

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 February 2005 (03.02.2005)

PCT

(10) International Publication Number
WO 2005/010039 A1

(51) International Patent Classification⁷: **C07K 14/435**,
14/47, C12Q 1/68

(74) Agents: **FULLER, Grover, F.** et al.; Pfizer Inc., P.O. Box
1027, St. Louis, MO 63006 (US).

(21) International Application Number:
PCT/IB2004/002471

(22) International Filing Date: 19 July 2004 (19.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/491,888 31 July 2003 (31.07.2003) US

(71) Applicant (for all designated States except US): **PFIZER
PRODUCTS INC.** [US/US]; Eastern Point Road, Groton,
CT 06340 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DENG, Ruitang**
[US/US]; Pfizer Global Research and Development,
Eastern Point Road, Groton, CT 06340. **WILLIS, Amy,**
Y. [US/US]; Pfizer Global Research and Development,
Eastern Point Road, Groton, CT 06340 (US). **XU, Jinghai**
[CN/US]; Pfizer Global Research & Development, Eastern
Point Road, Groton, CT 06340 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, 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, NA, 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 (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI,
SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

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.

(54) Title: BSEP POLYPEPTIDE VARIANTS AND USES THEREOF

(57) Abstract: The invention provides isolated nucleic acids encoding BSEP polypeptide variants, wherein the nucleic acids com-
prise a segment of nucleotides at a position which corresponds to the junction of Exon I and Exon II of a human BSEP polypeptide
nucleotide sequence, and the polypeptide encoded by these nucleic acids. The invention also provides vectors, host cells, and trans-
genic animals comprising a nucleic acid encoding a BSEP polypeptide variant. The invention provides antibodies that specifically
bind to the BSEP polypeptide variants of the invention. The invention also provides methods of identifying compounds that are sub-
strates and/or modulators of the BSEP polypeptide variants of the invention. Finally, the invention provides methods of inhibiting
the expression of a BSEP polypeptide variant of the invention.



WO 2005/010039 A1

BSEP POLYPEPTIDE VARIANTS AND USES THEREOF

FIELD OF THE INVENTION

5 This invention relates to the field of recombinant DNA technology and relates to nucleic acids encoding BSEP polypeptide variants, BSEP polypeptide variants and uses of these BSEP polypeptide variants and the nucleic acids encoding them.

BACKGROUND OF THE INVENTION

10 Bile salt export pump (BSEP) is an ATP-dependent bile acid transporter that is responsible for the active transport of bile acid across the hepatocyte canalicular membrane into the bile duct. The gene encoding BSEP is formally known as *ABCB11*. BSEP was formerly known as "sister of p-glycoprotein."

 The human, rat, pig and mouse BSEP genes have been isolated. See, e.g.,
15 Strautnieks *et al.*, *Nature Genetics* 20:233-38 (1998), Gerloff *et al.*, *J. Biol. Chem.*, 273:10046-50 (1998); Green *et al.*, *Gene*, 241:117-23 (2000); Childs *et al.*, *Cancer Res.*, 55:2029-34 (1995). The human BSEP gene encodes a 1321 amino acid polypeptide. It has been reported that the human BSEP gene is mutated in individuals having progressive familial intrahepatic cholestasis 2 (PFIC-2), a genetic
20 disease causing extreme pruritus, growth failure and progression to cirrhosis in the first decade of life. Strautnieks *et al.*, *Nature Genetics*, 20:233-38 (1998).

SUMMARY OF THE INVENTION

 This invention provides isolated nucleic acids encoding BSEP polypeptide
25 variants. In one embodiment the invention provides isolated nucleic acids encoding BSEP polypeptide insert variants, wherein the nucleic acids comprise a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

 This invention also provides isolated nucleic acids encoding a BSEP
30 polypeptide insert variant, wherein the nucleic acid is at least 80% identical to a nucleic acid encoding a human BSEP polypeptide insert variant, wherein the nucleic acid encoding the human BSEP polypeptide insert variant comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

The invention also provides vectors comprising nucleotides encoding a BSEP polypeptide variant. In one embodiment, the vectors comprise nucleotides encoding a BSEP polypeptide insert variant.

5 The invention also provides host cells comprising a nucleic acid encoding a BSEP polypeptide variant. In one embodiment the host cell comprises a nucleic acid encoding a BSEP polypeptide insert variant.

The invention also provides isolated BSEP polypeptide variants. In one embodiment the invention provides isolated BSEP polypeptide insert variants encoded by a nucleic acid comprising a segment of nucleotides at a position which
10 corresponds to the junction between Exon I and Exon II of a BSEP polypeptide nucleotide sequence.

The invention also provides methods of preparing a BSEP polypeptide variant comprising culturing host cells described herein under conditions that permit expression of the BSEP polypeptide variant; and isolating the BSEP polypeptide
15 variant, thereby preparing the BSEP polypeptide variant.

This invention provides antibodies, antibody chains, or fragments thereof, that specifically bind to a BSEP polypeptide variant, or a fragment of a BSEP polypeptide variant, and do not bind to a BSEP polypeptide.

The present invention also provides methods for detecting a BSEP
20 polypeptide variant in a sample comprising: (a) contacting a suitable sample with a compound that specifically binds to the BSEP polypeptide variant; and (b) determining whether any compound is bound to the BSEP polypeptide variant, where the presence of any compound bound to the BSEP polypeptide variant detects the BSEP polypeptide variant in the sample.

25 The invention also provides methods for detecting a nucleic acid which encodes a BSEP polypeptide variant comprising: (a) contacting a suitable sample with a compound capable of specifically binding a nucleic acid encoding a BSEP polypeptide variant; and (b) determining whether any compound is bound to the nucleic acid, where the presence of compound bound to the nucleic acid detects the
30 nucleic acid in the sample.

The present invention also provides for methods of determining whether a subject is expressing a BSEP polypeptide variant comprising: (a) contacting a suitable sample from the subject with a compound that specifically binds to the BSEP polypeptide variant; and (b) determining whether any compound is bound to the
35 BSEP polypeptide variant, wherein the presence of the compound bound to the

BSEP polypeptide variant indicates that the individual is expressing the BSEP polypeptide variant.

The invention also provides methods of detecting a nucleic acid which encodes a BSEP polypeptide variant comprising: (a) obtaining cDNA from mRNA
5 obtained from a suitable sample; (b) amplifying the cDNA corresponding to the nucleic acid encoding the BSEP polypeptide or a portion of said nucleic acid; (c) comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide, wherein a difference between the amplified cDNA and the DNA from a nucleic acid known to encode a BSEP polypeptide indicates the detection of a
10 nucleic acid encoding a BSEP polypeptide variant in the sample.

In one embodiment a suitable sample is obtained from a subject, and the detection of the amplified cDNA indicates that the subject is expressing the BSEP polypeptide insert variant. The invention also provides methods of detecting a nucleic acid which encodes a BSEP polypeptide insert variant comprising: (a)
15 obtaining cDNA from mRNA obtained from a suitable sample; (b) amplifying the cDNA corresponding to the insertion of the BSEP polypeptide insert variant or a portion of said insertion; and (c) comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide, wherein the presence of an insertion in the amplified cDNA indicates the detection of a nucleic acid encoding the
20 BSEP polypeptide insert variant.

The present invention provides methods of identifying compounds that are substrates and/or modulators of the BSEP polypeptide variants of the invention.

The invention provides methods for determining whether a compound is a substrate or modulator of a BSEP polypeptide variant comprising: (a) contacting the
25 compound with a BSEP polypeptide variant; and (b) determining whether any compound binds to the BSEP polypeptide variant, wherein the presence of compound bound to the BSEP polypeptide variant indicates that the compound is a substrate or modulator of the BSEP polypeptide variant.

The invention also provides methods for determining whether a compound is
30 a substrate or modulator of a BSEP polypeptide variant comprising: (a) contacting the BSEP polypeptide variant with a known binder and the compound; (b) measuring the amount of binder bound to the BSEP polypeptide variant; and (c) comparing the amount of binder bound to the BSEP polypeptide variant in step (b) with the amount of binder bound to the BSEP polypeptide variant in the absence of the compound,
35 and determining whether the presence of the compound increased or decreased the amount of binder bound to the BSEP polypeptide variant, wherein an increase or a

decrease in the amount of binder bound to the BSEP polypeptide variants indicates that the compound is a substrate or modulator of BSEP.

The invention also provides methods for determining whether a compound is a substrate of a BSEP polypeptide variant comprising: (a) contacting the compound with the BSEP polypeptide variant and ATP; and (b) detecting the ATPase activity of the BSEP polypeptide insert variant, wherein the detection of ATPase activity indicates that the compound is a substrate of the BSEP polypeptide variant.

The invention also provides methods for determining whether a compound is an modulator of a BSEP polypeptide variant comprising: (a) contacting the BSEP polypeptide variant with a substrate of the BSEP polypeptide variant, a compound and ATP; (b) measuring the ATPase activity of the BSEP polypeptide variant; and (c) comparing the ATPase activity of the BSEP polypeptide variant in step (b) with the ATPase activity of the BSEP polypeptide variant in the absence of the compound, and determining whether the compound modulated the ATPase activity of the BSEP polypeptide variant, wherein an increase or a decrease in the ATPase activity indicates that the compound is a substrate or modulator of the BSEP polypeptide variant.

The present invention also provides methods for determining whether a compound is a substrate of BSEP comprising: (a) contacting the compound with the BSEP polypeptide variant; (b) determining whether the compound is transported by the BSEP polypeptide variant, wherein the transport of the compound by the BSEP polypeptide variant indicates that the compound is a substrate of the BSEP polypeptide variant.

The invention also provides methods for determining whether a compound is a substrate or modulator of a human BSEP polypeptide variant comprising: (a) contacting the compound with the BSEP polypeptide variant and a known substrate of the BSEP polypeptide variant; (b) measuring the activity of the BSEP polypeptide variant with respect to the known substrate; and (c) comparing the activity of the BSEP polypeptide variant with respect to the known substrate in step (b) with the activity of the BSEP polypeptide with respect to the known substrate in the absence of the compound, wherein an increase or a decrease in transport of activity of the BSEP polypeptide insert variant with respect to the known substrate indicates that the compound is a substrate or modulator of the BSEP polypeptide insert variant.

The invention also provides for methods of inhibiting the expression of a BSEP polypeptide variant in a cell comprising contacting the cell with a compound capable of inhibiting the expression of the BSEP polypeptide variant. In one

embodiment, the compound capable of inhibiting the expression of a BSEP polypeptide variant is an antisense oligonucleotide capable of specifically hybridizing to the BSEP polypeptide variant. In another embodiment, the compound capable of inhibiting the expression of a BSEP polypeptide variant is an RNAi construct.

5 The invention further provides a non-human transgenic animal whose somatic and germ cells contain a heterologous nucleic acid encoding a BSEP polypeptide variant operably-linked to a promoter, said nucleic acid comprising a segment of nucleotide at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

10

DETAILED DESCRIPTION OF THE INVENTION

The articles "a," "an" and "the" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article.

Nucleic Acids

15 This invention provides isolated nucleic acids encoding BSEP polypeptide variants.

In one embodiment, the BSEP polypeptide variants will be human BSEP polypeptide variants. In another embodiment, the BSEP polypeptide variants will be rat BSEP polypeptide variants. In another embodiment, the BSEP polypeptide variants will be mouse BSEP polypeptide variants. In another embodiment, the BSEP polypeptide variants may be derived from other species including, but not limited to, dogs, guinea pigs, ferrets and rabbits.

In one embodiment the BSEP polypeptide variants will be insert variants.

25 In one embodiment the invention provides isolated nucleic acids encoding BSEP polypeptide insert variants, wherein the nucleic acids comprise a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

30 The term "isolated," as used herein, refers to material that is removed from its original or native environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated by human intervention from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of the environment in which it is found in nature.

The term "nucleic acid," as used herein, refers to either DNA or RNA. The nucleic acid can be isolated from natural sources or can be a product of chemical synthetic procedures. The term "nucleic acid" should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs. The nucleic acid can be single-stranded or double-stranded, and if single-stranded may be the coding strand or non-coding (antisense) strand.

"Nucleic acid," "nucleic acid sequence," "nucleotide sequence," or "polynucleotide sequence" refers to a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end.

The phrase "nucleic acid encoding," as used herein, refers to a nucleic acid, which directs the expression of a specific protein or peptide. The nucleic acid sequences include both the DNA strand sequence that is transcribed into RNA, and the RNA sequence that is translated into protein.

As used herein, the term "polypeptide" refers to two or more amino acids linked by a peptide bond between the α -carboxyl group of one amino acid and the α -amino group of the next amino acid. The polypeptide can be isolated from natural sources, can be produced by recombinant techniques from a eukaryotic or prokaryotic host, or can be a product of synthetic procedures.

The term "BSEP polypeptide," as used herein, refers to any BSEP polypeptide previously known. Examples of amino acid sequences of human BSEP polypeptides can be found in GenBank® (see, e.g., Accession Nos. AAC77455, NP_003733, AAD28285). All of these human BSEP polypeptides comprise 1321 amino acids. Examples of nucleic acid sequences encoding human BSEP polypeptides can be found in GenBank® (see, e.g., Accession Nos. AF091582, NM_003742 and AF136523). The junction between Exon I and Exon II in the nucleotide sequence disclosed in GenBank® Accession No. AF091582 is between nucleotides 202 and 203. The junction between Exon I and Exon II in the nucleotide sequence disclosed in GenBank® Accession No. NM_003742 is between nucleotides 202 and 203. The junction between Exon I and Exon II in the nucleotide sequence disclosed in GenBank® Accession No. AF136523 is between nucleotides 203 and 204.

As used herein, the phrase "BSEP polypeptide nucleotide sequence" refers to a nucleic acid encoding a BSEP polypeptide.

In one embodiment, the human BSEP polypeptide is encoded by the nucleotide sequence disclosed in SEQ ID NO:1. In SEQ ID NO:2, the junction

between Exon I and Exon II, corresponds to the position between nucleotides 76 and 77.

In one embodiment, the human BSEP polypeptide comprises the amino acid sequence disclosed in SEQ ID NO:2, or an amino acid sequence that substantially
5 corresponds to this sequence.

In one embodiment, the human BSEP polypeptide is encoded by the nucleotide sequence disclosed in SEQ ID NO:1, or a nucleotide sequence that substantially corresponds to this sequence.

A sequence that "substantially corresponds" to another sequence is a
10 sequence that allows single amino acid or nucleotide substitutions, deletions and/or insertions. In one embodiment, sequences that substantially correspond have 80% sequence identity. In another embodiment, sequences that substantially correspond have 85% sequence identity. In another embodiment, sequences that substantially correspond have 90% sequence identity. In another embodiment, sequences that
15 substantially correspond have 95% sequence identity. In another embodiment, sequences that substantially correspond have 97% sequence identity. In another embodiment, sequences that substantially correspond have 99% sequence identity.

As used herein, the term "% sequence identity" refers to the sequence relationship between two or more nucleic acid molecules or polypeptides. The terms
20 "identical" or percent "identity", in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison
25 algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence, provided that the test sequence has a designated or substantial identity to a reference sequence.

Comparisons between the sequences of two or more polynucleotides or polypeptides can be performed using the local homology algorithm of Smith and
30 Waterman, *Adv. Appl. Math.* 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970), or the method of Pearson and Lipman, *PNAS* 85: 2444 (1988). Computer programs implementing these methods can be used and include, BLAST, GAP, BESTFIT, FASTA, and TFASTA which are offered in the Wisconsin Genetics Software Package, Genetics Computer Group, 575
35 Science Dr., Madison, Wis.

The term "BSEP polypeptide variant," as used herein, refers to a BSEP polypeptide comprising amino acid insertions, deletions and/or substitutions with respect to the BSEP polypeptides previously known. Examples of previously known BSEP polypeptides can be found in GenBank® Accession Nos. AAC77455,
5 NP_003733 and AAD28285.

The term "BSEP polypeptide insert variant," as used herein, refers to a human BSEP polypeptide comprising an insertion of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence. In one embodiment, the human BSEP polypeptide
10 insert variant of the invention comprises more than 1321 amino acids.

In one embodiment of the nucleic acids described herein, the segment of nucleotides at the position that corresponds to the junction between Exon I and Exon II comprises nucleotides encoding the amino acid sequence set forth in SEQ ID NO:11, or such amino acid sequence comprising at least one conservative amino
15 acid substitution. In one embodiment, the segment of nucleotides encoding the amino acid sequence set forth in SEQ ID NO:1 comprises the nucleotide sequence set forth in SEQ ID NO:12.

In one embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:4. In
20 another embodiment, the nucleic acid encoding SEQ ID NO:4 comprises the nucleotide sequence set forth in SEQ ID NO:3.

In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:4 with a conservative amino acid substitution. In another embodiment, the nucleic acid
25 encodes a human BSEP polypeptide insert variant comprising 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 conservative amino acid substitutions with respect to SEQ ID NO:4. In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 conservative amino acid substitutions with respect to SEQ ID NO:4. In another embodiment, the nucleic acid
30 encodes a human BSEP polypeptide insert variant comprising 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 conservative amino acid substitutions with respect to SEQ ID NO:4. In yet another embodiment, the nucleic acid encodes a human BSEP polypeptide comprising more than 30 conservative amino acid substitutions with respect to SEQ ID NO:4.

35 In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:8. In

another embodiment, the nucleic acid encoding SEQ ID NO:8 comprises the nucleotide sequence set forth in SEQ ID NO:7.

In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:8 with a conservative amino acid substitution. In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 conservative amino acid substitutions with respect to SEQ ID NO:8. In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 conservative amino acid substitutions with respect to SEQ ID NO:8. In another embodiment, the nucleic acid encodes a human BSEP polypeptide insert variant comprising 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 conservative amino acid substitutions with respect to SEQ ID NO:8. In yet another embodiment, the nucleic acid encodes a human BSEP polypeptide comprising more than 30 conservative amino acid substitutions with respect to SEQ ID NO:8.

The phrase "conservative amino acid substitution" refers to the substitution of one amino acid with a different amino acid having certain common properties. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer, *Principles of Protein Structure*, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer, *Principles of Protein Structure*, Springer-Verlag). Examples of amino acid groups defined in this manner include: (i) a charged group, consisting of Glu and Asp, Lys, Arg and His; (ii) a positively-charged group, consisting of Lys, Arg and His; (iii) a negatively-charged group, consisting of Glu and Asp; (iv) an aromatic group, consisting of Phe, Tyr and Trp; (v) a nitrogen ring group, consisting of His and Trp; (vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile; (vii) a slightly-polar group, consisting of Met and Cys; (viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro; (ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys; and (x) a small hydroxyl group consisting of Ser and Thr.

In another embodiment, the nucleic acids described herein encode BSEP polypeptide insert variants comprising an amino acid sequence substantially

corresponding to the amino acid sequence set forth in SEQ ID NO:4 or SEQ ID NO:8.

The nucleic acids of the invention encoding BSEP polypeptide variants of the invention are not limited to the nucleic acids disclosed herein. Now that the existence of human BSEP polypeptide variants has been disclosed by the present document, additional nucleic acid sequences encoding other BSEP polypeptide variants can be discovered using methods known to a person having ordinary skill in the art. For example, additional BSEP polypeptide variants can be discovered by analysis of PCR (polymerase chain reaction) products from segments of total RNA and mRNA (see, e.g., discussion *infra*, regarding methods of detecting nucleic acids encoding BSEP polypeptide variants). Additional BSEP polypeptide variants can be also discovered by directly cloning, and sequencing, nucleic acids encoding BSEP polypeptides from a cDNA library. These methods could be used to identify BSEP polypeptide variants in other species.

The invention also provides isolated nucleic acids capable of hybridizing under high stringency conditions to a nucleic acid described herein.

The invention provides an isolated nucleic acid capable of hybridizing under high stringency conditions to a nucleic acid encoding a human BSEP polypeptide insert variant, wherein the nucleic acid comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to a nucleic acid encoding SEQ ID NO:6, or a fragment thereof.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to SEQ ID NO:5, or a fragment thereof.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to a nucleic acid encoding SEQ ID NO:10, or a fragment thereof.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to SEQ ID NO:9, or a fragment thereof.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to a nucleic acid encoding SEQ ID NO:11, or a fragment thereof.

This invention also provides an isolated nucleic acid capable of hybridizing under high stringency conditions to SEQ ID NO:12, or a fragment thereof.

As used here, "hybridization," means hydrogen bonding between complementary nucleoside or nucleotide bases. "Complementary," as used herein, refers to the capacity for precise pairing between two nucleosides. The oligonucleotide and the DNA or RNA are complementary to each other when a sufficient number of corresponding positions in each molecule are occupied by nucleotides which can hydrogen bond with each other. Thus, "specifically hybridizing to" and "complementary" are terms which are used to indicate a sufficient degree of complementarity or precise pairing such that stable and specific binding occurs between the oligonucleotide and the DNA or RNA target. Hybridization can be determined by using, e.g., standard nucleic acid hybridization techniques. Stringency can be controlled, e.g., by varying salt and/or denaturant concentrations and/or by varying the temperature. See, for example, *Molecular Cloning, a Laboratory Manual*, ed. by Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory Press, 1989.

High stringency hybridization conditions are selected at about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, high stringency conditions will be those in which the salt concentration is at least about 0.02 molar at pH 7 and the temperature is at least about 60°C. As other factors may significantly affect the stringency of hybridization, including, among others, base composition and size of the complementary strands, the presence of organic solvents, *i.e.* salt or formamide concentration, and the extent of base mismatching. When used in the claims, the term "high stringency conditions" refers to conditions involving hybridization at about 68°C in a 6x SSC solution, followed by washing at about 68°C in a 0.6x SSC solution or similar conditions (1x SSC is 0.15 M sodium chloride and 0.015 M sodium citrate). Those skilled in the art will recognize, based upon the present description, how such conditions can be varied to alter specificity and selectivity to achieve higher or lower stringency conditions.

This invention also provides isolated nucleic acids encoding a BSEP polypeptide insert variant, wherein the nucleic acid is at least 80% identical to a nucleic acid encoding a human BSEP polypeptide insert variant, wherein the nucleic acid encoding the human BSEP polypeptide insert variant comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence. In increasingly more preferred embodiments, rather than 80%, the percent identity is 85%, 90%, 95%, 97%, or 99%.

In one embodiment, the invention provides isolated nucleic acids encoding a BSEP polypeptide insert variant, wherein the nucleic acid is at least 80% identical to a nucleic acid encoding a human BSEP polypeptide comprising the amino acid sequence set forth in SEQ ID NO:4, wherein the nucleic acid encoding the human
5 BSEP polypeptide insert variant comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence. In increasingly more preferred embodiments, rather than 80%, the percent identity is 85%, 90%, 95%, 97%, or 99%.

In one embodiment, the invention provides isolated nucleic acids encoding a
10 BSEP polypeptide insert variant, wherein the nucleic acid is at least 80% identical to a nucleic acid encoding a human BSEP polypeptide comprising the amino acid sequence set forth in SEQ ID NO:8, wherein the nucleic acid encoding the human BSEP polypeptide insert variant comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP
15 polypeptide nucleotide sequence. In increasingly more preferred embodiments, rather than 80%, the percent identity is 85%, 90%, 95%, 97%, or 99%.

The nucleic acids described herein can be labeled with a detectable marker. Detectable markers include, but are not limited to: a radioactive marker, a colorimetric marker, a luminescent marker, an enzyme marker and a fluorescent
20 marker. Radioactive markers include, but are not limited to, ^3H , ^{14}C , ^{32}P , ^{33}P , ^{35}S , ^{36}Cl , ^{51}Cr , ^{57}Co , ^{59}Co , ^{59}Fe , ^{90}Y , ^{125}I , ^{131}I , and ^{186}Re . Fluorescent markers include, but are not limited to, fluorescein, rhodamine and auramine. Colorimetric markers include, but are not limited to, biotin and digoxigenin. Any suitable method for attaching markers to nucleic acids may be used with the nucleotides of the invention,
25 and many such methods are well known in the art.

Further, the invention provides nucleic acids complementary to the nucleic acids disclosed herein. By a nucleic acid sequence "homologous to" or "complementary to", it is meant a nucleic acid that selectively hybridizes, duplexes or binds to a target nucleic acid sequence. For example, adenine is complementary to
30 thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. A nucleic acid sequence, which is homologous to a target sequence, can include sequences, which are shorter or longer than the target sequence as long as they meet the functional test set forth.

It will be readily understood by those skilled in the art and it is intended here,
35 that when reference is made to particular sequence listings, such reference includes sequences which substantially correspond to its complementary sequence and those

described including allowances for minor sequencing errors, single base changes, deletions, substitutions and the like, such that any such sequence variation corresponds to the nucleic acid encoding the polypeptide to which the relevant sequence listing relates.

5 **Vectors**

The invention also provides vectors comprising nucleotides encoding a BSEP polypeptide variant. In one embodiment, the vectors comprise nucleotides encoding a human BSEP polypeptide variant. In one embodiment, the vectors comprise nucleotides encoding a BSEP polypeptide insert variant. In one embodiment, the
10 vectors comprise the nucleotide sequences described herein. The vectors include, but are not limited to, a virus, plasmid, cosmid, λ phage or a yeast artificial chromosome (YAC).

In accordance with the invention, numerous vector systems may be employed. For example, one class of vectors utilizes DNA elements which are
15 derived from animal viruses such as, for example, bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (Rous Sarcoma Virus, MMTV or MoMLV) or SV40 virus. Another class of vectors utilizes RNA elements derived from RNA viruses such as Semliki Forest virus, Eastern Equine Encephalitis virus and Flaviviruses.

20 Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy to an auxotrophic host, biocide resistance, (e.g., antibiotics), or resistance to heavy metals such as copper, or the like. The selectable marker gene can be
25 either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcriptional promoters, enhancers, and termination signals.

Exemplary vectors include, for example, those described by Okayama and
30 Berg, *Mol. Cell. Biol.* 3:280 (1983).

Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression vectors may be transfected or introduced into an appropriate host cell. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate
35 precipitation, electroporation, retroviral transduction, viral transfection, gene gun, lipid based transfection or other conventional techniques. In the case of protoplast fusion,

the cells are grown in media and screened for the appropriate activity. Expression of the gene encoding a BSEP polypeptide variant results in production of the BSEP polypeptide variant.

5 Methods and conditions for culturing the resulting transfected cells and for recovering the BSEP polypeptide variant so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed, based upon the present description.

Cells

10 The invention also provides host cells comprising a nucleic acid encoding a BSEP polypeptide variant. In one embodiment, the BSEP polypeptide variant is a human BSEP polypeptide variant. In one embodiment the BSEP polypeptide variant is an insert variant.

In one embodiment, the host cells are genetically engineered to comprise nucleic acids encoding BSEP polypeptide variants.

15 In another embodiment, the nucleic acid encoding the BSEP polypeptide variant can be situated on an episome in a host cell. An "episome" is any of a group of genetic elements consisting of DNA and capable of giving selective advantage to the bacteria in which they occur. Episomes may be attached to the bacterial cell membrane or become part of the chromosome. Cells with episomes act like males during conjugation, a mating process in certain bacteria. During conjugation, cells 20 lacking the episome may receive either the episome or the episome plus the genes to which it is attached. Experiments involving gene transfers from cells in which episomes have been incorporated in the chromosomes have been used to determine the locations of genes on the chromosome.

25 In one embodiment, the host cells are genetically engineered by using an expression cassette. The phrase "expression cassette," refers to nucleotide sequences, which are capable of affecting expression of a gene in hosts compatible with such sequences. Such cassettes may include a promoter, an open reading frame with or without introns, and a termination signal. Additional factors necessary 30 or helpful in effecting expression may also be used, such as, for example, an inducible promoter.

The invention also provides host cells comprising the vectors described herein.

35 The cell can be, but is not limited to, a eukaryotic cell, a bacterial cell, an insect cell, or a human cell. Suitable eukaryotic cells include, but are not limited to,

Vero cells, HeLa cells, COS cells, CHO cells, HEK293 cells, BHK cells and MDCKII cells. Suitable insect cells include, but are not limited to, Sf9 cells.

Host cells comprising any of the nucleic acids described herein are useful, e.g., for identifying substrates and modulators of the BSEP polypeptide variants, and/or for identifying potentially toxic compounds or drugs.

BSEP Polypeptides Variants

The invention also provides isolated BSEP polypeptide variants. As discussed above, the term "BSEP polypeptide variant," as used herein, refers to a BSEP polypeptide comprising amino acid insertions, deletions and/or substitutions with respect to the BSEP polypeptides previously known. In one embodiment, the invention provides isolated human BSEP polypeptide insert variants.

In one embodiment, the invention provides isolated BSEP polypeptide variants. In one embodiment, the invention provides isolated rat BSEP polypeptide insert variants. In one embodiment, the invention provides isolated mouse BSEP polypeptide insert variants. In other embodiments the isolated BSEP polypeptide variants will be derived from other species, including but not limited to, dogs, pigs, guinea pigs and rabbits.

In one embodiment the invention provides isolated BSEP polypeptide insert variants encoded by a nucleic acid comprising a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a nucleic acid encoding a BSEP polypeptide.

In one embodiment the invention provides isolated human BSEP polypeptide insert variants encoded by a nucleic acid comprising a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence. In one embodiment, the segment of nucleotides encodes the amino acid sequence set forth in SEQ ID NO:11. In another embodiment, the segment of nucleotides encodes an amino acid sequence having one or more conservative amino acid substitutions with respect to the amino acid sequence set forth in SEQ ID NO:11.

In one embodiment, the isolated BSEP polypeptide insert variant comprises the amino acid sequence set forth in SEQ ID NO:4. In one embodiment, the isolated human BSEP polypeptide comprises the amino acid sequence set forth in SEQ ID NO:8.

In another embodiment, the isolated BSEP polypeptide insert variant comprises an amino acid sequence having one or more conservative amino acid substitutions with respect to the sequence disclosed in SEQ ID NO:4. In one

embodiment, the isolated BSEP polypeptide variant comprises 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 conservative amino acid substitutions relative to SEQ ID NO:4. In another embodiment the isolated BSEP polypeptide variant comprises 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 conservative amino acid substitutions relative to SEQ ID NO:4. In another embodiment the isolated BSEP polypeptide variant comprises 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 conservative amino acid substitutions relative to SEQ ID NO:4. In yet another embodiment, the isolated BSEP polypeptide variant comprises more than 30 conservative amino acid substitutions relative to SEQ ID NO:4.

In another embodiment, the isolated BSEP polypeptide insert variant comprises an amino acid sequence having one or more conservative amino acid substitutions with respect to the sequence disclosed in SEQ ID NO:8. In one embodiment the isolated BSEP polypeptide variant comprises 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 conservative amino acid substitutions relative to SEQ ID NO:8. In another embodiment the isolated BSEP polypeptide variant comprises 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 conservative amino acid substitutions relative to SEQ ID NO:8. In another embodiment the isolated BSEP polypeptide variant comprises 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 conservative amino acid substitutions relative to SEQ ID NO:8. In yet another embodiment, the isolated BSEP polypeptide variant comprises more than 30 conservative amino acid substitutions relative to SEQ ID NO:1.

In another embodiment, the isolated BSEP polypeptide insert variant comprises an amino acid sequence substantially corresponding to the amino acid sequence set forth in SEQ ID NO:4 or SEQ ID NO:8.

The BSEP polypeptide variants can be isolated from natural sources, or can be a product of chemical synthetic procedures, or can be produced by recombinant techniques from a prokaryotic or eukaryotic host.

The BSEP polypeptide variants can be used, for example, during a drug screening and development process, e.g., to assess the potential toxic impact of a compound.

Methods of Preparing BSEP Polypeptide Variants

The invention also provides methods of preparing BSEP polypeptide variants comprising culturing host cells under conditions that permit expression of the BSEP polypeptide variant; and isolating the BSEP polypeptide variant, thereby preparing the BSEP polypeptide variant. In one embodiment, the invention provides a method of preparing a human BSEP polypeptide variant. In one embodiment, the invention provides a method of preparing a BSEP polypeptide insert variant. Procedures for preparing a polypeptide using the above describe method are well known to those

skilled in the art. See, for example, Membrane Protein Purification and Crystallization, ed. by Hunte, Jagow and Schagger, Academic Press, 2002.

Antibodies

This invention provides antibodies, antibody chains, or fragments thereof that
5 specifically bind to a BSEP polypeptide variant or a fragment of a BSEP polypeptide variant. In one embodiment the BSEP polypeptide variant to which the antibody binds is a human BSEP polypeptide variant. In one embodiment the BSEP polypeptide variant to which the antibody binds is BSEP polypeptide insert variant.

The term "specifically binds," as use herein, refers to a binding reaction that is
10 determinative of the presence of a target (such a specific polypeptide or nucleic acid) in a population of proteins and other biologics.

In one embodiment, the antibody specifically binds a BSEP polypeptide comprising the amino acid sequence set forth in SEQ ID NO:11. In one embodiment, the invention provides an antibody that specifically binds to a human BSEP
15 polypeptide insert comprising the amino acid sequence set forth in SEQ ID NO:4. In another embodiment, the invention provides an antibody that specifically binds to a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:8.

In the context of the present invention, antibodies specifically binding to the
20 BSEP polypeptide insert variants can be screened and tested to identify those antibodies that specifically recognize human BSEP polypeptide insert variants, but do not recognize the previously known human BSEP polypeptides. The antibodies can be screened using methods known to people skilled in the art such as, for example, Western blot, immunoprecipitation and ELISA (enzyme-linked immunosorbent
25 assay).

In one embodiment, the invention provides an antibody that specifically binds to the human BSEP polypeptide insert variant of SEQ ID NO:4, but does not bind to the human BSEP polypeptide of SEQ ID NO:2. In another embodiment, the invention provides for an antibody that specifically binds to the human BSEP
30 polypeptide insert variant of SEQ ID NO:4, but does not bind to the human BSEP polypeptide of SEQ ID NO:2 and does not bind to the human BSEP polypeptide insert variant of SEQ ID NO:8. In one embodiment, the invention provides an antibody that specifically binds to the human BSEP polypeptide insert variant of SEQ ID NO:8, but does not bind to the human BSEP polypeptide of SEQ ID NO:2. In
35 another embodiment, the invention provides an antibody that specifically binds to the human BSEP polypeptide insert variant of SEQ ID NO:8, but does not bind to the

human BSEP polypeptide of SEQ ID NO:2 and does not bind to the human BSEP polypeptide insert variant of SEQ ID NO:4.

The antibodies may be polyclonal or monoclonal. Monoclonal or polyclonal antibodies can be made using standard protocols (see, for example, *Antibodies: A Laboratory Manual*, ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A
5 subject such as a mammal, including but not limited to a mouse, a hamster, a rat, a goat, a rabbit or a human can be immunized with an immunogenic form of the human BSEP polypeptide variant. In one embodiment, the subject will be injected with a human BSEP polypeptide insert variant comprising the amino acid sequence set
10 forth in SEQ ID NO:4, or a fragment thereof. In another embodiment, the subject will be injected with a human BSEP polypeptide insert variant comprising the amino acid sequence set forth in SEQ ID NO:8, or a fragment thereof. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. See, for example, *Antibodies, a Laboratory*
15 *Manual*, ed. by Harlow and Lane, Cold Spring Harbor Press, 1988.

Following immunization of a subject with an antigenic preparation of a polypeptide, antisera can be obtained and, if desired, polyclonal antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized subject and fused by standard
20 somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, *Nature*, 256: 495-97 (1975)), the human B cell hybridoma technique (Kozbar et al., *Immunology Today*, 4: 72 (1983)), and the EBV-hybridoma technique to produce
25 human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96 (1985)). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a particular polypeptide and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

30 The antibodies may be of the IgM, IgA, IgE or IgG class or subclasses thereof. The above antibody fragment includes but is not limited to Fab, Fab', (Fab')₂, Fv and single chain antibodies.

This invention provides an isolated antibody light chain of the above antibody, or fragment or oligomer thereof. This invention also provides an isolated antibody
35 heavy chain of the above antibody, or fragment or oligomer thereof. This invention also provides one or more CDR regions of the above antibody. In one embodiment,

the antibody is derivatized. In another embodiment, the antibody is a human antibody. In one embodiment, antibody is humanized.

As used herein, "oligomer" means a complex of 2 or more subunits.

As used herein, "CDR" or "complementarity determining region" means a
5 highly variable sequence of amino acids in the variable domain of an antibody.

As used herein, a "derivatized" antibody is one that has been modified. Methods of derivatization include but are not limited to the addition of a fluorescent moiety, a radionucleotide, a toxin, an enzyme or an affinity ligand such as biotin.

As used herein, "humanized," describes antibodies wherein some, most or all
10 of the amino acids outside the CDR regions are replaced with corresponding amino acids derived from human immunoglobulin molecules. In one embodiment of the humanized forms of the antibodies, some, most or all of the amino acids outside the CDR regions have been replaced with amino acids from human immunoglobulin molecules but where some, most or all amino acids within one or more CDR regions
15 are unchanged. Small additions, deletions, insertions, substitutions or modifications of amino acids are permissible as long as they would not abrogate the ability of the antibody to bind a given antigen. Suitable human immunoglobulin molecules would include IgG1, IgG2, IgG3, IgG4, IgA, IgE and IgM molecules. A "humanized" antibody would retain a similar antigenic specificity as the original antibody. One skilled in the
20 art would know how to make the humanized antibodies of the subject invention. Various publications, several of which are hereby incorporated by reference into this application, also describe how to make humanized antibodies. For example, the methods described in United States Patent No. 4,816,567 comprise the production of chimeric antibodies having a variable region of one antibody and a constant region of
25 another antibody. See also Berger *et al.*, *American J. of the Medical Sciences* 324(1):14 -30, 2002.

United States Patent No. 5,225,539 describes another approach for the production of a humanized antibody. This patent describes the use of recombinant DNA technology to produce a humanized antibody wherein the CDRs of a variable
30 region of one immunoglobulin are replaced with the CDRs from an immunoglobulin with a different specificity such that the humanized antibody would recognize the desired target but would not be recognized in a significant way by the human subject's immune system. Specifically, site directed mutagenesis is used to graft the CDRs onto the framework.

35 Other approaches for humanizing an antibody are described in United States Patent Nos. 5,585,089 and 5,693,761, as well as in International Patent Publication

No. WO 90/07861, all of which describe methods for producing humanized immunoglobulins. These have one or more CDRs and possible additional amino acids from a donor immunoglobulin and a framework region from an accepting human immunoglobulin. These documents describe methods to increase the affinity
5 of an antibody for the desired antigen. Some amino acids in the framework are chosen to be the same as the amino acids at those positions in the donor rather than in the acceptor. Specifically, these documents describe the preparation of a humanized antibody that binds to a receptor by combining the CDRs of a mouse monoclonal antibody with human immunoglobulin framework and constant regions.
10 Human framework regions can be chosen to maximize homology with the mouse sequence. A computer model can be used to identify amino acids in the framework region which are likely to interact with the CDRs or the specific antigen and then mouse amino acids can be used at these positions to create the humanized antibody.

The above-cited documents also propose four possible criteria, which may be
15 used in designing the humanized antibodies. The first proposal was that for an acceptor, use a framework from a particular human immunoglobulin that is unusually homologous to the donor immunoglobulin to be humanized, or use a consensus framework from many human antibodies. The second proposal was that if an amino acid in the framework of the human immunoglobulin is unusual and the donor amino
20 acid at that position is typical for human sequences, then the donor amino acid rather than the acceptor may be selected. The third proposal was that in the positions immediately adjacent to the 3 CDRs in the humanized immunoglobulin chain, the donor amino acid rather than the acceptor amino acid may be selected. The fourth proposal was to use the donor amino acid residue at the framework positions at which
25 the amino acid is predicted to have a side chain atom within 3 angstrom units of the CDRs in a three dimensional model of the antibody and is predicted to be capable of interacting with the CDRs. The above methods are merely illustrative of some of the methods that one skilled in the art could employ to make humanized antibodies.

This invention provides an isolated nucleic acid molecule encoding the above
30 antibody. The nucleic acid molecule includes but is not limited to RNA, genomic DNA and cDNA.

This invention provides a labeled antibody. The antibody may be labeled with a detectable marker including, but not limited to: a radioactive marker, a colorimetric marker, a luminescent marker, an enzyme marker, a fluorescent marker, or gold.
35 Methods for attaching markers to antibodies are well known in the art.

Radioactive markers include, but are not limited to: ^3H , ^{14}C , ^{32}P , ^{33}P , ^{35}S , ^{36}Cl , ^{51}Cr , ^{57}Co , ^{59}Co , ^{59}Fe , ^{90}Y , ^{125}I , ^{131}I , and ^{186}Re .

Fluorescent markers include, but are not limited to, fluorescein, rhodamine and auramine, isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthalaldehyde, fluorescamine and Texas Red.

Colorimetric markers include, but are not limited to, biotin and digoxigenin.

Additionally, chemiluminescent compounds may be used as markers. Typical chemiluminescent compounds include, but are not limited to, luminol, isoluminol, aromatic acridinium esters, imidazoles, acridinium salts, and oxalate esters. Similarly, bioluminescent compounds may be utilized as markers, the bioluminescent compounds including luciferin, luciferase, and aequorin.

Enzyme markers include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase.

The antibody can also be labeled using fluorescence emitting metals such as ^{152}Eu , or other of the lanthanide series. These metals can be attached to the antibody using such metal groups as diethylenetriamine pentacetic acid (EDTA).

Once labeled, the antibody may be employed to identify and quantify the human BSEP polypeptide variants of the invention utilizing techniques well known to the art. See, for example, Immunology Methods Manual, ed. by Lefkovitis, Academic Press, 1996.

Methods of Detecting BSEP Polypeptide Variants

The invention also provides methods for detecting a BSEP polypeptide variant in a sample comprising: (a) contacting a suitable sample with a compound that specifically binds to the BSEP polypeptide variant; and (b) determining whether any compound is bound to the BSEP polypeptide variant, where the presence of any compound bound to the BSEP polypeptide variant detects the BSEP polypeptide variant in the sample. In one embodiment, any unbound compound is removed prior to determining whether any compound is bound to the BSEP polypeptide variant. In one embodiment, the BSEP polypeptide variant is an insert variant.

In one embodiment the compound is immobilized on a solid support. In one embodiment the solid support is a microtiter plate well. In another embodiment, the solid support is a membrane such as a cellulose membrane, a nitrocellulose

membrane, or a nylon membrane. In another embodiment, the solid support is a bead. In a further embodiment, the solid support is a surface plasmon resonance sensor chip. The surface plasmon resonance sensor chip can have pre-immobilized streptavidin. In one embodiment, the surface plasmon resonance sensor chip is a
5 BIAcore™ chip (Biacore, Piscataway, N.J.).

The term "suitable sample," as it refers to samples used for detecting polypeptides include, but is not limited to: cells, cell lysates, protein or membrane extracts of cells, body fluids, or tissue samples.

In one embodiment of the method described immediately above, the
10 compound capable of specifically binding the BSEP polypeptide variant is an antibody. In one embodiment the antibody specifically binds to the human BSEP polypeptide insert variant of SEQ ID NO:4. In another embodiment, the antibody specifically binds to the human BSEP polypeptide insert variant of SEQ ID NO:8.

In one embodiment the compound capable of specifically binding the BSEP
15 polypeptide variant is labeled with a detectable marker. In one embodiment, the detectable marker is a radioactive marker, a colorimetric marker, a luminescent marker, enzyme marker, or a fluorescent marker. When a compound is labeled with a detectable marker, the compound, and anything bound to the compound, can be detected by detecting the labeled marker. Methods for attaching markers to
20 compounds are well known in the art.

As used herein "determining whether any compound is bound to BSEP polypeptide variant" can be done by any of the methods known to those skilled in the art.

Methods such as surface plasmon resonance may be used to measure the
25 direct binding of the compound to the BSEP polypeptide variant using commercially available instruments, methods and reagents (Biacore, Piscataway, N.J.). Other methods that can be used to measure the binding of the compound to the BSEP polypeptide variant include, but are not limited to, Western blot, immunoprecipitation and ELISA.

30 When the compound capable of specifically binding the BSEP polypeptide variant is an antibody, "determining whether the compound is bound to BSEP polypeptide variant" can be accomplished using any of a variety of immunoassays, such as RIA (radioimmunoassay), ELISA, Western blot analysis and immunoprecipitation.

Methods of Detecting Nucleic Acids Encoding BSEP Polypeptide**Variants**

The invention also provides methods for detecting a nucleic acid which encodes a BSEP polypeptide variant comprising: (a) contacting a suitable sample
5 with a compound capable of specifically binding a nucleic acid encoding a BSEP polypeptide variant; and (b) determining whether any compound is bound to the nucleic acid, where the presence of compound bound to the nucleic acid in the sample. In one embodiment, any unbound compound is removed prior to determining whether the compound is removed prior to determining whether any
10 compound is bound to the nucleic acid. In one embodiment, the BSEP polypeptide variant is an insert variant.

In one embodiment the compound is immobilized on a solid support. In one embodiment the solid support is a microtiter plate well. In another embodiment, the solid support is a membrane such as a cellulose membrane, a nitrocellulose
15 membrane, or a nylon membrane. In another embodiment, the solid support is a bead. In a further embodiment, the solid support is a surface plasmon resonance sensor chip. The surface plasmon resonance sensor chip can have pre-immobilized streptavidin. In one embodiment, the surface plasmon resonance sensor chip is a BIAcore™ chip.

20 The term "suitable sample," as it refers to samples used for detecting nucleic acids includes, but is not limited to, cells, cell lysates, nucleic acids extracts of cells, tissue samples, or body fluids. Body fluids include, but are not limited to, blood, serum and saliva.

The term "specifically binding to a nucleic acid," as used herein, refers to a
25 binding reaction that is determinative of the presence of the nucleic acid in a population of nucleic acids or other material.

In one embodiment of the method described immediately above, the compound is a nucleic acid.

In one embodiment, the compound is a nucleic acid complementary to a
30 nucleic acid encoding SEQ ID NO:6 or a fragment thereof. In one embodiment, the compound is a nucleic acid complementary to SEQ ID NO:5 or a fragment thereof. In one embodiment, the compound is a nucleic acid complementary to a nucleic acid encoding SEQ ID NO:10 or a fragment thereof. In one embodiment, the compound is a nucleic acid complementary to SEQ ID NO:9 or a fragment thereof. In one
35 embodiment, the compound is a nucleic acid complementary to a nucleic acid

encoding SEQ ID NO:11 or a fragment thereof. In one embodiment, the compound is a nucleic acid complementary to a SEQ ID NO:12 or a fragment thereof.

In one embodiment the compound capable of specifically binding the nucleic acid encoding the BSEP polypeptide variant is labeled with a detectable marker. In one embodiment, the detectable marker is a radioactive label, a colorimetric marker, a luminescent marker, enzyme marker or a fluorescent marker. When a compound is labeled with a detectable marker, the compound, and anything bound to the compound, can be detected by detecting the labeled marker. Methods for attaching markers to compounds are well known in the art.

As used herein "determining whether any compound is bound to nucleic acid" can be done by any of the methods known to those skilled in the art. These methods include, but are not limited to, Southern blotting, Northern blotting and RNA protection assays.

Methods of Diagnosis

The present invention also provides methods of determining whether a subject is expressing a BSEP polypeptide variant comprising: (a) contacting a suitable sample from the subject with a compound that specifically binds to the BSEP polypeptide variant; and (b) determining whether any compound is bound to the BSEP polypeptide variant, wherein the presence of the compound bound to the BSEP polypeptide variant indicates that the individual is expressing the BSEP polypeptide variant. In one embodiment, any unbound compound is removed prior to determining whether any compound is bound to the BSEP polypeptide variant. In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide insert variant.

A "suitable sample" in connection to the above method of determining whether a subject is expressing a human BSEP polypeptide variant refers to any sample from the subject that could contain the BSEP polypeptide variant. Examples include, but are not limited to, body fluids and tissue samples. Examples of body fluids include, but are not limited to, blood, serum, urine and saliva.

In one embodiment the compound is an antibody that specifically binds to the BSEP polypeptide insert variant of SEQ ID NO:4. In another embodiment the compound is an antibody that specifically binds to the BSEP polypeptide insert variant of SEQ ID NO:8.

Methods of determining whether the compound is bound to the BSEP polypeptide variant are well known to a person having ordinary skill in the art. These methods include, but are not limited to, Western blot, immunoprecipitation and ELISA.

A "subject," as used herein means any animal or artificially modified animal expressing a BSEP polypeptide variant. The animals include but are not limited to mice, rats, dogs, guinea pigs, ferrets, rabbits, humans, and other primates.

Methods of Detecting Nucleic Acids Encoding BSEP Polypeptide

5 **Variants Using RT-PCR or PCR**

The invention also provides methods of detecting a nucleic acid which encodes a BSEP polypeptide variant comprising: (a) obtaining cDNA from mRNA obtained from a suitable sample; (b) amplifying the cDNA corresponding to the nucleic acid encoding the BSEP polypeptide or a portion of said nucleic acid; (c)
10 comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide, wherein a difference between the amplified cDNA and the DNA from a nucleic acid known to encode a BSEP polypeptide indicates the detection of a nucleic acid encoding a BSEP polypeptide variant in the sample.

The invention also provides methods for detecting a nucleic acid which
15 encodes a BSEP polypeptide variant comprising: (a) contacting a suitable sample with a compound capable of specifically binding a nucleic acid encoding a BSEP polypeptide variant; and (b) determining whether any compound is bound to the nucleic acid, where the presence of compound bound to the nucleic acid in the sample.

20 In one embodiment, the BSEP polypeptide variant is a human BSEP polypeptide variant. In one embodiment, the BSEP polypeptide variant is an insert variant.

As used herein, "mRNA" refers to messenger RNA. As used herein, "cDNA" refers to complementary DNA.

25 The term "suitable sample," as it refers to samples used for detecting nucleic acids includes, but is not limited to, cells, cell lysates, nucleic acids extracts of cells, tissue samples, or body fluids. Body fluids include, but are not limited to, blood, serum and saliva. In one embodiment, the suitable sample is obtained from a subject.

30 Methods of obtaining mRNA from a suitable sample are well known in the art. Further, methods of making cDNA from mRNA, such as reverse transcription, are also well known in the art.

As used herein, "amplifying" means increasing the numbers of copies of a specific DNA fragment. In one embodiment, the amplifying of the cDNA is carried out
35 using PCR (polymerase chain reaction).

In one embodiment, the amplifying of the cDNA is accomplished using primers flanking the entire reading frame of a gene encoding a BSEP polypeptide are used to amplify the cDNA. In another embodiment, the amplifying of the cDNA is accomplished out using primers flanking a portion of a nucleic acid encoding a BSEP polypeptide.

If the BSEP polypeptide variant is an insert variant, the amplifying of the cDNA may be accomplished using primers that flank the inserted sequence. In another embodiment, the primers used flank a portion of the cDNA which portion comprises the inserted sequence, but which is bigger than the inserted sequence. In one embodiment, the primers used flank the entire open reading frame a nucleic acid encoding the BSEP polypeptide variant.

In one embodiment, comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide is accomplished by comparing the sequence of the amplified cDNA to the sequence of a nucleic acid known to encode a BSEP polypeptide. The presence of additional nucleotides in the BSEP polypeptide variant will indicate that the BSEP polypeptide variant has an insertion.

In another embodiment, comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide is accomplished by comparing the size of the amplified cDNA to the size of the DNA of a gene known to encode a BSEP polypeptide. A difference in size will indicate that the amplified DNA encodes a BSEP polypeptide variant. For example, if the amplified DNA encodes more than 1321 amino acids, then the amplified DNA encodes a BSEP polypeptide insert variant.

Methods of Diagnosis using RT-PCR

The invention also provides methods of determining whether a subject is expressing a BSEP polypeptide insert variant comprising: (a) obtaining cDNA from mRNA obtained from a suitable sample from the subject; (b) amplifying the cDNA corresponding to the insertion of the BSEP polypeptide insert variant or a portion of said insertion; and (c) comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide, wherein the presence of an insertion in the amplified cDNA indicates that the subject is expressing the BSEP polypeptide insert variant.

A "suitable sample" in connection to the above method of determining whether a whether a subject is expressing a human BSEP polypeptide variant refers to any sample from the subject that could contain the BSEP polypeptide variant.

Examples include, but are not limited to, body fluids and tissue samples. Examples of body fluids include, but are not limited to, blood, serum, urine and saliva.

In one embodiment, the amplifying of the cDNA may be accomplished using primers that flank the inserted sequence. In another embodiment, the primers used
5 flank a bigger portion of the cDNA which portion comprises the inserted sequence. In one embodiment, the primers used flank the entire open reading frame of a nucleic acid encoding the BSEP polypeptide variant.

In one embodiment, comparing the amplified cDNA to the DNA of a nucleic acid known to encode a BSEP polypeptide is accomplished by comparing the
10 sequence of the amplified cDNA to the sequence of a nucleic acid known to encode a BSEP polypeptide. The presence of additional nucleotides in the BSEP polypeptide variant will indicate that the BSEP polypeptide variant has an insertion, and is a BSEP polypeptide insert variant.

In another embodiment, comparing the amplified cDNA to the DNA of a
15 nucleic acid known to encode a BSEP polypeptide is accomplished by comparing the size of the amplified cDNA to the size of the DNA of a gene known to encode a BSEP polypeptide. A difference in size will indicate that the amplified DNA encodes a BSEP polypeptide variant.

Methods of Identifying BSEP Substrates or Modulators

20 The present invention provides methods of identifying compounds that are substrates and/or modulators of the BSEP polypeptide variants of the invention. Compounds that are substrates and/or modulators of BSEP polypeptide variants could be potentially toxic to cells. Therefore, identifying compounds that are substrates and/or modulators of the BSEP polypeptide variants of the invention could
25 be useful in drug screening, to identify compounds that could be potentially toxic to a subject expressing a BSEP polypeptide variant.

The term "compound," as used herein, refers to can be any chemical or biological agent. Some examples of such test agents are synthetic chemicals, naturally occurring chemicals, proteins (e.g., antibodies), nucleotides (e.g., antisense
30 oligonucleotides, interference RNA oligonucleotides), etc.

The term "modulator," as used herein, refers to a compound that alters the function or activity of a BSEP polypeptide variant. The "modulator" can be an inhibitor, an activator, or an inducer a BSEP polypeptide variant, or a combination thereof. An "inhibitor" is a compound that can inhibit or decrease the transport
35 activity of the BSEP polypeptide variant. An "activator" is a compound that can increase the transport activity of the BSEP polypeptide variant. An "inducer" is a

compound that can increase the expression of the BSEP polypeptide variant, therefore increasing the transport activity of the BSEP polypeptide variant.

The term "substrate," as used herein with respect to a human BSEP polypeptide variant refers to molecule that is to be transported by a human BSEP polypeptide variant. Some examples of such transport activities are efflux from inside cells to outside cells, or uptake into inside-out membrane vesicles. The substrate of a BSEP polypeptide variant can be any compound capable of being transported by the BSEP polypeptide variant. In one embodiment, the substrate of a BSEP polypeptide variant is any bile acid or salt thereof, or conjugate thereof. In one embodiment, the substrate of a BSEP polypeptide variant is selected from the group consisting of: cholyltaurine, taurocholate, taurochenodeoxycholate, taurodeoxycholate, tauroursodeoxycholate, glycochenodeoxycholate, glycocholate, lithocholate, deoxycholate, chenodeoxycholate, cholyl lysyl fluorescein AM, cholyl lysyl fluorescein diacetate, and calcein-AM. In another embodiment the substrate of a BSEP polypeptide variant is a compound identified using the methods described herein.

Binding Assays

The present invention provides methods for determining whether a compound is a substrate or modulator of a BSEP polypeptide variant comprising: (a) contacting the compound with a BSEP polypeptide variant and (b) determining whether any compound binds to the BSEP polypeptide variant, wherein the presence of compound bound to the BSEP polypeptide variant indicates that the compound is a substrate or modulator of the BSEP polypeptide variant. In one embodiment, step (a) is carried out under conditions permitting binding of the compound to the BSEP polypeptide variant.

In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide insert variant. In one embodiment, the BSEP polypeptide variant is isolated or purified. In another embodiment, the BSEP polypeptide variant is in a cell that naturally contains the BSEP polypeptide variant. In another embodiment, the BSEP polypeptide variant is in a cell that has been genetically engineered to express a BSEP polypeptide variant. In another embodiment, the BSEP polypeptide variant is in a cell membrane preparation. In another embodiment, the BSEP polypeptide variant is in a cell membrane vesicle or an inside-out cell membrane vesicle.

In one embodiment, the BSEP polypeptide variant is immobilized on a solid support. In one embodiment of the method described herein, the solid support is a microtiter plate well. In another embodiment, the solid support is a membrane such

as a cellulose membrane, a nitrocellulose membrane, or a nylon membrane. In another embodiment, the solid support is a bead. In a further embodiment, the solid support is a surface plasmon resonance sensor chip. The surface plasmon resonance sensor chip can have pre-immobilized streptavidin. In one embodiment, the surface plasmon resonance sensor chip is a BIAcore™ chip.

As used herein, a "cell membrane preparation" refers to isolated or reconstituted cell membranes. In one embodiment, the cell membrane preparation is a cell membrane vesicle. In another embodiment, the cell membrane preparation is an inside-out cell membrane vesicle.

As used herein, a "cell membrane vesicle" refers to a closed membrane shell. As used here in, "inside-out cell membrane vesicle" refers to a closed membrane shell where the natural topography of the membrane is inverted. A person of ordinary skill in the art would know how to prepare cell membrane vesicles and inside-out cell membrane vesicles.

The amount of compound bound to the BSEP polypeptide variant can be measured using any method known in the art. In one embodiment, the amount of compound bound to the BSEP polypeptide variant can be determined by measuring the concentration of unbound or free compound, and comparing this concentration to the concentration of compound originally added. If binding occurs between the compound and the BSEP polypeptide variant, the concentration of unbound or free compound will be less than the concentration of compound originally added.

In one embodiment the compound is labeled with a detectable marker. The detectable marker can be, but is not limited to, a radioactive label, a colorimetric marker, a luminescent marker, enzymatic marker, a fluorescent marker or a marker that is capable of emitting electromagnetic energy.

In one embodiment, this method will be carried out using a high throughput assay. "High throughput assay" as used herein refers to an assay, which provides for multiple candidate agents or samples to be screened simultaneously. As further described below, examples of such assays may include the use of microtiter plates which are especially convenient because a large number of assays can be carried out simultaneously, using relatively small amounts of reagents and samples.

Competition Binding Assays

The present invention also provides s methods for determining whether a compound is a substrate or modulator of a BSEP polypeptide variant comprising: (a) contacting the BSEP polypeptide variant with a known binder and the compound; (b) measuring the amount of binder bound to the BSEP polypeptide variant; and (c)

comparing the amount of binder bound to the BSEP polypeptide variant in step (b) with the amount of binder bound to the BSEP polypeptide variant in the absence of the compound, and determining whether the presence of the compound increased or decreased the amount of binder bound to the BSEP polypeptide variant, wherein an increase or a decrease in the amount of binder bound to the BSEP polypeptide variants indicates that the compound is a substrate or modulator of BSEP. In one embodiment, step (a) is carried out under conditions permitting binding of the known binder to the BSEP polypeptide variant.

The invention also provides methods for determining whether a compound is a substrate or modulator of a BSEP polypeptide variant comprising: (a) contacting the BSEP polypeptide variant with the compound and a known binder of the BSEP polypeptide variant; (b) measuring the amount of compound bound to the BSEP polypeptide variant; and (c) comparing the amount of binding in step (b) with the amount of binding in the absence of the known binder, and determining whether the presence of the known binder increased or decreased the binding of the compound to the BSEP polypeptide variant, wherein an increase or a decrease in the amount of compound bound to the BSEP polypeptide variants indicates that the compound is a substrate or modulator of BSEP. In one embodiment, step (a) is carried out under conditions permitting binding of the known binder to the BSEP polypeptide variant.

In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide insert variant. In one embodiment, the BSEP polypeptide variant is isolated or purified. In another embodiment, the BSEP polypeptide variant is in a cell that naturally contains the BSEP polypeptide variant. In another embodiment, the BSEP polypeptide variant is in a cell that has been genetically engineered to express a BSEP polypeptide variant. In another embodiment, the BSEP polypeptide variant is in a cell membrane vesicle or an inside-out cell membrane vesicle. In another embodiment, the BSEP polypeptide variant is in a cell membrane preparation.

In one embodiment, the BSEP polypeptide variant is immobilized on a solid support. In one embodiment of the method described herein, the solid support is a microtiter plate well. In another embodiment, the solid support is a membrane such as a cellulose membrane, a nitrocellulose membrane, or a nylon membrane. In another embodiment, the solid support is a bead. In a further embodiment, the solid support is a surface plasmon resonance sensor chip. The surface plasmon resonance sensor chip can have pre-immobilized streptavidin. In one embodiment, the surface plasmon resonance sensor chip is a BIAcore™ chip.

As used herein, a "known binder," refers to a compound that is known to bind to the BSEP polypeptide variant. In one embodiment, the known binder is selected from the group consisting of: cholytaurine, taurocholate, taurochenodeoxycholate, taurodeoxycholate, tauroursodeoxycholate, glycochenodeoxycholate, glycocholate, lithocholate, deoxycholate, chenodeoxycholate, choly lysyl fluorescein AM, choly lysyl fluorescein diacetate, and calcein-AM. In another embodiment, the known binder is a binder identified in any of the methods described herein.

In one embodiment the compound and the known binder are introduced at the same time. In another embodiment, the known binder is introduced first and then the compound is added. In another embodiment, the compound is introduced first and then the known binder is added.

Determining whether a compound binds the BSEP polypeptide variant can be accomplished using any method known in the art including, but not limited to, equilibrium dialysis and rapid filtration.

The amount of compound or known binder bound to the BSEP polypeptide variant can be measured using any method known in the art. In one embodiment, the amount of compound or known binder bound to the BSEP polypeptide variant can be determined by measuring the concentration of unbound or free compound or known binder, and comparing this concentration to the concentration of compound or known binder originally added. If binding occurs between the compound or known binder and the BSEP polypeptide variant, the concentration of unbound or free compound or known binder will be less than the concentration of compound or known binder originally added.

In one embodiment, the compound is labeled with a detectable marker. In another embodiment, the known binder is labeled with a detectable marker. The detectable marker can be, but is not limited to, a radioactive label, a colorimetric marker, a luminescent marker, enzymatic marker, a fluorescent marker or a marker that is capable of emitting electromagnetic energy.

In one embodiment, invention methods may be carried out using high throughput assays. A high throughput assay for measuring the unbound or free concentration of a compound has been disclosed, for example, in European Patent EP1088589 (Banker MJ *et al.*). However, those skilled in the art are knowledgeable about many other high throughput assays that might be applied to the methods of the present invention.

ATPase Assays

Where a compound is a substrate of the BSEP polypeptide variant, ATP will be hydrolyzed. The energy derived from this reaction is used for the transport activity of the BSEP polypeptide variant. Therefore, a measure of the ATPase activity can be used to identify substrates and/or modulators of the BSEP polypeptide variants.

5 The invention also provides methods for determining whether a compound is a substrate of a BSEP polypeptide variant comprising: (a) contacting the compound with the BSEP polypeptide variant and ATP; and (b) detecting the ATPase activity of the BSEP polypeptide insert variant, wherein the detection of ATPase activity indicates that the compound is a substrate of the BSEP polypeptide variant.

10 The invention also provides methods for determining whether a compound is an modulator of a BSEP polypeptide variant comprising: (a) contacting the BSEP polypeptide variant with a substrate of the BSEP polypeptide variant, a compound and ATP; (b) measuring the ATPase activity of the BSEP polypeptide variant; and (c)
15 comparing the ATPase activity of the BSEP polypeptide variant in step (b) with the ATPase activity of the BSEP polypeptide variant in the absence of the compound, and determining whether the compound modulated the ATPase activity of the BSEP polypeptide variant, wherein an increase or a decrease in the ATPase activity indicates that the compound is a substrate or modulator of the BSEP polypeptide variant.

20 Where the compound is an activator or inducer of the BSEP polypeptide variant, the ATPase activity will increase in the presence of the compound. Where the compound is an inhibitor of the BSEP polypeptide variant, the ATPase activity will decrease in the presence of the compound.

 In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide
25 insert variant. In another embodiment, the BSEP polypeptide variant is in a cell that naturally contains the BSEP polypeptide variant. In another embodiment, the BSEP polypeptide variant is in a cell membrane vesicle or an inside-out cell membrane vesicle. In another embodiment, the BSEP polypeptide variant is in a cell membrane preparation. In one embodiment, the BSEP polypeptide variant is isolated. In one
30 embodiment, the BSEP polypeptide variant is purified.

 The term "ATPase activity," as used herein, refers to the ability of the polypeptide to hydrolyze ATP. The hydrolysis of ATP can be measured by either a decrease in chemical reactants (ATP), or an increase in chemical products (ADP and phosphate). Methods of detecting or measuring ATPase activity are well known in
35 the art. An ATPase assay for a human transporter is disclosed in Shimabuku *et al.*,

J. Biol. Chem. 267(7):4308 -11 (1992). A higher throughput ATPase assay was published in Szabo *et al.*, *J. Biol. Chem.* 273(17):10132 -38 (1998).

Transport Assays

5 The present invention also provides methods for determining whether a compound is a substrate of BSEP comprising: (a) contacting the compound with the BSEP polypeptide variant; (b) determining whether the compound is transported by the BSEP polypeptide variant, wherein the transport of the compound by the BSEP polypeptide variant indicates that the compound is a substrate of the BSEP polypeptide variant.

10 Determining whether the compound is transported by the BSEP polypeptide variant can be done using a variety of methods known to those skilled in the art. See, for example, Byrne, J.A., *et al.*, *Gastroenterology* 123(5):1649 - 58 (2002); Gerloff, T., *et al.*, *European Journal of Biochemistry* 269(14):3495 -503 (2002); Bode, K.A., *et al.*, *Biochemical Pharmacology* 64(1):151-58 (2002); Noe, J., *et al.*,
15 *Hepatology* 33(5):1223-31 (2001); Kullak-Ublick, G.A., *et al.*, *Seminars in Liver Disease* 20(3):273-92 (2000); Stieger, B., *et al.*, *Gastroenterology* 118(2):422-30 (2000); Ballatori, N., *et al.*, *American Journal of Physiology - Gastrointestinal & Liver Physiology* 278(1):G57-63 (2000).

In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide
20 insert variant. In another embodiment, the BSEP polypeptide variant is in a cell that has been genetically engineered to express a BSEP polypeptide variant. In one embodiment, the cells that contain BSEP polypeptide variants can be grown in culture, either in suspension, or attached to a surface (e.g., plastics or glass or porous surface). In another embodiment, the cells are grown on a porous surface
25 such that a compound can be added to the basal side of the cells and efflux of the compound can be directed to the apical surface where the BSEP polypeptide variants are located.

In another embodiment, the BSEP polypeptide variant is in a cell membrane preparation. In another embodiment, the BSEP polypeptide variant is in a cell
30 membrane vesicle or an inside-out cell membrane vesicle.

In another embodiment, the BSEP polypeptide variant is isolated. In another embodiment, the BSEP polypeptide variant is purified.

The "transport" or "transport activity" of a BSEP polypeptide or a BSEP polypeptide variant refers to the ability of a BSEP polypeptide to transport a
35 compound from the inside a cell to the outside a cell or cell membrane vesicle, or

from the outside of an "inside-out" cell membrane vesicle to the inside of a cell membrane vesicle.

In one embodiment, the compound is labeled with a detectable marker. The detectable marker can be, but is not limited to, a radioactive label, a colorimetric marker, a luminescent marker, enzymatic marker, a fluorescent marker or a marker
5 that is capable of emitting electromagnetic energy.

In one embodiment, the method will be carried out using high throughput assay.

In one embodiment, the transport activity can be detected by measuring
10 either the intracellular concentration of compound over time, or the extracellular concentration of compound over time, or both. If the compound is a substrate of the BSEP polypeptide variant and the assay is carried out in a cell, the intracellular concentration of the compound will decrease over time and the extracellular concentration of the compound will increase over time. If the compound is a
15 substrate of the BSEP polypeptide variant and the assay is carried out in an inside-out membrane vesicle, the compound will accumulate inside the inside-out cell membrane vesicle over time, and the concentration of the compound outside of the inside-out cell membrane vesicle will decrease over time.

20

Competition Transport Assays

The invention also relates to methods for determining whether a compound is a substrate or modulator of a human BSEP polypeptide variant comprising: (a) contacting the compound with the BSEP polypeptide variant and a known substrate
25 of the BSEP polypeptide variant; (b) measuring the activity of the BSEP polypeptide variant with respect to the known substrate; and (c) comparing the activity of the BSEP polypeptide variant with respect to the known substrate in step (b) with the activity of the BSEP polypeptide with respect to the known substrate in the absence of the compound, wherein an increase or a decrease in transport of activity of the
30 BSEP polypeptide insert variant with respect to the known substrate indicates that the compound is a substrate or modulator of the BSEP polypeptide insert variant. In one embodiment, the BSEP polypeptide variant is a BSEP polypeptide insert variant.

In one embodiment, the BSEP polypeptide variant is in a cell that naturally contains the BSEP polypeptide variant. In another embodiment, the BSEP
35 polypeptide variant is in a cell that has been genetically engineered to express a BSEP polypeptide variant. In one embodiment, the cells that contain BSEP

polypeptide variants can be grown in culture, either in suspension, or attached to a surface (e.g., plastics or glass or porous surface). In another embodiment, the cells are grown on a porous surface such that a compound can be added to the basal side of the cells and efflux of the compound can be directed to the apical surface where
5 the BSEP polypeptide variants are located.

In another embodiment, the BSEP polypeptide variant is in a cell membrane vesicle, or an inside-out cell membrane vesicle.

The "transport activity" of a BSEP polypeptide or a BSEP polypeptide variant refers to the ability of a BSEP polypeptide to transport a compound from the inside a
10 cell to the outside a cell or cell membrane vesicle, or from the outside of an "inside-out" cell membrane vesicle to the inside of a cell membrane vesicle.

In one embodiment the substrate is labeled with a detectable marker. The detectable marker can include, but is not limited to, a radioactive label, a colorimetric marker, a luminescent marker, enzymatic marker, a fluorescent marker or a marker
15 that is capable of emitting electromagnetic energy.

The known substrate can be any substrate of the BSEP polypeptide variant. In one embodiment, the known substrate is selected from the group consisting of: cholyllaurine, taurocholate, taurochenodeoxycholate, taurodeoxycholate, tauroursodeoxycholate, glycochenodeoxycholate, glycocholate, lithocholate,
20 deoxycholate, chenodeoxycholate, choyl lysyl fluorescein AM, choyl lysyl fluorescein diacetate, and calcein-AM. In another embodiment, the known substrate is a substrate identified in any of the methods described herein.

In one embodiment the known substrate can be added prior to the compound. In another embodiment, the known substrate is added simultaneously with the
25 compound. In another embodiment, the known substrate is added after the compound.

In one embodiment, the method will be carried out using a high throughput assay.

In one embodiment, the transport activity can be detected by measuring
30 either the intracellular concentration of compound over time, or the extracellular concentration of compound over time, or both.

If the compound is an inhibitor of the BSEP polypeptide variant, the transport activity of the BSEP polypeptide variant will decrease. If the compound is an activator or inducer of the BSEP polypeptide variant, the transport activity of the
35 BSEP polypeptide variant will increase. By comparing the transport activity of the

BSEP polypeptide variant in the presence and in the absence of the compound, the degree of inhibition, activation or inducement by the compound can be determined.

A method of determining the transport activity of a BSEP polypeptide, wherein the BSEP polypeptide is expressed in a host cell is described in Lecureur *et al.*,
5 *Molecular Pharmacology*, 57:24-35 (2000).

A method of determining the transport activity of a BSEP polypeptide, wherein the BSEP polypeptide is on a cell membrane vesicle, is described in Stieger *et al.*,
Gastroenterology, 118:422-30, (2000).

All of the methods disclosed herein to identify activators are similarly
10 expected to also identify inducers. However, as presently understood, the induction process usually takes a much longer time than the activation process. Therefore, to test whether a compound is an inducer of the polypeptide variant, the compound is preferably introduced to the assay system hours or even days before the known substrate is added. In addition, there are additional assay formats that can be used
15 to unambiguously identify a compound as an inducer. By definition, an inducer will increase messenger RNA and protein levels of BSEP polypeptide variant. Therefore, any of the diagnostic methods to measure the level of messenger RNA, such as quantitative PCR, Invader® assay (Third Wave Technologies, Madison, WI), Northern blots and RNase protection assays, and protein levels of the BSEP
20 polypeptide variants, such as Western blots, ELISA assays, cytoimmunofluorescence assays, can be used to identify whether a test agent is an inducer of the BSEP polypeptide variants.

All of the method of identifying compounds that are substrates and/or modulators of the BSEP polypeptide variants of the invention can be carried out in
25 24-well or higher density microarrays or microwells to enable higher assay throughput. In an embodiment where the compound or substrate is labeled with a fluorescent marker, the binding and the transport process of such molecules can be detected and measured automatically by quantitative fluorometry (e.g., automated fluorescence plate readers, automated and quantitative fluorescence microscope,
30 etc.). Examples of automated fluorescence plate readers that are capable of performing kinetic measurements are: Victor V from Perkin Elmer, FLIPR HTS reader from Molecular Devices, et al. Examples of the automated and quantitative fluorescence microscopes that are capable of performing kinetic measurements are: ArrayScan® Kinetic Scan™ (Cellomics, Pittsburgh, PA), Pathway HT™ system (Atto
35 Bioscience, Rockville, MD), and the like. The combination of the disclosed assay methods and modern day automated equipments is expected to deliver

unprecedented capacity and throughput to identify substrates, inhibitors, activators, and inducers of human BSEP polypeptide variants.

In one embodiment, any of the methods of identifying substrates and/or modulators of a BSEP polypeptide variant described herein can be used to characterize the transport function of the claimed BSEP polypeptide insert variants as compared to previously known human BSEP polypeptides. An increase or decrease in the transport activity of the BSEP polypeptide variants may be used to explain decreased or increased bile salt efflux in humans having the BSEP polypeptide variants.

10 **Methods of Inhibiting the Expression of a Human BSEP Polypeptide Variant**

The invention also provides methods of inhibiting the expression of a BSEP polypeptide variant in a cell comprising contacting the cell with a compound capable of inhibiting the expression of the BSEP polypeptide variant. In one embodiment the BSEP polypeptide variant is a human BSEP polypeptide insert variant.

In one embodiment the method will be carried out in vitro. In one embodiment the method will be carried out in vivo. These methods could be used in research, diagnosis and treatment.

In research, these methods could be used, for example, to elucidate the function of a BSEP polypeptide variant of the invention with respect to a human BSEP polypeptide or another BSEP polypeptide variant.

In one embodiment, the compound capable of inhibiting the expression of a BSEP polypeptide variant is an antisense oligonucleotide capable of specifically hybridizing to the BSEP polypeptide variant.

25 It is understood in the art that the sequence of an antisense compound need not be 100% complementary to that of its target nucleic acid to specifically hybridize to that sequence. An antisense compound specifically hybridizes to a target DNA or RNA sequence when binding of the compound to the target DNA or RNA sequence interferes with the normal function of the target DNA or RNA. This interference should cause a loss of utility, and there should be a sufficient degree of complementarity to avoid non-specific binding of the antisense compound to non-target sequences under conditions in which specific binding is desired, i.e., under physiological conditions in the case of in vivo assays or therapeutic treatment, and in case of in vitro assays, under conditions in which the assays are performed.

35 An "oligonucleotide," as used herein, refers to an oligomer or polymer of a ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof. This

term includes oligonucleotides composed of naturally occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as oligonucleotides having non-naturally-occurring portions which function similarly. Such modified or substituted oligonucleotides are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target and increased stability in the presence of nucleases.

The sequence of an antisense oligonucleotide capable of specifically hybridizing to a BSEP polypeptide variant can be identified through routine experimentation. In one embodiment the antisense oligonucleotide is capable of specifically hybridizing to a nucleic acid sequence from the group consisting of: (a) a nucleic acid encoding SEQ ID NO:6 or a fragment thereof; (b) SEQ ID NO:5 or a fragment thereof; (c) a nucleic acid encoding SEQ ID NO:10 or a fragment thereof; (d) SEQ ID NO:9 or a fragment thereof; (e) a nucleic acid encoding SEQ ID NO:11 or a fragment thereof; and (f) SEQ ID NO:12 or a fragment thereof.

In another embodiment, the compound capable of inhibiting the expression of a BSEP polypeptide variant is an RNAi construct. In one embodiment the RNAi construct is capable of specifically hybridizing to a nucleic acid sequence selected from the group consisting of: (a) a nucleic acid encoding SEQ ID NO:6 or a fragment thereof; (b) SEQ ID NO:5 or a fragment thereof; (c) a nucleic acid encoding SEQ ID NO:10 or a fragment thereof; (d) SEQ ID NO:9 or a fragment thereof; (e) a nucleic acid encoding SEQ ID NO:11 or a fragment thereof; and (f) SEQ ID NO:12 or a fragment thereof.

The antisense oligonucleotides and RNAi constructs can be used to specifically inhibit the expression of the claimed human BSEP polypeptide variants without inhibiting the human BSEP polypeptide. Using this technology, the specific function of each claimed human BSEP polypeptide variant can be studied. Further, antisense oligonucleotides and RNAi constructs may be used for disease treatment.

Antisense oligonucleotides are relatively short nucleic acids that are complementary (or antisense) to the coding strand (sense strand) of the mRNA encoding a particular protein. Although antisense oligonucleotides are typically RNA based, they can also be DNA based. Additionally, antisense oligonucleotides are often modified to increase their stability. See, for example, *Antisense Technology in Methods in Enzymology*, Vols. 313-314, ed. by Phillips, Abelson and Simon, Academic Press, 1999.

Without being bound by any particular theory, the binding of these relatively short oligonucleotides to the mRNA is believed to induce stretches of double

stranded RNA that trigger degradation of the messages by endogenous RNAses. Additionally, sometimes the oligonucleotides are specifically designed to bind near the promoter of the message, and under these circumstances, the antisense oligonucleotides may additionally interfere with translation of the message.

- 5 Regardless of the specific mechanism by which antisense oligonucleotides function, their administration to a cell or tissue allows the inactivation (at least partially) of the mRNA encoding a specific protein. Accordingly, antisense oligonucleotides decrease the translation of a particular protein.

- The oligonucleotides can be DNA or RNA, or chimeric mixtures or derivatives
10 or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve its stability, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors), or agents facilitating transport across the cell
15 membrane (see, e.g., Letsinger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-56 (1989); Lemaitre *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:648-52 (1987); International Patent Publication No. W0 88/09810) or the blood- brain barrier (see, e.g., International Patent Publication No. W0 89/10134), hybridization-triggered cleavage agents (see, e.g., Krol *et al.*, *BioTechniques* 6:958- 76 (1988)) or intercalating
20 agents. (see, e.g., Zon, *Pharm. Res.* 5:539-49 (1988)). To this end, the oligonucleotide may be conjugated to another molecule.

- The antisense oligonucleotide may comprise at least one modified base moiety which may be selected from the group including, but not limited to, 5-fluorouracil, 5- bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-
25 acetylcytosine, 5- (carboxyhydroxytriethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5- carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6- isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-
30 methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6- isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 5-
35 methyl-2-thiouracil, 3-(3-amino-3- N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

5 The antisense oligonucleotide can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, *e.g.*, in Perry-O'Keefe *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:14670 (1996) and in Eglom *et al.* *Nature* 365:566 (1993). One advantage of PNA oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another
10 embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

15 In yet a further embodiment, the antisense oligonucleotide is an alpha-anomeric oligonucleotide. An alpha-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual antiparallel orient, the strands run parallel to each other (Gautier *et al.*, *Nucl. Acids Res.* 15:6625-41 (1987)). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue *et al.*,
20 *Nucl. Acids Res.* 15:6131-48 (1987)), or a chimeric RNA-DNA analogue (Inoue *et al.*, *FEBS Lett.* 215:327-30 (1987)).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch Technologies, Inc. (Novato, CA), Applied
25 Biosystems (Foster City, CA), and others). As¹ examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein *et al.* (*Nucl. Acids Res.* 16:3209 (1988)), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-51 (1988)).

30 The selection of an appropriate oligonucleotide can be readily performed by one of skill in the art, based upon the present description. Given the nucleic acid encoding a particular protein, one of skill in the art can design antisense oligonucleotides that bind to that protein, and test these oligonucleotides in an in vitro or in vivo system to confirm that they bind to and mediate the degradation of the
35 mRNA encoding the particular protein. To design an antisense oligonucleotide that specifically binds to and mediates the degradation of a particular protein, it is

important that the sequence recognized by the oligonucleotide is unique or substantially unique to that particular protein. For example, sequences that are frequently repeated across proteins may not be an ideal choice for the design of an oligonucleotide that specifically recognizes and degrades a particular message. One
5 of skill in the art can design an oligonucleotide, and compare the sequence of that oligonucleotide to nucleic acid sequences that are deposited in publicly available databases to confirm that the sequence is specific, or substantially specific, for a particular protein.

In another example, it may be desirable to design an antisense
10 oligonucleotide that binds to and mediates the degradation of more than one message. In one example, the messages may encode related proteins such as isoforms or functionally redundant proteins. In such a case, one of skill in the art can align the nucleic acid sequences that encode these related proteins, and design an oligonucleotide that recognizes both messages.

15 A number of methods have been developed for delivering antisense DNA or RNA to cells, e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically to a subject. See, for
20 example, *Antisense Technology in Methods in Enzymology*, Vols. 313-314, ed. by Phillips, Abelson and Simon, Academic Press, 1999.

However, in some cases, it may be difficult to achieve intracellular concentrations of the antisense molecule sufficient to suppress translation on endogenous mRNAs in certain instances. Therefore another approach utilizes a
25 recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. For example, a vector can be introduced in vivo such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA.
30 Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, and preferably human, cells. Such promoters can be inducible or constitutive. Such
35 promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, *Nature* 290:304-10 (1981)), the promoter contained in the 3' long

terminal repeat of Rous sarcoma virus (Yamamoto *et al.*, *Cell* 22:787-97 (1980)), the herpes thymidine kinase promoter (Wagner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-45 (1981)), the regulatory sequences of the metallothionein gene (Brinster *et al.*, *Nature* 296:39-42 (1982), and the like. Any type of plasmid, cosmid, YAC or
5 viral vector can be used to prepare the recombinant DNA construct that can be introduced directly into the cells or tissue. Alternatively, viral vectors can be used which selectively infect the desired cells or tissue, in which case administration may be accomplished by another route (e.g., systemically).

RNAi constructs comprise double stranded RNA that can specifically block
10 expression of a target gene. "RNA interference" or "RNAi" is a term initially applied to a phenomenon observed in plants and worms where double-stranded RNA (dsRNA) blocks gene expression in a specific and post-transcriptional manner. Without being bound by any particular theory, RNAi appears to involve mRNA degradation; however, the biochemical mechanisms remain an active area of research.

15 As used herein, the term "dsRNA" refers to siRNA molecules, or other RNA molecules including a double stranded feature and able to be processed to siRNA in cells, such as hairpin RNA moieties.

As used herein, the phrase "mediates RNAi" refers to (indicates) the ability to distinguish, which RNAs are to be degraded by the RNAi process, e.g., degradation
20 occurs in a sequence-specific manner rather than by a sequence-independent dsRNA response, e.g., a PKR response.

As used herein, the term "RNAi construct" is a generic term used throughout the specification to include small interfering RNAs (siRNAs), hairpin RNAs, and other RNA species, which can be cleaved *in vivo* to form siRNAs. RNAi constructs herein
25 also include expression vectors (also referred to as RNAi expression vectors) capable of giving rise to transcripts which form dsRNAs or hairpin RNAs in cells, and/or transcripts, which can produce siRNAs *in vivo*.

"RNAi expression vector" (also referred to herein as a "dsRNA-encoding plasmid") refers to a replicable nucleic acid constructs used to express (transcribe)
30 RNA, which produces siRNA moieties in the cell in which the construct is expressed. Such vectors include a transcriptional unit comprising an assembly of (1) genetic element(s) having a regulatory role in gene expression, for example, promoters, operators, or enhancers, operatively linked to (2) a "coding" sequence which is transcribed to produce a double-stranded RNA (two RNA moieties that anneal in the
35 cell to form an siRNA, or a single hairpin RNA which can be processed to an siRNA), and (3) appropriate transcription initiation and termination sequences. The choice of

promoter and other regulatory elements generally varies according to the intended host cell. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to circular double stranded DNA loops, which, in their vector form are not bound to the chromosome. In the present
5 specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

The RNAi constructs contain a nucleotide sequence that hybridizes under
10 physiologic conditions of the cell to the nucleotide sequence of at least a portion of the mRNA transcript for the gene to be inhibited (i.e., the "target" gene). The double-stranded RNA need only be sufficiently similar to natural RNA that it has the ability to mediate RNAi. Thus, the invention has the advantage of being able to tolerate sequence variations that might be expected due to genetic mutation, strain
15 polymorphism or evolutionary divergence. The number of tolerated nucleotide mismatches between the target sequence and the RNAi construct sequence is no more than 1 in 5 basepairs, or 1 in 10 basepairs, or 1 in 20 basepairs, or 1 in 50 basepairs. Mismatches in the center of the siRNA duplex are most critical and may essentially abolish cleavage of the target RNA. In contrast, nucleotides at the 3' end
20 of the siRNA strand that is complementary to the target RNA do not significantly contribute to specificity of the target recognition.

The sequence identity between the RNAi construct and a target sequence may be optimized by sequence comparison and alignment algorithms known in the art (see, Gribskov and Devereux, *Sequence Analysis Primer*, Stockton Press, 1991,
25 and references cited therein) and by calculating the percent difference between the nucleotide sequences by, for example, the Smith-Waterman algorithm as implemented in the BESTFIT software program using default parameters (e.g., University of Wisconsin Genetic Computing Group). Greater than 90% sequence identity, or even 100% sequence identity, between the inhibitory RNA and the portion
30 of the target gene is preferred. Alternatively, the duplex region of the RNA may be defined functionally as a nucleotide sequence that is capable of hybridizing with a portion of the target gene transcript (e.g., using hybridization conditions such as 400 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 50 °C or 70 °C for 12-16 hours; followed by washing).

35 Production of RNAi constructs can be carried out by chemical synthetic methods or by recombinant nucleic acid techniques. Endogenous RNA polymerase

of the treated cell may mediate transcription in vivo, or cloned RNA polymerase can be used for transcription in vitro. The RNAi constructs may include modifications to either the phosphate-sugar backbone or the nucleoside, e.g., to reduce susceptibility to cellular nucleases, improve bioavailability, improve formulation characteristics, and/or change other pharmacokinetic properties. For example, the phosphodiester linkages of natural RNA may be modified to include at least one of nitrogen or sulfur heteroatom. Modifications in RNA structure may be tailored to allow specific genetic inhibition while avoiding a general response to dsRNA. Likewise, bases may be modified to block the activity of adenosine deaminase. The RNAi construct may be produced enzymatically or by partial/total organic synthesis, any modified ribonucleotide can be introduced by in vitro enzymatic or organic synthesis.

Methods of chemically modifying RNA molecules can be adapted for modifying RNAi constructs (see, e.g., Heidenreich *et al.*, *Nucleic Acids Res.* 25:776-80 (1997); Wilson *et al.*, *J. Mol. Recog.* 7:89-98 (1994); Chen *et al.*, *Nucleic Acids Res.* 23:2661-68 (1995); Hirschbein *et al.*, *Antisense Nucleic Acid Drug Dev.* 7:55-61 (1997)). Merely to illustrate, the backbone of an RNAi construct can be modified with phosphorothioates, phosphoramidate, phosphodithioates, chimeric methylphosphonate-phosphodiester, peptide nucleic acids, 5-propynyl-pyrimidine containing oligomers or sugar modifications (e.g., 2'-substituted ribonucleosides, a-configuration).

The double-stranded structure may be formed by a single self-complementary RNA strand or two complementary RNA strands. RNA duplex formation may be initiated either inside or outside the cell. The RNA may be introduced in an amount, which allows delivery of at least one copy per cell. Higher doses (e.g., at least 5, 10, 100, 500 or 1000 copies per cell) of double-stranded material may yield more effective inhibition, while lower doses may also be useful for specific applications. Inhibition is sequence-specific in that nucleotide sequences corresponding to the duplex region of the RNA are targeted for genetic inhibition.

In certain embodiments, the subject RNAi constructs is "small interfering RNAs" or "siRNAs." These nucleic acids may be around 19-30 nucleotides in length, and even more preferably 21-23 nucleotides in length, e.g., corresponding in length to the fragments generated by nuclease "dicing" of longer double-stranded RNAs. The siRNAs are understood to recruit nuclease complexes and guide the complexes to the target mRNA by pairing to the specific sequences. As a result, the target mRNA is degraded by the nucleases in the protein complex. In a particular

embodiment, the 21-23 nucleotide-long siRNA molecules comprise a 3' hydroxyl group.

5 The siRNA molecules of the present invention can be obtained using a number of techniques known to those of skill in the art. For example, the siRNA can be chemically synthesized or recombinantly produced using methods known in the art. For example, short sense and antisense RNA oligomers can be synthesized and annealed to form double-stranded RNA structures with 2-nucleotide overhangs at each end (Caplen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 98:9742-47 (2001); Elbashir *et al.*, *EMBO J.*, 20:6877-88 (2001)). These double-stranded siRNA structures can then
10 be directly introduced to cells, either by passive uptake or a delivery system of choice, such as described below.

In certain embodiments, the siRNA constructs can be generated by processing of longer double-stranded RNAs, for example, in the presence of the enzyme dicer. In one embodiment, the *Drosophila* in vitro system is used. In this
15 embodiment, dsRNA is combined with a soluble extract derived from *Drosophila* embryo, thereby producing a combination. The combination is maintained under conditions in which the dsRNA is processed to RNA molecules of about 21 to about 23 nucleotides.

The siRNA molecules can be purified using a number of techniques known to
20 those of skill in the art. For example, gel electrophoresis can be used to purify siRNAs. Alternatively, non-denaturing methods, such as non-denaturing column chromatography, can be used to purify the siRNA. In addition, chromatography (*e.g.*, size exclusion chromatography), glycerol gradient centrifugation, affinity purification with antibody can be used to purify siRNAs.

25 In certain preferred embodiments, at least one strand of the siRNA molecules has a 3' overhang from about 1 to about 6 nucleotides in length, preferably from 2 to 4 nucleotides in length. More preferably, the 3' overhangs are 1-3 nucleotides in length. In certain embodiments, one strand has a 3' overhang and the other strand is blunt-ended or also has an overhang. The length of the overhangs may be the same
30 or different for each strand. In order to further enhance the stability of the siRNA, the 3' overhangs can be stabilized against degradation. In one embodiment, the RNA is stabilized by including purine nucleotides, such as adenosine or guanosine nucleotides. Alternatively, substitution of pyrimidine nucleotides by modified analogues, *e.g.*, substitution of uridine nucleotide 3' overhangs by 2'-deoxythymidine
35 is tolerated and does not affect the efficiency of RNAi. The absence of a 2' hydroxyl

significantly enhances the nuclease resistance of the overhang in tissue culture medium and may be beneficial in vivo.

In other embodiments, the RNAi construct is in the form of a long double-stranded RNA. In certain embodiments, the RNAi construct is at least 25, 50, 100, 200, 300 or 400 bases. In certain embodiments, the RNAi construct is 400-800 bases in length. The double-stranded RNAs are digested intracellularly, *e.g.*, to produce siRNA sequences in the cell. However, in some embodiments, the uses of local delivery systems and/or agents, which reduce the effects of interferon or PKR, are preferred.

In certain embodiments, the RNAi construct is in the form of a hairpin structure (named as hairpin RNA). The hairpin RNAs can be synthesized exogenously or can be formed by transcribing from RNA polymerase III or RNA polymerase II promoters in vivo. Examples of making and using such hairpin RNAs for gene silencing in mammalian cells are described in, for example, Paddison *et al.*, *Genes Dev*, 16:948-58 (2002); McCaffrey *et al.*, *Nature*, 418:38-39 (2002); McManus *et al.*, *RNA* 8:842-50 (2002); Yu *et al.*, *Proc Natl Acad Sci U.S.A.*, 99:6047-52 (2002)). Preferably, such hairpin RNAs are engineered in cells or in an animal to ensure continuous and stable suppression of a desired gene. It is known in the art that siRNAs can be produced by processing a hairpin RNA in the cell. In yet other embodiments, a plasmid is used to deliver the double-stranded RNA, *e.g.*, as a transcriptional product. In such embodiments, the plasmid is designed to include a "coding sequence" for each of the sense and antisense strands of the RNAi construct. The coding sequences can be the same sequence, *e.g.*, flanked by inverted promoters, or can be two separate sequences each under transcriptional control of separate promoters. After the coding sequence is transcribed, the complementary RNA transcripts base-pair to form the double-stranded RNA.

International Patent Publication No. WO 01/77350 describes an exemplary vector for bi-directional transcription of a transgene to yield both sense and antisense RNA transcripts of the same transgene in a eukaryotic cell. Accordingly, in certain embodiments, the present invention provides a recombinant vector having the following unique characteristics: it comprises a viral replicon having two overlapping transcription units arranged in an opposing orientation and flanking a transgene for an RNAi construct of interest, wherein the two overlapping transcription units yield both sense and antisense RNA transcripts from the same transgene fragment in a host cell.

Exemplary RNAi constructs that specifically recognize a particular gene, or a particular family of genes can be selected using methodology outlined in detail herein with respect to the selection of antisense oligonucleotide. Similarly, methods of delivery RNAi constructs include the methods for delivery antisense oligonucleotides
5 outlined in detail herein.

Transgenic animals

The invention also provides transgenic non-human animals, e.g. mice, rats, rabbits, goats, sheep, dogs, cats, cows, or non-human primates, comprising, and preferably expressing, a transgene encoding a BSEP polypeptide variant. Such a
10 transgenic animal can serve as an animal model for screening compounds having a potential toxic effect on humans.

In one embodiment, the invention provides a non-human transgenic animal whose somatic and germ cells contain a heterologous nucleic acid encoding a BSEP polypeptide variant operably-linked to a promoter, said nucleic acid comprising a
15 segment of nucleotide at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.

As used herein, a "transgenic non-human animal" is a non-human animal in which one or more of the cells of the animal include a transgene. As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., a BSEP
20 polypeptide variant), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which
25 differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of a selected nucleic acid.

A nucleic acid is operably linked when it is placed into a functional
30 relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of said sequence.

It may be desirable to express the heterologous BSEP polypeptide variant transgene conditionally such that either the timing or the level of expression of the
35 BSEP polypeptide variant can be regulated. Such conditional expression can be provided using prokaryotic promoter sequences, which require prokaryotic proteins to

be simultaneously expressed in order to facilitate expression of the transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Pat. No. 4,833,080.

Transgenic animals exhibiting tissue specific expression can be generated, for example, by inserting a tissue specific regulatory element, such as an enhancer, into the transgene.

Transgenic animals containing an inducible BSEP polypeptide variant transgene can be generated using inducible regulatory elements (e.g. metallothionein promoter), which are well known in the art. BSEP polypeptide variant transgene expression can then be initiated in these animals by administering to the animal a compound that induces gene expression (e.g. heavy metals). Another preferred inducible system comprises a tetracycline-inducible transcriptional activator (U.S. Pat. No. 5,654,168 issued Aug. 5, 1997 to Bujard and Gossen and U.S. Pat. No. 5,650,298 issued Jul. 22, 1997 to Bujard *et al.*).

In general, transgenic animal lines can be obtained by generating transgenic animals having incorporated into their genome at least one transgene, selecting at least one founder from these animals and breeding the founder or founders to establish at least one line of transgenic animals having the selected transgene incorporated into their genome. Methods of obtaining transgenic animals are well known in the art see, e.g., U.S. Patent Nos. 4,736,866, 4,873,191, 5,602,229, 5,625,125, 5,387,742, 5,175,385, 5,175,384, 5,175,383, 6,300,540, and 5,633,425, and the references cited therein.

In one embodiment, the transgenic animals will lack the endogenous gene encoding the BSEP polypeptide. Methods for knocking-out an endogenous gene in a non-human transgenic animal are well known in the art.

Animals for obtaining eggs or other nucleated cells (e.g. embryonic stem cells) for generating transgenic animals can be obtained from standard commercial sources such as Charles River Laboratories (Wilmington, Mass.), Taconic (Germantown, N.Y.), Harlan Sprague Dawley (Indianapolis, Ind.).

Eggs can be obtained from suitable animals, e.g., by flushing from the oviduct or using techniques described in U.S. Pat. No. 5,489,742 issued Feb. 6, 1996 to Hammer and Taurog; U.S. Pat. No. 5,625,125 issued on Apr. 29, 1997 to Bennett *et al.*; Gordon *et al.*, *Proc. Natl. Acad. Sci. USA* 77:7380-84 (1980); Gordon & Ruddle, *Science* 214: 1244-46 (1981); U.S. Pat. No. 4,873,191 to T. E. Wagner and P. C. Hoppe; U.S. Pat. No. 5,604,131; Armstrong *et al.*, *J. of Reproduction*, 39:511 (1988) or International Patent Publication No. WO 94/00568 by Mehtali *et al.* Preferably, the

female is subjected to hormonal conditions effective to promote superovulation prior to obtaining the eggs.

Many techniques can be used to introduce DNA into an egg or other nucleated cell, including in vitro fertilization using sperm as a carrier of exogenous DNA ("sperm-mediated gene transfer", e.g., Lavitrano *et al.*, *Cell* 57: 717-23 (1989), microinjection, gene targeting (Thompson *et al.*, *Cell* 56: 313-21(1989)), electroporation (Lo, *Mol. Cell. Biol.* 3: 1803-14 (1983)), transfection, or retrovirus mediated gene transfer (Van der Putten *et al.*, *Proc. Natl. Acad. Sci. USA* 82: 6148-52 (1985)). For a review of such techniques, see Gordon, Transgenic Animals, *Intl. Rev. Cytol.* 115:171-29 (1989).

Except for sperm-mediated gene transfer, eggs should be fertilized in conjunction with (before, during or after) other transgene transfer techniques. A preferred method for fertilizing eggs is by breeding the female with a fertile male. However, eggs can also be fertilized by in vitro fertilization techniques.

Fertilized, transgene containing eggs can then be transferred to pseudopregnant animals, also termed "foster mother animals", using suitable techniques. Pseudopregnant animals can be obtained, for example, by placing 40-80 day old female animals, which are more than 8 weeks of age, in cages with infertile males, e.g., vasectomized males. The next morning females are checked for vaginal plugs. Females who have mated with vasectomized males are held aside until the time of transfer.

Recipient females can be synchronized, e.g. using the GNRH agonist (GnRH-a): des-gly10, (D-Ala6)-LH-RH Ethylamide, SigmaChemical Co., St. Louis, Mo. Alternatively, a unilateral pregnancy can be achieved by a brief surgical procedure involving the "peeling" away of the bursa membrane on the left uterine horn. Injected embryos can then be transferred to the left uterine horn via the infundibulum. Potential transgenic founders can typically be identified immediately at birth from the endogenous litter mates. For generating transgenic animals from embryonic stem cells, see e.g. *Teratocarcinomas and Embryonic Stem Cells, A Practical Approach*, ed. E. J. Robertson, (IRL Press 1987) or in Potter *et al.*, *Proc. Natl. Acad. Sci. USA* 81:7161 (1984), the teachings of which are incorporated herein by reference.

Founders that express the gene can then be bred to establish a transgenic line. Accordingly, founder animals can be bred, inbred, crossbred or outbred to produce colonies of animals of the present invention. Animals comprising multiple transgenes can be generated by crossing different founder animals (e.g. an HIV transgenic animal and a transgenic animal, which expresses human CD4), as well as by

introducing multiple transgenes into an egg or embryonic cell as described above. Furthermore, embryos from A-transgenic animals can be stored as frozen embryos, which are thawed and implanted into pseudo-pregnant animals when needed (see e.g. Hirabayashi *et al.*, *Exp Anim* 46: 111 (1997) and Anzai, *Jikken Dobutsu* 43: 247 (1994)).

5 The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals that carry the transgene in some, but not all cells, i.e., mosaic animals. The transgene can be integrated as a single transgene or in tandem, e.g., head to head tandems, or head to tail or tail-to-tail or as multiple copies.

10 The successful expression of the transgene can be detected by any of several means well known to those skilled in the art. Non-limiting examples include Northern blot, in situ hybridization of mRNA analysis, Western blot analysis, immunohistochemistry, and FACS analysis of protein expression.

15 The non-human transgenic animals of the present invention can serve as animal models for screening compounds having a potential toxic effect on a subject. In one embodiment, the screening assay comprises administering a test compound to a transgenic animal of the invention, and comparing a phenotypic change, such as increased toxicity or cholestasis, in the animal relative to a transgenic animal, which has not received the test compound.

20 In a further aspect, the invention features non-human animal cells containing a BSEP polypeptide variant transgene. For example, the animal cell (e.g. somatic cell or germ cell (i.e. egg or sperm)) can be obtained from the transgenic animal. Transgenic somatic cells or cell lines can be used, for example, in screening assays for identifying substrates and/or modulators of the BSEP polypeptide variants as described above.

25 Cells from the transgenic animals of the invention can be established in culture and immortalized to establish cell lines. For example, immortalized cell lines can be established from the livers of transgenic rats, as described in Bulera *et al.*, *Hepatology* 25: 1192 (1997). Cell lines from other types of cells can be established according to methods known in the art.

EXAMPLES

Example 1: Cloning and Sequencing of the human BSEP polypeptide variants

Two oligonucleotide primers flanking the entire open reading frame (ORF) of human BSEP were chemically synthesized (Invitrogen, Carlsbad, CA). The forward sense oligo hBSEP-1 has the following sequence: 5'-AATT **ACGCGT** ACCACC **ATG** TCT GAC TCA GTA ATT CTT CGA AGT ATA AAG AAA TTT GG-3' (SEQ ID NO:13). The reverse oligo hBSEP-2 has the following sequence: 5'-AATT **GCGGCCGC** GGG **TCA** ACT GAT GGG GGA TCC AGT GGT GAC TAG TTT G-3' (SEQ ID NO:14). To facilitate the cloning, Mlu I and Not I restriction enzyme sites were incorporated into the forward and reverse primer, respectively. The initiate codon in the forward primer and the stop codon in the reverse primer were bolded and underlined. The nucleic acid encoding the human BSEP polypeptide was PCR amplified using a human liver cDNA library (Clontech, Palo Alto, CA) as templates and Platinum Taq DNA polymerase high fidelity PCR amplification kit (Invitrogen, Carlsbad, CA). The PCR reaction was carried out as following: 94° C for 2 minutes, followed by 35 cycles of 94° C for 45 second, 54° C for 45 second and 68° C for 4.5 minutes, and finally extension at 68° C for 10 minutes.

The human BSEP PCR fragments were directly cloned into pcDNA5/FRT/V5-His-TOPO expression following the protocol provided by Invitrogen. Seventeen positive clones were sequenced by the Sequencing Lab at Pfizer.

Two of the clone's encoded BSEP polypeptide insert variants and contained a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence. The sequences of these two clones are disclosed herein in SEQ ID NOs:3 and 7. SEQ ID NO:3 has extra 42 nucleotides at the junction of Exon I and II of a human BSEP polypeptide nucleotide sequence. SEQ ID NO:7 has a 102 nucleotides insertion at the junction of Exon I and II of a human BSEP polypeptide nucleotide sequence .

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al.* U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds.

1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 5 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu *et al.* eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring 10 Harbor, N.Y., 1986).

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

Unless defined otherwise, all technical and scientific terms have the same meaning as is commonly understood by one of skill in the art to which this invention 15 belongs.

All of the publications cited herein are hereby incorporated by reference in their entirety to describe more fully the art to which the application pertains.

All patents and publications (of any sort, including sequences cited by 20 accession number) mentioned in the above specification are herein incorporated by reference in their entirety. Various modifications and variations of the described invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as 25 claimed is not limited to such specific embodiments.

CLAIMS

What is claimed is:

1. An isolated nucleic acid encoding a human BSEP polypeptide insert variant, wherein the nucleic acid comprises a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.
2. The isolated nucleic acid of claim 1, wherein the human BSEP polypeptide insert variant is selected from the group consisting of:
 - (a) a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO: 4; and
 - (b) a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:8.
3. The isolated nucleic acid of claim 1, wherein the nucleic acid is selected from the group consisting of:
 - (a) the nucleic acid set forth in SEQ ID NO:3; and
 - (b) the nucleic acid set forth in SEQ ID NO:7.
4. An isolated nucleic acid capable of hybridizing under high stringency conditions to a nucleic acid comprising a nucleic acid selected from the group consisting of:
 - a nucleic acid encoding SEQ ID NO:6 or a fragment thereof;
 - (a) SEQ ID NO:5 or a fragment thereof;
 - (b) a nucleic acid encoding SEQ ID NO:10 or a fragment thereof;
 - (c) SEQ ID NO:9 or a fragment thereof;
 - (d) a nucleic acid encoding SEQ ID NO:11 or a fragment thereof; and
 - (e) SEQ ID NO:12 or a fragment thereof.
5. An isolated human BSEP polypeptide insert variant encoded by a nucleic acid comprising a segment of nucleotides at a position which corresponds to the junction between Exon I and Exon II of a human BSEP polypeptide nucleotide sequence.
6. The isolated BSEP polypeptide insert variant of claim 5, selected from the group consisting of:
 - (a) the amino acid sequence set forth in SEQ ID NO: 4; and
 - (b) the amino acid sequence set forth in SEQ ID NO:8.
7. An isolated BSEP polypeptide insert variant comprising a conservative amino acid substitution relative to the isolated human BSEP polypeptide insert variant of claim 6.

8. The isolated BSEP polypeptide insert variant of claim 7, comprising 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 conservative amino acid substitutions.
9. The isolated BSEP polypeptide insert variant of claim 7, comprising 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 conservative amino acid substitutions.
10. The host cell comprising the nucleic acid of claim 1.
11. A method for detecting a BSEP polypeptide insert variant in a sample comprising:
 - (a) contacting a suitable sample with a compound that specifically binds to the BSEP polypeptide insert variant;
 - (b) determining whether any compound is bound to the BSEP polypeptide insert variant, wherein the presence of compound bound to the BSEP polypeptide insert variant detects the BSEP polypeptide insert variant in the sample.
12. A method of determining whether a subject is expressing a BSEP polypeptide insert variant comprising:
 - (a) contacting a suitable sample from the subject with a compound that specifically binds to the BSEP polypeptide insert variant;
 - (b) determining whether any compound is bound to the BSEP polypeptide insert variant, wherein the presence of compound bound to the BSEP polypeptide insert variant indicates that the subject is expressing the BSEP polypeptide insert variant.
13. A method for determining whether a compound is a substrate or modulator of a BSEP polypeptide insert variant comprising:
 - (a) contacting the compound with the BSEP polypeptide insert variant; and
 - (b) determining whether any compound binds to the BSEP polypeptide insert variant, wherein the presence of compound bound to the BSEP polypeptide insert variant indicates that the compound is a substrate or modulator of the BSEP polypeptide insert variant.
14. A method for determining whether a compound is a substrate or modulator of a BSEP polypeptide insert variant comprising:
 - (a) contacting the BSEP polypeptide insert variant with a known substrate of the BSEP polypeptide insert variant, a compound, and ATP;
 - (b) measuring the ATPase activity of the BSEP polypeptide insert variant; and

- (c) comparing the ATPase activity of the BSEP polypeptide insert variant in step (b) with the ATPase activity of the BSEP polypeptide insert variant in the absence of the compound, wherein an increase or a decrease in the ATPase activity indicates that the compound is a substrate or modulator of the BSEP polypeptide variant.
15. A method for determining whether a compound is a substrate or modulator of a BSEP polypeptide insert variant comprising:
- (a) contacting the compound with the BSEP polypeptide insert variant and a known substrate of the BSEP polypeptide insert variant;
 - (b) measuring the transport activity of the BSEP polypeptide insert variant with respect to the known substrate;
 - (c) comparing the transport activity of the BSEP polypeptide insert variant with respect to the known substrate in step (b) with the transport activity of the BSEP polypeptide with respect to the known substrate in the absence of the compound, wherein an increase or a decrease in transport of activity of the BSEP polypeptide insert variant with respect to the known substrate indicates that the compound is a substrate or modulator of the BSEP polypeptide insert variant.

SEQUENCE LISTING

<110> Pfizer Inc.
Ruitang Deng
Amy Y. Willis
Jinghai J. Xu

<120> BSEP POLYPEPTIDE VARIANTS AND USES THEREOF

<130> PC25434A

<140>

<141>

<160> 14

<170> PatentIn version 3.2

<210> 1

<211> 3966

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(3966)

<400> 1

atg tct gac tca gta att ctt cga agt ata aag aaa ttt gga gag gag	48
Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu	
1 5 10 15	

aat gat ggt ttt gag tca gat aaa tca tat aat aat gat aag aaa tca	96
Asn Asp Gly Phe Glu Ser Asp Lys Ser Tyr Asn Asn Asp Lys Lys Ser	
20 25 30	

agg tta caa gat gag aag aaa ggt gat ggc gtt aga gtt ggc ttc ttt	144
Arg Leu Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly Phe Phe	
35 40 45	

caa ttg ttt cgg ttt tct tca tca act gac att tgg ctg atg ttt gtg	192
Gln Leu Phe Arg Phe Ser Ser Thr Asp Ile Trp Leu Met Phe Val	
50 55 60	

gga agt ttg tgt gca ttt ctc cat gga ata gcc cag cca ggc gtg cta	240
Gly Ser Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly Val Leu	
65 70 75 80	

ctc att ttt ggc aca atg aca gat gtt ttt att gac tac gac gtt gag	288
Leu Ile Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp Val Glu	
85 90 95	

tta caa gaa ctc cag att cca gga aaa gca tgt gtg aat aac acc att	336
Leu Gln Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn Thr Ile	
100 105 110	

gta tgg act aac agt tcc ctc aac cag aac atg aca aat gga aca cgt	384
Val Trp Thr Asn Ser Ser Leu Asn Gln Asn Met Thr Asn Gly Thr Arg	
115 120 125	

tgt ggg ttg ctg aac atc gag agc gaa atg atc aaa ttt gcc agt tac	432
Cys Gly Leu Leu Asn Ile Glu Ser Glu Met Ile Lys Phe Ala Ser Tyr	
130 135 140	
tat gct gga att gct gtc gca gta ctt atc aca gga tat att caa ata	480
Tyr Ala Gly Ile Ala Val Ala Val Leu Ile Thr Gly Tyr Ile Gln Ile	
145 150 155 160	
tgc ttt tgg gtc att gcc gca gct cgt cag ata cag aaa atg aga aaa	528
Cys Phe Trp Val Ile Ala Ala Ala Arg Gln Ile Gln Lys Met Arg Lys	
165 170 175	
ttt tac ttt agg aga ata atg aga atg gaa ata ggg tgg ttt gac tgc	576
Phe Tyr Phe Arg Arg Ile Met Arg Met Glu Ile Gly Trp Phe Asp Cys	
180 185 190	
aat tca gtg ggg gag ctg aat aca aga ttc tct gat gat att aat aaa	624
Asn Ser Val Gly Glu Leu Asn Thr Arg Phe Ser Asp Asp Ile Asn Lys	
195 200 205	
atc aat gat gcc ata gct gac caa atg gcc ctt ttc att cag cgc atg	672
Ile Asn Asp Ala Ile Ala Asp Gln Met Ala Leu Phe Ile Gln Arg Met	
210 215 220	
acc tcg acc atc tgt ggt ttc ctg ttg gga ttt ttc agg ggt tgg aaa	720
Thr Ser Thr Ile Cys Gly Phe Leu Leu Gly Phe Phe Arg Gly Trp Lys	
225 230 235 240	
ctg acc ttg gtt att att tct gtc agc cct ctc att ggg att gga gca	768
Leu Thr Leu Val Ile Ile Ser Val Ser Pro Leu Ile Gly Ile Gly Ala	
245 250 255	
gcc acc att ggt ctg agt gtg tcc aag ttt acg gac tat gag ctg aag	816
Ala Thr Ile Gly Leu Ser Val Ser Lys Phe Thr Asp Tyr Glu Leu Lys	
260 265 270	
gcc tat gcc aaa gca ggg gtg gtg gct gat gaa gtc att tca tca atg	864
Ala Tyr Ala Lys Ala Gly Val Val Ala Asp Glu Val Ile Ser Ser Met	
275 280 285	
aga aca gtg gct gct ttt ggt ggt gag aaa aga gag gtt gaa agg tat	912
Arg Thr Val Ala Ala Phe Gly Gly Glu Lys Arg Glu Val Glu Arg Tyr	
290 295 300	
gag aaa aat ctt gtg ttc gcc cag cgt tgg gga att aga aaa gga ata	960
Glu Lys Asn Leu Val Phe Ala Gln Arg Trp Gly Ile Arg Lys Gly Ile	
305 310 315 320	
gtg atg gga ttc ttt act gga ttc gtg tgg tgt ctc atc ttt ttg tgt	1008
Val Met Gly Phe Phe Thr Gly Phe Val Trp Cys Leu Ile Phe Leu Cys	
325 330 335	
tat gca ctg gcc ttc tgg tac ggc tcc aca ctt gtc ctg gat gaa gga	1056
Tyr Ala Leu Ala Phe Trp Tyr Gly Ser Thr Leu Val Leu Asp Glu Gly	
340 345 350	
gaa tat aca cca gga acc ctt gtc cag att ttc ctc agt gtc ata gta	1104
Glu Tyr Thr Pro Gly Thr Leu Val Gln Ile Phe Leu Ser Val Ile Val	
355 360 365	

gga gct tta aat ctt ggc aat gcc tct cct tgt ttg gaa gcc ttt gca	1152
Gly Ala Leu Asn Leu Gly Asn Ala Ser Pro Cys Leu Glu Ala Phe Ala	
370 375 380	
act gga cgt gca gca gcc acc agc att ttt gag aca ata gac agg aaa	1200
Thr Gly Arg Ala Ala Thr Ser Ile Phe Glu Thr Ile Asp Arg Lys	
385 390 395 400	
ccc atc att gac tgc atg tca gaa gat ggt tac aag ttg gat cga atc	1248
Pro Ile Ile Asp Cys Met Ser Glu Asp Gly Tyr Lys Leu Asp Arg Ile	
405 410 415	
aag ggt gaa att gaa ttc cat aat gtg acc ttc cat tat cct tcc aga	1296
Lys Gly Glu Ile Glu Phe His Asn Val Thr Phe His Tyr Pro Ser Arg	
420 425 430	
cca gag gtg aag att cta aat gac ctc aac atg gtc att aaa cca ggg	1344
Pro Glu Val Lys Ile Leu Asn Asp Leu Asn Met Val Ile Lys Pro Gly	
435 440 445	
gaa atg aca gct ctg gta gga ccc agt gga gct gga aaa agt aca gca	1392
Glu Met Thr Ala Leu Val Gly Pro Ser Gly Ala Gly Lys Ser Thr Ala	
450 455 460	
ctg caa ctc att cag cga ttc tat gac ccc tgt gaa gga atg gtg acc	1440
Leu Gln Leu Ile Gln Arg Phe Tyr Asp Pro Cys Glu Gly Met Val Thr	
465 470 475 480	
gtg gat ggc cat gac att cgc tct ctt aac att cag tgg ctt aga gat	1488
Val Asp Gly His Asp Ile Arg Ser Leu Asn Ile Gln Trp Leu Arg Asp	
485 490 495	
cag att ggg ata gtg gag caa gag cca gtt ctg ttc tct acc acc att	1536
Gln Ile Gly Ile Val Glu Gln Glu Pro Val Leu Phe Ser Thr Thr Ile	
500 505 510	
gca gaa aat att cgc tat ggc aga gaa gat gca aca atg gaa gac ata	1584
Ala Glu Asn Ile Arg Tyr Gly Arg Glu Asp Ala Thr Met Glu Asp Ile	
515 520 525	
gtc caa gct gcc aag gag gcc aat gcc tac aac ttc atc atg gac ctg	1632
Val Gln Ala Ala Lys Glu Ala Asn Ala Tyr Asn Phe Ile Met Asp Leu	
530 535 540	
cca cag caa ttt gac acc ctt gtt gga gaa gga gga ggc cag atg agt	1680
Pro Gln Gln Phe Asp Thr Leu Val Gly Glu Gly Gly Gln Met Ser	
545 550 555 560	
ggt ggc cag aaa caa agg gta gct atc gcc aga gcc ctc atc cga aat	1728
Gly Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Ile Arg Asn	
565 570 575	
ccc aag att ctg ctt ttg gac atg gcc acc tca gct ctg gac aat gag	1776
Pro Lys Ile Leu Leu Leu Asp Met Ala Thr Ser Ala Leu Asp Asn Glu	
580 585 590	
agt gaa gcc atg gtg caa gaa gtg ctg agt aag att cag cat ggg cac	1824
Ser Glu Ala Met Val Gln Glu Val Leu Ser Lys Ile Gln His Gly His	
595 600 605	
aca atc att tca gtt gct cat cgc ttg tct acg gtc aga gct gca gat	1872

Thr Ile Ile Ser Val Ala His Arg Leu Ser Thr Val Arg Ala Ala Asp 610 615 620	
acc atc att ggt ttt gaa cat ggc act gca gtg gaa aga ggg acc cat Thr Ile Ile Gly Phe Glu His Gly Thr Ala Val Glu Arg Gly Thr His 625 630 635 640	1920
gaa gaa tta ctg gaa agg aaa ggt gtt tac ttc act cta gtg act ttg Glu Glu Leu Leu Glu Arg Lys Gly Val Tyr Phe Thr Leu Val Thr Leu 645 650 655	1968
caa agc cag gga aat caa gct ctt aat gaa gag gac ata aag gat gca Gln Ser Gln Gly Asn Gln Ala Leu Asn Glu Glu Asp Ile Lys Asp Ala 660 665 670	2016
act gaa gat gac atg ctt gcg agg acc ttt agc aga ggg agc tac cag Thr Glu Asp Asp Met Leu Ala Arg Thr Phe Ser Arg Gly Ser Tyr Gln 675 680 685	2064
gat agt tta agg gct tcc atc cgg caa cgc tcc aag tct cag ctt tct Asp Ser Leu Arg Ala Ser Ile Arg Gln Arg Ser Lys Ser Gln Leu Ser 690 695 700	2112
tac ctg gtg cac gaa cct cca tta gct gtt gta gat cat aag tct acc Tyr Leu Val His Glu Pro Pro Leu Ala Val Val Asp His Lys Ser Thr 705 710 715 720	2160
tat gaa gaa gat aga aag gac aag gac att cct gtg cag gaa gaa gtt Tyr Glu Glu Asp Arg Lys Asp Lys Asp Ile Pro Val Gln Glu Glu Val 725 730 735	2208
gaa cct gcc cca gtt agg agg att ctg aaa ttc agt gct cca gaa tgg Glu Pro Ala Pro Val Arg Arg Ile Leu Lys Phe Ser Ala Pro Glu Trp 740 745 750	2256
ccc tac atg ctg gta ggg tct gtg ggt gca gct gtg aac ggg aca gtc Pro Tyr Met Leu Val Gly Ser Val Gly Ala Ala Val Asn Gly Thr Val 755 760 765	2304
aca ccc ttg tat gcc ttt tta ttc agc cag att ctt ggg act ttt tca Thr Pro Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr Phe Ser 770 775 780	2352
att cct gat aaa gag gaa caa agg tca cag atc aat ggt gtg tgc cta Ile Pro Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val Cys Leu 785 790 795 800	2400
ctt ttt gta gca atg ggc tgt gta tct ctt ttc acc caa ttt cta cag Leu Phe Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe Leu Gln 805 810 815	2448
gga tat gcc ttt gct aaa tct ggg gag ctc cta aca aaa agg cta cgt Gly Tyr Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg Leu Arg 820 825 830	2496
aaa ttt ggt ttc agg gca atg ctg ggg caa gat att gcc tgg ttt gat Lys Phe Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp Phe Asp 835 840 845	2544
gac ctc aga aat agc cct gga gca ttg aca aca aga ctt gct aca gat Asp Leu Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala Thr Asp	2592

850	855	860	
gct tcc caa gtt caa ggg gct gcc ggc tct cag atc ggg atg ata gtc			2640
Ala Ser Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met Ile Val			
865	870	875	880
aat tcc ttc act aac gtc act gtg gcc atg atc att gcc ttc tcc ttt			2688
Asn Ser Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe Ser Phe			
	885	890	895
agc tgg aag ctg agc ctg gtc atc ttg tgc ttc ttc ccc ttc ttg gct			2736
Ser Trp Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe Leu Ala			
	900	905	910
tta tca gga gcc aca cag acc agg atg ttg aca gga ttt gcc tct cga			2784
Leu Ser Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala Ser Arg			
	915	920	925
gat aag cag gcc ctg gag atg gtg gga cag att aca aat gaa gcc ctc			2832
Asp Lys Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu Ala Leu			
	930	935	940
agt aac atc cgc act gtt gct gga att gga aag gag agg cgg ttc att			2880
Ser Asn Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg Phe Ile			
	945	950	955
gaa gca ctt gag act gag ctg gag aag ccc ttc aag aca gcc att cag			2928
Glu Ala Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala Ile Gln			
	965	970	975
aaa gcc aat att tac gga ttc tgc ttt gcc ttt gcc cag tgc atc atg			2976
Lys Ala Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln Cys Ile Met			
	980	985	990
ttt att gcg aat tct gct tcc tac aga tat gga ggt tac tta atc tcc			3024
Phe Ile Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly Tyr Leu Ile Ser			
	995	1000	1005
aat gag ggg ctc cat ttc agc tat gtg ttc agg gtg atc tct gca			3069
Asn Glu Gly Leu His Phe Ser Tyr Val Phe Arg Val Ile Ser Ala			
	1010	1015	1020
gtt gta ctg agt gca aca gct ctt gga aga gcc ttc tct tac acc			3114
Val Val Leu Ser Ala Thr Ala Leu Gly Arg Ala Phe Ser Tyr Thr			
	1025	1030	1035
cca agt tat gca aaa gct aaa ata tca gct gca cgc ttt ttt caa			3159
Pro Ser Tyr Ala Lys Ala Lys Ile Ser Ala Ala Arg Phe Phe Gln			
	1040	1045	1050
ctg ctg gac cga caa ccc cca atc agt gta tac aat act gca ggt			3204
Leu Leu Asp Arg Gln Pro Pro Ile Ser Val Tyr Asn Thr Ala Gly			
	1055	1060	1065
gaa aaa tgg gac aac ttc cag ggg aag att gat ttt gtt gat tgt			3249
Glu Lys Trp Asp Asn Phe Gln Gly Lys Ile Asp Phe Val Asp Cys			
	1070	1075	1080
aaa ttt aca tat cct tct cga cct gac tcg caa gtt ctg aat ggt			3294
Lys Phe Thr Tyr Pro Ser Arg Pro Asp Ser Gln Val Leu Asn Gly			
	1085	1090	1095

ctc tca	gtg tcg	att agt	cca	ggg cag	aca ctg	gcg	ttt gtt	ggg	3339
Leu Ser	Val Ser	Ile Ser	Pro	Gly Gln	Thr Leu		Phe Val	Gly	
1100			1105			1110			
agc agt	gga tgt	ggc aaa	agc	act agc	att cag	ctg	ttg gaa	cgt	3384
Ser Ser	Gly Cys	Gly Lys	Ser	Thr Ser	Ile Gln	Leu	Leu Glu	Arg	
1115			1120			1125			
ttc tat	gat cct	gat caa	ggg	aag gtg	atg ata	gat	ggg cat	gac	3429
Phe Tyr	Asp Pro	Asp Gln	Gly	Lys Val	Met Ile	Asp	Gly His	Asp	
1130			1135			1140			
agc aaa	aaa gta	aat gtc	cag	ttc ctc	cgc tca	aac	att gga	att	3474
Ser Lys	Lys Val	Asn Val	Gln	Phe Leu	Arg Ser	Asn	Ile Gly	Ile	
1145			1150			1155			
gtt tcc	cag gaa	cca gtg	ttg	ttt gcc	tgt agc	ata	atg gac	aat	3519
Val Ser	Gln Glu	Pro Val	Leu	Phe Ala	Cys Ser	Ile	Met Asp	Asn	
1160			1165			1170			
atc aag	tat gga	gac aac	acc	aaa gaa	att ccc	atg	gaa aga	gtc	3564
Ile Lys	Tyr Gly	Asp Asn	Thr	Lys Glu	Ile Pro	Met	Glu Arg	Val	
1175			1180			1185			
ata gca	gct gca	aaa cag	gct	cag ctg	cat gat	ttt	gtc atg	tca	3609
Ile Ala	Ala Ala	Lys Gln	Ala	Gln Leu	His Asp	Phe	Val Met	Ser	
1190			1195			1200			
ctc cca	gag aaa	tat gaa	act	aac gtt	ggg tcc	cag	ggg tct	caa	3654
Leu Pro	Glu Lys	Tyr Glu	Thr	Asn Val	Gly Ser	Gln	Gly Ser	Gln	
1205			1210			1215			
ctc tct	aga ggg	gag aaa	caa	cgc att	gct att	gct	cgg gcc	att	3699
Leu Ser	Arg Gly	Glu Lys	Gln	Arg Ile	Ala Ile	Ala	Arg Ala	Ile	
1220			1225			1230			
gta cga	gat cct	aaa atc	ttg	cta cta	gat gaa	gcc	act tct	gcc	3744
Val Arg	Asp Pro	Lys Ile	Leu	Leu Leu	Asp Glu	Ala	Thr Ser	Ala	
1235			1240			1245			
tta gac	aca gaa	agt gaa	aag	acg gtg	cag gtt	gct	cta gac	aaa	3789
Leu Asp	Thr Glu	Ser Glu	Lys	Thr Val	Gln Val	Ala	Leu Asp	Lys	
1250			1255			1260			
gcc aga	gag ggt	cgg acc	tgc	att gtc	att gcc	cat	cgc ttg	tcc	3834
Ala Arg	Glu Gly	Arg Thr	Cys	Ile Val	Ile Ala	His	Arg Leu	Ser	
1265			1270			1275			
acc atc	cag aac	gcg gat	atc	att gct	gtc atg	gca	cag ggg	gtg	3879
Thr Ile	Gln Asn	Ala Asp	Ile	Ile Ala	Val Met	Ala	Gln Gly	Val	
1280			1285			1290			
gtg att	gaa aag	ggg acc	cat	gaa gaa	ctg atg	gcc	caa aaa	gga	3924
Val Ile	Glu Lys	Gly Thr	His	Glu Glu	Leu Met	Ala	Gln Lys	Gly	
1295			1300			1305			
gcc tac	tac aaa	cta gtc	acc	act gga	tcc ccc	atc	agt tga		3966
Ala Tyr	Tyr Lys	Leu Val	Thr	Thr Gly	Ser Pro	Ile	Ser		
1310			1315			1320			

<210> 2
 <211> 1321
 <212> PRT
 <213> Homo sapiens

<400> 2

Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu
 1 5 10 15

Asn Asp Gly Phe Glu Ser Asp Lys Ser Tyr Asn Asn Asp Lys Lys Ser
 20 25 30

Arg Leu Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly Phe Phe
 35 40 45

Gln Leu Phe Arg Phe Ser Ser Ser Thr Asp Ile Trp Leu Met Phe Val
 50 55 60

Gly Ser Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly Val Leu
 65 70 75 80

Leu Ile Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp Val Glu
 85 90 95

Leu Gln Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn Thr Ile
 100 105 110

Val Trp Thr Asn Ser Ser Leu Asn Gln Asn Met Thr Asn Gly Thr Arg
 115 120 125

Cys Gly Leu Leu Asn Ile Glu Ser Glu Met Ile Lys Phe Ala Ser Tyr
 130 135 140

Tyr Ala Gly Ile Ala Val Ala Val Leu Ile Thr Gly Tyr Ile Gln Ile
 145 150 155 160

Cys Phe Trp Val Ile Ala Ala Ala Arg Gln Ile Gln Lys Met Arg Lys
 165 170 175

Phe Tyr Phe Arg Arg Ile Met Arg Met Glu Ile Gly Trp Phe Asp Cys
 180 185 190

Asn Ser Val Gly Glu Leu Asn Thr Arg Phe Ser Asp Asp Ile Asn Lys
 195 200 205

Ile Asn Asp Ala Ile Ala Asp Gln Met Ala Leu Phe Ile Gln Arg Met

210	215	220
Thr Ser Thr Ile Cys Gly Phe Leu Leu Gly Phe Phe Arg Gly Trp Lys 225 230 235 240		
Leu Thr Leu Val Ile Ile Ser Val Ser Pro Leu Ile Gly Ile Gly Ala 245 250 255		
Ala Thr Ile Gly Leu Ser Val Ser Lys Phe Thr Asp Tyr Glu Leu Lys 260 265 270		
Ala Tyr Ala Lys Ala Gly Val Val Ala Asp Glu Val Ile Ser Ser Met 275 280 285		
Arg Thr Val Ala Ala Phe Gly Gly Glu Lys Arg Glu Val Glu Arg Tyr 290 295 300		
Glu Lys Asn Leu Val Phe Ala Gln Arg Trp Gly Ile Arg Lys Gly Ile 305 310 315 320		
Val Met Gly Phe Phe Thr Gly Phe Val Trp Cys Leu Ile Phe Leu Cys 325 330 335		
Tyr Ala Leu Ala Phe Trp Tyr Gly Ser Thr Leu Val Leu Asp Glu Gly 340 345 350		
Glu Tyr Thr Pro Gly Thr Leu Val Gln Ile Phe Leu Ser Val Ile Val 355 360 365		
Gly Ala Leu Asn Leu Gly Asn Ala Ser Pro Cys Leu Glu Ala Phe Ala 370 375 380		
Thr Gly Arg Ala Ala Ala Thr Ser Ile Phe Glu Thr Ile Asp Arg Lys 385 390 395 400		
Pro Ile Ile Asp Cys Met Ser Glu Asp Gly Tyr Lys Leu Asp Arg Ile 405 410 415		
Lys Gly Glu Ile Glu Phe His Asn Val Thr Phe His Tyr Pro Ser Arg 420 425 430		
Pro Glu Val Lys Ile Leu Asn Asp Leu Asn Met Val Ile Lys Pro Gly 435 440 445		
Glu Met Thr Ala Leu Val Gly Pro Ser Gly Ala Gly Lys Ser Thr Ala 450 455 460		

Leu Gln Leu Ile Gln Arg Phe Tyr Asp Pro Cys Glu Gly Met Val Thr
 465 470 475 480

Val Asp Gly His Asp Ile Arg Ser Leu Asn Ile Gln Trp Leu Arg Asp
 485 490 495

Gln Ile Gly Ile Val Glu Gln Glu Pro Val Leu Phe Ser Thr Thr Ile
 500 505 510

Ala Glu Asn Ile Arg Tyr Gly Arg Glu Asp Ala Thr Met Glu Asp Ile
 515 520 525

Val Gln Ala Ala Lys Glu Ala Asn Ala Tyr Asn Phe Ile Met Asp Leu
 530 535 540

Pro Gln Gln Phe Asp Thr Leu Val Gly Glu Gly Gly Gly Gln Met Ser
 545 550 555 560

Gly Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Ile Arg Asn
 565 570 575

Pro Lys Ile Leu Leu Leu Asp Met Ala Thr Ser Ala Leu Asp Asn Glu
 580 585 590

Ser Glu Ala Met Val Gln Glu Val Leu Ser Lys Ile Gln His Gly His
 595 600 605

Thr Ile Ile Ser Val Ala His Arg Leu Ser Thr Val Arg Ala Ala Asp
 610 615 620

Thr Ile Ile Gly Phe Glu His Gly Thr Ala Val Glu Arg Gly Thr His
 625 630 635 640

Glu Glu Leu Leu Glu Arg Lys Gly Val Tyr Phe Thr Leu Val Thr Leu
 645 650 655

Gln Ser Gln Gly Asn Gln Ala Leu Asn Glu Glu Asp Ile Lys Asp Ala
 660 665 670

Thr Glu Asp Asp Met Leu Ala Arg Thr Phe Ser Arg Gly Ser Tyr Gln
 675 680 685

Asp Ser Leu Arg Ala Ser Ile Arg Gln Arg Ser Lys Ser Gln Leu Ser
 690 695 700

Tyr Leu Val His Glu Pro Pro Leu Ala Val Val Asp His Lys Ser Thr
 705 710 715 720

Tyr Glu Glu Asp Arg Lys Asp Lys Asp Ile Pro Val Gln Glu Glu Val
 725 730 735

Glu Pro Ala Pro Val Arg Arg Ile Leu Lys Phe Ser Ala Pro Glu Trp
 740 745 750

Pro Tyr Met Leu Val Gly Ser Val Gly Ala Ala Val Asn Gly Thr Val
 755 760 765

Thr Pro Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr Phe Ser
 770 775 780

Ile Pro Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val Cys Leu
 785 790 795 800

Leu Phe Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe Leu Gln
 805 810 815

Gly Tyr Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg Leu Arg
 820 825 830

Lys Phe Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp Phe Asp
 835 840 845

Asp Leu Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala Thr Asp
 850 855 860

Ala Ser Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met Ile Val
 865 870 875 880

Asn Ser Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe Ser Phe
 885 890 895

Ser Trp Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe Leu Ala
 900 905 910

Leu Ser Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala Ser Arg
 915 920 925

Asp Lys Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu Ala Leu
 930 935 940

Ser Asn Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg Phe Ile
945 950 955 960

Glu Ala Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala Ile Gln
965 970 975

Lys Ala Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln Cys Ile Met
980 985 990

Phe Ile Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly Tyr Leu Ile Ser
995 1000 1005

Asn Glu Gly Leu His Phe Ser Tyr Val Phe Arg Val Ile Ser Ala
1010 1015 1020

Val Val Leu Ser Ala Thr Ala Leu Gly Arg Ala Phe Ser Tyr Thr
1025 1030 1035

Pro Ser Tyr Ala Lys Ala Lys Ile Ser Ala Ala Arg Phe Phe Gln
1040 1045 1050

Leu Leu Asp Arg Gln Pro Pro Ile Ser Val Tyr Asn Thr Ala Gly
1055 1060 1065

Glu Lys Trp Asp Asn Phe Gln Gly Lys Ile Asp Phe Val Asp Cys
1070 1075 1080

Lys Phe Thr Tyr Pro Ser Arg Pro Asp Ser Gln Val Leu Asn Gly
1085 1090 1095

Leu Ser Val Ser Ile Ser Pro Gly Gln Thr Leu Ala Phe Val Gly
1100 1105 1110

Ser Ser Gly Cys Gly Lys Ser Thr Ser Ile Gln Leu Leu Glu Arg
1115 1120 1125

Phe Tyr Asp Pro Asp Gln Gly Lys Val Met Ile Asp Gly His Asp
1130 1135 1140

Ser Lys Lys Val Asn Val Gln Phe Leu Arg Ser Asn Ile Gly Ile
1145 1150 1155

Val Ser Gln Glu Pro Val Leu Phe Ala Cys Ser Ile Met Asp Asn
1160 1165 1170

Ile Lys Tyr Gly Asp Asn Thr Lys Glu Ile Pro Met Glu Arg Val

1175	1180	1185	
Ile Ala Ala Ala Lys Gln	Ala Gln Leu His Asp	Phe Val Met Ser	
1190	1195	1200	
Leu Pro Glu Lys Tyr Glu	Thr Asn Val Gly Ser	Gln Gly Ser Gln	
1205	1210	1215	
Leu Ser Arg Gly Glu Lys	Gln Arg Ile Ala Ile	Ala Arg Ala Ile	
1220	1225	1230	
Val Arg Asp Pro Lys Ile	Leu Leu Leu Asp Glu	Ala Thr Ser Ala	
1235	1240	1245	
Leu Asp Thr Glu Ser Glu	Lys Thr Val Gln Val	Ala Leu Asp Lys	
1250	1255	1260	
Ala Arg Glu Gly Arg Thr	Cys Ile Val Ile Ala	His Arg Leu Ser	
1265	1270	1275	
Thr Ile Gln Asn Ala Asp	Ile Ile Ala Val Met	Ala Gln Gly Val	
1280	1285	1290	
Val Ile Glu Lys Gly Thr	His Glu Glu Leu Met	Ala Gln Lys Gly	
1295	1300	1305	
Ala Tyr Tyr Lys Leu Val	Thr Thr Gly Ser Pro	Ile Ser	
1310	1315	1320	
<210> 3			
<211> 4008			
<212> DNA			
<213> Homo sapiens			
<220>			
<221> CDS			
<222> (1)..(4008)			
<400> 3			
atg tct gac tca gta att ctt cga agt ata aag aaa ttt gga gag gag	48		
Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu			
1 5 10 15			
aat gat ggt ttt gag tca gat aaa tca tgt gag tgg ctt ttt tcc ctc	96		
Asn Asp Gly Phe Glu Ser Asp Lys Ser Cys Glu Trp Leu Phe Ser Leu			
20 25 30			
act gca tct tgt aca agg aga gat aat aat gat aag aaa tca agg tta	144		
Thr Ala Ser Cys Thr Arg Arg Asp Asn Asn Asp Lys Lys Ser Arg Leu			
35 40 45			

caa gat gag aag aaa ggt gat ggc gtt aga gtt ggc ttc ttt caa ttg Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly Phe Phe Gln Leu 50 55 60	192
ttt cgg ttt tct tca tca act gac att tgg ctg atg ttt gtg gga agt Phe Arg Phe Ser Ser Ser Thr Asp Ile Trp Leu Met Phe Val Gly Ser 65 70 75 80	240
ttg tgt gca ttt ctc cat gga ata gcc cag cca ggc gtg cta ctc att Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly Val Leu Leu Ile 85 90 95	288
ttt ggc aca atg aca gat gtt ttt att gac tac gac gtt gag tta caa Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp Val Glu Leu Gln 100 105 110	336
gaa ctc cag att cca gga aaa gca tgt gtg aat aac acc att gta tgg Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn Thr Ile Val Trp 115 120 125	384
act aac agt tcc ctc aac cag aac atg aca aat gga aca cgt tgt ggg Thr Asn Ser Ser Leu Asn Gln Asn Met Thr Asn Gly Thr Arg Cys Gly 130 135 140	432
ttg ctg aac atc gag agc gaa atg atc aaa ttt gcc agt tac tat gct Leu Leu Asn Ile Glu Ser Glu Met Ile Lys Phe Ala Ser Tyr Tyr Ala 145 150 155 160	480
gga att gct gtc gca gta ctt atc aca gga tat att caa ata tgc ttt Gly Ile Ala Val Ala Val Leu Ile Thr Gly Tyr Ile Gln Ile Cys Phe 165 170 175	528
tgg gtc att gcc gca gct cgt cag ata cag aaa atg aga aaa ttt tac Trp Val Ile Ala Ala Ala Arg Gln Ile Gln Lys Met Arg Lys Phe Tyr 180 185 190	576
ttt agg aga ata atg aga atg gaa ata ggg tgg ttt gac tgc aat tca Phe Arg Arg Ile Met Arg Met Glu Ile Gly Trp Phe Asp Cys Asn Ser 195 200 205	624
gtg ggg gag ctg aat aca aga ttc tct gat gat att aat aaa atc aat Val Gly Glu Leu Asn Thr Arg Phe Ser Asp Asp Ile Asn Lys Ile Asn 210 215 220	672
gat gcc ata gct gac caa atg gcc ctt ttc att cag cgc atg acc tcg Asp Ala Ile Ala Asp Gln Met Ala Leu Phe Ile Gln Arg Met Thr Ser 225 230 235 240	720
acc atc tgt ggt ttc ctg ttg gga ttt ttc agg ggt tgg aaa ctg acc Thr Ile Cys Gly Phe Leu Leu Gly Phe Phe Arg Gly Trp Lys Leu Thr 245 250 255	768
ttg gtt att att tct gtc agc cct ctc att ggg att gga gca gcc acc Leu Val Ile Ile Ser Val Ser Pro Leu Ile Gly Ile Gly Ala Ala Thr 260 265 270	816
att ggt ctg agt gtg tcc aag ttt acg gac tat gag ctg aag gcc tat Ile Gly Leu Ser Val Ser Lys Phe Thr Asp Tyr Glu Leu Lys Ala Tyr 275 280 285	864

gcc aaa gca ggg gtg gtg gct gat gaa gtc att tca tca atg aga aca Ala Lys Ala Gly Val Val Ala Asp Glu Val Ile Ser Ser Met Arg Thr 290 295 300	912
gtg gct gct ttt ggt ggt gag aaa aga gag gtt gaa agg tat gag aaa Val Ala Ala Phe Gly Gly Glu Lys Arg Glu Val Glu Arg Tyr Glu Lys 305 310 315 320	960
aat ctt gtg ttc gcc cag cgt tgg gga att aga aaa gga ata gtg atg Asn Leu Val Phe Ala Gln Arg Trp Gly Ile Arg Lys Gly Ile Val Met 325 330 335	1008
gga ttc ttt act gga ttc gtg tgg tgt ctc atc ttt ttg tgt tat gca Gly Phe Phe Thr Gly Phe Val Trp Cys Leu Ile Phe Leu Cys Tyr Ala 340 345 350	1056
ctg gcc ttc tgg tac ggc tcc aca ctt gtc ctg gat gaa gga gaa tat Leu Ala Phe Trp Tyr Gly Ser Thr Leu Val Leu Asp Glu Gly Glu Tyr 355 360 365	1104
aca cca gga acc ctt gtc cag att ttc ctc agt gtc ata gta gga gct Thr Pro Gly Thr Leu Val Gln Ile Phe Leu Ser Val Ile Val Gly Ala 370 375 380	1152
tta aat ctt ggc aat gcc tct cct tgt ttg gaa gcc ttt gca act gga Leu Asn Leu Gly Asn Ala Ser Pro Cys Leu Glu Ala Phe Ala Thr Gly 385 390 395 400	1200
cgt gca gca gcc acc agc att ttt gag aca ata gac agg aaa ccc atc Arg Ala Ala Ala Thr Ser Ile Phe Glu Thr Ile Asp Arg Lys Pro Ile 405 410 415	1248
att gac tgc atg tca gaa gat ggt tac aag ttg gat cga atc aag ggt Ile Asp Cys Met Ser Glu Asp Gly Tyr Lys Leu Asp Arg Ile Lys Gly 420 425 430	1296
gaa att gaa ttc cat aat gtg acc ttc cat tat cct tcc aga cca gag Glu Ile Glu Phe His Asn Val Thr Phe His Tyr Pro Ser Arg Pro Glu 435 440 445	1344
gtg aag att cta aat gac ctc aac atg gtc att aaa cca ggg gaa atg Val Lys Ile Leu Asn Asp Leu Asn Met Val Ile Lys Pro Gly Glu Met 450 455 460	1392
aca gct ctg gta gga ccc agt gga gct gga aaa agt aca gca ctg caa Thr Ala Leu Val Gly Pro Ser Gly Ala Gly Lys Ser Thr Ala Leu Gln 465 470 475 480	1440
ctc att cag cga ttc tat gac ccc tgt gaa gga atg gtg acc gtg gat Leu Ile Gln Arg Phe Tyr Asp Pro Cys Glu Gly Met Val Thr Val Asp 485 490 495	1488
ggc cat gac att cgc tct ctt aac att cag tgg ctt aga gat cag att Gly His Asp Ile Arg Ser Leu Asn Ile Gln Trp Leu Arg Asp Gln Ile 500 505 510	1536
ggg ata gtg gag caa gag cca gtt ctg ttc tct acc acc att gca gaa Gly Ile Val Glu Gln Glu Pro Val Leu Phe Ser Thr Thr Ile Ala Glu 515 520 525	1584
aat att cgc tat ggc aga gaa gat gca aca atg gaa gac ata gtc caa	1632

Asn	Ile	Arg	Tyr	Gly	Arg	Glu	Asp	Ala	Thr	Met	Glu	Asp	Ile	Val	Gln	
530						535					540					
gct	gcc	aag	gag	gcc	aat	gcc	tac	aac	ttc	atc	atg	gac	ctg	cca	cag	1680
Ala	Ala	Lys	Glu	Ala	Asn	Ala	Tyr	Asn	Phe	Ile	Met	Asp	Leu	Pro	Gln	
545					550				555						560	
caa	ttt	gac	acc	ctt	gtt	gga	gaa	gga	gga	ggc	cag	atg	agt	ggt	ggc	1728
Gln	Phe	Asp	Thr	Leu	Val	Gly	Glu	Gly	Gly	Gly	Gln	Met	Ser	Gly	Gly	
				565					570					575		
cag	aaa	caa	agg	gta	gct	atc	gcc	aga	gcc	ctc	atc	cga	aat	ccc	aag	1776
Gln	Lys	Gln	Arg	Val	Ala	Ile	Ala	Arg	Ala	Leu	Ile	Arg	Asn	Pro	Lys	
			580				585						590			
att	ctg	ctt	ttg	gac	atg	gcc	acc	tca	gct	ctg	gac	aat	gag	agt	gaa	1824
Ile	Leu	Leu	Leu	Asp	Met	Ala	Thr	Ser	Ala	Leu	Asp	Asn	Glu	Ser	Glu	
		595				600						605				
gcc	atg	gtg	caa	gaa	gtg	ctg	agt	aag	att	cag	cat	ggg	cac	aca	atc	1872
Ala	Met	Val	Gln	Glu	Val	Leu	Ser	Lys	Ile	Gln	His	Gly	His	Thr	Ile	
	610					615					620					
att	tca	gtt	gct	cat	cgc	ttg	tct	acg	gtc	aga	gct	gca	gat	acc	atc	1920
Ile	Ser	Val	Ala	His	Arg	Leu	Ser	Thr	Val	Arg	Ala	Ala	Asp	Thr	Ile	
	625				630					635					640	
att	ggt	ttt	gaa	cat	ggc	act	gca	gtg	gaa	aga	ggg	acc	cat	gaa	gaa	1968
Ile	Gly	Phe	Glu	His	Gly	Thr	Ala	Val	Glu	Arg	Gly	Thr	His	Glu	Glu	
				645					650					655		
tta	ctg	gaa	agg	aaa	ggt	gtt	tac	ttc	act	cta	gtg	act	ttg	caa	agc	2016
Leu	Leu	Glu	Arg	Lys	Gly	Val	Tyr	Phe	Thr	Leu	Val	Thr	Leu	Gln	Ser	
			660				665						670			
cag	gga	aat	caa	gct	ctt	aat	gaa	gag	gac	ata	aag	gat	gca	act	gaa	2064
Gln	Gly	Asn	Gln	Ala	Leu	Asn	Glu	Glu	Asp	Ile	Lys	Asp	Ala	Thr	Glu	
		675					680					685				
gat	gac	atg	ctt	gcg	agg	acc	ttt	agc	aga	ggg	agc	tac	cag	gat	agt	2112
Asp	Asp	Met	Leu	Ala	Arg	Thr	Phe	Ser	Arg	Gly	Ser	Tyr	Gln	Asp	Ser	
		690				695					700					
tta	agg	gct	tcc	atc	cgg	caa	cgc	tcc	aag	tct	cag	ctt	tct	tac	ctg	2160
Leu	Arg	Ala	Ser	Ile	Arg	Gln	Arg	Ser	Lys	Ser	Gln	Leu	Ser	Tyr	Leu	
	705				710				715						720	
gtg	cac	gaa	cct	cca	tta	gct	gtt	gta	gat	cat	aag	tct	acc	tat	gaa	2208
Val	His	Glu	Pro	Pro	Leu	Ala	Val	Val	Asp	His	Lys	Ser	Thr	Tyr	Glu	
				725					730					735		
gaa	gat	aga	aag	gac	aag	gac	att	cct	gtg	cag	gaa	gaa	gtt	gaa	cct	2256
Glu	Asp	Arg	Lys	Asp	Lys	Asp	Ile	Pro	Val	Gln	Glu	Glu	Val	Glu	Pro	
			740				745						750			
gcc	cca	gtt	agg	agg	att	ctg	aaa	ttc	agt	gct	cca	gaa	tgg	ccc	tac	2304
Ala	Pro	Val	Arg	Arg	Ile	Leu	Lys	Phe	Ser	Ala	Pro	Glu	Trp	Pro	Tyr	
		755				760						765				
atg	ctg	gta	ggg	tct	gtg	ggt	gca	gct	gtg	aac	ggg	aca	gtc	aca	ccc	2352
Met	Leu	Val	Gly	Ser	Val	Gly	Ala	Ala	Val	Asn	Gly	Thr	Val	Thr	Pro	

770	775	780	
ttg tat gcc ttt tta ttc agc cag att ctt ggg act ttt tca att cct			2400
Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr Phe Ser Ile Pro			
785	790	795	800
gat aaa gag gaa caa agg tca cag atc aat ggt gtg tgc cta ctt ttt			2448
Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val Cys Leu Leu Phe			
	805	810	815
gta gca atg ggc tgt gta tct ctt ttc acc caa ttt cta cag gga tat			2496
Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe Leu Gln Gly Tyr			
	820	825	830
gcc ttt gct aaa tct ggg gag ctc cta aca aaa agg cta cgt aaa ttt			2544
Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg Leu Arg Lys Phe			
	835	840	845
ggg ttc agg gca atg ctg ggg caa gat att gcc tgg ttt gat gac ctc			2592
Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp Phe Asp Asp Leu			
	850	855	860
aga aat agc cct gga gca ttg aca aca aga ctt gct aca gat gct tcc			2640
Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala Thr Asp Ala Ser			
	865	870	875
caa gtt caa ggg gct gcc ggc tct cag atc ggg atg ata gtc aat tcc			2688
Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met Ile Val Asn Ser			
	885	890	895
ttc act aac gtc act gtg gcc atg atc att gcc ttc tcc ttt agc tgg			2736
Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe Ser Phe Ser Trp			
	900	905	910
aag ctg agc ctg gtc atc ttg tgc ttc ttc ccc ttc ttg gct tta tca			2784
Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe Leu Ala Leu Ser			
	915	920	925
gga gcc aca cag acc agg atg ttg aca gga ttt gcc tct cga gat aag			2832
Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala Ser Arg Asp Lys			
	930	935	940
cag gcc ctg gag atg gtg gga cag att aca aat gaa gcc ctc agt aac			2880
Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu Ala Leu Ser Asn			
	945	950	955
atc cgc act gtt gct gga att gga aag gag agg cgg ttc att gaa gca			2928
Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg Phe Ile Glu Ala			
	965	970	975
ctt gag act gag ctg gag aag ccc ttc aag aca gcc att cag aaa gcc			2976
Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala Ile Gln Lys Ala			
	980	985	990
aat att tac gga ttc tgc ttt gcc ttt gcc cag tgc atc atg ttt att			3024
Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln Cys Ile Met Phe Ile			
	995	1000	1005
gcg aat tct gct tcc tac aga tat gga ggt tac tta atc tcc aat			3069
Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly Tyr Leu Ile Ser Asn			
	1010	1015	1020

gag ggg	ctc cat ttc agc tat	gtg ttc agg gtg atc	tct gca gtt	3114
Glu Gly	Leu His Phe Ser Tyr	Val Phe Arg Val Ile	Ser Ala Val	
1025	1030	1035		
gta ctg	agt gca aca gct ctt	gga aga gcc ttc tct	tac acc cca	3159
Val Leu	Ser Ala Thr Ala Leu	Gly Arg Ala Phe Ser	Tyr Thr Pro	
1040	1045	1050		
agt tat	gca aaa gct aaa ata	tca gct gca cgc ttt	ttt caa ctg	3204
Ser Tyr	Ala Lys Ala Lys Ile	Ser Ala Ala Arg Phe	Phe Gln Leu	
1055	1060	1065		
ctg gac	cga caa ccc cca atc	agt gta tac aat act	gca ggt gaa	3249
Leu Asp	Arg Gln Pro Pro Ile	Ser Val Tyr Asn Thr	Ala Gly Glu	
1070	1075	1080		
aaa tgg	gac aac ttc cag ggg	aag att gat ttt gtt	gat tgt aaa	3294
Lys Trp	Asp Asn Phe Gln Gly	Lys Ile Asp Phe Val	Asp Cys Lys	
1085	1090	1095		
ttt aca	tat cct tct cga cct	gac tcg caa gtt ctg	aat ggt ctc	3339
Phe Thr	Tyr Pro Ser Arg Pro	Asp Ser Gln Val Leu	Asn Gly Leu	
1100	1105	1110		
tca gtg	tcg att agt cca ggg	cag aca ctg gcg ttt	gtt ggg agc	3384
Ser Val	Ser Ile Ser Pro Gly	Gln Thr Leu Ala Phe	Val Gly Ser	
1115	1120	1125		
agt gga	tgt ggc aaa agc act	agc att cag ctg ttg	gaa cgt ttc	3429
Ser Gly	Cys Gly Lys Ser Thr	Ser Ile Gln Leu Leu	Glu Arg Phe	
1130	1135	1140		
tat gat	cct gat caa ggg aag	gtg atg ata gat ggt	cat gac agc	3474
Tyr Asp	Pro Asp Gln Gly Lys	Val Met Ile Asp Gly	His Asp Ser	
1145	1150	1155		
aaa aaa	gta aat gtc cag ttc	ctc cgc tca aac att	gga att gtt	3519
Lys Lys	Val Asn Val Gln Phe	Leu Arg Ser Asn Ile	Gly Ile Val	
1160	1165	1170		
tcc cag	gaa cca gtg ttg ttt	gcc tgt agc ata atg	gac aat atc	3564
Ser Gln	Glu Pro Val Leu Phe	Ala Cys Ser Ile Met	Asp Asn Ile	
1175	1180	1185		
aag tat	gga gac aac acc aaa	gaa att ccc atg gaa	aga gtc ata	3609
Lys Tyr	Gly Asp Asn Thr Lys	Glu Ile Pro Met Glu	Arg Val Ile	
1190	1195	1200		
gca gct	gca aaa cag gct cag	ctg cat gat ttt gtc	atg tca ctc	3654
Ala Ala	Ala Lys Gln Ala Gln	Leu His Asp Phe Val	Met Ser Leu	
1205	1210	1215		
cca gag	aaa tat gaa act aac	gtt ggg tcc cag ggg	tct caa ctc	3699
Pro Glu	Lys Tyr Glu Thr Asn	Val Gly Ser Gln Gly	Ser Gln Leu	
1220	1225	1230		
tct aga	ggg gag aaa caa cgc	att gct att gct cgg	gcc att gta	3744
Ser Arg	Gly Glu Lys Gln Arg	Ile Ala Ile Ala Arg	Ala Ile Val	
1235	1240	1245		

cga gat cct aaa atc ttg cta cta gat gaa gcc act tct gcc tta 3789
 Arg Asp Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu
 1250 1255 1260

 gac aca gaa agt gaa aag acg gtg cag gtt gct cta gac aaa gcc 3834
 Asp Thr Glu Ser Glu Lys Thr Val Gln Val Ala Leu Asp Lys Ala
 1265 1270 1275

 aga gag ggt cgg acc tgc att gtc att gcc cat cgc ttg tcc acc 3879
 Arg Glu Gly Arg Thr Cys Ile Val Ile Ala His Arg Leu Ser Thr
 1280 1285 1290

 atc cag aac gcg gat atc att gct gtc atg gca cag ggg gtg gtg 3924
 Ile Gln Asn Ala Asp Ile Ile Ala Val Met Ala Gln Gly Val Val
 1295 1300 1305

 att gaa aag ggg acc cat gaa gaa ctg atg gcc caa aaa gga gcc 3969
 Ile Glu Lys Gly Thr His Glu Glu Leu Met Ala Gln Lys Gly Ala
 1310 1315 1320

 tac tac aaa cta gtc acc act gga tcc ccc atc agt tga 4008
 Tyr Tyr Lys Leu Val Thr Thr Gly Ser Pro Ile Ser
 1325 1330 1335

<210> 4
 <211> 1335
 <212> PRT
 <213> Homo sapiens

<400> 4

Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu
1 5 10 15

Asn Asp Gly Phe Glu Ser Asp Lys Ser Cys Glu Trp Leu Phe Ser Leu
20 25 30

Thr Ala Ser Cys Thr Arg Arg Asp Asn Asn Asp Lys Lys Ser Arg Leu
35 40 45

Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly Phe Phe Gln Leu
50 55 60

Phe Arg Phe Ser Ser Ser Thr Asp Ile Trp Leu Met Phe Val Gly Ser
65 70 75 80

Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly Val Leu Leu Ile
85 90 95

Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp Val Glu Leu Gln
100 105 110

Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn Thr Ile Val Trp

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

Thr Pro Gly Thr Leu Val Gln Ile Phe Leu Ser Val Ile Val Gly Ala
 370 375 380

Leu Asn Leu Gly Asn Ala Ser Pro Cys Leu Glu Ala Phe Ala Thr Gly
 385 390 395 400

Arg Ala Ala Ala Thr Ser Ile Phe Glu Thr Ile Asp Arg Lys Pro Ile
 405 410 415

Ile Asp Cys Met Ser Glu Asp Gly Tyr Lys Leu Asp Arg Ile Lys Gly
 420 425 430

Glu Ile Glu Phe His Asn Val Thr Phe His Tyr Pro Ser Arg Pro Glu
 435 440 445

Val Lys Ile Leu Asn Asp Leu Asn Met Val Ile Lys Pro Gly Glu Met
 450 455 460

Thr Ala Leu Val Gly Pro Ser Gly Ala Gly Lys Ser Thr Ala Leu Gln
 465 470 475 480

Leu Ile Gln Arg Phe Tyr Asp Pro Cys Glu Gly Met Val Thr Val Asp
 485 490 495

Gly His Asp Ile Arg Ser Leu Asn Ile Gln Trp Leu Arg Asp Gln Ile
 500 505 510

Gly Ile Val Glu Gln Glu Pro Val Leu Phe Ser Thr Thr Ile Ala Glu
 515 520 525

Asn Ile Arg Tyr Gly Arg Glu Asp Ala Thr Met Glu Asp Ile Val Gln
 530 535 540

Ala Ala Lys Glu Ala Asn Ala Tyr Asn Phe Ile Met Asp Leu Pro Gln
 545 550 555 560

Gln Phe Asp Thr Leu Val Gly Glu Gly Gly Gly Gln Met Ser Gly Gly
 565 570 575

Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Ile Arg Asn Pro Lys
 580 585 590

Ile Leu Leu Leu Asp Met Ala Thr Ser Ala Leu Asp Asn Glu Ser Glu
 595 600 605

Ala Met Val Gln Glu Val Leu Ser Lys Ile Gln His Gly His Thr Ile
610 615 620

Ile Ser Val Ala His Arg Leu Ser Thr Val Arg Ala Ala Asp Thr Ile
625 630 635 640

Ile Gly Phe Glu His Gly Thr Ala Val Glu Arg Gly Thr His Glu Glu
645 650 655

Leu Leu Glu Arg Lys Gly Val Tyr Phe Thr Leu Val Thr Leu Gln Ser
660 665 670

Gln Gly Asn Gln Ala Leu Asn Glu Glu Asp Ile Lys Asp Ala Thr Glu
675 680 685

Asp Asp Met Leu Ala Arg Thr Phe Ser Arg Gly Ser Tyr Gln Asp Ser
690 695 700

Leu Arg Ala Ser Ile Arg Gln Arg Ser Lys Ser Gln Leu Ser Tyr Leu
705 710 715 720

Val His Glu Pro Pro Leu Ala Val Val Asp His Lys Ser Thr Tyr Glu
725 730 735

Glu Asp Arg Lys Asp Lys Asp Ile Pro Val Gln Glu Glu Val Glu Pro
740 745 750

Ala Pro Val Arg Arg Ile Leu Lys Phe Ser Ala Pro Glu Trp Pro Tyr
755 760 765

Met Leu Val Gly Ser Val Gly Ala Ala Val Asn Gly Thr Val Thr Pro
770 775 780

Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr Phe Ser Ile Pro
785 790 795 800

Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val Cys Leu Leu Phe
805 810 815

Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe Leu Gln Gly Tyr
820 825 830

Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg Leu Arg Lys Phe
835 840 845

Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp Phe Asp Asp Leu
850 855 860

Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala Thr Asp Ala Ser
865 870 875 880

Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met Ile Val Asn Ser
885 890 895

Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe Ser Phe Ser Trp
900 905 910

Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe Leu Ala Leu Ser
915 920 925

Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala Ser Arg Asp Lys
930 935 940

Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu Ala Leu Ser Asn
945 950 955 960

Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg Phe Ile Glu Ala
965 970 975

Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala Ile Gln Lys Ala
980 985 990

Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln Cys Ile Met Phe Ile
995 1000 1005

Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly Tyr Leu Ile Ser Asn
1010 1015 1020

Glu Gly Leu His Phe Ser Tyr Val Phe Arg Val Ile Ser Ala Val
1025 1030 1035

Val Leu Ser Ala Thr Ala Leu Gly Arg Ala Phe Ser Tyr Thr Pro
1040 1045 1050

Ser Tyr Ala Lys Ala Lys Ile Ser Ala Ala Arg Phe Phe Gln Leu
1055 1060 1065

Leu Asp Arg Gln Pro Pro Ile Ser Val Tyr Asn Thr Ala Gly Glu
1070 1075 1080

Lys Trp Asp Asn Phe Gln Gly Lys Ile Asp Phe Val Asp Cys Lys

1085	1090	1095
Phe Thr Tyr Pro Ser Arg Pro Asp Ser Gln Val Leu Asn Gly Leu 1100 1105 1110		
Ser Val Ser Ile Ser Pro Gly Gln Thr Leu Ala Phe Val Gly Ser 1115 1120 1125		
Ser Gly Cys Gly Lys Ser Thr Ser Ile Gln Leu Leu Glu Arg Phe 1130 1135 1140		
Tyr Asp Pro Asp Gln Gly Lys Val Met Ile Asp Gly His Asp Ser 1145 1150 1155		
Lys Lys Val Asn Val Gln Phe Leu Arg Ser Asn Ile Gly Ile Val 1160 1165 1170		
Ser Gln Glu Pro Val Leu Phe Ala Cys Ser Ile Met Asp Asn Ile 1175 1180 1185		
Lys Tyr Gly Asp Asn Thr Lys Glu Ile Pro Met Glu Arg Val Ile 1190 1195 1200		
Ala Ala Ala Lys Gln Ala Gln Leu His Asp Phe Val Met Ser Leu 1205 1210 1215		
Pro Glu Lys Tyr Glu Thr Asn Val Gly Ser Gln Gly Ser Gln Leu 1220 1225 1230		
Ser Arg Gly Glu Lys Gln Arg Ile Ala Ile Ala Arg Ala Ile Val 1235 1240 1245		
Arg Asp Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu 1250 1255 1260		
Asp Thr Glu Ser Glu Lys Thr Val Gln Val Ala Leu Asp Lys Ala 1265 1270 1275		
Arg Glu Gly Arg Thr Cys Ile Val Ile Ala His Arg Leu Ser Thr 1280 1285 1290		
Ile Gln Asn Ala Asp Ile Ile Ala Val Met Ala Gln Gly Val Val 1295 1300 1305		
Ile Glu Lys Gly Thr His Glu Glu Leu Met Ala Gln Lys Gly Ala 1310 1315 1320		

Tyr Tyr Lys Leu Val Thr Thr Gly Ser Pro Ile Ser
 1325 1330 1335

<210> 5
 <211> 42
 <212> DNA
 <213> Homo sapiens

<400> 5
 gtgagtggct tttttccctc actgcatctt gtacaaggag ag 42

<210> 6
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 6

Cys Glu Trp Leu Phe Ser Leu Thr Ala Ser Cys Thr Arg Arg Asp
 1 5 10 15

<210> 7
 <211> 4068
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(4068)

<400> 7
 atg tct gac tca gta att ctt cga agt ata aag aaa ttt gga gag gag 48
 Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu
 1 5 10 15

aat gat ggt ttt gag tca gat aaa tca tgt gag tgg ctt ttt tcc ctc 96
 Asn Asp Gly Phe Glu Ser Asp Lys Ser Cys Glu Trp Leu Phe Ser Leu
 20 25 30

act gca tct tgt aca agg aga ggt gag aac aaa agt agg aca agc tgg 144
 Thr Ala Ser Cys Thr Arg Arg Gly Glu Asn Lys Ser Arg Thr Ser Trp
 35 40 45

tca agt ttc aag gag cag aaa aaa atc agc aac aat aat aat gat aag 192
 Ser Ser Phe Lys Glu Gln Lys Lys Ile Ser Asn Asn Asn Asn Asp Lys
 50 55 60

aaa tca agg tta caa gat gag aag aaa ggt gat ggc gtt aga gtt ggc 240
 Lys Ser Arg Leu Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly
 65 70 75 80

ttc ttt caa ttg ttt cgg ttt tct tca tca act gac att tgg ctg atg 288
 Phe Phe Gln Leu Phe Arg Phe Ser Ser Ser Thr Asp Ile Trp Leu Met
 85 90 95

ttt gtg gga agt ttg tgt gca ttt ctc cat gga ata gcc cag cca ggc	336
Phe Val Gly Ser Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly	
100 105 110	
gtg cta ctc att ttt ggc aca atg aca gat gtt ttt att gac tac gac	384
Val Leu Leu Ile Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp	
115 120 125	
gtt gag tta caa gaa ctc cag att cca gga aaa gca tgt gtg aat aac	432
Val Glu Leu Gln Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn	
130 135 140	
acc att gta tgg act aac agt tcc ctc aac cag aac atg aca aat gga	480
Thr Ile Val Trp Thr Asn Ser Ser Leu Asn Gln Asn Met Thr Asn Gly	
145 150 155 160	
aca cgt tgt ggg ttg ctg aac atc gag agc gaa atg atc aaa ttt gcc	528
Thr Arg Cys Gly Leu Leu Asn Ile Glu Ser Glu Met Ile Lys Phe Ala	
165 170 175	
agt tac tat gct gga att gct gtc gca gta ctt atc aca gga tat att	576
Ser Tyr Tyr Ala Gly Ile Ala Val Ala Val Leu Ile Thr Gly Tyr Ile	
180 185 190	
caa ata tgc ttt tgg gtc att gcc gca gct cgt cag ata cag aaa atg	624
Gln Ile Cys Phe Trp Val Ile Ala Ala Ala Arg Gln Ile Gln Lys Met	
195 200 205	
aga aaa ttt tac ttt agg aga ata atg aga atg gaa ata ggg tgg ttt	672
Arg Lys Phe Tyr Phe Arg Arg Ile Met Arg Met Glu Ile Gly Trp Phe	
210 215 220	
gac tgc aat tca gtg ggg gag ctg aat aca aga ttc tct gat gat att	720
Asp Cys Asn Ser Val Gly Glu Leu Asn Thr Arg Phe Ser Asp Asp Ile	
225 230 235 240	
aat aaa atc aat gat gcc ata gct gac caa atg gcc ctt ttc att cag	768
Asn Lys Ile Asn Asp Ala Ile Ala Asp Gln Met Ala Leu Phe Ile Gln	
245 250 255	
cgc atg acc tcg acc atc tgt ggt ttc ctg ttg gga ttt ttc agg ggt	816
Arg Met Thr Ser Thr Ile Cys Gly Phe Leu Leu Gly Phe Phe Arg Gly	
260 265 270	
tgg aaa ctg acc ttg gtt att att tct gtc agc cct ctc att ggg att	864
Trp Lys Leu Thr Leu Val Ile Ile Ser Val Ser Pro Leu Ile Gly Ile	
275 280 285	
gga gca gcc acc att ggt ctg agt gtg tcc aag ttt acg gac tat gag	912
Gly Ala Ala Thr Ile Gly Leu Ser Val Ser Lys Phe Thr Asp Tyr Glu	
290 295 300	
ctg aag gcc tat gcc aaa gca ggg gtg gtg gct gat gaa gtc att tca	960
Leu Lys Ala Tyr Ala Lys Ala Gly Val Val Ala Asp Glu Val Ile Ser	
305 310 315 320	
tca atg aga aca gtg gct gct ttt ggt ggt gag aaa aga gag gtt gaa	1008
Ser Met Arg Thr Val Ala Ala Phe Gly Gly Glu Lys Arg Glu Val Glu	
325 330 335	
agg tat gag aaa aat ctt gtg ttc gcc cag cgt tgg gga att aga aaa	1056

Arg	Tyr	Glu	Lys	Asn	Leu	Val	Phe	Ala	Gln	Arg	Trp	Gly	Ile	Arg	Lys	
			340					345					350			
gga	ata	gtg	atg	gga	ttc	ttt	act	gga	ttc	gtg	tgg	tgt	ctc	atc	ttt	1104
Gly	Ile	Val	Met	Gly	Phe	Phe	Thr	Gly	Phe	Val	Trp	Cys	Leu	Ile	Phe	
		355					360					365				
ttg	tgt	tat	gca	ctg	gcc	ttc	tgg	tac	ggc	tcc	aca	ctt	gtc	ctg	gat	1152
Leu	Cys	Tyr	Ala	Leu	Ala	Phe	Trp	Tyr	Gly	Ser	Thr	Leu	Val	Leu	Asp	
	370					375					380					
gaa	gga	gaa	tat	aca	cca	gga	acc	ctt	gtc	cag	att	ttc	ctc	agt	gtc	1200
Glu	Gly	Glu	Tyr	Thr	Pro	Gly	Thr	Leu	Val	Gln	Ile	Phe	Leu	Ser	Val	
385					390					395					400	
ata	gta	gga	gct	tta	aat	ctt	ggc	aat	gcc	tct	cct	tgt	ttg	gaa	gcc	1248
Ile	Val	Gly	Ala	Leu	Asn	Leu	Gly	Asn	Ala	Ser	Pro	Cys	Leu	Glu	Ala	
			405					410						415		
ttt	gca	act	gga	cgt	gca	gca	gcc	acc	agc	att	ttt	gag	aca	ata	gac	1296
Phe	Ala	Thr	Gly	Arg	Ala	Ala	Ala	Thr	Ser	Ile	Phe	Glu	Thr	Ile	Asp	
		420					425						430			
agg	aaa	ccc	atc	att	gac	tgc	atg	tca	gaa	gat	ggc	tac	aag	ttg	gat	1344
Arg	Lys	Pro	Ile	Ile	Asp	Cys	Met	Ser	Glu	Asp	Gly	Tyr	Lys	Leu	Asp	
		435				440						445				
cga	atc	aag	ggt	gaa	att	gaa	ttc	cat	aat	gtg	acc	ttc	cat	tat	cct	1392
Arg	Ile	Lys	Gly	Glu	Ile	Glu	Phe	His	Asn	Val	Thr	Phe	His	Tyr	Pro	
	450					455					460					
tcc	aga	cca	gag	gtg	aag	att	cta	aat	gac	ctc	aac	atg	gtc	att	aaa	1440
Ser	Arg	Pro	Glu	Val	Lys	Ile	Leu	Asn	Asp	Leu	Asn	Met	Val	Ile	Lys	
465					470					475					480	
cca	ggg	gaa	atg	aca	gct	ctg	gta	gga	ccc	agt	gga	gct	gga	aaa	agt	1488
Pro	Gly	Glu	Met	Thr	Ala	Leu	Val	Gly	Pro	Ser	Gly	Ala	Gly	Lys	Ser	
			485					490						495		
aca	gca	ctg	caa	ctc	att	cag	cga	ttc	tat	gac	ccc	tgt	gaa	gga	atg	1536
Thr	Ala	Leu	Gln	Leu	Ile	Gln	Arg	Phe	Tyr	Asp	Pro	Cys	Glu	Gly	Met	
		500						505					510			
gtg	acc	gtg	gat	ggc	cat	gac	att	cgc	tct	ctt	aac	att	cag	tgg	ctt	1584
Val	Thr	Val	Asp	Gly	His	Asp	Ile	Arg	Ser	Leu	Asn	Ile	Gln	Trp	Leu	
		515				520						525				
aga	gat	cag	att	ggg	ata	gtg	gag	caa	gag	cca	gtt	ctg	ttc	tct	acc	1632
Arg	Asp	Gln	Ile	Gly	Ile	Val	Glu	Gln	Glu	Pro	Val	Leu	Phe	Ser	Thr	
	530					535					540					
acc	att	gca	gaa	aat	att	cgc	tat	ggc	aga	gaa	gat	gca	aca	atg	gaa	1680
Thr	Ile	Ala	Glu	Asn	Ile	Arg	Tyr	Gly	Arg	Glu	Asp	Ala	Thr	Met	Glu	
545					550					555					560	
gac	ata	gtc	caa	gct	gcc	aag	gag	gcc	aat	gcc	tac	aac	ttc	atc	atg	1728
Asp	Ile	Val	Gln	Ala	Ala	Lys	Glu	Ala	Asn	Ala	Tyr	Asn	Phe	Ile	Met	
			565					570					575			
gac	ctg	cca	cag	caa	ttt	gac	acc	ctt	gtt	gga	gaa	gga	gga	ggc	cag	1776
Asp	Leu	Pro	Gln	Gln	Phe	Asp	Thr	Leu	Val	Gly	Glu	Gly	Gly	Gly	Gln	

580	585	590	
atg agt ggt ggc cag aaa caa agg gta gct atc gcc aga gcc ctc atc Met Ser Gly Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Ile 595 600 605			1824
cga aat ccc aag att ctg ctt ttg gac atg gcc acc tca gct ctg gac Arg Asn Pro Lys Ile Leu Leu Leu Asp Met Ala Thr Ser Ala Leu Asp 610 615 620			1872
aat gag agt gaa gcc atg gtg caa gaa gtg ctg agt aag att cag cat Asn Glu Ser Glu Ala Met Val Gln Glu Val Leu Ser Lys Ile Gln His 625 630 635 640			1920
ggg cac aca atc att tca gtt gct cat cgc ttg tct acg gtc aga gct Gly His Thr Ile Ile Ser Val Ala His Arg Leu Ser Thr Val Arg Ala 645 650 655			1968
gca gat acc atc att ggt ttt gaa cat ggc act gca gtg gaa aga ggg Ala Asp Thr Ile Ile Gly Phe Glu His Gly Thr Ala Val Glu Arg Gly 660 665 670			2016
acc cat gaa gaa tta ctg gaa agg aaa ggt gtt tac ttc act cta gtg Thr His Glu Glu Leu Leu Glu Arg Lys Gly Val Tyr Phe Thr Leu Val 675 680 685			2064
act ttg caa agc cag gga aat caa gct ctt aat gaa gag gac ata aag Thr Leu Gln Ser Gln Gly Asn Gln Ala Leu Asn Glu Glu Asp Ile Lys 690 695 700			2112
gat gca act gaa gat gac atg ctt gcg agg acc ttt agc aga ggg agc Asp Ala Thr Glu Asp Asp Met Leu Ala Arg Thr Phe Ser Arg Gly Ser 705 710 715 720			2160
tac cag gat agt tta agg gct tcc atc cgg caa cgc tcc aag tct cag Tyr Gln Asp Ser Leu Arg Ala Ser Ile Arg Gln Arg Ser Lys Ser Gln 725 730 735			2208
ctt tct tac ctg gtg cac gaa cct cca tta gct gtt gta gat cat aag Leu Ser Tyr Leu Val His Glu Pro Pro Leu Ala Val Val Asp His Lys 740 745 750			2256
tct acc tat gaa gaa gat aga aag gac aag gac att cct gtg cag gaa Ser Thr Tyr Glu Glu Asp Arg Lys Asp Lys Asp Ile Pro Val Gln Glu 755 760 765			2304
gaa gtt gaa cct gcc cca gtt agg agg att ctg aaa ttc agt gct cca Glu Val Glu Pro Ala Pro Val Arg Arg Ile Leu Lys Phe Ser Ala Pro 770 775 780			2352
gaa tgg ccc tac atg ctg gta ggg tct gtg ggt gca gct gtg aac ggg Glu Trp Pro Tyr Met Leu Val Gly Ser Val Gly Ala Ala Val Asn Gly 785 790 795 800			2400
aca gtc aca ccc ttg tat gcc ttt tta ttc agc cag att ctt ggg act Thr Val Thr Pro Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr 805 810 815			2448
ttt tca att cct gat aaa gag gaa caa agg tca cag atc aat ggt gtg Phe Ser Ile Pro Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val 820 825 830			2496

tgc cta ctt ttt gta gca atg ggc tgt gta tct ctt ttc acc caa ttt	2544
Cys Leu Leu Phe Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe	
835 840 845	
cta cag gga tat gcc ttt gct aaa tct ggg gag ctc cta aca aaa agg	2592
Leu Gln Gly Tyr Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg	
850 855 860	
cta cgt aaa ttt ggt ttc agg gca atg ctg ggg caa gat att gcc tgg	2640
Leu Arg Lys Phe Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp	
865 870 875 880	
ttt gat gac ctc aga aat agc cct gga gca ttg aca aca aga ctt gct	2688
Phe Asp Asp Leu Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala	
885 890 895	
aca gat gct tcc caa gtt caa ggg gct gcc ggc tct cag atc ggg atg	2736
Thr Asp Ala Ser Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met	
900 905 910	
ata gtc aat tcc ttc act aac gtc act gtg gcc atg atc att gcc ttc	2784
Ile Val Asn Ser Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe	
915 920 925	
tcc ttt agc tgg aag ctg agc ctg gtc atc ttg tgc ttc ttc ccc ttc	2832
Ser Phe Ser Trp Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe	
930 935 940	
ttg gct tta tca gga gcc aca cag acc agg atg ttg aca gga ttt gcc	2880
Leu Ala Leu Ser Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala	
945 950 955 960	
tct cga gat aag cag gcc ctg gag atg gtg gga cag att aca aat gaa	2928
Ser Arg Asp Lys Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu	
965 970 975	
gcc ctc agt aac atc cgc act gtt gct gga att gga aag gag agg cgg	2976
Ala Leu Ser Asn Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg	
980 985 990	
ttc att gaa gca ctt gag act gag ctg gag aag ccc ttc aag aca gcc	3024
Phe Ile Glu Ala Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala	
995 1000 1005	
att cag aaa gcc aat att tac gga ttc tgc ttt gcc ttt gcc cag	3069
Ile Gln Lys Ala Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln	
1010 1015 1020	
tgc atc atg ttt att gcg aat tct gct tcc tac aga tat gga ggt	3114
Cys Ile Met Phe Ile Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly	
1025 1030 1035	
tac tta atc tcc aat gag ggg ctc cat ttc agc tat gtg ttc agg	3159
Tyr Leu Ile Ser Asn Glu Gly Leu His Phe Ser Tyr Val Phe Arg	
1040 1045 1050	
gtg atc tct gca gtt gta ctg agt gca aca gct ctt gga aga gcc	3204
Val Ile Ser Ala Val Val Leu Ser Ala Thr Ala Leu Gly Arg Ala	
1055 1060 1065	

ttc tct tac acc cca agt tat gca aaa gct aaa ata tca gct gca	3249
Phe Ser Tyr Thr Pro Ser Tyr Ala Lys Ala Lys Ile Ser Ala Ala	
1070 1075 1080	
cgc ttt ttt caa ctg ctg gac cga caa ccc cca atc agt gta tac	3294
Arg Phe Phe Gln Leu Leu Asp Arg Gln Pro Pro Ile Ser Val Tyr	
1085 1090 1095	
aat act gca ggt gaa aaa tgg gac aac ttc cag ggg aag att gat	3339
Asn Thr Ala Gly Glu Lys Trp Asp Asn Phe Gln Gly Lys Ile Asp	
1100 1105 1110	
ttt gtt gat tgt aaa ttt aca tat cct tct cga cct gac tcg caa	3384
Phe Val Asp Cys Lys Phe Thr Tyr Pro Ser Arg Pro Asp Ser Gln	
1115 1120 1125	
gtt ctg aat ggt ctc tca gtg tcg att agt cca ggg cag aca ctg	3429
Val Leu Asn Gly Leu Ser Val Ser Ile Ser Pro Gly Gln Thr Leu	
1130 1135 1140	
gcg ttt gtt ggg agc agt gga tgt ggc aaa agc act agc att cag	3474
Ala Phe Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Ser Ile Gln	
1145 1150 1155	
ctg ttg gaa cgt ttc tat gat cct gat caa ggg aag gtg atg ata	3519
Leu Leu Glu Arg Phe Tyr Asp Pro Asp Gln Gly Lys Val Met Ile	
1160 1165 1170	
gat ggt cat gac agc aaa aaa gta aat gtc cag ttc ctc cgc tca	3564
Asp Gly His Asp Ser Lys Lys Val Asn Val Gln Phe Leu Arg Ser	
1175 1180 1185	
aac att gga att gtt tcc cag gaa cca gtg ttg ttt gcc tgt agc	3609
Asn Ile Gly Ile Val Ser Gln Glu Pro Val Leu Phe Ala Cys Ser	
1190 1195 1200	
ata atg gac aat atc aag tat gga gac aac acc aaa gaa att ccc	3654
Ile Met Asp Asn Ile Lys Tyr Gly Asp Asn Thr Lys Glu Ile Pro	
1205 1210 1215	
atg gaa aga gtc ata gca gct gca aaa cag gct cag ctg cat gat	3699
Met Glu Arg Val Ile Ala Ala Ala Lys Gln Ala Gln Leu His Asp	
1220 1225 1230	
ttt gtc atg tca ctc cca gag aaa tat gaa act aac gtt ggg tcc	3744
Phe Val Met Ser Leu Pro Glu Lys Tyr Glu Thr Asn Val Gly Ser	
1235 1240 1245	
cag ggg tct caa ctc tct aga ggg gag aaa caa cgc att gct att	3789
Gln Gly Ser Gln Leu Ser Arg Gly Glu Lys Gln Arg Ile Ala Ile	
1250 1255 1260	
gct cgg gcc att gta cga gat cct aaa atc ttg cta cta gat gaa	3834
Ala Arg Ala Ile Val Arg Asp Pro Lys Ile Leu Leu Leu Asp Glu	
1265 1270 1275	
gcc act tct gcc tta gac aca gaa agt gaa aag acg gtg cag gtt	3879
Ala Thr Ser Ala Leu Asp Thr Glu Ser Glu Lys Thr Val Gln Val	
1280 1285 1290	
gct cta gac aaa gcc aga gag ggt cgg acc tgc att gtc att gcc	3924

Ala Leu Asp Lys Ala Arg Glu Gly Arg Thr Cys Ile Val Ile Ala
 1295 1300 1305

cat cgc ttg tcc acc atc cag aac gcg gat atc att gct gtc atg 3969
 His Arg Leu Ser Thr Ile Gln Asn Ala Asp Ile Ile Ala Val Met
 1310 1315 1320

gca cag ggg gtg gtg att gaa aag ggg acc cat gaa gaa ctg atg 4014
 Ala Gln Gly Val Val Ile Glu Lys Gly Thr His Glu Glu Leu Met
 1325 1330 1335

gcc caa aaa gga gcc tac tac aaa cta gtc acc act gga tcc ccc 4059
 Ala Gln Lys Gly Ala Tyr Tyr Lys Leu Val Thr Thr Gly Ser Pro
 1340 1345 1350

atc agt tga 4068
 Ile Ser
 1355

<210> 8
 <211> 1355
 <212> PRT
 <213> Homo sapiens

<400> 8

Met Ser Asp Ser Val Ile Leu Arg Ser Ile Lys Lys Phe Gly Glu Glu
 1 5 10 15

Asn Asp Gly Phe Glu Ser Asp Lys Ser Cys Glu Trp Leu Phe Ser Leu
 20 25 30

Thr Ala Ser Cys Thr Arg Arg Gly Glu Asn Lys Ser Arg Thr Ser Trp
 35 40 45

Ser Ser Phe Lys Glu Gln Lys Lys Ile Ser Asn Asn Asn Asp Lys
 50 55 60

Lys Ser Arg Leu Gln Asp Glu Lys Lys Gly Asp Gly Val Arg Val Gly
 65 70 75 80

Phe Phe Gln Leu Phe Arg Phe Ser Ser Ser Thr Asp Ile Trp Leu Met
 85 90 95

Phe Val Gly Ser Leu Cys Ala Phe Leu His Gly Ile Ala Gln Pro Gly
 100 105 110

Val Leu Leu Ile Phe Gly Thr Met Thr Asp Val Phe Ile Asp Tyr Asp
 115 120 125

Val Glu Leu Gln Glu Leu Gln Ile Pro Gly Lys Ala Cys Val Asn Asn
 130 135 140

Thr Ile Val Trp Thr Asn Ser Ser Leu Asn Gln Asn Met Thr Asn Gly
145 150 155 160

Thr Arg Cys Gly Leu Leu Asn Ile Glu Ser Glu Met Ile Lys Phe Ala
165 170 175

Ser Tyr Tyr Ala Gly Ile Ala Val Ala Val Leu Ile Thr Gly Tyr Ile
180 185 190

Gln Ile Cys Phe Trp Val Ile Ala Ala Ala Arg Gln Ile Gln Lys Met
195 200 205

Arg Lys Phe Tyr Phe Arg Arg Ile Met Arg Met Glu Ile Gly Trp Phe
210 215 220

Asp Cys Asn Ser Val Gly Glu Leu Asn Thr Arg Phe Ser Asp Asp Ile
225 230 235 240

Asn Lys Ile Asn Asp Ala Ile Ala Asp Gln Met Ala Leu Phe Ile Gln
245 250 255

Arg Met Thr Ser Thr Ile Cys Gly Phe Leu Leu Gly Phe Phe Arg Gly
260 265 270

Trp Lys Leu Thr Leu Val Ile Ile Ser Val Ser Pro Leu Ile Gly Ile
275 280 285

Gly Ala Ala Thr Ile Gly Leu Ser Val Ser Lys Phe Thr Asp Tyr Glu
290 295 300

Leu Lys Ala Tyr Ala Lys Ala Gly Val Val Ala Asp Glu Val Ile Ser
305 310 315 320

Ser Met Arg Thr Val Ala Ala Phe Gly Gly Glu Lys Arg Glu Val Glu
325 330 335

Arg Tyr Glu Lys Asn Leu Val Phe Ala Gln Arg Trp Gly Ile Arg Lys
340 345 350

Gly Ile Val Met Gly Phe Phe Thr Gly Phe Val Trp Cys Leu Ile Phe
355 360 365

Leu Cys Tyr Ala Leu Ala Phe Trp Tyr Gly Ser Thr Leu Val Leu Asp
370 375 380

Glu Gly Glu Tyr Thr Pro Gly Thr Leu Val Gln Ile Phe Leu Ser Val
385 390 395 400

Ile Val Gly Ala Leu Asn Leu Gly Asn Ala Ser Pro Cys Leu Glu Ala
405 410 415

Phe Ala Thr Gly Arg Ala Ala Ala Thr Ser Ile Phe Glu Thr Ile Asp
420 425 430

Arg Lys Pro Ile Ile Asp Cys Met Ser Glu Asp Gly Tyr Lys Leu Asp
435 440 445

Arg Ile Lys Gly Glu Ile Glu Phe His Asn Val Thr Phe His Tyr Pro
450 455 460

Ser Arg Pro Glu Val Lys Ile Leu Asn Asp Leu Asn Met Val Ile Lys
465 470 475 480

Pro Gly Glu Met Thr Ala Leu Val Gly Pro Ser Gly Ala Gly Lys Ser
485 490 495

Thr Ala Leu Gln Leu Ile Gln Arg Phe Tyr Asp Pro Cys Glu Gly Met
500 505 510

Val Thr Val Asp Gly His Asp Ile Arg Ser Leu Asn Ile Gln Trp Leu
515 520 525

Arg Asp Gln Ile Gly Ile Val Glu Gln Glu Pro Val Leu Phe Ser Thr
530 535 540

Thr Ile Ala Glu Asn Ile Arg Tyr Gly Arg Glu Asp Ala Thr Met Glu
545 550 555 560

Asp Ile Val Gln Ala Ala Lys Glu Ala Asn Ala Tyr Asn Phe Ile Met
565 570 575

Asp Leu Pro Gln Gln Phe Asp Thr Leu Val Gly Glu Gly Gly Gly Gln
580 585 590

Met Ser Gly Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Ile
595 600 605

Arg Asn Pro Lys Ile Leu Leu Leu Asp Met Ala Thr Ser Ala Leu Asp
610 615 620

Asn Glu Ser Glu Ala Met Val Gln Glu Val Leu Ser Lys Ile Gln His
625 630 635 640

Gly His Thr Ile Ile Ser Val Ala His Arg Leu Ser Thr Val Arg Ala
645 650 655

Ala Asp Thr Ile Ile Gly Phe Glu His Gly Thr Ala Val Glu Arg Gly
660 665 670

Thr His Glu Glu Leu Leu Glu Arg Lys Gly Val Tyr Phe Thr Leu Val
675 680 685

Thr Leu Gln Ser Gln Gly Asn Gln Ala Leu Asn Glu Glu Asp Ile Lys
690 695 700

Asp Ala Thr Glu Asp Asp Met Leu Ala Arg Thr Phe Ser Arg Gly Ser
705 710 715 720

Tyr Gln Asp Ser Leu Arg Ala Ser Ile Arg Gln Arg Ser Lys Ser Gln
725 730 735

Leu Ser Tyr Leu Val His Glu Pro Pro Leu Ala Val Val Asp His Lys
740 745 750

Ser Thr Tyr Glu Glu Asp Arg Lys Asp Lys Asp Ile Pro Val Gln Glu
755 760 765

Glu Val Glu Pro Ala Pro Val Arg Arg Ile Leu Lys Phe Ser Ala Pro
770 775 780

Glu Trp Pro Tyr Met Leu Val Gly Ser Val Gly Ala Ala Val Asn Gly
785 790 795 800

Thr Val Thr Pro Leu Tyr Ala Phe Leu Phe Ser Gln Ile Leu Gly Thr
805 810 815

Phe Ser Ile Pro Asp Lys Glu Glu Gln Arg Ser Gln Ile Asn Gly Val
820 825 830

Cys Leu Leu Phe Val Ala Met Gly Cys Val Ser Leu Phe Thr Gln Phe
835 840 845

Leu Gln Gly Tyr Ala Phe Ala Lys Ser Gly Glu Leu Leu Thr Lys Arg
850 855 860

Leu Arg Lys Phe Gly Phe Arg Ala Met Leu Gly Gln Asp Ile Ala Trp

865		870		875		880
Phe Asp Asp Leu Arg Asn Ser Pro Gly Ala Leu Thr Thr Arg Leu Ala						
	885			890		895
Thr Asp Ala Ser Gln Val Gln Gly Ala Ala Gly Ser Gln Ile Gly Met						
	900			905		910
Ile Val Asn Ser Phe Thr Asn Val Thr Val Ala Met Ile Ile Ala Phe						
	915			920		925
Ser Phe Ser Trp Lys Leu Ser Leu Val Ile Leu Cys Phe Phe Pro Phe						
	930			935		940
Leu Ala Leu Ser Gly Ala Thr Gln Thr Arg Met Leu Thr Gly Phe Ala						
	945			950		955
						960
Ser Arg Asp Lys Gln Ala Leu Glu Met Val Gly Gln Ile Thr Asn Glu						
	965			970		975
Ala Leu Ser Asn Ile Arg Thr Val Ala Gly Ile Gly Lys Glu Arg Arg						
	980			985		990
Phe Ile Glu Ala Leu Glu Thr Glu Leu Glu Lys Pro Phe Lys Thr Ala						
	995			1000		1005
Ile Gln Lys Ala Asn Ile Tyr Gly Phe Cys Phe Ala Phe Ala Gln						
	1010			1015		1020
Cys Ile Met Phe Ile Ala Asn Ser Ala Ser Tyr Arg Tyr Gly Gly						
	1025			1030		1035
Tyr Leu Ile Ser Asn Glu Gly Leu His Phe Ser Tyr Val Phe Arg						
	1040			1045		1050
Val Ile Ser Ala Val Val Leu Ser Ala Thr Ala Leu Gly Arg Ala						
	1055			1060		1065
Phe Ser Tyr Thr Pro Ser Tyr Ala Lys Ala Lys Ile Ser Ala Ala						
	1070			1075		1080
Arg Phe Phe Gln Leu Leu Asp Arg Gln Pro Pro Ile Ser Val Tyr						
	1085			1090		1095
Asn Thr Ala Gly Glu Lys Trp Asp Asn Phe Gln Gly Lys Ile Asp						
	1100			1105		1110

Phe Val	Asp Cys Lys Phe Thr	Tyr Pro Ser Arg Pro	Asp Ser Gln
1115	1120	1125	
Val Leu	Asn Gly Leu Ser Val	Ser Ile Ser Pro Gly	Gln Thr Leu
1130	1135	1140	
Ala Phe	Val Gly Ser Ser Gly	Cys Gly Lys Ser Thr	Ser Ile Gln
1145	1150	1155	
Leu Leu	Glu Arg Phe Tyr Asp	Pro Asp Gln Gly Lys	Val Met Ile
1160	1165	1170	
Asp Gly	His Asp Ser Lys Lys	Val Asn Val Gln Phe	Leu Arg Ser
1175	1180	1185	
Asn Ile	Gly Ile Val Ser Gln	Glu Pro Val Leu Phe	Ala Cys Ser
1190	1195	1200	
Ile Met	Asp Asn Ile Lys Tyr	Gly Asp Asn Thr Lys	Glu Ile Pro
1205	1210	1215	
Met Glu	Arg Val Ile Ala Ala	Ala Lys Gln Ala Gln	Leu His Asp
1220	1225	1230	
Phe Val	Met Ser Leu Pro Glu	Lys Tyr Glu Thr Asn	Val Gly Ser
1235	1240	1245	
Gln Gly	Ser Gln Leu Ser Arg	Gly Glu Lys Gln Arg	Ile Ala Ile
1250	1255	1260	
Ala Arg	Ala Ile Val Arg Asp	Pro Lys Ile Leu Leu	Leu Asp Glu
1265	1270	1275	
Ala Thr	Ser Ala Leu Asp Thr	Glu Ser Glu Lys Thr	Val Gln Val
1280	1285	1290	
Ala Leu	Asp Lys Ala Arg Glu	Gly Arg Thr Cys Ile	Val Ile Ala
1295	1300	1305	
His Arg	Leu Ser Thr Ile Gln	Asn Ala Asp Ile Ile	Ala Val Met
1310	1315	1320	
Ala Gln	Gly Val Val Ile Glu	Lys Gly Thr His Glu	Glu Leu Met
1325	1330	1335	

Ala Gln Lys Gly Ala Tyr Tyr Lys Leu Val Thr Thr Gly Ser Pro
 1340 1345 1350

Ile Ser
 1355

<210> 9
 <211> 102
 <212> DNA
 <213> Homo sapiens

<400> 9
 gtgagtggct tttttccctc actgcatctt gtacaaggag aggtgagaac aaaagtagga 60
 caagctgggtc aagtttcaag gagcagaaaa aaatcagcaa ca 102

<210> 10
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 10

Cys Glu Trp Leu Phe Ser Leu Thr Ala Ser Cys Thr Arg Arg Gly Glu
 1 5 10 15

Asn Lys Ser Arg Thr Ser Trp Ser Ser Phe Lys Glu Gln Lys Lys Ile
 20 25 30

Ser Asn Asn
 35

<210> 11
 <211> 13
 <212> PRT
 <213> Homo sapiens

<400> 11

Glu Trp Leu Phe Ser Leu Thr Ala Ser Cys Thr Arg Arg
 1 5 10

<210> 12
 <211> 42
 <212> DNA
 <213> Homo sapiens

<400> 12
 gtgagtggct tttttccctc actgcatctt gtacaaggag ag 42

<210> 13
 <211> 57

<212> DNA
<213> Artificial Sequence

<220>
<223> forward sense oligo hBSEP-1

<400> 13
aattacgcgt accaccatgt ctgactcagt aattcttcga agtataaaga aatttgg 57

<210> 14
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse oligo hBSEP-2

<400> 14
aattgcggcc gcgggtcaac tgatggggga tccagtgggtg actagtttg 49

INTERNATIONAL SEARCH REPORT

IB2004/002471

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C07K14/435 C07K14/47 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, Sequence Search, PAJ, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SAITO SUSUMU ET AL: "Three hundred twenty-six genetic variations in genes encoding nine members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the Japanese population" JOURNAL OF HUMAN GENETICS, XX, XX, vol. 47, no. 1, 2002, pages 38-50, XP002210489 figures 1A-I; table 1I ----- -/--	1,5

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

° Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

19 October 2004

Date of mailing of the international search report

05/11/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Renggli-Zulliger, N

INTERNATIONAL SEARCH REPORT

IB2004/002471

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! EMBL; 30 July 1999 (1999-07-30), "Homo sapiens BAC clone RP11-527A7 from 2 , complete sequence" XP002301387 retrieved from EBI accession no. EMBL:AC008177 Database accession no. AC008177 abstract	4
A	----- STRAUTNIEKS SANDRA S ET AL: "A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis" NATURE GENETICS, vol. 20, no. 3, November 1998 (1998-11), pages 233-238, XP002301385 ISSN: 1061-4036 cited in the application figure 2; table 1 -& DATABASE EMBL 'Online! EMBL; 16 November 1998 (1998-11-16), "Homo sapiens bile salt export pump (BSEP) mRNA, complete cds" XP002301401 retrieved from EMBL accession no. EMBL:AF091582 Database accession no. AF091582 abstract	1-15
A	----- BYRNE JANE A ET AL: "The human bile salt export pump: characterization of substrate specificity and identification of inhibitors." GASTROENTEROLOGY. NOV 2002, vol. 123, no. 5, November 2002 (2002-11), pages 1649-1658, XP009038352 ISSN: 0016-5085 the whole document	1-15
A	----- LAMBA JATINDER KAUR ET AL: "Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons." HUMAN MOLECULAR GENETICS. 15 JAN 2003, vol. 12, no. 2, 15 January 2003 (2003-01-15), pages 99-109, XP002301386 ISSN: 0964-6906 the whole document -----	1-15