



US 20030166225A1

(19) **United States**

(12) **Patent Application Publication**

Ryan

(10) **Pub. No.: US 2003/0166225 A1**

(43) **Pub. Date:** Sep. 4, 2003

(54) **ISOLATED GENOMIC POLYNUCLEOTIDE
FRAGMENTS THAT ENCODE HUMAN
LIPOPROTEIN-ASSOCIATED
PHOSPHOLIPASE A2**

(76) Inventor: **James W. Ryan**, Augusta, GA (US)

Publication Classification

(51) **Int. Cl.⁷** C07H 21/04; C12P 21/04;
C12N 9/20; C12P 21/02; C12N 5/06
(52) **U.S. Cl.** 435/198; 435/69.1; 435/320.1;
435/325; 435/70.21; 536/23.2

Correspondence Address:
CHERYL H AGRIS PHD
PO BOX 806
PELHAM, NY 10803 (US)

(57)

ABSTRACT

The invention is directed to isolated genomic polynucleotide fragments that encode human lipoprotein-associated phospholipase A2, vectors and hosts containing the fragment and fragments hybridizing to noncoding regions as well as antisense oligonucleotides to these fragments. The invention is further directed to methods of using these fragments to obtain human lipoprotein-associated phospholipase A2 and to diagnose, treat, prevent and/or ameliorate a pathological disorder.

(21) Appl. No.: **10/161,127**

(22) Filed: **May 30, 2002**

Related U.S. Application Data

(60) Provisional application No. 60/294,404, filed on May 30, 2001.

**ISOLATED GENOMIC POLYNUCLEOTIDE
FRAGMENTS THAT ENCODE HUMAN
LIPOPROTEIN-ASSOCIATED PHOSPHOLIPASE
A2**

PRIORITY CLAIM

[0001] This invention claims priority from application serial No. 60/294,404, filed May 30, 2001, the contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The invention is directed to isolated genomic polynucleotide fragments that encode human lipoprotein-associated phospholipase A2, vectors and hosts containing these fragments and fragments hybridizing to noncoding regions as well as antisense oligonucleotides to these fragments. The invention is further directed to methods of using these fragments to obtain human lipoprotein-associated phospholipase A2 and to diagnose, treat, prevent and/or ameliorate a pathological and/or medical disorders.

BACKGROUND OF THE INVENTION

[0003] Human lipoprotein-associated phospholipase A2, also known in the art as Platelet Activating Factor Acetyl Hydrolase (PAF acetyl hydrolase), is one of a family of enzymes that catalyze release of fatty acids from membrane phospholipids and can thereby initiate synthesis of proinflammatory mediators. During the conversion of LDL to its oxidised form, lipoprotein-associated phospholipase A2 is responsible for hydrolysing the sn-2 ester of oxidatively modified phosphatidylcholine to give lyso-phosphatidylcholine and an oxidatively modified fatty acid. Both of these products of human lipoprotein-associated phospholipase A2 action are potent chemoattractants for circulating monocytes. The enzyme appears to play a central role in the development of atherosclerosis and is regarded as an independent risk factor for coronary artery disease (Caslake et al., Atherosclerosis 150: 413-19, 2000). Specifically, this enzyme is thought to be responsible for the accumulation of cells loaded with cholesterol ester in the arteries, causing the characteristic 'fatty streak' associated with the early stages of atherosclerosis. Recently, medicinal chemists have begun to design and prepare lipoprotein-associated phospholipase A2 inhibitors for use in preventing or inhibiting progression of atherosclerotic diseases (See, for example, U.S. Pat. Nos. 5,981,252 and 5,968,818; Boyd et al., Bioorg. Med. Chem. 10: 2557-61, 2000).

[0004] The level of PAF acetylhydrolase has been found to be altered in several disease states. For example, acquired deficiency of PAF acetylhydrolase activity has been reported in patients with systemic lupus erythematosus, stroke and asthma, and increased levels of PAF have been reported in children with acute asthmatic attacks (see, for example, Hiramoto et al., Stroke 28: 2417-2420, 1997; Kruse et al., Am. J. Hum. Genet. 66: 1522-1530, 2000; Stafforini et al., J. Clin. Invest. 97: 2784-2791, 1996). Miwa et al. (1988, J. Clin. Invest. 82:1983-1991) have also described an autosomal recessive form of PAF acetylhydrolase deficiency which has been observed only in the Japanese population. PAF acetylhydrolase activity was absent in 4% of the Japanese population. This inherited deficiency is the result of a point mutation in exon 9 and completely abolishes enzymatic

activity. These patients suffer from severe asthma. Results from further studies indicated that the variant allele thr198 was found to be highly associated with total IgE concentrations in an atopic population and with 'asthma' in an asthmatic population (Kruse et al., 2000, Am. J. Hum. Genet. 66:1522-1530). The variant allele val379 was found to be highly associated with 'specific sensitization' in the atopic population and with 'asthma' in the asthmatic population.

[0005] The full length cDNA clone has been isolated (see U.S. Pat. Nos. 5,981,252 and 5,968,818) and the DNA sequence has been determined. The complete amino acid sequence has been deduced from the DNA sequence. A gene encoding the human lipoprotein-associated phospholipase A2 polypeptide is located at gene map locus 6p21.2-p12.

OBJECTS OF THE INVENTION

[0006] Although cDNA encoding the above-disclosed protein, lipoprotein associated phospholipase A2, has been isolated (e.g. see accession no. AX006795 and NM_005084), its exact chromosome location and exon/intron/regulatory organization have not been determined. Furthermore, genomic DNA encoding the polypeptide has not been isolated. Noncoding sequences play a significant role in regulating the expression of polypeptides as well as the processing of RNA encoding these polypeptides.

[0007] There is clearly a need for obtaining genomic polynucleotide sequences encoding the lipoprotein-associated phospholipase A2 polypeptide. Therefore, it is an object of the invention to isolate such genomic polynucleotide sequences.

[0008] There is also a need to develop means for identifying mutations, duplications, translocations, polysomies and mosaicism as may affect the lipoprotein associated phospholipase A2 gene.

SUMMARY OF THE INVENTION

[0009] The invention is directed to an isolated genomic polynucleotide, said polynucleotide obtainable from human genomic clone library RP11-446F17 having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

[0010] (a) a polynucleotide encoding human lipoprotein-associated phospholipase A2 depicted in SEQ ID NO:1.

[0011] (b) a polynucleotide consisting of SEQ ID NO:2, which encodes human lipoprotein-associated phospholipase A2 depicted in SEQ ID NO:1

[0012] (c) a polynucleotide which is a variant of SEQ ID NO:2;

[0013] (d) a polynucleotide which is an allelic variant of SEQ ID NO:2;

[0014] (e) a polynucleotide which encodes a variant of SEQ ID NO:1;

[0015] (f) a polynucleotide which hybridizes to any one of the polynucleotides specified in (a)-(e) and

[0016] (g) a polynucleotide that is a reverse complement to the polynucleotides specified in (a)-(f) and

[0017] (h) containing at least 10 transcription factor binding sites selected from the group consisting of AP1-C, AP1_Q4, AP4-Q5, DELTAEF1-01, GATA1_04, GATA1-06, GATA2-02, GATA3_02, GATA-C, LMO2COM-02, LYF1-01, MYOD_Q6, MZF_01, NFAT_Q6, NKX25-01, S8-01, SOX5-01, TATA-C, and TCF11-01

[0018] as well as nucleic acid constructs, expression vectors and host cells containing these polynucleotide sequences.

[0019] The polynucleotides of the present invention may be used to modulate human lipoprotein-associated phospholipase A2 levels in a subject in need thereof. Specifically, it may be used for the manufacture of a gene therapy for the prevention, treatment or amelioration of a medical condition such as asthma or systemic lupus erythematosus by adding an amount of a composition comprising said polynucleotide effective to prevent, treat or ameliorate said medical condition.

[0020] The invention is further directed to obtaining these polypeptides by

[0021] (a) culturing host cells comprising these sequences under conditions that provide for the expression of said polypeptide and

[0022] (b) recovering said expressed polypeptide.

[0023] The polypeptides obtained may be used to produce antibodies by

[0024] (a) optionally conjugating said polypeptide to a carrier protein;

[0025] (b) immunizing a host animal with said polypeptide or peptide-carrier protein conjugate of step (a) with an adjuvant and

[0026] (c) obtaining antibody from said immunized host animal.

[0027] The invention is further directed to polynucleotides that hybridize to noncoding regions of said polynucleotide sequences as well as antisense oligonucleotides to these polynucleotides as well as antisense mimetics. The polynucleotides, antisense oligonucleotides or mimetics may be used to modulate human lipoprotein-associated phospholipase A2 levels in a subject in need thereof and for the manufacture of a medicament for prevention, treatment or amelioration of a medical condition, such as atherosclerosis. The invention is further directed to kits comprising these polynucleotides and kits comprising these antisense oligonucleotides or mimetics.

[0028] In a specific embodiment, the noncoding regions are transcription regulatory regions. The transcription regulatory regions may be used to produce a heterologous peptide by expressing in a host cell, said transcription regulatory region operably linked to a polynucleotide encoding the heterologous polypeptide and recovering the expressed heterologous polypeptide.

[0029] The polynucleotides and antisense oligonucleotides of the present invention may be used to diagnose a pathological condition such as asthma in a subject comprising

[0030] (a) determining the presence or absence of a mutation in the polynucleotides of the present invention or fragments thereof and

[0031] (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

DETAILED DESCRIPTION OF THE INVENTION

[0032] The invention is directed to isolated genomic polynucleotide fragments that encode human lipoprotein-associated phospholipase A2, which in a specific embodiment is the human lipoprotein-associated phospholipase A2 gene, as well as vectors and hosts containing these fragments and polynucleotide fragments hybridizing to noncoding regions, as well as antisense oligonucleotides to these fragments.

[0033] As defined herein, a "gene" is the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region, as well as intervening sequences (introns) between individual coding segments (exons).

[0034] As defined herein "isolated" refers to material removed from its original environment and is thus altered "by the hand of man" from its natural state. An isolated polynucleotide can be part of a vector, a composition of matter or can be contained within a cell as long as the cell is not the original environment of the polynucleotide.

[0035] The polynucleotides of the present invention may be in the form of RNA or in the form of DNA, which DNA includes genomic DNA and synthetic DNA. The DNA may be double-stranded or single-stranded and if single stranded may be the coding strand or non-coding strand.

[0036] The polynucleotides of the invention have at least a 95% identity and may have a 96%, 97%, 98% or 99% identity to the polynucleotides depicted in SEQ ID NO:2 as well as the polynucleotides in reverse sense orientation, or the polynucleotide sequences encoding the human lipoprotein-associated phospholipase A2 polypeptide depicted in SEQ ID NO:1.

[0037] A polynucleotide having 95% "identity" to a reference nucleotide sequence of the present invention, is identical to the reference sequence except that the polynucleotide sequence may include, on average, up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence, the ORF (open reading frame), or any fragment specified as described herein.

[0038] As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query

sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are:

[0039] Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cut-off Score=1, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

[0040] If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identify, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence are calculated for the purposes of manually adjusting the percent identity score.

[0041] For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total numbers of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for purposes of the present invention.

[0042] A polypeptide that has an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence is identical to the query sequence except that the subject polypeptide sequence may include on average, up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a

polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the referenced sequence or in one or more contiguous groups within the reference sequence.

[0043] A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Com. App. Biosci. (1990) 6:237-245). In a sequence alignment, the query and subject sequence are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are:

[0044] Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cut-off Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

[0045] If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

[0046] The invention also encompasses polynucleotides that hybridize to the polynucleotides depicted in SEQ ID NO: 2. A polynucleotide "hybridizes" to another polynucleotide, when a single-stranded form of the polynucleotide can anneal to the other polynucleotide under the appropriate conditions of temperature and solution ionic strength (see Sambrook et al., *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a temperature of 42° C., can be used, e.g., 5× SSC, 0.1% SDS, 0.25% milk, and no formamide; or 40% forma-

mide, 5× SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher temperature of 55° C., e.g., 40% formamide, with 5× or 6× SCC. High stringency hybridization conditions correspond to the highest temperature of 65° C., e.g., 50% formamide, 5× or 6× SSC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA.

[0047] Polynucleotide and Polypeptide Variants

[0048] The invention is directed to both polynucleotide and polypeptide variants. A “variant” refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar and in many regions, identical to the polynucleotide or polypeptide of the present invention.

[0049] The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred.

[0050] The invention also encompasses allelic variants of said polynucleotides. An allelic variant denotes any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

[0051] The amino acid sequences of the variant polypeptides may differ from the amino acid sequences depicted in SEQ ID NO:1 by an insertion or deletion of one or more amino acid residues and/or the substitution of one or more amino acid residues by different amino acid residues. Preferably, amino acid changes are of a minor nature, that is conservative amino acid substitutions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

[0052] Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histi-

dine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions which do not generally alter the specific activity are known in the art and are described, for example, by H. Neurath and R. L. Hill, 1979, In, *The Proteins*, Academic Press, New York. The most commonly occurring exchanges are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, as well as these in reverse.

[0053] Noncoding Regions

[0054] The invention is further directed to polynucleotide fragments containing or hybridizing to noncoding regions of the human lipoprotein-associated phospholipase A2 gene. These include but are not limited to an intron, a 5'- non-coding region, a 3'- non-coding region and splice junctions (see Table 1), as well as transcription factor binding sites (see Table 2). The polynucleotide fragments may be a short polynucleotide fragment which is between about 8 nucleotides to about 40 nucleotides in length. Such shorter fragments may be useful for diagnostic purposes. Such short polynucleotide fragments are also preferred with respect to polynucleotides containing or hybridizing to polynucleotides containing splice junctions. Alternatively larger fragments, e.g., of about 50, 150, 500, 600 or about 2000 nucleotides in length may be used.

TABLE 1

Exon/Intron Regions of the Lipoprotein-Associated Phospholipase A2 Gene in SEQ ID NO: 2 (42,611 base pairs, forward strand coding).

Exon	Nucleotide no.	Peptide Amino Acid no.
1	11967-12071	1-35
2	17761-17886	36-77
3	18333-18476	78-125
4	20307-20399	126-156
5	22522-22590	157-179
6	23240-23365	180-221
7	24203-24316	222-259
8	25443-25532	260-289
9	26698-26868	290-346
10	29558-29707	347-396
11	30164-30298	397-441
Stop codon	30299-30301	

[0055]

TABLE 2

Transcription Factor Binding Sites on the Lipoprotein-Associated Phospholipase A2 Gene.

Transcription Factor	No. of Binding Sites
AP1_C	14
AP1_Q4	4
AP4_Q5	4
DELTAEF1_01	5
GATA1_04	7
GATA1_06	8
GATA2_02	5
GATA3_02	4
GATA_C	7

TABLE 2-continued

Transcription Factor Binding Sites on the Lipoprotein-Associated Phospholipase A2 Gene.	
Transcription Factor	No. of Binding Sites
LMO2COM_02	4
LYF1_01	9
MYOD_Q6	9
MZF1_01	14
NFAT_Q6	10
NKX25_01	16
S8_01	7
SOX5_01	2
TATA_C	7
TCF11_01	46

[0056] In a specific embodiment, such noncoding sequences are expression control sequences. These include but are not limited to DNA regulatory sequences, such as promoters, enhancers, repressors, terminators, and the like, that provide for the regulation of expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are also control sequences.

[0057] In a more specific embodiment of the invention, the expression control sequences may be operatively linked to a polynucleotide encoding a heterologous polypeptide. Such expression control sequences may be about 50-200 nucleotides in length and specifically about 50, 100, 200, 500, 600, 1000 or 2000 nucleotides in length. A transcriptional control sequence is "operatively linked" to a polynucleotide encoding a heterologous polypeptide sequence when the expression control sequence controls and regulates the transcription and translation of that polynucleotide sequence. The term "operatively linked" includes having an appropriate start signal (e.g., ATG) in front of the polynucleotide sequence to be expressed and maintaining the correct reading frame to permit expression of the DNA sequence under the control of the expression control sequence and production of the desired product encoded by the polynucleotide sequence. If a gene that one desires to insert into a recombinant DNA molecule does not contain an appropriate start signal, such a start signal can be inserted upstream (5') of and in reading frame with the gene.

[0058] The invention is further directed to antisense oligonucleotides and mimetics to these polynucleotide sequences. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription or RNA processing (triple helix (see Lee et al., *Nucl. Acids Res.*, 6:3073 (1979); Cooney et al, *Science*, 241:456 (1988); and Dervan et al., *Science*, 251: 1360 (1991)), thereby preventing transcription and the production of said polypeptides.

[0059] Expression of Polypeptides

[0060] Isolated Polynucleotide Sequences

[0061] The human lipoprotein related phospholipase A2 gene is situated in genomic clone AC008104; contigs 21 and 20 of gi 8072415, and these two contigs are joined by clone AL591242 contig 9 of gi 14018186. Genomic clone AC008104 is reported by GENBANK to be derived from chromosome 18, whereas genomic clone AL591242 is reported by Genbank to be derived from chromosome 6. However, both Genbank reports indicate that the genomic clones were prepared from clone library RP11-446F17 (Osoegawa et al., 2001, *Genome Res.* 11:483-96).

[0062] The human genomic clone RP11-446F17 sequences of the gi versions noted above (gi 8072415 and gi 14018186) have been discovered to contain the human lipoprotein related phospholipase A2 gene by Genscan analysis (Burge et al., 1997, *J. Mol. Biol.* 268:78-94), BLAST2 and TBLASTN analysis (Altschul et al., 1997, *Nucl. Acids Res.* 25:3389-3402), in which the clone sequences were compared to the human lipoprotein related phospholipase A2 cDNA sequences, accession numbers AX006795 and NM_005084.

[0063] The cloning of the nucleic acid sequences of the present invention from such genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, e.g., Innis et al., 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleic acid sequence-based amplification (NASBA) or long range PCR may be used. In a specific embodiment, 5'- or 3'-non-coding portions of the gene may be identified by methods including but are not limited to, filter probing, clone enrichment using specific probes and protocols similar or identical to 5'- and 3'- "RACE" protocols which are well known in the art. For instance, a method similar to 5'-RACE is available for generating the missing 5'- end of a desired full-length transcript. (Fromont-Racine et al., 1993, *Nucl. Acids Res.* 21:1683-1684).

[0064] Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired human lipoprotein-associated phospholipase A2 gene may be accomplished in a number of ways. For example, if an amount of a portion of a human lipoprotein-associated phospholipase A2 gene or its specific RNA, or a fragment thereof, is available and can be purified and labeled, the generated DNA fragments may be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, 1977, *Science* 196:180; Grunstein and Hogness, 1975, *Proc. Natl. Acad. Sci. U.S.A.* 72:3961). The present invention provides such nucleic acid probes, which can be conveniently prepared from the specific sequences disclosed herein, e.g., a hybridizable probe having a nucleotide sequence corresponding to at least a 10, and preferably a 15, nucleotide fragment of the sequences depicted in SEQ ID NO:2 Preferably, a fragment is selected that is highly unique to the polynucleotides of the invention. Those DNA fragments with substantial homology to the probe will hybridize. As noted above, the greater the degree of homology, the more stringent hybridization conditions can be used. In one

embodiment, low stringency hybridization conditions are used to identify a homologous human lipoprotein-associated phospholipase A2 polynucleotide. However, in a preferred aspect, a nucleic acid encoding a polypeptide of the invention will hybridize to a nucleic acid derived from the polynucleotide sequence depicted in SEQ ID NO:2 or a hybridizable fragment thereof, under moderately stringent conditions; more preferably, it will hybridize under high stringency conditions.

[0065] Alternatively, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNAs, can be selected which produce a protein that, e.g., has similar or identical electrophoretic migration, isoelectric focusing behavior, proteolytic digestion maps, or antigenic properties as known for the human lipoprotein-associated phospholipase A2 polypeptide.

[0066] A gene encoding human lipoprotein-associated phospholipase A2 polypeptide can also be identified by mRNA selection, i.e., by nucleic acid hybridization followed by in vitro translation. In this procedure, fragments are used to isolate complementary mRNAs by hybridization. Immunoprecipitation analysis or functional assays of the in vitro translation products of the products of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments, that contain the desired sequences.

[0067] Nucleic Acid Constructs

[0068] The present invention also relates to nucleic acid constructs comprising a polynucleotide sequence containing the exon/intron segments of the human lipoprotein-associated phospholipase A2 gene operably linked to one or more control sequences which direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences. Expression will be understood to include any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0069] The invention is further directed to a nucleic acid construct comprising expression control sequences derived from SEQ ID NO: 2 and a heterologous polynucleotide sequence.

[0070] "Nucleic acid construct" is defined herein as a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acid which are combined and juxtaposed in a manner which would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term expression cassette when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention. The term "coding sequence" is defined herein as a portion of a nucleic acid sequence which directly specifies the amino acid sequence of its protein product. The boundaries of the coding sequence are generally determined by a ribosome binding site (prokaryotes) or by the ATG start codon (eukaryotes) located just upstream of the open reading frame at the 5'- end of the mRNA and a transcription terminator sequence located just downstream of the open reading frame at the 3'- end of the mRNA. A coding

sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleic acid sequences.

[0071] The isolated polynucleotide of the present invention may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the nucleic acid sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleic acid sequences utilizing recombinant DNA methods are well known in the art.

[0072] The control sequence may be an appropriate promoter sequence, a nucleic acid sequence which is recognized by a host cell for expression of the nucleic acid sequence. The promoter sequence contains transcriptional control sequences which regulate the expression of the polynucleotide. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

[0073] Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli* lac operon, the *Streptomyces coelicolor* agarase gene (dagA), the *Bacillus subtilis* levansucrase gene (sacB), the *Bacillus licheniformis* alpha-amylase gene (amyL), the *Bacillus stearothermophilus* malto-genic amylase gene (amyM), the *Bacillus amyloliquefaciens* alpha-amylase gene (amyQ), the *Bacillus licheniformis* penicillinase gene (penP), the *Bacillus subtilis* xylA and xylB genes, and the prokaryotic beta-lactamase gene (Villa-Komaroff et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), as well as the tac promoter (DeBoer et al., 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74-94; and in Sambrook et al., 1989, supra.

[0074] Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), NA2-tpi (a hybrid of the promoters from the genes encoding *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

[0075] In a yeast host, useful promoters are obtained from the *Saccharomyces cerevisiae* enolase (ENO-1) gene, the *Saccharomyces cerevisiae* galactokinase gene (GAL1), the *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase genes (ADH2/GAP), and the *Saccharomyces cerevisiae* 3-phosphoglycerate kinase gene. Other useful promoters for yeast host cells are described by Romanos et al., 1992, *Yeast* 8: 423-488.

[0076] The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a

host cell to terminate transcription. The terminator sequence is operably linked to the 3'- terminus of the nucleic acid sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used in the present invention. Preferred terminators for filamentous fungal host cells are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase, and *Fusarium oxysporum* trypsin-like protease.

[0077] Preferred terminators for yeast host cells are obtained from the genes encoding *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), or *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos et al., 1992, supra.

[0078] The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5'- terminus of the nucleic acid sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

[0079] Preferred leaders for filamentous fungal host cells are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

[0080] Suitable leaders for yeast host cells are obtained from the *Saccharomyces cerevisiae* enolase (ENO-1) gene, the *Saccharomyces cerevisiae* 3-phosphoglycerate kinase gene, the *Saccharomyces cerevisiae* alpha-factor, and the *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase genes (ADH2/GAP).

[0081] The control sequence may also be a polyadenylation sequence, a sequence which is operably linked to the 3'-terminus of the nucleic acid sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

[0082] Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, and *Aspergillus niger* alpha-glucosidase.

[0083] Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

[0084] The control sequence may also be a signal peptide coding region, which codes for an amino acid sequence linked to the amino terminus of the polypeptide which can direct the encoded polypeptide into the cell's secretory pathway. The 5'- end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5'- end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence

does not normally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to obtain enhanced secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

[0085] An effective signal peptide coding region for bacterial host cells is the signal peptide coding region obtained from the maltogenic amylase gene from *Bacillus* NCIB 11837, the *Bacillus stearothermophilus* alpha-amylase gene, the *Bacillus licheniformis* subtilisin gene, the *Bacillus licheniformis* beta-lactamase gene, the *Bacillus stearothermophilus* neutral proteases genes (nprT, nprS, nprM), or the *Bacillus subtilis* prsA gene. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

[0086] An effective signal peptide coding region for filamentous fungal host cells is the signal peptide coding region obtained from the *Aspergillus oryzae* TAKA amylase gene, *Aspergillus niger* neutral amylase gene, *Aspergillus niger* glucoamylase gene, *Rhizomucor miehei* aspartic proteinase gene, *Humicola lanuginosa* cellulase gene, or *Humicola lanuginosa* lipase gene.

[0087] Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding regions are described by Romanos et al., 1992, supra.

[0088] The control sequence may also be a propeptide coding region, which codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the *Bacillus subtilis* alkaline protease gene (aprE), the *Bacillus subtilis* neutral protease gene (nprT), the *Saccharomyces cerevisiae* alpha-factor gene, the *Rhizomucor miehei* aspartic proteinase gene, or the *Myceliophthora thermophila* laccase gene (WO 95/33836).

[0089] Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

[0090] It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems would include the lac, tac, and trp operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and the *Aspergillus oryzae* glucoamylase promoter may be used as regulatory sequences. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic

systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the polypeptide would be operably linked with the regulatory sequence.

[0091] Expression Vectors

[0092] The present invention also relates to recombinant expression vectors comprising a nucleic acid sequence of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the polynucleotide of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0093] The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

[0094] The vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

[0095] The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Examples of bacterial selectable markers are the dal genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. A selectable marker for use in a filamentous fungal host cell may be selected from the group including, but not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hygB (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5 \bar{O} -phosphate decarboxylase), sC (sulfate adenyltransferase), trpC (anthranilate synthase), as well as

equivalents from other species. Preferred for use in an *Aspergillus* cell are the amdS and pyrG genes of *Aspergillus nidulans* or *Aspergillus oryzae* and the bar gene of *Streptomyces hygroscopicus*.

[0096] The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into the host cell genome or autonomous replication of the vector in the cell independent of the genome of the cell.

[0097] For integration into the host cell genome, the vector may rely on the polynucleotide sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional polynucleotide sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0098] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM \pm 1 permitting replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1433).

[0099] More than one copy of a polynucleotide sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the polynucleotide sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0100] The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, *supra*).

[0101] Host Cells

[0102] The present invention also relates to recombinant host cells, comprising a nucleic acid sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. A vector comprising a nucleic acid sequence of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

[0103] The host cell may be a unicellular microorganism, e.g., a prokaryote, or a non-unicellular microorganism, e.g., a eukaryote. Useful unicellular cells are bacterial cells such as gram positive bacteria including, but not limited to, a *Bacillus* cell, e.g., *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus laetus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis*; or a Streptomyces cell, e.g., *Streptomyces lividans* or *Streptomyces murinus*, or gram negative bacteria such as *E. coli* and *Pseudomonas* sp. In a preferred embodiment, the bacterial host cell is a *Bacillus licheniformis*, *Bacillus stearothermophilus* or *Bacillus subtilis* cell. In another preferred embodiment, the *Bacillus* cell is an alkaliophilic *Bacillus*.

[0104] The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), using competent cells (see, e.g., Young and Spizizin, 1961, *Journal of Bacteriology* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, e.g., Kochler and Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

[0105] The host cell may be a eukaryote, such as a mammalian cell (e.g., human cell), an insect cell, a plant cell or a fungal cell. Mammalian host cells that could be used include but are not limited to human HeLa, 293, H9 and Jurkat cells, mouse NIH 3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese Hamster ovary (CHO) cells. These cells may be transfected with a vector containing a transcriptional regulatory sequence, a protein coding sequence and transcriptional termination sequences. Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

[0106] The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth et al., In, *Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth et al., 1995, supra, page 171) and all mitosporic

fungi (Hawksworth et al., 1995, supra). The fungal host cell may also be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, F. A., Passmore, S. M., and Davenport, R. R., eds, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980). The fungal host cell may also be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth et al., 1995, supra). The filamentous fungi are characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

[0107] Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known per se. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton et al., 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Suitable methods for transforming *Fusarium* species are described by Malardier et al., 1989, *Gene* 78: 147-156 and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J. N. and Simon, M. I., editors, *Guide to Yeast Genetics and Molecular Biology*, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, *Journal of Bacteriology* 153: 163; and Hinnen et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

[0108] Methods of Production

[0109] The present invention also relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a host cell under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

[0110] In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

[0111] The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These

detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. In a specific embodiment, an enzyme assay may be used to determine the activity of the polypeptide. For example, human lipoprotein-associated phospholipase A2 can be assayed by its ability to release fatty acids from phospholipids. Caslake et al (Atherosclerosis 150: 413-19, 2000) have described a specific immunoassay for the enzyme.

[0112] The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

[0113] The polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., *Protein Purification*, J. -C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

[0114] Antibodies

[0115] According to the invention, the human lipoprotein-associated phospholipase A2 polypeptide produced according to the method of the present invention may be used as an immunogen to generate any of these antibodies. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library.

[0116] Various procedures known in the art may be used for the production of antibodies. For the production of antibody, various host animals can be immunized by injection with the polypeptide thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the polypeptide or fragment thereof can optionally be conjugated to an immunogenic carrier, e.g., bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

[0117] For preparation of monoclonal antibodies directed toward the human lipoprotein-associated phospholipase A2 polypeptide, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be

produced in germ-free animals utilizing recent technology (PCT/US90/02545). According to the invention, human antibodies may be used and can be obtained by using human hybridomas (Cote et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030) or by transforming human B cells with EBV virus in vitro (Cole et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, pp. 77-96). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, J. Bacteriol. 159:870; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule specific for the human lipoprotein-associated phospholipase A2 polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention.

[0118] According to the invention, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778) can be adapted to produce polypeptide-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse et al., 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for the human lipoprotein-associated phospholipase A2 polypeptide.

[0119] Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

[0120] In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g., radioimmunoassay, ELISA (enzyme-linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of a particular polypeptide, one may assay generated hybridomas for a product which binds to particular polypeptide fragment containing such epitope. For selection of an antibody specific to a particular polypeptide from a particular species of animal, one can select on the basis of positive binding with the polypeptide expressed by or isolated from cells of that species of animal.

[0121] Immortal, antibody-producing cell lines can also be created by techniques other than fusion, such as direct

transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al., "Hybridoma Techniques" (1980); Hammerling et al., "Monoclonal Antibodies And T-cell Hybridomas" (1981); Kennett et al., "Monoclonal Antibodies" (1980); see also U.S. Pat. Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,451,570; 4,466,917; 4,472,500; 4,491,632; 4,493,890.

[0122] **Uses of Polynucleotides**

[0123] **Diagnostics**

[0124] Polynucleotides containing noncoding regions of SEQ ID NO:2 may be used as probes for detecting mutations from samples from a patient. Genomic DNA may be isolated from the patient. A mutation(s) may be detected by Southern blot analysis, specifically by hybridizing restriction digested genomic DNA to various probes and subjecting to agarose electrophoresis. Alternatively, these polynucleotides may be used as PCR primers and be used to amplify the genomic DNA isolated from the patients. Additionally, primers may be obtained by routine or long range PCR that yield products containing contiguous intron/exon sequence and products containing more than one exon with intervening intron. The sequence of the amplified genomic DNA from the patient may be determined using methods known in the art. Such probes may be between 10-100 nucleotides in length and may preferably be between 20-50 nucleotides in length.

[0125] Thus the invention is thus directed to kits comprising these polynucleotide probes. In a specific embodiment, these probes are labeled with a detectable substance.

[0126] **Antisense Oligonucleotides and Mimetics**

[0127] The antisense oligonucleotides or mimetics of the present invention may be used to decrease or modulate levels of a polypeptide. For example, human lipoprotein-associated phospholipase A2 has been found to be associated with atherosclerosis and diabetes. Therefore, the human lipoprotein-associated phospholipase A2 antisense oligonucleotides of the present invention could be used to inhibit progression of atherosclerosis, including coronary artery disease. The antisense oligonucleotides may also be used for diagnostic purposes as well, as probes for detecting mutations from samples from a patient as described above.

[0128] The antisense oligonucleotides of the present invention may be formulated into pharmaceutical compositions. These compositions may be administered in a number of ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration may be topical (including ophthalmic and to mucous membranes including vaginal and rectal delivery), pulmonary, e.g., by inhalation or insufflation of powders or aerosols, including by nebulizer; intratracheal, intranasal, epidermal and transdermal), oral or parenteral. Parenteral administration includes intravenous, intraarterial, subcutaneous, intraperitoneal or intramuscular injection or infusion; or intracranial, e.g., intrathecal or intraventricular, administration.

[0129] Pharmaceutical compositions and formulations for topical administration may include transdermal patches, ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.

[0130] Compositions and formulations for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, capsules, sachets or tablets. Thickeners, flavoring agents, diluents, emulsifiers, dispersing aids or binders may be desirable.

[0131] Compositions and formulations for parenteral, intrathecal or intraventricular administration may include sterile aqueous solutions which may also contain buffers, diluents and other suitable additives such as, but not limited to, penetration enhancers, carrier compounds and other pharmaceutically acceptable carriers or excipients.

[0132] Pharmaceutical compositions of the present invention include, but are not limited to, solutions, emulsions, and liposome-containing formulations. These compositions may be generated from a variety of components that include, but are not limited to, preformed liquids, self-emulsifying solids and self-emulsifying semisolids.

[0133] The pharmaceutical formulations of the present invention, which may conveniently be presented in unit dosage form, may be prepared according to conventional techniques well known in the pharmaceutical industry. Such techniques include the step of bringing into association the active ingredients with the pharmaceutical carrier(s) or excipient(s). In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredients with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

[0134] The compositions of the present invention may be formulated into any of many possible dosage forms such as, but not limited to, tablets, capsules, liquid syrups, soft gels, suppositories, and enemas. The compositions of the present invention may also be formulated as suspensions in aqueous, non-aqueous or mixed media. Aqueous suspensions may further contain substances which increase the viscosity of the suspension including, for example, sodium carboxymethylcellulose, sorbitol and/or dextran. The suspension may also contain stabilizers.

[0135] In one embodiment of the present invention, the pharmaceutical compositions may be formulated and used as foams. Pharmaceutical foams include formulations such as, but not limited to, emulsions, microemulsions, creams, jellies and liposomes. While basically similar in nature these formulations vary in the components and the consistency of the final product. The preparation of such compositions and formulations is generally known to those skilled in the pharmaceutical and formulation arts and may be applied to the formulation of the compositions of the present invention.

[0136] The formulation of therapeutic compositions and their subsequent administration is believed to be within the skill of those in the art. Dosing is dependent on severity and responsiveness of the disease state to be treated, with the course of treatment lasting from several days to several months, or until a cure is effected or a diminution of the disease state is achieved. Optimal dosing schedules can be calculated from measurements of drug accumulation in the body of the patient. Persons of ordinary skill can easily determine optimum dosages, dosing methodologies and repetition rates. Optimum dosages may vary depending on the relative potency of individual oligonucleotides, and can generally be estimated based on EC₅₀ as found to be effective in *in vitro* and *in vivo* animal models.

[0137] In general, dosage is from 0.01 ug to 10 g per kg of body weight, and may be given once or more daily, weekly, monthly or yearly, or even once every 2 to 20 years. Persons of ordinary skill in the art can easily estimate repetition rates for dosing based on measured residence times and concentrations of the drug in bodily fluids or tissues. Following successful treatment, it may be desirable to have the patient undergo maintenance therapy to prevent the recurrence of the disease state, wherein the oligonucleotide is administered in maintenance doses, ranging from 0.01 ug to 10 g per kg of body weight, once or more daily, to once every 20 years.

[0138] Gene Therapy

[0139] Human lipoprotein-associated phospholipase A2 deficiency states have been described (Yamada et al., Metabolism 47: 177-81, 1998). Therefore, the human lipoprotein-associated phospholipase A2 gene may be used to via gene therapy to correct any such deficiency state or disorders associated with such deficiency states (e.g., asthma).

[0140] As described herein, the polynucleotide of the present invention may be introduced into a patient's cells for therapeutic uses. As will be discussed in further detail below, cells can be transfected using any appropriate means, including viral vectors, as shown by the example, chemical transfectants, or physico-mechanical methods such as electroporation and direct diffusion of DNA. See, for example, Wolff, Jon A, et al., "Direct gene transfer into mouse muscle in vivo," *Science*, 247, 1465-1468, 1990; and Wolff, Jon A, "Human dystrophin expression in mdx mice after intramuscular injection of DNA constructs," *Nature*, 352, 815-818, 1991. As used herein, vectors are agents that transport the gene into the cell without degradation and include a promoter yielding expression of the gene in the cells into which it is delivered. As will be discussed in further detail below, promoters can be general promoters, yielding expression in a variety of mammalian cells, or cell specific, or even nuclear versus cytoplasmic specific. These are known to those skilled in the art and can be constructed using standard molecular biology protocols. Vectors have been divided into two classes:

[0141] a) Biological agents derived from viral, bacterial or other sources.

[0142] b) Chemical physical methods that increase the potential for gene uptake, directly introduce the gene into the nucleus or target the gene to a cell receptor.

[0143] Biological Vectors

[0144] Viral vectors have higher transaction (ability to introduce genes) abilities than do most chemical or physical methods to introduce genes into cells. Vectors that may be used in the present invention include viruses, such as adenoviruses, adeno associated virus (AAV), vaccinia, herpesviruses, baculoviruses and retroviruses, bacteriophages, cosmids, plasmids, fungal vectors and other recombination vehicles typically used in the art which have been described for expression in a variety of eukaryotic and prokaryotic hosts, and may be used for gene therapy as well as for simple protein expression. Polynucleotides are inserted into vector genomes using methods well known in the art.

[0145] Retroviral vectors are the vectors most commonly used in clinical trials, since they carry a larger genetic payload than other viral vectors. However, they are not useful in non-proliferating cells. Adenovirus vectors are

relatively stable and easy to work with, have high titers, and can be delivered in aerosol formulation. Pox viral vectors are large and have several sites for inserting genes, they are thermostable and can be stored at room temperature.

[0146] Examples of promoters are SP6, T4, T7, SV40 early promoter, cytomegalovirus (CMV) promoter, mouse mammary tumor virus (MMTV) steroid-inducible promoter, Moloney murine leukemia virus (MMLV) promoter, phosphoglycerate kinase (PGK) promoter, and the like. Alternatively, the promoter may be an endogenous adenovirus promoter, for example the E1 a promoter or the Ad2 major late promoter (MLP). Similarly, those of ordinary skill in the art can construct adenoviral vectors utilizing endogenous or heterologous poly A addition signals. Plasmids are not integrated into the genome and the vast majority of them are present only from a few weeks to several months, so they are typically very safe. However, they have lower expression levels than retroviruses and since cells have the ability to identify and eventually shut down foreign gene expression, the continuous release of DNA from the polymer to the target cells substantially increases the duration of functional expression while maintaining the benefit of the safety associated with non-viral transfections.

[0147] Chemical/Physical Vectors

[0148] Other methods to directly introduce genes into cells or exploit receptors on the surface of cells include the use of liposomes and lipids, ligands for specific cell surface receptors, cell receptors, and calcium phosphate and other chemical mediators, microinjections directly to single cells, electroporation and homologous recombination. Liposomes are commercially available from Gibco BRL, for example, as LIPOFECTIN® and LIPOFECTACE®, which are formed of cationic lipids such as N-[1-(2,3 dioleyloxy)-propyl]-n,n,n-trimethylammonium chloride (DOTMA) and dimethyl dioctadecylammonium bromide (DDAB). Numerous methods are also published for making liposomes, known to those skilled in the art.

[0149] For example, Nucleic acid-Lipid Complexes—Lipid carriers can be associated with naked nucleic acids (e.g., plasmid DNA) to facilitate passage through cellular membranes. Cationic, anionic, or neutral lipids can be used for this purpose. However, cationic lipids are preferred because they have been shown to associate better with DNA which, generally, has a negative charge. Cationic lipids have also been shown to mediate intracellular delivery of plasmid DNA (Felgner and Ringold, *Nature* 337:387 (1989)). Intravenous injection of cationic lipid-plasmid complexes into mice has been shown to result in expression of the DNA in lung (Brigham et al., *Am. J. Med. Sci.* 298:278 (1989)). See also, Osaka et al., *J. Pharm. Sci.* 85(6):612-618 (1996); San et al., *Human Gene Therapy* 4:781-788 (1993); Senior et al., *Biochimica et Biophysica Acta* 1070:173-179 (1991); Kabanov and Kabanov, *Bioconjugate Chem.* 6:7-20 (1995); Remy et al., *Bioconjugate Chem.* 5:647-654 (1994); Behr, J -P., *Bioconjugate Chem.* 5:382-389 (1994); Behr et al., *Proc. Natl. Acad. Sci., USA* 86:6982-6986 (1989); and Wyman et al., *Biochem.* 36:3008-3017 (1997).

[0150] Cationic lipids are known to those of ordinary skill in the art. Representative cationic lipids include those disclosed, for example, in U.S. Pat. No. 5,283,185; and e.g., U.S. Pat. No. 5,767,099. In a preferred embodiment, the cationic lipid is N⁴-spermine cholesteryl carbamate (GL-67)

disclosed in U.S. Pat. No. 5,767,099. Additional preferred lipids include N4-spermidine cholestryl carbamate (GL-53) and 1-(N4-spermidine)-2,3-dilaurylglycerol carbamate (GL-89).

[0151] The vectors of the invention may be targeted to specific cells by linking a targeting molecule to the vector. A targeting molecule is any agent that is specific for a cell or tissue type of interest, including for example, a ligand, antibody, sugar, receptor, or other binding molecule.

[0152] Invention vectors may be delivered to the target cells in a suitable composition, either alone, or complexed, as provided above, comprising the vector and a suitably acceptable carrier. The vector may be delivered to target cells by methods known in the art, for example, intravenous, intramuscular, intranasal, subcutaneous, intubation, lavage, and the like. The vectors may be delivered via in vivo or ex vivo applications. In vivo applications involve the direct administration of an adenoviral vector of the invention formulated into a composition to the cells of an individual. Ex vivo applications involve the transfer of the adenoviral vector directly to harvested autologous cells which are maintained in vitro, followed by readministration of the transduced cells to a recipient.

[0153] In a specific embodiment, the vector is transfected into antigen-presenting cells. Suitable sources of antigen-presenting cells (APCs) include, but are not limited to, whole cells such as dendritic cells or macrophages; purified MHC class I molecule complexed to beta2-microglobulin and foster antigen-presenting cells. In a specific embodiment, the vectors of the present invention may be introduced into T cells or B cells using methods known in the art (see, for example, Tsokos and Nepom, 2000, J. Clin. Invest. 106:181-183).

[0154] The invention described and claimed herein is not to be limited in scope by the specific embodiments herein disclosed, since these embodiments are intended as illustrations of several aspects of the invention. Any equivalent embodiments are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

[0155] Various references are cited herein, the disclosure of which are incorporated by reference in their entireties.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 2

<210> SEQ ID NO 1
<211> LENGTH: 441
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

```

Met Val Pro Pro Lys Leu His Val Leu Phe Cys Leu Cys Gly Cys Leu
1           5          10          15

Ala Val Val Tyr Pro Phe Asp Trp Gln Tyr Ile Asn Pro Val Ala His
20          25          30

Met Lys Ser Ser Ala Trp Val Asn Lys Ile Gln Val Leu Met Ala Ala
35          40          45

Ala Ser Phe Gly Gln Thr Lys Ile Pro Arg Gly Asn Gly Pro Tyr Ser
50          55          60

Val Gly Cys Thr Asp Leu Met Phe Asp His Thr Asn Lys Gly Thr Phe
65          70          75          80

Leu Arg Leu Tyr Tyr Pro Ser Gln Asp Asn Asp Arg Leu Asp Thr Leu
85          90          95

Trp Ile Pro Asn Lys Glu Tyr Phe Trp Gly Leu Ser Lys Phe Leu Gly
100         105         110

Thr His Trp Leu Met Gly Asn Ile Leu Arg Leu Leu Phe Gly Ser Met
115         120         125

Thr Thr Pro Ala Asn Trp Asn Ser Pro Leu Arg Pro Gly Glu Lys Tyr
130         135         140

Pro Leu Val Val Phe Ser His Gly Leu Gly Ala Phe Arg Thr Leu Tyr
145         150         155         160

Ser Ala Ile Gly Ile Asp Leu Ala Ser His Gly Phe Ile Val Ala Ala
165         170         175

```

-continued

Val Glu His Arg Asp Arg Ser Ala Ser Ala Thr Tyr Tyr Phe Lys Asp
 180 185 190
 Gln Ser Ala Ala Glu Ile Gly Asp Lys Ser Trp Leu Tyr Leu Arg Thr
 195 200 205
 Leu Lys Gln Glu Glu Glu Thr His Ile Arg Asn Glu Gln Val Arg Gln
 210 215 220
 Arg Ala Lys Glu Cys Ser Gln Ala Leu Ser Leu Ile Leu Asp Ile Asp
 225 230 235 240
 His Gly Lys Pro Val Lys Asn Ala Leu Asp Leu Lys Phe Asp Met Glu
 245 250 255
 Gln Leu Lys Asp Ser Ile Asp Arg Glu Lys Ile Ala Val Ile Gly His
 260 265 270
 Ser Phe Gly Gly Ala Thr Val Ile Gln Thr Leu Ser Glu Asp Gln Arg
 275 280 285
 Phe Arg Cys Gly Ile Ala Leu Asp Ala Trp Met Phe Pro Leu Gly Asp
 290 295 300
 Glu Val Tyr Ser Arg Ile Pro Gln Pro Leu Phe Phe Ile Asn Ser Glu
 305 310 315 320
 Tyr Phe Gln Tyr Pro Ala Asn Ile Ile Lys Met Lys Lys Cys Tyr Ser
 325 330 335
 Pro Asp Lys Glu Arg Lys Met Ile Thr Ile Arg Gly Ser Val His Gln
 340 345 350
 Asn Phe Ala Asp Phe Thr Phe Ala Thr Gly Lys Ile Ile Gly His Met
 355 360 365
 Leu Lys Leu Lys Gly Asp Ile Asp Ser Asn Val Ala Ile Asp Leu Ser
 370 375 380
 Asn Lys Ala Ser Leu Ala Phe Leu Gln Lys His Leu Gly Leu His Lys
 385 390 395 400
 Asp Phe Asp Gln Trp Asp Cys Leu Ile Glu Gly Asp Asp Glu Asn Leu
 405 410 415
 Ile Pro Gly Thr Asn Ile Asn Thr Thr Asn Gln His Ile Met Leu Gln
 420 425 430
 Asn Ser Ser Gly Ile Glu Lys Tyr Asn
 435 440

<210> SEQ ID NO 2

<211> LENGTH: 42611

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

```

cttttctttt ttttctctct ctctctcttt ctttcgttcc ttctttcttt cttttgaggc 60
ggagtctcat tttgtcgcca ggctggagtg cagaggcgcg atctcgactc actgcaacct 120
ctgtctcccc ggttaagca attctctgc ctcaacttcc cgagtagctg ggactgctgg 180
cgtgccgtat cacgtcgcc taatttttttaaaatggat tagagacggg gtttcgttgt 240
gttcgcacaca atggctctga tctctgtacc tcgtgattcg cccgcctccg cttccaaaa 300
tgctgggatt acaggcatga gccaccgcgc cccggccggaa gttttcttat ctgttagatct 360
cctacaggga caatctctct ctccgtctct ctctctctct ctctctctct 420
ctctctctct ctctctctct ctctcacaca cacacacacc actacacacc tctcactctg 480
cctcctctct ctctctccct ccgtctctc tctcacacac ccttctccct ctctccctct 540

```

-continued

ctctcaactc	cccttgagg	agtattgcgt	tttggagcc	cagttatgcg	tggctgctgt	600
aagaacaaaa	caaagcac	tttgcctac	ctggagcatc	cccctgccgg	ccagataaga	660
tcacagctgt	gtcaccaagt	caagggagag	gtcattttt	aaaaaattaa	gcaaaatcag	720
aaaacagaga	acactgtctt	cctcatcatc	tcctggagct	tgttaagaa	atttctgtca	780
ccaaaaacga	aaagtccgg	cttttattca	tgctgttgt	tcttggatt	actgcagtgg	840
tgatgaggaa	ggagagggtc	gtatggctca	taaatggcag	cttgcacact	ttaaaaggca	900
cacggattca	gatttggat	gtgttaacaa	atcaggctt	cacaattac	ttttgatcca	960
aatgagaatg	ctgagctgg	agactggat	atggtagga	gctggaaat	gtatcccctg	1020
aggtgttggc	aatcccaaga	ggctcagcac	gaaaccttga	gcttacattt	aattcagaga	1080
tctgagcatt	aagactggc	ctgatcaaga	agaagcagat	ggccttctgg	tcacaggcga	1140
ctggtagcc	tccatcaccc	aggcagattc	cctcacagag	gaattctgct	ggccctgcag	1200
aaagtggta	gacagcgcac	cacccggca	tttccagca	ggcctacttc	agctccttgc	1260
acttggcac	tgaaaatgt	tcccttccc	cagatattt	tgccacagaa	gctgctgt	1320
gccatgtgt	gatgtctcaa	tgtctttctt	cattgaactt	ggcccttcat	acccaaggct	1380
ggcaccatag	tctgtgactt	cctactgtac	agccagcaag	actgaaggag	ttgcagcatt	1440
aggagtgtt	tgaatctgtt	tcagccactc	attttacagt	tgagaaatcc	cttctgaggt	1500
gtggaaaggc	ttgcttagca	ccgcaggact	tttccctcag	acagaacctg	ggacagctgg	1560
gtccttattt	cctctccatt	gtgtctgtt	cccccttgc	cttttgtt	gaggaacata	1620
catggccta	aacacattcc	ctcaactctca	cctttgcaca	tgaatatgc	ttaccacact	1680
ctttgaagct	ctgtcagcac	agaacattaa	gaaaagtgt	aactgggtac	ctgagcagct	1740
ggagtgggag	ccccctggaa	agtcatgtt	ggagtcccct	ttatcttcct	tttcctgata	1800
cccaagaaga	tggctattt	ctggaaagga	aggtggat	tggattattt	acactctgtat	1860
tcctcttaag	cctaattttt	tttcaagatt	cgtgataata	agcctgactt	acgtctcaag	1920
atgttgagaa	atttccataa	aataaaagca	ttgaagagaa	agtactgagt	gaagggaaat	1980
tgaagagaaa	cattcagttc	taaaattgt	ttgtttcaga	atatgattt	gggttcttac	2040
tgacttttt	catgacycca	gccacctgtt	agacacatta	tgtgcctca	ctgttattt	2100
gctcaatttt	atggcgctga	ctcaacacta	acagaaacag	gctaatacat	acaaacatgt	2160
aggtttttgg	tttttaaaga	aacttagcag	ttgaggagaa	ataaaaagtc	tattgttaaga	2220
gtagatgctc	attgaattaa	gtggctagt	tgattcctag	tagaattt	tttatttt	2280
tcattcactt	caaaaaat	ttattgtca	gatcctaagt	gccaggact	gagtagcaca	2340
cagttat	gatagatat	ctctctgtc	tcacagtgt	cacattctt	aattgaggga	2400
tgaacagggt	gaattcatat	aatgagaaga	aaagcaagca	gattgctgca	gtatagttgt	2460
gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtatgtgtgt	gttggatata	ttggaggggc	2520
ctagctctt	ggatggatt	gtgaggagag	acttctctgc	aagaaggcct	ggagatgaaa	2580
cctcaggaag	atggcgac	gtccctgcac	agggaccac	aagcacaat	gtcctgctcc	2640
tggAACAGG	tgtgtatgtt	tggaggcaca	gttggatgc	cagtgtggct	ggagtttaggt	2700
gagggagaag	gagagtggaa	ggagttagct	ttagggcag	gagccagtc	aggaagaccc	2760
tcataggcca	aggaaacat	ttggatttt	ctcttacaa	ttagatgtt	caaaagagag	2820

-continued

tttgaagcag agatgtgaca taattgattt actttataaa gagatgactc tgggtgctat	2880
gtggagaatg gattacaggg gaaggagagg ggacagaaga aggccagaag cgagtttact	2940
gatgggggggt ggttagcagta gatggcgta ggctatttt tggagttaca gtggaaagtt	3000
aatgtttctc tgcgtcaat ttcccatcta caaaatggga atgagattgc tatacctgg	3060
ccacttgtaa aaactaataa acaaataaat atgcctgaa gcataatgta agacattatt	3120
accattatta tttgtcata ttgttattact ttgaattccc tgcgtttat cttaaaat	3180
attttaataa atttgatgtc ttattataattt attttcagta aacctaattt attatacata	3240
aaagtatctt tctctgttcg gagaactgaa aactcttaac taagatgttca aataaaaaac	3300
aagtcgtcac aaaatgcagg gaaattatttgc ctaaaataat agctggctta atttcaaaag	3360
ttgaatatgg tcagataaaa attcaaaaat agaataatta aagctggaa tggaaaaaaa	3420
atattttgga cagatttgc tttgtcctgg ctgtggcat gcagatatct gtgagtgc	3480
agatttttagt ataactgaaa atatctatataa ggcagcagaa ttctggaga agtctgtgg	3540
ccagagttgg gttaggttgtt ttccttagt gaagcattct ctctcagagg aagctcttgc	3600
attttcctc gtttatcact aagttttaga gtgtcagaa gttttaaagt tgtttaccaa	3660
ctaggatagt gacgttcttc gcagagacag gacaagtaac tgccacaata aggcaactct	3720
atttctcttt cctgactcac tctgtatgtc aagttacttag aaggagatgc gcctgttgc	3780
aggaaaaaaa gctcttcctc attttttagt tggcagcagaa aatgtcagc tcgtacttgc	3840
atagtcccta ggtcatgttacttgcaga tccccatataat tattttatgg tcagtgtgtc	3900
tgaatttcag cttgaagtgg ctgtctggt gattcaaaat gcagacaaaa tgtgatagta	3960
attctggat ctgtatgatt tatctcttgc catatcaacc accatatttt tcttagggct	4020
ggagagttggg ggtgtttggg tatcttaaag ggagaggttacatctctc tgaagccaga	4080
aatctatttg tttcactcac gttttgaaag gttgaatttag ttgctcctt aaaaatcagt	4140
ggatattaca tacaaatgta atttccagt tttcttggaa agcaatcaga tgatctagt	4200
acatggagct gccattccta catgaaagca attaactgga atagagtggt agttgccttc	4260
tttggatggt ccagtttgc ccatccaaat agaataagga cagttcacgc cagtccttat	4320
tctattgtcat gaacctgttt gattcattta aattacctgt ctggcccttg tagacttttgc	4380
atttttact tcaggttttag ggcaatcaag tgtaacctgt ctgttataact ttcacccat	4440
tgtaatgcag gaatctatatac tccagttatcc ctgtgttagt tccccaaacc tgcactttaa	4500
ctgtctctct gataaagaag ataatgttataa ttgttggaaa gctactgttg atggagctgt	4560
acagtctggg cctacataga acatgttttag gaactttgc caggactgtg gctgtcacat	4620
gcaacccatgg tttccacat taaaagttaa cattccatcat tcctcaaca gtcctcatg	4680
tgacttgctg ccttggccc tcatcattct ggttattttt gggttactgt ccctgtcct	4740
gaacccatgg tttttttat ctttaaatcc caaccaattc aataagatgt ggtctactgt	4800
gtgcagaatg cagtggagg gtggcctccc ttctctatg atgtcataact ttgtcatgc	4860
aactgaagat gtccttagat ttttcctggc cttatccatttgc tttttttaaa ttactttca	4920
gggcttattt tccagttaga ttgttccatgtt gaggatcag gagaactgctt caatgttacag	4980
cttggagagca atgggtttca gctgaggatg atttatgtccc ccactccagg ggacatctgg	5040
caatgtctgg agacgtataa tttagttgcc acaactgggg tgcgttatttgc tgcattttgt	5100

-continued

ggtagaggcc	agggatgtt	ctaaacaccc	tgcaatgcac	aggacagccc	ccacagccaa	5160
gacttatcag	cccctagtgt	cagcagtgt	gagcttaaga	attcctactt	taggtccgg	5220
gtctaggttc	agatcccttgt	ttagctactt	cctatccttg	tgaacttggg	cacgttattt	5280
aaccttttg	tcttcattat	ctcaactgta	tgtcaaagaa	ctctgaggac	taaatgagtt	5340
catttatatg	aaatgcttag	aaggtaact	gtatgtacaa	acactatctt	attttctac	5400
catattttat	tcatttttgt	gtgtacatat	gacatcagat	aagtgggtgg	gtaaagtttgg	5460
agggatagca	cttcatttct	aatcttgctt	atgggatggg	cactagtgca	tttggatgtt	5520
ttagggagct	aacagctttt	ctcagaaaat	ccttctctct	atggccccca	ttatcattttt	5580
tccggcattct	tttagtattt	tcctccatat	tggggcatat	atgtcagttt	aatactgcat	5640
cctcctggag	aatatggaca	ctgagttcct	ttagcgggag	agcatctgt	cctgttgctt	5700
gtacatccca	tttccgggtgt	gccccaggct	cctccacacc	ccagtcaccc	ttggcttgcac	5760
ctgcacaaca	gattcctggg	gctgtgtagc	aattggtaga	cttggataaa	taataatagc	5820
atttctccaa	gggctcacat	ccccataaga	accccattat	cagggttcct	ctttgacaga	5880
tgaggagaga	aaagcagaaaa	gtagttaaga	acttaactga	agtcacccaa	ctagtaagta	5940
gcagattcaa	gattcaaacc	ttggcagcca	gaaactggag	ctaaatttct	aaaccacttt	6000
attactctga	tacaggttaag	ttttctcaag	aagtatacc	atgaccttaa	cctaaccatc	6060
atgtagccat	ttatgcatac	aacactcaac	agatccagtt	actcttactc	ttactattgt	6120
gagtaataata	gttggggggac	tgagtcttag	agtaagatca	ggaacttaca	gtatagcagc	6180
aggctgccc	ttgatcttag	acaactttta	ctttttatga	cctttaagcc	atatagaaca	6240
ccaaacattt	ggacatttgc	acatggatg	cagattaatt	cccatttcca	aaattgtgag	6300
tagaatcaac	taattatatg	ttacttgcag	gttttgcac	tcatgcac	gcctccctag	6360
aatgaaataa	aatttcttcc	ttattaaaca	aggatagagt	tgccccagct	tttcttagac	6420
cagggtttct	caacccatc	agtattgaca	tttttagatt	cctctatgt	tagggctgtc	6480
ctgtgcacta	cagatgccta	acagcatccc	ctgactgcta	ctcactagat	gccagtagca	6540
gccccctcct	ggttgcacaa	cgcaaaaata	tctccagaca	ttaataagtg	tcttcttggg	6600
ggccaaaaat	gccttagtt	agagacactg	ctgtagagtg	atgttctagc	tgcatctcc	6660
agccttcatc	ctatgtaaa	atatccaaag	aagattttat	cttcttcagc	aatcatttgg	6720
tgtggtaatg	ggtcagttt	gaggtattt	agccacttcc	tttgacttta	gttggaaactg	6780
aattgggctta	ttatggctt	tttgtctcat	tcagttactt	tatatgctcc	atattagttt	6840
tctattgctg	ctgtaaaaat	tttctacaaa	ttttagtca	taaagcaaca	taatttatta	6900
tctcatggtt	ctgaaagtca	gaagtctgc	atggccttac	taggctaaaa	taacagggtt	6960
agtgaacctt	cagaggtcag	ctgtctgt	ttgctttcct	gaggctctag	agagaatctg	7020
tctccattacc	tttccaaat	tctagaggcc	atccacatgt	tatagtctcat	gcccccttcc	7080
ttcattttca	aaaccagcca	tgggggttg	agttctcaca	tcacatcatt	cagatcttt	7140
cttctgcctat	cctcttatac	tttaaggac	tcttgcatt	acattgggtc	agataatcca	7200
tgataatctc	cctattttaa	ggtcaactga	ttaacattaa	ttccatttgc	atcttaattc	7260
ttttttgcca	tggAACATTA	catattcaca	tggctgggg	tttagaaat	ggacatttgg	7320
gtgggaggtgc	attattcaac	ctgcacatg	ccccccattcc	aaaggtaat	caggaagaat	7380

-continued

ttggaaagag agtactgata cttAACAAAT gAGCATTTC TTTCTGACT CTTTACTAAT	7440
aacacaagag ctaaggccca CCTGCCCTAA TCTCTTGCCTA GGACTGGGTT TTACTTGGAT	7500
ggTTTATCTG TCAGCTTTc ACAGAGATTt CAAACTCGAG ATTGGACATT TTATCTATTG	7560
TTTTGGGTTG TCTTTGTtT CGAAAAGGGA AATGATGCCT GATGAGAAATT TAACAGTGAC	7620
AAATAGGGGA AATCGGGATT TGTGGCAGTT ATGGTTGTGT ATGGAGACTA CAACTGAGAT	7680
CTGAAGAACCA GGTAGGGCAC TGTGGTTCA GTGGGAGAAA TGAAGTCTGC AGGGAGAGG	7740
CAGAAAGGTA GAATCCGTTT AAATGGAATG TATTATCTGA GTGCTCATTG AAAGTCTTCT	7800
TGATTGCCAG GATGATATTt TATCTAGTGG TCAAGGGCC AAATTGAAGA TAGTTATTT	7860
CAGGCCTTAC TATTGAAAGA AAATCATTGA TTTTAAAGT GAACATTCCC ATTGATTAG	7920
AATGTTGGTG GCTCATGCCT GTAATCCCAG CACTTGAGA GGTAGAGGTG AGCGGATCAC	7980
ATGAGGTCAg GAGTCATAA GCAGCCTGGC CAACATGGAG AACCTTGTc TCTACTAAA	8040
ATACAAAAAT TAGCCAGGCA TTATGGTGGG TGCTGTAAAT CCCAGCTACT CAGAAGGCTG	8100
AGGCAGGAGA ATCACTTGAa CCTGGAAGGC GGAGGTTGCA GTGAGCTGAG GTGGTGCTAT	8160
TGCACTCTTA GGTGACAGAG TGAGACTCCA TCTCCAAAAA AAAAAGTTT GTGTTTTGT	8220
ATTGACCAGT GATTGGGGTT ATGGTAATAG ATTTCATAC AATATGATGA TGAAGACGGC	8280
TGTATCAAGG GTTTAGTGT AGTCCTGTG GATTCTCTAG TTCTAAGAGA ATGTCTAAAG	8340
TTACCACTAA GTTTTCAGG GTCCTAGTTC TAAATATAA ATACATTCT AGCTAACTCA	8400
TATCCATGG TTGAGAAAAA GGCACCTTAC TGTGTAAGA GAACTTTAA AGTTAAAAGA	8460
CTTCTTATGA ATTATTCCtG TGTGTTcATA TTCAGAAGCC TCTGTGGTCA TTGTGATTAT	8520
TACATGTTAA TATGGTCCtT GGCTCATAAC TCCCGTTGCC CTGCTACAG TCTTGGTTA	8580
TAATATTGGA GTGCTTAgA CCCAGAACG AGGTCTCAGA AAGCAGAAcT TTTCTCTGT	8640
ACCTTTCCtT GCCTTCCtT CTACTGCCA AGGCAGGACT CTAATCTGAT TGTAGGTTAT	8700
AAGACCTTA TTCAGGGAAG GGTCTCTTA CCTTGGAAAA AGGAATGCTG CACAGACAGA	8760
CCAAGAAAGA TCTGAATAGA CAGGGCTTGC TGGGTTGCC CTCTGAGTCT ATTACCTTA	8820
GAGCTTACCC TTACTGTCCA GTCATATTcC TACACAGTG TTCATACTTT GTGAAcCTA	8880
TGCTAAAGAAA TAGACAGTTT CCCCCTGTAC TTTGGGTCTT CATTCTGAAG GTCCTTGT	8940
AATACACATT AAATAAAATA TCTGCCCTTT ACAAGTTGAT TTTCTGTAA CCCTTCAGAG	9000
GTCCAAGGGG AATGOCACCC GTGGCCCCCT TGGCCCTTC AGTGTGCTAG TATTGAGGCC	9060
TAGAAATGCC TCTGGAAAAA AATCTGTGAG CAAGTCATTA TTTGCGAAGA GATAACTCAA	9120
TTTACCTTT TCTGTATGT TTGGATTTT GAAATCATGT ATTAACCTTA TAATATAATT	9180
AAATGTAATA CTGAAATAAG ATCAAATAAA ATCAACCAA ATTATAATA ATCTAAAGA	9240
GTTGCACATA TATACCTTT TTTCTGTAC AAAAGAAAGA TGTGAAATAA GGAAATACAG	9300
TCCCTCTGCA AGTAGTTAA GAAGTTTGG CTGGACACGG CGGCTGTGA CTGTAATCCC	9360
AGCACTTGG GAGGCCAAGG TGGCAGATC ATGAGGTCAg GAGATCAAGA CCACTCTGGC	9420
CAACATGGTG AAACCCCATC TCTACTAAA ATACAAAAAT TAGCCGGCA TGGAGGTGGG	9480
CACCTATAGT CCCAGCTACT CGAGGGCTG AGGCAGGAGA ATTGCCTGAA CCCAGGAGGC	9540
AGAGGTGCA ATGAGCCGAG ATCACACTAC TGCATTCCAG CCTGGGTGAC AGAGTAAAC	9600
TCTGTCAAAG AAAAAGAAAAA ATTAGAGCT CCAGCAACAC ATTACCAcG TGCCTATCTG	9660

-continued

ggctttctca gcctgtacaa taggagctgc ctttacacga aagccccatc tctttgaaat 9720
ctttctgtct acattccctgt tgctctcaga acttgtacca tatcatctaa gaattataaa 9780
atgcccctgt aatttgcttt cttccctgtat tatattgtta atatatgcta gtcttgcctc 9840
tgaaacacaa ttgttaaactc tctgagagta gggattgtgt caaattctct ggtacttacc 9900
acagttgttg ccataggaga gggcatatag tatttgcctta agaaataccc tttgattgaa 9960
tgactgcaga tctgataatc aggataggcc atgctgtgct ccattaaggc aataacagct 10020
tcacctggaa atttcagtgg cttaacacag caaagggtta tttccctt gtgaaaggc 10080
atttgaatttgcaggc ttctctgttc agtactagg tatccaggcg tctttcactt 10140
caagactgca ctacatcagc acatgatttc taagtttgcgtt gtgtcaggga aaggaaggca 10200
tgaaggagat acagtgcactc tgttgcagac gacttctgtt caaatttttg tttctcaggaa 10260
attagccccca tgggcccggc ccaaccacag gggcggtaa gaaatgttagg gaagcacata 10320
gatgttcagg agcactcatt gtctgtccca cacatctata ttacatcttc ccgcaaaaata 10380
gctatactaa cgagtgcctc attagtgtca gtgggacacc ggatcattaa acttggctcg 10440
gtttgttatta gtaaccagat gtgttagatt tatttcactca gcaggcattt actgagtgcc 10500
tctgtacacc aggcaactt ctaggtctt ggaatacagc agggggaaaa accatcaaaa 10560
atcccttacac tcacagaact aagttttagt tggaagggtga taaaacaacta acaaataaca 10620
tacaaggtac gtttagttgtt gataaatattt ggagaaaaat aaagttaggg gggaaatagg 10680
gaatgtattt gtgtgaagaa ggaggttaga ttgcaattttt aaatagttt aaccggggaaa 10740
gtgtcactga gaaagtatca ctggagccaa aaccggagg tgagaggacc agtcattttgg 10800
gtgtcagagc aaagagcatt cccataaaaag ggaacacatt caaagccctc aaagtggggaa 10860
tggaagggtgg gtgtggagta ctggcatgt ttggtaatg gcaaagaagc aagtgtgatt 10920
gcagcagggtg tccaagaggg gctaggggaa ggattggag tggttgagt aaaggttatg 10980
tagatcatttta taagtttttgggtgaatg gcttctcccc cgtaacatattt aacaccattt 11040
gggggttttg tgcaagagg tgacacaatc actctggctg ctgtgttggg gatagatttt 11100
agggtcaagg gcaaaggcag ggagactcgtaggtt ggcattttt aggttaagaga 11160
tgatgttggc ttggacctttt tctatccccca acaaaagttt aaggggccaaa tttaggggg 11220
gaaaatggat gcacaaatgc aagatagttt ttgctataaa gatgttagca tacagactaa 11280
ttataatttttta ttattaccac acatctcata ttctcttgcattt tcattttactg attcaacaag 11340
catctatgca atactacttc ttccaggccc tctatttaggc actgagacaa caaataagac 11400
ctgacttctg ctttttggaa acttatatca tagtaggatt atagaaatattt gcaagaatattt 11460
aacatcatat taacacccaa aacaaccaca gcagcagaaa tatagtactt atataattaa 11520
atgatggaaa tagccaaaat ttatgaatta ttactatgt gctagacata cctatggata 11580
tttattatgtt gatgacagat ggtttttataa ttgtttagt tttgttgctt atttttatag 11640
ctgaccatac tggatataaag ttggaaaact agtgacccaa aacttaatct aactgtctga 11700
aagtattttt tctctgtctt acaatccat gtaatcctta gtaatgttagg taatggtaat 11760
cctacattttttaa accattgttag taccaatcta aaacccagca cagaaaataac 11820
atgttttttttttccaaatgtt gttacttagt cctcagcctt tcttgattttt tcaatgtttttt 11880
taaggccttcttcattgcata cttctttttt ctttaatca tctgcttcga aggagactaa 11940

-continued

gctgaaactg ctgctcagct cccaaagatgg tgccacccaa attgcattgtg cttttctgcc 12000
tctgcggctg cctggctgtg gtttatcctt ttgactggca atacataat cctgttgc 12060
atatgaaatc atcaggtaag aggtgtatgg gttcaaggc ttgagcaact gatctgtcgc 12120
catactcaa gtggggccca agaagttgca catctgcaca tctaaacaag tcctatgg 12180
aggcttatgg agatcctgta ttcttcgttga gtaggggaga ggttctcaact ttttcctcat 12240
ttcaactgaa attgaagaag cacttgttgc ctccatgtc aggactgcg gacagaacca 12300
ggatgggaca ttttaagata tactggacgc tcagcttgca atcaccacgg agaacctcct 12360
taatgcacc ccttcctct ctgactgccc taacgcattgc ccacacaagc atagtaatgc 12420
aactggctcg gagtccactc caagacagcc tcaaaacctg ctccatgag gtgaaaaaga 12480
agcagggtgc ttgttcaaca ggcatactcc tggtaatag tgaaagtca tgtgagataa 12540
caatttcagg tcttggatc caaaaggaaa caatcttcc tacctgataa tagagagaag 12600
tagggacaag taggtgctag ggagttccag atgagaatgg atgtgatcac tcatttttc 12660
ttgccttgc gtaagtgggg aaagaaaagag gtgaaggca tagactaaac ccaggaggag 12720
tggccaaac gggacagcca tcttagacg gctacattta atgtcactaa gagagacctt 12780
aattggaaaa agattgttcc tggagcttgc caaattccctg ctcatcccagg aggattgtga 12840
tgcataactact tctctttcca gtctcctatg gagagcgtg taggtggctt ctgaatgg 12900
tcataactgac tatacctaca gagctgtgca cctggatagt tgtgggtcac tctggatcta 12960
attgttcaag cttgtatgggactcttac gatagcattt gcaaaagagc tttcataatt 13020
cagtagcttca aattttgcct tactacataa gcaggaaaat gaatgcatta ggagaggtt 13080
tatatottat atctacattt gcataggaaa tagataataa tctcacaggt ccattcatcc 13140
tataacgcattt caggaagtta agtcatcaag gatacactgg tgactaagac aataaatgtg 13200
tagtcaaattt caagggcttgc acatcttgc tgacattaag ttcatcttc ctactacctt 13260
tttcatggat tgagaacagg taatatctta gatgggtggaa aaatcatcat tcggaaaca 13320
tttataatcct tcacttttaa ttggtacagg gtaaaaattt aaacattttt aaaatccctt 13380
tgtctttaaa aataatttggaa aggctgtgc tggactgaca gaaacttgag gacagggacg 13440
ttgttttatac ttatagtc cagcatttgc cccaatgca tgccattctt aatgttcaa 13500
aaatgaatgc aggccgggca tggggctcg tgccctgatctt ctgagcactt tggggggccg 13560
agggtgggtgg atcacttgag gtcaggagtt cgagaccagc ctggccaaca tggtaaaacc 13620
ccgtctatataaaaaaaa aatacataaa attagctggg tgggtgggtg cacacctgt 13680
gtccttagcta cttgggagac tgaggcagga gaattgttgc aacccagaag aaggagggtt 13740
cagtagctg agattgcacc actgcactcc agcctggctt acagagcggactctgtc 13800
aaaacaaaca aaaaaagaat gcatatataat tggttaaag tattagaaac agtttattaa 13860
ataagattaa aataataggt gattttttt aatttcttac taaatgataa cacagaaaat 13920
atggctatgc aggttaaaag gttttagga atggctaga gatattcttctt tactaatgg 13980
aggacataaa gccatataag agtctaatttgc tccatgtc tatacagttc atctccaaac 14040
ccaagactga acatcaggaa tatttgacaa atgaaaacca ccaaaaatataat gtaataaaat 14100
taatagcatg ccatagggttc caatcctctt attaaaataa agttaagaa aactaagaag 14160
agacatttat aaattctgaa ccacaaaactt gttgtggatc tcatttgatc tgatgatgt 14220

-continued

cctccttcca cttccatataat gatgactcct gacagatgg tagagaatgg gaaatccagg 14280
cccataaccg ggtcaagctgt aggctatgg a gcctgttcca c agccaccc c caggcacct 14340
ggcactggtg cataggaccc catagaccta t accaaaacc tggatcagcc tgatgattcc 14400
tgaggcttc tcagaggta ggtgcctgg a acaggcaga c ctcacccctg acctgaatgc 14460
tcctgaatgc tgagtggaga aacttccag ggagaacagg t gttggacaa ccagggact 14520
ccctaacttat tggaaaatgt tccctgtcc tttggaaattt agatctcaca agcggaggct 14580
tgtgtttggg t gaaaccctc ccatgcctcc agaactgagc ctccatttct agaggtatgt 14640
tctccaacag c agctgtggc cctggggcctc cttagcgagt t g t g t g a g c c 14700
ctgcagcaag ccctgcacag gcacaccatg cctgcctagg ggtggggttt c agtagcacc 14760
aggctgttta ccacagaaga agaagcagg t g c c t t a c t t a a g g t t g g a 14820
cttcacatct ccctgttaca ccagaagcat ggctcagtag atacttttt aatggaaagaa 14880
gagcctttt g taattaaaa tagaccctt a a g t t c a c c t t a g c a g g t a a a 14940
aaagtttctg gggtttgaat ctacttgggg agaggtaga cctgtctaaa cttccctaaag 15000
gactcagcac agatttacact agcaagagta aatggactct ggggaaggag taaacagcat 15060
tattttgtcta ccccttatgt gtcatgtttt g t a c t a a g g a a c t t g a c c t 15120
catctttaca actgtactgt gggctggta a g t g c t t t a t g t c a t t t a a g 15180
tcagagaggg gagcaactag gccagaggat tata gctt g t g a a g g t a c a c t 15240
gaaatagaac catgtccctc caattcacag ggttggta c t c t c a t c c c t t 15300
gcctttctt g a a c t t a c t c c t a c c a t t t g a c c a t g g a t t 15360
ggtggcttgg a a c t t c a g c c t a g t g t c t t a a a c t t t a a t g g a 15420
gagtgaaattt g t g g t g a a g t c c a g g c t g c a t t t a a a a g g t t 15480
tgctggtcca tggatcagct ttgatcagct g a g c c t t a t g t a a a a g g a 15540
ggcagggtt t c a c a c a t g a g t c a c c a t g t g g a g g c t g t a a 15600
ctcaaaaactt ttctggcaa t g t t c a t g g a t t t a a a a a a a a a a a a 15660
tcttgatagc tggatcagc a g a g t a a t c a c t c c a c c a t g a a a a a 15720
gacaagctga atatgtcaag a a a a g t g a g a a g t c a a t g a a t t t t 15780
tatggat t t g a g g t t t t a g a a a g g a g a a c c g g a t g a a g g a t t 15840
gcaccagaga g a g a a a a g t a 15900
ggtttaaagg a a a a c a a g a g a g a g a g a g a g a g a g a g a g a g a 15960
agtatttgg t g a g c a t t g a c a t g a t g c a t g a t g t a a a a a a a a 16020
atatggctct g a a g a a a g a g a g a g a g a g a g a g a g a g a g a g a 16080
ggaagaatga ttacatgtgt g t a t a c a c a c a t g a c a t a g a t t t a a 16140
tgaaatgatt a a g t c t g t c t g g g a a a t c a g g a g g a g a g a g a 16200
ccaaggattt a a a g t g t g a a t g g a g t t t g a a a g g g g a t t t a a 16260
ttacaaagg t g c a g t g g c a t t a a c c a g a c c c a t g t g t g t c t a c a a 16320
ggaagggtt a a c t t a a g g a g a a g a g a g a g a g a g a g a g a g a g a 16380
tggagggtt t t t t g c a g t g c a g t g c a g t g c a c c t g t g a a c c t 16440
agtatgtttc a a c c c a g a g a g a g a g a g a g a g a g a g a g a g a g a 16500

-continued

ttggcttagg aaagcatgtg tagatttaaa ttcctcctac aaactgcctt ccaaaaagat	16560
tacaatacat tcctccctca agggaaaac acttgctgat agggaggcat tgaaggattt	16620
taagaattgt ctggtgtat ttgtgcttta aatccctcca ggagctgtgt aaaggatgca	16680
tttgaagggg tcaagatcca acgtgggtag aacagttagg agagagagca acagtctgg	16740
caagagacag tgagggcata aatgagacag ttgagctaga gggteagcga tatattgtgg	16800
agatcattct aagctaaaaa ctaatgttag caatagaggt ttggggctgg gcgtgggtgg	16860
tcacgcctgt aaaccacgca ctttgggagg ctgagggtgg cggatcacga ggtcaggaga	16920
tcaagaccat tctggccaaac gtggtgaaac cccgtctcta ctaaaaatac aaaaatttagc	16980
tggcatggt ggcacatgcc tacacctgta atcccagcta ctcaggaggc tgaggcagga	17040
gaatcacttg aaccaggag tcggagggtt cagtgagcca agattgcacc actgcactcc	17100
agcctggcaa cacggtgaga ctctgtctca aaaaaaaaaa aaaagaaaaa gaaaaaagaa	17160
agaaaataga ggctccgtt aatgtagggt ggttggtgag aaagaggttag gagaatagga	17220
tgattcccaag gtttctggct tgggtgcag gagatgagga gttcttgagt tgcaggctgt	17280
gtgggtggaa agtgagaagg gagatgttga agatgtatcat ttgtttcat ttggaaact	17340
agtttcattt tgacattctt atgggaagga catcaggtag agatgtctaa ggagaagttt	17400
taggcacaaa tctgaaatgt gaaataaaatg tcatgttttag gaaaaacaga ttgtgaagt	17460
catcttaat ttactctgga tagaagttaga atgagagggc aaggctgcaa acattaggag	17520
gtAACAGTCC aaggcagctt gagagaaaagg ctatgtctac ttcatctct ttaccctcca	17580
aaaccctac acagtgtttc aaacagagca gaccctcaat aattgcataat cttacttgg	17640
aggttgagaa agaaaagg ccagataact atggaaagta acttgattcc gttggattc	17700
ttttgcataa taaaatctga tatgtaatgg atgacaatgt agataatatt tacctgtttt	17760
tcagcatggg tcaacaaaat acaagtactg atggctgctg caagcttgg ccaaactaaa	17820
atccccggg gaaatgggcc ttattccgtt ggttgatcag acttaatgtt tgatcacact	17880
aataaggtaa tgcttgatt tatacaactt atccgtatac tctaataattt tctgtcgcta	17940
tggaccacta gaagggttca aaatgtgacc ttgcctcac ctgagaatga ctcatttgc	18000
aattttgtatt gtttcatatg aaggcttttta ctatgttgc cattcctcaa ttctttgtta	18060
ttgtctgatt aatttctcta taaaaccttat ttccacttc ctaataacctt gaagccaggc	18120
tgcttgatt ttcccttcac tgagatagaa tattgtttt ctgtttctct ttcatgacta	18180
tcttcaatca ccacacgacg ctaaaaaggctt cttagacactt ttgtgttgc acagaggtat	18240
ttgagttcccc actaattaaa tatgcaaaat agctgttgc atatgtttga gacacaactt	18300
ctctaaaaatgt gcattaatattt ctttcttaac agggcacctt ctgcgttta tattatccat	18360
cccaagataa tgatgcctt gacaccctt ggatccaaa taaaatattt tttggggc	18420
tttagcaaaat tcttggaaaca cactggctt tggcaacat ttgggttta ctctttggta	18480
agatttctgt tgatccttot ttgtggctc ttgcgttat gaaaaccttg aaaacaacaa	18540
gaacttcaag tagttaaagac caaagtagat ttttcttcag tccaaatagc tcctaaaatg	18600
ataaggaaag tatttctta aagccaggc aactacgaca gaatcaaggt tctcattttgc	18660
tccattctga gttggatggg agtggccgag agtacgttgc acatctgttgc atctttctgt	18720
ggctgtctt tagtttcat ctgacatacc atggagaagg caataccgtg gtgagaatag	18780

-continued

caggttattt ggtgtggcca ctatccgca tgctctgtgc taaagttaga caaaaagaga 18840
aagaaaatgt aagccatcaa actctccaga tccaaaacaa aggagtcaaa gctaatgcct 18900
ttcttgcata acatgaagag actctactct cattccatc acctacccc atcgctaaga 18960
caatctgtat tgcatttc acttactgtc tacattctgc taatatagac ttttcattcc 19020
ttatctgttt gactagacac tactgttgg a ctgtatgtt aa ctgtgttcc acatttaccc 19080
ctcaaaacga tctacaatt ttgggattca tcttaaacat cggctttaa ttttctttaaa 19140
aaccacttat aacagacatg aaaatagcat aactctgtat atttatgtt gaaaataatt 19200
tataatcac ttccatagca atcatcacat tcagtcattt tgactgcac tgcattttat 19260
agatgagaaa atgaaacaca aatgattgt gataataatg gtgacaataa aaattgagaa 19320
gtgaaagaag acaatatttc cattttttga gcccttact atgtgttgg ttgcaccact 19380
gcattacata tattaccaa ttaaataaaa ttctcacaaa aaccctatga aagagatatc 19440
attttgcata gtttacagaa gagaaatctg aagtacagaa agttttagga acttgcacca 19500
gatccatagt taggaagtgg cagagctgca tccatgactc agtataatgt cagtggtttc 19560
tctacccattt cctgcataatgat tatgaccaaa tcataatgaa cagagtctct gtttctttaga 19620
ttctggctct tagtgcataatg tatccaaggat tttatggta gtttggtttc agaacattcc 19680
ctttaaaatt ttccaaaac ctggaaatgg cagtaatgt cagcccttgcataatgatc 19740
ataagtggaa aaatctctca ccttaagtcc agagctgtca gtaagaactc atcgatataatc 19800
atccatgtgt ctgaggaaag aacttacgccc ggtgcgggt gctcacgcgt gtaatcccc 19860
cactttggaa ggctgaggca ggtggatcac ctgagggtcgagttcacgca ccagcctgac 19920
caacatggag aaacccatc tctactaaaa atataaaaatt agccaggcgt ggtggcgtat 19980
gcattgtatc tcagtcataatc gggaggctga ggccaggagaa tcgcttgcac ccggggaggca 20040
gagtttgcag tgagccgaga ttgtgccatt gcactccagc ctggcaaca agagcaaaac 20100
tccatctca aaaaaaaaaaaa aaaaacttaca gttcagctct ttgggttggt ggtatctatgt 20160
agcagtcattt ttaatgtatc tactattcat ccataaaaaa gttagatataa atcagatggg 20220
tctgcatttt atgctaatgtatc gatataatc aaatttacta gcaacactca gagaaaacct 20280
taactataac ctccattgt tgcattgtt caatgacaaac tcctgcacac tggaaattccc 20340
ctctggggcc ttggaaaaaa tatccacttg ttgttttttc tcatggctt ggggatttca 20400
ggtaatgttt gagaggttga acaattttgg ctccaggaa taaatgacaa tttttttatt 20460
caagaaagaa atagoagatgtt ttggaaatgtc atgcaggccc ttgtctggag gagttgggtt 20520
tcctcaataaa ttggctgtgg gtctattgtat cagtcctaga cctgtctggt caagtagttt 20580
tttccctact atcagctcat ttggatttgcat ctcacagcag agaagaaagg gtgttgcatt 20640
ttctatagtt gtccttcattt attgtatataatc ttacactctt aaaattatcc tctgtaaatgt 20700
tttagaacttt tgagaaatccca attttagattt agctggaaa gactgttattt aagagaaata 20760
tctgtgaagg caagagaagg agcaggaaaa gatggggaga accttggac catgacacag 20820
gttggacacc tgaaaagggt atggggaaagg gagaaggatt ggataagaag agcttcagaa 20880
tgaagcatgg ttctaaagaaa tattgagccaa ggccaaatggaa aggagtgcgt gactatattt 20940
gttccatcaag agtccatgtat ctcataggaa tggctaaatg ttgttccatca tgctgtgc 21000
agtctctggg agcagctcac tggatgctt a gttccactt gaccatggcgt atggatacag 21060

-continued

aggggcagga gctggatagg gtcttttaga agctatgcag tctgttagcat gatagctgag 21120
tggcacat tt cataacaac aaccaccctc ttcaacttct tccaccacctt tctcacctgt 21180
cacccctctt ctacttcaag ggattgtaaa actcagaaaa taagagagat ggcttttagc 21240
tgggaagaag ttgggtagca cttgcaataa agaattctct tcagcactgt aaatgttagtt 21300
cagctagcta tttgatttct cttacattag tgactggcaa attagaaccc aaggccaaa 21360
tccagccac tgccttatttc tataaagttt tattagaaca cagccacacc cattcatttt 21420
tgtacactgt ctatggctac ttttgccta caacagcaga gttcagtagt ggtgacaaag 21480
actatggccc acaaaccctgg aaatatttc tctctggct tcaacagaaa aagtttgcta 21540
atccctgtgc tataatcatc cccctcttg aaaaggctt gatatctcg caataaataa 21600
atagtttcta gctacactct ttttattctc tggacatt tcctattgaa tgcattgtat 21660
gaggaaaagg ttgtgcaa at gaagctgtgg aagaggcat aaatgcagtg cttacacat 21720
gattttcaa ctatttcacc cagtcacatc atttaataa tagctaacat gtgttaagta 21780
tttacatgtt tctaataataa gtaactgaaa tgtattaatt caattgtatcc tcactgtat 21840
ggtagcatac tattatagtg tcaccat tttt atagataagg aaccaaggca gaaaaatata 21900
tgtgacttat taaaggtaca cagctaacaa gtggtagaac cagggttcaa actcagtggt 21960
cagactgcag gtctcattct tgaactaca atactgtact gctttctat atacattatc 22020
ttagcaaaac cttacaacta ccctaaaaca aatgcatttac tcttaggccc attttacagg 22080
tgaggacaaa aatgcagtaa ttacctaaaa tcatatagct gggaaatggc agagccaaga 22140
ttaaaaaccc aaaactctca aaggcaat at tataatcac tgccttgggt tttctcaact 22200
ttagtagtct tgacat tttt gaccagataa tctcttgggt tgaagggttgc ccctgtcat 22260
tgttagatgc ttgacagcat ccctttacc actagacacc agaaatacc tgccttgcacat 22320
tcatgataat gaaaatacc ccagacattt tcaaatatcc cctaaggaaac aaaatcaccc 22380
tcccttaaga accgttagtcc tccactccac cccaggac tactacaggg gtgatggc 22440
ctccatgttc ccagttttagt tagtgactc agccttgc tcatgactgg tagttgtat 22500
tcttcctct ttttgggtt aaggacactt tattctgtt tggcatttgc cctggcatct 22560
catgggttta tagttgtgc tggatggac acgtatgtt cctgtatataa ttgggttctt 22620
tggcaacta cagggatgt caatgcctt aactatgtt ctaatttca taaaagttta 22680
tttaaaatgt tggatggact ttcaagtatg gtaacatcat gagcaaaaaa ggagattgag 22740
ttttagtact taaaagactt aacatgttca ggtatgtt tgccttgcacat 22800
gtcctatgcg aatgtttt aatacgccat ttcaatttgc ttcatttgc cttcataaca atcctatgag 22860
atggtttata ttatttttgc tggatggatggac atatagaggt gataaagggt 22920
aatttgcacca aagtccacaca actatgttgc ggcagtgcgg ggatttgc tgccttgcacat 22980
agtttcaaca tggcttaat tgccttgc aatttgcacaca atatagggatggac ttttttttttgc 23040
ctgcaacttgc atgatagaca ataaatgtt ttttgc tgccttgc aatttgcacaca atatagggatggac 23100
attttgcacca aatgttgc tggatggatggac atatagggatggac aatgttgc tgccttgc 23160
gtcaaggatg gaactgagaa acatgggtca gatgagggatggac ggaaggagca tgcataaataa 23220
ataatgttgc ttgttattataa gagatagatc tgcatttgc aatttgcacaca atatagggatggac 23280
atctgttgc gaaatagggg acaagtcttgc gctcttgc aatgttgc aatgttgc aatgttgc 23340

-continued

ggagacacat atacgaaatg agcaggtaca ttgcagtgaa aggagaggtg gttggtgacc 23400
taaaaagcatg tacaaaagga tgacatttgc taatttatc ttacacctgg caagtttatgc 23460
tccttagctct cctatccc attcccaaaa gatctgtcaa tagattcctg gagcagtaaa 23520
atcccttaa tggatatctc agttcatgt aaaaacaaag gcaaatacaa aaatttggga 23580
gatgacagt aatattcaga attcccttga attaaaaactt ctaattttag aatctaaaaa 23640
gtgctgaaaa ataaatataat agattcttc tcacagacaa tataaaaaga ctactttca 23700
gagaggcagt atacaataga attaagagta tgggccttag aacccaactt tctatgtttg 23760
aatccctggct ccaccactta attttgtt agaataatga atagaaacat tattttgtctt 23820
acagagctat tagggcagtt aagttatgt aagcatattt gtttccttctt tatagtgggt 23880
gttttacatg tttttatgt tttgagtccca cccacattt ctttaggttag gaattttat 23940
acaatcaatg gaaactcaga gaggttaagt aaatcgctg aagtacata gtaggtaaac 24000
aacagagcca ggattggac taagctata ctatgtccaa agctggggcc tgggttcatta 24060
tggtagcaag taatagtac taatcagatt tccagttt aactgaccaa cgatttttcc 24120
caaatacagc ttctacctaa actttaaaaat aagtgttata actttttact ttgtcatttc 24180
cttcttctaa taatttatatt aggtacggca aagagcaaaa gaatgttccc aagctctcag 24240
tctgattctt gacattgatc atggaaagcc agtgaagaat gcatttagatt taaagtttgc 24300
tatgaaacaa ctgaaggtaa gctataaaaaa gtaatttttc tcttgcctca cagttcttta 24360
ttgttttttgc tcatatattttt ttcgttata ttgcaaggta caatatgata aagggtgtca 24420
accagcccccc tccccatgc gcacacacag acacacaaag cagttacaggta aagtatttgc 24480
agcaatgaag aatgcattat ctggacttag atatgaaatt gccaaaagtt agtcaatttt 24540
gatctacaaa aacagcaatg tcatatgtt caactcaact cctgtggaaat tattttttttt 24600
tcatggacag ctcctgcattttttaagct tgatcttgc aacttgacac aagttgggaa 24660
gcgggtggaa aaatatcaga cagagtggc acatgtacaa aagactcaaa aagtgtgtgt 24720
gcatgtgaag tgcagggatt agggggactt agagaataag atgatggctt gttctccaga 24780
tacaatgaag aaatttttta tccctcaagc accactctc ttctgttatcc ttgcctttgt 24840
acatgtgttc cccttggctt gacacacccct tcccttgcc taactcttac ctaatttca 24900
gactccagttt gacatcacc tccctctaaga agctttcttgc gaccccaata cccacttctg 24960
gactgggctc gctgtgttc atgtgtgttcc ctttgcattca ctgtactgtt ttgcattttt 25020
cctctgtata acttttttcc ctgtgtgttcc gcaactcac tggaaatgata ctgcgttacc 25080
tggcacagag taggtactca ataaatactc atgaaatgaa caaacaata aacatgggt 25140
gaggagaggc agaagtgcaga actgtgttttgc aagttttccag tgggtgtgac tacaagaac 25200
attaagtttta ctttcaaaacc tttacatgtt ttatataatgt gtgtaaatgt gttttatgt 25260
tgtatataatgt tttatgttgc tttatgttgc tttatgttgc tttatgttgc 25320
tattttataa gaaatcaatg tatttaacca tccctatgaa atgaaacattt atatgatttgc 25380
caaaatcatt tccctctaaca ccacaaataa gctataaattt tatatcatgc tttttcaat 25440
aggactctat tgatagggaa aaaatagcag taattggaca ttctttgggtt ggagcaacgg 25500
ttatttgcac tcttagtggaa gatcagatgttgc tttatgttgc tttatgttgc tttatgttgc 25560
agaatagtaa attattggaa gaaatttatattt tttttttat tcaatttctt 25620

-continued

agtgaagaag ggatctctt gagttataa ggctattctt ttgccccat aaaatactct 25680
atatacattt tcctaggcta aaacatctac ctctcctgct attaaaatct ccccctactc 25740
ccataagttt tccctcatta ttcttggtta cccaaggggt taacactttt cactgaaaaa 25800
tttatcttta tataattttt tgcacataa tgattgtat aataatattg tcatacttaac 25860
agtattgcca gacttctcca gtgggcagga catttttaa tttaggtttt cccccttcca 25920
cctttaagat ggagctggat ctAAAattac aattatttgg gtaagagttg ggacaagacc 25980
ccagttttct agcttgcagg aacattttct ttcttaaacc ttcatccag actgatgaga 26040
tcacatagaa ccaaccgtac agagagtggg caggatcatt ttgaaaactt cacctcaaatt 26100
ttcttaaaag tccatttcag tgtaagaatt ctcttcattt atccattttac caaaagatca 26160
ggaaacatag gataaaactta cttttaaaaa ggaagttgt ttcatcttca aatgatatgc 26220
tccatatgtt ttcttcaac ctcagacag ttttcagct gaagtgcctc taaaataaaa 26280
gtaatggcc caatttgttac tctaataaag ctataataaa tacactacaa ggtgcataaaa 26340
agctctgaaa acaaagggtt atttattata ctttttcaga tactccacag tgttgaagaa 26400
attattatgtt attctaaaat ttatTTTgaa taaaaagttt taaggatata tcactcaaatt 26460
ggttgcaatt cctaaaggca tataaagaaa atgattatataaaattccag gtctgctact 26520
ttacaagttt aatcatatcc cttttccaca ttgaagttatg atacctctt attccaatga 26580
gataaccat aataaactgg tatggtgcgt gtccaccaat cctagcatta ttaggatgtc 26640
ctcaatgtt gcttagtatgtt aaccagttaa atttcatcat tgcacaaaa tatctacaga 26700
tgcgttattt ccctggatgc atggatgtt ccactgggtg atgaagtata ttccagaatt 26760
cctcagcccc tctttttat caactctgaa tatttccaat atcctgttaa tatctaaaaa 26820
atgaaaaat gctactcacc tgataaaagaa agaaagatga ttacaatcg gtaagtatta 26880
gtgacttattt tcattatgtt aaacaaactt gaagcttggg taaaatcaa tcgatcat 26940
ttggtaacta ttaaaagaatt gctgaattgg ttgttttagac ttcaataag gagagaattt 27000
gataatctca gtttctaagt acatttagtc ttactttt taaaatggga atgttaacgt 27060
atatagtata tataactggttt atattatgtt gttcttgcatt tgctatgg aaatacctga 27120
gactggtaaa tttataaaga aaagagggtt cattggctca cagttctgca gcctgtacag 27180
gaagcatgat gctggcatct gctggcttc tggggaggcc tcaggaaact tacaattatg 27240
gtggaggca accggggcat gagoacttca catggccaga gccggaggaa gagagggatg 27300
ggtaaggta ctacacactt taaaataacc aggtctcaca agaatttact gtcacaatga 27360
cagcaccaag ggggatggtg taaaaccatg agaaaactgtt cccatgtatcc aatcacctcc 27420
caccagttcc tgcctctgac actggggattt ccaatttgcattt atgtgatggt gttggggacac 27480
agatccaaac tataactgtt gtaattaaaa tttagcattt actacatgtt acttcaatct 27540
aaacaccaga atatgcctac agatTTTgg ggttagagct agggggaaatgatttcat 27600
taggctgggg taggaactt taaaagaaaa agtcagatggatc tcgactgaga cctgcaat 27660
actaacctgt tgagaaaaa tatgatTTTaaatcccc aagctataa tggggatcca 27720
tgacgtgttgc gcaattcttgcattt tgcatttttttgcattt gttggggatggatggatgg 27780
ctctcgatgtt gatattacaa agcaaaatgat gacctgaaact cacccttccatctgtaaa 27840
tctggtttat ttaggagaag agaagaggca ggaaggaaat atatggatgatc taggcaataa 27900

-continued

acactatggg ctcagagaat agaggggtct ctccaacag aaaaggattc atagatgaca 27960
ctaaaattgg attttgcata atcaagagag ttttagatgt aaggaagggt ttaggtatTTT 28020
caggcaagga ctccatTTT cttatgtgta tgtcaactc ttcttaattc taaaacggaa 28080
ataatacaga acacctaacc ttccctcaac cctatttcct tactaagctc tcacccTTG 28140
ttttgtccct tctctgttag aaaccttcaa agaatagcca atattagtca ttgcattgtgc 28200
ttacttctat tcacttTTG actcttgcTTG tctagcttcc tggccccca cgggcatgcc 28260
aaaacaaaga cactaaaggc ctgcgcctg ccacaccaggc agccttcaac gagtctcaac 28320
cttgctctct gcatctgttc aacattgtcg accacccttc ttccctgtcc cttggctact 28380
atgaccacc tctttgtgc ttctccgtc atttatacca ctccTTTcat ttcttctcct 28440
ctggcagtTC tctacagtca gatcatccaa gttctgtc gcaattctct ttttccctg 28500
ctgttatct tctttatgtt tctgacattc aagtttccata ctacaatgtg gggaaagatg 28560
agagggagtt aagtggccct gtttgagct tagatgacta aaagaatgtat gggttgctta 28620
taagaaccag cagactcaag aggagagca gtttggca ggactggaa gatgtcttt 28680
aacttcacaa ttcttagct tgagaactga ggaatctgtc cacagatgtc tggccacaca 28740
ttggatgatt acataattta catgataaca cagatgaaac atttgaagta gaagagattg 28800
caaagtgaga aaagatgg tctgccagta gaatcctaag gaatgcctac attcagccaa 28860
tggaaatgg gtgaagagcc agtgaagaac atagagtgg tagcgtcaga agggaaaggT 28920
caatgttaag gaaacgtaaa gaaaaaaagt gcaagaaggc catcaagtgt caaatgtccc 28980
agtgggtgta gaaaacgagg actttaaaaaa ggccattcta ttgacgtta ggctattgt 29040
gaatttacag agcagtgcAA atgatattta taaagtgcTT acagttaat actggggaaa 29100
gtgggttttT aaatctgttt cctctaaggc ttAAatctaa agtggattaa atttAAatgt 29160
actagcatca aatacatacc acgttcagtg gtgagggcag gtgcggct ctggctctga 29220
gttcaggggac cctcaatac agaacacatt ccagtattca aactggaaagt attccaattc 29280
actaaaaAGC aagaatcatt tcttctaaaa tcaagatacc aagcaagaac aagattcttt 29340
gagttgtatt tctagaggga agaagaatAT actctggat ccctaaacaa acagcctgtg 29400
acccttggaaa cacatctaag tagatcaaAT tacaagtttt atttcttctt tggtttcag 29460
taaacagacc aacaagacca gtaccttct tacactctaa ctaaaaaat aataatttta 29520
tcaaaacaaatg tgactttaa atgtcttgcTT ctcttttagg ggttcagtcc accagaattt 29580
tgctgacttc actttgcAA ctggcaaaat aattggcacAC atgctcaat taaaggaga 29640
catagattca aatgttagcta ttgatcttag caacaaagct tcattagcat tcttacaaa 29700
gcatttaggt aagaaactat tttttcatg acctaaacca gatgaatctc aggacaaAGC 29760
tgtctatctt aatacagctt tagtactatt taaactatTT ccagttggTT tacaatggaa 29820
caaaggAGTA tatcaatttg aaaacagaaa tttgagaaag tcaatTTGc tgctttacat 29880
cctctatATC atagaaAGCA aatccaACTG ttaaaggtaa tattcttGT atgaAGCCTA 29940
gagttggactt ccattttgag gatactgaca gcaggTTGCC tcactcctat cccgtttgca 30000
ttcagctgct aaagcagCCA tgaggcagCT gatacagAGC acatcgTCtC taccatCCTA 30060
acggaaCTG tgtaatTTGT aaatctttat tgccacCTAG gggcacAAA ctgtttaaTG 30120
ctctcaaaAG TTTAATATGT tgatTAACAC tttatTTTtT ataggacttc ataaagattt 30180

-continued

tgatcatgtt gactgcttga ttgaaggaga tggatgagaat cttattccag ggaccaacat 30240
taacacaacc aatcaacaca tcatgttaca gaactcttca ggaatagaga aatacaatta 30300
ggattaaaat aggttttta aaagtcttgc ttcaaaactg tctaaaattt tgggtgtgt 30360
tgtgtgtgt tgggtgtgt tgagagagag agagagagag agagagagag 30420
tttaatgtt tttccaaaaa ggactcataat tttaaatgtt aggctatact gtaatcgta 30480
ttgaagctt gactaagaat ttttcctt tagatgtaaa gaaagaatac agtatacaat 30540
attcatatca gcctaaattt taatttaaa gatgattcct tttcagtgtc gaagttaaaa 30600
actgtttta cattactttt acagacaagt agattaaaac aggccaaatc ccagtggaaaa 30660
cctgttgc aa tgatacaaga ctccctaaac atagagtaaa aaacaatttc ttgcctttta 30720
ttatctacta tggcagtgg agttaattt tagcaacatg atattctagg tagaattttgg 30780
cagctttctt ctttgacttt tggcacggt aagaaaaatt agaacaagca aaagccattt 30840
tttcagcaca ggatgggtgg tactacacaa ttcaatgtt tgactaaaag gagtagagaa 30900
gggttaagaa tattagagat gaggcaagg cagtgattt aatttgcac 30960
tttccttaat tcttaaggc attctacttg cttaagaaag gttcctcagg tattacagca 31020
tatagcaggg attccagttt gaatacatta gttactcaca caaaggctgtt gaaggataat 31080
aaaacattca gaattacatt tagttcttc atattccagt taattttttttaatgtttt 31140
aattctat ttagtagtaa aacagaataa cacgattttt tggaggtaaa aatatacgat 31200
aagaatgtt acttataat atggaaatca tggaaaatgtt ttttattttcc caggggacat 31260
aaaaattttt accactcatt tggcaccta gtcctatctt agtgggtgtca ggtgcaaaat 31320
cactcattttt gagaatttt acttctgttag ccattcagat taaaattttttaataggaa 31380
gttaaactca gcttggaaattt gaaggcttcc tataatctga tcctaatgtca tcttcctgt 31440
gctttgcctt tggcaattca cattatatta tctatatgag taagccaaac ttgtctattt 31500
attattccag gaataagtaa tataactcattt tacctcaattt actttataatg caaacctctc 31560
tcccttctac ttacggaaat tataatgggt ttttaagggtt cataattctt tttccatttt 31620
gcttcagagt ctgataatgc ttccacagaga tacaccttta ccccagtctt attttagcaaa 31680
cctttccctt attcatttttta gaaataagca ttccaaactgtt atggatgtcc cattcaacac 31740
ttttttgtttt tttttttttt tacagtctttt tgggtgtggc gcatgctgtt gctgaaaggg 31800
gaaaataactt attgaaaata tctaataatgtt taagggtttttt aagaaaataac ctaacatagt 31860
gcacactgag ttttttcag ggtaaagaaa gacacattta tcaaattgcctt actcctttttt 31920
agggacttta tatcaattttt ttctcatctt aaaaaacac tggcaagaaaat gttttttttt 31980
caccacttta cagatgaaga tcaaaggtaa aaatattaca taatttgcctt atgtcacata 32040
tctgttaagc agcatatagc tcagatctgtt ttctaaaggc taagatctat gctttgc 32100
atcatgttttta aaaaatgttgg ggcttcaaaat ggtttttttttaatgtt gcttcaattt 32160
gagactcatc agctatctga ctttaggcaat gtatgtttaac ctctttatgtt agcttcaattt 32220
ttctcaccta taaaatgttgg ataataatgtt tggatgttgc tggatgttgc taaatgtt 32280
aaaaatgttgc taaaatgttgg agttaatgtt tggatgttgc tggatgttgc taaatgtt 32340
ctatgttttcccaatccctt cattacagag gagatgttgc accaaggataat tttttttttt 32400
ctatccctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 32460

-continued

actttattca aatcttgtat tagcatactt caaattcaaa cagaatgtta gtcaaatgtt 32520
 tagacaaaaca ggcaaagatt cagagtcaag aaaaaattc tgaaggagga tctgctaata 32580
 agctatctaa gtagaatcaa ataatgtact tgtcatgagc ccttattttt ctattnaat 32640
 ttaaacactg tatttggaaag gtaccaaata agttsaatgaa atgtaatgtg ttagtaaact 32700
 gtttagcaaa ataaatccctc ccaactaattt agaaaacgtaa gacaataaag acyaaagatc 32760
 ccagatctgt atatataaca ggatctactt aaacacaccc caccatcaac ctgagaccta 32820
 atttgttttcc cacatatata gaagtttttg aaaatgacaa cttcttggtt ttacatgttt 32880
 tcaattttct caaatctct gtttgtcgtc gtaagatgcc acttgttcaaa gggcagcttc 32940
 tgactgattt gcccacagcta tagatccacg gttaaatctc catcacctcc aaacccttt 33000
 tctgaaacaa aaagagaaaa acttcaactt gagaagaaaa catacaaaca aaaaatata 33060
 tttccatggg tgettaccatg acaaatgtaa tggtgtctaa aaggagagg ggacagatga 33120
 gtattttaat tactataggg cattattttc atagaaacaa tgggttattttt tcatttaat 33180
 attttgttcc tcacactgtat ccatcaatc aagtttctctt aaaaattatg aaattggcat 33240
 attttatattt ttgcttagata atgggtgaa caaggctgtat aatcaatttag ttctttttcc 33300
 ttttgatttta gtatataattt ttacatgtat attaaagcta ccattagctt atttattttt 33360
 caatgtgcct caattctaaag accactgca ttatggatca acaagaatgt gcaataccac 33420
 gaaccagcca ataaatttcaa ttcttataaa actaaagattt cagaaattcc aagagatttt 33480
 ttaacaaaata tcattgtacat atatggatca attaaacagag aaaaattttt aaaaatgtgaa 33540
 tttatcataat tcaattttaa actttcaaga ttatttcaat taataatattt ctatatttcc 33600
 aagaaatccct catcttgggt tattttctca cataagtata tttaattttt ttacatctt 33660
 gtggtttttc ttccatgtca cctatgtcca aagaaatttga taaccatattt ttgtttgagg 33720
 gattcagcat aacaatattt ttaagaaattt gtaatttctgg gtgataaaag tattttgttt 33780
 taaaacacaca agttaccattt aataatggta tcaccacat gatgaaatgtt aagatcttc 33840
 cctactctttt gggatacagc ctaacaaaagt gaacccttga attgtatttgc attctttatc 33900
 acgagtacaa aataatttcaaa atcagaaatgtt gaaccacaga taaggccagaa aagatgtaaa 33960
 caacaaaatg tattttcttcc caaaatatttga caatcttgcattt ttataatggat gaatcatgtc 34020
 aagtactttt tctatttccatca caatcttgcattt ttataatggat gatcatgtc 34080
 cacaggagttt aaagtaccctt ttccatgtt gtagtctca aatccccaaac aactctggta 34140
 attgggtgtttaa gaggggggtaa catcttgcattt caagggggctt ttatttttttccacatcc 34200
 tggatgtttaa cagatgctttt attaaggaca atgaaaaggaa aaatattttctt ctatatttcc 34260
 tttaaacaac caatcatgtc atgatgtatca attatgtat gatcatgtt acaaaaactt 34320
 ctatgtttaaa aataaaaataa ctcagatgtt aaaccattttt tttaaaaaaaa gtaattttttt 34380
 tccttagatca tagtgcattt gatgtatcaattt aaaaatgtt cttatgtt 34440
 ctttggtaagg aaagtatgtt catttgcattt cttttgcattt gagttaagag cttccctcgaa 34500
 ttttgagggaa ggatattttaa atcttcatttgcattt aacccaaatc aaaaatttttaa gatgtttttt 34560
 cttatgtttaaa atgtatgtt cttatgtt cttttgcattt gatgtttttt 34620
 ttacttatc agatataaa ttttatgtt cttatgtt cttttgcattt gatgtttttt 34680
 cttatgtttaaa atgtatgtt cttatgtt cttttgcattt gatgtttttt 34740

-continued

ttcttcataa ttaatggatg gagacaaccc aaatgtaaaa agtttagta aatactgatg 34800
caaaaagtgt agtaggctgg tattttgc ctaagcaaa aagtaattaa cataaagaga 34860
gatattcatt gatttttat agtaaaaatt tcaaatacgacc taaatatcca acactatgg 34920
aatggtttatacattatg gctcactagt gtcatgcagc attatggc tattaaatat 34980
tgtttatggc cttctgggtg cagccaagat ggaataagct gactataacc tcttcgctg 35040
attacaatga aaaattctag acagaatata aaaagcaact actcaaagac tcaaaagtga 35100
acagaaatag gttagattaag aattaaagta aaaacgtgaa taatgactca caacaggagt 35160
gagtttcata ggcctttct tttcttttt ttacttctt tgctcctga ttttgatctg 35220
agggtgggcc aaatcaaaac tgcctagcta gcgttagacag caaaactctg agagaaattt 35280
cttattttata gccagtggac taggataaa ggaaccctga atgcttagaga gtaagaaata 35340
ggtaggaaat ctcttgctg tatttggtg ttatgtttat ttcttatttt tttctggcc 35400
ctgcctgag ggtggctaac cccagttact gcctgggtac agctatggca acataggcac 35460
ctacaatgt gagagatgac ccgtctctct agcagtgga accaggaaaa gagctttctg 35520
tcatctgaag aatagaataa ggagattgtt tttgttatct cttaatctgc ttcgcttta 35580
cgggcaaaaca gaactgtgtg gtggagcaga ggctgagaga aaaatactgt ctggcttaggt 35640
gaccagaaaa agggccatt ggaggctggta taagagagga tcctggagc cagaaaatgt 35700
ggaggacatc tcagagagga aaggagttga aaatgtttaa ctccataatgc tacatgtgaa 35760
ctctaagtct caggattgcc cctgagctat gcacatgcag aacagagccaa aataaggcaca 35820
gcaaagtctt tgacacttaa ctactatata aaccactgac caaagtccct gactaccctg 35880
ggtggottat gtacctggaa gacctgaaca acaaaggctt tgaaaactaa actgacagga 35940
caatcaccaa ccacagaagg agaaagagaa catgttagat gaattgaaac cagctgatca 36000
tctgctgtaa caaaaacttc aacattctcc agattttAAC aggacccat gtctaaaac 36060
ataatatgca aatgttcag gataataatcc aaaattactc aacatacaaa gaaccaagaa 36120
aagctgacca tttctctagg gaaaagaaaa tcaacagcta caaaccctaa gatgaccaga 36180
tattaaact atcagacaaa gattgtgtg tgagaagtcc aagaagtcca accaaccctaa 36240
aacaggataa actcaaatgc ccagacacat cataatgtca tggcaaattt cccaccatata 36300
agtcttaaat gaagaaaaac agaatataaa actatacata ggatgaaatc aattgtgtct 36360
aaaatatgtat gatgaaatac ttatattaca tataaacaga aaaattataa aagaaaacca 36420
aactggtcaa cctatttgtc tctggatggg tatattatcg gtgattgttta ctttctcatt 36480
tttaatatttgc cagttttctt gcatctaata tgatgtttat tagagtcatg 36540
aagaccaaaa cattttctaa tcaaggatgaa aacatataat aactatgtttt ctcagaagaa 36600
tacttggcat aaggactaaa atcactgatt ttgttattgc tggttttaac tattgtcaca 36660
ggatttttctt cccatTTTC caaacacaag gactttggaa atagtttgtt aaatcacctt 36720
tcaaaaatgaa gaacaatgtt ctcactaattt actatattata gaattccata ctcaggtgg 36780
actcttatcc aatttggta tgcattccca gacattctca gggttccacta tctcctccat 36840
accatttgaa aggttcaaaa ctcacaggag aacacaaattt tttttaaagt gtccaattca 36900
aaaaataaac tgatttatgtt accagagcaat ttatTTTAA aatgaaatac tattattttt 36960
attattttt tgagataggg tcttgctcta ttgcccagga tggagggcag tggcaggatc 37020

-continued

agagctcact gtgacacctga cctccctggc tcaggcaatc tgcccatctc agcctccctga 37080
ataactggga ctacaggcgc acaccaccaat gcccagctaa tttttgtat ttggataaaa 37140
gacagggtt cactatgctg cccagtctgg tctcaaactc cttggctcaa gcgatccacc 37200
tgcccttggcc tcccaaagtg ctgggattac aggcataga cactgcaccc ggccctaaaa 37260
tgaaaactt taaaagacaa tgaaaatgaa tttattgagc ttcaacctta tgttcaaac 37320
tccatgctac ttcagttcc caaacatatg tagagtatct atgactgggt atcaaagatg 37380
gataaaattt caacatcaa tagttcaaaa tcttagtggag aaaagacctc taagcagata 37440
attacaatac accataggaa gttcagactt ctgttagtagg aggtatagga ttgtatacaca 37500
gaaataaggc ttttttggg atggggaaag gttgaggaca agaataaatt tagaaaataa 37560
catcagtgaa tgctccctgg tagaggcaat gggtaagttt attggatta accagtttgg 37620
taacaggatg gaatatacgataaaatagaac tgcaatattaa aatatatcga aagagcaagc 37680
agatggata tgtaagagaa gtcaagaattt ttgatgtcca tttaatataa agtacaaaat 37740
gtatagaattt agggccaaag agagtacata aacacttaac cctagaggat ttccatgcc 37800
atgttaatga gcttgacttt tatgttccag gccagggtc tccaattttt atggaatatg 37860
aatgcgtattt ctggaaatcctt aaccccagaa attctgaattt agcagttccag ggcgcaggcgt 37920
aggaatctat aatacaattt aaaaaacacgc cgggcgcgtt ggctcatgcc tgtaatcccc 37980
gcactttggg aggctgaggtt gggggatca caaggtcagg agatcgagac catcctcgct 38040
aacaagggga aaccccgctt ctacaaaaaa tacaaaaaaat tagctggcg tcgtgggg 38100
cgcctgttgtt cccagctact tggaggctg aggcaagaga atgggtgtgaa ctgggaggca 38160
gagcttgcag tgagccaaga tgggccact gcactccagc ctgggtgaca gagtttgacc 38220
ctgccccaaa aaaaaaaaaa aaaaaaagag ggaaatataat atataatataat atatacacac 38280
acacacacac acacacacac acacatataat acatataatataat acacatataat 38340
atatacatat atattacata tacatataataataataataataataataata 38400
tatataacc aagtaactct cgaacaggtt gtttcaat cacgggagaa acacgtcaaa 38460
caggctccctt aataagttt acacaataat gaagtaatca aattttacattt ttacaaactt 38520
caccctgaca ccatggacaa tcaatttaag gaagactaga agaagcaaac aaaatagaag 38580
gttattaaatg tagttgtgaa atgactcaac taaggttagtg gcatttagata taaaaattt 38640
caactaagag ataagatgtg taaagtgtgg aacttaggtga ctaatagcat ttagaaagag 38700
taagggagat agatgataca gcctaggatg atttcttgct acgattttta tataatggaga 38760
ttttaaaat atgtgcatta catcaagtttta tcagggggaa caagcaagta aacatttttt 38820
caatcgcggtt ctcttagtctg acagagtctg gtaaactttt cttatagtc tgcaaataca 38880
tttggagata aagtacttaa acatcactta cccttggcc ttcttcagct gtttctaaaa 38940
tctcacattt agaccatgtt tttcttagac ttgaccacac aacacattta gatccaacag 39000
taaaggctttt cagagtgtg gtgtctgtt gttgagctgc caaatggat tgagtaagac 39060
tcaatgcata tctcttaattt ccttaatgtt aacatcacaa gcaaaggattt ttccatataat 39120
gcattattttc tataatcatca atatcatgtt aattttttttt gttttttttt gttttttttt 39180
atttttttta aaattttttt acgtgcccccccccataataat gttttttttt gttttttttt 39240
accaacaaac caaacaaaaaa atcaccacaa ggtatgtaat ttgatttcttta acagtgaaaa 39300

-continued

tcgctaataa aatatggta tattaagaca agtgggttt ttataactta caacagccat 39360
ttctataaga aagcagcatg ttacaatga gcagacattc ttatcacaga gttaatacta 39420
ttattaaata tgacaatgag gaataaatat ctaatgcaag ggtataactt aaataactta 39480
tgtgttaccc ataggatggc ataccatgc attaaaatat tttgaaagca ttttaaggg 39540
actaaaatgc aggattcaat actgtatgat agtatgatct aaattgttt taaaactata 39600
tagtagacat acacaaagac agtaaaagac aacaaaaaga aaatacatca tggtgtaac 39660
attgatcaact tcctgttgc tttcattctc actattccat tatttcctt attttacaa 39720
tggttcaaata ttaataatat aaaaattgc attaaaaag tatagcattt attataaaaa 39780
gctaaaagtt acccatttaa caattcttt taagtagcaa acagtagtaa aaataattta 39840
taaaaaagtt aataaatgtt ctccaagatt tacacatgaa acagttttt attaataatag 39900
gctacctaaa tctatTTTT ttctttaagg gtttctaattg aatttcaaga tcaatgcatt 39960
atacagttaa gtctaaaaga ccctgttgc ataattttaa gtgttaattt atgcctgttag 40020
tagtgcata aatttggca gggataccca tgTTTtataa aacttacca cacatgccct 40080
cttgattata aatgtgatta cattttaaat gtgattacct ttggagact aataggattt 40140
aagcattcaa ctTTTaaaca caaaactaat atttcagtgt ttcttccagc acatatttt 40200
cactgatttc acaatataatt ttcccataa taaaatgcc cttcataaaa tccaatctga 40260
ttcactttcc aggttttaa tctcttacaa aaaatttattt tttagaaatc catataagaa 40320
ttatttcctt ttctttaaaa caaaaacaa ctctggaaag tcactatgtt aaacttaaaa 40380
gtaaaagcac acTTTggag gctgaggcag aaggactgct tgagcccagg agtttgagac 40440
caacctggat aacatgacaa accctgtctc tacaAAAAAc agaaaaataa gctgggtgt 40500
gtggcacaca cctgtgttcc cagctacttg ggaggctgag gtggggaaat cactagagcc 40560
tgggaggcag aggctgcagt gagcaagat cacaccactg cactccagcc taggtgacag 40620
agcaagaccc tgtctaaaaa aaaaaaaaaa aaaaaagttaa aaaggaaaaaa aaattcacat 40680
acctgagata tgatctggta aaccattatt gtgttgcattt ccattactgc ttcttcctc 40740
agagaacaaa ggcataatg cggaaatggc atccctgttt ttataactcta caatttcctc 40800
ttcacatata ttcatTCAT tctgtttt aggtacacaa tctgttccat gtagatgcaa 40860
tgacatccctg cgctggcatc caaaagattc tacacaactt gactttctca tgcctcagt 40920
acacatggat tcctgtgtt tctgactaac tcctaaagac aaaggatcca tcttgcattc 40980
taaaggcagc tgggtgttag gtgttctgat ttctggctgt ttctcagctt cacagctgag 41040
aggaagctgc agtgtaaata gctccatggc gcctttgtg acacactctt gggagagtgg 41100
cacattcggt ggcacccggc ttagctcccc tttttcttca tccaccacca gagaattctg 41160
taactcaata gattccagtt ctaagaattc ttttgattca tcatcaggag aaagcggcac 41220
ctcaagtgaat ttcagtgcata gtatctccattt tggatcattt gcatgtggta gtaatgtgtt 41280
aaatccgtatc atcagtttattt tttttttttt tttttttttt tttttttttt tttttttttt 41340
ttcacattcc agttccctcag tatcaatttt atctttgcag gggttttcaaa acccttcaca 41400
aatgttatca gtttttaggtt cacgaaagtt acttggtttt gttccataaa tttttttttt 41460
ttttttagttt atttcagcaaa gctcatctgt ctgttgcattt atatataattt tttttttttt 41520
agcttttattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 41580

-continued

```

tatctctgag tcaatatttt catagggaag agcacttttta ataccagtct tagaatattt 41640
ctccatTTTC tcattaatac ttttacCTTt gctttcaag tcaacaatttgc ctaaagggt 41700
atcacaaaca tcccttctga tttcttagtat tgTTtatttcc aacacatctt ctgttattat 41760
ttcatagaaa tagtcatttgc cctcttgaga acataatcct tctgaaatgt taaaccctga 41820
gaggeaacaat ggaaaggcTTt gcatgggaac cgacagaagt tcagaaggaa tggcccagag 41880
tgcttttggg tccacacagt cttaaatgtt tccaaagtcc acaaggctga cagatacaag 41940
ataatcttca caaatatttag tgataagtgc cctataataa tgtccatctt ctctgtatct 42000
tactatacaa ggatctccaa tataaggaca tgggatacaa ttTCTCCTGT ctgttacctg 42060
ttctccagca gtcgtactt ctacttctaa acactgaagt ttctccgtat cagcaaactg 42120
acacaaaaag tactcaggcTc catctatcac agtggcataa gctcttatca ttttttttc 42180
tggattatac cagttaaagaa atactgaagt gtcaatgtct gatTTGTTAA cagactttga 42240
actggactt ttaattactt gggtagaaag ttctacttga gatTTTTCAC tgagagcata 42300
cctgctaattc atatcatctg ctatgatccc atgttcatca gcaagaataa cttcccatct 42360
gtcttggaaat ttaacaaatt cacatcttgc tgcagcctcg ctggTCCGTT gggaaaagta 42420
atgcatttcatt ttcttagaat ttttattgtc aggaacctca aatccctgca aggagcaatg 42480
aatgcacaac cccggcaata ttgcattaaC aaggtcaagc ctaccttattt tgTTtagtATG 42540
aaccacagaa acattGCCat aatctataaa ctgcacagag agaaggcTcat tgggttGTT 42600
ctcccttgcata 42611

```

What is claimed is:

1. An isolated genomic polynucleotide, said polynucleotide obtainable from human genomic clone library RP11-446F17 having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polynucleotide encoding human lipoprotein-associated phospholipase A2 depicted in SEQ ID NO:1;
 - (b) a polynucleotide consisting of SEQ ID NO:2 which encodes human lipoprotein-associated phospholipase A2 depicted in SEQ ID NO:1;
 - (c) a polynucleotide which is a variant of SEQ ID NO:2;
 - (d) a polynucleotide which is an allelic variant of SEQ ID NO:2;
 - (e) a polynucleotide which encodes a variant of SEQ ID NO:1;
 - a polynucleotide which hybridizes to any one of the polynucleotides specified in (a)-(e)
 - a polynucleotide that is a reverse complement of the polynucleotides specified in (a)-(f) and containing at least 10 transcription factor binding sites selected from the group consisting of AP1-C, AP1_Q4, AP4-Q5, DELTAEF1-01, GATA1_04, GATA1-06, GATA2-02, GATA3_02, GATA-C, LMO2COM-02, LYF1-01, MYOD_Q6, MZF_01, NFAT_Q6, NKX25-01, S8-01, SOX5-01, TATA-C, and TCF11-01
- as well as nucleic acid constructs, expression vectors and host cells containing these polynucleotide sequences.
- 2. A nucleic acid construct comprising the polynucleotide of claim 1.**
 - 3. An expression vector comprising the polynucleotide of claim 1.**
 - 4. A recombinant host cell comprising the polynucleotide of claim 1.**
 - 5. A method for obtaining a polypeptide encoded by a polynucleotide obtainable from human genomic clone library RP11-446F17, said polypeptide consisting of human lipoprotein-associated phospholipase A2**
 - (a) culturing the recombinant host cell of claim 4 under conditions that provide for the expression of said polypeptide and
 - (b) recovering said expressed polypeptide.
 - 6. A method for preparing an antibody specific to a polypeptide consisting of human lipoprotein-associated phospholipase A2 comprising:**
 - (a) obtaining a polypeptide according to the method of claim 5;
 - (b) optionally conjugating said polypeptide to a carrier protein;
 - (c) immunizing a host animal with said polypeptide or polypeptide-carrier protein conjugate of step (b) with an adjuvant and
 - (d) obtaining antibody from said immunized host animal.
 - 7. An isolated polynucleotide which hybridizes to a non-coding region of SEQ ID NO:2, which non-coding region is**

selected from the group consisting of an intron, a splice junction, a 5'- non-coding region, a transcription factor binding region and a 3'- non-coding region.

8. An antisense oligonucleotide or mimetic to the polynucleotide of claim 7.

9. A method of diagnosing a pathological condition or susceptibility to a pathological condition such as asthma in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1 or fragment thereof and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

10. A composition comprising the polynucleotide of claim 1 and a carrier.

11 A composition comprising the polynucleotide of claim 7 and a carrier.

12. A composition comprising the antisense oligonucleotide of claim 8 and a carrier.

13. A method for modulating human lipoprotein phospholipase A levels in a subject in need thereof, comprising administering to said subject an amount of the composition of claim 10 effective to modulate said human lipoprotein phospholipase A levels in said subject.

14. A kit comprising the polynucleotide of claim 7.

15. The kit according to claim 14, in which the polynucleotide is labeled with a detectable substance.

16. A kit comprising the antisense oligonucleotide or mimetic of claim 8.

17. The kit according to claim 17, in which the antisense oligonucleotide is labeled with a detectable substance.

18. An isolated polynucleotide which hybridizes to a transcriptional regulatory region of SEQ ID NO:2

19. A nucleic acid construct comprising the polynucleotide sequence of claim 18 operably linked to a polynucleotide sequence encoding a heterologous polypeptide.

20. An expression vector comprising the polynucleotide sequence of claim 18.

21. A recombinant host cell comprising the polynucleotide sequence of claim 18.

22. A method for expressing a heterologous polypeptide sequence comprising

(a) culturing the recombinant host cell of claim 21 under conditions that provide for the expression of said polypeptide and

(b) recovering said expressed polypeptide.

* * * * *