1935	1940
ttg gcc atg ggc ttg	ttc atg gtg aga ccc ctg gcc acc gag tac	8073

Leu Ala Met Gly Leu Phe Met Val Arg Pro Leu Ala Thr Glu Tyr	
1945 1950 1955	
cct ccc ctc aga ctc aca cct gga cat tac cag cgg gcc gag acc	8118
Pro Pro Leu Arg Leu Thr Pro Gly His Tyr Gln Arg Ala Glu Thr	
1960 1965 1970	
tac ttt ttc agc agt ggg ggc gac aac ttg gac ctc acc cgt gtg	8163
Tyr Phe Phe Ser Ser Gly Gly Asp Asn Leu Asp Leu Thr Arg Val	
1975 1980 1985	
ctt ctg cgg aag ttt aga gat caa gat ttg ccc tgt gca gat tta	8208
Leu Leu Arg Lys Phe Arg Asp Gln Asp Leu Pro Cys Ala Asp Leu	
1990 1995 2000	
aac cca cgc cag aag aat tct tca tgc tgg cgc aca gat ccc ttt	8253
Asn Pro Arg Gln Lys Asn Ser Ser Cys Trp Arg Thr Asp Pro Phe	
2005 2010 2015	
tct cac cca gaa ttc cag gat tca tgt ggc tgc ctg aag tgt cca	8298
Ser His Pro Glu Phe Gln Asp Ser Cys Gly Cys Leu Lys Cys Pro	
2020 2025 2030	
aat aga agt gct agt gct ccc tac ctg acc aac cac ctg ggc cac	8343
Asn Arg Ser Ala Ser Ala Pro Tyr Leu Thr Asn His Leu Gly His	
2035 2040 2045	
aca ctg ttg aat ctc tca ggc ttc aat atg gag gag tac ttg ctg	8388
Thr Leu Leu Asn Leu Ser Gly Phe Asn Met Glu Glu Tyr Leu Leu	
2050 2055 2060	
gca cca tct gaa aaa cca agg ctt gga ggt tgg tct ttt gga tta	8433
Ala Pro Ser Glu Lys Pro Arg Leu Gly Gly Trp Ser Phe Gly Leu	
2065 2070 2075	
aaa atc ccc agt gaa gct gga ggt gca aat gga aac ata tca aaa	8478
Lys Ile Pro Ser Glu Ala Gly Gly Ala Asn Gly Asn Ile Ser Lys	
2080 2085 2090	
ccc cca act ctg gca aag gtg tgg tat aat cag aag ggt ttt cat	8523
Pro Pro Thr Leu Ala Lys Val Trp Tyr Asn Gln Lys Gly Phe His	
2095 2100 2105	
tcc cta cct tcc tac tta aat cat cta aac aac ctt att ttg tgg	8568
Ser Leu Pro Ser Tyr Leu Asn His Leu Asn Asn Leu Ile Leu Trp	
2110 2115 2120	
cag cac cta ccc cct act gtg gac tgg aga caa tac gga ata aca	8613
Gln His Leu Pro Pro Thr Val Asp Trp Arg Gln Tyr Gly Ile Thr	
2125 2130 2135	
ctc tac agc cac cca tat gga ggg gcc ttg ctg aac gag gac aag	8658
Leu Tyr Ser His Pro Tyr Gly Gly Ala Leu Leu Asn Glu Asp Lys	
2140 2145 2150	
atc ctg gag agc atc cgt cag tgt gga gtg gcc ctc tgc atc gtg	8703
Ile Leu Glu Ser Ile Arg Gln Cys Gly Val Ala Leu Cys Ile Val	
2155 2160 2165	
ctg gga ttc tcc atc ctg tct gca tcc atc ggc agc tct gtg gtg	8748
Leu Gly Phe Ser Ile Leu Ser Ala Ser Ile Gly Ser Ser Val Val	

2170	2175	2180	
agg gac agg gtg att gga gcc aaa agg ttg cag cac ata agt ggc			8793
Arg Asp Arg Val Ile Gly Ala Lys Arg Leu Gln His Ile Ser Gly			
2185	2190	2195	
ctt ggc tac agg atg tac tgg ttc aca aac ttc cta tat gac atg			8838
Leu Gly Tyr Arg Met Tyr Trp Phe Thr Asn Phe Leu Tyr Asp Met			
2200	2205	2210	
ctc ttt tac ttg gtt tcc gtc tgc ctg tgt gtt gcc gtt att gtc			8883
Leu Phe Tyr Leu Val Ser Val Cys Leu Cys Val Ala Val Ile Val			
2215	2220	2225	
gcc ttc cag tta aca gct ttt act ttc cgc aag aac ttg gca gcc			8928
Ala Phe Gln Leu Thr Ala Phe Thr Phe Arg Lys Asn Leu Ala Ala			
2230	2235	2240	
acg gcc ctc ctg ctg tca ctt ttc gga tat gca act ctt cca tgg			8973
Thr Ala Leu Leu Leu Ser Leu Phe Gly Tyr Ala Thr Leu Pro Trp			
2245	2250	2255	
atg tac ctg atg tcc aga atc ttt tcc agt tcg gac gtg gct ttc			9018
Met Tyr Leu Met Ser Arg Ile Phe Ser Ser Ser Asp Val Ala Phe			
2260	2265	2270	
att tcc tat gtc tca cta aac ttc atc ttt ggc ctt tgt acc atg			9063
Ile Ser Tyr Val Ser Leu Asn Phe Ile Phe Gly Leu Cys Thr Met			
2275	2280	2285	
ccc ata acc att atg ccc cgg ttg cta gcc atc atc tcc aaa gct			9108
Pro Ile Thr Ile Met Pro Arg Leu Leu Ala Ile Ile Ser Lys Ala			
2290	2295	2300	
aag aat tta cag aat atc tat gat gtc ctc aag tgg gtc ttt act			9153
Lys Asn Leu Gln Asn Ile Tyr Asp Val Leu Lys Trp Val Phe Thr			
2305	2310	2315	
att ttt cct caa ttc tgt ctt ggt caa gga ctg gta gaa ctc tgc			9198
Ile Phe Pro Gln Phe Cys Leu Gly Gln Gly Leu Val Glu Leu Cys			
2320	2325	2330	
tat aat cag atc aaa tat gac ctg acc cac aac ttc ggc att gat			9243
Tyr Asn Gln Ile Lys Tyr Asp Leu Thr His Asn Phe Gly Ile Asp			
2335	2340	2345	
tcc tat gtg agt ccc ttt gag atg aac ttt ctg ggc tgg atc ttc			9288
Ser Tyr Val Ser Pro Phe Glu Met Asn Phe Leu Gly Trp Ile Phe			
2350	2355	2360	
gtg caa ctg gcc tcg cag ggc aca gta ctt ctc ctc ttg agg gtt			9333
Val Gln Leu Ala Ser Gln Gly Thr Val Leu Leu Leu Leu Arg Val			
2365	2370	2375	
ctg cta cac tgg gac ctt ctg cga tgg cca agg ggt cat tct act			9378
Leu Leu His Trp Asp Leu Leu Arg Trp Pro Arg Gly His Ser Thr			
2380	2385	2390	
ctc caa ggc aca gtc aaa tct tct aag gat aca gat gtt gaa aaa			9423
Leu Gln Gly Thr Val Lys Ser Ser Lys Asp Thr Asp Val Glu Lys			
2395	2400	2405	

gag gaa aag aga gtg ttt gaa gga agg acc aat gga gac att ctt	9468
Glu Glu Lys Arg Val Phe Glu Gly Arg Thr Asn Gly Asp Ile Leu	
2410 2415 2420	
gtg tta tac aac ctt agt aaa cat tat cga cgc ttt ttc cag aat	9513
Val Leu Tyr Asn Leu Ser Lys His Tyr Arg Arg Phe Phe Gln Asn	
2425 2430 2435	
att att gct gtg caa gat att agt ttg ggc ata cca aaa gga gag	9558
Ile Ile Ala Val Gln Asp Ile Ser Leu Gly Ile Pro Lys Gly Glu	
2440 2445 2450	
tgc ttt gga ctt cta ggg gtg aat gga gct ggg aag agc acg act	9603
Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Ser Thr Thr	
2455 2460 2465	
ttc aaa atg ctg aat ggt gaa gtt tct cta act tca gga cat gct	9648
Phe Lys Met Leu Asn Gly Glu Val Ser Leu Thr Ser Gly His Ala	
2470 2475 2480	
atc atc agg act ccc atg gga gac gcc gtg gac ctg tct tct gct	9693
Ile Ile Arg Thr Pro Met Gly Asp Ala Val Asp Leu Ser Ser Ala	
2485 2490 2495	
ggc acg gca ggc gtg ctc att ggc tac tgt ccc cag cag gat gcc	9738
Gly Thr Ala Gly Val Leu Ile Gly Tyr Cys Pro Gln Gln Asp Ala	
2500 2505 2510	
ctg gac gag ctt ctg act ggt tgg gaa cat ctc tat tat tac tgt	9783
Leu Asp Glu Leu Leu Thr Gly Trp Glu His Leu Tyr Tyr Tyr Cys	
2515 2520 2525	
agc tta cgc ggg att cca agg cag tgc atc cct gag gtt gct gga	9828
Ser Leu Arg Gly Ile Pro Arg Gln Cys Ile Pro Glu Val Ala Gly	
2530 2535 2540	
gac ctc atc agg cgc tta cac ctc gaa gcc cac gcg gac aaa cct	9873
Asp Leu Ile Arg Arg Leu His Leu Glu Ala His Ala Asp Lys Pro	
2545 2550 2555	
gtg gcc acc tac agt ggg gga acc aag cgg aaa ctc tct aca gcc	9918
Val Ala Thr Tyr Ser Gly Gly Thr Lys Arg Lys Leu Ser Thr Ala	
2560 2565 2570	
ctg gcc ctg gtg ggg aaa cct gac att ctt tta ttg gat gag ccc	9963
Leu Ala Leu Val Gly Lys Pro Asp Ile Leu Leu Leu Asp Glu Pro	
2575 2580 2585	
agc tct ggg atg gat ccc tgc tct aag cgg tac ctg tgg caa aca	10008
Ser Ser Gly Met Asp Pro Cys Ser Lys Arg Tyr Leu Trp Gln Thr	
2590 2595 2600	
ata atg aag gag gtt cgg gaa ggc tgt gct gcg gtg ctg acc tcc	10053
Ile Met Lys Glu Val Arg Glu Gly Cys Ala Ala Val Leu Thr Ser	
2605 2610 2615	
cac agc atg gag gag tgt gag gct ctt tgc aca aga ctg gcc ata	10098
His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile	
2620 2625 2630	

```

atg gtt aac ggc agc   ttc aaa tgt ctt ggt   tct cct cag cac atc       10143
Met Val Asn Gly Ser   Phe Lys Cys Leu Gly   Ser Pro Gln His Ile
                2635                2640                2645

aaa aat agg tgc gtt   gaa gtt ctc ctt ttg   ctt aga tat tta cta       10188
Lys Asn Arg Cys Val   Glu Val Leu Leu Leu   Leu Arg Tyr Leu Leu
                2650                2655                2660

tga aatcaatagc atgatgcagg gttatgactg gagatgcttt agaggcaagt       10241

atgttccaat tacctgctgg cataaaatag gagatgaagg ctctgagggg ataaaatact 10301

gtggctagta agtcaaacac atagcaactc ccaaacacag tctagtgttc tccctgtaag 10361

ggagacattt tgctttttct cttgccaggt agacattgct ccctccttgt agtgacattg 10421

gtaggggtga gaataagatt gcagagtcaa gatgatgaca ggacacaaaa tgtcctcttt 10481

acacacaaaa tgccttatga aaaggctgtg ctgatatgtc cctcatctgc agtgaatcaa 10541

agttcccat ttatcctacc aaacactatc cttttgaaaa agaccatcaa ttataaggg 10601

taggagaggg ttgttatggg tgctttccat ttggattatg aaagactcaa gacttctttc 10661

ctgtgctcat tgatccttta ttttcttcc tggtaattc tctcatgta ttttctttat 10721

ttcctttttg gaatcttaat aattttgttg tcaatttgtg tgagcattgt tacatcctgg 10781

tataacatta aaaaaattaa agttgcctaa atacacgtcc aattttcctt tgatttgttt 10841

aacctgtcat tcttatgtag aataatttat tcatttatgt tgataaatct cttattttca 10901

tttatcagt gtttttaagg ttaggattct caagcagaga tttggtaa atgtcatatgt 10961

aaattttaaa tactttta attatggagaa aatagctgat gctcaggatt gagatgacat 11021

agtagggagt gagtgtccat ttttagagtt aggcaactgct gggatgggtca gcctcacagt 11081

tgcttgctgc ttcggtgcaa ccactgggtca gatgcaccta ctggaaacta ttctcggtat 11141

gtactctgcc catctttaat aagtatcatc aaaagaaaaa taatatttca tggctctcaa 11201

cctttatgaa tgcctacata aagttttata acccaaaaaa aaaaaaaaaa aaaaaaaa 11259

```

<210> 25

<211> 2660

<212> PRT

<213> Homo Sapiens

<220>

<221> misc_feature

<222> (376)..(376)

<223> The 'Xaa' at location 376 stands for Arg, Gly, or Trp.

<400> 25

```

Met Ser Phe Val Pro Lys Asp Lys Ile Leu Glu Ile Leu Lys Leu Asp
1           5           10          15

```

Gln Phe Leu Thr Leu Met Ile Gln Asp Arg Leu Met Asn Ile Phe Ser
 20 25 30

Ser Leu Lys Glu Thr Ile Tyr His Leu Met Lys Ser Ser Phe Ile Leu
 35 40 45

Asp Asn Gly Glu Phe Tyr Phe Asp Thr His Gln Gly Leu Lys Phe Met
 50 55 60

Gln Asp Leu Phe Asn Ala Leu Leu Arg Glu Thr Ser Met Lys Asn Lys
 65 70 75 80

Thr Glu Asn Asn Ile Asp Phe Phe Thr Val Val Ser Gln Leu Phe Phe
 85 90 95

His Val Asn Lys Ser Glu Asp Leu Phe Lys Leu Asn Gln Asp Leu Gly
 100 105 110

Ser Ala Leu His Leu Val Arg Glu Cys Ser Thr Glu Met Ala Arg Leu
 115 120 125

Leu Asp Thr Ile Leu His Ser Pro Asn Lys Asp Phe Tyr Ala Leu Tyr
 130 135 140

Pro Thr Leu Gln Glu Val Ile Leu Ala Asn Leu Thr Asp Leu Leu Phe
 145 150 155 160

Phe Ile Asn Asn Ser Phe Pro Leu Arg Asn Arg Ala Thr Leu Glu Ile
 165 170 175

Thr Lys Arg Leu Val Gly Ala Ile Ser Arg Ala Ser Glu Glu Ser His
 180 185 190

Val Leu Lys Pro Leu Leu Glu Met Ser Gly Thr Leu Val Met Leu Leu
 195 200 205

Asn Asp Ser Ala Asp Leu Arg Asp Leu Ala Thr Ser Met Asp Ser Ile
 210 215 220

Val Lys Leu Leu Lys Leu Val Lys Lys Val Ser Gly Lys Met Ser Thr
 225 230 235 240

Val Phe Lys Thr His Phe Ile Ser Asn Thr Lys Asp Ser Val Lys Phe
 245 250 255

Phe Asp Thr Leu Tyr Ser Ile Met Gln Gln Ser Val Gln Asn Leu Val

260	265	270
Lys Glu Ile Ala Thr Leu Lys Lys Ile Asp His Phe Thr Phe Glu Lys 275 280 285		
Ile Asn Asp Leu Leu Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile 290 295 300		
Gly Val Glu Pro Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser 305 310 315 320		
Pro Ser Ile Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile 325 330 335		
Leu Glu Glu Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu 340 345 350		
Asp Met Arg Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe 355 360 365		
Ser Met Asp Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu Val 370 375 380		
Pro Ile Asn Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro 385 390 395 400		
Ile Ser Thr His Ser Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His 405 410 415		
Glu Val Ile Leu Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu 420 425 430		
Ile Gly Ser Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu 435 440 445		
Trp Lys Asn Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met 450 455 460		
Thr Phe Thr Gln His Pro Asn Asn Leu Leu Lys Thr Ile Glu Thr Val 465 470 475 480		
Leu Glu Ala Ser Ser Gly Ile Lys Ser Asp Tyr Glu Gly Asp Leu Asn 485 490 495		
Lys Ser Leu Tyr Phe Asp Thr Pro Leu Ser Gln Asn Ile Thr His His 500 505 510		

Gln Leu Glu Lys Ala Ile His Asn Val Leu Ser Arg Ile Ala Leu Trp
 515 520 525

Arg Lys Gly Leu Arg Phe Asn Asn Ser Glu Trp Ile Thr Ser Thr Arg
 530 535 540

Thr Leu Phe Gln Pro Leu Phe Glu Ile Phe Ile Lys Ala Thr Thr Gly
 545 550 555 560

Lys Asn Val Thr Ser Glu Lys Glu Glu Arg Thr Glu Lys Glu Met Ile
 565 570 575

Asp Phe Pro Tyr Ser Phe Lys Pro Phe Phe Cys Leu Glu Lys Tyr Leu
 580 585 590

Gly Gly Leu Phe Val Leu Thr Lys Tyr Trp Gln Gln Ile Pro Leu Thr
 595 600 605

Asp Gln Ser Val Val Glu Ile Cys Glu Val Phe Gln Gln Thr Val Lys
 610 615 620

Pro Ser Glu Ala Met Glu Met Leu Gln Lys Val Lys Met Met Val Val
 625 630 635 640

Arg Val Leu Thr Ile Val Ala Glu Asn Pro Ser Trp Thr Lys Asp Ile
 645 650 655

Leu Cys Ala Thr Leu Ser Cys Lys Gln Asn Gly Ile Arg His Leu Ile
 660 665 670

Leu Ser Ala Ile Gln Gly Val Thr Leu Ala Gln Asp His Phe Gln Glu
 675 680 685

Ile Glu Lys Ile Trp Ser Ser Pro Asn Gln Leu Asn Cys Glu Ser Leu
 690 695 700

Ser Lys Asn Leu Ser Ser Thr Leu Glu Ser Phe Lys Ser Ser Leu Glu
 705 710 715 720

Asn Ala Thr Gly Gln Asp Cys Thr Ser Gln Pro Arg Leu Glu Thr Val
 725 730 735

Gln Gln His Leu Tyr Met Leu Ala Lys Ser Leu Glu Glu Thr Trp Ser
 740 745 750

Ser Gly Asn Pro Ile Met Thr Phe Leu Ser Asn Phe Thr Val Thr Glu
 755 760 765

Asp Val Lys Ile Lys Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu
 770 775 780

Glu Leu Arg Ser Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile
 785 790 795 800

Leu Glu Ala Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr
 805 810 815

Val Ala Leu Ser Gly Lys Cys Asp Gln Glu Ile Leu His Leu Leu Leu
 820 825 830

Thr Phe Pro Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys
 835 840 845

Ser Leu Pro Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg
 850 855 860

Asn Leu Asp Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu
 865 870 875 880

Ala Asn Gly Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser
 885 890 895

Ala Leu Leu Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe
 900 905 910

Leu His Thr Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln
 915 920 925

Gln Val Ser Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe
 930 935 940

Gln Phe Val Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser
 945 950 955 960

Asp Ser Asn Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu Glu
 965 970 975

Asp Asp Lys Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys
 980 985 990

Leu Lys Leu Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val
 995 1000 1005

Trp Thr Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr
 1010 1015 1020

Pro Asn Thr Pro Glu Ile Asn Lys Val Ile Gln Lys Ala Asn Tyr
 1025 1030 1035

Thr Phe Tyr Ile Val Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu
 1040 1045 1050

Leu Glu Met Ser Ser Leu Phe Gln Arg Ser Gly Ser Gly Gln Met
 1055 1060 1065

Phe Asn Gln Leu Gln Glu Ala Leu Arg Asn Lys Phe Val Arg Asn
 1070 1075 1080

Phe Val Glu Asn Gln Leu His Ile Asp Val Asp Lys Leu Thr Glu
 1085 1090 1095

Lys Leu Gln Thr Tyr Gly Gly Leu Leu Asp Glu Met Phe Asn His
 1100 1105 1110

Ala Gly Ala Gly Arg Phe Arg Phe Leu Gly Ser Ile Leu Val Asn
 1115 1120 1125

Leu Ser Ser Cys Val Ala Leu Asn Arg Phe Gln Ala Leu Gln Ser
 1130 1135 1140

Val Asp Ile Leu Glu Thr Lys Ala His Glu Leu Leu Gln Gln Asn
 1145 1150 1155

Ser Phe Leu Ala Ser Ile Ile Phe Ser Asn Ser Leu Phe Asp Lys
 1160 1165 1170

Asn Phe Arg Ser Glu Ser Val Lys Leu Pro Pro His Val Ser Tyr
 1175 1180 1185

Thr Ile Arg Thr Asn Val Leu Tyr Ser Val Arg Thr Asp Val Val
 1190 1195 1200

Lys Asn Pro Ser Trp Lys Phe His Pro Gln Asn Leu Pro Ala Asp
 1205 1210 1215

Gly	Phe	Lys	Tyr	Asn	Tyr	Val	Phe	Ala	Pro	Leu	Gln	Asp	Met	Ile
1220						1225					1230			
Glu	Arg	Ala	Ile	Ile	Leu	Val	Gln	Thr	Gly	Gln	Glu	Ala	Leu	Glu
1235						1240					1245			
Pro	Ala	Ala	Gln	Thr	Gln	Ala	Ala	Pro	Tyr	Pro	Cys	His	Thr	Ser
1250						1255					1260			
Asp	Leu	Phe	Leu	Asn	Asn	Val	Gly	Phe	Phe	Phe	Pro	Leu	Ile	Met
1265						1270					1275			
Met	Leu	Thr	Trp	Met	Val	Ser	Val	Ala	Ser	Met	Val	Arg	Lys	Leu
1280						1285					1290			
Val	Tyr	Glu	Gln	Glu	Ile	Gln	Ile	Glu	Glu	Tyr	Met	Arg	Met	Met
1295						1300					1305			
Gly	Val	His	Pro	Val	Ile	His	Phe	Leu	Ala	Trp	Phe	Leu	Glu	Asn
1310						1315					1320			
Met	Ala	Val	Leu	Thr	Ile	Ser	Ser	Ala	Thr	Leu	Ala	Ile	Val	Leu
1325						1330					1335			
Lys	Thr	Ser	Gly	Ile	Phe	Ala	His	Ser	Asn	Thr	Phe	Ile	Val	Phe
1340						1345					1350			
Leu	Phe	Leu	Leu	Asp	Phe	Gly	Met	Ser	Val	Val	Met	Leu	Ser	Tyr
1355						1360					1365			
Leu	Leu	Ser	Ala	Phe	Phe	Ser	Gln	Ala	Asn	Thr	Ala	Ala	Leu	Cys
1370						1375					1380			
Thr	Ser	Leu	Val	Tyr	Met	Ile	Ser	Phe	Leu	Pro	Tyr	Ile	Val	Leu
1385						1390					1395			
Leu	Val	Leu	His	Asn	Gln	Leu	Ser	Phe	Val	Asn	Gln	Thr	Phe	Leu
1400						1405					1410			
Cys	Leu	Leu	Ser	Thr	Thr	Ala	Phe	Gly	Gln	Gly	Val	Phe	Phe	Ile
1415						1420					1425			
Thr	Phe	Leu	Glu	Gly	Gln	Glu	Thr	Gly	Ile	Gln	Trp	Asn	Asn	Met
1430						1435					1440			
Tyr	Gln	Ala	Leu	Glu	Gln	Gly	Gly	Met	Thr	Phe	Gly	Trp	Val	Cys

1445		1450		1455
Trp Met Ile Leu Phe Asp Ser Ser Leu Tyr Phe Leu Cys Gly Trp				
1460		1465		1470
Tyr Leu Ser Asn Leu Ile Pro Gly Thr Phe Gly Leu Arg Lys Pro				
1475		1480		1485
Trp Tyr Phe Pro Phe Thr Ala Ser Tyr Trp Lys Ser Val Gly Phe				
1490		1495		1500
Leu Val Glu Lys Arg Gln Tyr Phe Leu Ser Ser Ser Leu Phe Phe				
1505		1510		1515
Phe Asn Glu Asn Phe Asp Asn Lys Gly Ser Ser Leu Gln Asn Arg				
1520		1525		1530
Glu Gly Glu Leu Glu Gly Ser Ala Pro Gly Val Thr Leu Val Ser				
1535		1540		1545
Val Thr Lys Glu Tyr Glu Gly His Lys Ala Val Val Gln Asp Leu				
1550		1555		1560
Ser Leu Thr Phe Tyr Arg Asp Gln Ile Thr Ala Leu Leu Gly Thr				
1565		1570		1575
Asn Gly Ala Gly Lys Thr Thr Ile Ile Ser Met Leu Thr Gly Leu				
1580		1585		1590
His Pro Pro Thr Ser Gly Thr Ile Ile Ile Asn Gly Lys Asn Leu				
1595		1600		1605
Gln Thr Asp Leu Ser Arg Val Arg Met Glu Leu Gly Val Cys Pro				
1610		1615		1620
Gln Gln Asp Ile Leu Leu Asp Asn Leu Thr Val Arg Glu His Leu				
1625		1630		1635
Leu Leu Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr Lys Lys Glu				
1640		1645		1650
Leu His Gln Gln Val Asn Gln Thr Leu Gln Asp Val Asp Leu Thr				
1655		1660		1665
Gln His Gln His Lys Gln Thr Arg Ala Leu Ser Gly Gly Leu Lys				
1670		1675		1680

Arg	Lys	Leu	Ser	Leu	Gly	Ile	Ala	Phe	Met	Gly	Met	Ser	Arg	Thr
1685						1690					1695			
Val	Val	Leu	Asp	Glu	Pro	Thr	Ser	Gly	Val	Asp	Pro	Cys	Ser	Arg
1700						1705					1710			
His	Ser	Leu	Trp	Asp	Ile	Leu	Leu	Lys	Tyr	Arg	Glu	Gly	Arg	Thr
1715						1720					1725			
Ile	Ile	Phe	Thr	Thr	His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser
1730						1735					1740			
Asp	Arg	Val	Ala	Val	Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly
1745						1750					1755			
Pro	Pro	Phe	Cys	Leu	Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu
1760						1765					1770			
Thr	Leu	Thr	Arg	Gln	Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys
1775						1780					1785			
Asp	Met	Ala	Cys	Val	Thr	Ser	Leu	Ile	Lys	Ile	Tyr	Ile	Pro	Gln
1790						1795					1800			
Ala	Phe	Leu	Lys	Asp	Ser	Ser	Gly	Ser	Glu	Leu	Thr	Tyr	Thr	Ile
1805						1810					1815			
Pro	Lys	Asp	Thr	Asp	Lys	Ala	Cys	Leu	Lys	Gly	Leu	Phe	Gln	Ala
1820						1825					1830			
Leu	Asp	Glu	Asn	Leu	His	Gln	Leu	His	Leu	Thr	Gly	Tyr	Gly	Ile
1835						1840					1845			
Ser	Asp	Thr	Thr	Leu	Glu	Glu	Val	Phe	Leu	Met	Leu	Leu	Gln	Asp
1850						1855					1860			
Ser	Asn	Lys	Lys	Ser	His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu
1865						1870					1875			
Gln	Asn	His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser
1880						1885					1890			
Leu	Ala	Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu	Leu	Arg	Ala
1895						1900					1905			

Gln Val 1910	Ala Ala Ile Leu 1915	Ala Arg Arg Leu Arg Arg 1920	Thr Leu Arg
Ala Gly 1925	Lys Ser Thr Leu 1930	Ala Asp Leu Leu Leu Pro 1935	Val Leu Phe
Val Ala 1940	Leu Ala Met Gly 1945	Leu Phe Met Val Arg Pro 1950	Leu Ala Thr
Glu Tyr 1955	Pro Pro Leu Arg 1960	Leu Thr Pro Gly His Tyr 1965	Gln Arg Ala
Glu Thr 1970	Tyr Phe Phe Ser 1975	Ser Gly Gly Asp Asn Leu 1980	Asp Leu Thr
Arg Val 1985	Leu Leu Arg Lys 1990	Phe Arg Asp Gln Asp Leu 1995	Pro Cys Ala
Asp Leu 2000	Asn Pro Arg Gln 2005	Lys Asn Ser Ser Cys Trp 2010	Arg Thr Asp
Pro Phe 2015	Ser His Pro Glu 2020	Phe Gln Asp Ser Cys Gly 2025	Cys Leu Lys
Cys Pro 2030	Asn Arg Ser Ala 2035	Ser Ala Pro Tyr Leu Thr 2040	Asn His Leu
Gly His 2045	Thr Leu Leu Asn 2050	Leu Ser Gly Phe Asn Met 2055	Glu Glu Tyr
Leu Leu 2060	Ala Pro Ser Glu 2065	Lys Pro Arg Leu Gly Gly 2070	Trp Ser Phe
Gly Leu 2075	Lys Ile Pro Ser 2080	Glu Ala Gly Gly Ala Asn 2085	Gly Asn Ile
Ser Lys 2090	Pro Pro Thr Leu 2095	Ala Lys Val Trp Tyr Asn 2100	Gln Lys Gly
Phe His 2105	Ser Leu Pro Ser 2110	Tyr Leu Asn His Leu Asn 2115	Asn Leu Ile
Leu Trp 2120	Gln His Leu Pro 2125	Pro Thr Val Asp Trp Arg 2130	Gln Tyr Gly

Ile Thr	Leu Tyr Ser His	Pro Tyr Gly Gly Ala	Leu Leu Asn Glu
2135		2140	2145
Asp Lys	Ile Leu Glu Ser	Ile Arg Gln Cys Gly	Val Ala Leu Cys
2150		2155	2160
Ile Val	Leu Gly Phe Ser	Ile Leu Ser Ala Ser	Ile Gly Ser Ser
2165		2170	2175
Val Val	Arg Asp Arg Val	Ile Gly Ala Lys Arg	Leu Gln His Ile
2180		2185	2190
Ser Gly	Leu Gly Tyr Arg	Met Tyr Trp Phe Thr	Asn Phe Leu Tyr
2195		2200	2205
Asp Met	Leu Phe Tyr Leu	Val Ser Val Cys Leu	Cys Val Ala Val
2210		2215	2220
Ile Val	Ala Phe Gln Leu	Thr Ala Phe Thr Phe	Arg Lys Asn Leu
2225		2230	2235
Ala Ala	Thr Ala Leu Leu	Leu Ser Leu Phe Gly	Tyr Ala Thr Leu
2240		2245	2250
Pro Trp	Met Tyr Leu Met	Ser Arg Ile Phe Ser	Ser Ser Asp Val
2255		2260	2265
Ala Phe	Ile Ser Tyr Val	Ser Leu Asn Phe Ile	Phe Gly Leu Cys
2270		2275	2280
Thr Met	Pro Ile Thr Ile	Met Pro Arg Leu Leu	Ala Ile Ile Ser
2285		2290	2295
Lys Ala	Lys Asn Leu Gln	Asn Ile Tyr Asp Val	Leu Lys Trp Val
2300		2305	2310
Phe Thr	Ile Phe Pro Gln	Phe Cys Leu Gly Gln	Gly Leu Val Glu
2315		2320	2325
Leu Cys	Tyr Asn Gln Ile	Lys Tyr Asp Leu Thr	His Asn Phe Gly
2330		2335	2340
Ile Asp	Ser Tyr Val Ser	Pro Phe Glu Met Asn	Phe Leu Gly Trp
2345		2350	2355
Ile Phe	Val Gln Leu Ala	Ser Gln Gly Thr Val	Leu Leu Leu

2360	2365	2370
Arg Val Leu Leu His Trp Asp 2375	Leu Leu Arg Trp Pro 2380	Arg Gly His 2385
Ser Thr Leu Gln Gly Thr Val 2390	Lys Ser Ser Lys Asp 2395	Thr Asp Val 2400
Glu Lys Glu Glu Lys Arg Val 2405	Phe Glu Gly Arg Thr 2410	Asn Gly Asp 2415
Ile Leu Val Leu Tyr Asn Leu 2420	Ser Lys His Tyr Arg 2425	Arg Phe Phe 2430
Gln Asn Ile Ile Ala Val Gln 2435	Asp Ile Ser Leu Gly 2440	Ile Pro Lys 2445
Gly Glu Cys Phe Gly Leu Leu 2450	Gly Val Asn Gly Ala 2455	Gly Lys Ser 2460
Thr Thr Phe Lys Met Leu Asn 2465	Gly Glu Val Ser Leu 2470	Thr Ser Gly 2475
His Ala Ile Ile Arg Thr Pro 2480	Met Gly Asp Ala Val 2485	Asp Leu Ser 2490
Ser Ala Gly Thr Ala Gly Val 2495	Leu Ile Gly Tyr Cys 2500	Pro Gln Gln 2505
Asp Ala Leu Asp Glu Leu Leu 2510	Thr Gly Trp Glu His 2515	Leu Tyr Tyr 2520
Tyr Cys Ser Leu Arg Gly Ile 2525	Pro Arg Gln Cys Ile 2530	Pro Glu Val 2535
Ala Gly Asp Leu Ile Arg Arg 2540	Leu His Leu Glu Ala 2545	His Ala Asp 2550
Lys Pro Val Ala Thr Tyr Ser 2555	Gly Gly Thr Lys Arg 2560	Lys Leu Ser 2565
Thr Ala Leu Ala Leu Val Gly 2570	Lys Pro Asp Ile Leu 2575	Leu Leu Asp 2580
Glu Pro Ser Ser Gly Met Asp 2585	Pro Cys Ser Lys Arg 2590	Tyr Leu Trp 2595

Gln Thr Ile Met Lys Glu Val Arg Glu Gly Cys Ala Ala Val Leu
 2600 2605 2610

Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu
 2615 2620 2625

Ala Ile Met Val Asn Gly Ser Phe Lys Cys Leu Gly Ser Pro Gln
 2630 2635 2640

His Ile Lys Asn Arg Cys Val Glu Val Leu Leu Leu Leu Arg Tyr
 2645 2650 2655

Leu Leu
 2660

<210> 26
 <211> 15779
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> CDS
 <222> (26)..(14884)

<220>
 <221> misc_feature
 <222> (873)..(873)
 <223> n = any nucleic acid residue

<220>
 <221> misc_feature
 <222> (8045)..(8045)
 <223> n = any nucleic acid residue

<220>
 <221> misc_feature
 <222> (14979)..(14979)
 <223> Xaa = any amino acid

<400> 26
 ggactgagag cagggagcag caggc atg ggg cat gcc ggg tgc cag ttc aaa 52
 Met Gly His Ala Gly Cys Gln Phe Lys
 1 5
 gcc ctg ctg tgg aag aat tgg ctc tgc aga ctc agg aac ccg gtc ctt 100
 Ala Leu Leu Trp Lys Asn Trp Leu Cys Arg Leu Arg Asn Pro Val Leu
 10 15 20 25
 ttc ctt gct gaa ttc ttc tgg cct tgt atc ctg ttt gta att ctg aca 148
 Phe Leu Ala Glu Phe Phe Trp Pro Cys Ile Leu Phe Val Ile Leu Thr
 30 35 40
 gtt ctt cgt ttt caa gaa cct ccc aga tac aga gac att tgt tat ttg 196

Val	Leu	Arg	Phe	Gln	Glu	Pro	Pro	Arg	Tyr	Arg	Asp	Ile	Cys	Tyr	Leu	
			45					50					55			
cag	ccc	cga	gat	cta	ccc	agc	tgt	ggc	gtt	atc	ccc	ttt	gtt	caa	agc	244
Gln	Pro	Arg	Asp	Leu	Pro	Ser	Cys	Gly	Val	Ile	Pro	Phe	Val	Gln	Ser	
		60					65				70					
ctt	ctt	tgt	aac	act	gga	tca	agg	tgt	agg	aac	ttc	agc	tat	gaa	ggg	292
Leu	Leu	Cys	Asn	Thr	Gly	Ser	Arg	Cys	Arg	Asn	Phe	Ser	Tyr	Glu	Gly	
	75					80				85						
tca	atg	gag	cat	cat	ttt	cgt	ttg	tct	agg	ttc	caa	act	gca	gct	gac	340
Ser	Met	Glu	His	His	Phe	Arg	Leu	Ser	Arg	Phe	Gln	Thr	Ala	Ala	Asp	
90					95					100					105	
ccc	aag	aaa	gtc	aac	aac	ctg	gcc	ttt	tta	aaa	gag	ata	caa	gac	ctg	388
Pro	Lys	Lys	Val	Asn	Asn	Leu	Ala	Phe	Leu	Lys	Glu	Ile	Gln	Asp	Leu	
				110					115					120		
gca	gag	gaa	att	cat	gga	atg	atg	gac	aag	gca	aaa	aac	tta	aaa	aga	436
Ala	Glu	Glu	Ile	His	Gly	Met	Met	Asp	Lys	Ala	Lys	Asn	Leu	Lys	Arg	
			125					130					135			
ctt	tgg	gta	gaa	cga	tcc	aac	act	cca	gat	tct	tct	tat	ggc	tcc	agt	484
Leu	Trp	Val	Glu	Arg	Ser	Asn	Thr	Pro	Asp	Ser	Ser	Tyr	Gly	Ser	Ser	
	140						145					150				
ttt	ttt	aca	atg	gat	ctc	aat	aag	acc	gag	gag	gta	ata	ttg	aaa	ctg	532
Phe	Phe	Thr	Met	Asp	Leu	Asn	Lys	Thr	Glu	Glu	Val	Ile	Leu	Lys	Leu	
	155					160					165					
gaa	agc	ctc	cat	cag	cag	cct	cat	atc	tgg	gat	ttt	cta	ctt	tta	ctg	580
Glu	Ser	Leu	His	Gln	Gln	Pro	His	Ile	Trp	Asp	Phe	Leu	Leu	Leu	Leu	
170					175					180					185	
ccg	aga	cta	cac	aca	agc	cat	gat	cat	gtg	gaa	gat	ggc	atg	gat	gtt	628
Pro	Arg	Leu	His	Thr	Ser	His	Asp	His	Val	Glu	Asp	Gly	Met	Asp	Val	
				190					195					200		
gca	gtg	aac	ctt	ctc	cag	acc	att	ttg	aat	tcc	tta	ata	tcc	cta	gaa	676
Ala	Val	Asn	Leu	Leu	Gln	Thr	Ile	Leu	Asn	Ser	Leu	Ile	Ser	Leu	Glu	
			205					210					215			
gat	tta	gat	tgg	ctt	cca	ctc	aac	caa	act	ttt	tcc	cag	gtt	tct	gaa	724
Asp	Leu	Asp	Trp	Leu	Pro	Leu	Asn	Gln	Thr	Phe	Ser	Gln	Val	Ser	Glu	
		220					225					230				
ctt	gta	ctg	aat	gtg	acc	att	tcg	aca	ctg	aca	ttt	ctg	cag	caa	cat	772
Leu	Val	Leu	Asn	Val	Thr	Ile	Ser	Thr	Leu	Thr	Phe	Leu	Gln	Gln	His	
	235					240					245					
gga	gta	gca	gtc	acc	gag	cca	gtt	tac	cac	ctg	tcc	atg	cag	aat	ata	820
Gly	Val	Ala	Val	Thr	Glu	Pro	Val	Tyr	His	Leu	Ser	Met	Gln	Asn	Ile	
250					255					260					265	
gtg	tgg	gat	cca	cag	aaa	gtc	cag	tat	gat	ctc	aaa	tcc	cag	ttt	ggc	868
Val	Trp	Asp	Pro	Gln	Lys	Val	Gln	Tyr	Asp	Leu	Lys	Ser	Gln	Phe	Gly	
				270					275					280		
ttt	gnt	gat	ctt	cac	acg	gaa	cag	atc	ctg	aac	tct	tca	gct	gaa	ctg	916
Phe	Xaa	Asp	Leu	His	Thr	Glu	Gln	Ile	Leu	Asn	Ser	Ser	Ala	Glu	Leu	

285						290						295						
aag	gag	att	ccc	aca	gac	act	tcc	ttg	gag	aag	atg	gtg	tgt	tca	gtc	964		
Lys	Glu	Ile	Pro	Thr	Asp	Thr	Ser	Leu	Glu	Lys	Met	Val	Cys	Ser	Val			
		300					305					310						
ttg	tct	agc	aca	tca	gag	gat	gaa	gct	gag	aaa	tgg	ggc	cac	gtt	gga	1012		
Leu	Ser	Ser	Thr	Ser	Glu	Asp	Glu	Ala	Glu	Lys	Trp	Gly	His	Val	Gly			
		315				320					325							
ggc	tgc	cac	cct	aag	tgg	tca	gaa	gcc	aaa	aac	tat	ctt	gtc	cat	gca	1060		
Gly	Cys	His	Pro	Lys	Trp	Ser	Glu	Ala	Lys	Asn	Tyr	Leu	Val	His	Ala			
330					335					340					345			
gtc	agc	tgg	ctg	cga	gtc	tac	caa	cag	gtg	ttt	gtt	cag	tgg	caa	cag	1108		
Val	Ser	Trp	Leu	Arg	Val	Tyr	Gln	Gln	Val	Phe	Val	Gln	Trp	Gln	Gln			
				350					355					360				
ggc	agc	ctg	ctt	cag	aag	aca	ctc	aca	ggc	atg	ggc	cat	agt	ctg	gag	1156		
Gly	Ser	Leu	Leu	Gln	Lys	Thr	Leu	Thr	Gly	Met	Gly	His	Ser	Leu	Glu			
			365					370					375					
gct	ctc	agg	aat	cag	ttt	gaa	gaa	gag	agc	aag	ccc	tgg	aag	gtg	gtg	1204		
Ala	Leu	Arg	Asn	Gln	Phe	Glu	Glu	Glu	Ser	Lys	Pro	Trp	Lys	Val	Val			
		380					385					390						
gaa	gct	ctg	cac	act	gca	ctg	ctc	ctg	ctg	aat	gac	agc	ttg	tca	gca	1252		
Glu	Ala	Leu	His	Thr	Ala	Leu	Leu	Leu	Leu	Asn	Asp	Ser	Leu	Ser	Ala			
	395					400					405							
gat	ggc	cca	aaa	gat	aat	cat	aca	ttt	cca	aag	ata	tta	cag	cat	ctg	1300		
Asp	Gly	Pro	Lys	Asp	Asn	His	Thr	Phe	Pro	Lys	Ile	Leu	Gln	His	Leu			
410					415					420					425			
tgg	aaa	ttg	caa	agc	ttg	ctg	caa	aac	ctg	ccc	cag	tgg	ccg	gca	ctg	1348		
Trp	Lys	Leu	Gln	Ser	Leu	Leu	Gln	Asn	Leu	Pro	Gln	Trp	Pro	Ala	Leu			
				430					435					440				
aag	aga	ttt	ctt	cag	ctt	gat	gga	gct	ctc	aga	aat	gcg	ata	gct	cag	1396		
Lys	Arg	Phe	Leu	Gln	Leu	Asp	Gly	Ala	Leu	Arg	Asn	Ala	Ile	Ala	Gln			
			445					450					455					
aat	tta	cat	ttt	gtc	caa	gaa	gtc	ctc	att	tgc	ctg	gag	aca	tca	gct	1444		
Asn	Leu	His	Phe	Val	Gln	Glu	Val	Leu	Ile	Cys	Leu	Glu	Thr	Ser	Ala			
		460					465					470						
aat	gat	ttt	aaa	tgg	ttt	gaa	ctt	aac	caa	ttg	aaa	ctg	gaa	aag	gat	1492		
Asn	Asp	Phe	Lys	Trp	Phe	Glu	Leu	Asn	Gln	Leu	Lys	Leu	Glu	Lys	Asp			
	475					480					485							
gtg	ttc	ttt	tgg	gag	ctg	aaa	cag	atg	ttg	gcg	aag	aat	gct	gtc	tgc	1540		
Val	Phe	Phe	Trp	Glu	Leu	Lys	Gln	Met	Leu	Ala	Lys	Asn	Ala	Val	Cys			
490					495					500					505			
ccg	aat	ggc	cgt	ttc	tct	gag	aag	gag	gtc	ttt	ttg	ccg	cct	gga	aac	1588		
Pro	Asn	Gly	Arg	Phe	Ser	Glu	Lys	Glu	Val	Phe	Leu	Pro	Pro	Gly	Asn			
				510					515					520				
tcc	agc	ata	tgg	ggc	ggc	ctc	cag	gga	ctg	ttg	tgc	tat	tgt	aac	tcc	1636		
Ser	Ser	Ile	Trp	Gly	Gly	Leu	Gln	Gly	Leu	Leu	Cys	Tyr	Cys	Asn	Ser			
			525					530					535					

tct gag acg agt gtt tta aac aag cta ctt ggt tca gta gag gat gct	1684
Ser Glu Thr Ser Val Leu Asn Lys Leu Leu Gly Ser Val Glu Asp Ala	
540 545 550	
gat cgt att ttg caa gag gtc att act tgg cac aaa aat atg tca gtt	1732
Asp Arg Ile Leu Gln Glu Val Ile Thr Trp His Lys Asn Met Ser Val	
555 560 565	
tta ata cct gaa gaa tat ttg gac tgg cag gaa ctt gag atg cag ctg	1780
Leu Ile Pro Glu Glu Tyr Leu Asp Trp Gln Glu Leu Glu Met Gln Leu	
570 575 580 585	
tca gaa gca agc ctt tcc tgt act cgg ctc ttc ctg ctg ctg gga gct	1828
Ser Glu Ala Ser Leu Ser Cys Thr Arg Leu Phe Leu Leu Leu Gly Ala	
590 595 600	
gat ccc tct cct gag aat gat gtc ttt tct agt gac tgt aag cac cag	1876
Asp Pro Ser Pro Glu Asn Asp Val Phe Ser Ser Asp Cys Lys His Gln	
605 610 615	
ctt gtc tcc aca gtg ata ttt cat aca ctt gaa aaa aca caa ttt ttc	1924
Leu Val Ser Thr Val Ile Phe His Thr Leu Glu Lys Thr Gln Phe Phe	
620 625 630	
ctg gaa caa gca tat tat tgg aaa gcc ttc aaa aag ttt atc agg aag	1972
Leu Glu Gln Ala Tyr Tyr Trp Lys Ala Phe Lys Lys Phe Ile Arg Lys	
635 640 645	
act tgc gaa gtg gcc caa tat gta aat atg caa gag agt ttc cag aac	2020
Thr Cys Glu Val Ala Gln Tyr Val Asn Met Gln Glu Ser Phe Gln Asn	
650 655 660 665	
aga cta ttg gct ttt cct gag gaa tct cct tgt ttt gaa gaa aac atg	2068
Arg Leu Leu Ala Phe Pro Glu Glu Ser Pro Cys Phe Glu Glu Asn Met	
670 675 680	
gat tgg aaa atg atc agt gat aat tat ttt caa ttt ttg aat aac tta	2116
Asp Trp Lys Met Ile Ser Asp Asn Tyr Phe Gln Phe Leu Asn Asn Leu	
685 690 695	
ctc aag tct cca aca gct tcc ata tcc agg gct tta aat ttc aca aag	2164
Leu Lys Ser Pro Thr Ala Ser Ile Ser Arg Ala Leu Asn Phe Thr Lys	
700 705 710	
cac ctt cta atg atg gaa aag aag ttg cac acc ctt gag gat gaa caa	2212
His Leu Leu Met Met Glu Lys Lys Leu His Thr Leu Glu Asp Glu Gln	
715 720 725	
atg aac ttt ctt tta tca ttt gtg gaa ttt ttt gag aaa tta ttg ttg	2260
Met Asn Phe Leu Leu Ser Phe Val Glu Phe Phe Glu Lys Leu Leu Leu	
730 735 740 745	
cct aat ctt ttt gac tcc tcc att gtt ccc agt ttc cac agc ctc cca	2308
Pro Asn Leu Phe Asp Ser Ser Ile Val Pro Ser Phe His Ser Leu Pro	
750 755 760	
tct ctc aca gag gat att ctg aat ata agt tct ctg tgg aca aat cat	2356
Ser Leu Thr Glu Asp Ile Leu Asn Ile Ser Ser Leu Trp Thr Asn His	
765 770 775	

tta aaa agt tta aag aga gac cca tct gcc act gat gct cag aaa ctc Leu Lys Ser Leu Lys Arg Asp Pro Ser Ala Thr Asp Ala Gln Lys Leu 780 785 790	2404
ttg gaa ttt ggc aac gaa gtg att tgg aaa atg cag act ctc gga agt Leu Glu Phe Gly Asn Glu Val Ile Trp Lys Met Gln Thr Leu Gly Ser 795 800 805	2452
cac tgg ata agg aag gaa cca aaa aat ctt ttg aga ttc ata gaa tta His Trp Ile Arg Lys Glu Pro Lys Asn Leu Leu Arg Phe Ile Glu Leu 810 815 820 825	2500
ata ctt ttt gaa att aat ccc aaa tta cta gaa tta tgg gcc tat ggc Ile Leu Phe Glu Ile Asn Pro Lys Leu Leu Glu Leu Trp Ala Tyr Gly 830 835 840	2548
att tca aaa gga aaa aga gct aaa ttg gaa aac ttc ttt aca ctt tta Ile Ser Lys Gly Lys Arg Ala Lys Leu Glu Asn Phe Phe Thr Leu Leu 845 850 855	2596
aat ttt tct gtt cca gaa aat gag att ctg agt aca agt ttt aac ttt Asn Phe Ser Val Pro Glu Asn Glu Ile Leu Ser Thr Ser Phe Asn Phe 860 865 870	2644
tcc cag ttg ttc cat tca gat tgg cct aaa tca cca gct atg aac ata Ser Gln Leu Phe His Ser Asp Trp Pro Lys Ser Pro Ala Met Asn Ile 875 880 885	2692
gat ttt gta cgt tta agt gag gct ata ata act agt ctc cat gaa ttt Asp Phe Val Arg Leu Ser Glu Ala Ile Ile Thr Ser Leu His Glu Phe 890 895 900 905	2740
gga ttt ttg gag cag gaa cag atc tca gaa gct ctg aac aca gtc tac Gly Phe Leu Glu Gln Glu Gln Ile Ser Glu Ala Leu Asn Thr Val Tyr 910 915 920	2788
gct atc agg aat gca tct gat ctt ttc tca gcc ctt tct gaa cca caa Ala Ile Arg Asn Ala Ser Asp Leu Phe Ser Ala Leu Ser Glu Pro Gln 925 930 935	2836
aaa caa gaa gtt gat aaa att ttg act cac ata cac cta aat gtc ttc Lys Gln Glu Val Asp Lys Ile Leu Thr His Ile His Leu Asn Val Phe 940 945 950	2884
cag gac aag gat tca gct tta ctt ctg caa att tat tct tca ttt tac Gln Asp Lys Asp Ser Ala Leu Leu Gln Ile Tyr Ser Ser Phe Tyr 955 960 965	2932
cga tat att tat gaa tta ttg aat att cag agt aga ggc tct tcg ttg Arg Tyr Ile Tyr Glu Leu Leu Asn Ile Gln Ser Arg Gly Ser Ser Leu 970 975 980 985	2980
act ttc ctt aca caa atc tca aaa cac att ttg gat atc ata aaa caa Thr Phe Leu Thr Gln Ile Ser Lys His Ile Leu Asp Ile Ile Lys Gln 990 995 1000	3028
ttt aat ttc caa aac atc agt aaa gca ttt gca ttt tta ttt aag Phe Asn Phe Gln Asn Ile Ser Lys Ala Phe Ala Phe Leu Phe Lys 1005 1010 1015	3073
aca gca gag gtt ctt ggg gga att tct aat gta tct tac tgt cag	3118

Thr	Ala	Glu	Val	Leu	Gly	Gly	Ile	Ser	Asn	Val	Ser	Tyr	Cys	Gln		
			1020					1025					1030			
caa	ttg	ctt	tca	att	ttt	aac	ttt	ttg	gag	ctt	cag	gcc	caa	tcc	3163	
Gln	Leu	Leu	Ser	Ile	Phe	Asn	Phe	Leu	Glu	Leu	Gln	Ala	Gln	Ser		
			1035					1040					1045			
ttc	atg	tct	aca	gag	ggc	caa	gaa	ctg	gaa	gtg	atc	cac	act	act	3208	
Phe	Met	Ser	Thr	Glu	Gly	Gln	Glu	Leu	Glu	Val	Ile	His	Thr	Thr		
			1050					1055					1060			
ttg	aca	ggc	ctc	aaa	cag	ctg	ctc	ata	att	gat	gaa	gat	ttt	cgt	3253	
Leu	Thr	Gly	Leu	Lys	Gln	Leu	Leu	Ile	Ile	Asp	Glu	Asp	Phe	Arg		
			1065					1070					1075			
att	tct	tta	ttt	caa	tat	atg	agc	caa	ttc	ttc	aac	agt	tca	gta	3298	
Ile	Ser	Leu	Phe	Gln	Tyr	Met	Ser	Gln	Phe	Phe	Asn	Ser	Ser	Val		
			1080					1085					1090			
gaa	gac	cta	ttg	gat	aat	aaa	tgc	ttg	att	tcg	gac	aat	aaa	cac	3343	
Glu	Asp	Leu	Leu	Asp	Asn	Lys	Cys	Leu	Ile	Ser	Asp	Asn	Lys	His		
			1095					1100					1105			
att	tct	tcc	gta	aat	tat	tca	aca	agt	gag	gag	tct	tca	ttt	gtt	3388	
Ile	Ser	Ser	Val	Asn	Tyr	Ser	Thr	Ser	Glu	Glu	Ser	Ser	Phe	Val		
			1110					1115					1120			
ttt	cca	ttg	gca	caa	att	ttt	tca	aac	ctc	tca	gca	aat	gtc	agt	3433	
Phe	Pro	Leu	Ala	Gln	Ile	Phe	Ser	Asn	Leu	Ser	Ala	Asn	Val	Ser		
			1125					1130					1135			
gtg	ttc	aac	aag	ttt	atg	tcc	att	cac	tgt	acc	gtt	tca	tgg	ctt	3478	
Val	Phe	Asn	Lys	Phe	Met	Ser	Ile	His	Cys	Thr	Val	Ser	Trp	Leu		
			1140					1145					1150			
caa	atg	tgg	act	gaa	atc	tgg	gaa	acc	ata	tct	caa	tta	ttt	aag	3523	
Gln	Met	Trp	Thr	Glu	Ile	Trp	Glu	Thr	Ile	Ser	Gln	Leu	Phe	Lys		
			1155					1160					1165			
ttt	gac	atg	aat	gtt	ttc	aca	tct	ctt	cat	cat	ggc	ttc	act	cag	3568	
Phe	Asp	Met	Asn	Val	Phe	Thr	Ser	Leu	His	His	Gly	Phe	Thr	Gln		
			1170					1175					1180			
ctt	ttg	gat	gaa	ttg	gaa	gat	gat	gtg	aaa	gtc	tct	aaa	agc	tgc	3613	
Leu	Leu	Asp	Glu	Leu	Glu	Asp	Asp	Val	Lys	Val	Ser	Lys	Ser	Cys		
			1185					1190					1195			
cag	ggc	ata	ctt	ccc	acc	cat	aat	gtt	gct	aga	ctc	ata	tta	aat	3658	
Gln	Gly	Ile	Leu	Pro	Thr	His	Asn	Val	Ala	Arg	Leu	Ile	Leu	Asn		
			1200					1205					1210			
ttg	ttt	aaa	aat	gta	act	caa	gcc	aat	gac	ttc	cat	aat	tgg	gag	3703	
Leu	Phe	Lys	Asn	Val	Thr	Gln	Ala	Asn	Asp	Phe	His	Asn	Trp	Glu		
			1215					1220					1225			
gac	ttc	ctg	gat	ctc	agg	gat	ttt	ttg	gta	gct	tta	ggc	aat	gca	3748	
Asp	Phe	Leu	Asp	Leu	Arg	Asp	Phe	Leu	Val	Ala	Leu	Gly	Asn	Ala		
			1230					1235					1240			
tta	gtt	tca	gta	aaa	aaa	ctt	aac	ttg	gag	caa	gtg	gag	aaa	tcc	3793	
Leu	Val	Ser	Val	Lys	Lys	Leu	Asn	Leu	Glu	Gln	Val	Glu	Lys	Ser		

1245	1250	1255	
ctt ttc acc atg gaa gct gcc ctg cat cag ttg aag aca ttt cca			3838
Leu Phe Thr Met Glu Ala Ala Leu His Gln Leu Lys Thr Phe Pro			
1260	1265	1270	
ttc aac gaa agt aca agc aga gag ttt tta aat tct ctg ctt gaa			3883
Phe Asn Glu Ser Thr Ser Arg Glu Phe Leu Asn Ser Leu Leu Glu			
1275	1280	1285	
gtt ttc att gag ttt agc agt acc tca gaa tat ata gtc aga aat			3928
Val Phe Ile Glu Phe Ser Ser Thr Ser Glu Tyr Ile Val Arg Asn			
1290	1295	1300	
cta gat tca ata aat gac ttt ctt tca aat aat ctc aca aat tat			3973
Leu Asp Ser Ile Asn Asp Phe Leu Ser Asn Asn Leu Thr Asn Tyr			
1305	1310	1315	
gga gaa aaa ttt gaa aat atc atc act gag cta aga gaa gca ata			4018
Gly Glu Lys Phe Glu Asn Ile Ile Thr Glu Leu Arg Glu Ala Ile			
1320	1325	1330	
gta ttt ctt aga aat gta tca cat gat cga gat ttg ttt tcc tgt			4063
Val Phe Leu Arg Asn Val Ser His Asp Arg Asp Leu Phe Ser Cys			
1335	1340	1345	
gct gat att ttc caa aat gtt act gag tgt att tta gaa gat ggc			4108
Ala Asp Ile Phe Gln Asn Val Thr Glu Cys Ile Leu Glu Asp Gly			
1350	1355	1360	
ttt tta tat gta aat acc tca cag agg atg tta cgt att cta gac			4153
Phe Leu Tyr Val Asn Thr Ser Gln Arg Met Leu Arg Ile Leu Asp			
1365	1370	1375	
acg tta aat tcc aca ttt tcc tct gag aac aca att agc agt ctg			4198
Thr Leu Asn Ser Thr Phe Ser Ser Glu Asn Thr Ile Ser Ser Leu			
1380	1385	1390	
aaa gga tgc att gta tgg tta gat gtc ata aac cat ttg tat ttg			4243
Lys Gly Cys Ile Val Trp Leu Asp Val Ile Asn His Leu Tyr Leu			
1395	1400	1405	
ttg tct aac tcc agt ttt tca caa ggt cgt ctt caa aat att ttg			4288
Leu Ser Asn Ser Ser Phe Ser Gln Gly Arg Leu Gln Asn Ile Leu			
1410	1415	1420	
ggg aat ttc aga gat ata gaa aac aaa atg aac tct ata tta aaa			4333
Gly Asn Phe Arg Asp Ile Glu Asn Lys Met Asn Ser Ile Leu Lys			
1425	1430	1435	
att gta act tgg gtg tta aat ata aaa aaa cct ctt tgt tca tca			4378
Ile Val Thr Trp Val Leu Asn Ile Lys Lys Pro Leu Cys Ser Ser			
1440	1445	1450	
aat ggc tca cat ata aat tgt gtc aat att tac ttg aaa gat gta			4423
Asn Gly Ser His Ile Asn Cys Val Asn Ile Tyr Leu Lys Asp Val			
1455	1460	1465	
act gac ttt cta aat att gta ctt act aca gtc ttt gaa aaa gag			4468
Thr Asp Phe Leu Asn Ile Val Leu Thr Thr Val Phe Glu Lys Glu			
1470	1475	1480	

aag aaa cct aaa ttt gag att tta tta gct ctt tta aat gat tcc	4513
Lys Lys Pro Lys Phe Glu Ile Leu Leu Ala Leu Leu Asn Asp Ser	
1485 1490 1495	
aca aag caa gta agg atg agt atc aac aac tta aca aca gac ttt	4558
Thr Lys Gln Val Arg Met Ser Ile Asn Asn Leu Thr Thr Asp Phe	
1500 1505 1510	
gat ttt gca tct cag tcc aat tgg aga tat ttt act gaa tta att	4603
Asp Phe Ala Ser Gln Ser Asn Trp Arg Tyr Phe Thr Glu Leu Ile	
1515 1520 1525	
cta aga cca ata gaa atg tca gat gaa att cct aat cag ttt caa	4648
Leu Arg Pro Ile Glu Met Ser Asp Glu Ile Pro Asn Gln Phe Gln	
1530 1535 1540	
aat att tgg ctt cat tta ata aca ctg ggg aag gaa ttt cag aag	4693
Asn Ile Trp Leu His Leu Ile Thr Leu Gly Lys Glu Phe Gln Lys	
1545 1550 1555	
ctt gta aaa ggt att tat ttt aac atc ctg gaa aat aat tcc tct	4738
Leu Val Lys Gly Ile Tyr Phe Asn Ile Leu Glu Asn Asn Ser Ser	
1560 1565 1570	
tct aaa act gaa aac ttg tta aac ata ttt gcc acc agt cca aaa	4783
Ser Lys Thr Glu Asn Leu Leu Asn Ile Phe Ala Thr Ser Pro Lys	
1575 1580 1585	
gaa aag gat gta aac agt gta ggc aat tcc att tat cac tta gct	4828
Glu Lys Asp Val Asn Ser Val Gly Asn Ser Ile Tyr His Leu Ala	
1590 1595 1600	
agt tac ctt gcc ttc agc tta tct cat gac ctc caa aat tca cca	4873
Ser Tyr Leu Ala Phe Ser Leu Ser His Asp Leu Gln Asn Ser Pro	
1605 1610 1615	
aaa ata ata att tca cct gaa ata atg aaa gct aca ggt ctt ggt	4918
Lys Ile Ile Ile Ser Pro Glu Ile Met Lys Ala Thr Gly Leu Gly	
1620 1625 1630	
att caa ctg ata agg gat gtg ttc aac tcc tta atg cct gta gtt	4963
Ile Gln Leu Ile Arg Asp Val Phe Asn Ser Leu Met Pro Val Val	
1635 1640 1645	
cat cac act agt cca caa aat gca ggt tat atg caa gct ttg aag	5008
His His Thr Ser Pro Gln Asn Ala Gly Tyr Met Gln Ala Leu Lys	
1650 1655 1660	
aag gta act tct gtc atg cgt acc ctt aag aaa gca gac ata gac	5053
Lys Val Thr Ser Val Met Arg Thr Leu Lys Lys Ala Asp Ile Asp	
1665 1670 1675	
ctt tta gtg gat cag ctt gaa caa gtt agt gta aac cta atg gat	5098
Leu Leu Val Asp Gln Leu Glu Gln Val Ser Val Asn Leu Met Asp	
1680 1685 1690	
ttc ttt aag aat atc agt agt gtg gga act ggc aat tta gtg gtc	5143
Phe Phe Lys Asn Ile Ser Ser Val Gly Thr Gly Asn Leu Val Val	
1695 1700 1705	

aat ttg ctt gtt	ggc ttg atg gaa aaa	ttt gca gac agc tca cat	5188
Asn Leu Leu Val	Gly Leu Met Glu Lys	Phe Ala Asp Ser Ser His	
1710	1715	1720	
tct tgg aat gtt	aat cat ctg ctg cag	ctc tca cgc ctg ttt cct	5233
Ser Trp Asn Val	Asn His Leu Leu Gln	Leu Ser Arg Leu Phe Pro	
1725	1730	1735	
aaa gat gtt gtg	gat gct gtg ata gat	gtg tac tat gtg ctt cct	5278
Lys Asp Val Val	Asp Ala Val Ile Asp	Val Tyr Tyr Val Leu Pro	
1740	1745	1750	
cat gct gta agg	ctc ctg cag gga gta	cct ggt aaa aac atc act	5323
His Ala Val Arg	Leu Leu Gln Gly Val	Pro Gly Lys Asn Ile Thr	
1755	1760	1765	
gaa ggc ctc aag	gat gtc tac agc ttc	aca ctc ctt cat ggc ata	5368
Glu Gly Leu Lys	Asp Val Tyr Ser Phe	Thr Leu Leu His Gly Ile	
1770	1775	1780	
acc att tca aat	atc acc aag gaa gac	ttc gca att gtg ata aaa	5413
Thr Ile Ser Asn	Ile Thr Lys Glu Asp	Phe Ala Ile Val Ile Lys	
1785	1790	1795	
att ctt ttg gat	aca att gaa tta gta	tca gat aag cca gat att	5458
Ile Leu Leu Asp	Thr Ile Glu Leu Val	Ser Asp Lys Pro Asp Ile	
1800	1805	1810	
att tca gag gct	tta gct tgt ttt cct	gtg gtt tgg tgc tgg aat	5503
Ile Ser Glu Ala	Leu Ala Cys Phe Pro	Val Val Trp Cys Trp Asn	
1815	1820	1825	
cac aca aat tct	gga ttt cgg cag aat	tca aag ata gac ccc tgc	5548
His Thr Asn Ser	Gly Phe Arg Gln Asn	Ser Lys Ile Asp Pro Cys	
1830	1835	1840	
aat gtc cat ggg	ctc atg tct tct tcc	ttt tat ggc aaa gtg gcc	5593
Asn Val His Gly	Leu Met Ser Ser Ser	Phe Tyr Gly Lys Val Ala	
1845	1850	1855	
agt ata ctt gat	cat ttc cac ctg tct	ccc caa ggt gaa gat tca	5638
Ser Ile Leu Asp	His Phe His Leu Ser	Pro Gln Gly Glu Asp Ser	
1860	1865	1870	
cca tgt tca aat	gaa agc tcc cga atg	gaa ata act agg aaa gtg	5683
Pro Cys Ser Asn	Glu Ser Ser Arg Met	Glu Ile Thr Arg Lys Val	
1875	1880	1885	
gtc tgc ata att	cat gaa tta gtg gac	tgg aat tct att ctt ctg	5728
Val Cys Ile Ile	His Glu Leu Val Asp	Trp Asn Ser Ile Leu Leu	
1890	1895	1900	
gag ctc tct gaa	gtc ttc cat gtt aac	att tct ctt gtg aaa act	5773
Glu Leu Ser Glu	Val Phe His Val Asn	Ile Ser Leu Val Lys Thr	
1905	1910	1915	
gtg cag aaa ttt	tgg cat aag ata tta	ccg ttt gtc cca cct tca	5818
Val Gln Lys Phe	Trp His Lys Ile Leu	Pro Phe Val Pro Pro Ser	
1920	1925	1930	
ata aat caa act	agg gat agc atc tct	gaa ctc tgt cct agt ggt	5863

Ile	Asn	Gln	Thr	Arg	Asp	Ser	Ile	Ser	Glu	Leu	Cys	Pro	Ser	Gly	
			1935					1940						1945	
tcc	ata	aag	caa	gtt	gct	ttg	caa	atc	ata	gaa	aaa	ctt	aaa	aat	5908
Ser	Ile	Lys	Gln	Val	Ala	Leu	Gln	Ile	Ile	Glu	Lys	Leu	Lys	Asn	
			1950					1955						1960	
gtc	aac	ttt	aca	aaa	gtt	aca	tca	ggc	gaa	aat	att	ctt	gac	aaa	5953
Val	Asn	Phe	Thr	Lys	Val	Thr	Ser	Gly	Glu	Asn	Ile	Leu	Asp	Lys	
			1965					1970						1975	
cta	agt	agt	tta	aac	aag	atc	ctt	aac	att	aat	gaa	gac	aca	gag	5998
Leu	Ser	Ser	Leu	Asn	Lys	Ile	Leu	Asn	Ile	Asn	Glu	Asp	Thr	Glu	
			1980					1985						1990	
aca	tct	gtt	caa	aat	att	att	tcc	tca	aat	ttg	gaa	agg	aca	gta	6043
Thr	Ser	Val	Gln	Asn	Ile	Ile	Ser	Ser	Asn	Leu	Glu	Arg	Thr	Val	
			1995					2000						2005	
caa	ttg	att	tct	gaa	gac	tgg	agc	cta	gaa	aaa	agt	acg	cat	aat	6088
Gln	Leu	Ile	Ser	Glu	Asp	Trp	Ser	Leu	Glu	Lys	Ser	Thr	His	Asn	
			2010					2015						2020	
cta	ctc	tct	tta	ttc	atg	atg	ctc	cag	aat	gca	aat	gtc	aca	ggc	6133
Leu	Leu	Ser	Leu	Phe	Met	Met	Leu	Gln	Asn	Ala	Asn	Val	Thr	Gly	
			2025					2030						2035	
agc	agt	tta	gaa	gca	tta	tca	agt	ttt	att	gaa	aaa	agt	gaa	aca	6178
Ser	Ser	Leu	Glu	Ala	Leu	Ser	Ser	Phe	Ile	Glu	Lys	Ser	Glu	Thr	
			2040					2045						2050	
cct	tac	aac	ttt	gaa	gaa	cta	tgg	ccc	aag	ttt	caa	caa	atc	atg	6223
Pro	Tyr	Asn	Phe	Glu	Glu	Leu	Trp	Pro	Lys	Phe	Gln	Gln	Ile	Met	
			2055					2060						2065	
aaa	gac	cta	acc	caa	gat	ttt	aga	atc	aga	cac	ctg	ctt	tct	gaa	6268
Lys	Asp	Leu	Thr	Gln	Asp	Phe	Arg	Ile	Arg	His	Leu	Leu	Ser	Glu	
			2070					2075						2080	
atg	aac	aaa	gga	atc	aaa	agt	ata	aat	tca	atg	gct	ctt	caa	aag	6313
Met	Asn	Lys	Gly	Ile	Lys	Ser	Ile	Asn	Ser	Met	Ala	Leu	Gln	Lys	
			2085					2090						2095	
ata	act	ttg	cag	ttt	gcc	cat	ttc	ctg	gaa	atc	ctg	gat	tca	ccg	6358
Ile	Thr	Leu	Gln	Phe	Ala	His	Phe	Leu	Glu	Ile	Leu	Asp	Ser	Pro	
			2100					2105						2110	
tca	ttg	aag	aca	tta	gaa	att	att	gaa	gat	ttt	cta	ttg	gtc	aca	6403
Ser	Leu	Lys	Thr	Leu	Glu	Ile	Ile	Glu	Asp	Phe	Leu	Leu	Val	Thr	
			2115					2120						2125	
aaa	aac	tgg	ctt	cag	gaa	tat	gca	aat	gag	gat	tac	tcc	aga	atg	6448
Lys	Asn	Trp	Leu	Gln	Glu	Tyr	Ala	Asn	Glu	Asp	Tyr	Ser	Arg	Met	
			2130					2135						2140	
ata	gaa	aca	tta	ttc	att	cct	gtg	acc	aat	gag	agt	tca	act	gaa	6493
Ile	Glu	Thr	Leu	Phe	Ile	Pro	Val	Thr	Asn	Glu	Ser	Ser	Thr	Glu	
			2145					2150						2155	
gat	ata	gct	ttg	tta	gcc	aaa	gct	att	gct	act	ttt	tgg	ggc	tct	6538
Asp	Ile	Ala	Leu	Leu	Ala	Lys	Ala	Ile	Ala	Thr	Phe	Trp	Gly	Ser	

2160	2165	2170	
tta aaa aat ata tct aga gca ggc aat ttt gat gtt gcc ttt ctt Leu Lys Asn Ile Ser Arg Ala Gly Asn Phe Asp Val Ala Phe Leu 2175 2180 2185			6583
acc cat ctg cta aat caa gaa cag ctg act aat ttc tca gtt gtt Thr His Leu Leu Asn Gln Glu Gln Leu Thr Asn Phe Ser Val Val 2190 2195 2200			6628
cag ctg ctt ttt gaa aac atc cta att aat ttg atc aat aac tta Gln Leu Leu Phe Glu Asn Ile Leu Ile Asn Leu Ile Asn Asn Leu 2205 2210 2215			6673
gct ggg aat tct cag gaa gca gct tgg aac tta aat gat act gac Ala Gly Asn Ser Gln Glu Ala Ala Trp Asn Leu Asn Asp Thr Asp 2220 2225 2230			6718
ctt caa ata atg aat ttc att aac ctt atc ttg aac cat atg cag Leu Gln Ile Met Asn Phe Ile Asn Leu Ile Leu Asn His Met Gln 2235 2240 2245			6763
tca gaa act agt agg aaa aca gtt ctc tct ctg aga agc ata gta Ser Glu Thr Ser Arg Lys Thr Val Leu Ser Leu Arg Ser Ile Val 2250 2255 2260			6808
gat ttc aca gaa cag ttt ttg aaa aca ttc ttc tcc ctt ttt cta Asp Phe Thr Glu Gln Phe Leu Lys Thr Phe Phe Ser Leu Phe Leu 2265 2270 2275			6853
aag gaa gat tct gag aac aaa ata tct ctt ctg ctg aaa tat ttc Lys Glu Asp Ser Glu Asn Lys Ile Ser Leu Leu Leu Lys Tyr Phe 2280 2285 2290			6898
cac aaa gat gtt att gca gag atg agt ttt gtc cca aaa gat aaa His Lys Asp Val Ile Ala Glu Met Ser Phe Val Pro Lys Asp Lys 2295 2300 2305			6943
att cta gaa att ctg aaa ctg gat caa ttt ctt acc ctg atg ata Ile Leu Glu Ile Leu Lys Leu Asp Gln Phe Leu Thr Leu Met Ile 2310 2315 2320			6988
caa gac aga ttg atg aac att ttt tca agt tta aag gag act ata Gln Asp Arg Leu Met Asn Ile Phe Ser Ser Leu Lys Glu Thr Ile 2325 2330 2335			7033
tat cac cta atg aaa agt tca ttt ata tta gac aat gga gaa ttt Tyr His Leu Met Lys Ser Ser Phe Ile Leu Asp Asn Gly Glu Phe 2340 2345 2350			7078
tat ttt gat act cat caa gga ctg aag ttc atg caa gat tta ttt Tyr Phe Asp Thr His Gln Gly Leu Lys Phe Met Gln Asp Leu Phe 2355 2360 2365			7123
aat gcc ctt ctc agg gaa act tca atg aaa aat aag act gaa aat Asn Ala Leu Leu Arg Glu Thr Ser Met Lys Asn Lys Thr Glu Asn 2370 2375 2380			7168
aat ata gac ttt ttc aca gtg gtg agt cag ttg ttt ttc cat gtg Asn Ile Asp Phe Phe Thr Val Val Ser Gln Leu Phe Phe His Val 2385 2390 2395			7213

aat aag tct gag	gac ctc ttc aaa ctc	aat caa gat ctt ggg tca	7258
Asn Lys Ser Glu	Asp Leu Phe Lys Leu	Asn Gln Asp Leu Gly Ser	
2400	2405	2410	
gct ctt cac ctt	gta aga gaa tgt tca	aca gag atg gca aga ctt	7303
Ala Leu His Leu	Val Arg Glu Cys Ser	Thr Glu Met Ala Arg Leu	
2415	2420	2425	
ctg gat aca att	tta cac tct cct aat	aag gac ttc tat gct ttg	7348
Leu Asp Thr Ile	Leu His Ser Pro Asn	Lys Asp Phe Tyr Ala Leu	
2430	2435	2440	
tat cct acc ctc	caa gaa gtt ata ctt	gct aat cta acg gat ttg	7393
Tyr Pro Thr Leu	Gln Glu Val Ile Leu	Ala Asn Leu Thr Asp Leu	
2445	2450	2455	
ctt ttc ttt ata	aat aat tca ttc cct	cta aga aac aga gca aca	7438
Leu Phe Phe Ile	Asn Asn Ser Phe Pro	Leu Arg Asn Arg Ala Thr	
2460	2465	2470	
tta gaa att act	aag aga tta gtt ggt	gct att tca aga gca agt	7483
Leu Glu Ile Thr	Lys Arg Leu Val Gly	Ala Ile Ser Arg Ala Ser	
2475	2480	2485	
gaa gaa agt cac	gtc ctg aaa ccc ctc	tta gaa atg tct ggg act	7528
Glu Glu Ser His	Val Leu Lys Pro Leu	Leu Glu Met Ser Gly Thr	
2490	2495	2500	
ctg gtc atg ctg	ttg aat gac agt gct	gac ctg aga gat ctt gcc	7573
Leu Val Met Leu	Leu Asn Asp Ser Ala	Asp Leu Arg Asp Leu Ala	
2505	2510	2515	
aca tca atg gac	tcc att gtg aaa ctt	ctt aag ctg gtc aag aaa	7618
Thr Ser Met Asp	Ser Ile Val Lys Leu	Leu Lys Leu Val Lys Lys	
2520	2525	2530	
gtt tcg ggg aag	atg tcc aca gtt ttt	aaa act cat ttt atc tcc	7663
Val Ser Gly Lys	Met Ser Thr Val Phe	Lys Thr His Phe Ile Ser	
2535	2540	2545	
aat acc aag gac	agt gtg aaa ttc ttt	gac act ctg tat tcc atc	7708
Asn Thr Lys Asp	Ser Val Lys Phe Phe	Asp Thr Leu Tyr Ser Ile	
2550	2555	2560	
atg caa caa agt	gtt caa aat ctt gtg	aaa gaa ata gct act tta	7753
Met Gln Gln Ser	Val Gln Asn Leu Val	Lys Glu Ile Ala Thr Leu	
2565	2570	2575	
aaa aaa ata gat	cat ttc aca ttt gaa	aag ata aat gat ttg ttg	7798
Lys Lys Ile Asp	His Phe Thr Phe Glu	Lys Ile Asn Asp Leu Leu	
2580	2585	2590	
gtg cca ttt ctt	gac ttg gcc ttt gaa	atg att ggg gta gaa cct	7843
Val Pro Phe Leu	Asp Leu Ala Phe Glu	Met Ile Gly Val Glu Pro	
2595	2600	2605	
tat ata tca tca	aac tct gat att ttc	agt atg tca cct agc ata	7888
Tyr Ile Ser Ser	Asn Ser Asp Ile Phe	Ser Met Ser Pro Ser Ile	
2610	2615	2620	

ctc tca tat atg aac caa tot aag gac ttt tct gat att ttg gaa	7933
Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile Leu Glu	
2625 2630 2635	
gaa att gct gaa ttt tta aca tct gtg aaa atg aac ttg gaa gat	7978
Glu Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu Asp	
2640 2645 2650	
atg agg agt ctt gcg gta gca ttt aac aat gag act caa aca ttt	8023
Met Arg Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe	
2655 2660 2665	
tct atg gat tct gtc aac tta ngg gaa gaa att ctg ggt tgc tta	8068
Ser Met Asp Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu	
2670 2675 2680	
gtt cct ata aat aac atc acc aac caa atg gac ttc tta tac cct	8113
Val Pro Ile Asn Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro	
2685 2690 2695	
aat cca att tcc act cat agt ggc cct caa gat ata aaa tgg gaa	8158
Asn Pro Ile Ser Thr His Ser Gly Pro Gln Asp Ile Lys Trp Glu	
2700 2705 2710	
ata att cat gaa gtg atc ctt ttt ttg gat aaa ata tta tca caa	8203
Ile Ile His Glu Val Ile Leu Phe Leu Asp Lys Ile Leu Ser Gln	
2715 2720 2725	
aac agc aca gaa ata gga tct ttc ttg aaa atg gtg atc tgt ctc	8248
Asn Ser Thr Glu Ile Gly Ser Phe Leu Lys Met Val Ile Cys Leu	
2730 2735 2740	
acc tta gaa gct ctt tgg aaa aac tta aag aaa gat aat tgg aat	8293
Thr Leu Glu Ala Leu Trp Lys Asn Leu Lys Lys Asp Asn Trp Asn	
2745 2750 2755	
gtt tct aat gtg ttg atg acg ttt act cag cat cca aat aac ctt	8338
Val Ser Asn Val Leu Met Thr Phe Thr Gln His Pro Asn Asn Leu	
2760 2765 2770	
ttg aaa acc ata gaa aca gtt tta gag gcc tcc agt gga att aaa	8383
Leu Lys Thr Ile Glu Thr Val Leu Glu Ala Ser Ser Gly Ile Lys	
2775 2780 2785	
agt gac tat gaa ggt gat ttg aat aaa agt tta tat ttt gac aca	8428
Ser Asp Tyr Glu Gly Asp Leu Asn Lys Ser Leu Tyr Phe Asp Thr	
2790 2795 2800	
cct ttg agt cag aat ata act cat cat caa ctt gaa aaa gca atc	8473
Pro Leu Ser Gln Asn Ile Thr His His Gln Leu Glu Lys Ala Ile	
2805 2810 2815	
cat aat gtt tta agt aga ata gct ctc tgg agg aaa gga ctt cgt	8518
His Asn Val Leu Ser Arg Ile Ala Leu Trp Arg Lys Gly Leu Arg	
2820 2825 2830	
ttt aac aac tct gaa tgg ata act tcc aca aga act ttg ttt cag	8563
Phe Asn Asn Ser Glu Trp Ile Thr Ser Thr Arg Thr Leu Phe Gln	
2835 2840 2845	
cca ctt ttt gag att ttc att aaa gca acc acc gga aag aat gtc	8608

Pro	Leu	Phe	Glu	Ile	Phe	Ile	Lys	Ala	Thr	Thr	Gly	Lys	Asn	Val	
			2850					2855					2860		
aca	tca	gaa	aaa	gaa	gag	aga	acc	gag	aaa	gag	atg	att	gac	ttt	8653
Thr	Ser	Glu	Lys	Glu	Glu	Arg	Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	
			2865					2870					2875		
cct	tat	agt	ttc	aaa	cca	ttt	ttc	tgt	ttg	gag	aaa	tac	ctg	gga	8698
Pro	Tyr	Ser	Phe	Lys	Pro	Phe	Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	
			2880					2885					2890		
gga	tta	ttt	gta	ttg	act	aaa	tac	tgg	caa	caa	atc	cca	cta	aca	8743
Gly	Leu	Phe	Val	Leu	Thr	Lys	Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	
			2895					2900					2905		
gat	caa	agt	gtt	gtt	gag	att	tgt	gaa	gtt	ttc	cag	cag	act	gtg	8788
Asp	Gln	Ser	Val	Val	Glu	Ile	Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	
			2910					2915					2920		
aag	ccc	tca	gaa	gcc	atg	gag	atg	ctg	cag	aaa	gtg	aag	atg	atg	8833
Lys	Pro	Ser	Glu	Ala	Met	Glu	Met	Leu	Gln	Lys	Val	Lys	Met	Met	
			2925					2930					2935		
gtc	gta	cgt	gtg	ctc	acc	atc	gtt	gca	gaa	aac	cct	tcc	tgg	acc	8878
Val	Val	Arg	Val	Leu	Thr	Ile	Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	
			2940					2945					2950		
aag	gac	att	ttg	tgt	gct	act	ctg	agt	tgc	aag	caa	aat	ggg	ata	8923
Lys	Asp	Ile	Leu	Cys	Ala	Thr	Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	
			2955					2960					2965		
agg	cat	ctc	att	tta	tct	gct	ata	caa	ggg	gtc	act	ttg	gcg	cag	8968
Arg	His	Leu	Ile	Leu	Ser	Ala	Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	
			2970					2975					2980		
gac	cac	ttc	cag	gaa	att	gaa	aag	ata	tgg	tcc	tcg	ccg	aat	cag	9013
Asp	His	Phe	Gln	Glu	Ile	Glu	Lys	Ile	Trp	Ser	Ser	Pro	Asn	Gln	
			2985					2990					2995		
cta	aat	tgt	gaa	agt	ctt	agc	aag	aat	ctt	tct	agc	acc	ttg	gag	9058
Leu	Asn	Cys	Glu	Ser	Leu	Ser	Lys	Asn	Leu	Ser	Ser	Thr	Leu	Glu	
			3000					3005					3010		
agc	ttc	aag	agc	agc	ttg	gaa	aat	gcc	act	ggc	cag	gac	tgc	aca	9103
Ser	Phe	Lys	Ser	Ser	Leu	Glu	Asn	Ala	Thr	Gly	Gln	Asp	Cys	Thr	
			3015					3020					3025		
agc	cag	ccg	agg	ctg	gag	acg	gtg	cag	cag	cac	ttg	tac	atg	ttg	9148
Ser	Gln	Pro	Arg	Leu	Glu	Thr	Val	Gln	Gln	His	Leu	Tyr	Met	Leu	
			3030					3035					3040		
gcc	aaa	agc	ctc	gag	gaa	act	tgg	tca	tca	ggg	aat	ccc	atc	atg	9193
Ala	Lys	Ser	Leu	Glu	Glu	Thr	Trp	Ser	Ser	Gly	Asn	Pro	Ile	Met	
			3045					3050					3055		
act	ttt	ctc	agc	aat	ttc	aca	gta	act	gag	gat	gta	aaa	ata	aaa	9238
Thr	Phe	Leu	Ser	Asn	Phe	Thr	Val	Thr	Glu	Asp	Val	Lys	Ile	Lys	
			3060					3065					3070		
gat	ttg	atg	aag	aat	atc	acc	aag	ttg	act	gag	gag	ctt	cgc	tct	9283
Asp	Leu	Met	Lys	Asn	Ile	Thr	Lys	Leu	Thr	Glu	Glu	Leu	Arg	Ser	

3075	3080	3085	
tcc atc caa atc Ser Ile Gln Ile 3090	tcg aat gag act atc Ser Asn Glu Thr 3095	cat agc att cta gaa gca His Ser Ile Leu Glu Ala 3100	9328
aat att tcc cac Asn Ile Ser His 3105	tcc aag gtt ctc ttc Ser Lys Val Leu Phe 3110	agt gcc ctc acc gta gct Ser Ala Leu Thr Val Ala 3115	9373
ctg tct gga aag Leu Ser Gly Lys 3120	tgt gat cag gaa atc Cys Asp Gln Glu Ile 3125	ctt cat ctc ctg ctg aca Leu His Leu Leu Leu Thr 3130	9418
ttt ccc aaa ggg Phe Pro Lys Gly 3135	gaa aaa tct tgg atc Glu Lys Ser Trp Ile 3140	gca gcg gag gaa ctc tgt Ala Ala Glu Glu Leu Cys 3145	9463
agc ctg cca ggg Ser Leu Pro Gly 3150	tca aaa gtg tat tct Ser Lys Val Tyr Ser 3155	ctg att gtg ttg ctg agt Leu Ile Val Leu Leu Ser 3160	9508
cga aac ttg gat Arg Asn Leu Asp 3165	gtg cga gct ttc att Val Arg Ala Phe Ile 3170	tac aag act ctg atg cct Tyr Lys Thr Leu Met Pro 3175	9553
tct gaa gca aat Ser Glu Ala Asn 3180	ggc ttg ctc aac tcc Gly Leu Leu Asn Ser 3185	ttg ctg gat ata gtt tcc Leu Leu Asp Ile Val Ser 3190	9598
agc ctc agc gcc Ser Leu Ser Ala 3195	ttg ctt gcc aaa gcc Leu Leu Ala Lys Ala 3200	cag cac gtc ttt gag tat Gln His Val Phe Glu Tyr 3205	9643
ctt cct gag ttt Leu Pro Glu Phe 3210	ctt cac aca ttt aaa Leu His Thr Phe Lys 3215	atc act gcc ttg cta gaa Ile Thr Ala Leu Leu Glu 3220	9688
acc ctg gac ttt Thr Leu Asp Phe 3225	caa cag gtt tca caa Gln Gln Val Ser Gln 3230	aat gtc cag gcc aga agt Asn Val Gln Ala Arg Ser 3235	9733
tca gct ttt ggt Ser Ala Phe Gly 3240	tct ttc cag ttt gtg Ser Phe Gln Phe Val 3245	atg aag atg gtt tgc aag Met Lys Met Val Cys Lys 3250	9778
gac caa gca tca Asp Gln Ala Ser 3255	ttc ctt agc gat tct Phe Leu Ser Asp Ser 3260	aat atg ttt att aat ttg Asn Met Phe Ile Asn Leu 3265	9823
ccc aga gtt aag Pro Arg Val Lys 3270	gaa ctc ttg gaa gat Glu Leu Leu Glu Asp 3275	gac aaa gaa aaa ttc aac Asp Lys Glu Lys Phe Asn 3280	9868
att cct gaa gat Ile Pro Glu Asp 3285	tca aca ccg ttt tgc Ser Thr Pro Phe Cys 3290	ttg aag ctt tat cag gaa Leu Lys Leu Tyr Gln Glu 3295	9913
att cta caa ttg Ile Leu Gln Leu 3300	cca aat ggt gct ttg Pro Asn Gly Ala Leu 3305	gtg tgg acc ttc cta aaa Val Trp Thr Phe Leu Lys 3310	9958

ccc ata ttg cat	gga aaa ata cta tac	aca cca aac act cca gaa	10003
Pro Ile Leu His	Gly Lys Ile Leu Tyr	Thr Pro Asn Thr Pro Glu	
3315	3320	3325	
att aac aag gtc	att caa aag gct aat	tac acc ttt tat att gtg	10048
Ile Asn Lys Val	Ile Gln Lys Ala Asn	Tyr Thr Phe Tyr Ile Val	
3330	3335	3340	
gac aaa cta aaa	act tta tca gaa aca	ctg ctg gaa atg tcc agc	10093
Asp Lys Leu Lys	Thr Leu Ser Glu Thr	Leu Leu Glu Met Ser Ser	
3345	3350	3355	
ctt ttc cag aga	agt gga agt ggc cag	atg ttc aac cag ctg cag	10138
Leu Phe Gln Arg	Ser Gly Ser Gly Gln	Met Phe Asn Gln Leu Gln	
3360	3365	3370	
gag gcc ctg aga	aac aaa ttt gta aga	aac ttt gta gaa aac cag	10183
Glu Ala Leu Arg	Asn Lys Phe Val Arg	Asn Phe Val Glu Asn Gln	
3375	3380	3385	
ttg cac att gat	gta gac aaa ctt act	gaa aaa ctc cag aca tac	10228
Leu His Ile Asp	Val Asp Lys Leu Thr	Glu Lys Leu Gln Thr Tyr	
3390	3395	3400	
gga ggg ctg ctg	gat gag atg ttt aac	cat gca ggc gct gga cgc	10273
Gly Gly Leu Leu	Asp Glu Met Phe Asn	His Ala Gly Ala Gly Arg	
3405	3410	3415	
ttc cgt ttc ttg	ggc agc atc ttg gtc	aat ctc tct tcc tgc gtg	10318
Phe Arg Phe Leu	Gly Ser Ile Leu Val	Asn Leu Ser Ser Cys Val	
3420	3425	3430	
gca ctg aac cgt	ttc cag gct ctg cag	tct gtc gac atc ctg gag	10363
Ala Leu Asn Arg	Phe Gln Ala Leu Gln	Ser Val Asp Ile Leu Glu	
3435	3440	3445	
act aaa gca cat	gaa ctc ttg cag cag	aac agc ttc ttg gcc agt	10408
Thr Lys Ala His	Glu Leu Leu Gln Gln	Asn Ser Phe Leu Ala Ser	
3450	3455	3460	
atc att ttc agc	aat tcc tta ttc gac	aag aac ttc aga tca gag	10453
Ile Ile Phe Ser	Asn Ser Leu Phe Asp	Lys Asn Phe Arg Ser Glu	
3465	3470	3475	
tct gtc aaa ctg	cca ccc cat gtc tca	tac aca atc cgg acc aat	10498
Ser Val Lys Leu	Pro Pro His Val Ser	Tyr Thr Ile Arg Thr Asn	
3480	3485	3490	
gtg tta tac agc	gtg cga aca gat gtg	gta aaa aac cct tct tgg	10543
Val Leu Tyr Ser	Val Arg Thr Asp Val	Val Lys Asn Pro Ser Trp	
3495	3500	3505	
aag ttc cac cct	cag aat cta cca gct	gat ggg ttc aaa tat aac	10588
Lys Phe His Pro	Gln Asn Leu Pro Ala	Asp Gly Phe Lys Tyr Asn	
3510	3515	3520	
tac gtc ttt gcc	cca ctg caa gac atg	atc gaa aga gcc atc att	10633
Tyr Val Phe Ala	Pro Leu Gln Asp Met	Ile Glu Arg Ala Ile Ile	
3525	3530	3535	

ttg gtg cag act	ggg cag gaa gcc ctg	gaa cca gca gca cag act	10678
Leu Val Gln Thr	Gly Gln Glu Ala Leu	Glu Pro Ala Ala Gln Thr	
3540	3545	3550	
cag gcg gcc cct	tac ccc tgc cat acc	agc gac cta ttc ctg aac	10723
Gln Ala Ala Pro	Tyr Pro Cys His Thr	Ser Asp Leu Phe Leu Asn	
3555	3560	3565	
aac gtt ggt ttc	ttt ttt cca ctg ata	atg atg ctg acg tgg atg	10768
Asn Val Gly Phe	Phe Phe Pro Leu Ile	Met Met Leu Thr Trp Met	
3570	3575	3580	
gtg tct gtg gcc	agc atg gtc aga aag	ttg gtg tat gag cag gag	10813
Val Ser Val Ala	Ser Met Val Arg Lys	Leu Val Tyr Glu Gln Glu	
3585	3590	3595	
ata cag ata gaa	gag tat atg cgg atg	atg gga gtg cat cca gtg	10858
Ile Gln Ile Glu	Glu Tyr Met Arg Met	Met Gly Val His Pro Val	
3600	3605	3610	
atc cat ttc ctg	gcc tgg ttc ctg gag	aac atg gct gtg ttg acc	10903
Ile His Phe Leu	Ala Trp Phe Leu Glu	Asn Met Ala Val Leu Thr	
3615	3620	3625	
ata agc agt gct	act ctg gcc atc gtt	ctg aaa aca agt ggc atc	10948
Ile Ser Ser Ala	Thr Leu Ala Ile Val	Leu Lys Thr Ser Gly Ile	
3630	3635	3640	
ttt gca cac agc	aat acc ttt att gtt	ttc ctc ttt ctc ttg gat	10993
Phe Ala His Ser	Asn Thr Phe Ile Val	Phe Leu Phe Leu Leu Asp	
3645	3650	3655	
ttt ggg atg tca	gtc gtc atg ctg agc	tac ctc ttg agt gca ttt	11038
Phe Gly Met Ser	Val Val Met Leu Ser	Tyr Leu Leu Ser Ala Phe	
3660	3665	3670	
ttc agc caa gct	aat aca gcg gcc ctt	tgt acc agc ctg gtg tac	11083
Phe Ser Gln Ala	Asn Thr Ala Ala Leu	Cys Thr Ser Leu Val Tyr	
3675	3680	3685	
atg atc agc ttt	ctg ccc tac ata gtt	cta ttg gtt cta cat aac	11128
Met Ile Ser Phe	Leu Pro Tyr Ile Val	Leu Leu Val Leu His Asn	
3690	3695	3700	
caa tta agt ttt	gtt aat cag aca ttt	ctg tgc ctt ctt tcg aca	11173
Gln Leu Ser Phe	Val Asn Gln Thr Phe	Leu Cys Leu Leu Ser Thr	
3705	3710	3715	
acc gcc ttt gga	caa ggg gta ttt ttt	att aca ttc ctg gaa gga	11218
Thr Ala Phe Gly	Gln Gly Val Phe Phe	Ile Thr Phe Leu Glu Gly	
3720	3725	3730	
caa gag aca ggg	att caa tgg aat aat	atg tac cag gct ctg gaa	11263
Gln Glu Thr Gly	Ile Gln Trp Asn Asn	Met Tyr Gln Ala Leu Glu	
3735	3740	3745	
caa ggg ggc atg	aca ttt ggc tgg gtt	tgc tgg atg att ctt ttt	11308
Gln Gly Gly Met	Thr Phe Gly Trp Val	Cys Trp Met Ile Leu Phe	
3750	3755	3760	
gat tca agc ctt	tat ttt ttg tgt gga	tgg tac ttg agc aac ttg	11353

Asp	Ser	Ser	Leu	Tyr	Phe	Leu	Cys	Gly	Trp	Tyr	Leu	Ser	Asn	Leu	
			3765					3770					3775		
att	cct	gga	aca	ttt	ggt	tta	cgg	aaa	cca	tgg	tat	ttc	ccc	ttt	11398
Ile	Pro	Gly	Thr	Phe	Gly	Leu	Arg	Lys	Pro	Trp	Tyr	Phe	Pro	Phe	
			3780					3785					3790		
act	gcc	tca	tat	tgg	aag	agt	gtg	ggt	ttc	ttg	gtg	gag	aaa	agg	11443
Thr	Ala	Ser	Tyr	Trp	Lys	Ser	Val	Gly	Phe	Leu	Val	Glu	Lys	Arg	
			3795					3800					3805		
caa	tac	ttt	cta	agt	tct	agt	ctg	ttc	ttc	ttc	aat	gag	aac	ttt	11488
Gln	Tyr	Phe	Leu	Ser	Ser	Ser	Leu	Phe	Phe	Phe	Asn	Glu	Asn	Phe	
			3810					3815					3820		
gac	aat	aaa	ggg	tca	tca	ctg	caa	aac	agg	gaa	gga	gag	ctt	gaa	11533
Asp	Asn	Lys	Gly	Ser	Ser	Leu	Gln	Asn	Arg	Glu	Gly	Glu	Leu	Glu	
			3825					3830					3835		
gga	agt	gcc	ccg	gga	gtc	acc	ctg	gtg	tct	gtg	acc	aag	gaa	tat	11578
Gly	Ser	Ala	Pro	Gly	Val	Thr	Leu	Val	Ser	Val	Thr	Lys	Glu	Tyr	
			3840					3845					3850		
gag	ggc	cac	aag	gct	gtg	gtc	caa	gac	ctc	agc	ctg	acc	ttc	tac	11623
Glu	Gly	His	Lys	Ala	Val	Val	Gln	Asp	Leu	Ser	Leu	Thr	Phe	Tyr	
			3855					3860					3865		
aga	gac	caa	atc	acc	gcc	ctg	ctg	ggg	aca	aac	ggt	gcc	ggg	aaa	11668
Arg	Asp	Gln	Ile	Thr	Ala	Leu	Leu	Gly	Thr	Asn	Gly	Ala	Gly	Lys	
			3870					3875					3880		
acc	act	atc	ata	tcc	atg	ttg	acg	ggg	ctc	cac	cct	ccc	act	tct	11713
Thr	Thr	Ile	Ile	Ser	Met	Leu	Thr	Gly	Leu	His	Pro	Pro	Thr	Ser	
			3885					3890					3895		
gga	acc	atc	atc	atc	aat	ggc	aag	aac	cta	cag	aca	gac	ctg	tcg	11758
Gly	Thr	Ile	Ile	Ile	Asn	Gly	Lys	Asn	Leu	Gln	Thr	Asp	Leu	Ser	
			3900					3905					3910		
agg	gtc	aga	atg	gag	ctt	ggt	gtg	tgt	ccg	cag	cag	gac	atc	ctg	11803
Arg	Val	Arg	Met	Glu	Leu	Gly	Val	Cys	Pro	Gln	Gln	Asp	Ile	Leu	
			3915					3920					3925		
ttg	gac	aac	ctc	acc	gtc	cgg	gaa	cat	ttg	ctg	ctc	ttt	gct	tcc	11848
Leu	Asp	Asn	Leu	Thr	Val	Arg	Glu	His	Leu	Leu	Leu	Phe	Ala	Ser	
			3930					3935					3940		
ata	aag	gcg	cct	cag	tgg	acc	aag	aag	gag	ctg	cat	cag	caa	gtc	11893
Ile	Lys	Ala	Pro	Gln	Trp	Thr	Lys	Lys	Glu	Leu	His	Gln	Gln	Val	
			3945					3950					3955		
aat	caa	act	ctt	cag	gat	gtg	gac	tta	act	cag	cat	cag	cac	aaa	11938
Asn	Gln	Thr	Leu	Gln	Asp	Val	Asp	Leu	Thr	Gln	His	Gln	His	Lys	
			3960					3965					3970		
cag	acc	cga	gct	ctg	tct	gga	ggc	ctg	aag	agg	aag	ctc	tcc	ctt	11983
Gln	Thr	Arg	Ala	Leu	Ser	Gly	Gly	Leu	Lys	Arg	Lys	Leu	Ser	Leu	
			3975					3980					3985		
ggc	att	gct	ttc	atg	ggc	atg	tcg	agg	acc	gtg	gtt	ctg	gat	gag	12028
Gly	Ile	Ala	Phe	Met	Gly	Met	Ser	Arg	Thr	Val	Val	Leu	Asp	Glu	

3990	3995	4000	
ccc acc agt ggg Pro Thr Ser Gly 4005	gtg gac cct tgc tcc Val Asp Pro Cys Ser 4010	cgg cat agc ctg tgg gac Arg His Ser Leu Trp Asp 4015	12073
att ctg ctc aag Ile Leu Leu Lys 4020	tac cga gaa ggt cgt Tyr Arg Glu Gly Arg 4025	acg atc atc ttc aca acc Thr Ile Ile Phe Thr Thr 4030	12118
cac cac ctg gat His His Leu Asp 4035	gaa gcc gaa gcg ctg Glu Ala Glu Ala Leu 4040	agt gac cgc gtg gcc gtc Ser Asp Arg Val Ala Val 4045	12163
ctc cag cat ggg Leu Gln His Gly 4050	agg ctc agg tgc tgc Arg Leu Arg Cys Cys 4055	ggt cct ccc ttc tgc ctg Gly Pro Pro Phe Cys Leu 4060	12208
aag gag gca tat Lys Glu Ala Tyr 4065	ggc cag ggg ctc cgc Gly Gln Gly Leu Arg 4070	ctg aca ctc acg agg cag Leu Thr Leu Thr Arg Gln 4075	12253
cct tct gtt ctg Pro Ser Val Leu 4080	gag gcc cat gat ctg Glu Ala His Asp Leu 4085	aaa gac atg gct tgt gtt Lys Asp Met Ala Cys Val 4090	12298
aca tcc ctg ata Thr Ser Leu Ile 4095	aag atc tat att cca Lys Ile Tyr Ile Pro 4100	caa gca ttt ctc aaa gac Gln Ala Phe Leu Lys Asp 4105	12343
agc agt gga agt Ser Ser Gly Ser 4110	gag ctg acc tac acc Glu Leu Thr Tyr Thr 4115	att cca aag gac aca gac Ile Pro Lys Asp Thr Asp 4120	12388
aag gcc tgc ttg Lys Ala Cys Leu 4125	aaa ggg ctc ttc cag Lys Gly Leu Phe Gln 4130	gcc ctg gat gag aac ctg Ala Leu Asp Glu Asn Leu 4135	12433
cat cag ctg cac His Gln Leu His 4140	ctg acg ggc tat ggg Leu Thr Gly Tyr Gly 4145	atc tca gac acc acc tta Ile Ser Asp Thr Thr Leu 4150	12478
gaa gag gtg ttt Glu Glu Val Phe 4155	ttg atg ctt ttg caa Leu Met Leu Leu Gln 4160	gat tcc aac aag aaa tct Asp Ser Asn Lys Lys Ser 4165	12523
cac att gcc ctg His Ile Ala Leu 4170	ggg act gag tca gag Gly Thr Glu Ser Glu 4175	ctg cag aac cac agg cct Leu Gln Asn His Arg Pro 4180	12568
aca gga cat ctg Thr Gly His Leu 4185	tct ggc tac tgt ggc Ser Gly Tyr Cys Gly 4190	tcc cta gca cgg ccc gca Ser Leu Ala Arg Pro Ala 4195	12613
act gtg cag ggc Thr Val Gln Gly 4200	gtc cag ctg ctc cgc Val Gln Leu Leu Arg 4205	gca caa gtg gcc gcg atc Ala Gln Val Ala Ala Ile 4210	12658
ctg gcc cgg agg Leu Ala Arg Arg 4215	ctc cgc cgc acg ctg Leu Arg Arg Thr Leu 4220	cgc gcc ggg aag agc acc Arg Ala Gly Lys Ser Thr 4225	12703

ctc gcc gac ctg	ctg ctg cca gtc ctc	ttc gtg gcc ttg gcc atg	12748
Leu Ala Asp Leu	Leu Leu Pro Val Leu	Phe Val Ala Leu Ala Met	
4230	4235	4240	
ggc ttg ttc atg	gtg aga ccc ctg gcc	acc gag tac cct ccc ctc	12793
Gly Leu Phe Met	Val Arg Pro Leu Ala	Thr Glu Tyr Pro Pro Leu	
4245	4250	4255	
aga ctc aca cct	gga cat tac cag cgg	gcc gag acc tac ttt ttc	12838
Arg Leu Thr Pro	Gly His Tyr Gln Arg	Ala Glu Thr Tyr Phe Phe	
4260	4265	4270	
agc agt ggg ggc	gac aac ttg gac ctc	acc cgt gtg ctt ctg cgg	12883
Ser Ser Gly Gly	Asp Asn Leu Asp Leu	Thr Arg Val Leu Leu Arg	
4275	4280	4285	
aag ttt aga gat	caa gat ttg ccc tgt	gca gat tta aac cca cgc	12928
Lys Phe Arg Asp	Gln Asp Leu Pro Cys	Ala Asp Leu Asn Pro Arg	
4290	4295	4300	
cag aag aat tct	tca tgc tgg cgc aca	gat ccc ttt tct cac cca	12973
Gln Lys Asn Ser	Ser Cys Trp Arg Thr	Asp Pro Phe Ser His Pro	
4305	4310	4315	
gaa ttc cag gat	tca tgt ggc tgc ctg	aag tgt cca aat aga agt	13018
Glu Phe Gln Asp	Ser Cys Gly Cys Leu	Lys Cys Pro Asn Arg Ser	
4320	4325	4330	
gct agt gct ccc	tac ctg acc aac cac	ctg ggc cac aca ctg ttg	13063
Ala Ser Ala Pro	Tyr Leu Thr Asn His	Leu Gly His Thr Leu Leu	
4335	4340	4345	
aat ctc tca ggc	ttc aat atg gag gag	tac ttg ctg gca cca tct	13108
Asn Leu Ser Gly	Phe Asn Met Glu Glu	Tyr Leu Leu Ala Pro Ser	
4350	4355	4360	
gaa aaa cca agg	ctt gga ggt tgg tct	ttt gga tta aaa atc ccc	13153
Glu Lys Pro Arg	Leu Gly Gly Trp Ser	Phe Gly Leu Lys Ile Pro	
4365	4370	4375	
agt gaa gct gga	ggt gca aat gga aac	ata tca aaa ccc cca act	13198
Ser Glu Ala Gly	Gly Ala Asn Gly Asn	Ile Ser Lys Pro Pro Thr	
4380	4385	4390	
ctg gca aag gtg	tgg tat aat cag aag	ggt ttt cat tcc cta cct	13243
Leu Ala Lys Val	Trp Tyr Asn Gln Lys	Gly Phe His Ser Leu Pro	
4395	4400	4405	
tcc tac tta aat	cat cta aac aac ctt	att ttg tgg cag cac cta	13288
Ser Tyr Leu Asn	His Leu Asn Asn Leu	Ile Leu Trp Gln His Leu	
4410	4415	4420	
ccc cct act gtg	gac tgg aga caa tac	gga ata aca ctc tac agc	13333
Pro Pro Thr Val	Asp Trp Arg Gln Tyr	Gly Ile Thr Leu Tyr Ser	
4425	4430	4435	
cac cca tat gga	ggg gcc ttg ctg aac	gag gac aag atc ctg gag	13378
His Pro Tyr Gly	Gly Ala Leu Leu Asn	Glu Asp Lys Ile Leu Glu	
4440	4445	4450	

agc atc cgt cag	tgt gga gtg gcc ctc	tgc atc gtg ctg gga ttc	13423
Ser Ile Arg Gln	Cys Gly Val Ala Leu	Cys Ile Val Leu Gly Phe	
4455	4460	4465	
tcc atc ctg tct	gca tcc atc ggc agc	tct gtg gtg agg gac agg	13468
Ser Ile Leu Ser	Ala Ser Ile Gly Ser	Ser Val Val Arg Asp Arg	
4470	4475	4480	
gtg att gga gcc	aaa agg ttg cag cac	ata agt ggc ctt ggc tac	13513
Val Ile Gly Ala	Lys Arg Leu Gln His	Ile Ser Gly Leu Gly Tyr	
4485	4490	4495	
agg atg tac tgg	ttc aca aac ttc cta	tat gac atg ctc ttt tac	13558
Arg Met Tyr Trp	Phe Thr Asn Phe Leu	Tyr Asp Met Leu Phe Tyr	
4500	4505	4510	
ttg gtt tcc gtc	tgc ctg tgt gtt gcc	gtt att gtc gcc ttc cag	13603
Leu Val Ser Val	Cys Leu Cys Val Ala	Val Ile Val Ala Phe Gln	
4515	4520	4525	
tta aca gct ttt	act ttc cgc aag aac	ttg gca gcc acg gcc ctc	13648
Leu Thr Ala Phe	Thr Phe Arg Lys Asn	Leu Ala Ala Thr Ala Leu	
4530	4535	4540	
ctg ctg tca ctt	ttc gga tat gca act	ctt cca tgg atg tac ctg	13693
Leu Leu Ser Leu	Phe Gly Tyr Ala Thr	Leu Pro Trp Met Tyr Leu	
4545	4550	4555	
atg tcc aga atc	ttt tcc agt tcg gac	gtg gct ttc att tcc tat	13738
Met Ser Arg Ile	Phe Ser Ser Ser Asp	Val Ala Phe Ile Ser Tyr	
4560	4565	4570	
gtc tca cta aac	ttc atc ttt ggc ctt	tgt acc atg ccc ata acc	13783
Val Ser Leu Asn	Phe Ile Phe Gly Leu	Cys Thr Met Pro Ile Thr	
4575	4580	4585	
att atg ccc cgg	ttg cta gcc atc atc	tcc aaa gct aag aat tta	13828
Ile Met Pro Arg	Leu Leu Ala Ile Ile	Ser Lys Ala Lys Asn Leu	
4590	4595	4600	
cag aat atc tat	gat gtc ctc aag tgg	gtc ttt act att ttt cct	13873
Gln Asn Ile Tyr	Asp Val Leu Lys Trp	Val Phe Thr Ile Phe Pro	
4605	4610	4615	
caa ttc tgt ctt	ggc caa gga ctg gta	gaa ctc tgc tat aat cag	13918
Gln Phe Cys Leu	Gly Gln Gly Leu Val	Glu Leu Cys Tyr Asn Gln	
4620	4625	4630	
atc aaa tat gac	ctg acc cac aac ttc	ggc att gat tcc tat gtg	13963
Ile Lys Tyr Asp	Leu Thr His Asn Phe	Gly Ile Asp Ser Tyr Val	
4635	4640	4645	
agt ccc ttt gag	atg aac ttt ctg ggc	tgg atc ttc gtg caa ctg	14008
Ser Pro Phe Glu	Met Asn Phe Leu Gly	Trp Ile Phe Val Gln Leu	
4650	4655	4660	
gcc tcg cag ggc	aca gta ctt ctc ctc	ttg agg gtt ctg cta cac	14053
Ala Ser Gln Gly	Thr Val Leu Leu Leu	Leu Arg Val Leu Leu His	
4665	4670	4675	
tgg gac ctt ctg	cga tgg cca agg ggt	cat tct act ctc caa ggc	14098

Trp	Asp	Leu	Leu	Arg	Trp	Pro	Arg	Gly	His	Ser	Thr	Leu	Gln	Gly	
			4680					4685					4690		
aca	gtc	aaa	tct	tct	aag	gat	aca	gat	gtt	gaa	aaa	gag	gaa	aag	14143
Thr	Val	Lys	Ser	Ser	Lys	Asp	Thr	Asp	Val	Glu	Lys	Glu	Glu	Lys	
			4695					4700					4705		
aga	gtg	ttt	gaa	gga	agg	acc	aat	gga	gac	att	ctt	gtg	tta	tac	14188
Arg	Val	Phe	Glu	Gly	Arg	Thr	Asn	Gly	Asp	Ile	Leu	Val	Leu	Tyr	
			4710					4715					4720		
aac	ctt	agt	aaa	cat	tat	cga	cgc	ttt	ttc	cag	aat	att	att	gct	14233
Asn	Leu	Ser	Lys	His	Tyr	Arg	Arg	Phe	Phe	Gln	Asn	Ile	Ile	Ala	
			4725					4730					4735		
gtg	caa	gat	att	agt	ttg	ggc	ata	cca	aaa	gga	gag	tgc	ttt	gga	14278
Val	Gln	Asp	Ile	Ser	Leu	Gly	Ile	Pro	Lys	Gly	Glu	Cys	Phe	Gly	
			4740					4745					4750		
ctt	cta	ggg	gtg	aat	gga	gct	ggg	aag	agc	acg	act	ttc	aaa	atg	14323
Leu	Leu	Gly	Val	Asn	Gly	Ala	Gly	Lys	Ser	Thr	Thr	Phe	Lys	Met	
			4755					4760					4765		
ctg	aat	ggg	gaa	gtt	tct	cta	act	tca	gga	cat	gct	atc	atc	agg	14368
Leu	Asn	Gly	Glu	Val	Ser	Leu	Thr	Ser	Gly	His	Ala	Ile	Ile	Arg	
			4770					4775					4780		
act	ccc	atg	gga	gac	gcc	gtg	gac	ctg	tct	tct	gct	ggc	acg	gca	14413
Thr	Pro	Met	Gly	Asp	Ala	Val	Asp	Leu	Ser	Ser	Ala	Gly	Thr	Ala	
			4785					4790					4795		
ggc	gtg	ctc	att	ggc	tac	tgt	ccc	cag	cag	gat	gcc	ctg	gac	gag	14458
Gly	Val	Leu	Ile	Gly	Tyr	Cys	Pro	Gln	Gln	Asp	Ala	Leu	Asp	Glu	
			4800					4805					4810		
ctt	ctg	act	ggg	tgg	gaa	cat	ctc	tat	tat	tac	tgt	agc	tta	cgc	14503
Leu	Leu	Thr	Gly	Trp	Glu	His	Leu	Tyr	Tyr	Tyr	Cys	Ser	Leu	Arg	
			4815					4820					4825		
ggg	att	cca	agg	cag	tgc	atc	cct	gag	gtt	gct	gga	gac	ctc	atc	14548
Gly	Ile	Pro	Arg	Gln	Cys	Ile	Pro	Glu	Val	Ala	Gly	Asp	Leu	Ile	
			4830					4835					4840		
agg	cgc	tta	cac	ctc	gaa	gcc	cac	gcg	gac	aaa	cct	gtg	gcc	acc	14593
Arg	Arg	Leu	His	Leu	Glu	Ala	His	Ala	Asp	Lys	Pro	Val	Ala	Thr	
			4845					4850					4855		
tac	agt	ggg	gga	acc	aag	cgg	aaa	ctc	tct	aca	gcc	ctg	gcc	ctg	14638
Tyr	Ser	Gly	Gly	Thr	Lys	Arg	Lys	Leu	Ser	Thr	Ala	Leu	Ala	Leu	
			4860					4865					4870		
gtg	ggg	aaa	cct	gac	att	ctt	tta	ttg	gat	gag	ccc	agc	tct	ggg	14683
Val	Gly	Lys	Pro	Asp	Ile	Leu	Leu	Leu	Asp	Glu	Pro	Ser	Ser	Gly	
			4875					4880					4885		
atg	gat	ccc	tgc	tct	aag	cgg	tac	ctg	tgg	caa	aca	ata	atg	aag	14728
Met	Asp	Pro	Cys	Ser	Lys	Arg	Tyr	Leu	Trp	Gln	Thr	Ile	Met	Lys	
			4890					4895					4900		
gag	gtt	cgg	gaa	ggc	tgt	gct	gcg	gtg	ctg	acc	tcc	cac	agc	atg	14773
Glu	Val	Arg	Glu	Gly	Cys	Ala	Ala	Val	Leu	Thr	Ser	His	Ser	Met	

4905	4910	4915	
gag gag tgt gag gct ctt tgc aca	aga ctg gcc ata atg gtt aac	14818	
Glu Glu Cys Glu Ala Leu Cys Thr	Arg Leu Ala Ile Met Val Asn		
4920	4925	4930	
ggc agc ttc aaa tgt ctt ggt tct cct	cag cac atc aaa aat agg	14863	
Gly Ser Phe Lys Cys Leu Gly Ser Pro	Gln His Ile Lys Asn Arg		
4935	4940	4945	
aag tta aat aag aca aac tga agtcaactgg	tgaccaagca catggagaag	14914	
Lys Leu Asn Lys Thr Asn			
4950			
gcagcatggg tgcaggagac gctgaagaac	acgataatag aactttacag ccggctgac	14974	
tgagnttgag aaaaggaggt gcctttgggt	ttgggacatg gaaagatagt gatgtcttta	15034	
gagaaacaga agttgggtgg gctgtgaggg	tgaggaggat tctgggtggt gagtatgttc	15094	
ctagggacac tgatgggctg tttggagaat	tcctacgagg aagggtggt gatattcatg	15154	
ttctctttgc tatggcaaga agcaggagag	aaaacttaaa gttgggggtg attattgctg	15214	
atggttgaga cggaggaata agtagctgaa	ctcatcacia ctttgaagca atatttataa	15274	
actctgaagt agagtgaagt gagctcagat	tcattgttgct gcattttatt taattattac	15334	
taaatatagg caatttgcca aatgctttct	gtgcatggtc ttattttaatc tttattacgt	15394	
aatggctatt attgtaccca cttcacaata	aagaaactga gacctaaaga atttaagcag	15454	
ttggcccca aacagagtca gtattgagaa	aagcagtttt caaactcaat ttcttgactc	15514	
taagtttctt attctttcta taccatttca	atagtctcaa ccaaccttat gtataactga	15574	
ctaaagaaca aagctttgct tcaagtacat	gtcgtgggca cctgctgat gaaacctaca	15634	
cactttactc tttgttgaat ttacttcctg	tgtatttacc gagaaaacag actatcttct	15694	
tttaaaaaga tgcacagact ctgctttgaa	agcctggctt gatgaaataa aagatattta	15754	
gagttaaaaa aaaaaaaaaa aaaaa		15779	

<210> 27

<211> 4952

<212> PRT

<213> Homo Sapiens

<220>

<221> misc_feature

<222> (283)..(283)

<223> The 'Xaa' at location 283 stands for Asp, Gly, Ala, or Val.

<220>

<221> misc_feature

<222> (2674)..(2674)

<223> The 'Xaa' at location 2674 stands for Arg, Gly, or Trp.

<400> 27

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

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

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

Val Glu Phe Phe Glu Lys Leu Leu Leu Pro Asn Leu Phe Asp Ser Ser
 740 745 750

Ile Val Pro Ser Phe His Ser Leu Pro Ser Leu Thr Glu Asp Ile Leu
 755 760 765

Asn Ile Ser Ser Leu Trp Thr Asn His Leu Lys Ser Leu Lys Arg Asp
 770 775 780

Pro Ser Ala Thr Asp Ala Gln Lys Leu Leu Glu Phe Gly Asn Glu Val
 785 790 795 800

Ile Trp Lys Met Gln Thr Leu Gly Ser His Trp Ile Arg Lys Glu Pro
 805 810 815

Lys Asn Leu Leu Arg Phe Ile Glu Leu Ile Leu Phe Glu Ile Asn Pro
 820 825 830

Lys Leu Leu Glu Leu Trp Ala Tyr Gly Ile Ser Lys Gly Lys Arg Ala
 835 840 845

Lys Leu Glu Asn Phe Phe Thr Leu Leu Asn Phe Ser Val Pro Glu Asn
 850 855 860

Glu Ile Leu Ser Thr Ser Phe Asn Phe Ser Gln Leu Phe His Ser Asp
 865 870 875 880

Trp Pro Lys Ser Pro Ala Met Asn Ile Asp Phe Val Arg Leu Ser Glu
 885 890 895

Ala Ile Ile Thr Ser Leu His Glu Phe Gly Phe Leu Glu Gln Glu Gln
 900 905 910

Ile Ser Glu Ala Leu Asn Thr Val Tyr Ala Ile Arg Asn Ala Ser Asp
 915 920 925

Leu Phe Ser Ala Leu Ser Glu Pro Gln Lys Gln Glu Val Asp Lys Ile
 930 935 940

Leu Thr His Ile His Leu Asn Val Phe Gln Asp Lys Asp Ser Ala Leu
 945 950 955 960

Leu Leu Gln Ile Tyr Ser Ser Phe Tyr Arg Tyr Ile Tyr Glu Leu Leu
 965 970 975

Asn Ile Gln Ser Arg Gly Ser Ser Leu Thr Phe Leu Thr Gln Ile Ser
 980 985 990

Lys His Ile Leu Asp Ile Ile Lys Gln Phe Asn Phe Gln Asn Ile Ser
 995 1000 1005

Lys Ala Phe Ala Phe Leu Phe Lys Thr Ala Glu Val Leu Gly Gly
 1010 1015 1020

Ile Ser Asn Val Ser Tyr Cys Gln Gln Leu Leu Ser Ile Phe Asn
 1025 1030 1035

Phe Leu Glu Leu Gln Ala Gln Ser Phe Met Ser Thr Glu Gly Gln
 1040 1045 1050

Glu Leu Glu Val Ile His Thr Thr Leu Thr Gly Leu Lys Gln Leu
 1055 1060 1065

Leu Ile Ile Asp Glu Asp Phe Arg Ile Ser Leu Phe Gln Tyr Met
 1070 1075 1080

Ser Gln Phe Phe Asn Ser Ser Val Glu Asp Leu Leu Asp Asn Lys
 1085 1090 1095

Cys Leu Ile Ser Asp Asn Lys His Ile Ser Ser Val Asn Tyr Ser
 1100 1105 1110

Thr Ser Glu Glu Ser Ser Phe Val Phe Pro Leu Ala Gln Ile Phe
 1115 1120 1125

Ser Asn Leu Ser Ala Asn Val Ser Val Phe Asn Lys Phe Met Ser
 1130 1135 1140

Ile His Cys Thr Val Ser Trp Leu Gln Met Trp Thr Glu Ile Trp
 1145 1150 1155

Glu Thr Ile Ser Gln Leu Phe Lys Phe Asp Met Asn Val Phe Thr
 1160 1165 1170

Ser Leu His His Gly Phe Thr Gln Leu Leu Asp Glu Leu Glu Asp
 1175 1180 1185

Asp Val Lys Val Ser Lys Ser Cys Gln Gly Ile Leu Pro Thr His
 1190 1195 1200

Asn Val	Ala Arg	Leu Ile	Leu	Asn Leu	Phe Lys	Asn	Val Thr	Gln
1205			1210			1215		
Ala Asn	Asp Phe	His Asn	Trp	Glu Asp	Phe Leu	Asp	Leu Arg	Asp
1220			1225			1230		
Phe Leu	Val Ala	Leu Gly	Asn	Ala Leu	Val Ser	Val	Lys Lys	Leu
1235			1240			1245		
Asn Leu	Glu Gln	Val Glu	Lys	Ser Leu	Phe Thr	Met	Glu Ala	Ala
1250			1255			1260		
Leu His	Gln Leu	Lys Thr	Phe	Pro Phe	Asn Glu	Ser	Thr Ser	Arg
1265			1270			1275		
Glu Phe	Leu Asn	Ser Leu	Leu	Glu Val	Phe Ile	Glu	Phe Ser	Ser
1280			1285			1290		
Thr Ser	Glu Tyr	Ile Val	Arg	Asn Leu	Asp Ser	Ile	Asn Asp	Phe
1295			1300			1305		
Leu Ser	Asn Asn	Leu Thr	Asn	Tyr Gly	Glu Lys	Phe	Glu Asn	Ile
1310			1315			1320		
Ile Thr	Glu Leu	Arg Glu	Ala	Ile Val	Phe Leu	Arg	Asn Val	Ser
1325			1330			1335		
His Asp	Arg Asp	Leu Phe	Ser	Cys Ala	Asp Ile	Phe	Gln Asn	Val
1340			1345			1350		
Thr Glu	Cys Ile	Leu Glu	Asp	Gly Phe	Leu Tyr	Val	Asn Thr	Ser
1355			1360			1365		
Gln Arg	Met Leu	Arg Ile	Leu	Asp Thr	Leu Asn	Ser	Thr Phe	Ser
1370			1375			1380		
Ser Glu	Asn Thr	Ile Ser	Ser	Leu Lys	Gly Cys	Ile	Val Trp	Leu
1385			1390			1395		
Asp Val	Ile Asn	His Leu	Tyr	Leu Leu	Ser Asn	Ser	Ser Phe	Ser
1400			1405			1410		
Gln Gly	Arg Leu	Gln Asn	Ile	Leu Gly	Asn Phe	Arg	Asp Ile	Glu
1415			1420			1425		

Asn Lys Met Asn Ser Ile Leu Lys Ile Val Thr Trp Val Leu Asn		
1430	1435	1440
Ile Lys Lys Pro Leu Cys Ser Ser Asn Gly Ser His Ile Asn Cys		
1445	1450	1455
Val Asn Ile Tyr Leu Lys Asp Val Thr Asp Phe Leu Asn Ile Val		
1460	1465	1470
Leu Thr Thr Val Phe Glu Lys Glu Lys Lys Pro Lys Phe Glu Ile		
1475	1480	1485
Leu Leu Ala Leu Leu Asn Asp Ser Thr Lys Gln Val Arg Met Ser		
1490	1495	1500
Ile Asn Asn Leu Thr Thr Asp Phe Asp Phe Ala Ser Gln Ser Asn		
1505	1510	1515
Trp Arg Tyr Phe Thr Glu Leu Ile Leu Arg Pro Ile Glu Met Ser		
1520	1525	1530
Asp Glu Ile Pro Asn Gln Phe Gln Asn Ile Trp Leu His Leu Ile		
1535	1540	1545
Thr Leu Gly Lys Glu Phe Gln Lys Leu Val Lys Gly Ile Tyr Phe		
1550	1555	1560
Asn Ile Leu Glu Asn Asn Ser Ser Ser Lys Thr Glu Asn Leu Leu		
1565	1570	1575
Asn Ile Phe Ala Thr Ser Pro Lys Glu Lys Asp Val Asn Ser Val		
1580	1585	1590
Gly Asn Ser Ile Tyr His Leu Ala Ser Tyr Leu Ala Phe Ser Leu		
1595	1600	1605
Ser His Asp Leu Gln Asn Ser Pro Lys Ile Ile Ile Ser Pro Glu		
1610	1615	1620
Ile Met Lys Ala Thr Gly Leu Gly Ile Gln Leu Ile Arg Asp Val		
1625	1630	1635
Phe Asn Ser Leu Met Pro Val Val His His Thr Ser Pro Gln Asn		
1640	1645	1650
Ala Gly Tyr Met Gln Ala Leu Lys Lys Val Thr Ser Val Met Arg		

1655	1660	1665
Thr Leu Lys Lys Ala Asp 1670	Ile Asp Leu Leu Val 1675	Asp Gln Leu Glu 1680
Gln Val Ser Val Asn Leu Met 1685	Asp Phe Phe Lys 1690	Asn Ile Ser Ser 1695
Val Gly Thr Gly Asn Leu Val 1700	Val Asn Leu Leu Val 1705	Gly Leu Met 1710
Glu Lys Phe Ala Asp Ser Ser 1715	His Ser Trp Asn Val 1720	Asn His Leu 1725
Leu Gln Leu Ser Arg Leu Phe 1730	Pro Lys Asp Val Val 1735	Asp Ala Val 1740
Ile Asp Val Tyr Tyr Val Leu 1745	Pro His Ala Val Arg 1750	Leu Leu Gln 1755
Gly Val Pro Gly Lys Asn Ile 1760	Thr Glu Gly Leu Lys 1765	Asp Val Tyr 1770
Ser Phe Thr Leu Leu His Gly 1775	Ile Thr Ile Ser Asn 1780	Ile Thr Lys 1785
Glu Asp Phe Ala Ile Val Ile 1790	Lys Ile Leu Leu Asp 1795	Thr Ile Glu 1800
Leu Val Ser Asp Lys Pro Asp 1805	Ile Ile Ser Glu Ala 1810	Leu Ala Cys 1815
Phe Pro Val Val Trp Cys Trp 1820	Asn His Thr Asn Ser 1825	Gly Phe Arg 1830
Gln Asn Ser Lys Ile Asp Pro 1835	Cys Asn Val His Gly 1840	Leu Met Ser 1845
Ser Ser Phe Tyr Gly Lys Val 1850	Ala Ser Ile Leu Asp 1855	His Phe His 1860
Leu Ser Pro Gln Gly Glu Asp 1865	Ser Pro Cys Ser Asn 1870	Glu Ser Ser 1875
Arg Met Glu Ile Thr Arg Lys 1880	Val Val Cys Ile Ile 1885	His Glu Leu 1890

Val Asp Trp Asn Ser Ile Leu Leu Glu Leu Ser Glu Val Phe His
 1895 1900 1905
 Val Asn Ile Ser Leu Val Lys Thr Val Gln Lys Phe Trp His Lys
 1910 1915 1920
 Ile Leu Pro Phe Val Pro Pro Ser Ile Asn Gln Thr Arg Asp Ser
 1925 1930 1935
 Ile Ser Glu Leu Cys Pro Ser Gly Ser Ile Lys Gln Val Ala Leu
 1940 1945 1950
 Gln Ile Ile Glu Lys Leu Lys Asn Val Asn Phe Thr Lys Val Thr
 1955 1960 1965
 Ser Gly Glu Asn Ile Leu Asp Lys Leu Ser Ser Leu Asn Lys Ile
 1970 1975 1980
 Leu Asn Ile Asn Glu Asp Thr Glu Thr Ser Val Gln Asn Ile Ile
 1985 1990 1995
 Ser Ser Asn Leu Glu Arg Thr Val Gln Leu Ile Ser Glu Asp Trp
 2000 2005 2010
 Ser Leu Glu Lys Ser Thr His Asn Leu Leu Ser Leu Phe Met Met
 2015 2020 2025
 Leu Gln Asn Ala Asn Val Thr Gly Ser Ser Leu Glu Ala Leu Ser
 2030 2035 2040
 Ser Phe Ile Glu Lys Ser Glu Thr Pro Tyr Asn Phe Glu Glu Leu
 2045 2050 2055
 Trp Pro Lys Phe Gln Gln Ile Met Lys Asp Leu Thr Gln Asp Phe
 2060 2065 2070
 Arg Ile Arg His Leu Leu Ser Glu Met Asn Lys Gly Ile Lys Ser
 2075 2080 2085
 Ile Asn Ser Met Ala Leu Gln Lys Ile Thr Leu Gln Phe Ala His
 2090 2095 2100
 Phe Leu Glu Ile Leu Asp Ser Pro Ser Leu Lys Thr Leu Glu Ile
 2105 2110 2115

Ile	Glu	Asp	Phe	Leu	Leu	Val	Thr	Lys	Asn	Trp	Leu	Gln	Glu	Tyr
2120						2125					2130			
Ala	Asn	Glu	Asp	Tyr	Ser	Arg	Met	Ile	Glu	Thr	Leu	Phe	Ile	Pro
2135						2140					2145			
Val	Thr	Asn	Glu	Ser	Ser	Thr	Glu	Asp	Ile	Ala	Leu	Leu	Ala	Lys
2150						2155					2160			
Ala	Ile	Ala	Thr	Phe	Trp	Gly	Ser	Leu	Lys	Asn	Ile	Ser	Arg	Ala
2165						2170					2175			
Gly	Asn	Phe	Asp	Val	Ala	Phe	Leu	Thr	His	Leu	Leu	Asn	Gln	Glu
2180						2185					2190			
Gln	Leu	Thr	Asn	Phe	Ser	Val	Val	Gln	Leu	Leu	Phe	Glu	Asn	Ile
2195						2200					2205			
Leu	Ile	Asn	Leu	Ile	Asn	Asn	Leu	Ala	Gly	Asn	Ser	Gln	Glu	Ala
2210						2215					2220			
Ala	Trp	Asn	Leu	Asn	Asp	Thr	Asp	Leu	Gln	Ile	Met	Asn	Phe	Ile
2225						2230					2235			
Asn	Leu	Ile	Leu	Asn	His	Met	Gln	Ser	Glu	Thr	Ser	Arg	Lys	Thr
2240						2245					2250			
Val	Leu	Ser	Leu	Arg	Ser	Ile	Val	Asp	Phe	Thr	Glu	Gln	Phe	Leu
2255						2260					2265			
Lys	Thr	Phe	Phe	Ser	Leu	Phe	Leu	Lys	Glu	Asp	Ser	Glu	Asn	Lys
2270						2275					2280			
Ile	Ser	Leu	Leu	Leu	Lys	Tyr	Phe	His	Lys	Asp	Val	Ile	Ala	Glu
2285						2290					2295			
Met	Ser	Phe	Val	Pro	Lys	Asp	Lys	Ile	Leu	Glu	Ile	Leu	Lys	Leu
2300						2305					2310			
Asp	Gln	Phe	Leu	Thr	Leu	Met	Ile	Gln	Asp	Arg	Leu	Met	Asn	Ile
2315						2320					2325			
Phe	Ser	Ser	Leu	Lys	Glu	Thr	Ile	Tyr	His	Leu	Met	Lys	Ser	Ser
2330						2335					2340			

Phe	Ile	Leu	Asp	Asn	Gly	Glu	Phe	Tyr	Phe	Asp	Thr	His	Gln	Gly
2345						2350					2355			
Leu	Lys	Phe	Met	Gln	Asp	Leu	Phe	Asn	Ala	Leu	Leu	Arg	Glu	Thr
2360						2365					2370			
Ser	Met	Lys	Asn	Lys	Thr	Glu	Asn	Asn	Ile	Asp	Phe	Phe	Thr	Val
2375						2380					2385			
Val	Ser	Gln	Leu	Phe	Phe	His	Val	Asn	Lys	Ser	Glu	Asp	Leu	Phe
2390						2395					2400			
Lys	Leu	Asn	Gln	Asp	Leu	Gly	Ser	Ala	Leu	His	Leu	Val	Arg	Glu
2405						2410					2415			
Cys	Ser	Thr	Glu	Met	Ala	Arg	Leu	Leu	Asp	Thr	Ile	Leu	His	Ser
2420						2425					2430			
Pro	Asn	Lys	Asp	Phe	Tyr	Ala	Leu	Tyr	Pro	Thr	Leu	Gln	Glu	Val
2435						2440					2445			
Ile	Leu	Ala	Asn	Leu	Thr	Asp	Leu	Leu	Phe	Phe	Ile	Asn	Asn	Ser
2450						2455					2460			
Phe	Pro	Leu	Arg	Asn	Arg	Ala	Thr	Leu	Glu	Ile	Thr	Lys	Arg	Leu
2465						2470					2475			
Val	Gly	Ala	Ile	Ser	Arg	Ala	Ser	Glu	Glu	Ser	His	Val	Leu	Lys
2480						2485					2490			
Pro	Leu	Leu	Glu	Met	Ser	Gly	Thr	Leu	Val	Met	Leu	Leu	Asn	Asp
2495						2500					2505			
Ser	Ala	Asp	Leu	Arg	Asp	Leu	Ala	Thr	Ser	Met	Asp	Ser	Ile	Val
2510						2515					2520			
Lys	Leu	Leu	Lys	Leu	Val	Lys	Lys	Val	Ser	Gly	Lys	Met	Ser	Thr
2525						2530					2535			
Val	Phe	Lys	Thr	His	Phe	Ile	Ser	Asn	Thr	Lys	Asp	Ser	Val	Lys
2540						2545					2550			
Phe	Phe	Asp	Thr	Leu	Tyr	Ser	Ile	Met	Gln	Gln	Ser	Val	Gln	Asn
2555						2560					2565			
Leu	Val	Lys	Glu	Ile	Ala	Thr	Leu	Lys	Lys	Ile	Asp	His	Phe	Thr

2570		2575		2580
Phe Glu Lys Ile Asn Asp Leu Leu Val Pro Phe Leu Asp Leu Ala				
2585		2590		2595
Phe Glu Met Ile Gly Val Glu Pro Tyr Ile Ser Ser Asn Ser Asp				
2600		2605		2610
Ile Phe Ser Met Ser Pro Ser Ile Leu Ser Tyr Met Asn Gln Ser				
2615		2620		2625
Lys Asp Phe Ser Asp Ile Leu Glu Glu Ile Ala Glu Phe Leu Thr				
2630		2635		2640
Ser Val Lys Met Asn Leu Glu Asp Met Arg Ser Leu Ala Val Ala				
2645		2650		2655
Phe Asn Asn Glu Thr Gln Thr Phe Ser Met Asp Ser Val Asn Leu				
2660		2665		2670
Xaa Glu Glu Ile Leu Gly Cys Leu Val Pro Ile Asn Asn Ile Thr				
2675		2680		2685
Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro Ile Ser Thr His Ser				
2690		2695		2700
Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His Glu Val Ile Leu				
2705		2710		2715
Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu Ile Gly Ser				
2720		2725		2730
Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu Trp Lys				
2735		2740		2745
Asn Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met Thr				
2750		2755		2760
Phe Thr Gln His Pro Asn Asn Leu Leu Lys Thr Ile Glu Thr Val				
2765		2770		2775
Leu Glu Ala Ser Ser Gly Ile Lys Ser Asp Tyr Glu Gly Asp Leu				
2780		2785		2790
Asn Lys Ser Leu Tyr Phe Asp Thr Pro Leu Ser Gln Asn Ile Thr				
2795		2800		2805

His	His	Gln	Leu	Glu	Lys	Ala	Ile	His	Asn	Val	Leu	Ser	Arg	Ile
2810						2815					2820			
Ala	Leu	Trp	Arg	Lys	Gly	Leu	Arg	Phe	Asn	Asn	Ser	Glu	Trp	Ile
2825						2830					2835			
Thr	Ser	Thr	Arg	Thr	Leu	Phe	Gln	Pro	Leu	Phe	Glu	Ile	Phe	Ile
2840						2845					2850			
Lys	Ala	Thr	Thr	Gly	Lys	Asn	Val	Thr	Ser	Glu	Lys	Glu	Glu	Arg
2855						2860					2865			
Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	Pro	Tyr	Ser	Phe	Lys	Pro	Phe
2870						2875					2880			
Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	Gly	Leu	Phe	Val	Leu	Thr	Lys
2885						2890					2895			
Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	Asp	Gln	Ser	Val	Val	Glu	Ile
2900						2905					2910			
Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	Lys	Pro	Ser	Glu	Ala	Met	Glu
2915						2920					2925			
Met	Leu	Gln	Lys	Val	Lys	Met	Met	Val	Val	Arg	Val	Leu	Thr	Ile
2930						2935					2940			
Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	Lys	Asp	Ile	Leu	Cys	Ala	Thr
2945						2950					2955			
Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	Arg	His	Leu	Ile	Leu	Ser	Ala
2960						2965					2970			
Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	Asp	His	Phe	Gln	Glu	Ile	Glu
2975						2980					2985			
Lys	Ile	Trp	Ser	Ser	Pro	Asn	Gln	Leu	Asn	Cys	Glu	Ser	Leu	Ser
2990						2995					3000			
Lys	Asn	Leu	Ser	Ser	Thr	Leu	Glu	Ser	Phe	Lys	Ser	Ser	Leu	Glu
3005						3010					3015			
Asn	Ala	Thr	Gly	Gln	Asp	Cys	Thr	Ser	Gln	Pro	Arg	Leu	Glu	Thr
3020						3025					3030			

Val	Gln	Gln	His	Leu	Tyr	Met	Leu	Ala	Lys	Ser	Leu	Glu	Glu	Thr
3035						3040					3045			
Trp	Ser	Ser	Gly	Asn	Pro	Ile	Met	Thr	Phe	Leu	Ser	Asn	Phe	Thr
3050						3055					3060			
Val	Thr	Glu	Asp	Val	Lys	Ile	Lys	Asp	Leu	Met	Lys	Asn	Ile	Thr
3065						3070					3075			
Lys	Leu	Thr	Glu	Glu	Leu	Arg	Ser	Ser	Ile	Gln	Ile	Ser	Asn	Glu
3080						3085					3090			
Thr	Ile	His	Ser	Ile	Leu	Glu	Ala	Asn	Ile	Ser	His	Ser	Lys	Val
3095						3100					3105			
Leu	Phe	Ser	Ala	Leu	Thr	Val	Ala	Leu	Ser	Gly	Lys	Cys	Asp	Gln
3110						3115					3120			
Glu	Ile	Leu	His	Leu	Leu	Leu	Thr	Phe	Pro	Lys	Gly	Glu	Lys	Ser
3125						3130					3135			
Trp	Ile	Ala	Ala	Glu	Glu	Leu	Cys	Ser	Leu	Pro	Gly	Ser	Lys	Val
3140						3145					3150			
Tyr	Ser	Leu	Ile	Val	Leu	Leu	Ser	Arg	Asn	Leu	Asp	Val	Arg	Ala
3155						3160					3165			
Phe	Ile	Tyr	Lys	Thr	Leu	Met	Pro	Ser	Glu	Ala	Asn	Gly	Leu	Leu
3170						3175					3180			
Asn	Ser	Leu	Leu	Asp	Ile	Val	Ser	Ser	Leu	Ser	Ala	Leu	Leu	Ala
3185						3190					3195			
Lys	Ala	Gln	His	Val	Phe	Glu	Tyr	Leu	Pro	Glu	Phe	Leu	His	Thr
3200						3205					3210			
Phe	Lys	Ile	Thr	Ala	Leu	Leu	Glu	Thr	Leu	Asp	Phe	Gln	Gln	Val
3215						3220					3225			
Ser	Gln	Asn	Val	Gln	Ala	Arg	Ser	Ser	Ala	Phe	Gly	Ser	Phe	Gln
3230						3235					3240			
Phe	Val	Met	Lys	Met	Val	Cys	Lys	Asp	Gln	Ala	Ser	Phe	Leu	Ser
3245						3250					3255			

Asp	Ser	Asn	Met	Phe	Ile	Asn	Leu	Pro	Arg	Val	Lys	Glu	Leu	Leu
3260						3265					3270			
Glu	Asp	Asp	Lys	Glu	Lys	Phe	Asn	Ile	Pro	Glu	Asp	Ser	Thr	Pro
3275						3280					3285			
Phe	Cys	Leu	Lys	Leu	Tyr	Gln	Glu	Ile	Leu	Gln	Leu	Pro	Asn	Gly
3290						3295					3300			
Ala	Leu	Val	Trp	Thr	Phe	Leu	Lys	Pro	Ile	Leu	His	Gly	Lys	Ile
3305						3310					3315			
Leu	Tyr	Thr	Pro	Asn	Thr	Pro	Glu	Ile	Asn	Lys	Val	Ile	Gln	Lys
3320						3325					3330			
Ala	Asn	Tyr	Thr	Phe	Tyr	Ile	Val	Asp	Lys	Leu	Lys	Thr	Leu	Ser
3335						3340					3345			
Glu	Thr	Leu	Leu	Glu	Met	Ser	Ser	Leu	Phe	Gln	Arg	Ser	Gly	Ser
3350						3355					3360			
Gly	Gln	Met	Phe	Asn	Gln	Leu	Gln	Glu	Ala	Leu	Arg	Asn	Lys	Phe
3365						3370					3375			
Val	Arg	Asn	Phe	Val	Glu	Asn	Gln	Leu	His	Ile	Asp	Val	Asp	Lys
3380						3385					3390			
Leu	Thr	Glu	Lys	Leu	Gln	Thr	Tyr	Gly	Gly	Leu	Leu	Asp	Glu	Met
3395						3400					3405			
Phe	Asn	His	Ala	Gly	Ala	Gly	Arg	Phe	Arg	Phe	Leu	Gly	Ser	Ile
3410						3415					3420			
Leu	Val	Asn	Leu	Ser	Ser	Cys	Val	Ala	Leu	Asn	Arg	Phe	Gln	Ala
3425						3430					3435			
Leu	Gln	Ser	Val	Asp	Ile	Leu	Glu	Thr	Lys	Ala	His	Glu	Leu	Leu
3440						3445					3450			
Gln	Gln	Asn	Ser	Phe	Leu	Ala	Ser	Ile	Ile	Phe	Ser	Asn	Ser	Leu
3455						3460					3465			
Phe	Asp	Lys	Asn	Phe	Arg	Ser	Glu	Ser	Val	Lys	Leu	Pro	Pro	His
3470						3475					3480			

Val	Ser	Tyr	Thr	Ile	Arg	Thr	Asn	Val	Leu	Tyr	Ser	Val	Arg	Thr
3485						3490					3495			
Asp	Val	Val	Lys	Asn	Pro	Ser	Trp	Lys	Phe	His	Pro	Gln	Asn	Leu
3500						3505					3510			
Pro	Ala	Asp	Gly	Phe	Lys	Tyr	Asn	Tyr	Val	Phe	Ala	Pro	Leu	Gln
3515						3520					3525			
Asp	Met	Ile	Glu	Arg	Ala	Ile	Ile	Leu	Val	Gln	Thr	Gly	Gln	Glu
3530						3535					3540			
Ala	Leu	Glu	Pro	Ala	Ala	Gln	Thr	Gln	Ala	Ala	Pro	Tyr	Pro	Cys
3545						3550					3555			
His	Thr	Ser	Asp	Leu	Phe	Leu	Asn	Asn	Val	Gly	Phe	Phe	Phe	Pro
3560						3565					3570			
Leu	Ile	Met	Met	Leu	Thr	Trp	Met	Val	Ser	Val	Ala	Ser	Met	Val
3575						3580					3585			
Arg	Lys	Leu	Val	Tyr	Glu	Gln	Glu	Ile	Gln	Ile	Glu	Glu	Tyr	Met
3590						3595					3600			
Arg	Met	Met	Gly	Val	His	Pro	Val	Ile	His	Phe	Leu	Ala	Trp	Phe
3605						3610					3615			
Leu	Glu	Asn	Met	Ala	Val	Leu	Thr	Ile	Ser	Ser	Ala	Thr	Leu	Ala
3620						3625					3630			
Ile	Val	Leu	Lys	Thr	Ser	Gly	Ile	Phe	Ala	His	Ser	Asn	Thr	Phe
3635						3640					3645			
Ile	Val	Phe	Leu	Phe	Leu	Leu	Asp	Phe	Gly	Met	Ser	Val	Val	Met
3650						3655					3660			
Leu	Ser	Tyr	Leu	Leu	Ser	Ala	Phe	Phe	Ser	Gln	Ala	Asn	Thr	Ala
3665						3670					3675			
Ala	Leu	Cys	Thr	Ser	Leu	Val	Tyr	Met	Ile	Ser	Phe	Leu	Pro	Tyr
3680						3685					3690			
Ile	Val	Leu	Leu	Val	Leu	His	Asn	Gln	Leu	Ser	Phe	Val	Asn	Gln
3695						3700					3705			
Thr	Phe	Leu	Cys	Leu	Leu	Ser	Thr	Thr	Ala	Phe	Gly	Gln	Gly	Val

3710	3715	3720
Phe Phe Ile Thr Phe Leu Glu Gly Gln Glu Thr Gly Ile Gln Trp 3725 3730 3735		
Asn Asn Met Tyr Gln Ala Leu Glu Gln Gly Gly Met Thr Phe Gly 3740 3745 3750		
Trp Val Cys Trp Met Ile Leu Phe Asp Ser Ser Leu Tyr Phe Leu 3755 3760 3765		
Cys Gly Trp Tyr Leu Ser Asn Leu Ile Pro Gly Thr Phe Gly Leu 3770 3775 3780		
Arg Lys Pro Trp Tyr Phe Pro Phe Thr Ala Ser Tyr Trp Lys Ser 3785 3790 3795		
Val Gly Phe Leu Val Glu Lys Arg Gln Tyr Phe Leu Ser Ser Ser 3800 3805 3810		
Leu Phe Phe Phe Asn Glu Asn Phe Asp Asn Lys Gly Ser Ser Leu 3815 3820 3825		
Gln Asn Arg Glu Gly Glu Leu Glu Gly Ser Ala Pro Gly Val Thr 3830 3835 3840		
Leu Val Ser Val Thr Lys Glu Tyr Glu Gly His Lys Ala Val Val 3845 3850 3855		
Gln Asp Leu Ser Leu Thr Phe Tyr Arg Asp Gln Ile Thr Ala Leu 3860 3865 3870		
Leu Gly Thr Asn Gly Ala Gly Lys Thr Thr Ile Ile Ser Met Leu 3875 3880 3885		
Thr Gly Leu His Pro Pro Thr Ser Gly Thr Ile Ile Ile Asn Gly 3890 3895 3900		
Lys Asn Leu Gln Thr Asp Leu Ser Arg Val Arg Met Glu Leu Gly 3905 3910 3915		
Val Cys Pro Gln Gln Asp Ile Leu Leu Asp Asn Leu Thr Val Arg 3920 3925 3930		
Glu His Leu Leu Leu Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr 3935 3940 3945		

Lys Lys Glu Leu His Gln Gln Val Asn Gln Thr Leu Gln Asp Val
 3950 3955 3960
 Asp Leu Thr Gln His Gln His Lys Gln Thr Arg Ala Leu Ser Gly
 3965 3970 3975
 Gly Leu Lys Arg Lys Leu Ser Leu Gly Ile Ala Phe Met Gly Met
 3980 3985 3990
 Ser Arg Thr Val Val Leu Asp Glu Pro Thr Ser Gly Val Asp Pro
 3995 4000 4005
 Cys Ser Arg His Ser Leu Trp Asp Ile Leu Leu Lys Tyr Arg Glu
 4010 4015 4020
 Gly Arg Thr Ile Ile Phe Thr Thr His His Leu Asp Glu Ala Glu
 4025 4030 4035
 Ala Leu Ser Asp Arg Val Ala Val Leu Gln His Gly Arg Leu Arg
 4040 4045 4050
 Cys Cys Gly Pro Pro Phe Cys Leu Lys Glu Ala Tyr Gly Gln Gly
 4055 4060 4065
 Leu Arg Leu Thr Leu Thr Arg Gln Pro Ser Val Leu Glu Ala His
 4070 4075 4080
 Asp Leu Lys Asp Met Ala Cys Val Thr Ser Leu Ile Lys Ile Tyr
 4085 4090 4095
 Ile Pro Gln Ala Phe Leu Lys Asp Ser Ser Gly Ser Glu Leu Thr
 4100 4105 4110
 Tyr Thr Ile Pro Lys Asp Thr Asp Lys Ala Cys Leu Lys Gly Leu
 4115 4120 4125
 Phe Gln Ala Leu Asp Glu Asn Leu His Gln Leu His Leu Thr Gly
 4130 4135 4140
 Tyr Gly Ile Ser Asp Thr Thr Leu Glu Glu Val Phe Leu Met Leu
 4145 4150 4155
 Leu Gln Asp Ser Asn Lys Lys Ser His Ile Ala Leu Gly Thr Glu
 4160 4165 4170

Ser	Glu	Leu	Gln	Asn	His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr
4175						4180					4185			
Cys	Gly	Ser	Leu	Ala	Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu
4190						4195					4200			
Leu	Arg	Ala	Gln	Val	Ala	Ala	Ile	Leu	Ala	Arg	Arg	Leu	Arg	Arg
4205						4210					4215			
Thr	Leu	Arg	Ala	Gly	Lys	Ser	Thr	Leu	Ala	Asp	Leu	Leu	Leu	Pro
4220						4225					4230			
Val	Leu	Phe	Val	Ala	Leu	Ala	Met	Gly	Leu	Phe	Met	Val	Arg	Pro
4235						4240					4245			
Leu	Ala	Thr	Glu	Tyr	Pro	Pro	Leu	Arg	Leu	Thr	Pro	Gly	His	Tyr
4250						4255					4260			
Gln	Arg	Ala	Glu	Thr	Tyr	Phe	Phe	Ser	Ser	Gly	Gly	Asp	Asn	Leu
4265						4270					4275			
Asp	Leu	Thr	Arg	Val	Leu	Leu	Arg	Lys	Phe	Arg	Asp	Gln	Asp	Leu
4280						4285					4290			
Pro	Cys	Ala	Asp	Leu	Asn	Pro	Arg	Gln	Lys	Asn	Ser	Ser	Cys	Trp
4295						4300					4305			
Arg	Thr	Asp	Pro	Phe	Ser	His	Pro	Glu	Phe	Gln	Asp	Ser	Cys	Gly
4310						4315					4320			
Cys	Leu	Lys	Cys	Pro	Asn	Arg	Ser	Ala	Ser	Ala	Pro	Tyr	Leu	Thr
4325						4330					4335			
Asn	His	Leu	Gly	His	Thr	Leu	Leu	Asn	Leu	Ser	Gly	Phe	Asn	Met
4340						4345					4350			
Glu	Glu	Tyr	Leu	Leu	Ala	Pro	Ser	Glu	Lys	Pro	Arg	Leu	Gly	Gly
4355						4360					4365			
Trp	Ser	Phe	Gly	Leu	Lys	Ile	Pro	Ser	Glu	Ala	Gly	Gly	Ala	Asn
4370						4375					4380			
Gly	Asn	Ile	Ser	Lys	Pro	Pro	Thr	Leu	Ala	Lys	Val	Trp	Tyr	Asn
4385						4390					4395			

Gln	Lys	Gly	Phe	His	Ser	Leu	Pro	Ser	Tyr	Leu	Asn	His	Leu	Asn
4400						4405					4410			
Asn	Leu	Ile	Leu	Trp	Gln	His	Leu	Pro	Pro	Thr	Val	Asp	Trp	Arg
4415						4420					4425			
Gln	Tyr	Gly	Ile	Thr	Leu	Tyr	Ser	His	Pro	Tyr	Gly	Gly	Ala	Leu
4430						4435					4440			
Leu	Asn	Glu	Asp	Lys	Ile	Leu	Glu	Ser	Ile	Arg	Gln	Cys	Gly	Val
4445						4450					4455			
Ala	Leu	Cys	Ile	Val	Leu	Gly	Phe	Ser	Ile	Leu	Ser	Ala	Ser	Ile
4460						4465					4470			
Gly	Ser	Ser	Val	Val	Arg	Asp	Arg	Val	Ile	Gly	Ala	Lys	Arg	Leu
4475						4480					4485			
Gln	His	Ile	Ser	Gly	Leu	Gly	Tyr	Arg	Met	Tyr	Trp	Phe	Thr	Asn
4490						4495					4500			
Phe	Leu	Tyr	Asp	Met	Leu	Phe	Tyr	Leu	Val	Ser	Val	Cys	Leu	Cys
4505						4510					4515			
Val	Ala	Val	Ile	Val	Ala	Phe	Gln	Leu	Thr	Ala	Phe	Thr	Phe	Arg
4520						4525					4530			
Lys	Asn	Leu	Ala	Ala	Thr	Ala	Leu	Leu	Leu	Ser	Leu	Phe	Gly	Tyr
4535						4540					4545			
Ala	Thr	Leu	Pro	Trp	Met	Tyr	Leu	Met	Ser	Arg	Ile	Phe	Ser	Ser
4550						4555					4560			
Ser	Asp	Val	Ala	Phe	Ile	Ser	Tyr	Val	Ser	Leu	Asn	Phe	Ile	Phe
4565						4570					4575			
Gly	Leu	Cys	Thr	Met	Pro	Ile	Thr	Ile	Met	Pro	Arg	Leu	Leu	Ala
4580						4585					4590			
Ile	Ile	Ser	Lys	Ala	Lys	Asn	Leu	Gln	Asn	Ile	Tyr	Asp	Val	Leu
4595						4600					4605			
Lys	Trp	Val	Phe	Thr	Ile	Phe	Pro	Gln	Phe	Cys	Leu	Gly	Gln	Gly
4610						4615					4620			
Leu	Val	Glu	Leu	Cys	Tyr	Asn	Gln	Ile	Lys	Tyr	Asp	Leu	Thr	His

4625		4630		4635
Asn Phe Gly Ile Asp Ser Tyr Val Ser Pro Phe Glu Met Asn Phe				
4640		4645		4650
Leu Gly Trp Ile Phe Val Gln Leu Ala Ser Gln Gly Thr Val Leu				
4655		4660		4665
Leu Leu Leu Arg Val Leu Leu His Trp Asp Leu Leu Arg Trp Pro				
4670		4675		4680
Arg Gly His Ser Thr Leu Gln Gly Thr Val Lys Ser Ser Lys Asp				
4685		4690		4695
Thr Asp Val Glu Lys Glu Glu Lys Arg Val Phe Glu Gly Arg Thr				
4700		4705		4710
Asn Gly Asp Ile Leu Val Leu Tyr Asn Leu Ser Lys His Tyr Arg				
4715		4720		4725
Arg Phe Phe Gln Asn Ile Ile Ala Val Gln Asp Ile Ser Leu Gly				
4730		4735		4740
Ile Pro Lys Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala				
4745		4750		4755
Gly Lys Ser Thr Thr Phe Lys Met Leu Asn Gly Glu Val Ser Leu				
4760		4765		4770
Thr Ser Gly His Ala Ile Ile Arg Thr Pro Met Gly Asp Ala Val				
4775		4780		4785
Asp Leu Ser Ser Ala Gly Thr Ala Gly Val Leu Ile Gly Tyr Cys				
4790		4795		4800
Pro Gln Gln Asp Ala Leu Asp Glu Leu Leu Thr Gly Trp Glu His				
4805		4810		4815
Leu Tyr Tyr Tyr Cys Ser Leu Arg Gly Ile Pro Arg Gln Cys Ile				
4820		4825		4830
Pro Glu Val Ala Gly Asp Leu Ile Arg Arg Leu His Leu Glu Ala				
4835		4840		4845
His Ala Asp Lys Pro Val Ala Thr Tyr Ser Gly Gly Thr Lys Arg				
4850		4855		4860

Lys Leu Ser Thr Ala Leu Ala Leu Val Gly Lys Pro Asp Ile Leu
 4865 4870 4875

Leu Leu Asp Glu Pro Ser Ser Gly Met Asp Pro Cys Ser Lys Arg
 4880 4885 4890

Tyr Leu Trp Gln Thr Ile Met Lys Glu Val Arg Glu Gly Cys Ala
 4895 4900 4905

Ala Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys
 4910 4915 4920

Thr Arg Leu Ala Ile Met Val Asn Gly Ser Phe Lys Cys Leu Gly
 4925 4930 4935

Ser Pro Gln His Ile Lys Asn Arg Lys Leu Asn Lys Thr Asn
 4940 4945 4950

<210> 28
 <211> 11068
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> misc_feature
 <222> (873)..(873)
 <223> Xaa = any amino acid

<220>
 <221> CDS
 <222> (2209)..(10173)

<220>
 <221> misc_feature
 <222> (3334)..(3334)
 <223> n = any nucleic acid residue

<220>
 <221> misc_feature
 <222> (10268)..(10268)
 <223> Xaa = any amino acid

<400> 28
 ggactgagag cagggagcag caggcatggg gcatgccggg tgccagttca aagccctgct 60
 gtggaagaat tggctctgca gactcaggaa cccggtcctt ttcttctgctg aattcttctg 120
 gccttgatc ctgtttgtaa ttctgacagt tcttcgtttt caagaacctc ccagatacag 180
 agacatttgt tatttgcagc cccgagatct acccagctgt ggtgttatcc cctttgttca 240

aagccttctt	tgtaacactg	gatcaaggtg	taggaacttc	agctatgaag	ggtcaatgga	300
gcatcatttt	cgtttgtcta	ggttccaaac	tgagctgac	cccaagaaag	tcaacaacct	360
ggccttttta	aaagagatac	aagacctggc	agaggaaatt	catggaatga	tggaacaaggc	420
aaaaaactta	aaaagacttt	gggtagaacg	atccaacact	ccagattctt	cttatggttc	480
cagttttttt	acaatggatc	tcaataagac	cgaggaggta	atattgaaac	tggaagacct	540
ccatcagcag	cctcatactt	gggattttct	acttttactg	ccgagactac	acacaagcca	600
tgatcatgtg	gaagatggca	tggtatgtgc	agtgaacctt	ctccagacca	ttttgaattc	660
cttaatatcc	ctagaagatt	tagattggct	tccactcaac	caaacttttt	cccaggtttc	720
tgaacttgta	ctgaatgtga	ccatttcgac	actgacattt	ctgcagcaac	atggagtagc	780
agtcaccgag	ccagtttacc	acctgtccat	gcagaatata	gtgtgggac	cacagaaagt	840
ccagtatgat	ctcaaattcc	agtttggtct	tgntgatctt	cacacggaac	agatcctgaa	900
ctcttcagct	gaactgaagg	agattcccac	agacaacttc	ttggagaaga	tggtgtgttc	960
agtcttgtct	agcacatcag	aggatgaagc	tgagaaatgg	ggccacgttg	gaggctgcca	1020
ccctaagtgg	tcagaagcca	aaaactatct	tgtccatgca	gtcagctggc	tgcgagtcta	1080
ccaacagggtg	tttgttcagt	ggcaacaggg	tagcctgctt	cagaagacac	tcacaggcat	1140
gggccatagt	ctggaggctc	tcaggaatca	gtttgaagaa	gagagcaagc	cctggaaggt	1200
ggtggaagct	ctgcacactg	cactgctcct	gctgaatgac	agcttgtcag	cagatggccc	1260
aaaagataat	catacatctt	caaagatatt	acagcatctg	tggaatttgc	aaagcttgct	1320
gcaaaacctg	ccccagtggc	cggcactgaa	gagatttctt	cagcttgatg	gagctctcag	1380
aaatgcgata	gctcagaatt	tacattttgt	ccaagaagtc	ctcatttgcc	tggaagacac	1440
agctaattgat	tttaaattgt	ttgaacttaa	ccaattgaaa	ctggaaaagg	atgtgttctt	1500
ttgggagctg	aaacagatgt	tggcgaagaa	tgctgtctgc	ccgaatggtc	gtttctctga	1560
gaaggaggctc	tttttgccgc	ctggaaactc	cagcatatgg	ggtggtctcc	agggactggt	1620
gtgctattgt	aactcctctg	agacgagtgt	tttaaacaag	ctacttggtt	cagtagagga	1680
tgctgatcgt	attttgcaag	aggtcattac	ttggcacaaa	aatatgtcag	ttttaatacc	1740
tgaagaatat	ttggactggc	aggaacttga	gatgcagctg	tcagaagcaa	gcctttcctg	1800
tactcggctc	ttcctgctgc	tgggagctga	tccctctcct	gagaatgatg	tcttttctag	1860
tgactgtaag	caccagcttg	tctccacagt	gatatttcat	acacttgaaa	aaacacaatt	1920
tttcctggaa	caagcatatt	attggaaagc	cttcaaaaag	tttatcagga	agacttgcca	1980
agtgggcccaa	tatgtaaata	tgcaagagag	ttccagaac	agactattgg	cttttcctga	2040

ggaatctcct tgttttgaag aaaacatgga ttggaaaatg atcagtgata attatttttca	2100
attttttgaat aacttactca agtctccaac agcttccata tccaggaaga ttctgagaac	2160
aaaatatctc ttctgtctgaa atattttccac aaagatgtta ttgcagag atg agt ttt	2217
Met Ser Phe	
1	
gtc cca aaa gat aaa att cta gaa att ctg aaa ctg gat caa ttt ctt	2265
Val Pro Lys Asp Lys Ile Leu Glu Ile Leu Lys Leu Asp Gln Phe Leu	
5 10 15	
acc ctg atg ata caa gac aga ttg atg aac att ttt tca agt tta aag	2313
Thr Leu Met Ile Gln Asp Arg Leu Met Asn Ile Phe Ser Ser Leu Lys	
20 25 30 35	
gag act ata tat cac cta atg aaa agt tca ttt ata tta gac aat gga	2361
Glu Thr Ile Tyr His Leu Met Lys Ser Ser Phe Ile Leu Asp Asn Gly	
40 45 50	
gaa ttt tat ttt gat act cat caa gga ctg aag ttc atg caa gat tta	2409
Glu Phe Tyr Phe Asp Thr His Gln Gly Leu Lys Phe Met Gln Asp Leu	
55 60 65	
ttt aat gcc ctt ctc agg gaa act tca atg aaa aat aag act gaa aat	2457
Phe Asn Ala Leu Leu Arg Glu Thr Ser Met Lys Asn Lys Thr Glu Asn	
70 75 80	
aat ata gac ttt ttc aca gtg gtg agt cag ttg ttt ttc cat gtg aat	2505
Asn Ile Asp Phe Phe Thr Val Val Ser Gln Leu Phe Phe His Val Asn	
85 90 95	
aag tct gag gac ctc ttc aaa ctc aat caa gat ctt ggg tca gct ctt	2553
Lys Ser Glu Asp Leu Phe Lys Leu Asn Gln Asp Leu Gly Ser Ala Leu	
100 105 110 115	
cac ctt gta aga gaa tgt tca aca gag atg gca aga ctt ctg gat aca	2601
His Leu Val Arg Glu Cys Ser Thr Glu Met Ala Arg Leu Leu Asp Thr	
120 125 130	
att tta cac tct cct aat aag gac ttc tat gct ttg tat cct acc ctc	2649
Ile Leu His Ser Pro Asn Lys Asp Phe Tyr Ala Leu Tyr Pro Thr Leu	
135 140 145	
caa gaa gtt ata ctt gct aat cta acg gat ttg ctt ttc ttt ata aat	2697
Gln Glu Val Ile Leu Ala Asn Leu Thr Asp Leu Leu Phe Phe Ile Asn	
150 155 160	
aat tca ttc cct cta aga aac aga gca aca tta gaa att act aag aga	2745
Asn Ser Phe Pro Leu Arg Asn Arg Ala Thr Leu Glu Ile Thr Lys Arg	
165 170 175	
tta gtt ggt gct att tca aga gca agt gaa gaa agt cac gtc ctg aaa	2793
Leu Val Gly Ala Ile Ser Arg Ala Ser Glu Glu Ser His Val Leu Lys	
180 185 190 195	
ccc ctc tta gaa atg tct ggg act ctg gtc atg ctg ttg aat gac agt	2841
Pro Leu Leu Glu Met Ser Gly Thr Leu Val Met Leu Leu Asn Asp Ser	
200 205 210	
gct gac ctg aga gat ctt gcc aca tca atg gac tcc att gtg aaa ctt	2889

Ala Asp Leu Arg Asp Leu Ala Thr Ser Met Asp Ser Ile Val Lys Leu	
215 220 225	
ctt aag ctg gtc aag aaa gtt tgc ggg aag atg tcc aca gtt ttt aaa	2937
Leu Lys Leu Val Lys Lys Val Ser Gly Lys Met Ser Thr Val Phe Lys	
230 235 240	
act cat ttt atc tcc aat acc aag gac agt gtg aaa ttc ttt gac act	2985
Thr His Phe Ile Ser Asn Thr Lys Asp Ser Val Lys Phe Phe Asp Thr	
245 250 255	
ctg tat tcc atc atg caa caa agt gtt caa aat ctt gtg aaa gaa ata	3033
Leu Tyr Ser Ile Met Gln Gln Ser Val Gln Asn Leu Val Lys Glu Ile	
260 265 270 275	
gct act tta aaa aaa ata gat cat ttc aca ttt gaa aag ata aat gat	3081
Ala Thr Leu Lys Lys Ile Asp His Phe Thr Phe Glu Lys Ile Asn Asp	
280 285 290	
ttg ttg gtg cca ttt ctt gac ttg gcc ttt gaa atg att ggg gta gaa	3129
Leu Leu Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile Gly Val Glu	
295 300 305	
cct tat ata tca tca aac tct gat att ttc agt atg tca cct agc ata	3177
Pro Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser Pro Ser Ile	
310 315 320	
ctc tca tat atg aac caa tct aag gac ttt tct gat att ttg gaa gaa	3225
Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile Leu Glu Glu	
325 330 335	
att gct gaa ttt tta aca tct gtg aaa atg aac ttg gaa gat atg agg	3273
Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu Asp Met Arg	
340 345 350 355	
agt ctt gcg gta gca ttt aac aat gag act caa aca ttt tct atg gat	3321
Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe Ser Met Asp	
360 365 370	
tct gtc aac tta ngg gaa gaa att ctg ggt tgc tta gtt cct ata aat	3369
Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu Val Pro Ile Asn	
375 380 385	
aac atc acc aac caa atg gac ttc tta tac cct aat cca att tcc act	3417
Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro Ile Ser Thr	
390 395 400	
cat agt ggc cct caa gat ata aaa tgg gaa ata att cat gaa gtg atc	3465
His Ser Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His Glu Val Ile	
405 410 415	
ctt ttt ttg gat aaa ata tta tca caa aac agc aca gaa ata gga tct	3513
Leu Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu Ile Gly Ser	
420 425 430 435	
ttc ttg aaa atg gtg atc tgt ctc acc tta gaa gct ctt tgg aaa aac	3561
Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu Trp Lys Asn	
440 445 450	
tta aag aaa gat aat tgg aat gtt tct aat gtg ttg atg acg ttt act	3609
Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met Thr Phe Thr	

455										460										465									
cag	cat	cca	aat	aac	ctt	ttg	aaa	acc	ata	gaa	aca	gtt	tta	gag	gcc														
Gln	His	Pro	Asn	Asn	Leu	Leu	Lys	Thr	Ile	Glu	Thr	Val	Leu	Glu	Ala														3657
		470					475					480																	
tcc	agt	gga	att	aaa	agt	gac	tat	gaa	ggt	gat	ttg	aat	aaa	agt	tta														3705
Ser	Ser	Gly	Ile	Lys	Ser	Asp	Tyr	Glu	Gly	Asp	Leu	Asn	Lys	Ser	Leu														
		485				490					495																		
tat	ttt	gac	aca	cct	ttg	agt	cag	aat	ata	act	cat	cat	caa	ctt	gaa														3753
Tyr	Phe	Asp	Thr	Pro	Leu	Ser	Gln	Asn	Ile	Thr	His	His	Gln	Leu	Glu														
500					505					510					515														
aaa	gca	atc	cat	aat	gtt	tta	agt	aga	ata	gct	ctc	tgg	agg	aaa	gga														3801
Lys	Ala	Ile	His	Asn	Val	Leu	Ser	Arg	Ile	Ala	Leu	Trp	Arg	Lys	Gly														
				520					525					530															
ctt	cgt	ttt	aac	aac	tct	gaa	tgg	ata	act	tcc	aca	aga	act	ttg	ttt														3849
Leu	Arg	Phe	Asn	Asn	Ser	Glu	Trp	Ile	Thr	Ser	Thr	Arg	Thr	Leu	Phe														
			535					540					545																
cag	cca	ctt	ttt	gag	att	ttc	att	aaa	gca	acc	acc	gga	aag	aat	gtc														3897
Gln	Pro	Leu	Phe	Glu	Ile	Phe	Ile	Lys	Ala	Thr	Thr	Gly	Lys	Asn	Val														
		550					555					560																	
aca	tca	gaa	aaa	gaa	gag	aga	acc	gag	aaa	gag	atg	att	gac	ttt	cct														3945
Thr	Ser	Glu	Lys	Glu	Glu	Arg	Thr	Glu	Lys	Glu	Met	Ile	Asp	Phe	Pro														
		565				570					575																		
tat	agt	ttc	aaa	cca	ttt	ttc	tgt	ttg	gag	aaa	tac	ctg	gga	gga	tta														3993
Tyr	Ser	Phe	Lys	Pro	Phe	Phe	Cys	Leu	Glu	Lys	Tyr	Leu	Gly	Gly	Leu														
580					585				590					595															
ttt	gta	ttg	act	aaa	tac	tgg	caa	caa	atc	cca	cta	aca	gat	caa	agt														4041
Phe	Val	Leu	Thr	Lys	Tyr	Trp	Gln	Gln	Ile	Pro	Leu	Thr	Asp	Gln	Ser														
				600					605					610															
gtt	gtt	gag	att	tgt	gaa	gtt	ttc	cag	cag	act	gtg	aag	ccc	tca	gaa														4089
Val	Val	Glu	Ile	Cys	Glu	Val	Phe	Gln	Gln	Thr	Val	Lys	Pro	Ser	Glu														
			615					620					625																
gcc	atg	gag	atg	ctg	cag	aaa	gtg	aag	atg	atg	gtc	gta	cgt	gtg	ctc														4137
Ala	Met	Glu	Met	Leu	Gln	Lys	Val	Lys	Met	Met	Val	Val	Arg	Val	Leu														
		630					635					640																	
acc	atc	gtt	gca	gaa	aac	cct	tcc	tgg	acc	aag	gac	att	ttg	tgt	gct														4185
Thr	Ile	Val	Ala	Glu	Asn	Pro	Ser	Trp	Thr	Lys	Asp	Ile	Leu	Cys	Ala														
		645				650					655																		
act	ctg	agt	tgc	aag	caa	aat	ggg	ata	agg	cat	ctc	att	tta	tct	gct														4233
Thr	Leu	Ser	Cys	Lys	Gln	Asn	Gly	Ile	Arg	His	Leu	Ile	Leu	Ser	Ala														
660					665				670				675																
ata	caa	ggg	gtc	act	ttg	gcg	cag	gac	cac	ttc	cag	gaa	att	gaa	aag														4281
Ile	Gln	Gly	Val	Thr	Leu	Ala	Gln	Asp	His	Phe	Gln	Glu	Ile	Glu	Lys														
				680					685				690																
ata	tgg	tcc	tcg	ccg	aat	cag	cta	aat	tgt	gaa	agt	ctt	agc	aag	aat														4329
Ile	Trp	Ser	Ser	Pro	Asn	Gln	Leu	Asn	Cys	Glu	Ser	Leu	Ser	Lys	Asn														
			695				700						705																

ctt tct agc acc ttg gag agc ttc aag agc agc ttg gaa aat gcc act	4377
Leu Ser Ser Thr Leu Glu Ser Phe Lys Ser Ser Leu Glu Asn Ala Thr	
710 715 720	
ggc cag gac tgc aca agc cag ccg agg ctg gag acg gtg cag cag cac	4425
Gly Gln Asp Cys Thr Ser Gln Pro Arg Leu Glu Thr Val Gln Gln His	
725 730 735	
ttg tac atg ttg gcc aaa agc ctc gag gaa act tgg tca tca ggg aat	4473
Leu Tyr Met Leu Ala Lys Ser Leu Glu Glu Thr Trp Ser Ser Gly Asn	
740 745 750 755	
ccc atc atg act ttt ctc agc aat ttc aca gta act gag gat gta aaa	4521
Pro Ile Met Thr Phe Leu Ser Asn Phe Thr Val Thr Glu Asp Val Lys	
760 765 770	
ata aaa gat ttg atg aag aat atc acc aag ttg act gag gag ctt cgc	4569
Ile Lys Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu Glu Leu Arg	
775 780 785	
tct tcc atc caa atc tcg aat gag act atc cat agc att cta gaa gca	4617
Ser Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile Leu Glu Ala	
790 795 800	
aat att tcc cac tcc aag gtt ctc ttc agt gcc ctc acc gta gct ctg	4665
Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr Val Ala Leu	
805 810 815	
tct gga aag tgt gat cag gaa atc ctt cat ctc ctg ctg aca ttt ccc	4713
Ser Gly Lys Cys Asp Gln Glu Ile Leu His Leu Leu Leu Thr Phe Pro	
820 825 830 835	
aaa ggg gaa aaa tct tgg atc gca gcg gag gaa ctc tgt agc ctg cca	4761
Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys Ser Leu Pro	
840 845 850	
ggg tca aaa gtg tat tct ctg att gtg ttg ctg agt cga aac ttg gat	4809
Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg Asn Leu Asp	
855 860 865	
gtg cga gct ttc att tac aag act ctg atg cct tct gaa gca aat ggc	4857
Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu Ala Asn Gly	
870 875 880	
ttg ctc aac tcc ttg ctg gat ata gtt tcc agc ctc agc gcc ttg ctt	4905
Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser Ala Leu Leu	
885 890 895	
gcc aaa gcc cag cac gtc ttt gag tat ctt cct gag ttt ctt cac aca	4953
Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe Leu His Thr	
900 905 910 915	
ttt aaa atc act gcc ttg cta gaa acc ctg gac ttt caa cag gtt tca	5001
Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln Gln Val Ser	
920 925 930	
caa aat gtc cag gcc aga agt tca gct ttt ggt tct ttc cag ttt gtg	5049
Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe Gln Phe Val	
935 940 945	

atg aag atg gtt tgc aag gac caa gca tca ttc ctt agc gat tct aat	5097
Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser Asp Ser Asn	
950 955 960	
atg ttt att aat ttg ccc aga gtt aag gaa ctc ttg gaa gat gac aaa	5145
Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu Glu Asp Asp Lys	
965 970 975	
gaa aaa ttc aac att cct gaa gat tca aca ccg ttt tgc ttg aag ctt	5193
Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys Leu Lys Leu	
980 985 990 995	
tat cag gaa att cta caa ttg cca aat ggt gct ttg gtg tgg acc	5238
Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val Trp Thr	
1000 1005 1010	
ttc cta aaa ccc ata ttg cat gga aaa ata cta tac aca cca aac	5283
Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr Pro Asn	
1015 1020 1025	
act cca gaa att aac aag gtc att caa aag gct aat tac acc ttt	5328
Thr Pro Glu Ile Asn Lys Val Ile Gln Lys Ala Asn Tyr Thr Phe	
1030 1035 1040	
tat att gtg gac aaa cta aaa act tta tca gaa aca ctg ctg gaa	5373
Tyr Ile Val Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu Leu Glu	
1045 1050 1055	
atg tcc agc ctt ttc cag aga agt gga agt ggc cag atg ttc aac	5418
Met Ser Ser Leu Phe Gln Arg Ser Gly Ser Gly Gln Met Phe Asn	
1060 1065 1070	
cag ctg cag gag gcc ctg aga aac aaa ttt gta aga aac ttt gta	5463
Gln Leu Gln Glu Ala Leu Arg Asn Lys Phe Val Arg Asn Phe Val	
1075 1080 1085	
gaa aac cag ttg cac att gat gta gac aaa ctt act gaa aaa ctc	5508
Glu Asn Gln Leu His Ile Asp Val Asp Lys Leu Thr Glu Lys Leu	
1090 1095 1100	
cag aca tac gga ggg ctg ctg gat gag atg ttt aac cat gca ggc	5553
Gln Thr Tyr Gly Gly Leu Leu Asp Glu Met Phe Asn His Ala Gly	
1105 1110 1115	
gct gga cgc ttc cgt ttc ttg ggc agc atc ttg gtc aat ctc tct	5598
Ala Gly Arg Phe Arg Phe Leu Gly Ser Ile Leu Val Asn Leu Ser	
1120 1125 1130	
tcc tgc gtg gca ctg aac cgt ttc cag gct ctg cag tct gtc gac	5643
Ser Cys Val Ala Leu Asn Arg Phe Gln Ala Leu Gln Ser Val Asp	
1135 1140 1145	
atc ctg gag act aaa gca cat gaa ctc ttg cag cag aac agc ttc	5688
Ile Leu Glu Thr Lys Ala His Glu Leu Leu Gln Gln Asn Ser Phe	
1150 1155 1160	
ttg gcc agt atc att ttc agc aat tcc tta ttc gac aag aac ttc	5733
Leu Ala Ser Ile Ile Phe Ser Asn Ser Leu Phe Asp Lys Asn Phe	
1165 1170 1175	
aga tca gag tct gtc aaa ctg cca ccc cat gtc tca tac aca atc	5778

Arg Ser Glu Ser Val	Lys Leu Pro Pro His	Val Ser Tyr Thr Ile	
1180	1185	1190	
cgg acc aat gtg tta	tac agc gtg cga aca	gat gtg gta aaa aac	5823
Arg Thr Asn Val Leu	Tyr Ser Val Arg Thr	Asp Val Val Lys Asn	
1195	1200	1205	
cct tct tgg aag ttc	cac cct cag aat cta	cca gct gat ggg ttc	5868
Pro Ser Trp Lys Phe	His Pro Gln Asn Leu	Pro Ala Asp Gly Phe	
1210	1215	1220	
aaa tat aac tac gtc	ttt gcc cca ctg caa	gac atg atc gaa aga	5913
Lys Tyr Asn Tyr Val	Phe Ala Pro Leu Gln	Asp Met Ile Glu Arg	
1225	1230	1235	
gcc atc att ttg gtg	cag act ggg cag gaa	gcc ctg gaa cca gca	5958
Ala Ile Ile Leu Val	Gln Thr Gly Gln Glu	Ala Leu Glu Pro Ala	
1240	1245	1250	
gca cag act cag gcg	gcc cct tac ccc tgc	cat acc agc gac cta	6003
Ala Gln Thr Gln Ala	Ala Pro Tyr Pro Cys	His Thr Ser Asp Leu	
1255	1260	1265	
ttc ctg aac aac gtt	ggg ttc ttt ttt cca	ctg ata atg atg ctg	6048
Phe Leu Asn Asn Val	Gly Phe Phe Phe Pro	Leu Ile Met Met Leu	
1270	1275	1280	
acg tgg atg gtg tct	gtg gcc agc atg gtc	aga aag ttg gtg tat	6093
Thr Trp Met Val Ser	Val Ala Ser Met Val	Arg Lys Leu Val Tyr	
1285	1290	1295	
gag cag gag ata cag	ata gaa gag tat atg	cgg atg atg gga gtg	6138
Glu Gln Glu Ile Gln	Ile Glu Glu Tyr Met	Arg Met Met Gly Val	
1300	1305	1310	
cat cca gtg atc cat	ttc ctg gcc tgg ttc	ctg gag aac atg gct	6183
His Pro Val Ile His	Phe Leu Ala Trp Phe	Leu Glu Asn Met Ala	
1315	1320	1325	
gtg ttg acc ata agc	agt gct act ctg gcc	atc gtt ctg aaa aca	6228
Val Leu Thr Ile Ser	Ser Ala Thr Leu Ala	Ile Val Leu Lys Thr	
1330	1335	1340	
agt ggc atc ttt gca	cac agc aat acc ttt	att gtt ttc ctc ttt	6273
Ser Gly Ile Phe Ala	His Ser Asn Thr Phe	Ile Val Phe Leu Phe	
1345	1350	1355	
ctc ttg gat ttt ggg	atg tca gtc gtc atg	ctg agc tac ctc ttg	6318
Leu Leu Asp Phe Gly	Met Ser Val Val Met	Leu Ser Tyr Leu Leu	
1360	1365	1370	
agt gca ttt ttc agc	caa gct aat aca gcg	gcc ctt tgt acc agc	6363
Ser Ala Phe Phe Ser	Gln Ala Asn Thr Ala	Ala Leu Cys Thr Ser	
1375	1380	1385	
ctg gtg tac atg atc	agc ttt ctg ccc tac	ata gtt cta ttg gtt	6408
Leu Val Tyr Met Ile	Ser Phe Leu Pro Tyr	Ile Val Leu Leu Val	
1390	1395	1400	
cta cat aac caa tta	agt ttt gtt aat cag	aca ttt ctg tgc ctt	6453
Leu His Asn Gln Leu	Ser Phe Val Asn Gln	Thr Phe Leu Cys Leu	

1405	1410	1415	
ctt tcg aca acc gcc	ttt gga caa ggg gta	ttt ttt att aca ttc	6498
Leu Ser Thr Thr Ala	Phe Gly Gln Gly Val	Phe Phe Ile Thr Phe	
1420	1425	1430	
ctg gaa gga caa gag	aca ggg att caa tgg	aat aat atg tac cag	6543
Leu Glu Gly Gln Glu	Thr Gly Ile Gln Trp	Asn Asn Met Tyr Gln	
1435	1440	1445	
gct ctg gaa caa ggg	ggc atg aca ttt ggc	tgg gtt tgc tgg atg	6588
Ala Leu Glu Gln Gly	Gly Met Thr Phe Gly	Trp Val Cys Trp Met	
1450	1455	1460	
att ctt ttt gat tca	agc ctt tat ttt ttg	tgt gga tgg tac ttg	6633
Ile Leu Phe Asp Ser	Ser Leu Tyr Phe Leu	Cys Gly Trp Tyr Leu	
1465	1470	1475	
agc aac ttg att cct	gga aca ttt ggt tta	cgg aaa cca tgg tat	6678
Ser Asn Leu Ile Pro	Gly Thr Phe Gly Leu	Arg Lys Pro Trp Tyr	
1480	1485	1490	
ttc ccc ttt act gcc	tca tat tgg aag agt	gtg ggt ttc ttg gtg	6723
Phe Pro Phe Thr Ala	Ser Tyr Trp Lys Ser	Val Gly Phe Leu Val	
1495	1500	1505	
gag aaa agg caa tac	ttt cta agt tct agt	ctg ttc ttc ttc aat	6768
Glu Lys Arg Gln Tyr	Phe Leu Ser Ser Ser	Leu Phe Phe Phe Asn	
1510	1515	1520	
gag aac ttt gac aat	aaa ggg tca tca ctg	caa aac agg gaa gga	6813
Glu Asn Phe Asp Asn	Lys Gly Ser Ser Leu	Gln Asn Arg Glu Gly	
1525	1530	1535	
gag ctt gaa gga agt	gcc ccg gga gtc acc	ctg gtg tct gtg acc	6858
Glu Leu Glu Gly Ser	Ala Pro Gly Val Thr	Leu Val Ser Val Thr	
1540	1545	1550	
aag gaa tat gag ggc	cac aag gct gtg gtc	caa gac ctc agc ctg	6903
Lys Glu Tyr Glu Gly	His Lys Ala Val Val	Gln Asp Leu Ser Leu	
1555	1560	1565	
acc ttc tac aga gac	caa atc acc gcc ctg	ctg ggg aca aac ggt	6948
Thr Phe Tyr Arg Asp	Gln Ile Thr Ala Leu	Leu Gly Thr Asn Gly	
1570	1575	1580	
gcc ggg aaa acc act	atc ata tcc atg ttg	acg ggg ctc cac cct	6993
Ala Gly Lys Thr Thr	Ile Ile Ser Met Leu	Thr Gly Leu His Pro	
1585	1590	1595	
ccc act tct gga acc	atc atc atc aat ggc	aag aac cta cag aca	7038
Pro Thr Ser Gly Thr	Ile Ile Ile Asn Gly	Lys Asn Leu Gln Thr	
1600	1605	1610	
gac ctg tcg agg gtc	aga atg gag ctt ggt	gtg tgt ccg cag cag	7083
Asp Leu Ser Arg Val	Arg Met Glu Leu Gly	Val Cys Pro Gln Gln	
1615	1620	1625	
gac atc ctg ttg gac	aac ctc acc gtc cgg	gaa cat ttg ctg ctc	7128
Asp Ile Leu Leu Asp	Asn Leu Thr Val Arg	Glu His Leu Leu Leu	
1630	1635	1640	

ttt gct tcc ata aag gcg cct cag tgg acc aag aag gag ctg cat	7173
Phe Ala Ser Ile Lys Ala Pro Gln Trp Thr Lys Lys Glu Leu His	
1645 1650 1655	
cag caa gtc aat caa act ctt cag gat gtg gac tta act cag cat	7218
Gln Gln Val Asn Gln Thr Leu Gln Asp Val Asp Leu Thr Gln His	
1660 1665 1670	
cag cac aaa cag acc cga gct ctg tct gga ggc ctg aag agg aag	7263
Gln His Lys Gln Thr Arg Ala Leu Ser Gly Gly Leu Lys Arg Lys	
1675 1680 1685	
ctc tcc ctt ggc att gct ttc atg ggc atg tcg agg acc gtg gtt	7308
Leu Ser Leu Gly Ile Ala Phe Met Gly Met Ser Arg Thr Val Val	
1690 1695 1700	
ctg gat gag ccc acc agt ggg gtg gac cct tgc tcc cgg cat agc	7353
Leu Asp Glu Pro Thr Ser Gly Val Asp Pro Cys Ser Arg His Ser	
1705 1710 1715	
ctg tgg gac att ctg ctc aag tac cga gaa ggt cgt acg atc atc	7398
Leu Trp Asp Ile Leu Leu Lys Tyr Arg Glu Gly Arg Thr Ile Ile	
1720 1725 1730	
ttc aca acc cac cac ctg gat gaa gcc gaa gcg ctg agt gac cgc	7443
Phe Thr Thr His His Leu Asp Glu Ala Glu Ala Leu Ser Asp Arg	
1735 1740 1745	
gtg gcc gtc ctc cag cat ggg agg ctc agg tgc tgc ggt cct ccc	7488
Val Ala Val Leu Gln His Gly Arg Leu Arg Cys Cys Gly Pro Pro	
1750 1755 1760	
ttc tgc ctg aag gag gca tat ggc cag ggg ctc cgc ctg aca ctc	7533
Phe Cys Leu Lys Glu Ala Tyr Gly Gln Gly Leu Arg Leu Thr Leu	
1765 1770 1775	
acg agg cag cct tct gtt ctg gag gcc cat gat ctg aaa gac atg	7578
Thr Arg Gln Pro Ser Val Leu Glu Ala His Asp Leu Lys Asp Met	
1780 1785 1790	
gct tgt gtt aca tcc ctg ata aag atc tat att cca caa gca ttt	7623
Ala Cys Val Thr Ser Leu Ile Lys Ile Tyr Ile Pro Gln Ala Phe	
1795 1800 1805	
ctc aaa gac agc agt gga agt gag ctg acc tac acc att cca aag	7668
Leu Lys Asp Ser Ser Gly Ser Glu Leu Thr Tyr Thr Ile Pro Lys	
1810 1815 1820	
gac aca gac aag gcc tgc ttg aaa ggg ctc ttc cag gcc ctg gat	7713
Asp Thr Asp Lys Ala Cys Leu Lys Gly Leu Phe Gln Ala Leu Asp	
1825 1830 1835	
gag aac ctg cat cag ctg cac ctg acg ggc tat ggg atc tca gac	7758
Glu Asn Leu His Gln Leu His Leu Thr Gly Tyr Gly Ile Ser Asp	
1840 1845 1850	
acc acc tta gaa gag gtg ttt ttg atg ctt ttg caa gat tcc aac	7803
Thr Thr Leu Glu Glu Val Phe Leu Met Leu Leu Gln Asp Ser Asn	
1855 1860 1865	

aag	aaa	tct	cac	att	gcc	ctg	ggg	act	gag	tca	gag	ctg	cag	aac	7848
Lys	Lys	Ser	His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu	Gln	Asn	
				1870					1875					1880	
cac	agg	cct	aca	gga	cat	ctg	tct	ggc	tac	tgt	ggc	tcc	cta	gca	7893
His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser	Leu	Ala	
				1885					1890					1895	
cgg	ccc	gca	act	gtg	cag	ggc	gtc	cag	ctg	ctc	cgc	gca	caa	gtg	7938
Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu	Leu	Arg	Ala	Gln	Val	
				1900					1905					1910	
gcc	gcg	atc	ctg	gcc	cgg	agg	ctc	cgc	cgc	acg	ctg	cgc	gcc	ggg	7983
Ala	Ala	Ile	Leu	Ala	Arg	Arg	Leu	Arg	Arg	Thr	Leu	Arg	Ala	Gly	
				1915					1920					1925	
aag	agc	acc	ctc	gcc	gac	ctg	ctg	ctg	cca	gtc	ctc	ttc	gtg	gcc	8028
Lys	Ser	Thr	Leu	Ala	Asp	Leu	Leu	Leu	Pro	Val	Leu	Phe	Val	Ala	
				1930					1935					1940	
ttg	gcc	atg	ggc	ttg	ttc	atg	gtg	aga	ccc	ctg	gcc	acc	gag	tac	8073
Leu	Ala	Met	Gly	Leu	Phe	Met	Val	Arg	Pro	Leu	Ala	Thr	Glu	Tyr	
				1945					1950					1955	
cct	ccc	ctc	aga	ctc	aca	cct	gga	cat	tac	cag	cgg	gcc	gag	acc	8118
Pro	Pro	Leu	Arg	Leu	Thr	Pro	Gly	His	Tyr	Gln	Arg	Ala	Glu	Thr	
				1960					1965					1970	
tac	ttt	ttc	agc	agt	ggg	ggc	gac	aac	ttg	gac	ctc	acc	cgt	gtg	8163
Tyr	Phe	Phe	Ser	Ser	Gly	Gly	Asp	Asn	Leu	Asp	Leu	Thr	Arg	Val	
				1975					1980					1985	
ctt	ctg	cgg	aag	ttt	aga	gat	caa	gat	ttg	ccc	tgt	gca	gat	tta	8208
Leu	Leu	Arg	Lys	Phe	Arg	Asp	Gln	Asp	Leu	Pro	Cys	Ala	Asp	Leu	
				1990					1995					2000	
aac	cca	cgc	cag	aag	aat	tct	tca	tgc	tgg	cgc	aca	gat	ccc	ttt	8253
Asn	Pro	Arg	Gln	Lys	Asn	Ser	Ser	Cys	Trp	Arg	Thr	Asp	Pro	Phe	
				2005					2010					2015	
tct	cac	cca	gaa	ttc	cag	gat	tca	tgt	ggc	tgc	ctg	aag	tgt	cca	8298
Ser	His	Pro	Glu	Phe	Gln	Asp	Ser	Cys	Gly	Cys	Leu	Lys	Cys	Pro	
				2020					2025					2030	
aat	aga	agt	gct	agt	gct	ccc	tac	ctg	acc	aac	cac	ctg	ggc	cac	8343
Asn	Arg	Ser	Ala	Ser	Ala	Pro	Tyr	Leu	Thr	Asn	His	Leu	Gly	His	
				2035					2040					2045	
aca	ctg	ttg	aat	ctc	tca	ggc	ttc	aat	atg	gag	gag	tac	ttg	ctg	8388
Thr	Leu	Leu	Asn	Leu	Ser	Gly	Phe	Asn	Met	Glu	Glu	Tyr	Leu	Leu	
				2050					2055					2060	
gca	cca	tct	gaa	aaa	cca	agg	ctt	gga	ggc	tgg	tct	ttt	gga	tta	8433
Ala	Pro	Ser	Glu	Lys	Pro	Arg	Leu	Gly	Gly	Trp	Ser	Phe	Gly	Leu	
				2065					2070					2075	
aaa	atc	ccc	agt	gaa	gct	gga	ggc	gca	aat	gga	aac	ata	tca	aaa	8478
Lys	Ile	Pro	Ser	Glu	Ala	Gly	Gly	Ala	Asn	Gly	Asn	Ile	Ser	Lys	
				2080					2085					2090	
ccc	cca	act	ctg	gca	aag	gtg	tgg	tat	aat	cag	aag	ggc	ttt	cat	8523

Pro	Pro	Thr	Leu	Ala	Lys	Val	Trp	Tyr	Asn	Gln	Lys	Gly	Phe	His	
				2095					2100					2105	
tcc	cta	cct	tcc	tac	tta	aat	cat	cta	aac	aac	ctt	att	ttg	tgg	8568
Ser	Leu	Pro	Ser	Tyr	Leu	Asn	His	Leu	Asn	Asn	Leu	Ile	Leu	Trp	
				2110					2115					2120	
cag	cac	cta	ccc	cct	act	gtg	gac	tgg	aga	caa	tac	gga	ata	aca	8613
Gln	His	Leu	Pro	Pro	Thr	Val	Asp	Trp	Arg	Gln	Tyr	Gly	Ile	Thr	
				2125					2130					2135	
ctc	tac	agc	cac	cca	tat	gga	ggg	gcc	ttg	ctg	aac	gag	gac	aag	8658
Leu	Tyr	Ser	His	Pro	Tyr	Gly	Gly	Ala	Leu	Leu	Asn	Glu	Asp	Lys	
				2140					2145					2150	
atc	ctg	gag	agc	atc	cgt	cag	tgt	gga	gtg	gcc	ctc	tgc	atc	gtg	8703
Ile	Leu	Glu	Ser	Ile	Arg	Gln	Cys	Gly	Val	Ala	Leu	Cys	Ile	Val	
				2155					2160					2165	
ctg	gga	ttc	tcc	atc	ctg	tct	gca	tcc	atc	ggc	agc	tct	gtg	gtg	8748
Leu	Gly	Phe	Ser	Ile	Leu	Ser	Ala	Ser	Ile	Gly	Ser	Ser	Val	Val	
				2170					2175					2180	
agg	gac	agg	gtg	att	gga	gcc	aaa	agg	ttg	cag	cac	ata	agt	ggc	8793
Arg	Asp	Arg	Val	Ile	Gly	Ala	Lys	Arg	Leu	Gln	His	Ile	Ser	Gly	
				2185					2190					2195	
ctt	ggc	tac	agg	atg	tac	tgg	ttc	aca	aac	ttc	cta	tat	gac	atg	8838
Leu	Gly	Tyr	Arg	Met	Tyr	Trp	Phe	Thr	Asn	Phe	Leu	Tyr	Asp	Met	
				2200					2205					2210	
ctc	ttt	tac	ttg	gtt	tcc	gtc	tgc	ctg	tgt	gtt	gcc	gtt	att	gtc	8883
Leu	Phe	Tyr	Leu	Val	Ser	Val	Cys	Leu	Cys	Val	Ala	Val	Ile	Val	
				2215					2220					2225	
gcc	ttc	cag	tta	aca	gct	ttt	act	ttc	cgc	aag	aac	ttg	gca	gcc	8928
Ala	Phe	Gln	Leu	Thr	Ala	Phe	Thr	Phe	Arg	Lys	Asn	Leu	Ala	Ala	
				2230					2235					2240	
acg	gcc	ctc	ctg	ctg	tca	ctt	ttc	gga	tat	gca	act	ctt	cca	tgg	8973
Thr	Ala	Leu	Leu	Leu	Ser	Leu	Phe	Gly	Tyr	Ala	Thr	Leu	Pro	Trp	
				2245					2250					2255	
atg	tac	ctg	atg	tcc	aga	atc	ttt	tcc	agt	tcg	gac	gtg	gct	ttc	9018
Met	Tyr	Leu	Met	Ser	Arg	Ile	Phe	Ser	Ser	Ser	Asp	Val	Ala	Phe	
				2260					2265					2270	
att	tcc	tat	gtc	tca	cta	aac	ttc	atc	ttt	ggc	ctt	tgt	acc	atg	9063
Ile	Ser	Tyr	Val	Ser	Leu	Asn	Phe	Ile	Phe	Gly	Leu	Cys	Thr	Met	
				2275					2280					2285	
ccc	ata	acc	att	atg	ccc	cgg	ttg	cta	gcc	atc	atc	tcc	aaa	gct	9108
Pro	Ile	Thr	Ile	Met	Pro	Arg	Leu	Leu	Ala	Ile	Ile	Ser	Lys	Ala	
				2290					2295					2300	
aag	aat	tta	cag	aat	atc	tat	gat	gtc	ctc	aag	tgg	gtc	ttt	act	9153
Lys	Asn	Leu	Gln	Asn	Ile	Tyr	Asp	Val	Leu	Lys	Trp	Val	Phe	Thr	
				2305					2310					2315	
att	ttt	cct	caa	ttc	tgt	ctt	ggc	caa	gga	ctg	gta	gaa	ctc	tgc	9198
Ile	Phe	Pro	Gln	Phe	Cys	Leu	Gly	Gln	Gly	Leu	Val	Glu	Leu	Cys	

2320	2325	2330	
tat aat cag atc aaa Tyr Asn Gln Ile Lys 2335	tat gac ctg acc cac Tyr Asp Leu Thr His 2340	aac ttc ggc att gat Asn Phe Gly Ile Asp 2345	9243
tcc tat gtg agt ccc Ser Tyr Val Ser Pro 2350	ttt gag atg aac ttt Phe Glu Met Asn Phe 2355	ctg ggc tgg atc ttc Leu Gly Trp Ile Phe 2360	9288
gtg caa ctg gcc tcg Val Gln Leu Ala Ser 2365	cag ggc aca gta ctt Gln Gly Thr Val Leu 2370	ctc ctc ttg agg gtt Leu Leu Leu Arg Val 2375	9333
ctg cta cac tgg gac Leu Leu His Trp Asp 2380	ctt ctg cga tgg cca Leu Leu Arg Trp Pro 2385	agg ggt cat tct act Arg Gly His Ser Thr 2390	9378
ctc caa ggc aca gtc Leu Gln Gly Thr Val 2395	aaa tct tct aag gat Lys Ser Ser Lys Asp 2400	aca gat gtt gaa aaa Thr Asp Val Glu Lys 2405	9423
gag gaa aag aga gtg Glu Glu Lys Arg Val 2410	ttt gaa gga agg acc Phe Glu Gly Arg Thr 2415	aat gga gac att ctt Asn Gly Asp Ile Leu 2420	9468
gtg tta tac aac ctt Val Leu Tyr Asn Leu 2425	agt aaa cat tat cga Ser Lys His Tyr Arg 2430	cgc ttt ttc cag aat Arg Phe Phe Gln Asn 2435	9513
att att gct gtg caa Ile Ile Ala Val Gln 2440	gat att agt ttg ggc Asp Ile Ser Leu Gly 2445	ata cca aaa gga gag Ile Pro Lys Gly Glu 2450	9558
tgc ttt gga ctt cta Cys Phe Gly Leu Leu 2455	ggg gtg aat gga gct Gly Val Asn Gly Ala 2460	ggg aag agc acg act Gly Lys Ser Thr Thr 2465	9603
ttc aaa atg ctg aat Phe Lys Met Leu Asn 2470	ggg gaa gtt tct cta Gly Glu Val Ser Leu 2475	act tca gga cat gct Thr Ser Gly His Ala 2480	9648
atc atc agg act ccc Ile Ile Arg Thr Pro 2485	atg gga gac gcc gtg Met Gly Asp Ala Val 2490	gac ctg tct tct gct Asp Leu Ser Ser Ala 2495	9693
ggc acg gca ggc gtg Gly Thr Ala Gly Val 2500	ctc att ggc tac tgt Leu Ile Gly Tyr Cys 2505	ccc cag cag gat gcc Pro Gln Gln Asp Ala 2510	9738
ctg gac gag ctt ctg Leu Asp Glu Leu Leu 2515	act ggt tgg gaa cat Thr Gly Trp Glu His 2520	ctc tat tat tac tgt Leu Tyr Tyr Tyr Cys 2525	9783
agc tta cgc ggg att Ser Leu Arg Gly Ile 2530	cca agg cag tgc atc Pro Arg Gln Cys Ile 2535	cct gag gtt gct gga Pro Glu Val Ala Gly 2540	9828
gac ctc atc agg cgc Asp Leu Ile Arg Arg 2545	tta cac ctc gaa gcc Leu His Leu Glu Ala 2550	cac gcg gac aaa cct His Ala Asp Lys Pro 2555	9873

gtg gcc acc tac agt	ggg gga acc aag cgg	aaa ctc tct aca gcc	9918
Val Ala Thr Tyr Ser	Gly Gly Thr Lys Arg	Lys Leu Ser Thr Ala	
2560	2565	2570	
ctg gcc ctg gtg ggg	aaa cct gac att ctt	tta ttg gat gag ccc	9963
Leu Ala Leu Val Gly	Lys Pro Asp Ile Leu	Leu Leu Asp Glu Pro	
2575	2580	2585	
agc tct ggg atg gat	ccc tgc tct aag cgg	tac ctg tgg caa aca	10008
Ser Ser Gly Met Asp	Pro Cys Ser Lys Arg	Tyr Leu Trp Gln Thr	
2590	2595	2600	
ata atg aag gag gtt	cgg gaa ggc tgt gct	gcg gtg ctg acc tcc	10053
Ile Met Lys Glu Val	Arg Glu Gly Cys Ala	Ala Val Leu Thr Ser	
2605	2610	2615	
cac agc atg gag gag	tgt gag gct ctt tgc	aca aga ctg gcc ata	10098
His Ser Met Glu Glu	Cys Glu Ala Leu Cys	Thr Arg Leu Ala Ile	
2620	2625	2630	
atg gtt aac ggc agc	ttc aaa tgt ctt ggt	tct cct cag cac atc	10143
Met Val Asn Gly Ser	Phe Lys Cys Leu Gly	Ser Pro Gln His Ile	
2635	2640	2645	
aaa aat agg aag tta	aat aag aca aac tga	agtcaactgg tgaccaagca	10193
Lys Asn Arg Lys Leu	Asn Lys Thr Asn		
2650			
catggagaag gcagcatggg	tcgaggagac gctgaagaac	acgataatag aactttacag	10253
cgggctgatac tgagnttgag	aaaaggaggt gcctttgggt	ttgggacatg gaaagatagt	10313
gatgtottta gagaaacaga	agttgggttg gctgtgaggg	tgaggaggat tctgggtgtt	10373
gagtatgttc ctagggacac	tgatgggctg tttggagaat	tcctacgagg aagggctgtt	10433
gatattcatg ttctctttgc	tatggcaaga agcaggagag	aaaacttaaa gttgggggtg	10493
attattgogt atggttgaga	cggaggaata agtagctgaa	ctcatcacia ctttgaagca	10553
atattttataa actctgaagt	agagtgaact gagctcagat	tcattgttgc gcattttattt	10613
taattattac taaatatagg	caatttgcca aatgctttct	gtgcatgggc ttattttaatc	10673
tttattacgt aatggctatt	attgtacca cttcacaata	aagaaactga gacctaaaga	10733
atttaagcag ttggcccaa	aacagagtca gtattgagaa	aagcagtttt caaactcaat	10793
ttcttgactc taagtttctt	attctttcta taccattca	atagtctcaa ccaaccttat	10853
gtataactga ctaaagaaca	aagctttgct tcaagtacat	gtcgtgggca cctgcgtgat	10913
gaaacctaca cactttactc	tttgttgaat ttacttctctg	tgtattttacc gagaaaacag	10973
actatcttct tttaaaaaga	tgacagact ctgctttgaa	agcctggctt gatgaaataa	11033
aagatattta gagttaaaaa	aaaaaaaaaa aaaaaa		11068

<210> 29

<211> 2654
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> misc_feature

<222> (376)..(376)
 <223> The 'Xaa' at location 376 stands for Arg, Gly, or Trp.

<400> 29

Met Ser Phe Val Pro Lys Asp Lys Ile Leu Glu Ile Leu Lys Leu Asp
 1 5 10 15

Gln Phe Leu Thr Leu Met Ile Gln Asp Arg Leu Met Asn Ile Phe Ser
 20 25 30

Ser Leu Lys Glu Thr Ile Tyr His Leu Met Lys Ser Ser Phe Ile Leu
 35 40 45

Asp Asn Gly Glu Phe Tyr Phe Asp Thr His Gln Gly Leu Lys Phe Met
 50 55 60

Gln Asp Leu Phe Asn Ala Leu Leu Arg Glu Thr Ser Met Lys Asn Lys
 65 70 75 80

Thr Glu Asn Asn Ile Asp Phe Phe Thr Val Val Ser Gln Leu Phe Phe
 85 90 95

His Val Asn Lys Ser Glu Asp Leu Phe Lys Leu Asn Gln Asp Leu Gly
 100 105 110

Ser Ala Leu His Leu Val Arg Glu Cys Ser Thr Glu Met Ala Arg Leu
 115 120 125

Leu Asp Thr Ile Leu His Ser Pro Asn Lys Asp Phe Tyr Ala Leu Tyr
 130 135 140

Pro Thr Leu Gln Glu Val Ile Leu Ala Asn Leu Thr Asp Leu Leu Phe
 145 150 155 160

Phe Ile Asn Asn Ser Phe Pro Leu Arg Asn Arg Ala Thr Leu Glu Ile
 165 170 175

Thr Lys Arg Leu Val Gly Ala Ile Ser Arg Ala Ser Glu Glu Ser His
 180 185 190

Val Leu Lys Pro Leu Leu Glu Met Ser Gly Thr Leu Val Met Leu Leu

195	200	205
Asn Asp Ser Ala Asp Leu Arg Asp Leu Ala Thr Ser Met Asp Ser Ile 210 215 220		
Val Lys Leu Leu Lys Leu Val Lys Lys Val Ser Gly Lys Met Ser Thr 225 230 235 240		
Val Phe Lys Thr His Phe Ile Ser Asn Thr Lys Asp Ser Val Lys Phe 245 250 255		
Phe Asp Thr Leu Tyr Ser Ile Met Gln Gln Ser Val Gln Asn Leu Val 260 265 270		
Lys Glu Ile Ala Thr Leu Lys Lys Ile Asp His Phe Thr Phe Glu Lys 275 280 285		
Ile Asn Asp Leu Leu Val Pro Phe Leu Asp Leu Ala Phe Glu Met Ile 290 295 300		
Gly Val Glu Pro Tyr Ile Ser Ser Asn Ser Asp Ile Phe Ser Met Ser 305 310 315 320		
Pro Ser Ile Leu Ser Tyr Met Asn Gln Ser Lys Asp Phe Ser Asp Ile 325 330 335		
Leu Glu Glu Ile Ala Glu Phe Leu Thr Ser Val Lys Met Asn Leu Glu 340 345 350		
Asp Met Arg Ser Leu Ala Val Ala Phe Asn Asn Glu Thr Gln Thr Phe 355 360 365		
Ser Met Asp Ser Val Asn Leu Xaa Glu Glu Ile Leu Gly Cys Leu Val 370 375 380		
Pro Ile Asn Asn Ile Thr Asn Gln Met Asp Phe Leu Tyr Pro Asn Pro 385 390 395 400		
Ile Ser Thr His Ser Gly Pro Gln Asp Ile Lys Trp Glu Ile Ile His 405 410 415		
Glu Val Ile Leu Phe Leu Asp Lys Ile Leu Ser Gln Asn Ser Thr Glu 420 425 430		
Ile Gly Ser Phe Leu Lys Met Val Ile Cys Leu Thr Leu Glu Ala Leu 435 440 445		

Trp Lys Asn Leu Lys Lys Asp Asn Trp Asn Val Ser Asn Val Leu Met
 450 455 460

Thr Phe Thr Gln His Pro Asn Asn Leu Leu Lys Thr Ile Glu Thr Val
 465 470 475 480

Leu Glu Ala Ser Ser Gly Ile Lys Ser Asp Tyr Glu Gly Asp Leu Asn
 485 490 495

Lys Ser Leu Tyr Phe Asp Thr Pro Leu Ser Gln Asn Ile Thr His His
 500 505 510

Gln Leu Glu Lys Ala Ile His Asn Val Leu Ser Arg Ile Ala Leu Trp
 515 520 525

Arg Lys Gly Leu Arg Phe Asn Asn Ser Glu Trp Ile Thr Ser Thr Arg
 530 535 540

Thr Leu Phe Gln Pro Leu Phe Glu Ile Phe Ile Lys Ala Thr Thr Gly
 545 550 555 560

Lys Asn Val Thr Ser Glu Lys Glu Glu Arg Thr Glu Lys Glu Met Ile
 565 570 575

Asp Phe Pro Tyr Ser Phe Lys Pro Phe Phe Cys Leu Glu Lys Tyr Leu
 580 585 590

Gly Gly Leu Phe Val Leu Thr Lys Tyr Trp Gln Gln Ile Pro Leu Thr
 595 600 605

Asp Gln Ser Val Val Glu Ile Cys Glu Val Phe Gln Gln Thr Val Lys
 610 615 620

Pro Ser Glu Ala Met Glu Met Leu Gln Lys Val Lys Met Met Val Val
 625 630 635 640

Arg Val Leu Thr Ile Val Ala Glu Asn Pro Ser Trp Thr Lys Asp Ile
 645 650 655

Leu Cys Ala Thr Leu Ser Cys Lys Gln Asn Gly Ile Arg His Leu Ile
 660 665 670

Leu Ser Ala Ile Gln Gly Val Thr Leu Ala Gln Asp His Phe Gln Glu
 675 680 685

Ile Glu Lys Ile Trp Ser Ser Pro Asn Gln Leu Asn Cys Glu Ser Leu
 690 695 700
 Ser Lys Asn Leu Ser Ser Thr Leu Glu Ser Phe Lys Ser Ser Leu Glu
 705 710 715 720
 Asn Ala Thr Gly Gln Asp Cys Thr Ser Gln Pro Arg Leu Glu Thr Val
 725 730 735
 Gln Gln His Leu Tyr Met Leu Ala Lys Ser Leu Glu Glu Thr Trp Ser
 740 745 750
 Ser Gly Asn Pro Ile Met Thr Phe Leu Ser Asn Phe Thr Val Thr Glu
 755 760 765
 Asp Val Lys Ile Lys Asp Leu Met Lys Asn Ile Thr Lys Leu Thr Glu
 770 775 780
 Glu Leu Arg Ser Ser Ile Gln Ile Ser Asn Glu Thr Ile His Ser Ile
 785 790 795 800
 Leu Glu Ala Asn Ile Ser His Ser Lys Val Leu Phe Ser Ala Leu Thr
 805 810 815
 Val Ala Leu Ser Gly Lys Cys Asp Gln Glu Ile Leu His Leu Leu Leu
 820 825 830
 Thr Phe Pro Lys Gly Glu Lys Ser Trp Ile Ala Ala Glu Glu Leu Cys
 835 840 845
 Ser Leu Pro Gly Ser Lys Val Tyr Ser Leu Ile Val Leu Leu Ser Arg
 850 855 860
 Asn Leu Asp Val Arg Ala Phe Ile Tyr Lys Thr Leu Met Pro Ser Glu
 865 870 875 880
 Ala Asn Gly Leu Leu Asn Ser Leu Leu Asp Ile Val Ser Ser Leu Ser
 885 890 895
 Ala Leu Leu Ala Lys Ala Gln His Val Phe Glu Tyr Leu Pro Glu Phe
 900 905 910
 Leu His Thr Phe Lys Ile Thr Ala Leu Leu Glu Thr Leu Asp Phe Gln
 915 920 925

Gln Val Ser Gln Asn Val Gln Ala Arg Ser Ser Ala Phe Gly Ser Phe
 930 935 940

Gln Phe Val Met Lys Met Val Cys Lys Asp Gln Ala Ser Phe Leu Ser
 945 950 955 960

Asp Ser Asn Met Phe Ile Asn Leu Pro Arg Val Lys Glu Leu Leu Glu
 965 970 975

Asp Asp Lys Glu Lys Phe Asn Ile Pro Glu Asp Ser Thr Pro Phe Cys
 980 985 990

Leu Lys Leu Tyr Gln Glu Ile Leu Gln Leu Pro Asn Gly Ala Leu Val
 995 1000 1005

Trp Thr Phe Leu Lys Pro Ile Leu His Gly Lys Ile Leu Tyr Thr
 1010 1015 1020

Pro Asn Thr Pro Glu Ile Asn Lys Val Ile Gln Lys Ala Asn Tyr
 1025 1030 1035

Thr Phe Tyr Ile Val Asp Lys Leu Lys Thr Leu Ser Glu Thr Leu
 1040 1045 1050

Leu Glu Met Ser Ser Leu Phe Gln Arg Ser Gly Ser Gly Gln Met
 1055 1060 1065

Phe Asn Gln Leu Gln Glu Ala Leu Arg Asn Lys Phe Val Arg Asn
 1070 1075 1080

Phe Val Glu Asn Gln Leu His Ile Asp Val Asp Lys Leu Thr Glu
 1085 1090 1095

Lys Leu Gln Thr Tyr Gly Gly Leu Leu Asp Glu Met Phe Asn His
 1100 1105 1110

Ala Gly Ala Gly Arg Phe Arg Phe Leu Gly Ser Ile Leu Val Asn
 1115 1120 1125

Leu Ser Ser Cys Val Ala Leu Asn Arg Phe Gln Ala Leu Gln Ser
 1130 1135 1140

Val Asp Ile Leu Glu Thr Lys Ala His Glu Leu Leu Gln Gln Asn
 1145 1150 1155

Ser Phe Leu Ala Ser Ile Ile Phe Ser Asn Ser Leu Phe Asp Lys

1160		1165		1170
Asn Phe Arg Ser Glu Ser Val Lys Leu Pro Pro His Val Ser Tyr				
1175		1180		1185
Thr Ile Arg Thr Asn Val Leu Tyr Ser Val Arg Thr Asp Val Val				
1190		1195		1200
Lys Asn Pro Ser Trp Lys Phe His Pro Gln Asn Leu Pro Ala Asp				
1205		1210		1215
Gly Phe Lys Tyr Asn Tyr Val Phe Ala Pro Leu Gln Asp Met Ile				
1220		1225		1230
Glu Arg Ala Ile Ile Leu Val Gln Thr Gly Gln Glu Ala Leu Glu				
1235		1240		1245
Pro Ala Ala Gln Thr Gln Ala Ala Pro Tyr Pro Cys His Thr Ser				
1250		1255		1260
Asp Leu Phe Leu Asn Asn Val Gly Phe Phe Phe Pro Leu Ile Met				
1265		1270		1275
Met Leu Thr Trp Met Val Ser Val Ala Ser Met Val Arg Lys Leu				
1280		1285		1290
Val Tyr Glu Gln Glu Ile Gln Ile Glu Glu Tyr Met Arg Met Met				
1295		1300		1305
Gly Val His Pro Val Ile His Phe Leu Ala Trp Phe Leu Glu Asn				
1310		1315		1320
Met Ala Val Leu Thr Ile Ser Ser Ala Thr Leu Ala Ile Val Leu				
1325		1330		1335
Lys Thr Ser Gly Ile Phe Ala His Ser Asn Thr Phe Ile Val Phe				
1340		1345		1350
Leu Phe Leu Leu Asp Phe Gly Met Ser Val Val Met Leu Ser Tyr				
1355		1360		1365
Leu Leu Ser Ala Phe Phe Ser Gln Ala Asn Thr Ala Ala Leu Cys				
1370		1375		1380
Thr Ser Leu Val Tyr Met Ile Ser Phe Leu Pro Tyr Ile Val Leu				
1385		1390		1395

Leu Val	Leu His Asn Gln	Leu	Ser Phe Val Asn Gln	Thr Phe Leu
1400		1405		1410
Cys Leu	Leu Ser Thr Thr	Ala	Phe Gly Gln Gly Val	Phe Phe Ile
1415		1420		1425
Thr Phe	Leu Glu Gly Gln	Glu	Thr Gly Ile Gln Trp	Asn Asn Met
1430		1435		1440
Tyr Gln	Ala Leu Glu Gln	Gly	Gly Met Thr Phe Gly	Trp Val Cys
1445		1450		1455
Trp Met	Ile Leu Phe Asp	Ser	Ser Leu Tyr Phe Leu	Cys Gly Trp
1460		1465		1470
Tyr Leu	Ser Asn Leu Ile	Pro	Gly Thr Phe Gly Leu	Arg Lys Pro
1475		1480		1485
Trp Tyr	Phe Pro Phe Thr	Ala	Ser Tyr Trp Lys Ser	Val Gly Phe
1490		1495		1500
Leu Val	Glu Lys Arg Gln	Tyr	Phe Leu Ser Ser Ser	Leu Phe Phe
1505		1510		1515
Phe Asn	Glu Asn Phe Asp	Asn	Lys Gly Ser Ser Leu	Gln Asn Arg
1520		1525		1530
Glu Gly	Glu Leu Glu Gly	Ser	Ala Pro Gly Val Thr	Leu Val Ser
1535		1540		1545
Val Thr	Lys Glu Tyr Glu	Gly	His Lys Ala Val Val	Gln Asp Leu
1550		1555		1560
Ser Leu	Thr Phe Tyr Arg	Asp	Gln Ile Thr Ala Leu	Leu Gly Thr
1565		1570		1575
Asn Gly	Ala Gly Lys Thr	Thr	Ile Ile Ser Met Leu	Thr Gly Leu
1580		1585		1590
His Pro	Pro Thr Ser Gly	Thr	Ile Ile Ile Asn Gly	Lys Asn Leu
1595		1600		1605
Gln Thr	Asp Leu Ser Arg	Val	Arg Met Glu Leu Gly	Val Cys Pro
1610		1615		1620

Gln	Gln	Asp	Ile	Leu	Leu	Asp	Asn	Leu	Thr	Val	Arg	Glu	His	Leu
1625						1630					1635			
Leu	Leu	Phe	Ala	Ser	Ile	Lys	Ala	Pro	Gln	Trp	Thr	Lys	Lys	Glu
1640						1645					1650			
Leu	His	Gln	Gln	Val	Asn	Gln	Thr	Leu	Gln	Asp	Val	Asp	Leu	Thr
1655						1660					1665			
Gln	His	Gln	His	Lys	Gln	Thr	Arg	Ala	Leu	Ser	Gly	Gly	Leu	Lys
1670						1675					1680			
Arg	Lys	Leu	Ser	Leu	Gly	Ile	Ala	Phe	Met	Gly	Met	Ser	Arg	Thr
1685						1690					1695			
Val	Val	Leu	Asp	Glu	Pro	Thr	Ser	Gly	Val	Asp	Pro	Cys	Ser	Arg
1700						1705					1710			
His	Ser	Leu	Trp	Asp	Ile	Leu	Leu	Lys	Tyr	Arg	Glu	Gly	Arg	Thr
1715						1720					1725			
Ile	Ile	Phe	Thr	Thr	His	His	Leu	Asp	Glu	Ala	Glu	Ala	Leu	Ser
1730						1735					1740			
Asp	Arg	Val	Ala	Val	Leu	Gln	His	Gly	Arg	Leu	Arg	Cys	Cys	Gly
1745						1750					1755			
Pro	Pro	Phe	Cys	Leu	Lys	Glu	Ala	Tyr	Gly	Gln	Gly	Leu	Arg	Leu
1760						1765					1770			
Thr	Leu	Thr	Arg	Gln	Pro	Ser	Val	Leu	Glu	Ala	His	Asp	Leu	Lys
1775						1780					1785			
Asp	Met	Ala	Cys	Val	Thr	Ser	Leu	Ile	Lys	Ile	Tyr	Ile	Pro	Gln
1790						1795					1800			
Ala	Phe	Leu	Lys	Asp	Ser	Ser	Gly	Ser	Glu	Leu	Thr	Tyr	Thr	Ile
1805						1810					1815			
Pro	Lys	Asp	Thr	Asp	Lys	Ala	Cys	Leu	Lys	Gly	Leu	Phe	Gln	Ala
1820						1825					1830			
Leu	Asp	Glu	Asn	Leu	His	Gln	Leu	His	Leu	Thr	Gly	Tyr	Gly	Ile
1835						1840					1845			

Ser	Asp	Thr	Thr	Leu	Glu	Glu	Val	Phe	Leu	Met	Leu	Leu	Gln	Asp
1850						1855					1860			
Ser	Asn	Lys	Lys	Ser	His	Ile	Ala	Leu	Gly	Thr	Glu	Ser	Glu	Leu
1865						1870					1875			
Gln	Asn	His	Arg	Pro	Thr	Gly	His	Leu	Ser	Gly	Tyr	Cys	Gly	Ser
1880						1885					1890			
Leu	Ala	Arg	Pro	Ala	Thr	Val	Gln	Gly	Val	Gln	Leu	Leu	Arg	Ala
1895						1900					1905			
Gln	Val	Ala	Ala	Ile	Leu	Ala	Arg	Arg	Leu	Arg	Arg	Thr	Leu	Arg
1910						1915					1920			
Ala	Gly	Lys	Ser	Thr	Leu	Ala	Asp	Leu	Leu	Leu	Pro	Val	Leu	Phe
1925						1930					1935			
Val	Ala	Leu	Ala	Met	Gly	Leu	Phe	Met	Val	Arg	Pro	Leu	Ala	Thr
1940						1945					1950			
Glu	Tyr	Pro	Pro	Leu	Arg	Leu	Thr	Pro	Gly	His	Tyr	Gln	Arg	Ala
1955						1960					1965			
Glu	Thr	Tyr	Phe	Phe	Ser	Ser	Gly	Gly	Asp	Asn	Leu	Asp	Leu	Thr
1970						1975					1980			
Arg	Val	Leu	Leu	Arg	Lys	Phe	Arg	Asp	Gln	Asp	Leu	Pro	Cys	Ala
1985						1990					1995			
Asp	Leu	Asn	Pro	Arg	Gln	Lys	Asn	Ser	Ser	Cys	Trp	Arg	Thr	Asp
2000						2005					2010			
Pro	Phe	Ser	His	Pro	Glu	Phe	Gln	Asp	Ser	Cys	Gly	Cys	Leu	Lys
2015						2020					2025			
Cys	Pro	Asn	Arg	Ser	Ala	Ser	Ala	Pro	Tyr	Leu	Thr	Asn	His	Leu
2030						2035					2040			
Gly	His	Thr	Leu	Leu	Asn	Leu	Ser	Gly	Phe	Asn	Met	Glu	Glu	Tyr
2045						2050					2055			
Leu	Leu	Ala	Pro	Ser	Glu	Lys	Pro	Arg	Leu	Gly	Gly	Trp	Ser	Phe
2060						2065					2070			
Gly	Leu	Lys	Ile	Pro	Ser	Glu	Ala	Gly	Gly	Ala	Asn	Gly	Asn	Ile

2075	2080	2085
Ser Lys Pro Pro Thr Leu Ala Lys Val Trp Tyr Asn Gln Lys Gly 2090 2095 2100		
Phe His Ser Leu Pro Ser Tyr Leu Asn His Leu Asn Asn Leu Ile 2105 2110 2115		
Leu Trp Gln His Leu Pro Pro Thr Val Asp Trp Arg Gln Tyr Gly 2120 2125 2130		
Ile Thr Leu Tyr Ser His Pro Tyr Gly Gly Ala Leu Leu Asn Glu 2135 2140 2145		
Asp Lys Ile Leu Glu Ser Ile Arg Gln Cys Gly Val Ala Leu Cys 2150 2155 2160		
Ile Val Leu Gly Phe Ser Ile Leu Ser Ala Ser Ile Gly Ser Ser 2165 2170 2175		
Val Val Arg Asp Arg Val Ile Gly Ala Lys Arg Leu Gln His Ile 2180 2185 2190		
Ser Gly Leu Gly Tyr Arg Met Tyr Trp Phe Thr Asn Phe Leu Tyr 2195 2200 2205		
Asp Met Leu Phe Tyr Leu Val Ser Val Cys Leu Cys Val Ala Val 2210 2215 2220		
Ile Val Ala Phe Gln Leu Thr Ala Phe Thr Phe Arg Lys Asn Leu 2225 2230 2235		
Ala Ala Thr Ala Leu Leu Leu Ser Leu Phe Gly Tyr Ala Thr Leu 2240 2245 2250		
Pro Trp Met Tyr Leu Met Ser Arg Ile Phe Ser Ser Ser Asp Val 2255 2260 2265		
Ala Phe Ile Ser Tyr Val Ser Leu Asn Phe Ile Phe Gly Leu Cys 2270 2275 2280		
Thr Met Pro Ile Thr Ile Met Pro Arg Leu Leu Ala Ile Ile Ser 2285 2290 2295		
Lys Ala Lys Asn Leu Gln Asn Ile Tyr Asp Val Leu Lys Trp Val		

2300	2305	2310
Phe Thr Ile Phe Pro Gln Phe Cys Leu Gly Gln Gly Leu Val Glu 2315 2320 2325		
Leu Cys Tyr Asn Gln Ile Lys Tyr Asp Leu Thr His Asn Phe Gly 2330 2335 2340		
Ile Asp Ser Tyr Val Ser Pro Phe Glu Met Asn Phe Leu Gly Trp 2345 2350 2355		
Ile Phe Val Gln Leu Ala Ser Gln Gly Thr Val Leu Leu Leu Leu 2360 2365 2370		
Arg Val Leu Leu His Trp Asp Leu Leu Arg Trp Pro Arg Gly His 2375 2380 2385		
Ser Thr Leu Gln Gly Thr Val Lys Ser Ser Lys Asp Thr Asp Val 2390 2395 2400		
Glu Lys Glu Glu Lys Arg Val Phe Glu Gly Arg Thr Asn Gly Asp 2405 2410 2415		
Ile Leu Val Leu Tyr Asn Leu Ser Lys His Tyr Arg Arg Phe Phe 2420 2425 2430		
Gln Asn Ile Ile Ala Val Gln Asp Ile Ser Leu Gly Ile Pro Lys 2435 2440 2445		
Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Ser 2450 2455 2460		
Thr Thr Phe Lys Met Leu Asn Gly Glu Val Ser Leu Thr Ser Gly 2465 2470 2475		
His Ala Ile Ile Arg Thr Pro Met Gly Asp Ala Val Asp Leu Ser 2480 2485 2490		
Ser Ala Gly Thr Ala Gly Val Leu Ile Gly Tyr Cys Pro Gln Gln 2495 2500 2505		
Asp Ala Leu Asp Glu Leu Leu Thr Gly Trp Glu His Leu Tyr Tyr 2510 2515 2520		
Tyr Cys Ser Leu Arg Gly Ile Pro Arg Gln Cys Ile Pro Glu Val 2525 2530 2535		

Ala Gly Asp Leu Ile Arg Arg Leu His Leu Glu Ala His Ala Asp
2540 2545 2550

Lys Pro Val Ala Thr Tyr Ser Gly Gly Thr Lys Arg Lys Leu Ser
2555 2560 2565

Thr Ala Leu Ala Leu Val Gly Lys Pro Asp Ile Leu Leu Leu Asp
2570 2575 2580

Glu Pro Ser Ser Gly Met Asp Pro Cys Ser Lys Arg Tyr Leu Trp
2585 2590 2595

Gln Thr Ile Met Lys Glu Val Arg Glu Gly Cys Ala Ala Val Leu
2600 2605 2610

Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu
2615 2620 2625

Ala Ile Met Val Asn Gly Ser Phe Lys Cys Leu Gly Ser Pro Gln
2630 2635 2640

His Ile Lys Asn Arg Lys Leu Asn Lys Thr Asn
2645 2650

<210> 30
<211> 142
<212> PRT
<213> Homo Sapiens

<400> 30

Met Ala Cys Trp Pro Gln Leu Arg Leu Leu Leu Trp Lys Asn Leu Thr
1 5 10 15

Phe Arg Arg Arg Gln Thr Cys Gln Leu Leu Leu Glu Val Ala Trp Pro
20 25 30

Leu Phe Ile Phe Leu Ile Leu Ile Ser Val Arg Leu Ser Tyr Pro Pro
35 40 45

Tyr Glu Gln His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
50 55 60

Gly Thr Leu Pro Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln
65 70 75 80

Gln Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser
 85 90 95

Arg Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys
 100 105 110

Ile Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro
 115 120 125

Ser Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr
 130 135 140

<210> 31
 <211> 142
 <212> PRT
 <213> Homo Sapiens

<400> 31

Met Gly Phe Leu His Gln Leu Gln Leu Leu Trp Lys Asn Val Thr
 1 5 10 15

Leu Lys Arg Arg Ser Pro Trp Val Leu Ala Phe Glu Ile Phe Ile Pro
 20 25 30

Leu Val Leu Phe Phe Ile Leu Leu Gly Leu Arg Gln Lys Lys Pro Thr
 35 40 45

Ile Ser Val Lys Glu Val Pro Phe Tyr Thr Ala Ala Pro Leu Thr Ser
 50 55 60

Ala Gly Ile Leu Pro Ser Arg Lys Leu Asp Gly Gly Trp Leu Lys Val
 65 70 75 80

Arg Gln Phe His Gly Leu Leu Val Lys Arg Phe His Cys Ala Arg Arg
 85 90 95

Asn Ser Lys Ala Leu Phe Ser Gln Ile Leu Leu Pro Ala Phe Phe Val
 100 105 110

Cys Val Ala Met Thr Val Ala Leu Ser Val Pro Glu Ile Gly Asp Leu
 115 120 125

Pro Pro Leu Val Leu Ser Pro Ser Gln Tyr His Asn Tyr Thr
 130 135 140

<210> 32
 <211> 140

<212> PRT

<213> Homo Sapiens

<400> 32

Met Ala Val Leu Arg Gln Leu Ala Leu Leu Trp Lys Asn Tyr Thr
 1 5 10 15

Leu Gln Lys Arg Lys Val Leu Val Thr Val Leu Glu Leu Phe Leu Pro
 20 25 30

Leu Leu Phe Pro Gly Ile Leu Ile Trp Leu Arg Leu Lys Ile Gln Ser
 35 40 45

Glu Asn Val Pro Asn Ala Thr Ile Tyr Pro Gly Gln Ser Ile Gln Glu
 50 55 60

Leu Pro Val Lys Leu Asn Thr Gly Leu Ala Leu His Cys Gln Gln Phe
 65 70 75 80

Trp Ala Met Phe Leu Lys Lys Ala Ala Tyr Ser Trp Arg Glu Trp Lys
 85 90 95

Met Val Ala Ala Gln Val Leu Val Pro Leu Thr Cys Val Thr Leu Ala
 100 105 110

Leu Leu Ala Ile Asn Tyr Ser Ser Glu Leu Phe Asp Asp Pro Met Leu
 115 120 125

Arg Leu Thr Leu Gly Glu Tyr Gly Arg Thr Val Val
 130 135 140

<210> 33

<211> 142

<212> PRT

<213> Homo Sapiens

<400> 33

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50

55

60

Gly Met Leu Pro Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln
 65 70 75 80

His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser
 85 90 95

His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe
 100 105 110

Leu Ala Leu Met Leu Ser Ile Val Ile Leu Pro Phe Gly Glu Tyr Pro
 115 120 125

Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr
 130 135 140

<210> 34
 <211> 145
 <212> PRT
 <213> Homo Sapiens

<400> 34

Met Ser Thr Ala Ile Arg Glu Val Gly Val Trp Arg Gln Thr Arg Thr
 1 5 10 15

Leu Leu Leu Lys Asn Tyr Leu Ile Lys Cys Arg Thr Lys Lys Ser Ser
 20 25 30

Val Gln Glu Ile Leu Phe Pro Leu Phe Phe Leu Phe Trp Leu Ile Leu
 35 40 45

Ile Ser Met Met His Pro Asn Lys Lys Tyr Glu Glu Val Pro Asn Ile
 50 55 60

Glu Leu Asn Pro Met Asp Ala Ser Leu Val Ser Thr Met Ser Leu Trp
 65 70 75 80

Lys Gln Gln Met Tyr Thr Ile Ala Lys Phe His Phe Phe Thr Leu Lys
 85 90 95

Arg Glu Ser Lys Ser Val Arg Ser Val Leu Leu Leu Leu Leu Ile Phe
 100 105 110

Phe Thr Val Gln Ile Phe Met Phe Leu Val His His Ser Phe Lys Asn
 115 120 125

Ala Val Val Pro Ile Lys Leu Val Pro Asp Leu Tyr Phe Leu Lys Pro
 130 135 140

Gly
 145

<210> 35
 <211> 142
 <212> PRT
 <213> Homo Sapiens

<400> 35

Met Asn Met Lys Gln Lys Ser Val Tyr Gln Gln Thr Lys Ala Leu Leu
 1 5 10 15

Cys Lys Asn Phe Leu Lys Lys Trp Arg Met Lys Arg Glu Ser Leu Leu
 20 25 30

Glu Trp Gly Leu Ser Ile Leu Leu Gly Leu Cys Ile Ala Leu Phe Ser
 35 40 45

Ser Ser Met Arg Asn Val Gln Phe Pro Gly Met Ala Pro Gln Asn Leu
 50 55 60

Gly Arg Val Asp Gln Thr Ala Val Ser Asp Met Gly Leu Trp Arg Met
 65 70 75 80

Gln Val Phe Ala Met Ala Arg Leu Arg Phe Leu Lys Leu Lys Arg Gln
 85 90 95

Thr Lys Val Leu Leu Thr Leu Leu Leu Val Phe Gly Ile Ala Ile Phe
 100 105 110

Pro Leu Ile Val Glu Asn Ile Ile Tyr Ala Met Leu Asn Glu Lys Ile
 115 120 125

Asp Trp Glu Phe Lys Asn Glu Leu Tyr Phe Leu Ser Pro Gly
 130 135 140

<210> 36
 <211> 152
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> MISC_FEATURE
 <222> (3)..(9)

<223> Xaa = any amino acid

<220>

<221> MISC_FEATURE

<222> (68)..(68)

<223> Xaa = any amino acid

<220>

<221> MISC_FEATURE

<222> (125)..(125)

<223> Xaa = any amino acid

<220>

<221> MISC_FEATURE

<222> (137)..(137)

<223> Xaa = any amino acid

<400> 36

Met Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Trp Thr Gln Leu Met Leu
1 5 10 15

Leu Leu Trp Lys Asn Phe Met Tyr Arg Arg Arg Gln Pro Val Gln Leu
20 25 30

Leu Val Glu Leu Leu Trp Pro Leu Phe Leu Phe Phe Ile Leu Val Ala
35 40 45

Val Arg His Ser His Pro Pro Leu Glu His His Glu Cys His Phe Pro
50 55 60

Asn Lys Pro Xaa Leu Pro Ser Ala Gly Thr Val Pro Val Gly Arg Val
65 70 75 80

Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu Lys
85 90 95

Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln Ile
100 105 110

Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Xaa Phe Ser Leu
115 120 125

Ile Val Pro Pro Phe Gly His Tyr Xaa Pro Ala Leu Arg Leu Ser Pro
130 135 140

Thr Met Tyr Gly Ala Gln Val Ser
145 150

<210> 37

<211> 142

<212> PRT

<213> Homo Sapiens

<400> 37

Met Arg Lys Arg Lys Ile Ser Val Cys Gln Gln Thr Trp Ala Leu Leu
 1 5 10 15

Cys Lys Asn Phe Leu Lys Lys Trp Arg Met Lys Arg Glu Ser Leu Met
 20 25 30

Glu Trp Leu Asn Ser Leu Leu Leu Leu Leu Cys Leu Tyr Ile Tyr Pro
 35 40 45

His Ser His Gln Val Asn Asp Phe Ser Ser Leu Leu Thr Met Asp Leu
 50 55 60

Gly Arg Val Asp Arg Lys Thr Ile Gly Gly Val Ala Leu Trp Arg Gln
 65 70 75 80

Gln Ile Cys Ala Ile Ala Arg Val Arg Leu Leu Lys Leu Lys His Glu
 85 90 95

Arg Lys Ala Leu Leu Ala Leu Leu Leu Ile Leu Met Ala Gly Phe Cys
 100 105 110

Pro Leu Leu Val Glu Tyr Thr Met Val Lys Ile Tyr Gln Asn Ser Tyr
 115 120 125

Thr Trp Glu Leu Ser Pro His Leu Tyr Phe Leu Ala Pro Gly
 130 135 140

<210> 38

<211> 143

<212> PRT

<213> Homo Sapiens

<400> 38

Met Ser Lys Arg Arg Met Ser Val Gly Gln Gln Thr Trp Ala Leu Leu
 1 5 10 15

Cys Lys Asn Cys Leu Lys Lys Trp Arg Met Lys Arg Gln Thr Leu Leu
 20 25 30

Glu Trp Leu Phe Ser Phe Leu Leu Val Leu Phe Leu Tyr Leu Phe Phe
 35 40 45

Ser Asn Leu His Gln Val His Asp Thr Pro Gln Met Ser Ser Met Asp

50 55 60
 Leu Gly Arg Val Asp Arg Lys Thr Ile Ser Gly Val Ala Leu Trp Arg
 65 70 75 80
 Gln Gln Val Cys Ala Ile Ala Lys Val Arg Phe Leu Lys Leu Lys Lys
 85 90 95
 Glu Arg Lys Ser Leu Trp Thr Ile Leu Leu Leu Phe Gly Ile Ser Phe
 100 105 110
 Ile Pro Gln Leu Leu Glu His Leu Phe Tyr Glu Ser Tyr Gln Lys Ser
 115 120 125
 Tyr Pro Trp Glu Leu Ser Pro Asn Thr Tyr Phe Leu Ser Pro Gly
 130 135 140

<210> 39
 <211> 76
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> MISC_FEATURE
 <222> (49)..(49)
 <223> Xaa = any amino acid

<220>
 <221> MISC_FEATURE
 <222> (61)..(61)
 <223> Xaa = any amino acid

<400> 39

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

<210> 40
 <211> 143
 <212> PRT
 <213> Homo Sapiens

<400> 40

Met Ala Ser Leu Phe His Gln Leu Gln Ile Leu Val Trp Lys Asn Trp
 1 5 10 15

Leu Gly Val Lys Arg Gln Pro Leu Trp Thr Leu Val Leu Ile Leu Trp
 20 25 30

Pro Val Ile Ile Phe Ile Ile Leu Ala Ile Thr Arg Thr Lys Phe Pro
 35 40 45

Pro Thr Ala Lys Pro Thr Cys Tyr Leu Ala Pro Arg Asn Leu Pro Ser
 50 55 60

Thr Gly Phe Phe Pro Gly Glu Arg Leu Asp Gly Phe Gly Leu Leu Leu
 65 70 75 80

Lys Lys Ile Met Ala Ile Leu Ile Lys Arg Phe His His Thr Arg Arg
 85 90 95

Asn Trp Lys Gly Leu Ile Ala Gln Val Ile Leu Pro Ile Val Phe Val
 100 105 110

Thr Thr Ala Met Gly Leu Gly Thr Leu Arg Asn Ser Ser Asn Ser Tyr
 115 120 125

Pro Glu Ile Gln Ile Ser Pro Ser Leu Tyr Gly Thr Ser Glu Gln
 130 135 140

<210> 41
 <211> 143
 <212> PRT
 <213> Homo Sapiens

<400> 41

Met Gly His Ala Gly Cys Gln Phe Lys Ala Leu Leu Trp Lys Asn Trp
 1 5 10 15

Leu Cys Arg Leu Arg Asn Pro Val Leu Phe Leu Ala Glu Phe Phe Trp
 20 25 30

Pro Cys Ile Leu Phe Val Ile Leu Thr Val Leu Arg Phe Gln Glu Pro
 35 40 45

Pro Arg Tyr Arg Asp Ile Cys Tyr Leu Gln Pro Arg Asp Leu Pro Ser
 50 55 60

Cys Gly Val Ile Pro Pro Ala Thr Val Gln Gly Val Gln Leu Leu Arg
 65 70 75 80

Ala Gln Val Ala Ala Ile Leu Ala Arg Arg Leu Arg Arg Thr Leu Arg
 85 90 95

Ala Gly Lys Ser Thr Leu Ala Asp Leu Leu Leu Pro Val Leu Phe Val
 100 105 110

Ala Leu Ala Met Gly Leu Phe Met Val Arg Pro Leu Ala Thr Glu Tyr
 115 120 125

Pro Pro Leu Arg Leu Thr Pro Gly His Tyr Gln Arg Ala Glu Thr
 130 135 140

<210> 42

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 42

aagtaggtct cggcccgctg

20

<210> 43

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 43

aggggaggggt actcggtggc

20

<210> 44

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 44

gtatgagaca tgggggtggca gtttg

25

<210> 45
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 45
ctctgatctg aagttcttgt cgaataagg 29

<210> 46
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 46
ctgtcattca acagcatgac cagag 25

<210> 47
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 47
attttgaaca ctttggtgca tgatggaata c 31

<210> 48
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 48
ccaaggcagt gcatccctga gg 22

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 49
ctgaaaaagt aggtctcggc 20

<210> 50
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 50
gccctggacg agcttctg 18

<210> 51
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 51
gagatgttcc caaccagtca g 21

<210> 52
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 52
gctccaccct cccacttc 18

<210> 53
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 53
gacagcagtg gaagtgag 18

<210> 54
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 54
gaaattcatg gaatgatgga c 21

<210> 55

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 55
tagtgactgt aagcaccagc 20

<210> 56
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 56
gaaatgtcag tgtcgaaatg 20

<210> 57
<211> 24

<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 57
aaagtttggc tctgtaagga agca 24

<210> 58
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 58
gaattcctgg aaaataaagc ttcaag 26

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 59
attgcactgt ttctgaccac 20

<210> 60

<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 60
tgggatgggc agcctcaca 19

<210> 61
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 61
gtgcatctga ccagtgggtg a 21

<210> 62
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 62
ttgcttgctg cttcg 15

<210> 63
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 63
gtcaactggc gaccaagcac at 22

<210> 64
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 64
gccggctgta aagttctatt atcg 24

<210> 65
<211> 25

<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 65
ttcttcagcg tctcctgcac ccatg 25

<210> 66
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 66
cgctggacgc ttccggt 17

<210> 67
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 67
gacagactgc agagcctgga a 21

<210> 68
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 68
cttgggcagc atcttgggtca atctctct 28

<210> 69
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic oligonucleotide

<400> 69
gctctggaag aactggattt g 21

<210> 70
<211> 22

<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic oligonucleotide
<400> 70
cgaaaatgat gatccatgga ac 22

<210> 71
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic oligonucleotide
<400> 71
gcaccaacct tcaaatagca g 21

<210> 72
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic oligonucleotide
<400> 72
catcagtttt cactgtggtc tc 22

<210> 73
<211> 27
<212> DNA
<213> artificial sequence
<220>
<223> synthetic oligonucleotide
<400> 73
tgactattta ccaaattctt gcttgag 27

<210> 74
<211> 22
<212> DNA
<213> artificial sequence
<220>
<223> synthetic oligonucleotide
<400> 74
tattcctccg tctcaaccat ac 22

<210> 75
<211> 25

<212> DNA

<213> artificial sequence

<220>

<223> synthetic oligonucleotide

<400> 75

attcaacaaa gagtaaagtg tgtag

25