(19) World Intellectual Property Organization International Bureau
(43) International Publication Date 24 May 2007 (24.05.2007)


PCT
(10) International Publication Number WO 2007/058896 A3
(51) International Patent Classification: C12Q $1 / 68$ (2006.01)
(21) International Application Number:

PCT/US2006/043656
(22) International Filing Date:

9 November 2006 (09.11.2006)
(25) Filing Language:

English
(26) Publication Language:

English
(30) Priority Data:

60/736,083 10 November 2005 (10.11.2005) US
(71) Applicant (for all designated States except US): GOVERNMENT OF THE UNITED STATES OF AMERICA, REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, 6011 Executive Boulevard, Suite 325, Rockville, MD 20852 (US).
(72) Inventors; and
(75) Inventors/Applicants (for US only): FIGG, William D. [US/US]; 10451 Breckinridge LN, Fairfax, VA 22030 (US). SPARREBOOM, Alexander [NL/US]; 187 Ridgefield Road, Memphis, TN 38111 (US). SISSUNG, Tristan M. [US/US]; 4708 Randolph Court, Annandale, VA 22003 (US). PIEKARZ, Richard L. [US/US]; 8708 First Ave., Silver Spring, MD 20910 (US). BATES, Susan E. [US/US]; 5404 Alta Vista Road, Bethesda, MD 20814 (US).
(74) Agents: PILLAI, Xavier et al.; LEYDIG, VOIT \& MAYER, LTD., Two Prudential Plaza, Suite 4900, 180 N. Stetson Avenue, Chicago, IL 60601-6780 (US).
(81) Designated States (unless otherwise indicated, for every kind of national protection available): $\mathrm{AE}, \mathrm{AG}, \mathrm{AL}, \mathrm{AM}$, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, $\mathrm{CO}, \mathrm{CR}, \mathrm{CU}, \mathrm{CZ}, \mathrm{DE}, \mathrm{DK}, \mathrm{DM}, \mathrm{DZ}, \mathrm{EC}, \mathrm{EE}, \mathrm{EG}, \mathrm{ES}, \mathrm{FI}$, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
(88) Date of publication of the international search report: 4 October 2007

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
(54) Title: MATERIALS AND METHODS FOR ABCB1 POLYMORPHIC VARIANT SCREENING, DIAGNOSIS, AND TREATMENT
(57) Abstract: The invention provides methods and materials for screening for polymorphic variants in ABCB and diagnosing altered susceptibilities for drug-induced heart rhythm irregularities based on the same. These methods allow better treatment regimens for using drugs that bind a protein encoded by the ABCB 1 and/or induce heart rhythm irregularities such as the anti-cancer drug FK228.

# MATERIALS AND METHODS FOR ABCB1 POLYMORPHIC VARIANT SCREENING, DIAGNOSIS, AND TREATMENT 

## BACKGROUND OF THE INVENTION

[0001] Drugs that have tremendous benefits in ameliorating human suffering unfortunately can also have undesirable, and potentially dangerous, side effects. For example, treatment with FK228 (romidepsin), an anti-cancer drug, has been associated with cardiac toxicities in preclinical models, including ST/T wave flattening and asymptomatic dysrhythmias, and with reversible ECG changes. Other drugs also have negative side effects on the heart. Complicating matters, the side effects a drug has can vary between individuals. There has been and continues to be a search for ways of identifying how a drug will affect a given individual, and once that identification is made, ways of treating that individual. Accordingly, there exists a need for materials and methods for identifying individuals' susceptibility for drug induced effects on the heart and associated means of treatment.

## BRIEF SUMMARY OF THE INVENTION

[0002] The invention provides methods and materials for screening for polymorphic variants in the ABCB 1 gene and diagnosing altered susceptibilities for drug-induced heart rhythm irregularities based on the same. In one aspect, a method of screening for an altered susceptibility for a drug-induced heart rhythm irregularity is provided. A sample from a subject is screened to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB 1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB 1 gene. A diagnosis for the altered susceptibility of the subject for the heart rhythm irregularity as induced by the drug is rendered based on the presence or absence of the polymorphic variant of the ABCB 1 gene. In one aspect, the polymorphism comprises a polymorphism identified as rs1128503, rs2032582, rs1045642, or a combination thereof. In one aspect, the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236, 2677, or 3435 of SEQ ID NO: 2; or a combination thereof. In another aspect, a
method of screening for a decreased susceptibility for a depsipeptide, e.g., FK228,induced QT interval prolongation is provided. A sample from a subject is screened to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB 1 gene, wherein the polymorphic variant is associated with a decreased susceptibility for QT interval prolongation induced by the depsipeptide, and wherein the polymorphic variant comprises a thymine at position 2677 of SEQ ID NO: 2, or a thymine at position 3435 of SEQ ID NO: 2, or a combination thereof. A diagnosis of a decreased susceptibility of the subject for QT interval prolongation as induced by FK228 is rendered based on the presence or absence of the polymorphic variant of the ABCB1 gene.
[0003] Kits compatible with the methods are also provided. In one aspect, a kit is provided that includes a nucleic acid and a drug that binds a protein encoded by ABCB1. The nucleic acid is for use in screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB1 gene, and wherein the nucleic acid specifically binds to ABCB 1 sequence comprising the at least one polymorphism or a sequence adjacent to ABCB 1 sequence comprising the at least one polymorphism. In one aspect, the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2; or a combination thereof. In another aspect, the drug is FK228.
[0004] Use of a drug such as FK228 to manufacture a medicament is also provided. In one aspect, there is a use of a drug that binds a protein encoded by the ABCB1 gene to manufacture a medicament to treat a subject that that has been screened for the presence or absence of at least one polymorphic variant in at least one polymorphism of the ABCB1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by the drug. In another aspect, the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2, or a combination thereof.

## BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S)

[0005] Figure 1 shows relationships between the area under the curve (AUC) of FK228 and the percentage decrease in platelet count at nadir (PLC) following FK228 treatment. Each symbol represents an individual patient. Data were fit to a sigmoidal maximum effect model (solid line) with $95 \%$ confidence intervals (dotted lines).
[0006] Figure 2 shows relationships between ABCB1 genotypes and the baseline corrected QTc interval following FK228 treatment. Fig. 2A shows ABCB1 2677G>T/A genotypes: 1) GG genotype; 2) GT genotype; 3) TT genotype; 4) GA genotype. Fig. 2B shows ABCB1 $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}-3435 \mathrm{C}>$ T genotypes: 1) homozygous variant TT-TT diplotype; 2) a homozygous variant TT genotype at either the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ or the $3435 \mathrm{C}>\mathrm{T}$ locus; 3) any other $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}-3435 \mathrm{C}>$ T diplotype that does not correspond to 1) or 2 ). Each symbol represents an individual patient, and horizontal lines represent median values.
[0007] Figure 3 shows clearance data related to plasma concentration versus time curves for FK228 as a function of $\mathrm{ABCB} 12677 \mathrm{G}>$ T/A genotype [1) GG genotype; 2) GT genotype; 3) TT genotype; 4) GA genotype] (Fig. 3A), CYP3A4*1B genotype [1), wildtype; 2), heterozygous or homozygous variant] (Fig. 3B), and (C) CYP3A5*3C genotype [1), wild-type or heterozygous; 2), homozygous variant] (Fig. 3C). Each symbol represents an individual patient, and horizontal lines represent median values.
[0008] Figure 4A shows the relationships between ABCB 1 genotypes and the baseline corrected QTc interval following FK228 treatment for $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ allele combination in group $1(\mathrm{P}=.011)$.
[0009] Figure 4B shows the relationships between ABCB1 genotypes and the baseline corrected QTc interval following FK228 treatment for ABCB1 2677G>T/A and 3435C>T allele combination in group $2(\mathrm{P}=.07)$.
[0010] Figure 5A shows the relationships between ABCB 1 genotypes and the baseline corrected QTc interval following FK228 treatment for (B) ABCB1 3435C $>$ T genotype in group $1(\mathrm{P}=.15)$.
[0011] Figure 5B shows the relationships between ABCB 1 genotypes and the baseline corrected QTc interval following FK228 treatment for ABCB1 3435C>T genotype in group $2(\mathrm{P}=.028)$.
[0012] Figure 6A shows the relationships between ABCB 1 genotypes and the baseline corrected QTc interval following FK228 treatment for ABCB1 2677G>A/T genotype in group $1(\mathrm{P}=.0046)$.
[0013] Figure 6B shows the relationships between ABCB1 genotypes and the baseline corrected QTc interval following FK228 treatment for ABCB1 $2677 \mathrm{G}>\mathrm{A} / \mathrm{T}$ genotype in group $2(\mathrm{P}=.015)$. Each symbol represents an individual patient, and horizontal lines represent median values.
[0014] Figure 7A shows the clearance of FK228 as a function of ABCB1 2677G $>$ T/A and $3435 \mathrm{C}>\mathrm{T}$ allele combination in group $1(\mathrm{P}=.51)$. Each symbol represents an individual patient, and horizontal lines represent median values.
[0015] Figure 7B shows the clearance of FK228 as a function of ABCB1 2677G $>$ T/A and $3435 \mathrm{C}>\mathrm{T}$ allele combination in group $2(\mathrm{P}=.46)$. Each symbol represents an individual patient, and horizontal lines represent median values.

## DETAILED DESCRIPTION OF THE INVENTION

[0016] A method of screening for an altered susceptibility for a drug-induced heart rhythm irregularity is provided. The method comprises screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB 1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB1 gene, and wherein the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2; or a combination thereof. These polymorphisms are also identified as rs1128503, rs2032582, and rs1045642, respectively. The method further comprises diagnosing the altered susceptibility of the subject for the heart rhythm irregularity as induced by the drug based on the presence or absence of the polymorphic variant of the ABCB 1 gene. Detecting such a variant does not require detecting the chromosomal DNA or the actual gene. Detection can be of any indicator of such a variant such as any one of, or a combination of, the genome, a genomic fragment, mRNA, a mRNA fragment, cDNA, a cDNA fragment, an encoded polypeptide, and a polypeptide fragment thereof. In an embodiment, the polymorphic variant is associated with an increase or decrease in the expression of ABCB1. In an embodiment, the polymorphic variant is associated with an
increase or decrease in an activity of a protein encoded by the ABCB 1 gene. That change in activity can be in form of an increased or decreased ability to transport a drug such as FK228. That change can be the result of an alteration of one or more amino acid residues. Such amino acid changes can alter the active site and/or the conformation of the ABCB1 gene product resulting in a more or less efficient drug effluxer. In some embodiments, the polymorphic variant is associated with both a change in expression and a change in an activity of ABCB 1 .
[0017] As used herein, a "gene" is a sequence of DNA present in a cell that directs the expression of a "gene product," most commonly by transcription to produce RNA and translation to produce protein. An "allele" is a particular form of a gene. The term allele is relevant when there are two or more forms of a particular gene. Genes and alleles are not limited to the open reading frame of the genomic sequence or the cDNA sequence corresponding to processed RNA. A gene and allele can also include sequences upstream and downstream of the genomic sequence such as promoters and enhancers. The term "gene product" or "polymorphic variant allele product" refer to a product resulting from transcription of a gene. Gene and polymorphic variant allele products include partial, precursor, mature transcription products such as pre-mRNA and mRNA, and translation products with or without further processing including, without limitation, lipidation, phosphorylation, glycosylation, other modifications known in the art, and combinations of such processing. RNA may be modified without limitation by complexing with proteins, polyadenylation, splicing, capping or export from the nucleus.
[0018] A "polymorphism" is a site in the genome that varies between two or more individuals or within an individual in the case of a heterozygote. The frequency of the variation can be defined above a specific value for inclusion of variations generally observed in a population as opposed to random mutations. Polymorphisms that can be screened according to the invention include variation both inside and outside the open reading frame. When outside the reading frame the polymorphism can occur within 200 , $500,1000,2000,3000,5000$, or more of either the $5^{\prime}$ or $3^{\prime}$ end of the open reading frame. When inside the reading frame, the polymorphism may occur within an exon or intron, or overlapping an exon/intron boundary. A polymorphism could also overlap the open reading frame and a sequence outside of that frame. Many polymorphisms have been
given a "rs" designation in the SNP database of NCBI's Entrez, some of these designations have been provided herein.
[0019] A "polymorphic variant" is a particular form or embodiment of a polymorphism. For example, if the polymorphism is a single nucleotide polymorphism, a particular variant could potentially be an "A" (adenosine), "G" (guanine), "T" (thymine), and " C " (cytosine). When the variant is a " T ", it is understood that a " U " can occur in those instances wherein the relevant nucleic acid molecule is RNA, and vice versa in respect to DNA. The convention "PositionNUC1>NUC2" is used to indicate a polymorphism contrasting one variant from another. For example, $242 \mathrm{~A}>\mathrm{C}$ would refer to a cytosine instead of an adenosine occurring at position 242 of a particular nucleic acid sequence. In some cases, the variation can be to two or more different bases, e.g., $242 \mathrm{~A}>\mathrm{C} / \mathrm{T}$. When $242 \mathrm{~A}>\mathrm{C}$ is used in respect to a mRNA/cDNA, it can also be used to represent the polymorphism as it occurs in the genomic DNA with the understanding that the position number will likely be different in the genome. Sequence and polymorphic location information for both coding domain sequence and genomic sequence is described herein for the genes relevant to the invention. "Polymorphic variant allele" refers to an allele comprising a particular polymeric variant or a particular set of polymorphic variants corresponding to a particular set of polymorphisms. Two alleles can both be considered the same polymorphic variant allele if they share the same variant or set of variants defined by the polymorphic variant allele even though they may differ in respect to other polymorphisms or variation outside the definition. For a mutation at the amino acid level, the convention "AA1PositionAA2" is used. For example, in the context of amino acid sequence, M726L, would indicate that the underlying, nucleotide level polymorphism(s) has resulted in a change from a methionine to a leucine at position 726 in the amino acid sequence.
[0020] A "genotype" can refer to a characterization of an individual's genome in respect to one or both alleles and/or one or more polymorphic variants within that allele. A subject can be characterized at the level that the subject contains a particular allele, or at the level of identifying both members of an allelic pair, the corresponding alleles on the set of two chromosomes. One can also be characterized at the level of having one or more polymorphic variants. The term "haplotype" refers to a cis arrangement of two or more polymorphic variants, on a particular chromosome such as in a particular gene. The
haplotype preserves the information of the phase of the polymorphic nucleotides-that is, which set of polymorphic variants were inherited from one parent, and which from the other. Wherein methods, materials, and experiments are described for the invention in respect to polymorphic variants, one will understand that can also be adapted for use with an analogous haplotype. A "diplotype" is a haplotype that includes two polymorphisms. [0021] A single nucleotide polymorphism (SNPs) refers to a variation at a single nucleotide location. In some cases the variations at the position could be any one of the four nucleotide bases, in others the variation is some subset of the four bases. For example, the variation could be between either purine base or either pyrimidine base. Simple-sequence length polymophisms (SSLPs) or short tandem repeat polymorphisms (STRPs) involve the repeat of a particular sequence of one or more nucleotides. A restriction fragment length polymorphism (RFLP) is a variation in the genetic sequence that results in the appearance or disappearance of an enzymatic cleavage site depending on which base(s) are present in a particular allele.
[0022] A diagnosis for a given susceptibility in accordance with this invention includes detection of homozygosity and/or heterozygosity for a given polymorphism(s). Heterozygosity and homozygosity are relevant wherein the cell, or extract thereof, tested has two chromosomal copies. In other contexts, such as in a sperm or egg, only a single chromosome is present so that the issue of homozygosity or heterozygosity does not directly present itself. In the some embodiments, such as those involving cancer, homozygosity or heterozygosity can be lost or at least obscured because of deletion or inactivation of one of the two gene copies.
[0023] In those embodiments where a sample is screened to detect the presence or absence of more than one polymorphic variant associated with a given condition, the combination of the polymorphic variants can be additive, synergistic, or even antagonists in regards to correlative strength-although not overly antagonistic if the susceptibility or drug effect probability is lost. When screening for multiple polymorphisms all can be heterozygous, all can be homozygous, or a combination with one or more polymorphism homozygous, and one or more polymorphism heterozygous, depending on the particular susceptibility relationship for a given set of polymorphic variants and a condition or drug response.
[0024] The polymorphic variants described herein can be associated with an altered susceptibility to one or more complications and/or therapeutic treatments. How a polymorphism is mechanistically associated with this susceptibility need not be known for the usefulness and operability of the invention. The polymorphism need not actually cause or contribute to etiology or severity of the condition. In some embodiments, the polymorphism can cause or contribute to the condition. In some embodiments, the polymorphism can serve as a marker for another polymorphism(s) responsible for causing or contributing to the condition. In such a situation, the polymorphism(s) screened for can be in linkage disequilibrium with the responsible polymorphism(s).
[0025] In those embodiments where the screened for polymorphic variant(s) is responsible in part or whole for the condition(s), the polymorphic variant(s) can result in a change in the steady state level of $m R N A$, for example, through a decrease in transcription and/or mRNA stability. Some polymorphic variants can alter the exon/intron boundary and/or effect how splicing occurs. When the polymorphic variant occurs within or overlaps with the protein-encoding sequence of the gene, the polymorphic variant may be silent resulting in no change at the amino acid level, result in a change of one or more amino acid residues, a deletion of one or more amino acids, addition of one or more amino acids, or some combination of such changes. For some polymorphic variants, the result is premature termination of translation. The effect may be neutral, beneficial, or detrimental, or both beneficial and detrimental, depending on the circumstances. Polymorphic variants occurring in noncoding regions can exert phenotypic effects indirectly via influence on replication, transcription, and/or translation. Polymorphic variants in DNA can affect the basal transcription or regulated transcription of a gene locus. Such polymorphic variants may be located in any part of the gene but are most likely to be located in the promoter region, the first intron, or in 5' or 3' flanking DNA, where enhancer or silencer elements may be located. A single polymorphism can affect more than one phenotypic trait. A single phenotypic trait may be affected by polymorphisms in different genes. Some polymorphisms predispose an individual to a distinct mutation that is causally related to a certain phenotype.
[0026] RNA polymorphic variants can affect a wide range of processes including RNA splicing, polyadenylation, capping, export from the nucleus, interaction with translation intiation, elongation or termination factors, or the ribosome, or interaction with
cellular factors including regulatory proteins, or factors that may affect mRNA half life. An effect of polymorphic variants on RNA function can ultimately be measurable as an effect on RNA levels-either basal levels or regulated levels or levels in some abnormal cell state. One method for assessing the effect of RNA polymorphic variants on RNA function is to measure the levels of RNA produced by different alleles in one or more conditions of cell or tissue growth. Such measuring can be done by conventional methods such as Northern blots or RNAase protection assays, which can employ kits available from Ambion, Inc., or by methods such as the Taqman assay, or by using arrays of oligonucleotides or arrays of cDNAs or other nucleic acids attached to solid surfaces, such as a multiplex chip. Systems for arraying cDNAs are available commercially from companies such as Nanogen and General Scanning. Complete systems for gene expression analysis are available from companies such as Molecular Dynamics. See also supplement to volume 21 of Nature Genetics entitled "The Chipping Forecast." Additional methods for analyzing the effect of polymorphic variants on RNA include secondary structure probing, and direct measurement of half life or turnover. Secondary structure can be determined by techniques such as enzymatic probing with use of enzymes such as $\mathrm{T} 1, \mathrm{~T} 2$, and S 1 nuclease, chemical probing or RNAase H probing using oligonucleotides. Some RNA structural assays can be performed in vitro or on cell extracts.
[0027] To determine if one or more polymorphic variants have an effect on protein levels and/or activity, a variety of techniques may be employed. The in vitro protein activity can be determined by transcription or translation in bacteria, yeast, baculovirus, COS cells (transient), CHO, or study directly in human cells. Further, one can perform pulse chase experiments for the determination of changes in protein stability such as half life measurements. One can manipulate the cell assay to address grouping the cells by genotypes or phenotypes. For example, identification of cells with different genotypes and phenotype can be performed using standardized laboratory molecular biological protocols. After identification and grouping, one skilled in the art could determine whether there exists a correlation between cellular genotype and cellular phenotype.
[0028] Correlation between one or more polymorphic variants can be performed for a population of individuals who have been screened for particular polymorphic variants. Correlation can be performed by standard statistical methods including, but not limited to,
a chi-squared test. Analyses of polymorphic variants, parametric linkage analysis, nonparametric linkage analysis, etc. and statistically significant correlations between polymorphic form(s) and phenotypic characteristics also can be used.
[0029] ATP-binding cassette, sub-family B (MDR/TAP), member 1 (ABCB1) is a member of the ATP-binding cassette (ABC) family of transporters that couple ATP hydrolysis to active transport of substrates out of the cell. ABCB 1 has been shown to serve a protective function in several tissues including heart, hematopoietic stem cells, and other tissues, where it effluxes endogenous and exogenous toxins. ABCB 1 has the further aliases HGNC:40, ABC20, CD243, CLCS, GP170, MDR1, P-gp, PGY1. ABCB1 has the further designations: P-glycoprotein 1; multidrug resistance 1 ; colchicin sensitivity; doxorubicin resistance; MDR-1 and multidrug resistance 1. ABCB1 has been assigned Gene ID 5243, and is positioned on chromosome 7 at locus 7 q 21.1 . Further information for ABCB 1 is found on the NCBI wesite in the Entrez Gene database and Online Mendelian Inheritance in Man (OMIM) website under entry "*171050."
[0030] ABCB1 nucleic acid and amino acid sequences relevant to the invention include genomic, cDNA, and fragments thereof. The particular sequences identified herein by sequence identification number and/or accession number are representative of ABCB 1 sequences. One of skill in the art can appreciate that there can be variability in the gene or gene fragment distinct from the polymorphism(s) of interest and that such allelic variants still fall within the scope of the invention. As the polymorphism will be reflected in both strands of the DNA, the screening in the context of the invention can involve one or both of the strand sequences. Accordingly, where the sequence for a given strand is provided, the invention also includes the use of its complement.
[0031] ABCB1 polymorphisms of particular interest include those known in the art as the 1236,2677 , and 3435 polymorphisms as well as the particular polymorphic variants $1236 \mathrm{C}>\mathrm{T}, 2677 \mathrm{G}>\mathrm{A} / \mathrm{T}$, and $3435 \mathrm{C}>\mathrm{T}$. Other variants of these polymorphisms are also provided as are other polymorphisms in the ABCB1 gene. Polymorphic variants of adenosine (A), guanine (G), cytosine (C), thymine (T), uracil (U) and other applicable nucleotides of each polymorphism are provided. Such is provided not just for ABCB1 polymorphisms, but also for polymorphisms of other genes described herein as well. Other polymorphic variants of these polymorphisms as well as other polymorphisms can also be screened for. The 1236,2677 , and 3435 polymorphisms are given the designations
rs1128503, rs2032582, and rs1045642 respectively in the SNP database of NCBI's Entrez. These polymorphisms and particular variants are exemplary and other ABCB1 polymorphisms and variants may also be screened for in accordance with the present invention. The following are representative genomic and cDNA sequences for $\mathrm{ABCB1}$. [0032] The ABCB1 genomic sequence is provided in SEQ ID NO: 1, derived from AY910577 from position 114998 to position 210947 inclusive. The 1236, 2677, and 3435 polymorphisms occur at positions 49,$910 ; 68,894$; and 90,871 of SEQ ID NO: 1 (corresponding to positions 164,$900 ; 183884$, and 205,861 respectively in AY910577). Screening with a genomic ABCB1 fragment of at least $5,10,20,25,30,35,40$, and 50 nucleic acids is within the scope of the invention, as well as, smaller, larger, and intermediate fragments. Fragments can comprise the relevant polymorphism(s) and provide a sequence unique in the human genome. Examples of fragments include the following. SEQ ID NO: 3 comprises the " 1236 polymorphism" at position 7. SEQ ID NO: 4 comprises the "2677 polymorphism" at position 7. SEQ ID NO: 5 comprises the "3435 polymorphism" at position 1. SEQ ID NO: 6 comprises the 1236 and 2677 polymorphisms at positions 1 and 18,895 respectively. SEQ ID NO: 7 comprises the 2677 and 3435 polymorphisms at positions 1 and 21,978 respectively. SEQ ID NO: 8 comprises the 1236,2677 , and 3435 polymorphisms at positions $1 ; 18,895$; and 40,962 respectively. Other relevant genomic sequence information includes AF016534, AY910577, CH236949, M29422, M29423, M29424, M29425, M29426, M29427, M29428, M29429, M29430, M29431, M29432, M29433, M29434, M29435, M29436, M29437, M29438, M29439, M29440, M29441, M29442, M29443, M29444, M29445, M29446, M29447, M37724, M37725, X58723, fragments thereof, and sequences comprising the same.
[0033] The ABCB1 cDNA sequence is provided in SEQ ID NO: 2, derived from NM_000927. The 1236, 2677, and 3435 polymorphisms occur at positions 1236, 2677, and 3435 of SEQ ID NO: 2. Screening with a cDNA ABCB1 fragment of at least 5, 10, $20,25,30,35,40$, and 50 nucleic acids is within the scope of the invention, as well as, smaller, larger, and intermediate fragments. Fragments can comprise the relevant polymorphism(s) and provide a sequence unique in the human genome. Examples of fragments include the following. SEQ ID NO: 9 comprises the 1236 polymorphism at position 7. SEQ ID NO: 10 comprises the 2677 polymorphism at position 7. SEQ ID NO:

11 comprises the 3435 polymorphism at position 507. SEQ ID NO: 12 comprises the 1236 and 2677 polymorphisms at positions 1 and 1,442 respectively. SEQ ID NO: 13 comprises the 2677 and 3435 polymorphisms at positions 1 and 759 respectively. SEQ ID NO: 14 comprises the 1236,2677 , and 3435 polymorphisms at positions $1,1,442$, and 2,200 respectively. Other relevant sequence information include mRNA sequences AB208970, AF016535, AY425005, AY425006, BQ720763, BQ882401, BX509020, CB164676, M14758, fragments thereof, and sequences comprising the same.
[0034] The translation of the ABCB1 cDNA coding region is provided in SEQ ID NO: 15. Position 893 of SEQ ID NO: 15 can be amino acids such as alanine, serine, or threonine corresponding to the polymorphic variants of the 2677 polymorphism. Position 893 can also be any other amino acid. Fragments of the ABCB 1 polypeptide sequence are also within the scope of the invention such as fragment recognized by ABCB 1 specific antibodies and fragments recognized by antibodies specific to particular variants as manifested in the polypeptide sequence. Other relevant ABCB 1 polypeptide sequence information includes AAB70218, AAW82430, EAL24173, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA59576, AAA88047, AAA88048, CAA41558, BAD92207, AAB69423, AAR91621, AAR91622, AAA59575, P08183, Q59GY9, Q6TBL4, fragments thereof, and sequences comprising the same.
[0035] In one aspect the polymorphic variant screened for is present in a single chromosomal copy of the gene, and wherein heterozygosity is associated with an altered susceptibility for the heart rhythm irregularity. In some embodiments, the heterozygosity for polymorphic variants of two or more polymorphisms is associated with an altered susceptibility for the heart rhythm irregularity. In another aspect, the polymorphic variant is present in both chromosomal copies of the gene, wherein homozygosity of the polymorphic variant is associated with an altered susceptibility for the heart rhythm irregularity if homozygosity of the polymorphic variant is detected. In some embodiments, homozygosity for polymorphic variants of two or more polymorphisms is associated with an altered susceptibility for the heart rhythm irregularity.
[0036] In one aspect, the method of screening is performed on a sample comprising a nucleic acid selected from the group consisting of (a) a nucleic acid encoding ABCB1, (b) a fragment of (a) comprising at least $5,10,15,20,25,30,35,40,45,50,75,100,150$, $200,250,500,1000$, or 10,000 contiguous nucleotides of (a) wherein the contiguous nucleotides comprise the polymorphism, (c) a complement of (a) or (b), and (d) a combination of two or more of (a), (b), and (c). In some embodiments, the nucleic acid encoding ABCB1 comprises SEQ ID NOS: 1,2 , or a combination thereof. The polymorphism can be a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1; or 1236, 2677, or 3435 of SEQ ID NO: 2; or a combination thereof.
[0037] The method can be performed by screening for one or more polymorphmic variants of a single polymorphism of ABCB1. In some embodiments, the polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1 ; or 1236 of SEQ ID NO: 2 , or a combination thereof. In such cases, the nucleic acid can comprise the sequence of SEQ ID NOS: 3,9 , or a combination thereof. In some embodiments, the polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1, or 2677 of SEQ ID NO: 2, or a combination thereof. In such cases, the nucleic acid can comprise the sequence of SEQ ID NOS: 4,10 , or a combination thereof. In some embodiments, the polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof. In such cases, the nucleic acid can comprise the sequence of SEQ ID NOS: 5,11 , or a combination thereof.
[0038] The method can be performed by screening for one or more polymorphmic variants of two or more polymorphisms of ABCB 1 . In some embodiments, the nucleic acid comprises first and second polymorphisms wherein the first polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1 ; or 1236 of SEQ ID NO: 2, or a combination thereof and the second polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1, or 2677 of SEQ ID NO: 2, or a combination thereof. In some such cases, the nucleic acid comprises the sequence of SEQ ID NO: 6,12 , or a combination thereof. In some embodiments, the nucleic acid comprises first and second polymorphisms wherein the first polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1, 2677, of SEQ ID NO: 2 , or a combination thereof the second polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID

NO: 2, or a combination thereof. In such cases, the nucleic acid can comprise the sequence of SEQ ID NOS: 7, 13, or a combination thereof.
[0039] In some embodiments, the nucleic acid comprises first, second and third polymorphisms wherein the first polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1; or 1236 of SEQ ID NO: 2, or a combination thereof, the second polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1 , or 2677 of SEQ ID NO: 2 , or a combination thereof, and the third polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof. In such cases, the nucleic acid can comprise the sequence of SEQ ID NOS: 8,14 , or a combination thereof.
[0040] The method can be performed by screening wherein the polymorphic variant screened for is a thymine at at least one polymorphism. In some embodiments, the polymorphism comprises a polymorphism at position 49,910 of SEQ ID NO: 1; or 1236 of SEQ ID NO: 2, or a combination thereof, and the subject is homozygous for thymine at that position. In some embodiments, the polymorphism comprises a polymorphism at position 68,894 of SEQ ID NO: 1, or 2677 of SEQ ID NO: 2 , or a combination thereof and the subject is homozygous for thymine at that position. In some embodiments, the polymorphism comprises first, second, and third polymorphisms wherein the first polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1,2677 , of SEQ ID NO: 2 , or a combination thereof the second polymorphism is 2677 , and the third polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof, and wherein the subject is homozygous for thymine at both positions.
[0041] Polymorphic variants to be screened for are principally located in or in close proximity to the ABCB 1 gene. Representative, polymorphic variants that can be tested for in addition to ABCB 1 variant(s), include those associated with the following described genes without limitation to polymorphic variant, polymorphism, allelic variant, or gene. In some embodiments, the screened for polymorphic variants are correlated with the same disease. In some embodiments, the screened for polymorphic variants are correlated with different diseases.
[0042] The invention provides screening for polymorphic variants in genes and sequence other than ABCB 1 sequences. In some embodiments, the additional variant is in
a sequence associated with another drug resistance related gene. In some embodiments, one or more variant in one or more organic anion transporting protein (OATP) family members and/or multidrug resistance associated protein ABCC 1 (MRP1) are screened for. In some embodiments, the additional polymorphic variant is in a cytochrome P 450 gene. The polymorphic variant can be associated with altered metabolism of the drug.
[0043] Cytochrome P450, Family 3, Subfamily A, Polypeptide 4 (CYP3A4) is a P450 enzyme for which FK228 is a substrate. CYP3A4 has the further alias HGNC:2637, CP33, CP34, CYP3A, CYP3A3, HLP, NF-25, P450C3, and P450PCN1. CYP3A4 has the further designations P450-III, steroid inducible; cytochrome P450, subfamily IIIA (niphedipine oxidase), polypeptide 3; cytochrome P450, subfamily IIIA (niphedipine oxidase), polypeptide 4 ; cytochrome P 450 , subfamily IIIA, polypeptide 4 ; glucocorticoidinducible P450; and nifedipine oxidase. CYP3A4 has been assigned Gene ID 1576, and is positioned on chromosome 7 at locus 7 q 21.1 . Further information for CYP3A4 is found on the NCBI website in the Entrez Gene database and Online Mendelian Inheritance in Man (OMIM) website under entry *124010. Polymorphic variants that can be screened for in addition to one or more of the ABCB 1 polymorphic variants relevant to the invention include the polymorphic variant CYP3A4*1B.
[0044] CYP3A4 nucleic acid and amino acid sequences relevant to the invention include genomic, cDNA, and fragments thereof. The particular sequences identified herein by sequence identification number and/or accession number are representative of CYP3A4 sequences. One of skill in the art can appreciate that there can be variability in the gene or gene fragment distinct from the polymorphism(s) of interest and that such allelic variants still fall within the scope of the invention. As the polymorphism will be reflected in both strands of the DNA, the screening in the context of the invention can involve one or both of the strand sequences. Accordingly, where the sequence for a given strand is provided, the invention also includes the use of its complement. Screening with a CYP3A4 nucleic acid fragment of at least $5,10,20,25,30,35,40$, and 50 nucleic acids is within the scope of the invention, as well as, smaller, larger, and intermediate fragments. Fragments can comprise the relevant polymorphism(s) and provide a sequence unique in the human genome. Examples of relevant cytochromes include CYP3A4 and CYP3A5. In some embodiments, the allelic variant CYP3A4*1B is screened for. In some embodiments, the alleleic variant CYP3A5*3C is screened for. Examples of CYP3A4
genomic sequences include AF209389, AF280107, AF307089, CH236956, D11131, fragments thereof, and sequences comprising the same. Examples of CYP3A4 mRNA sequences include AF182273, AJ563375, AJ563376, AJ563377, BC069418, D00003, J04449, M13785, M14096, M18907, X12387, fragments thereof, and sequences comprising the same. Examples of CYP3A4 amino acid sequences include AAF21034, AAG32290, AAG53948, EAL23866, AAF13598, CAD91343, CAD91645, CAD91345, AAH69418, BAA00001, AAA35747, AAA35742, AAA35744, AAA35745, CAA30944, P05184, P08684, Q6GRK0, Q7Z448, Q86SK2, Q86SK3, Q9BZM0, fragments thereof, and sequences comprising the same.
[0045] The following are representative sequences for CYP3A4. CYP3A4 has a 5' genomic flanking sequence (SEQ ID NO: 16 as derived from D11131) and a genomic sequence beginning with exon 1 (SEQ ID NO: 17 as derived from positions 148,895 to 176,090 of NG_000004). CYP3A4*1B is the allelic variant of CYP3A4 of particular relevance to the present invention. This allelic variant is found in the $5^{\prime}$ genomic flanking sequence at position 810 of SEQ ID NO: 16, and is the result of an $A>G$ variance from the consensus sequence to the varient. Other nucleotides can also be at this position. The polymorphism at this position has been designated rs2740574. SEQ ID NO: 18 provides the cDNA sequence for CYP3A4. This sequence is derived from the complete CYP3A4 cDNA sequence, coding strand which has the Accession \#M18907. The CYP3A4*1B polymorphism is not found in this sequence as it is prior to the transcription start site and is not found expressed in the mRNA. SEQ ID NO: 19 provides the polypeptide sequence for CYP3A4. This sequence is derived from the complete CYP3A4 protein sequence, which has the Accession \#NP_059488.
[0046] Cytochrome P450, Family 3, Subfamily A, Polypeptide 5 (CYP3A5) is a P450 enzyme for which FK228 is a substrate. CYP3A5 has the further aliases HGNC:2638, CP35, P450PCN3, and PCN3. CYP3A5 has the further designations aryl hydrocarbon hydroxylase; cytochrome P-450; cytochrome P450, subfamily IIIA (niphedipine oxidase), polypeptide 5; flavoprotein-linked monooxygenase; microsomal monooxygenase; niphedipine oxidase; and xenobiotic monooxygenase. CYP3A5 has been assigned Gene ID 1577, and is positioned on chromosome 7 at locus 7 q 21.1 . Further information for CYP3A5 is found on the NCBI website in the Entrez Gene database and Online Mendelian Inheritance in Man (OMIM) website under entry *605325. Polymorphic
variants that can be screened for in addition to one or more of the ABCB 1 polymorphic variants relevant to the invention include the polymorphic variant CYP3A5*3C.
[0047] CYP3A5 nucleic acid and amino acid sequences relevant to the invention include genomic, cDNA, and fragments thereof. The particular sequences identified herein by sequence identification number and/or accession number are representative of CYP3A5 sequences. One of skill in the art can appreciate that there can be variability in the gene or gene fragment distinct from the polymorphism(s) of interest and that such allelic variants still fall within the scope of the invention. As the polymorphism will be reflected in both strands of the DNA, the screening in the context of the invention can involve one or both of the strand sequences. Accordingly, where the sequence for a given strand is provided, the invention also includes the use of its complement. Screening with a CYP3A5 nucleic acid fragment of at least $5,10,20,25,30,35,40$, and 50 nucleic acids is within the scope of the invention, as well as, smaller, larger, and intermediate fragments. Fragments can comprise the relevant polymorphism(s) and provide a sequence unique in the human genome. Examples of CYP3A5 genomic sequences include AC005020, AF280107, AF355803, CH236956, L35912, fragments thereof, and sequences comprising the same. Examples of CYP3A5 mRNA sequences include AF355801, AJ563378, AJ563379, AK223008, BC022298, BC025176, BC026255, BC033862, BX537676, J04813, L26985, fragments thereof, and sequences comprising the same. Examples of CYP3A5 amino acid sequences include AAS02016, AAG32288, AAK73691, EAL23868, AAB00083, AAK73689, CAD91347, CAD91647, CAD91649, BAD96728, AAH33862, CAD97807, AAA02993, P20815, Q53GC3, Q75MV0, Q7Z3N0, Q7Z446, Q7Z447, Q86SK1, Q96RK6, fragments thereof, and sequences comprising the same.
[0048] The following are representative sequences for CYP3A5. The genomic DNA for CYP3A5 is shown in SEQ ID NO: 20 (corresponding to positions 253,080-288,849. The cDNA for CYP3A5 is provided in SEQ ID NO: 21 as derived from BC033862. CYP3A5*1B is the allelic variant of CYP3A5 of particular relevance to the present invention. The cDNA sequence for CYP3A5*1B is provided in SEQ ID NO: 22. The CYP3A5*3C allelic variant is a result of an $A>G$ variance at position 7087 of SEQ ID NO: 20 (260167 of NG_000004). Other nucleotides can also be at this position. The polymorphism at this position has been designated rs776746. The CYP3A5*3C polymorphism is contained in an intron and is not found expressed in the consensus
mRNA sequence. However, the CYP3A5*3C polymorphic variant results in the inclusion of intron 3 in the spliced mRNA as it is contained within a cryptic splice site. The mRNA and cDNA corresponding to the CYP3A5*3C polymorphism therefore includes intron 3 (bases 258551-260403 in the CYP3A5 genomic DNA sequence; Accession \#NG_000004) between bases 307 and 308 in SEQ ID NO: 21. The CYP3A5*3C polymorphism in the cDNA sequence, SEQ ID NO: 22, occurs at position 1923. [0049] Amino acid sequences for CYP3A5 and CYP3A5*1B are provided in SEQ ID NOS: 23 and 24 respectively. The following sequence contains a total of 502 amino acids. This sequence is derived from the complete CYP3A5 protein sequence, which has the Accession \# NP_000768. The protein is not expressed in individuals homozygous for the CYP3A5*3C polymorphism as the incorporation of intronic DNA results in premature truncation of the protein after amino acid 102 due to the presence of a stop codon within intron 3.
[0050] The invention also includes use of other polymorphic variants of the genes and proteins described herein. Use of both the nucleic acids described herein and their complements are within the scope of the invention. In connection with the provision and description of nucleic acid sequences, the references herein to gene names and to GenBank and OMIM reference numbers provide the relevant sequences, recognizing that the described sequences will, in most cases, also have other corresponding allelic variants. Although the referenced sequences may contain sequencing error, such error does not interfere with identification of a relevant gene or portion of a gene, and can be readily corrected by redundant sequencing of the relevant sequence (preferably using both strands of DNA). Nucleic acid molecules or sequences can be readily obtained or determined utilizing the reference sequences. Molecules such as nucleic acid hybridization probes and amplification primers can be provided and are described by the selected portion of the reference sequence with correction if appropriate. In some embodiments, probes comprise $5,6,10,12,13,14,15,16,17,18,19,20,23,25,27,30,35,40,45,50$, or more nucleotides.
[0051] The terms "disease" or "condition" are commonly recognized in the art and designate the presence of signs and/or symptoms in an individual or patient that are generally recognized as abnormal. Unless indicated as otherwise, the terms "disease," "disease state," condition," "disorder," and "complication" can be used interchangeably.

Diseases or conditions can be diagnosed and categorized based on pathological changes. Signs can include any objective evidence of a disease such as changes that are evident by physical examination of a patient or the results of diagnostic tests which may include, among others, laboratory tests to determine the presence of polymorphic variants or variant forms of certain genes in a patient. Symptoms can include a patient's perception of an abnormal condition that differs from normal function, sensation, or appearance, which may include, for example, physical disabilities, morbidity, pain, and other changes from the normal condition experienced by an individual. Various diseases or conditions include, but are not limited to, those categorized in medical texts.
[0052] Unless otherwise indicated, the term "suffering from a disease or condition" can refer to a person that currently has signs and symptoms, or is more likely to develop such signs and symptoms than a normal person in the population. For example, a person suffering from a condition can include a developing fetus, a person subject to a treatment or environmental condition that enhances the likelihood of developing the signs or symptoms of a condition, or a person who is being given or will be given a treatment that increases the likelihood of the person developing a particular condition. Methods of the invention relating to treatments of patients can include primary treatments directed to a presently active disease or condition, secondary treatments that are intended to cause a biological effect relevant to a primary treatment, and prophylactic treatments intended to delay, reduce, or prevent the development of a disease or condition, as well as treatments intended to cause the development of a condition different from that which would have been likely to develop in the absence of the treatment.
[0053] Combined detection of several polymorphic variants typically increases the probability of an accurate diagnosis. Analysis of the polymorphisms of the invention can be combined with that of other polymorphisms or other risk factors such as family history. Polymorphisms can be used to diagnose a disease at the pre-symptomatic stage, as a method of post-symptomatic diagnosis, as a method of confirmation of diagnosis or as a post-mortem diagnosis. Ethical issues to be considered in screening and diagnosis are discussed generally in Reich, et al., Genet. Med., 5:133-143 (2003).
[0054] In some embodiments, the sample screened is from a subject who has previously experienced a heart rhythm irregularity. In some embodiment, the heart rhythm irregularity is a cardiac arrhythmia. The heart rhythm irregularity comprises at
least one member selected from the group consisting of asymptomatic dysrhythmias and ventricular arrthymias. The heart rhythm irregularity can be characterized by at least one of ST/T wave flattening, torsade de pointes, and QT interval prolongation.
[0055] "Prolonged QT interval," "QT interval prolongation" or "QT interval elongation" refers to the QT interval measured from QRS onset to T wave offset (QTo) and from QRS onset to $T$ wave peak (QTm) adjusted to a heart rate of 60 beats per minute, which is QTc. "QTc" is also referred to as the Bazett corrected QT interval. See, e.g., Kligfield et al., J. Am. Coll. Cardiol, 28: 1547-55 (1996). Prolonged QT intervals can be induced directly or indirectly by one or more polymorphic variant of one or more polymorphism.
[0056] "Torsades de Pointes" or "TdP" is an uncommon variant of ventricular tachycardia (VT). The underlying etiology and management of TdP can be different from the more common ventricular tachycardia. TdP is a polymorphous ventricular tachycardia in which the morphology of the QRS complexes vary from beat to beat. The ventricular rate can range from about $150 / \mathrm{min}$ to about $250 / \mathrm{min}$. In some cases, there is a constantly changing wave form, but there may not be regularity to the axis changes. Q-T interval can be markedly increased (usually to 600 msec or greater). Cases of polymorphic VT, which are not associated with a prolonged Q-T interval, can be treated as generic VT. TdP can occur in bursts that are not sustained. Accordingly, one can employ a rhythm strip showing the patient's base-line Q-T prolongation
[0057] Any applicable method or combination of methods can be used to screen for polymorphic variants in a sample. Screening methods can utilize one or more of a nucleic acid array, allele-specific-oligonucleotide (ASO) hybridization, PCR-RFLP analysis, PCR., a single-strand conformation polymorphic variant (SSCP) technique, an amplification refractory mutation system (ARMS) technique, nucleotide sequencing, an antibody specific to a polypeptide encoded by the polymorphic variant containing gene, mass spectrometry, and combinations thereof. The sample screened can comprise at least one of genomic DNA, cDNA, mRNA, other DNA, other RNA, a fragment thereof, and a combination thereof. The sample screened can be derived from any number of single or combined sample and/or cell or tissue sources. In some embodiments, the screened sample comprises blood. The sample need not be directly from a subject. One or more steps can be performed on the sample prior to, subsequent to, and/or as part of the
screening. For example, one or more of the following: mRNA from a subject can be converted to cDNA, cDNA can be amplified using PCR, amplified DNA can be sequenced and/or assayed with one or more restriction enzymes, etc.
[0058] The molecules and probes relevant to the invention can be used in screening techniques. A variety of screening techniques are known in the art for detecting the presence of one or more copies of one or more polymorphic variants in a sample or from a subject. Many of these assays have been reviewed by Landegren et al., Genome Res., 8:769-776, 1998. Determination of polymorphic variants within a particular nucleotide sequence among a population can be determined by any method known in the art, for example and without limitation, direct sequencing, restriction length fragment polymorphism (RFLP), single-strand conformational analysis (SSCA), denaturing gradient gel electrophoresis (DGGE) [see, e.g., Van Orsouw et al., Genet Anal., 14(5-6):205-13 (1999)], heteroduplex analysis (HET) [see, e.g., Ganguly A, et al., Proc Natl Acad Sci USA. 90 (21):10325-9 (1993)], chemical cleavage analysis (CCM) [see, e.g., Ellis T P, et al., Human Mutation 11(5):345-53 (1998)] (either enzymatic as with T4 Endonuclease 7, or chemical as with osmium tetroxide and hydroxylamine) and ribonuclease cleavage. Screening for polymorphic variants can be performed when a polymorphic variant is already known to be associated with a particular disease or condition. In some embodiments, the screening is performed in pursuit of identifying one or more polymorphic variants and determining whether they are associated with a particular disease or condition.
[0059] In respect to DNA, polymorphic variant screening can include genomic DNA screening and/or cDNA screening. Genomic polymorphic variant detection can include screening the entire genomic segment spanning the gene from the transcription start site to the polyadenylation site. In some embodiments, genomic polymorphic variant detection can include the exons and some region around them containing the splicing signals, for example, but not all of the intronic sequences. In addition to screening introns and exons for polymorphic variants, regulatory DNA sequences can be screened for polymorphic variants. Promoter, enhancer, silencer and other regulatory elements have been described in human genes. The promoter is generally proximal to the transcription start site, although there may be several promoters and several transcription start sites. Enhancer, silencer and other regulatory elements can be intragenic or can lie outside the introns and
exons, possibly at a considerable distance, such as 100 kb away. Polymorphic variants in such sequences can affect basal gene expression or regulation of gene expression.
[0060] The presence or absence of the at least one polymorphic variant can be determined by nucleotide sequencing. Sequencing can be carried out by any suitable method, for example, dideoxy sequencing [Sanger et al., Proc. Natl. Acad. Sci. USA, 74:5463-5467 (1977)], chemical sequencing [Maxam and Gilbert, Proc. Natl. Acad. Sci. USA, 74:560-564, (1977)] or variations thereof. Methods for sequencing can also be found in Ausubel et al., eds., Short Protocols in Molecular Biology, .3rd ed., Wiley, 1995 and Sambrook et al., Molecular Cloning, 2nd ed., Chap. 13, Cold Spring Harbor Laboratory Press, 1989. The sequencing can involve sequencing of a portion or portions of a gene and/or portions of a plurality of genes that includes at least one polymorphic variant site, and can include a plurality of such sites. The portion can be of sufficient length to discern whether the polymorphic variant(s) of interest is present. In some embodiments the portion is $500,250,100,75,65,50,45,35,25$ nucleotides or less in length. Sequencing can also include the use of dye-labeled dideoxy nucleotides, and the use of mass spectrometric methods. Mass spectrometric methods can also be used to determine the nucleotide present at a polymorphic variant site.
[0061] RFLP analysis is useful for detecting the presence of genetic variants at a locus in a population when the variants differ in the size of a probed restriction fragment within the locus, such that the difference between the variants can be visualized by electrophoresis [see, e.g. U.S. Pat. Nos. 5,324,631 and 5,645,995]. Such differences will occur when a variant creates or eliminates a restriction site within the probed fragment. RFLP analysis is also useful for detecting a large insertion or deletion within the probed fragment. RFLP analysis is useful for detecting, for example, an Alu or other sequence insertion or deletion.
[0062] Single-strand conformational polymorphisms (SSCPs) can be detected in $<220$ bp PCR amplicons with high sensitivity. SSCP is usually paired with a DNA sequencing method, because the SSCP method does not provide the nucleotide identity of polymorphic variants. The SSCP technique can be used on genomic DNA as well as PCR amplified DNA as well. [Orita et al, Proc. Natl. Acad. Sci. USA, 86:2766-2770, 1989; Warren et al., In: Current Protocols in Human Genetics, Dracopoli et al., eds, Wiley, 1994, 7.4.1-7.4.6.]
[0063] Another method for detecting polymorphic variants is the T4 endonuclease VII (T4E7) mismatch cleavage method: T4E7 specifically cleaves heteroduplex DNA containing single base mismatches, deletions or insertions. Denaturing gradient gel electrophoresis (DGGE) can detect single base mutations based on differences in migration between homoduplexes and heteroduplexes [Myers et al., Nature, 313:495-498 (1985)]. In heteroduplex analysis (HET) [Keen et al., Trends Genet. 7:5 (1991)], genomic DNA is amplified by the polymerase chain reaction followed by an additional denaturing step that increases the chance of heteroduplex formation in heterozygous individuals. The PCR products are then separated on Hydrolink gels where the presence of the heteroduplex is observed as an additional band. Chemical cleavage analysis (CCM) is based on the chemical reactivity of thymine ( T ) when mismatched with cytosine, guanine or thymine and the chemical reactivity of cytosine(C) when mismatched with thymine, adenine or cytosine [Cotton et al., Proc. Natl. Acad. Sci. USA, 85:4397-4401 (1988)]. Ribonuclease cleavage involves enzymatic cleavage of RNA at a single base mismatch in an RNA:DNA hybrid (Myers et al., Science 230:1242-1246, 1985).
[0064] In addition to the physical methods described herein and others known to those skilled in the art, see, for example, Housman, U.S. Pat. No. 5,702,890; Housman et al., U.S. Patent Application No. 09/045,053, polymorphisms can be detected using computational methods, involving computer comparison of sequences from two or more different biological sources, which can be obtained in various ways, for example from public sequence databases. The term "polymorphic variant scanning" refers to a process of identifying sequence polymorphic variants using computer-based comparison and analysis of multiple representations of at least a portion of one or more genes. Computational polymorphic variant detection involves a process to distinguish true polymorphic variants from sequencing errors or other artifacts, and thus does not require perfectly accurate sequences. Such scanning can be performed in a variety of ways as known to those skilled in the art, preferably, for example, as described in U.S. Patent Application No. 09/300,747. The "gene" and "SNP" databases of Pubmed Entrez can also be utilized for identifying polymorphisms.
[0065] Genomic and cDNA sequences can both or in the alternative be used in identifying polymorphisms. Genomic sequences are useful where the detection of polymorphism in or near splice sites is sought, such polymorphism can be in introns,
exons, or overlapping intron/exon boundaries. Nucleic acid sequences analyzed may represent full or partial genomic DNA sequences for a gene or genes. Partial cDNA sequences can also be utilized although this is less preferred. As described herein, the polymorphic variant scanning analysis can utilize sequence overlap regions, even from partial sequences. While the present description is provided by reference to DNA, for example, cDNA, some sequences can be provided as RNA sequences, for example, mRNA sequences.
[0066] Interpreting the location of the polymorphic variant in the gene can depend on the correct assignment of the initial ATG of the encoded protein (the translation start site). The correct ATG can be incorrect in GenBank, but that one skilled in the art will know how to carry out experiments to definitively identify the correct translation initiation codon (which is not always an ATG). In the event of any potential question concerning the proper identification of a gene or part of a gene, due for example, to an error in recording an identifier or the absence of one or more of the identifiers, the priority for use to resolve the ambiguity is GenBank accession number, OMIM identification number, HUGO identifier, common name identifier.
[0067] Allele and genotype frequencies can be compared between cases and controls using statistical software (for example, SAS PROC NLMIXED). The odds ratios can be calculated using a log linear model by the delta method [Agresti, New York: John Wiley \& Sons (1990)] and statistical significance is assessed via the chi-square test. Likelihood ratios (G2) were used to assess goodness of fit of different models i.e., G2 provides a measure of the reliability of the odds ratio; small G2 P-values indicate a poor fit to the model being tested. Combined genotypes can be analyzed by estimating, maximum likelihood estimation, the gamete frequencies in cases and controls using a model of the four combinations of alleles as described by Weir, Sunderland, MA: Sinauer (1996). Gene-gene interactive effects can be tested using a series of non-hierarchical logistic models [Piegorsch et al., Stat. Med. 13:153-162 (1994)] to estimate interactive dominant and recessive effects. A sample size as large as possible from a relatively homogenous population to minimize variables outside the focus of the study.
[0068] Genomic DNA can be extracted from cases and controls using the QIAamp DNA Blood Mini Kit from Qiagen, UK. Genotyping of polymorphisms can be performed using PCR-RFLP (Restriction Fragment Length Polymorphism) using appropriate
restriction sites for the gene(s) being studied [Frosst et al., Nature Genet., 10:111-113 (1995); Hol et al., Clin. Genet., 53:119-125 (1998); Brody et al., Am. J. Hum. Genet., 71:1207-1215 (2002)]. A polymorphism may be genotyped using an allele-specific primer extension assay and scored by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (Sequenom, San Diego). Appropriate controls should be included in all assays. genotyping consistency can be tested by analyzing between 10 $15 \%$ of samples in duplicate.
[0069] One type of assay has been termed an array hybridization assay, an example of which is the multiplexed allele-specific diagnostic assay (MASDA) (U.S. Pat. No. 5,834,181; Shuber et al., Hum. Molec. Genet., 6:337-347 (1997). In MASDA, samples from multiplex PCR are immobilized on a solid support. A single hybridization is conducted with a pool of labeled allele specific oligonucleotides (ASO). The support is then washed to remove unhybridized ASOs remaining in the pool. Labeled ASO remaining on the support are detected and eluted from the support. The eluted ASOs are then sequenced to determine the mutation present.
[0070] Two assays depend on hybridization-based allele-discrimination during PCR. The TaqMan assay (U.S. Pat. No. 5,962,233; Livak et al., Nature Genet., 9:341-342, 1995) uses allele specific (ASO) probes with a donor dye on one end and an acceptor dye on the other end such that the dye pair interact via fluorescence resonance energy transfer (FRET).
[0071] An alternative to the TaqMan assay is the molecular beacons assay [U.S. Pat. No. 5,925,517; Tyagi et al., Nature Biotech., 16:49-53 (1998)]. High throughput screening for SNPs that affect restriction sites can be achieved by Microtiter Array Diagonal Gel Electrophoresis (MADGE)(Day and Humphries, Anal. Biochem., 222:389395, 1994).
[0072] Additional assays depend on mismatch distinction by polymerases and ligases. The polymerization step in PCR places high stringency requirements on correct base pairing of the 3 ' end of the hybridizing primers. This has allowed the use of PCR for the rapid detection of single base changes in DNA by using specifically designed oligonucleotides in a method variously called PCR amplification of specific alleles (PASA) [Sommer et al., Mayo Clin. Proc., 64:1361-1372 (1989); Sarker et al., Anal. Biochem. (1990), allele-specific amplification (ASA), allele-specific PCR, and
amplification refractory mutation system (ARMS) [Newton et al., Nuc. Acids Res. (1989); Nichols et al., Genomics (1989); Wu et al., Proc. Natl. Acad. Sci. USA, (1989)]. In these methods, an oligonucleotide primer is designed that perfectly matches one allele but mismatches the other allele at or near the $3^{\prime}$ end. This results in the preferential amplification of one allele over the other. By using three primers that produce two differently sized products, it can be determine whether an individual is homozygous or heterozygous for the mutation [Dutton and Sommer, Bio Techniques, 11:700-702 (1991)]. In another method, termed bi-PASA, four primers are used; two outer primers that bind at different distances from the site of the SNP and two allele specific inner primers [Liu et al., Genome Res., 7:389-398 (1997)].
[0073] Another technique is the oligonucleotide ligation assay [Landegren et al., Science, 241:1077-1080 (1988)] and the ligase chain reaction [LCR; Barany, Proc. Natl. Acad. Sci. USA, 88:189-193 (1991)]. In OLA, the sequence surrounding the SNP is first amplified by PCR, whereas in LCR, genomic DNA can by used as a template. In one method for mass screening based on the OLA, amplified DNA templates are analyzed for their ability to serve as templates for ligation reactions between labeled oligonucleotide probes [Samotiaki et al., Genomics, 20:238-242, (1994)]. In alternative gel-based OLA assays, polymorphic variants can be detected simultaneously using multiplex PCR and multiplex ligation [U.S. Pat. No. 5,830,711; Day et al., Genomics, 29:152-162 (1995); Grossman et al., Nuc. Acids Res., 22:4527-4534, (1994)]. A further modification of the ligation assay has been termed the dye-labeled oligonucleotide ligation (DOL) assay [U.S. Pat. No. 5,945,283; Chen et al., Genome Res., 8:549-556 (1998)].
[0074] In another method for the detection of polymorphic variants termed minisequencing, the target-dependent addition by a polymerase of a specific nucleotide immediately downstream ( $3^{\prime}$ ) to a single primer is used to determine which allele is present (U.S. Pat. No. 5,846,710). Using this method, several variants can be analyzed in parallel by separating locus specific primers on the basis of size via electrophoresis and determining allele specific incorporation using labeled nucleotides. Determination of individual variants using solid phase minisequencing has been described by Syvanen et al., Am. J. Hum. Genet., 52:46-59 (1993). Minisequencing has also been adapted for use with microarrays [Shumaker et al., Human Mut., 7:346-354 (1996)]. In a variation of this method suitable for use with multiplex PCR, extension is accomplished with the use of the
appropriate labeled ddNTP and unlabeled ddNTPs [Pastinen et al., Genome Res., 7:606614 (1997)]. Solid phase minisequencing has also been used to detect multiple polymorphic nucleotides from different templates in an undivided sample [Pastinen et al., Clin. Chem., 42:1391-1397 (1996)]. Fluorescence resonance energy transfer (FRET) has been used in combination with minisequencing to detect polymorphic variants [U.S. Pat. No. 5,945,283; Chen et al., Proc. Natl. Acad. Sci. USA, 94:10756-10761 (1997)].
[0075] Many of the methods described involve amplification of DNA from target samples. This can be accomplished by e.g., PCR. Other suitable amplification methods include the ligase chain reaction (LCR) [see Wu and Wallace, Genomics 4, 560 (1989), Landegren et al., Science 241, 1077 (1988)], transcription amplification [Kwoh et al., Proc. Natl. Acad. Sci. USA 86, 1173 (1989)], self-sustained sequence replication [Guatelli et al., Proc. Nat. Acad. Sci. USA, 87, 1874 (1990)] and nucleic acid based sequence amplification (NASBA).
[0076] Single base extension methods are described by e.g., U.S. Pat. No. 5,846,710, U.S. Pat. No. 6,004,744, U.S. Pat. No. 5,888,819 and U.S. Pat. No. 5,856,092.

Amplification products generated using the polymerase chain reaction can be analyzed by the use of denaturing gradient gel electrophoresis. Different alleles can be identified based on the different sequence-dependent melting properties and electrophoretic migration of DNA in solution. [Erlich, ed., PCR Technology, Principles and Applications for DNA Amplification, (W. H. Freeman and Co, New York, (1992)), Chapter 7.]
[0077] Arrays provide a high throughput technique that can assay a large number of polynucleotides in a sample. Techniques for constructing arrays and methods of using these arrays are described in, for example, Schena et al., (1996) Proc Natl Acad Sci USA. 93(20):10614-9; Schena et al., (1995) Science 270(5235):467-70; Shalon et al., (1996) Genome Res. 6(7):639-45, U.S. Pat. No. 5,807,522, EP 799 897; WO 97/29212; WO 97/27317; EP 785 280; WO 97/02357; U.S. Pat. No. 5,593,839; U.S. Pat. No. 5,578,832; EP 728 520; U.S. Pat. No. 5,599,695; EP 721 016; U.S. Pat. No. 5,556,752; WO 95/22058; and U.S. Pat. No. 5,631,734.
[0078] Screening may also be based on the functional or antigenic characteristics of the protein. Immunoassays designed to detect predisposing polymorphisms in proteins relevant to the invention can be used in screening. Antibodies specific for a polymorphism variant or gene products may be used in screening immunoassays. A
sample is taken from a subject. Samples, as used herein, include biological fluids such as tracheal lavage, blood, cerebrospinal fluid, tears, saliva, lymph, dialysis fluid and the like; organ or tissue culture derived fluids; and fluids extracted from physiological tissues. Samples can also include derivatives and fractions of such fluids. In some embodiments, the sample is derived from a biopsy. The number of cells in a sample will generally be at least about $10^{3}$, usually at least $10^{4}$ more usually at least about $10^{5}$. The cells can be dissociated, in the case of solid tissues, or tissue sections may be analyzed. Alternatively a lysate of the cells can be prepared.
[0079] In some embodiments, detection utilizes staining of cells or histological sections, performed in accordance with conventional methods. An alternative method for diagnosis depends on the in vitro detection of binding between antibodies and protein encoded by the polymorphic variant in a lysate. Other immunoassays are known in the art and may find use as diagnostics. Ouchterlony plates provide a simple determination of antibody binding. Western blots can be performed on protein gels or protein spots on filters, using a detection system specific for polymorphic variant protein as desired, conveniently using a labeling method as described for the sandwich assay.
[0080] The invention provides a method for determining a genotype of an individual in relation to one or more polymorphic variants in one or more of the genes identified in above aspects by using mass spectrometric determination of a nucleic acid sequence that is a portion of a gene identified for other aspects of this invention or a complementary sequence. Such mass spectrometric methods are known to those skilled in the art.
[0081] The detection of the presence or absence of a polymorphic variant can involve contacting a nucleic acid sequence corresponding to one of the genes identified above or a product of such a gene with a probe. The probe is able to distinguish a particular form of the gene, gene product, polymorphic variant allele product, or allele product, or the presence or a particular polymorphic variant or polymorphic variants, for example, by differential binding or hybridization. The term "probe" refers to a molecule that can detectably distinguish between target molecules differing in structure. Detection can be accomplished in a variety of different ways depending on the type of probe used and the type of target molecule. Thus, for example, detection may be based on discrimination of activity levels of the target molecule, but preferably is based on detection of specific binding. Examples of such specific binding include antibody binding and nucleic acid
probe hybridization. Probes can comprise one or more of the following, a protein, carbohydrate, polymer, or small molecule, that is capable of binding to one polymorphic variant or variant form of the gene or gene product to a greater extent than to a form of the gene having a different base at one or more polymorphic variant sites, such that the presence of the polymorphic variant or variant form of the gene can be determined. A probe can incorporate one or more markers including, but not limited to, radioactive labels, such as radionuclides, fluorophores or fluorochromes, peptides, enzymes, antigens, antibodies, vitamins or steroids. A probe can distinguishe at least one of the polymeric variant described herein. The probe can also have specificity for the particular gene or gene product, at least to an extent such that binding to other genes or gene products does not prevent use of the assay to identify the presence or absence of the particular polymorphic variant or polymorphic variants of interest.
[0082] The nucleic acid molecules relevant to the invention can readily be obtained in a variety of ways, including, without limitation, chemical synthesis, cDNA or genomic library screening, expression library screening, and/or PCR amplification of cDNA. These methods and others useful for isolating such DNA are set forth, for example, by Sambrook, et al., "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), by Ausubel, et al., eds., "Current Protocols In Molecular Biology," Current Protocols Press (1994), and by Berger and Kimmel, "Methods In Enzymology: Guide To Molecular Cloning Techniques," vol. 152, Academic Press, Inc., San Diego, Calif. (1987). Nucleic acid sequences are mammalian sequences. In some embodiments, the nucleic acid sequences are human, rat, and mouse.
[0083] Chemical synthesis of a nucleic acid molecule can be accomplished using methods well known in the art, such as those set forth by Engels et al., Angew. Chem. Intl. Ed., 28:716-734 (1989). These methods include, inter alia, the phosphotriester, phosphoramidite and H-phosphonate methods of nucleic acid synthesis. Nucleic acids larger than about 100 nucleotides in length can be synthesized as several fragments, each fragment being up to about 100 nucleotides in length. The fragments can then be ligated together to form a full length nucleic acid encoding the polypeptide. A preferred method is polymer-supported synthesis using standard phosphoramidite chemistry.
[0084] Alternatively, the nucleic acid may be obtained by screening an appropriate cDNA library prepared from one or more tissue source(s) that express the polypeptide, or
a genomic library from any subspecies. The source of the genomic library may be any tissue or tissues from any mammalian or other species believed to harbor a gene encoding a protein relevant to the invention. The library can be screened for the presence of a cDNA/gene using one or more nucleic acid probes (oligonucleotides, cDNA or genomic DNA fragments that possess an acceptable level of homology to the gene or gene homologue cDNA or gene to be cloned) that will hybridize selectively with the gene or gene homologue $\mathrm{cDNA}(\mathrm{s})$ or gene(s) that is(are) present in the library. The probes preferably are complementary to or encode a small region of the DNA sequence from the same or a similar species as the species from which the library can be prepared.
Alternatively, the probes may be degenerate, as discussed below. After hybridization, the blot containing the library is washed at a suitable stringency, depending on several factors such as probe size, expected homology of probe to clone, type of library being screened, number of clones being screened, and the like. Stringent washing solutions are usually low in ionic strength and are used at relatively high temperatures.
[0085] Another suitable method for obtaining a nucleic acid in accordance with the invention is the polymerase chain reaction (PCR). In this method, poly(A)+RNA or total RNA is extracted from a tissue that expresses the gene product. cDNA is then prepared from the RNA using the enzyme reverse transcriptase. Two primers typically complementary to two separate regions of the cDNA (oligonucleotides) are then added to the cDNA along with a polymerase such as Taq polymerase, and the polymerase amplifies the cDNA region between the two primers.
[0086] The invention provides for the use of isolated, purified or enriched nucleic acid sequences of 15 to 500 nucleotides in length, 15 to 100 nucleotides in length, 15 to 50 nucleotides in length, and 15 to 30 nucleotides in length, which have sequence that corresponds to a portion of one of the genes identified for aspects above. In some embodiments the nucleic acid is at least $17,20,22$, or 25 nucleotides in length. In some embodiments, the nucleic acid sequence is 30 to 300 nucleotides in length, or 45 to 200 nucleotides in length, or 45 to 100 nucleotides in length. In some embodiments, the probe is a nucleic acid probe at least $15,1720,2225,30,35,40$, or more nucleotides in length, or $500,250,200,100,50,40,30$ or fewer nucleotides in length. In preferred embodiments, the probe has a length in a range from any one of the above lengths to any other of the above lengths including endpoints. The nucleic acid sequence includes at
least one polymorphic variant site. Such sequences can, for example, be amplification products of a sequence that spans or includes a polymorphic variant site in a gene identified herein. A nucleic acid with such a sequence can be utilized as a primer or amplification oligonucleotide that is able to bind to or extend through a polymorphic variant site in such a gene. Another example is a nucleic acid hybridization probe comprised of such a sequence. In such probes, primers, and amplification products, the nucleotide sequence can contain a sequence or site corresponding to a polymorphic variant site or sites, for example, a polymorphic variant site identified herein. The design and use of allele-specific probes for analyzing polymorphisms is known generally in the art, see, for example, Saiki et al., Nature 324:163-166 (1986); Dattagupta, EP 235,726, Saiki, WO 89/11548. Allele-specific probes can be designed that hybridize to a segment of target DNA from one individual but do not hybridize to the corresponding segment from another individual due to the presence of different polymorphic forms in the respective segments from the two individuals. A nucleic acid hybridization probe may span two or more polymorphic variant sites. Unless otherwise specified, a nucleic acid probe can include one or more nucleic acid analogs, labels or other substituents or moieties so long as the base-pairing function is retained. The nucleic acid sequence includes at least one polymorphic variant site. The probe may also comprise a detectable label, such as a radioactive or fluorescent label. A variety of other detectable labels are known to those skilled in the art. Nucleic acid probe can also include one or more nucleic acid analogs. [0087] In connection with nucleic acid probe hybridization, the term "specifically hybridizes" indicates that the probe hybridizes to a sufficiently greater degree to the target sequence than to a sequence having a mismatched base at at least one polymorphic variant site to allow distinguishing of such hybridization. The term "specifically hybridizes" means that the probe hybridizes to the target sequence, and not to non-target sequences, at a level which allows ready identification of probe/target sequence hybridization under selective hybridization conditions. "Selective hybridization conditions" refer to conditions that allow such differential binding. Similarly, the terms "specifically binds" and "selective binding conditions" refer to such differential binding of any type of probe, and to the conditions that allow such differential binding. Hybridization reactions to determine the status of variant sites in patient samples can be carried out with two different probes,
one specific for each of the possible variant nucleotides. The complementary information derived from the two separate hybridization reactions is useful in corroborating the results.
[0088] A variety of variables can be adjusted to optimize the discrimination between two variant forms of a gene, including changes in salt concentration, temperature, pH and addition of various compounds that affect the differential affinity of GC vs. AT base pairs, such as tetramethyl ammonium chloride. [See Current Protocols in Molecular Biology, Ausubel et al. (Editors), John Wiley \& Sons.] Hybridization conditions should be sufficiently stringent such that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. Hybridizations are usually performed under stringent conditions that allow for specific binding between an oligonucleotide and a target nucleic acid containing one of the polymorphic sites described herein or identified using the techniques described herein. Stringent conditions are defined as any suitable buffer concentrations and temperatures that allow specific hybridization of the oligonucleotide to highly homologous sequences spanning at least one polymorphic site and any washing conditions that remove non-specific binding of the oligonucleotide. For example, conditions of $5 \mathrm{xSSPE}(750 \mathrm{mM} \mathrm{NaCl}, 50 \mathrm{mM}$ Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of $25-30^{\circ} \mathrm{C}$ are suitable for allele-specific probe hybridizations. The washing conditions usually range from room temperature to $60^{\circ} \mathrm{C}$. Some probes are designed to hybridize to a segment of target DNA such that the polymorphic site aligns with a central position of the probe. This probe design achieves good discrimination in hybridization between different allelic forms.
[0089] Allele-specific probes are can be used in pairs, one member of a pair showing a perfect match to a reference form of a target sequence and the other member showing a perfect match to a variant form. Several pairs of probes can then be immobilized on the same support for simultaneous analysis of multiple polymorphisms within the same target sequence. The polymorphisms can also be identified by hybridization to nucleic acid arrays, some examples of which are described by WO 95/11995. Arrays may be provided in the form of a multiplex chip.
[0090] One use of probe(s) is as a primer(s) that hybridizes to a nucleic acid sequence containing at least one sequence polymorphic variant. Preferably such primers hybridize to a sequence not more than 300 nucleotides, more preferably not more than 200
nucleotides, still more preferably not more than 100 nucleotides, and most preferably not more than 50 nucleotides away from a polymorphic variant site which is to be analyzed. Preferably, a primer is 100 nucleotides or fewer in length, more preferably 50 nucleotides or fewer, still more preferable 30 nucleotides or fewer, and most preferably 20 or fewer nucleotides in length. In some embodiments, the set includes primers or amplification oligonucleotides adapted to bind to or extend through a plurality of sequence polymorphic variants in a gene(s) identified herein. In some embodiments, the plurality of polymorphic variants comprises a haplotype. In certain embodiments, the oligonucleotides are designed and selected to provide polymorphic variant-specific amplification.
[0091] Another type of probe is a peptide or protein, for example, an antibody or antibody fragment that specifically or preferentially binds to a polypeptide expressed by a particular form of a gene as characterized by the presence or absence of at least one polymorphic variant. Such antibodies may be polyclonal or monoclonal antibodies, and can be prepared by methods well-known in the art.
[0092] Antibodies can be used to probe for presence of a given polymorphism variant for those polymorphism variants that have an effect on the polypeptide encoded by the gene. For example, an antibody can recognize a change in one or more amino acid residues in the resulting protein. In some embodiments, the antibody is used to recognize polypeptides encoded by differential splice variants. If the polymorphism introduces or eliminates a surface feature of the protein such as a glycosylation site, lipid modification, etc., an antibody can also be used to identify a particular variant.
[0093] Polyclonal and/or monoclonal antibodies and antibody fragments capable of binding to a portion of the gene product relevant for identifying a given polymorphism variant are provided. Antibodies can be made by injecting mice or other animals with the variant gene product or synthetic peptide fragments thereof. Monoclonal antibodies are screened as are described, for example, in Harlow \& Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Press, New York (1988); Goding, Monoclonal antibodies, Principles and Practice (2d ed.) Academic Press, New York (1986). Monoclonal antibodies are tested for specific immunoreactivity with a variant gene product and lack of immunoreactivity to the corresponding prototypical gene product. These antibodies are useful in diagnostic assays for detection of the variant form, or as an active ingredient in a pharmaceutical composition.
[0094] The invention provides methods for choosing a relevant therapeutic strategy based on the detection of the presence or absence of one or more polymorphic variants. General methods of testing effects of a polymorphic variant for an effect on drug efficacy are known to those of skill in the art and are provided in various sources such as U.S. Patent Nos. 6,537,759; 6,664,062; and 6,759,200.
[0095] A therapeutic agent, which can be a compound and/or a composition, relevant to the invention can comprise a small molecule, a nucleic acid, a protein, an antibody, or any other agent with one or more therapeutic property. The therapeutic agent can be formulated in any pharmaceutically acceptable manner. In some embodiments, the therapeutic agent is prepared in a depot form to allow for release into the body to which it is administered is controlled with respect to time and location within the body (see, for example, U.S. Patent No. 4,450,150). Depot forms of therapeutic agents can be, for example, an implantable composition comprising the therapeutic agent and a porous or non-porous material, such as a polymer, wherein the therapeutic agent is encapsulated by or diffused throughout the material and/or degradation of the non-porous material. The depot is then implanted into the desired location within the body and the therapeutic agent is released from the implant at a predetermined rate.
[0096] The therapeutic agent that is used in the invention can be formed as a composition, such as a pharmaceutical composition comprising a carrier and a therapeutic compound. Pharmaceutical compositions containing the therapeutic agent can comprise more than one therapeutic agent. The pharmaceutical composition can alternatively comprise a therapeutic agent in combination with other pharmaceutically active agents or drugs, such as chemotherapeutic agents, for example, a cancer drug.
[0097] The carrier can be any suitable carrier. Preferably, the carrier is a pharmaceutically acceptable carrier. With respect to pharmaceutical compositions, the carrier can be any of those conventionally used and is limited only by chemico physical considerations, such as solubility and lack of reactivity with the active compound(s), and by the route of administration. In addition to the following described pharmaceutical composition, the therapeutic compounds of the present inventive methods can be formulated as inclusion complexes, such as cyclodextrin inclusion complexes, or liposomes.
[0098] The pharmaceutically acceptable carriers described herein, for example, vehicles, adjuvants, excipients, and diluents, are well-known to those skilled in the art and are readily available to the public. The pharmaceutically acceptable carrier can be chemically inert to the active agent(s) and one which has no detrimental side effects or toxicity under the conditions of use. The choice of carrier can be determined in part by the particular therapeutic agent, as well as by the particular method used to administer the therapeutic compound. There are a variety of suitable formulations of the pharmaceutical composition of the invention. The following formulations for oral, aerosol, parenteral, subcutaneous, transdermal, transmucosal, intestinal, parenteral, intramedullary injections, direct intraventricular, intravenous, intranasal, intraocular, intramuscular, intraarterial, intrathecal, interperitoneal, rectal, and vaginal administration are exemplary and are in no way limiting. More than one route can be used to administer the therapeutic agent, and in some instances, a particular route can provide a more immediate and more effective response than another route. Depending on the specific conditions being treated, such agents can be formulated and administered systemically or locally. Techniques for formulation and administration may be found in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, Pa. (1990).
[0099] Formulations suitable for oral administration can include (a) liquid solutions, such as an effective amount of the inhibitor dissolved in diluents, such as water, saline, or orange juice; (b) capsules, sachets, tablets, lozenges, and troches, each containing a predetermined amount of the active ingredient, as solids or granules; (c) powders; (d) suspensions in an appropriate liquid; and (e) suitable emulsions. Liquid formulations may include diluents, such as water and alcohols, for example, ethanol, benzyl alcohol, and the polyethylene alcohols, either with or without the addition of a pharmaceutically acceptable surfactant. Capsule forms can be of the ordinary hard or soft shelled gelatin type containing, for example, surfactants, lubricants, and inert fillers, such as lactose, sucrose, calcium phosphate, and corn starch. Tablet forms can include one or more of lactose, sucrose, mannitol, corn starch, potato starch, alginic acid, microcrystalline cellulose, acacia, gelatin, guar gum, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, calcium stearate, zinc stearate, stearic acid, and other excipients, colorants, diluents, buffering agents, disintegrating agents, moistening agents, preservatives, flavoring agents, and other pharmacologically compatible excipients.

Lozenge forms can comprise the inhibitor in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the inhibitor in an inert base, such as gelatin and glycerin, or sucrose and acacia, emulsions, gels, and the like containing, in addition to, such excipients as are known in the art.
[0100] Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.
[0101] The therapeutic agent, alone or in combination with other suitable components, can be made into aerosol formulations to be administered via inhalation. These aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. They also can be formulated as pharmaceuticals for non pressured preparations, such as in a nebulizer or an atomizer. Such spray formulations also may be used to spray mucosa. Topical formulations are well known to those of skill in the art. Such formulations are particularly suitable in the context of the invention for application to the skin.
[0102] Injectable formulations are in accordance with the invention. The parameters for effective pharmaceutical carriers for injectable compositions are well-known to those of ordinary skill in the art [see, e.g., Pharmaceutics and Pharmacy Practice, J.B. Lippincott Company, Philadelphia, PA, Banker and Chalmers, eds., pages 238250 (1982), and ASHP Handbook on Injectable Drugs, Toissel, 4th ed., pages 622630 (1986)]. For injection, the agents of the invention can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For such transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.
[0103] Formulations suitable for parenteral administration include aqueous and non aqueous, isotonic sterile injection solutions, which can contain anti oxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the
intended recipient, and aqueous and non aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The therapeutic agent can be administered in a physiologically acceptable diluent in a pharmaceutical carrier, such as a sterile liquid or mixture of liquids, including water, saline, aqueous dextrose and related sugar solutions, an alcohol, such as ethanol or hexadecyl alcohol, a glycol, such as propylene glycol or polyethylene glycol, dimethylsulfoxide, glycerol, ketals such as 2,2-dimethyl-1,3-dioxolane-4-methanol, ethers, poly(ethyleneglycol) 400, oils, fatty acids, fatty acid esters or glycerides, or acetylated fatty acid glycerides with or without the addition of a pharmaceutically acceptable surfactant, such as a soap or a detergent, suspending agent, such as pectin, carbomers, methylcellulose, hydroxypropylmethylcellulose, or carboxymethylcellulose, or emulsifying agents and other pharmaceutical adjuvants.
[0104] Oils, which can be used in parenteral formulations include petroleum, animal, vegetable, or synthetic oils. Specific examples of oils include peanut, soybean, sesame, cottonseed, corn, olive, petrolatum, and mineral. Suitable fatty acids for use in parenteral formulations include oleic acid, stearic acid, and isostearic acid. Ethyl oleate and isopropyl myristate are examples of suitable fatty acid esters.
[0105] Suitable soaps for use in parenteral formulations include fatty alkali metal, ammonium, and triethanolamine salts, and suitable detergents include (a) cationic detergents such as, for example, dimethyl dialkyl ammonium halides, and alkyl pyridinium halides, (b) anionic detergents such as, for example, alkyl, aryl, and olefin sulfonates, alkyl, olefin, ether, and monoglyceride sulfates, and sulfosuccinates, (c) nonionic detergents such as, for example, fatty amine oxides, fatty acid alkanolamides, and polyoxyethylenepolypropylene copolymers, (d) amphoteric detergents such as, for example, alkyl- $\beta$-aminopropionates, and 2-alkyl-imidazoline quaternary ammonium salts, and (e) mixtures thereof.
[0106] The parenteral formulations will typically contain from about $0.5 \%$ to about $25 \%$ by weight of the drug in solution. Preservatives and buffers may be used. In order to minimize or eliminate irritation at the site of injection, such compositions may contain one or more nonionic surfactants having a hydrophile-lipophile balance (HLB) of from about 12 to about 17. The quantity of surfactant in such formulations will typically range from about $5 \%$ to about $15 \%$ by weight. Suitable surfactants include polyethylene glycol
sorbitan fatty acid esters, such as sorbitan monooleate and the high molecular weight adducts of ethylene oxide with a hydrophobic base, formed by the condensation of propylene oxide with propylene glycol. The parenteral formulations can be presented in unit-dose or multi-dose sealed containers, such as ampoules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid excipient, for example, water, for injections, immediately prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described.
[0107] The therapeutic agent can be made into suppositories by mixing with a variety of bases, such as emulsifying bases or water-soluble bases. Formulations suitable for vaginal administration can be presented as pessaries, tampons, creams, gels, pastes, foams, or spray formulas containing, in addition to the active ingredient, such carriers as are known in the art to be appropriate.
[0108] The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. [See, e.g., Fingl et. al., in The Pharmacological Basis of Therapeutics, 1975, Ch. 1p.1]. The attending physician can determine when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician can also adjust treatment to higher levels if the clinical response were not adequate, precluding toxicity. The magnitude of an administrated dose in the management of disorder of interest will vary with the severity of the condition to be treated and the route of administration. The severity of the condition may, for example, be evaluated, in part, by standard prognostic evaluation methods. The dose and perhaps dose frequency, can vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above can be used in veterinary medicine.
[0109] Use of pharmaceutically acceptable carriers to formulate the compounds herein disclosed for the practice of the invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions relevant to the invention, in particular, those formulated as solutions, can be administered parenterally, such as by intravenous injection. The compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers
enable the compounds relevant to the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, tablets, dragees, solutions, suspensions and the like, for oral ingestion by a patient to be treated.
[0110] Agents intended to be administered intracellularly may be administered using techniques well known to those of ordinary skill in the art. For example, such agents may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are efficiently delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, small organic molecules may be directly administered intracellularly.
[0111] The altered susceptibility can be either an increased or decreased susceptibility for a drug-induced heart rhythm irregularity. The relative susceptibility can be measured according to any acceptable medical parameters. Generally, the susceptibility is gauged relative to a subject that lacks the polymorphic variant or is heterozygous for the polymorphic variant. In some embodiments, the measure would be homozygous for the polymorphic variant or heterozygous for the polymorphic variant relative to a subject that is homozygous lacking the polymorphic variant. In some embodiments, two or more polymorphic variants for a give polymorphism are taken to be equivalent to each other relative to two or more polymorphic variants for the polymorphism.
[0112] According to one aspect, the method comprises not only screening and diagnosing steps, but also prescribing a treatment regimen based on the diagnosis. In some embodiments, the treatment regimen comprises increasing dosage of the drug in the presence of a polymorphic variant associated with a decreased susceptibility for the heart rhythm irregularity. In some embodiments, the treatment regimen comprises increasing dosage of the drug in the absence of a polymorphic variant associated with an increased susceptibility for the heart rhythm irregularity. In some embodiments, the treatment regimen comprises decreasing dosage of the drug in the presence of a polymorphic variant associated with an increased susceptibility for the heart rhythm irregularity. In some embodiments, the treatment regimen comprises decreasing dosage of the drug in the absence of a polymorphic variant associated with a decreased susceptibility for the heart
rhythm irregularity. For example, one could decide based on the screening and diagnosis to not administer the heart rhythm irregularity inducing drug. In some such cases, a different drug is administered. In some embodiments, the drug does not bind ABCB1. In some embodiments, the treatment regimen comprises increased heart monitoring.
[0113] In another aspect, the screening and diagnosis result in the administration of one or more additional drug is administered. In some embodiments, the second drug ameliorates the heart rhythm irregularity.
[0114] The invention provides selecting a method of administration of an agent to a patient suffering from a disease or condition, by determining the presence or absence of at least one polymorphic variant in cells of the patient, where such presence or absence is indicative of an appropriate method of administration of the agent. The selection of a treatment regimen can involve selecting a dosage level or frequency of administration or route of administration of the agent(s) or combinations of those parameters. In some embodiments, two or more agents are administered, and the selecting involves selecting a method of administration for one, two, or more than two of the agents, jointly, concurrently, or separately. As understood by those skilled in the art, such plurality of agents is often used in combination therapy, and thus may be formulated in a single drug, or may be separate drugs administered concurrently, serially, or separately. Other embodiments are as indicated above for selection of second treatment methods, methods of identifying polymorphic variants, and methods of treatment as described for aspects above. The frequency of administration is generally selected to achieve a pharmacologically effective average or peak serum level without excessive deleterious effects. In some embodiments, the serum level of the drug is maintained within a therapeutic window of concentrations for the greatest percentage of time possible without such deleterious effects as would cause a prudent physician to reduce the frequency of administration for a particular dosage level. Administration of a particular treatment, for example, administration of a therapeutic compound or combination of compounds, is chosen depending on the disease or condition which is to be treated. In some embodiments, the disease or condition is one for which administration of a treatment is expected to provide a therapeutic benefit. In embodiments involving selection of a patient for a treatment, selection of a method or mode of administration of a treatment, and selection of a patient for a treatment or a method of treatment, the selection can be positive
selection or negative selection. The methods can include modifying or eliminating a treatment for a patient, modifying or eliminating a method or mode of administration of a treatment to a patient, or modification or elimination of a patient for a treatment or method of treatment. A patient can be selected for a method of administration of a treatment, by detecting the presence or absence of at least one polymorphic variant in a gene as identified herein, where the presence or absence of the at least one polymorphic variant is indicative that the treatment or method of administration will be effective in the patient. If the at least one polymorphic variant is present in the patient's cells, then the patient is selected for administration of the treatment.
[0115] The term "drug" or "therapeutic agent" as used herein refers to a chemical entity or biological product, or combination of chemical entities or biological products, administered to a person to treat or prevent or control a disease or condition. In some embodiments, the chemical entity or biological product is a low molecular weight compound. A "low molecular weight compound" has a molecular weight $<5,000 \mathrm{Da}$, $<2500 \mathrm{Da},<1000 \mathrm{Da}$, or $<700 \mathrm{Da}$. In some embodiments, the chemical entity is a larger compound, for example, an oligomer of nucleic acids, amino acids, or carbohydrates including without limitation proteins, oligonucleotides, ribozymes, DNAzymes, glycoproteins, lipoproteins, and modifications and combinations thereof. In some embodiments, the biological product is a monoclonal or polyclonal antibody or fragment thereof such as a variable chain fragment cells; or an agent or product arising from recombinant technology, such as, without limitation, a recombinant protein, recombinant vaccine, or DNA construct developed for therapeutic use. The term "drug" or "therapeutic agent" can include, without limitation, compounds that are approved for sale as pharmaceutical products by government regulatory agencies such as the U.S. Food and Drug Administration (USFDA or FDA), the European Medicines Evaluation Agency (EMEA), and a world regulatory body governing the Intemation Conference of Harmonization (ICH) rules and guidelines, compounds that do not require approval by government regulatory agencies, food additives or supplements including compounds commonly characterized as vitamins, natural products, and completely or incompletely characterized mixtures of chemical entities including natural compounds or purified or partially purified natural products. In some embodiments, the drug is approved by a government agency for treatment of a specific disease or condition.
[0116] In treating a patient exhibiting a disorder of interest, a therapeutically effective amount of a agent or agents is administered. A therapeutically effective dose refers to that amount of the compound that results in amelioration of one or more symptoms or a prolongation of survival in a patient. The amount or dose of the therapeutic compound administered should be sufficient to affect a therapeutic response in the subject or animal over a reasonable time frame. For example, in the case of cancer, the dose of the therapeutic compound should be sufficient to inhibit metastasis, prevent metastasis, treat or prevent cancer in a period of from about 2 hours or longer, e.g., 12 to 24 or more hours, from the time of administration. In certain embodiments, the time period could be even longer. The dose can be determined by the efficacy of the particular therapeutic agent and the condition of the subject, as well as the body weight of the subject to be treated. Many assays for determining an administered dose are known in the art.
[0117] The dose of the therapeutic compound can also be determined by the existence, nature and extent of any adverse side effects that might accompany the administration of a particular therapeutic compound. The attending physician can decide the dosage of the inhibitor relevant to the invention with which to treat each individual patient using the correlation between polymorphic variant and disease and/or drug efficacies provided by the invention and taking into consideration a variety of factors, such as age, body weight, general health, diet, sex, inhibitor to be administered, route of administration, and the severity of the condition being treated. In some embodiments, the dose of the therapeutic compound is about 0.001 to about $1000 \mathrm{mg} / \mathrm{kg}$ body weight of the subject being treated/day, from about 0.01 to about $10 \mathrm{mg} / \mathrm{kg}$ body weight/day, about 0.01 mg to about $1 \mathrm{mg} / \mathrm{kg}$ body weight/day.
[0118] Toxicity and therapeutic efficacy of therapeutic agents can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, for example, for determining the $\mathrm{LD}_{50}$ and the $\mathrm{ED}_{50}$. The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio $\mathrm{LD}_{50} / E D_{50}$. In some embodiments, compounds that exhibit large therapeutic indices are used. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds can lie within a range of circulating concentrations that can include the $\mathrm{ED}_{50}$ with little or no toxicity. The dosage can vary within this range depending upon the dosage form and route of
administration utilized. The therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the $\mathrm{IC}_{50}$ as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by HPLC.
[0119] In connection with the administration of a drug, a drug which is "effective against" a disease or condition indicates that administration in a clinically appropriate manner results in a beneficial effect for at least a statistically significant fraction of patients, such as a improvement of symptoms, a cure, a reduction in disease load, reduction in tumor mass or cell numbers, extension of life, improvement in quality of life, or other effect generally recognized as positive by those of skill in the art.
[0120] In some embodiments, the drug is an anti-cancer agent. Examples of anticancer agents include actinomycin D , daunorubicin, docetaxel, doxorubicin, erlotinib, etoposide, gefitinib, imatinib, irinotecan, mitomycin c, mitoxantrone, paclitaxel, $\mathrm{SN}-38$, teniposide, topotecan, vinblastine, vincristine, a prodrug thereof, a salt thereof, or a combination thereof. Another applicable cancer drug is a depsipeptide, e.g., FK228, as well as prodrugs, salts and combination thereof. FK228 is also known as romidepsin. In some embodiments, the FK228 is the isomer FR901228, which is (E)-(1S, 4S, 10S, 21R)-7-[(Z)-ethylidene]-4,21-diisopropyl-2-oxa-12,13-dithia-5,8,20,23-tetraazabicyclo [8,7,6]-tricos-16-ene-3,6,19,22-pentanone (NSC 630176). FK228 compounds, salts, prodrugs, formulation, method of preparation, dosage, administration, and other FK228 parameters can be used in accordance with the materials and method of this invention. The salt of FK228, e.g., FR901228, is a biologically acceptable salt, which is generally non-toxic, and is exemplified by salts with base or acid addition salts, inclusive of salts with inorganic base such as alkali metal salt (e.g., a sodium salt, a potassium salt, etc.), alkaline earth metal salt (e.g., calcium salt, magnesium salt, etc.), ammonium salt, salts with organic base such as organic amine salt (e.g., triethylamine salt, diisopropylethylamine salt, pyridine salt, picoline salt, ethanolamine salt, diethanolamine salt, triethanolamine salt, dicyclohexylamine salt, $\mathrm{N}, \mathrm{N}^{\prime}$-dibenzylethylenediamine salt, etc.), inorganic acid salt (e.g., hydrochloride, hydrobromide, sulfate, phosphate, etc.), organic carboxylic or sulfonic acid salt (e.g., formate, acetate, trifluoroacetate, maleate, tartrate, fumarate, methanesulfonate, benzenesulfonate, toulenesulfonate, etc.), salt with basic or acid amino acid (e.g., arginine,
aspartic acid, glutamic acid, etc.), and the like. Examples of relevant FK228 parameters, as well as parameters for other depsipeptides and histone deacetylase inhibitors (HDIs), applicable to the invention are provided in U.S. Provisional Application Nos. 60/226,234 and 60/709,553; WO 02/15921; WO 03/084611; and WO 02/055688.
[0121] Drugs applicable to the method are not limited to anti-cancer drugs. The heart rhythm irregularity inducing drug can be an antacid. Examples of antacids include cimetidine, ranitidine, a prodrug thereof, a salt thereof, or a combination thereof. In some embodiments, the heart rhythm inducing drug is an antiarrhythmic. Examples of such antiarrthymics include amiodarone, digoxin, propafenone, quinidine, verapamil, a prodrug thereof, a salt thereof, or a combination thereof. The heart rhythm irregularity inducing drug can be an antibiotic. Examples of such antibiotics include clarithromycin, erythromycin, levofloxacin, rifampin, sparfloxacin, tetracycline, a prodrug thereof, a salt thereof, or a combination thereof. In some embodiments, the drug is an antidepressant, such as amitriptyline, fluoxetine, paroxetine, sertraline, St John's wort, a prodrug thereof, a salt thereof, or a combination thereof. The drug can be an antiemetic. Examples of such antiemetics include domperidon, ondansetron, a prodrug thereof, a salt thereof, or a combination thereof. In some embodiments, the drug is an antiepileptic such as phenobarbital, phenytoin, a prodrug thereof, a salt thereof, or a combination thereof. The drug can also be an antihypertensive. Examples of antihypertensives include carvedilol, celiprolol, diltiazem, losartan, nicardipine, reserpine, talinolol, a prodrug thereof, a salt thereof, or a combination thereof.
[0122] In some embodiments, the heart rhythm irregularity inducing drug is an antimycotic. Examples of such antimycotics include itraconazole, ketoconazole, a prodrug thereof, a salt thereof, or a combination thereof. The drug can be an antiviral agent. Examples of antiviral agents include amprenavir, indinavir, nelfinavir, ritonavir, saquinavir, a prodrug thereof, a salt thereof, or a combination thereof. The drug can be a glucocorticoid such as aldosterone, cortisol, dexamethasone, methylprednisolone, a prodrug thereof, a salt thereof, or a combination thereof. In some embodiments, the drug is an immunosuppressant. Examples of such immunosuppressants include cyclosporine, sirolimus, tacrolimus, valspodar, a prodrug thereof, a salt thereof, or a combination thereof. The drug can also be a neuroleptic such as chloropromazine, flupenthixol, phenothiazine, a prodrug thereof, a salt thereof, or a combination thereof. In some
embodiments, the drug is an opioid. Examples of such opioid include methadone, morphine, pentazocine, a prodrug thereof, a salt thereof, or a combination thereof.
[0123] In some embodiments, the heart rhythm irregularity inducing drug is selected from the group consisting of torvastatin, bromocriptine, colchicine, dipyridamole, emetine, fexofenadine, ivermectin, loperamide, mefloquine, progesterone, retinoic acid, rhodamine 123, spironolactone, terfenadine, vecuronium, a prodrug thereof, a salt thereof, or a combination thereof.
[0124] Kits compatible with the methods are also provided. In one aspect, a kit is provided that includes a nucleic acid and a drug that binds a protein encoded ABCB1. The nucleic acid is for use in screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB1 gene, and wherein the nucleic acid specifically binds to $A B C B 1$ sequence comprising the at least one polymorphism or a sequence adjacent to ABCB 1 sequence comprising the at least one polymorphism. In one aspect, the polymorphism comprises polymorphism identified as rs1128503, rs2032582, rs1045642, or a combination thereof. In one aspect, the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1; or 1236, 2677, or 3435 of SEQ ID NO: 2; or a combination thereof. In another aspect, the drug is FK228 and/or another drug described herein. In some embodiments, the kit's nucleic acid comprises the nucleotide sequence of any one of SEQ ID NOS: 25-36 or a compliment thereof or a combination thereof.
[0125] The invention includes kits for the detection of polymorphic variants associated with disease states, conditions or complications. The kits can comprise a polynucleotide of at least 30 contiguous nucleotides of one of the variants described herein. In one embodiment, the polynucleotide contains at least one polymorphism of the invention. Alternatively, the $3^{\prime}$ end of the polynucleotide is immediately 5 ' to a polymorphic site, preferably a polymorphic site of the invention. In one embodiment, the polymorphic site contains a genetic variant. In still another embodiment, the genetic variant is located at the $3^{\prime}$ end of the polynucleotide. In yet another embodiment, the polynucleotide of the kit contains a detectable label. Suitable labels include, but are not limited to, radioactive labels, such as radionucleotides, fluorophores or fluorochromes,
peptides, enzymes, antigens, antibodies, vitamins or steroids. The kit may also contain additional materials for detection of the polymorphisms. A kit can contain one or more of the following: buffer solutions, enzymes, nucleotide triphosphates, and other reagents and materials useful for the detection of genetic polymorphisms. Kits can contain instructions for conducting analyses of samples for the presence of polymorphisms and for interpreting the results obtained.
[0126] In some embodiments, the kit contains one or more pairs of allele-specific oligonucleotides hybridizing to different forms of a polymorphism. In some embodiments, the kit contains at least one probe or at least one primer or both corresponding to a gene or genes relevant to the invention. The kit can be adapted and configured to be suitable for identification of the presence or absence of one or more polymorphic variants. The kit can contain a plurality of either or both of such probes and/or primers, for example, $2,3,4,5,6$, or more of such probes and/or primers. The plurality of probes and/or primers are adapted to provide detection of a plurality of different sequence polymorphic variants in a gene or plurality of genes, for example, in 2, $3,4,5$, or more genes or to sequence a nucleic acid sequence including at least one polymorphic variant site in a gene or genes. In some embodiments, the kit contains components for detection of a plurality of polymorphic variants indicative of the effectiveness of a treatment or treatment against a plurality of diseases. Additional kit components can include one or more of the following: a buffer or buffers, such as amplification buffers and hybridization buffers, which may be in liquid or dry form, a DNA polymerase, such as a polymerase suitable for carrying out PCR, and deoxy nucleotide triphosphases (dNTPs). Preferably a probe includes a detectable label, for example, a fluorescent label, enzyme label, light scattering label, or other label. Additional components of the kit can also include restriction enzymes, reversetranscriptase or polymerase, the substrate nucleoside triphosphates, means used to label, for example, an avidin-enzyme conjugate and enzyme substrate and chromogen if the label is biotin, and the appropriate buffers for reverse transcription, PCR, or hybridization reactions.
[0127] In some kits, the allele-specific oligonucleotides are provided immobilized to a substrate. For example, the same substrate can comprise allele-specific oligonucleotide probes for detecting any or all of the polymorphism variants described herein.

Accordingly, the kit may comprise an array including a nucleic acid array and/or a polypeptide array. The array can include a plurality of different antibodies, a plurality of different nucleic acid sequences. Sites in the array can allow capture and/or detection of nucleic acid sequences or gene products corresponding to different polymorphic variants in one or more different genes. The array can be arranged to provide polymorphic variant detection for a plurality of polymorphic variants in one or more genes which correlate with the effectiveness of one or more treatments of one or more diseases.
[0128] The kit also can contain instructions for carrying out the methods. In some embodiments, the instructions include a listing of the polymorphic variants correlating with a particular treatment or treatments for a disease of diseases. The kit components can be selected to allow detection of a polymorphic variant described herein, and/or detection of a polymorphic variant indicative of a treatment, for example, administration of a drug. [0129] Uses of a drugs such as FK228 to manufacture a medicament are also provided. In one aspect, there is a use of a drug that binds a protein encoded by the ABCB 1 gene to manufacture a medicament to treat a subject that that has been screened for the presence or absence of at least one polymorphic variant of at least one polymorphism of the $A B C B 1$ gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by the drug. In one aspect, the polymorphism comprises polymorphism identified as rs1128503, rs2032582, rs1045642, or a combination thereof. In another aspect, the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2, or a combination thereof. Other uses such as uses analogous to the methods described herein are also provided.
[0130] The following examples further illustrate the invention but, of course, should not be construed as in any way limiting its scope.

## EXAMPLE 1

[0131] This example demonstrates that individuals with certain polymorphic variants in the ABCB 1 gene encounter fewer heart rhythm irregularities typically induced by FK228 treatment.
[0132] Subject eligibility criteria used are in accordance with those described in Piekarz et al, Blood 98:2865-8 (2001). Eligible patients have a confirmed diagnosis of
cutaneous T-cell lymphoma or relapsed peripheral T-cell lymphoma. Additional common eligibility criteria include: (i) a life expectancy of $\geq 12$ weeks; (ii) an Eastern Cooperative Group performance status $\leq 2$; (iii) no chemotherapy, hormonal therapy or radiotherapy, within four weeks prior to treatment; (iv) age above 18 years; (v) adequate contraception for women of child-bearing potential; and (vi) adequate bone marrow function (absolute neutrophil count, $>1.0 \times 10^{9} / \mathrm{L}$; platelets, platelet count, $>100 \times 10^{9} / \mathrm{L}$ ), renal function [serum creatinine, $\leq 1.5 \times$ the upper limit of normal (ULN)], and hepatic function (serum bilirubin, $\leq 1.5 \times$ ULN; and aspartate aminotransferase, $<3.0 \times$ ULN, unless impairment is due to organ involvement by lymphoma). The study protocol is approved by the local ethical review board, and all patients are provided written informed consent before study entry.
[0133] FK228 is supplied as a lyophilized powder by the Pharmaceutical Management Branch, Cancer Therapy Evaluation Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute (Bethesda, Md) in sterile vials containing 10 mg of drug and 20 mg of povidine as a bulking agent. Immediately prior to drug administration, FK228 is reconstituted in 2 mL of a diluent containing a mixture of propylene glycol and ethanol ( $4: 1, \mathrm{vol} / \mathrm{vol}$ ). This $5-\mathrm{mg} / \mathrm{mL}$ solution is diluted in 500 mL or 1000 mL of sodium chloride for injection, USP. FK228 is administered as a 4-hour continuous infusion on days 1,8 , and 15 via a portable infusion pump, with cycles repeated every 21 days. Provided toxic effects are not prohibitive, patients are eligible to continue treatment until there is evidence of progressive disease.
[0134] Complete blood cell counts with differential are obtained immediately prior to FK228 administration and on days 2, 9, and 16 to evaluate FK228-related myelosuppression. Multiple surface electrocardiograms (ECGs) are obtained immediately before FK228 administration, and at 4 hours after the start of FK228 administration, to evaluate the ability of FK228 to delay cardiac repolarization. This effect is manifested on the ECG as prolongation of the QT interval. The QT interval is transformed into the heartrate independent corrected value known as the QTc interval. Prolongation of the QTc interval is the electrocardiographic finding associated with increased susceptibility to the development of cardiac arrythmias, including ventricular arrhythmias such as Torsade de Pointes. Because measurement of the baseline value is a factor that critically influences the observed variability in the mean QTc interval, values are computed as the mean of
multiple ECGs to enhance the precision of the measurement. This computation is performed by collecting drug-free ECGs on three or more different days. The on-study time point for obtaining an ECG are selected to coincide with the maximum plasma concentration of FK228, as recommended in the preliminary FDA concept paper: The Clinical Evaluation Of Qt/Qtc Interval Prolongation And Proarrhythmic Potential For Non-Antiarrhythmic Drugs (November 15, 2002) available at: http://www.fda.gov/ohrms/dockets/ac/03/briefing/pubs\\prelim.pdf.
[0135] To examine the pharmacokinetic profile of FK228 following its intravenous administration, blood samples are collected following the first administration from a peripheral site contra lateral to the venous access used for drug infusion, and immediately placed in an ice water bath. Samples are obtained before drug administration and at serial time points after the start of drug administration, including at the end of infusion (4 hours), and at $2,7,9,11,14$, and 21 hours after the end of infusion. All samples are centrifuged in a refrigerated centrifuge, and then stored at or below $-20^{\circ} \mathrm{C}$ until the time of analytical analysis. FK228 concentrations in samples from patients treated with FK228 are quantitated by liquid chromatography with single-quadrupole mass spectrometric detection over the concentration range of $0.5 \mathrm{ng} / \mathrm{mL}$ to $100 \mathrm{ng} / \mathrm{mL}$, according to a validated, previously published procedure. Hwang, et al, J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 809:81-6 (2004). The values for precision and percent deviation from nominal (accuracy) are $\leq 7.88 \%$ and $<3.33 \%$, respectively.
[0136] Estimates of pharmacokinetic parameters for FK228 are derived from individual concentration-time data sets using model independent methods as implemented in the computer software program WinNonlin v5.0 (Pharsight Corporation, Mountain View, Calif). The maximum plasma concentration (Cmax) and the time of maximum plasma concentration (Tmax) are the observed values. The area under the concentrationtime curve (AUC) from time zero to the time of the final quantifiable sample (AUC[tf]) is calculated using the log-linear trapezoidal method. In addition, the AUC from time zero to infinity (AUC[inf]) is extrapolated to infinity by dividing the last measured concentration by the terminal rate constant, $\lambda_{z}$, which is determined from the slope of the terminal phase of the concentration-time curve using weighted least-squares as the estimation procedure, and inverse variance of the output error (linear) as the weighting option. In view of the linear pharmacokinetic profile of FK228 within the tested dose
range, see Sandor et al., Br. J. Cancer 83:817-25 (2000), individual values for Cmax and AUC[inf] are normalized to a dose of $14 \mathrm{mg} / \mathrm{m}^{2}$. The terminal half-life $\left(\mathrm{t}_{1 / 2, \mathrm{z}}\right)$ is calculated as 0.693 divided by $\lambda_{z}$. Additional pharmacokinetic parameters include the volume of distribution at steady-state ( $\mathrm{V}_{\mathrm{ss}}$ ) and the systemic clearance (CL), which is calculated as dose divided by AUC[inf], with dose expressed in mg. The clearance is also calculated in units of $\mathrm{L} / \mathrm{h} / \mathrm{m}^{2}$, by dividing CL by each patient's body-surface area (BSA).
[0137] Relationships between various exposure measures, for example, plasma AUC, and hematological and cardiac toxicity are evaluated by sigmoid maximum-effect models. Cardiac functional assessment is evaluated using base-line corrected QTc interval values ( $\triangle \mathrm{QTc}$ ), as described by Sandor et al., Br. J. Cancer 83:817-25 (2000). Hematological pharmacodynamics are evaluated by analysis of the absolute nadir values of platelet counts or the relative thrombocytopenia, that is, the percent decrease in platelet count. Data are fitted to a sigmoid maximum-effect model based on the modified Hill equation, as follows: $\mathrm{E}=\mathrm{E}_{0}+\mathrm{E}_{\max } \mathrm{X}\left[\left(\mathrm{KP}^{\gamma}\right) /\left(\mathrm{KP}^{\gamma}+\mathrm{KP}_{50}{ }^{\gamma}\right)\right]$. In this equation, $\mathrm{E}_{0}$ is the minimum reduction possible, $\mathrm{E}_{\text {max }}$ is the maximum response (fixed to a value of 100 ), KP is the pharmacokinetic parameter of interest, $\mathrm{KP}_{50}$ the value of the pharmacokinetic parameter predicted to result in half of the maximum response, and $\gamma$ is the Hill constant, which describes the sigmoidicity of the curve. Models are evaluated for goodness of fit by minimization of sums of the squared residuals and by reduction of the estimated coefficient of variation for fitted parameters. Significance of the relationships are assessed by construction of contingency tables with subsequent chi-squared analysis.
[0138] Genomic deoxyribonucleic acid (DNA) is extracted from 1 mL of plasma using the QIAamp DNA Blood midi kit (Qiagen Inc, Valencia, CA), following the manufacturers instructions, and is reconstituted in a buffer containing 10 mM Tris ( pH 7.6) and 1 mM EDTA. For analysis of ABCB 1 variants, a $50-\mu \mathrm{L}$ reaction is prepared for polymerase chain reaction (PCR) amplification using the PCR primer combinations listed in Table I. The reaction consists of 1 PCR buffer, 2 mM of each of the four deoxynucleotide triphosphates (dNTPs), 1.5 mM magnesium chloride, and 1 unit of Platinum Taq DNA polymerase from Invitrogen (Carlsbad, Calif). PCR conditions are as follows: $94^{\circ} \mathrm{C}$ for 5 minutes, followed by 40 cycles of $94^{\circ} \mathrm{C}$ for 30 seconds, $68^{\circ} \mathrm{C}$ for 30 seconds, and $72^{\circ} \mathrm{C}$ for 30 seconds, with a final 7 minute cycle at $72^{\circ} \mathrm{C}$. Direct nucleotide sequencing PCR is conducted using the Big Dye Terminator Cycle Sequencing Ready

Reaction kit V1.1 (Applied Biosystems) using the sequencing primers listed in Table I. Sequences are generated on an ABI Prism 310 Genetic Analyzer. Variations in CYP3A4 (CYP3A4*1B) and CYP3A5 (CYP3A5*3C) are also analyzed using direct nucleotide sequencing, as described by Lepper et al., Clin Cancer Res., 11(20):7398-404 (2005). The genotype is called variant if it differed from the Refseq consensus sequence (rs) for the SNP position. Refseqs are available at http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html.
Table I. Primers used for $A B C B 1$ amplification and sequencing.

| $1236 \mathrm{C}>\mathrm{T}$ | F GTTCACTTCAGTTACCCATCTCG | (SEQ ID NO: 25)F GTCAGTTCCTATATCCTGTGTCTG | (SEQ ID NO: 31) |
| :---: | :---: | :---: | :---: |
|  | R TATCCTGTCCATCAACACTGACC | (SEQ ID NO: 26)R TCCTGTCCATCAACACTGACCTG | (SEQ ID NO: 32) |
| $2677 \mathrm{G}>\mathrm{A} / \mathrm{T}$ | F AGGCTATAGGTTCCAGGCTTGC | (SEQ ID NO: 27)F CCCATCATTGCAATAGCAGGAG | (SEQ ID NO: 33) |
|  | R AGAACAGTGTGAAGACAATGGCC | (SEQ ID NO: 28)R GAACAGTGTGAAGACAATGGCCT | (SEQ ID NO: 34) |
| $3435 \mathrm{C}>\mathrm{T}$ | F ATCTCACAGTAACTTGGCAGTTTC | (SEQ ID NO: 29)F GCTGGTCCTGAAGTTGATCTGTG | (SEQ ID NO: 35) |
|  | R AACCCAAACAGGAAGTGTGGCC | (SEQ ID NO: 30)R AAACAGGAAGTGTGGCCAGATGC | (SEQ ID NO: 36) |

[0139] All data are reported as median values with range, unless specified otherwise. Interindividual pharmacokinetic variability is calculated as the coefficient of variation, and expressed as a percentage. Genotype-frequency analysis of Hardy-Weinberg equilibrium is carried out using Clump version 1.9. The linkage between each pair of SNPs is determined in terms of the classical statistic $D^{\prime}$. The absolute value for $D^{\prime}\left(\left|D^{\prime}\right|\right)$ of 1 denotes complete linkage disequilibrium, while a value of 0 denotes complete linkage equilibrium. The effects of the variant genotypes on $\Delta \mathrm{QTc}$, relative thrombocytopenia, dose-normalized AUC, apparent oral clearance, half-life, volume of distribution at steadystate are evaluated statistically with the nonparametric Kruskal-Wallis test. A post-hoc distribution-free multiple comparison procedure is performed using the Dunn test with Bonferroni correction to test pairs of median observations. All statistical analyses are performed using the NCSS software program (version 2001; NCSS, Kaysville, Utah). The a priori level of significance is set at 0.05 .
[0140] FK228 is administered to 42 patients with T-cell lymphoma (17 female, 25 male) as a 4-hour continuous infusion at a dose of $14 \mathrm{mg} / \mathrm{m}^{2}(\mathrm{n}=37)$ or $18 \mathrm{mg} / \mathrm{m}^{2}(\mathrm{n}=5)$. The median age of the patients is 56 years (range, $27-79$ years) and the median BSA is $1.93 \mathrm{~m}^{2}$ (range, $1.43-2.46 \mathrm{~m}^{2}$ ). Thirty-three patients (79\%) are Caucasian, 8 are AfricanAmerican (19\%), and 1 is Hispanic (2\%). Pharmacokinetic data are available from all 42, patients; complete baseline and on-study measurements on blood cell counts and $\Delta \mathrm{QTc}$ from 34 and 29 patients, respectively.
[0141] With the data from all patients combined, the mean ( $\pm$ standard deviation) values for FK228 clearance and terminal half-life are $17.5 \pm 12.7 \mathrm{~L} / \mathrm{h}$ and $7.23 \pm 3.0$ hours, respectively. This is within the range of values observed previously in patients treated with FK228 at doses of $12.7 \mathrm{mg} / \mathrm{m} 2$ and $17.8 \mathrm{mg} / \mathrm{m} 2$ as described in Sandor et al., Br. J. Cancer 83:817-25 (2000). The interindividual variability in drug clearance is relatively high, with a percent coefficient of variation of approximately $72 \%$. Pharmacokinetic parameters of FK228 are not significantly different between men and women $(P>0.12)$. The AUC of FK228 is weakly associated with the percentage decrease in platelet count ( P $<0.001$; Fig 1) using a sigmoid maximum effect model, but not with interindividual $\triangle Q T c$ interval following FK228 treatment $(\mathrm{P}=0.62)$.
[0142] The observed ABCB1, CYP3A4, and CYP3A5 genotype frequencies are in Hardy-Weinberg equilibrium ( $\mathrm{P}>0.13$ ) (Table II). [Cascorbi et al, Clin. Pharmacol. Ther., 69:169-74 (2001); Lamba et al., Adv. Drug Deliv. Rev. 54:1271-94 (2002); Xie et al., Pharmacogenomics 5:243-72 (2004).] Strong linkage is observed between the 3 SNPs in ABCB 1 , with a $\mathrm{D}^{\prime}$ of 0.88 for the $1236 \mathrm{C}>\mathrm{T}$ and $2677 \mathrm{C}>\mathrm{T} / \mathrm{A}$ loci ( $\mathrm{P}<0.001$ ); a $\mathrm{D}^{\prime}$ of 0.66 ( $\mathrm{P}<0.001$ ) for the $1236 \mathrm{C}>\mathrm{T}$ and $3435 \mathrm{C}>\mathrm{T}$ loci; and a $\mathrm{D}^{\prime}$ of 0.65 for the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ loci $(\mathrm{P}<0.001)$. The overall linkage for the three loci is about $57 \%$. The most frequently observed haplotypes in our population are C-G-C ( $44.3 \%$; haplotype 1), T-T-T (31.4\%; haplotype 2), and C-G-T (12.0\%; haplotype 3), although in total 8 different haplotypes are observed.
55
Table II. Genotype and allele frequencies of the studied variants.

| Polymorphism ${ }^{\text {c }}$ | Nomenclature | Effect ${ }^{\text {d }}$ | Genotype frequencies ${ }^{\text {a }}$ |  | Var | Allele frequencies ${ }^{\text {b }}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Wt ${ }^{\text {e }}$ | Het |  | p | q |
| Caucasians |  |  |  |  |  |  |  |
| ABCB1 1236C>T | N/a | G411G | 10 (33.6) | 14 (46.7) | 6 (20.0) | 0.567 | 0.433 |
| $A B C B 12677 \mathrm{G}>\mathrm{T}$ | N/a | A893S | 9 (30.0) | 13 (43.3) | 6 (20.0) | 0.517 | 0.417 |
| $A B C B 12677 \mathrm{G}>\mathrm{A}$ | N/a | A893T | 9 (30.0) | 2 (3.3) | 0 (0) | 0.517 | 0.033 |
| $A B C B 13435 \mathrm{C}>\mathrm{T}$ | N/a | I1145I | 8 (26.7) | 14 (46.7) | 8 (26.7) | 0.500 | 0.500 |
| CYP3A4-392A>G | CYP3A4*1B | Promoter | 25 (78.2) | 3 (9.4) | 4 (12.5) | 0.828 | 0.172 |
| CYP3A5 6986A>G | CYP3A5*3C | Splice variant | 4 (12.5) | 6 (18.8) | 22 (68.8) | 0.219 | 0.781 |
| African Americans |  |  |  |  |  |  |  |
| ABCB1 1236C>T | N/a | G411G | 5 (62.5) | 1 (12.5) | 2 (20.0) | 0.590 | 0.410 |
| $A B C B 12677 \mathrm{G}>\mathrm{T}$ | N/a | A892S | 6 (75.0) | 1 (12.5) | 1 (12.5) | 0.813 | 0.187 |
| $A B C B 12677 \mathrm{G}>\mathrm{A}$ | N/a | A893T | 0 (0) | 0 (0) | 0 (0) | 0.813 | 0.000 |
| $A B C B 13435 \mathrm{C}>\mathrm{T}$ | N/a | I1145I | 1 (12.5) | 4 (50.0) | 3 (37.5) | 0.375 | 0.625 |
| CYP3A4-392A>G | CYP3A4*1B | M445T | 5 (62.5) | 0 (0) | 3 (37.5) | 0.625 | 0.375 |
| CYP3A5 6986A>G | CYP3A5*3C | Splice variant | 2 (25.0) | 1 (12.5) | 5 (62.5) | 0.312 | 0.688 |
| ${ }^{\text {a }}$ Number represent number of patients with percentage in parenthesis; the difference in the total number of patients is due to the fact th yield sequencing data or showed PCR amplification; ${ }^{b}$ Hardy-Weinberg notation for allele frequencies (p, frequency for wild type alle for variant allele); ${ }^{\text {c }}$ Number represents position in nucleotide sequence; ${ }^{d}$ Number represents amino acid codon; ${ }^{e}$ Wt, Homozygous wild Het, Heterozygous patient; Var, Homozygous variant patient. |  |  |  |  |  |  |  |

[0143] A significant association between $\triangle \mathrm{QTc}$ at four hours and ABCB 1 genotype at the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ locus is observed $(\mathrm{P}=.024)$ (Fig 2A). Patients carrying the $2677 \mathrm{~T} / \mathrm{T}$ genotype have a significantly lower $\triangle \mathrm{QTc}$ (median $\triangle \mathrm{QTc},-5 \mathrm{msec}$; range, $-12.5-3.25$ $\mathrm{msec} ; \mathrm{n}=4)$ as compared to those with the $2677 \mathrm{GG}(\Delta \mathrm{QTc}, 18.3 \mathrm{msec}$; range, $-1-22.7$ $\mathrm{msec} ; \mathrm{n}=10$ ), $2677 \mathrm{GT}(\Delta \mathrm{QTc}, 16.5 \mathrm{msec}$; range, $2.75-28.2 \mathrm{msec} ; \mathrm{n}=14$ ) or 2677 GA genotypes ( $\Delta \mathrm{QTc}, 17.8 \mathrm{msec} ; \mathrm{n}=1$ ). A trend for similar observation is noted for the $1236 \mathrm{C}>\mathrm{T}(\mathrm{P}=0.10)$ and $3435 \mathrm{C}>\mathrm{T}$ loci $(\mathrm{P}=0.079)$, although for these SNPs the associations are not statistically significant. Additional analyses indicate that consideration of haplotype 2 in this group of patients does not result in improved associations as compared to the single-phased SNPs $(\mathrm{P}=0.033)$. However, patients homozygous for the $\mathrm{ABCB} 12677 \mathrm{TT} / 3435 \mathrm{TT}$ diplotype ( $\triangle \mathrm{QTc},-5.0 \mathrm{msec}$; range, -12.5 $3.25 ; \mathrm{n}=3)$ have a significantly lower $\triangle \mathrm{QTc}(\mathrm{P}=0.0084)$ compared with carriers of the heterozygote ( $\Delta \mathrm{QTc}, 11.3 \mathrm{msec}$; range, $-7-17.8 \mathrm{msec} ; \mathrm{n}=7$ ) or homozygote diplotype ( $\Delta \mathrm{QTc}, 18.5 \mathrm{msec}$; range, $-1=28.2 \mathrm{msec} ; \mathrm{n}=19$ ) (Fig 2 B ).
[0144] None of the variant ABCB 1 or any of the ABCB 1 haplotypes is significantly associated with the relative hematologic toxicity or FK228 clearance. The CYP3A4*1B and CYP3A5*3C alleles are also not statistically significantly associated with any measure of toxicity or FK228 clearance (Fig 3). Differences in other pharmacokinetic parameters are also not statistically significantly different between the different genotype groups.

## EXAMPLE 2

[0145] This example further demonstrates that individuals with certain polymorphic variants of the ABCB 1 gene, e.g., $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>$ T, encounter fewer heart rhythm irregularities typically induced by FK228 (romidepsin, a cyclic depsipeptide) treatment and that QT and QTc interval prolongation associated with romidepsin treatment is linked to ABCB1 variants. This effect is unrelated to an altered plasma pharmacokinetic profile. Romidepsin is used as a model substrate for ABCB1.
[0146] Data from patients with T-cell lymphoma participating on a phase II clinical trial of romidepsin are initially evaluated (group 1). Eligibility criteria are consistent with those described in Example 1 and patients with evidence of heart disease are excluded from the trial. Toxicities are reported using the NCI Common Toxicity Criteria, version 2.0. The Inclusion Criteria required measurable disease; an age of 18 years or older; an

Eastern Cooperative Oncology Group performance status of 0,1 , or 2 ; and a life expectancy of $>12$ weeks. Eligible laboratory values can include $\mathrm{AGC} \geq 1,000 / \mathrm{AL}$, platelets $\geq 100,000 / \mathrm{AL}$, bilirubin $<1.5 \mathrm{x}$ the institutional upper limit of normal, aspartate aminotransferase $<3 \mathrm{x}$ upper limit of normal, and creatinine $<1.5 \mathrm{x}$ upper limit of normal. Patients with a myocardial infarction within the previous 6 months, a left ventricular ejection fraction (LVEF) below normal ( $<45 \%$ if done by MUGA, or $<50 \%$ if done by echocardiogram or cardiac magnetic resonance imaging), a corrected QT interval of $>500$ milliseconds, unstable angina, or third-degree heart block (unless with pacemaker) are excluded. Patients can be premedicated with ondansetron.
[0147] Confirmatory analysis (group 2) utilizes data from two sources: a) patients participating on the same multi-institutional trial as the initial analysis, but who are treated at institutions other than the NCI ; and b) patients treated on the single-agent Phase I clinical trial of romidepsin previously conducted at the National Cancer Institute [Sandor et al., Clin Cancer Res 8:718-28 (2002)]. The common eligibility criteria are as described above for group 1, except that patients with malignancies other than T-cell lymphoma are also eligible.
[0148] Electrocardiograms (ECGs) are obtained immediately before romidepsin administration, and at 4 hours after the start of romidepsin administration (at the end of infusion and within 1 hour thereafter). Electrocardiograms can be obtained using an HP Pagewriter XLi or a GE Marquette MAC1200 and recorded at $25 \mathrm{~mm} / \mathrm{s}$, with an amplitude of $10 \mathrm{~mm} / \mathrm{mV}$ and with $60-\mathrm{Hz}$ filtering. They can be analyzed using Pagewriter A. 04.01 electrocardiogram analysis software (Philips Medical Systems, Andover, MA). The QT interval measurement in this program can be made by averaging the five longest QT intervals with a T or $\mathrm{T}^{\prime}$ wave amplitude of $>0.15 \mathrm{mV}$. The heart rate-corrected QT interval (QTc), indicating repolarization time, is calculated using Bazett's formula (QT divided by the square root of the preceding R-R interval) using the electrocardiogram machine software. QTc as calculated by Friderica's formula is the QT divided by the cubed root of the preceding R-R interval. QTc intervals of 480 ms or greater are independently reviewed by a cardiologist. Because measurement of the baseline value is a factor that critically influences the observed variability in the mean QTc interval, the initial analysis utilized baseline values that are computed as the mean of multiple ECGs to enhance the precision of the measurement. The on-study time point for obtaining an ECG
is selected to coincide with the maximum plasma concentration of romidepsin, and multiple baseline ECGs are measured as recommended by the official guidelines of the FDA [Guidance for Industry E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhytmic Potential for Non-Antiarrhythmic Drugs; U.S. Department of Health and Human Services: Food and Drug Administration: Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) (October 2005), available at http://www.fda.gov/cber/gdlns/iche14qtc.pdf]. Confirmatory analysis utilizes the same design, but with only a single baseline ECG measurement obtained prior to administration of romidepsin as is conducted in most clinics. A clinical scoring system is also utilized wherein ECG abnormalities following romidepsin treatment are graded. A score of 0 indicates no change in the ECG wave, a score of 1 indicates T-wave flattening, and a score of 2 indicates ST segment depression of 2 mm or greater. Accordingly, grade 1 toxicity can be defined as nonspecific T-wave abnormalities (flattening or inversion without ST segment abnormalities), and grade 2 can be defined as ST segment depression of at least 1 mm in at least two leads. If both are observed, then the ECG is assigned a grade 2 toxicity.
[0149] Blood samples are obtained before drug administration, at the end of infusion ( 4 hours), and at $2,7,9,11,14$, and 21 hours after the end of infusion. All samples are immediately centrifuged, and then stored at or below $-20^{\circ} \mathrm{C}$ until analysis. Romidepsin concentrations in plasma samples are determined by a validated method based on liquid chromatography with single-quadrupole mass spectrometric detection [Hwang et al., J Chromatogr. B Analyt. Technol. Biomed. Life Sci., 809:81-6 (2004)]. Pharmacokinetic parameters for romidepsin are derived using non-compartmental analysis using WinNonlin v5.0 (Pharsight Corporation, Mountain View, Calif). Since romidepsin delineates a linear pharmacokinetic profile within the tested dose range [Sandor et al., Clin. Cancer Res., 8:718-28 (2002)], individual values for peak concentration (Cmax) and $\mathrm{AUC}_{[\text {inf }]}$ are normalized to a dose of $14 \mathrm{mg} / \mathrm{m}^{2}$ in order to eliminate drug dose as a variable affecting the parameter estimates.
[0150] Genomic deoxyribonucleic acid (DNA) is extracted from 1 mL of plasma using the QIAamp DNA Blood midi kit (Qiagen Inc, Valencia, Calif), following the manufacturers instructions, and is reconstituted in a buffer containing 10 mM Tris ( pH 7.6) and 1 mM EDTA. Variants in the ABCB 1 and CYP3A5 genes are analyzed as
described in Example 1. The reference genotype is defined as the Refseq consensus sequence for the SNP position, and allelic variants are those differing from the consensus sequence. Genotype-frequency analysis of Hardy-Weinberg equilibrium and inference of haplotypes is conducted using Helix Tree Software v4.4.1 (Golden Helix Inc., Montana). The linkage between each pair of SNPs is determined in terms of the classical statistic D'. [0151] All data are reported as median values with range, unless specified otherwise. Changes in QTc interval from baseline ( $\Delta \mathrm{QTc}$ ) as well as drug clearance are evaluated with respect to the presence of a trend in the association of these parameters according to the number of reference alleles in individual variant genotypes using the JonckheereTerpstra trend test. [Hollander et al., Nonparametric Statistical Methods, Second Edition. New York, John Wiley and Sons, Inc., (1999)]. Because of limited numbers of observations, subsequent analyses are based on grouping patients on the basis of the number of reference alleles in multiple loci, with these resulting two group statistical comparisons being evaluated using an exact Wilcoxon rank sum test, with a standard Bonferroni adjustment used for multiple comparisons in these evaluations. The simultaneous effects of genetic variants and clearance on $\triangle Q T c$ are evaluated using a regression analysis using a backward selection algorithm, and should be interpreted as an exploratory finding because of limited power. Again, because of relatively limited amounts of data for analysis, comparisons between the distribution of clinical toxicity scores vs. categorized genotypes are performed using Mehta's modification to Fisher's exact test [Mehta et al., J. Am. Stat. Assoc., 78:427-34 (1983)].
[0152] The characteristics of all patients are reported in Table III. In the initial analysis ("group 1"), romidepsin is administered to 45 patients ( 42 patients as in Example 1 and 3 additional patients) with T-cell lymphoma. In the confirmatory analysis ("group $2^{\prime \prime}$ ), romidepsin is administered to 29 patients. The 17 patients with T-cell lymphoma receive the same therapeutic regimen as the original 45 patients in group 1 , while the remaining 12 patients receive FK288 at a dose of either $12.7 \mathrm{mg} / \mathrm{m} 2(\mathrm{~N}=3), 17.8 \mathrm{mg} / \mathrm{m} 2$ ( $\mathrm{N}=7$ ), or $24.3 \mathrm{mg} / \mathrm{m} 2(\mathrm{~N}=2$; on a day 1 and 5 schedule). Pharmacokinetic data are available in all patients in both groups.

Table III. Patient Demographics and Dosages

| Parameter $^{\mathrm{a}}$ | Group 1 <br> $(N=45)$ | Group 2 <br> $(N=29)$ |
| :--- | :--- | :--- |
|  |  |  |
| Age $^{\mathrm{b}}$ | $56(27-79)$ | $63(40-77)$ |
| Male/Female | $28 / 17$ | $18 / 11$ |

Race:

| Caucasian | $34(76 \%)$ | $28(97 \%)$ |
| :--- | :--- | :--- |
| African American | $9(20 \%)$ | $1(3 \%)$ |
| Hispanic | $1(2 \%)$ | 0 |
| Unknown | $1(2 \%)$ | 0 |

Dose:

| $12.7 \mathrm{mg} / \mathrm{m}^{2}$ | 0 | 3 |
| :--- | :--- | :--- |
| $14.0 \mathrm{mg} / \mathrm{m}^{2}$ | 41 | 17 |
| $17.8 \mathrm{mg} / \mathrm{m}^{2}$ | 0 | 7 |
| $18.0 \mathrm{mg} / \mathrm{m}^{2}$ | 4 | 0 |
| $24.3 \mathrm{mg} / \mathrm{m}^{2}$ | 0 | 2 |

[^0][0153] A summary of the pharmacokinetic parameter estimates is reported in Table IV. The observed values for romidepsin clearance are within the range observed previously in patients treated with romidepsin at doses of $12.7 \mathrm{mg} / \mathrm{m}^{2}$ and $17.8 \mathrm{mg} / \mathrm{m}^{2}$. [Sandor et al., Clin. Cancer Res., 8:718-28 (2002)] The interindividual variability in drug clearance is relatively high, with a percent coefficient of variation of approximately $72 \%$. Pharmacokinetic parameters of romidepsin are not statistically significantly different between men and women (all P > .10).

Table IV. Summary of plasma pharmacokinetic parameter estimates

|  | Group 1 <br> $(N=45)$ | Group 2 <br> $(N=29)$ | All <br> $(N=74)$ |
| :--- | :--- | :--- | :--- |
| Parameter | $15.1(3.8-70.3)$ | $13.9(2.7-35.8)$ | $14.3(2.7-70.3)$ |
| Clearance (L/h) | $1760(358-6072)$ | $1008(391-5237)$ | $1501(358-6072)$ |
| AUC (ng h/mL) | $501(88.0-1599)$ | $322(113-1213)$ | $431(88.0-1599)$ |
| $\mathrm{C}_{\text {max }}(\mathrm{ng} / \mathrm{mL})$ | $6.8(2.2-15.0)$ | $3.8(1.0-8.8)$ | $6.0(1.0-15.0)$ |
| $\mathrm{T}_{1 / 2}(\mathrm{~h})$ | $129(30.8-621)$ | $64.9(15.0-329)$ | $93.6(15.0-621)$ |

Data are presented as median with range in parenthesis.
Abbreviations: AUC, area under the concentration-time curve extrapolated to infinity normalized to a dose of $14 \mathrm{mg} / \mathrm{m}^{2} ; \mathrm{C}_{\text {max }}$, peak plasma concentration normalized to a dose of $14 \mathrm{mg} / \mathrm{m}^{2} ; \mathrm{T}_{1 / 2}$, half-life of the terminal phase; Vss, volume of distribution at steady-state.
[0154] For the Caucasian population, the observed ABCB 1 and CYP3A5 genotype frequencies are in Hardy-Weinberg equilibrium ( $\mathrm{P}>.15$ ) (Table V). Strong linkage is observed between the 3 SNPs in ABCB1 in the Group 1 cohort, with a linkage statistic (D') value of 0.90 for the $1236 \mathrm{C}>\mathrm{T}$ and $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ loci $(\mathrm{P}<.001)$; a $\mathrm{D}^{\prime}$ of 0.56 ( $\mathrm{P}<$ .001) for the $1236 \mathrm{C}>\mathrm{T}$ and $3435 \mathrm{C}>\mathrm{T}$ loci; and a $\mathrm{D}^{\prime}$ of 0.68 for the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ loci $(\mathrm{P}<.001)$. The most frequently observed ABCB 1 haplotypes in the Caucasian population are the 1236T-2677T-3435T (T-T-T; 37.0\%; haplotype 1), C-G-C ( $33.6 \%$; haplotype 2), and C-G-T ( $18.0 \%$; haplotype 3 ), although in total 7 different haplotypes are observed. The variant genotypes observed in the African American patients are also in Hardy-Weinberg equilibrium ( $\mathrm{P}>.13$ ) (Table V). Strong linkage is also observed between the 3 SNPs in ABCB1 in the Group 2 cohort, with a D' of 1.0 for the $1236 \mathrm{C}>\mathrm{T}$ and $2677 \mathrm{C}>\mathrm{T} / \mathrm{A}$ loci $(\mathrm{P}=.002)$; a $\mathrm{D}^{\prime}$ of $0.89(\mathrm{P}=.007)$ for the $1236 \mathrm{C}>\mathrm{T}$ and $3435 \mathrm{C}>\mathrm{T}$ loci; and a $\mathrm{D}^{\prime}$ of 1.0 for the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ loci $(\mathrm{P}=.012)$. The
predominant haplotypes observed in the African American population are haplotype 2 (66.1\%), haplotype 1 (33.3\%), and haplotype 3 (5.6\%).
Table V. Genotype and allele frequencies of the studied variants

| Allelic variant ${ }^{\text {c }}$ | Effect ${ }^{\text {d }}$ | Genotype frequencies ${ }^{\text {a }}$ |  |  |  | Allele frequencies ${ }^{\text {b }}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $N^{e}$ | Ref' | Het | Var | $p$ | $q$ |
| Caucasians $(\mathrm{N}=62)^{\text {g }}$ |  |  |  |  |  |  |  |
| $A B C B 11236 \mathrm{C}>\mathrm{T}$ | G411G | 55 | 19 (34.5) | 22 (40.0) | 14 (25.5) | 0.545 | 0.455 |
| $A B C B 12677 \mathrm{G}>\mathrm{T}$ | A893S | 54 | 15 (27.8) | 22 (40.7) | 15 (27.8) | 0.481 | 0.481 |
| $A B C B 12677 \mathrm{G}>\mathrm{A}$ | A893T ${ }^{\text {h }}$ | 54 | 15 (27.8) | 2 (3.7) | 0 (0) | 0.481 | 0.019 |
| $A B C B 13435 \mathrm{C}>\mathrm{T}$ | I1145I | 62 | 13 (21.0) | 28 (45.2) | 21 (33.9) | 0.435 | 0.565 |
| CYP3A5 6986A $>\mathrm{G}^{\mathrm{i}}$ | Splice variant | 55 | 1 (1.8) | 9 (16.4) | 45 (81.8) | 0.100 | 0.900 |
| African Americans ( $\mathrm{N}=10$ ) |  |  |  |  |  |  |  |
| ABCB1 1236C>T | G411G | 9 | 5 (55.6) | 1 (11.1) | 3 (3.33) | 0.611 | 0.389 |
| $A B C B 12677 \mathrm{G}>\mathrm{T}$ | A893S | 9 | 6 (66.7) | 1 (11.1) | 2 (22.2) | 0.722 | 0.278 |
| $A B C B 12677 \mathrm{G}>\mathrm{A}$ | A893T | 9 | 0 (0) | 0 (0) | 0 (0) | 0.722 | 0.000 |
| ABCB1 3435C>T | I1145I | 10 | 5 (50.0) | 1 (10.0) | 4 (40.0) | 0.550 | 0.450 |
| CYP3A5 6986A>G ${ }^{\text {i }}$ | Splice variant | 8 | 5 (62.5) | 2 (25.0) | 1 (12.5) | 0.750 | 0.250 |

${ }^{\text {a }}$ Number represent number of patients with percentage in parenthesis; the difference in the total number of patients is due to the fact that not all samples yielded sequencing data or showed PCR amplification; ${ }^{b}$ Hardy-Weinberg notation for allele frequencies ( $p$, frequency for wild type allele and $q$, frequency for variant allele); ${ }^{c}$ Number represents position in nucleotide sequence; ${ }^{\text {d }}$ Number represents amino acid codon; ${ }^{\text {e }}$ genotype data are not available in all patients as not all samples yield sufficient DNA or PCR amplified; ${ }^{f}$ Ref, Homozygous reference allele patient; Het, Heterozygous patient; Var, Homozygous variant patient; ${ }^{\mathrm{g}}$ A single Hispanic male is also included, and his genotype is $1236 \mathrm{C}>\mathrm{T}$, unknown; $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$, wild-type; $3435 \mathrm{C}>\mathrm{T}$, wild-type; ${ }^{\text {h }}$ The $2677 \mathrm{G}>$ T/A polymorphism is triallelic and two different SNPs are therefore presented; ${ }^{1}$ The CYP3A5 6986A>G transition is also known as the CYP3A5*3C polymorphism.
[0155] There is no association between the dosage of romidepsin and the $\triangle Q T c$ in either group 1 ( $\mathrm{P}=0.38$ by Wilcoxon rank sum test comparing two dose levels), or in group 2 ( $\mathrm{P}=0.30$ by Wilcoxon rank sum test comparing doses up through $14 \mathrm{mg} / \mathrm{m} 2$, $\mathrm{n}=18$, vs. doses of $17.8 \mathrm{mg} / \mathrm{m} 2$ and $24.9 \mathrm{mg} / \mathrm{m} 2, \mathrm{n}=7$ ); thus, comparisons between genotype and $\triangle Q T c$ are therefore made by grouping patients receiving different doses. In group 1, a significant trend toward increasing $\triangle Q T c$ (i.e. the difference between pre- and post-treatment QT intervals at 4 hours) and increasing number of reference alleles of the ABCB 1 genotype at the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ loci is observed ( $\mathrm{P}=.011$; Fig 4 A ). Patients carrying a copy number of 0 reference alleles (i.e. "wild-type" alleles) at both loci have a significantly shorter $\Delta \mathrm{QTc}$ (median $\Delta \mathrm{QTc},-1 \mathrm{msec}$; range, -12.5 to $+21.6 \mathrm{msec} ; \mathrm{N}$ $=4)$ as compared to those patients with only a single reference allele at either locus ( $\Delta \mathrm{QTc}, 9.7 \mathrm{msec}$; range, -7.3 to $+38.8 \mathrm{msec} ; \mathrm{N}=6$ ), or two or more reference allele copy numbers ( $\Delta \mathrm{QTc}, 18.5 \mathrm{msec}$; range, -1.0 to $+39.5 \mathrm{msec} ; \mathrm{N}=28$ ). A similar, although weaker, trend is noted for the association of reference alleles of $\mathrm{ABCB} 13435 \mathrm{C}>\mathrm{T}$ locus and $\Delta \mathrm{QTc}$ when it is considered separately $(\mathrm{P}=0.15$; Fig 5 A$)$. Additionally, patients carrying the 3435 TT variant genotype have a higher median $\triangle \mathrm{QTc}$ than patients carrying the 2677 TT genotype suggesting that 2677 alleles have a greater impact on the association with $\Delta \mathrm{QTc}$. When the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ allele is considered independently of the others with respect to its association with $\triangle \mathrm{QTc}$, a significant relationship is observed ( $\mathrm{P}=$ .0046, after adjustment for multiple comparisons). Those patients carrying no reference alleles at the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ locus have a significantly shorter $\triangle \mathrm{QTc}$ (median $\triangle \mathrm{QTc}$, -2.0 msec ; range, -12.5 to $+21.6 \mathrm{msec} ; \mathrm{N}=6$ ) compared to patients carrying one or more reference alleles (median $\Delta \mathrm{QTc}, 18.2 \mathrm{msec}$; range, -1.0 to $+39.5 \mathrm{msec} ; \mathrm{N}=32$ ) (Fig 6A).
[0156] Similar trends are noted in group 2 , wherein those patients carrying either 0 or 1 reference alleles at both the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ loci trend towards a smaller $\Delta \mathrm{QTc}$ than those with 2-4 reference alleles $(\mathrm{P}=.07$; Fig 4 B$)$. When the ABCB 1 $3435 \mathrm{C}>\mathrm{T}$ allele is considered alone in association with $\Delta \mathrm{QTc}$ in group 2, a statistically significant trend is noted whereby those patients carrying fewer copy numbers of the reference allele have a smaller $\Delta \mathrm{QTc}$ after treatment with romidepsin ( $\mathrm{P}=.028$; Fig 5 B ). Similar results are also observed with patients carrying either 0 or 1 reference alleles at the $A B C B 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ locus; these individuals have a statistically significant smaller $\triangle \mathrm{QTc}$ $(\mathrm{P}=.015$, after adjustment for multiple comparisons; Fig 6B). Those patients carrying 0
or 1 reference alleles at $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ have a significantly smaller $\triangle \mathrm{QTc}$ (median $\Delta \mathrm{QTc}, 4 \mathrm{msec}$; range -5 to $+21 \mathrm{msec} ; \mathrm{N}=14$ ) as compared to patients carrying more than 1 reference allele (median $\Delta Q T c, 24.5 \mathrm{msec}$; range 17 to $+30 \mathrm{msec} ; N=4$ ). Neither analysis includes the $\mathrm{ABCB} 11236 \mathrm{C}>\mathrm{T}$ transition as this SNP is in very strong linkage with the $2677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ transition, and there is no evidence that the $1236 \mathrm{C}>\mathrm{T}$ is involved in differential $A B C B 1$ expression in heart tissue.
[0157] Neither the T-wave flattening nor the ST segment depression is associated with ABCB 1 allelic variation based on the clinical scoring system utilized in this study. Based upon results from a generalized Fisher's exact test, the ABCB1 2677G $>$ T/A allele is not associated with the scores obtained at baseline ( $\mathrm{P}=0.46$ for group 1 ; all scored 0 for group 2), or at 4-hours post treatment in either Groups $1(p=0.86)$ or $2(p=0.18)$. Similar results at pre-treatment $(\mathrm{P}=0.086$ for group $1 ; \mathrm{P}=1.00$ for group 2 ), or 4 -hours $(\mathrm{P}=0.45$ for group $1 ; \mathrm{P}=0.47$ for group 2) post treatment are observed with the $\mathrm{ABCB} 13435 \mathrm{C}>\mathrm{T}$ polymorphism. When the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ and $3435 \mathrm{C}>\mathrm{T}$ polymorphisms are considered in combination, the pre-treatment $(\mathrm{P}=0.067$ for group 1 ; all score zero in group 2) toxicity score is marginally associated in group 1 , while the post-treatment value at 4hours (Group 1, $\mathrm{P}=0.10$; Group 2, $\mathrm{P}=0.024$ ) post treatment is found to be associated with the ECG abnormality score in Group 2.
[0158] None of the variant ABCB1 SNPs, or combinations thereof is significantly associated with romidepsin clearance $(\mathrm{P}=0.51$ for Group 1 and $\mathrm{P}=0.46$ for Group 2; Figs. 7A \& 7B). Based on linear regression modelling using a backward selection algorithm, the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ reference allele copy number is the sole parameter remaining in the model, and found to be a potentially important parameter in the determination of $\Delta \mathrm{QTc}(\mathrm{P}=0.0004$ by t -test for whether parameter estimate is equal to zero). Systemic drug clearance is eliminated as a parameter for consideration in the model, with $\mathrm{P}>0.25$ after adjusting for the $\mathrm{ABCB} 12677 \mathrm{G}>\mathrm{T} / \mathrm{A}$ reference allele copy number. The CYP3A5*3C allele is also not statistically significantly associated with any measure of toxicity or romidepsin clearance ( $\mathrm{P}>.05$ ). Differences in other pharmacokinetic parameters are also not statistically significantly different between the different genotype groups.
[0159] The use of the terms " $a$ " and "an" and "the" and similar referents in the context of describing the invention (especially in the context of the following claims) are
to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms "comprising," "having," "including," and "containing" are to be construed as open-ended terms (i.e., meaning "including, but not limited to,") unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.
[0160] Preferred embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those preferred embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

CLAIM(S):

1. A method of screening for an altered susceptibility for a drug-induced heart rhythm irregularity, the method comprising:
(a) screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the $A B C B 1$ gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB 1 gene, and wherein the polymorphism comprises a polymorphism at position 49,910, 68,894, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2 ; or a combination thereof; and
(b) diagnosing the altered susceptibility of the subject for the heart rhythm irregularity as induced by the drug based on the presence or absence of the polymorphic variant of the ABCB 1 gene.
2. The method of claim 1 , wherein the drug is an anti-cancer agent.
3. The method of claim 1 or 2, wherein the drug is FK228, a prodrug thereof, a salt thereof, or a combination thereof.
4. The method of claim 3, wherein the drug is FR901228, a prodrug thereof, a salt thereof, or a combination thereof.
5. The method of any one of claims 1-4, wherein the polymorphic variant is associated with an increase or decrease in the expression of the ABCB1 gene.
6. The method of any one of claims $1-4$, wherein the polymorphic variant is associated with an increase or decrease in an activity of a protein encoded by the ABCB1 gene.
7. The method of any one of claims 1-4, wherein the polymorphic variant is associated with an increased susceptibility for a drug-induced heart rhythm irregularity.
8. The method of any one of claims 1-4, wherein the polymorphic variant is associated with a decreased susceptibility for a drug-induced heart rhythm irregularity.
9. The method of any one of claims $1-4$, wherein the method further comprises prescribing a treatment regimen based on the diagnosis.
10. The method of claim 9 , wherein the treatment regimen comprises increasing dosage of the drug in the presence of a polymorphic variant associated with a decreased susceptibility for the heart rhythm irregularity.
11. The method of claim 9 , wherein the treatment regimen comprises decreasing dosage of the drug in the absence of a polymorphic variant associated with a decreased susceptibility for the heart rhythm irregularity.
12. The method of claim 11, wherein the drug is not administered.
13. The method of claim 12 , wherein a different drug is administered.
14. The method of claim 13, wherein the drug does not bind a protein expressed by the ABCB 1 gene.
15. The method of claim 9 , wherein the treatment regimen comprises increased heart monitoring.
16. The method of claim 9 , wherein a second, additional drug is administered.
17. The method of claim 16 , wherein the second drug ameliorates the heart rhythm irregularity.
18. The method of any one of claims 1-4, wherein the subject has previously experienced a heart rhythm irregularity.
19. The method of any one of claims 1-4, wherein the heart rhythm irregularity is a cardiac arrhythmia.
20. The method of any one of claims 1-4, wherein the heart rhythm irregularity comprises at least one member selected from the group consisting of asymptomatic dysrhythmias and ventricular arrthymias.
21. The method of any one of claims $1-4$, wherein the heart rhythm irregularity is characterized by at least one of ST/T wave flattening, torsade de pointes, and QT interval prolongation.
22. The method of any one of claims 1-4, wherein the sample comprises blood.
23. The method of any one of claims 1-4, wherein the polymorphism variant is a single nucleotide polymorphism (SNP).
24. The method of any one of claims 1-4, wherein the polymorphic variant is present in a single chromosomal copy of the gene, and wherein heterozygosity is associated with an altered susceptibility for the heart rhythm irregularity.
25. The method of claim 24 , wherein heterozygosity for polymorphic variants of two or more polymorphisms is associated with an altered susceptibility for the heart rhythm irregularity.
26. The method of any one of claims 1-4, wherein the polymorphic variant is present in both chromosomal copies of the gene, wherein homozygosity of the polymorphic variant is associated with an altered susceptibility for the heart rhythm irregularity if homozygosity of the polymorphic variant is detected.
27. The method of claim 26, wherein homozygosity for polymorphic variants of two or more polymorphisms is associated with an altered susceptibility for the heart rhythm irregularity.
28. The method of claim 1, wherein the sample comprises a nucleic acid selected from the group consisting of (a) a nucleic acid encoding ABCB , (b) a fragment of (a) comprising at least 20 contiguous nucleotides of (a) wherein the 20 contiguous nucleotides comprise the polymorphism, (c) a complement of (a) or (b), and (d) a combination of two or more of (a), (b), and (c).
29. The method of any one of claim 28 , wherein the nucleic acid encoding ABCB1 comprises SEQ ID NOS: 1,2 , or a combination thereof.
30. The method of claim 28 , wherein the polymorphism is a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2 ; or a combination thereof.
31. The method of claim 28 , wherein the polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1 ; or 1236 of SEQ ID NO: 2 , or a combination thereof.
32. The method of claim 31, wherein the nucleic acid comprises the sequence of SEQ ID NOS: 3,9 , or a combination thereof.
33. The method of claim 28 , wherein the polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1 , or 2677 of SEQ ID NO: 2 , or a combination thereof.
34. The method of claim 33, wherein the nucleic acid comprises the sequence of SEQ ID NOS: 4,10 , or a combination thereof.
35. The method of claim 28, wherein the polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof.
36. The method of claim 35, wherein the nucleic acid comprises the sequence of SEQ ID NOS: 5,11 , or a combination thereof.
37. The method of claim 28, wherein the nucleic acid comprises first and second polymorphisms wherein the first polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1; or 1236 of SEQ ID NO: 2, or a combination thereof and the second
polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1 , or 2677 of SEQ ID NO: 2 , or a combination thereof.
38. The method of claim 37 , wherein the nucleic acid comprises the sequence of SEQ ID NO: 6,12 , or a combination thereof.
39. The method of claim 28 , wherein the nucleic acid comprises first and second polymorphisms wherein the first polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1,2677 , of SEQ ID NO: 2 , or a combination thereof the second polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof.
40. The method of claim 39 , wherein the nucleic acid comprises the sequence of SEQ ID NOS: 7, 13, or a combination thereof.
41. The method of claim 28 , wherein the nucleic acid comprises first, second and third polymorphisms wherein the first polymorphism is a polymorphism at position 49,910 of SEQ ID NO: 1; or 1236 of SEQ ID NO: 2 , or a combination thereof, the second polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1 , or 2677 of SEQ ID NO: 2 , or a combination thereof, and the third polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof.
42. The method of claim 39, wherein the nucleic acid comprises the sequence of SEQ ID NOS: 8,14 , or a combination thereof.
43. The method of any one of claims 28-42, wherein the polymorphic variant is a thymine at at least one polymorphism.
44. The method of any one of claims 28-42, wherein the polymorphism comprises a polymorphism at position 68,894 of SEQ ID NO: 1, or 2677 of SEQ ID NO: 2 , or a combination thereof and the subject is homozygous for thymine at that position.
45. The method of any one of claims 28-42, wherein the polymorphism comprises first, second, and third polymorphisms wherein the first polymorphism is a polymorphism at position 68,894 of SEQ ID NO: 1,2677 , of SEQ ID NO: 2 , or a combination thereof the second polymorphism is 2677 , and the third polymorphism is a polymorphism at position 90,871 of SEQ ID NO: 1,3435 of SEQ ID NO: 2 , or a combination thereof, and wherein the subject is homozygous for thymine at both positions.
46. The method of any one of claims $1-4$, wherein the sample comprises genomic DNA, cDNA, mRNA, a fragment thereof, or a combination thereof.
47. The method of any one of claims $1-4$, wherein the sample is screened using a nucleic acid array.
48. The method of any one of claims $1-4$, wherein the sample is screened using allele-specific-oligonucleotide (ASO) hybridization.
49. The method of any one of claims $1-4$, wherein the sample is screened using PCRRFLP analysis.
50. The method of any one of claims 1-4, wherein the sample is screened using PCR.
51. The method of any one of claims $1-4$, wherein the sample is screened using a single-strand conformation polymorphic variant (SSCP) technique.
52. The method of any one of claims 1-4, wherein the sample is screened using an amplification refractory mutation system (ARMS) technique.
53. The method of any one of claims 1-4, wherein the sample is screened using nucleotide sequencing.
54. The method of any one of claims $1-4$, wherein the sample is screened using an antibody specific to a polypeptide encoded by the polymorphic variant containing gene.
55. The method of any one of claims 1-4, wherein the sample is screened using mass spectrometry.
56. A kit comprising:
(a) a nucleic acid for use in screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB1 gene, wherein the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236, 2677, or 3435 of SEQ ID NO: 2 or a combination thereof, and wherein the nucleic acid specifically binds to ABCB 1 sequence comprising the at least one polymorphism or a sequence adjacent to ABCB 1 sequence comprising the at least one polymorphism.
(b) a drug that binds a protein encoded by ABCB 1.
57. The kit of claim 56, wherein the drug is FK228.
58. The kit of claim 57, wherein the drug is FR901228.
59. The kit of any one of claims 56-58, wherein the nucleic acid comprises the nucleotide sequence of any one of SEQ ID NOS: 25-36 or a complement thereof or a combination thereof.
60. Use of a drug that binds a protein encoded by the ABCB 1 gene to manufacture a medicament to treat a subject that that has been screened for the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB 1 gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by the drug, and wherein the polymorphism comprises a polymorphism at position $49,910,68,894$, or 90,871 of SEQ ID NO: 1 ; or 1236,2677 , or 3435 of SEQ ID NO: 2, or a combination thereof.
61. A method of screening for a decreased susceptibility for FK228-induced QTc interval prolongation, the method comprising:
(a) screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the ABCB 1 gene, wherein the polymorphic variant is associated with a decreased susceptibility for QTc interval prolongation induced by FK228, and wherein the polymorphic variant comprises a thymine at position 2677 of SEQ ID NO: 2, or a thymine at position 3435 of SEQ ID NO: 2 , or a combination thereof; and
(b) diagnosing decreased susceptibility of the subject for QTc interval prolongation as induced by FK228 based on the presence or absence of the polymorphic variant of the ABCB 1 gene.
62. A method of screening for an altered susceptibility for a drug-induced heart rhythm irregularity, the method comprising:
(a) screening a sample from a subject to detect the presence or absence of at least one polymorphic variant of at least one polymorphism of the $A B C B 1$ gene, wherein the polymorphic variant is associated with an altered susceptibility for a heart rhythm irregularity induced by a drug that binds a protein encoded by the ABCB 1 gene, and wherein the polymorphism comprises a polymorphism identified as rs1128503, rs2032582, rs1045642, or a combination thereof; and
(b) diagnosing the altered susceptibility of the subject for the heart rhythm irregularity as induced by the drug based on the presence or absence of the polymorphic variant of the ABCB 1 gene.

FIG. 1


## 2/8




3/8


FIG. 3C


5/8



6/8






SEQUENCE LISTING

```
<110>
    GOVERNMENT OF THE UNITED STATES OF AMERICA, REPRESENTED BY
        THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES
        Figg, william D.
        Sparreboom, Alexendar
        Sissung, Tristan M.
        Piekarz, Richard L.
        Bates, Susan E.
<120> MATERIALS AND METHODS FOR ABCB1 POLYMORPHIC VARIANT SCREENING,
        DIAGNOSIS, AND TREATMENT
<130> 701083
<150> US 60/736,083
<151> 2005-11-10
<160> 39
<170> PatentIn version 3.3
<210> 1
<211> 95957
<212> DNA
<213> Homo sapiens
<220>
<221> misc feature
<222> (49910)..(49910)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (68894)..(68894)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (90871)..(90871)
<223> n may be any nucleotide
<400> 1
ggtcgggatg gatcttgaag gggaccgcaa tggaggagca aagaagaaga acttttttaa 60
actgaacaat aaaaggtaac tagcttgttt cattttcata gtttacatag ttgcgagatt 120
tgagtaattt atttctagcc tccagctctg aaataaatga catgttgttg tttttaatta 180
tttttaagaa acgcaagcta gcctttggaa tcaatatccc tgcttagagc agaagtttgt 240
tggctgagtg gagcacagca tatgcatttt ccctgtcttt tttgttcttt cttttaatga 300
tacataatat tttacatatt tatgaaatgg ggtacatgga agcgtttttt acatgcccgg 360
aatgtgtaat gatcaagtcc gggtatttga aggatacatc accttaggta tatttcattt 420
ctatgtgttg ataacatttt aagtcttcta gctactttga aatatacaat atattgctaa 480
ctgtagtcac cctcgtctgc tatcgaacat tggaacttat ttgtcctatc caaccgttct 540
tagtcattca ccaacctctt ttcatttcac ctttttaccc ttcccggcct ctttccctta 600
gtcttggtgt gcctctttct cagctttcct gccccagaca ggcggatgct catatgtgtt 660
tctgtcttat gaacttctgc ttttcaagtg gtgttggtcg cccacacgtg agccatatgc 720
```

tgctggtgat ctgctctgtg gtccaggctc ttgcttccgg taaatggcta tgtaaacatc
gcgtttgtgg cctggctgat gagacagaag gtcaaaagta catttaggtt gttaactggcaataaatatc tgtatataat attggtaatg taatcatata gggaaaataa ttatttaaagtaaattttga tcatggtgct ctgcctttat agaatattta aaacttcact aaatagattcattgttagta gtaaattgta aaatagacta gtaagtttaa taatattaga aactgtaatgtaaattataa gataaattag taaacacatt aatattataa gaaaccaagc ttttcagtgtaagagaaaaa atacaaatgt ggaaatcaaa tacattttta aaaataatgt taagtttgaattagaaattt caatatgaat tcataatctt ttaatagttc attttcagtc cactgaaagggacagtaaca atgagcactg ttagtaccag atctgggttt ctaaatacca ttcctccctaaaagaaatga acccctcagt agctaatttc agccaggtct gggacaggaa aaatagaaagtgagcctgga gtatcttgtg gtacctgaaa ataaggaagt ggagtttaat ggagtagtcgaaagcacact gaagctgtct ggaagatgct tccaatggcc aactctggga caatttgagcatcaaagaac atcgtaactg taactgattt tgagatactg aatgctactg ctgaatgcttaatattctga gagaagggaa actgaaagaa agtgtgagag gtagagaaaa taagcaatgatcttttaga gaaaaattcc agctatcatt ttggaacttc cattgttata atagattcaggcaaggattg tcaatgaact ctaaactatc agaaaaatgt tgtatctaag aattttcttcagaatgttta ctaattggaa ggagaatatg cttctccaat taaagatttt gcagttagcaccttaaccga gtgatcagac ctggtatctc tcacggtaga atggactgta gtaaatgactcatcacttgg attttgagta aacaacatca cctagcaaat attaatatat ctgcaaaacatttagtctg tatacaatta aacttgtaga ctggggctgt ccaatggata tatcatgtgagccacatttg taatttaaaa ttttctagta gcaacagtaa aaaatgaaaa taaaagataaagtaatttt gtaaatatat tttatagaac tcaatatgtc caaaagattg gattttctacatgtaatcaa aatgattaat aagatagttt acattctctt ttatgctaaa tctcctgtttgatgtgtatt ttacatttat tgcacctccc atttgtattc agtttgatac aggtatacccaagatgtggg acattctgta agacagctgt tcttgttcct ttacaaaagt caatgttaaaaaaataaaag gaggacttct agattaatag agactaaaga gatataacta ccaaacgtagtgtgtgaaca ttacaaaaca aaacaaaaac aatctaaaag ttcaacaagg aaaaaatcctatggaaaaca ttcttgggag atttgaatgt gcattagata ttagccaata tcagaattgtaatcactttt cttagatgtg ataatggtag gagaacatgt aattgcaatt gaaaatttttgtaacaaaaa tgttcatata aaccatgaaa aagactccat agagtggagg gtaggtattccactttttgt ttttaacttt gtatttttga aaacttgata gcttaaagaa ttatgtacttatgatacaaa catttaggta gatttaacta agaaatattg agggtatgtg aagagaaacagtaaagaaaa ttgtatgatc tttcacatat ttctattatt ggttctatca tggcctatagactttttttt tcaaattaga gattatattt taaagtagtt cttattaaat tgtgttcatt
gccattgatt taaatctatt ttaacatggg attgtaaaga agcatatggc tttggtaata ..... 2820
gataatgctg ggataatagt cctatttgga atatgactat tagcaaagat tctttattaa ..... 2880
aatgatgttt gatgaatgac ttgtctttct aagcactgtt ttttgtgcta atggggataa ..... 2940
gtaagttatc tagtaggggt gagagtttgt gtgaaagtgc attttaatgt gatgcgatgc ..... 3000
agtatctcag aaatctgagg ttgcacctta tagtttggtg gtcagggaac agacttggca ..... 3060
tttaaatatt ggcttagtta tagctgtgaa ccttggacaa gttaactctt ctcactcttc ..... 3120
tacgtctcta aaagaacaaa aatgtgtctt cctcatagaa ttgtttatga attgtgtgtg ..... 3180
ggaagtgctt acttagccag tgcctggaac attataactt ctccagagta gcagctgttt ..... 3240
tagagaacaa aaaaataaat aaaaggctta gagctaaaac tcaactattt atggctattt ..... 3300
ttctcctttc atcctggttc cagggatact gaaaccatgc ctttactggg aattgggtgg ..... 3360
gaccaaacct gaagagttgt gtttgtgtat tttatgtctg tctaacatta ctccaaagcc ..... 3420
aaatgggtaa actctggatt tttttcttta gaagtctcct cacctctttt gacctcactt ..... 3480
agtgtaaaga acaaagacaa agatgaattt accttatgaa cttaaaaacc gtgtaaaaaa ..... 3540
taacacaaat cttttctaaa atagtttttc ttttatacct acaaaaatac agatgaggca ..... 3600
gatttgttac ggttggtttg ctttcacatc ctaggtagca gttacacaga gtagaaaatg ..... 3660
ggctacagag ttagattcat ttgtatctgg gctcaaaata atagttgacc ttatgcctga ..... 3720
ggagaacttt ttgacaattt atagaacatc atttctgacc atagctttct agcgtgcatt ..... 3780
tatttcataa tggtccattt aggactccag acattttttt cacaattaga tttgcttata ..... 3840
aggagtaatt taattttctt ctgaggctag ttatgcccag catctaattg ccacttctct ..... 3900
tcactagaaa gggaaggatt aagaataata tctgggtagt ggatggataa aatgtctact ..... 3960
gccatcatta atcaaatgac aattaagggg attcataatt aaggggaagc atgaatgatt ..... 4020
ttctaattag aataaaaaca gggaactctt ttcagtatgt atatttttct aattgcaaag ..... 4080
ggtgtatgta tgtattcatc atagcaaatc aggaaatatc aaaggcacaa ggaagaaaat ..... 4140
atttttagag agattccagc actcagtgat aaccacgtat gtagatttct aagatatatg ..... 4200
taaacatatt taaacattat aaatggggtc atactattat catcctgttt tttaactgta ..... 4260
taatatagac atatttcttc ttcctattag tcattaaata tacattctat gttaatttgg ..... 4320
gtaatttaaa aaatatgtac tttttaaagt tttatgtaat tccattgctg ttttgcagtg ..... 4380
aaaaagataa gaaggaaaag aaaccaactg tcagtgtatt ttcaatggtg agttttgaat ..... 4440
ttattaacta ttcaaaatac ttcggaaatt tgacatctcc ttacatggaa aagagatatt ..... 4500
tcatctgatg taagattttc gtatagggta tgttaatgga gatgcaaaat aaattggttt ..... 4560
gatttagctt attttcagga gaatgctgat cataacttgc tatctatatt actaatctac ..... 4620
agccaccatt cctgacttag atttcaattc ttctaatgaa ttatgttgcc agtgatccat ..... 4680
tacctacaca agcacctctc ccatctgcat agaggaaacc aatcacatag aatgctgtct ..... 4740
ttaatggctc tgaaatttgg caggtttttt aggtcaaagt cagttgctat caataaaagt ..... 4800
tagcaatgga ttatttatta ctatgcaaat tttaggtcat ttaggttagt ccattgcaaa ..... 4860
ttttagttca aaatccatta ctaacaataa aaggtagcaa tggattattt attactatga ..... 4920 ..... 4980 ..... 5040 ..... 5100516052205280
accttgtttc aacatttaca gaaacacaaa agtagaataa aatctaaagc aagtgaggaa ..... 6900
atgtgacaca tagagctggg ctgtaatttc tacacaaaat gaacttctta gagattatac ..... 6960
actcttggtc aggtgctgtg gctcacgcct gtaatcccag aactttggaa ggctaaggtg ..... 7020
ggcagatcac tagagccctg aagttcaaga ccagcctggg caacatgaca aaaccctgtc ..... 7080
tatgtgaaaa gtacaaaaat tagctgggca tggtggcaca tgcctgtagt ctcagctact ..... 7140
cgggaggctg aggtgggagg atcacttgaa cccaggaggt ggagatgcag tgagccaaga ..... 7200
ttgtgctact gcactctagc ctgggtgaca gagtgagaca ctgtctcaaa aaaaaaaaaa ..... 7260
aaagattat atattcttgc aagaattgtg cactgtagat tgaactctga gctttctatt ..... 7320
atgtgtgtct tagggaaaca ttaatttttt tctaaataga taatacctaa caatgttaag ..... 7380
tgtttccctt gtaccaggca tttgttaaat accctatgtt tattgtgtga cataatcttc ..... 7440
atatagggac ataagtacct gcaatttagg aacttgtgtt attcttgtgt tccagatgaa ..... 7500
gaacttgact tagagaagtt aagtaatgca ttcaaaatca cagggtcaca tagctactaa ..... 7560
gcagggtaga ctcagatatg cccagttgct tttaacaact gtaatcattt atgactggtt ..... 7620
aaactgttat ctctctggaa ttgaagagtt ctttattgct gctatctttt gatcttatga ..... 7680
ccatctcatc atcaaagact acgccacctg gctcagtatc ctccctttcc actttggtgc ..... 7740
tccaaccaac ttgggtgact tttatggtta tatggataat cactccttga gcttcacagc ..... 7800
tgcaataatc ataaaaacct acctgtgtct gaatctgttt cctcctcctt ctttccagtt ..... 7860
atgacagagt tttagtctta agactaatca ttcatacttt ggctcctatc tcctctgcct ..... 7920
tctcagggac cttacatcat tattaccact cattctcttc tgtatcttcc atctttccct ..... 7980
ctcaactgga tctgtcccat tggcattgac acatcaaggc actgatgcce ttcatggagc ..... 8040
cacccctgcc agctcccttc acacacaaaa gtggcctttg cgttctgtct tgatttcgcc ..... 8100
attgacttag ttctcaacct attgaagttt ggctcctgtt cctgtctctc tctacattca ..... 8160
gtggacattt taaaattcta tcttgatctc tcagcagcac ttagactttc tcctatattt ..... 8220
tttgtatttg cttctgttat gtcacggttt gctgattttc ctccagctct ttggcttttt ..... 8280
cttgtgctgc cctcgcctgt aaagttaagt tcctcaaccc cttatcacag gagtcagaga ..... 8340
tgatgtggac attcctatag ttccacaag ccattttcaa atttttttta ttctcatgca ..... 8400
aactcctttg tcatttgagt ttcattctgc ataattatac tatcttctat ctattttcgg ..... 8460
tggtatcctc gaagtatacc cccaccccca ttccttgagg actttggcat ttgattgctt ..... 8520
ctacttcatt caccttatgt tcctgttatt cttctgagta ttttggtgac cttacaaata ..... 8580
agtctcaaca ttctctgact gcttcagttc caacaacgac gctccataaa ttacatgagt ..... 8640
accttagtaa attgcacgtg tggacaggat cttgggctta ttgtctgggt gagctttcct ..... 8700
ctatgaaggt ataaacacac acacacgtgt atcagtcttt gaccacaacc tgtgattcct ..... 8760
tccagctctg ccctgcctct ttatcatccc tcttttgaca ttgtgtattc atttcctatt ..... 8820
gctgctgtag gaaaatagca ctcacttagt ggcttaaaac aacacaagtt tattacctta ..... 8880

gtaaggtttt ctgtgtgtat gaatattttc ctcattttgt aaattcagaa actgagactc
10980 aaggttatta agattttttt cccctaaagt ctccttccta ataaactaca gagtgagggc atttgactga ggcttttgtc ctactcagca ttatgttacc actttcttca aatcttctca cccccttctc tctagaaatc aatcttgctt catgtcatta agaaatatga ggttctagca tttataagga acctaaacaa atttacaaag aaaaaacaaa caacccaata aaaagatggg caaaggacat gaatagagac ttcttaaaag acatgtggcc aacaattata taaaaaaagc tcaacatcac tgatcattag agaaatgcaa atcaaaacca ctgtgagaaa ccatttaaca ccagtcagaa tggctattat tacaaagtaa aaaaataaca gatgctggca agattgtgga gaaaaaggaa catttataca ctgttcagcc attgtggaag acagtgtggt gattcctcaa agacctaaag acagaaatac catttgacct attactgggt gtatacccaa aggaatataa atcattgtat tataagaca catgcattgt atgttcattg cagcactatt cacaatagca aagacatgga atcaacctaa atggccatca attgtagact ggataaagaa aatgtggtac acatagagca ttgaacacca tgcagccata aaajgaacg agatcacgtc ctttgcaggg catggctgga gctggaggcc attatcctta gcaaactaat acaggaacag aaaaccaaac accacatgtt ctcacttaaa agtgggagct aaatgatgag aacacatgga cacctagagg ggaacaacac acattggggc ctttcggagg gtggagtgtg ggaggaggga gaggatcagg aaaataact aacgagtact aggcttaata gctgggtgat gaaataatct gtacaacaac cccccatgaa acaagtttac gtatgtaaca aacctgcatt tgtacccatg aacttaaaag ttaaaataaa caaataaaaa ataataaaaa aaaaaacaaa aggaaatatg aggttctgaa ctctttcaat tttctgtctt tattcctgga aaattagtac tcactttaat ctccttcctt ctagtctaag attaagagat gtaattccaa ttttcccagc tcagtgaggg gtgtcttttg atcagtcagc tatttctgtt tccacacttt acaccccaaa cttgagggtc cccagtcttt tttaagcaca ttgtgctctt ttctgtctca gtttactttt ctttttggct taaatgcctt tttttttttt ttttttttga gacagagtct cactctgtcg cccaggctgg agtgcagtgg tgcgatctcg gctcactgca agctccgcct cccaggttca caccattctc ctgcctcagc ctcctgagta gctgggatta caggcgcctg ccaccacatc cggctaattt ctttttttgt attttaatag agacggggtt tcactgtgtt agccaggatg gtattgatct cctgacctcg tgatctgccc gccttggcct cccaaagtgc tgggattaca ggtgtgagcc actgcaccca gctacacttg gtaaactttc aaagctggaa ggaagcatca cctcttttat catatgaagc cftttcacaa aggagggaat gatgattgac tcagttttgt gtctattcta tacactgtac tgtcaaagca tgcagagcat gtattgcata tacatttttc tttaacctct ctgaaggcag ggttatatct ttggcattta tctctggatt ttatctgtca ttgaataggc attcagtaaa tgtctacttt catgatccac ttatttttat tatggccagt gacagttctg gaggtaaaat gtcctgtgtt ggtggtgatg ggagaaattt tttcatatca actatcttct cttcctgtac

11040
11100
11160
11220
11280
11340
11400
11460
11520
11580
11640
11700
11760
11820
11880
11940
12000
12060
12120
12180
12240
12300
12360
12420
12480
12540
12600
12660
12720
12780
12840
12900
12960

atttcctaac agattgaaca tgggtgtaag acgaaaagtc ggggatgatt ctaaagtttt ..... 15060
tagccctaga aactgggagg atgcagttac tatcaacaga aatgagagat ggagaaggag ..... 15120
agtggatttg agcaataaat cccagaaatt tagttttgga cattgttata tgttgggctt ..... 15180
tctggaagta gatgcaggaa tggaaaagtg gggcaaaatt gtgtgtgtgt gtgtgggggg ..... 15240
gtattttagg ataggataaa taatagcttg tagcttgttt tgtgtgttaa tggaaatgat ..... 15300
ccaaaagaaa ggagaaagtg atgtttgaat gaaagatgga agagatttgc tgagtgatat ..... 15360
ccttgagaaa aaatggaatc cagtgctgag gaggaaaaca gtaacaggtg ggaaaagtga ..... 15420
atttgagtgt cagtacccag tggtgggtag atgagaaggt gagaatctgt ggaatttttc ..... 15480
tcctgataac tttagttttc tcagtgaagt aggaagcagg agaattggct gatagagaca ..... 15540
gagaaaacag atgtaacaca gctctcaagg aaagtgggaa agtgaataga ctacagaaat ..... 15600
aatctgggtg taacatacat attctttcct ctctctccct ctctctctct ctttctcgat ..... 15660
tggggaggaa tgggttgata tctctgagtt tcctaggtta cctctggtca agcctagatt ..... 15720
gatttaaact gagctactca aaaacttaca atccagtttg cactcaggtg gtcaaatgtt ..... 15780
gaaaatgcct gtttgataac gtacttgaaa atgtagtgat accacttagc ttacaaggac ..... 15840
ctttcaactt ttgaactttc aagtttctga tgaatgcaat gaattgcttc cccagaaaca ..... 15900
tgcacatgaa tattcatacc aatctcgggt atatacggat cttttgatac ttgtcagtgg ..... 15960
gtcctcagaa ccacttcatg accctaaaaa gtacattgtt taccctgggg tagaactact ..... 16020
tcaattctgt ggatgcaggt aatgagctgg gccagaaaaa tttttgtatg tttgagctat ..... 16080
gtaaaattct tatgctttat tacatccttt cacggcatcc aaagagacct gcagaaagat ..... 16140
gctgcaccac ctctgaactg gtacctttct tttacaccat tcaccagagg aagtcctcaa ..... 16200
tgctttctga attcattagt tcctatactt gctatattgg ccaatattct atggcgttag ..... 16260
ctctcttact gcttcatagt ggaagaatca aatacttcat cactctctgg tggctcagca ..... 16320
ggcctaatgg gtacatctct gtccatccag atcttgctca ctttttaatg acttcccagc ..... 16380
cttcagaact gtgggaaata aatttttatt gtttgtatgt taatacatac agtttatgta ..... 16440
ttttgttata tcatcccaaa tggactaaga cacacagtaa actcacttac ttaatagcat ..... 16500
gcatgagctc ccaaacagat gatgatctag ttgtgaagac tagctccttt ctggataggc ..... 16560
ccatgtgtct gtatgcctat caaaatctcc taggttaaac aggggccaca tcaccacagc ..... 16620
ttgccagctg gtaatttaga tttccatctt catagatatc tagtttctag gtgaggttgt ..... 16680
aattcacatt gcagtccttg ccagtcattc tcattcagct cactcttcta gtcttgagaa ..... 16740
cgggaaagag aaagcttaga gaaatatttg aaagtttact tttgttattc tcttaagagg ..... 16800
tggtatattt agtggttagc gaacatactc tggagttgga ctgcctggct caaaaaccac ..... 16860
acccttacca atttttaatt ttgaggttta gggaatttat tcatacttta gttattaagt ..... 16920
tggataacaa tagaacctac ttctcaggat ggttgtgata tttaagtgga ttcattcatt ..... 16980
gaagaatcac agaacaaagt ctagtcataa gaagtactca acaaatatta gttactgtta ..... 17040
ttatcaacag gctttcatgt tcctcacaaa tgaaattaca taactttcaa tgatgatatg ..... 17100
gttccttata taatttcttc tttacaaaat tctgtgatac tacctcaatt ttcatagttg ..... 17160
ttgaaagttt gggacatacc caaagttact cagtacctat tatatgatgt gaataatcat ..... 17220
ccctttaagt agaatttcct tggttaccag ccatgccttt cagggaaggt aggtatcctc ..... 17280
caactatgac ccctgggcta atttggcctg ctgcctcttt ttgtacaggg tgtgaactaa ..... 17340
gaatggattt tacattttga aatcattgaa aaaaaatcaa aagaagaaat tatattttgt ..... 17400
gacccgtgga aattatgtga aatttaaatt tcagtacctg taaataaagt ttaatgaaat ..... 17460
gtagctacac ctattcattt acattttatg aatgtaagca gcctatggct gcttctatgc ..... 17520
cacaatgaca cagtagagta gttgtacaaa ccgtatggtc tgcagagcct aagatattta ..... 17580
ctatctggcc ctttatagaa aaaggttgcc aatctctgaa gtagggtatt gagtgagggg ..... 17640
ttgttatctt cagtcaatcc atcaattaat ttgtattaga gcattctgtg tgccagtcac ..... 17700
agtacatgcc ctcatcattc ccaactcttg aggagcttag tgtagtggga gggatggaca ..... 17760
agggtgaacc aataaataca gtcgaatttg attagtgata taatacagag ataaataaaa ..... 17820
tgcattggaa aaagagagaa aggacacctg actccatata gaggaatcat gggaggcttc ..... 17880
caagaggagg tgataaccaa atatggtcag catctatctc ctacacaaag tgtacctatt ..... 17940
ctgtgtaaca gatgaaacca tttcaaaata aagattaaat aagtaagatt aatatccttc ..... 18000
cctcactgcc actccttatt tccactcccc tcaccattgt ctatccttcc atgcctttcc ..... 18060
aagacagca tgctatcttt actgtgattt tcactttagg acgcttacta cactatcact ..... 18120
aatctgcttg tctgcagatt cagactgcca gataccctca agtctgcctg gaaagcaggg ..... 18180
aatagcatca aaggttttga tgcttgtccc agagatgagg gcaaacatgc tctgtcatgc ..... 18240
ctgtggaggg agaaccagcc tactaaattc ttggggcccg agatacagtc agatcattct ..... 18300
tctttttgcc cttccctagg cagaacctca ttatctctcc atgggggtgg actagatacc ..... 18360
agagcctcct tattggaaaa gtagtccttg cttaagggga cttggacttg cggaggctct ..... 18420
tgaaagcctg ggggttttta caataaacat aggatctaat tagacagctc taaaaccttc ..... 18480
ttcagtaagc agatccttct ctgttgctgc aatagaaccc ttgtaagttt tctcgtccca ..... 18540
atactttgcc ttaattaaca gacatcagct tttgttctgc tgttttactg tggaccagat ..... 18600
ttcctgacag ccactcccat ctatctctgg tttaattata gaatatgtgt agcctttgcc ..... 18660
ttaaaatctc tgatgactgg agcctaggga actctctgat tctgtcatat cacttttttt ..... 18720
tttgtaaaca aagcttaatc actgatgaat ttctgtaata ttctgtgaat tctcatcttc ..... 18780
cttatgagac tggccactat ctggaaatga tcctagataa aggcgattgt ggttttaaat ..... 18840
gtgttgtggt gaagacggca ggtggggtgc ctgtaaaatg acatccaggc agttgaagaa ..... 18900
atgacactaa atcttggcat aaagaacagg actagaagct gggtgtggtg gtgtgtgcct ..... 18960
gcagtcccag ctattggaga ggctgaggtg ggaggatcac ttgaggtcag aagttcaagg ..... 19020
ctagcctggg caagatagtg agagtttgtc tctaataaaa cttaaaaggt taaaaaatga ..... 19080
taaaatgtaa ggaggaggtg gaacaagata gccaaacaaa agcctccaga aatcatcccc 19140
ccctccccca tccctacagg aacaccaaat tgaacaacta ttcacaaaac aaagcacctt 19200
cataagaact aaaaaccagg tgagcgacca cactatttgg ttttaacatt atattaagga 19260
aagaggcact gaagaaggta ggaaagagag tcttgaattg cctgcaccac tcttcccctc 19320
ccccaccagc agcacatact gaagaatgaa tcacagtctc ttaatagcaa aattgatcaa 19380
gcagaaggaa gaattagtga gcttgaagac aggctatttg aaatacacac actcagagga 19440
gacaaaagaa aaaagaatga agcatgccca caagacctac aaaatagcct taaaagggta 19500
aatccaagag ttattggect taaataggag gtagacagag agattagagt agaaagttta 19560
tttaaaaggg taataacaca gaatttccca agcctaaaga aagatatcag tattcaagta 19620
cgagaatgtt gtgggacacc aagcatattt aactcaaata agactacctc aagatattta 19680
acaatcaaat tcccaaaagt caaggataaa gaaaggatcc taaaaggagc aagagaaaag 19740
gaacaaataa catacaaagg agctccaaca tgtctggaag cagacttctc agtggaaact 19800
ttacaggtct gaagagagta gcgtgacata tttaaagtgc tgaaagaaaa acacttttta 19860
tcctaaaata gaatatccaa tgaaaatatc cttcaaacat gaaggagaaa taaatacttt 19920
ctcagatgaa caaaagctga gggatttatt caacaccaga cctgtctaga caagaaatgc 19980 taaagggagt tctttaatca gaaagaaaat gatgttaatg aacaataaga aatcatctga 20040 aggtacaaaa ctcactggta ataataagta ctcagaaaaa cacagaatat tacaacactg 20100 taattgtggt gtataactta ctcatatctt gagtagaaag gctaaaaaat gaacctataa 20160 aaaataataa ctacaataac tgttcaagac acagtctaat aagatataaa cagatacaga 20220 tataaaaagt agggagctga agttaaagtg tagagtttgt attagttttc cctttgcttg 20280 tttgtgtatg taatcagcat taagttgcca tcagtttaaa aatgagttat atgatatcat 20340 ttgcacacct gatggtaacc tcaaataaaa taacatacag cagatacata aaaaattaaa 20400 agcaagacat taaaacatac caccagagaa gatcaccttc actaaaagga agacaagaaa 20460 gaacgaaaga aggaagagaa gatcacaaaa aacaacaaac aaaatggcag gagtaagtct 20520 tcacttatca ataataacat tgaatctaaa tggactaaat tctctaatca aaagacatag 20580 agtggctgaa tggattaaaa aaaacacatt aaaatattat ttaagaccaa atggggttta 20640 tcccagggat gcaaggattg ttcatcgtat gcaaactaat caatgtgata tgtcatatca 20700 acagaatgaa ggatataaaa tatatgataa tttcagttga tgatgaaaag gcctttgata 20760 aaattcacca cccttcatga taaaaactgg gtatagaagg aacattcata aacaaaataa 20820 aagccacata aacagaccca cagctagtat cacaccaaat gaggaaaaac tgaaagcctt 20880
tccttcaaga tctggaacaa gacaaggatg ctcactttca ctagttattc agcataatac 20940
tgaaagtcct agcgagggca atcagacaag agaaagaaaa aaaaaggcat ctgaattgga 21000
aaggaagaag tcaaattatc cctgttggca gatgatacaa tcttatattt ggaaaaacct 21060
aaagactcca acaaataact attaggagtg ataaatgaat tcagttaagt tgcaagatac 21120
aaaatcaaca tacaacaatc agtagcattt ctacatggtg acagcaaaca atctgaaaag
21180
21240
21300
gggaggccga ggcaggtgga tcacgaggtc aggagatcga gactatcccg gctaaaacgg tgaaaccctg tctctactaa aaatacaaaa aattagccgg gcgtggtggc gggcgcctgt
agtcccagct acttgggagg ctgaggcagg agaatggcgt gaacccggga ggcggagctt
gcagtgagcc gagatcccgc cactgcactc cagcctgggc gacagagcga gactccgtct
21480
caaaaaaaa aaaaaaaaa gaaagtagtc tcatttacta ttgctacaca tgaaataaat 21540
acctaggaat aacttttact caataaatga gagatctcta caatgaaaac tgtagaacat 21600 tgatgtaaaa aattgaagag gacacaaaaa atggaaagat actccatgtt cattgattgg 21660 aagaatggat atttttaaaa tgtccatact accaaagcaa tctacagaca gtgcaatcgc 21720 tatcaaaata ccagtgacat tcttcacaga aatagaaaaa agaatcctaa catttacatg 21780 gaaccacgaa aaaaacagaa tagccaaagc tatcctcagc aaaaagaaca aaactggaga 21840 aatcacatta cctgacttca aattatacta cagagctagt aaccaaaaca ccatggtaca 21900 aagaaaagag acacatagac caatggaaca gaatagagaa tgcagacata aatccacaaa 21960 tctacagtga actcattttt gacaaagttt ctaagaatgt atattgggga aaagaccgtc 22020 tcttcaaaaa atggtggtgg gagaactgga tatccatatg cataaaaatg aaactaaatc 22080 cctttctctc accatataca aaaataaat caatatggat taaagactta aatctatgac 22140 ctgaaactat gaaactacta aaagaagaca ctggaaaaac tctctaggac attagattgg 22200 gcaaagattt cttgagtaat acccctacaa gcacagggaa ccaaaacaaa aatggacaaa 22260 tggatcacat caagttaaaa accttctgga ctgcagagga aacaatcagc aaagtgaaga 22320 gacaacctgc aaaatgggag aaagtatttg caagctatcc atctgacatc agattaataa 22380 ccagaatata tgaggaactc aaacaactca ataggaaaaa tctccaatta acaaatgggc 22440 aaagatctg aatagacatt tctcaaaaga agctatacaa atggcaaaca ggtatatgaa 22500 aatatgctca acatcattga tcatcagaga aatgcaagtc aaaactataa tgcgatatca 22560 tttcacccca gttaaaatgg cttttagcca aagataggca ataacaaatg ctggcgagga 22620 tgtggagaaa agaaaaccet tgtacactgt tgtacaactt gtacattttt gagaatgtaa 22680 attagtacaa ccatgatgga gaacagtatg ggggttcctc aaaaaatgga aaattaaact 22740 accatatgat ccagttatcc ctctgctggg gaaatactca aagaaagga aaccagtata 22800 tcgaacctgc actcccatgt ttcttgcatc actattcaca atagccagga tttgggagca 22860 acctaagtgt ccatcagcaa gtaaatggat gaagaaatta tgatacatac acacaatgga 22920 gtactattca tccttaaaaa aatgagatcc tgtcatttgc aacaacaaaa atggaactgg 22980 aggagattat attaagtgaa ataagccagg cacagaaaga cgaacttcac atttctcact 23040 tatttgtggg agctaaaaaa ttaaacaat tggacttata gagatagtag aatgacattt 23100 accagaggcc gggaagtgta gtggggtggc agaggtagaa gtggagatgg ttaaaggata 23160 Page 12
caaaaatata tagttagcta gaatgaatat catctagtat ttgatcacac aacagggtga 23220
ctacagtcaa caacaattta ttgcacattt aaaaataact aaagggtat aactggattg 23280
tttgtaacac aaagaaagga taaatttttg aggtgataga taccccattc accctaatgt 23340
gattattatc cattgtatgc ctatgtcaaa atatctcata caccctgtaa atatatacat 23400
ctactatgta ccacccaaat taaaataaa ttaaaattaa aattaaataa atttaaaaat 23460
aaaactaaaa aaaccaggac tatcaacttg gctttattta agttgaatgt gatgttaaat 23520
cctctgggat agaatatgga gaaaagtaga aagaatggga gccccaagcc ttgaggaaga 23580
actcaactgt tagcaataac tgaagccaac cctcattgaa catttactat ttgctgggtc 23640
ttgttcttgg agctttacat gcagtaactc ttttaattct taaatacaca tgagattaat 23700
actattatat atgtgtgtaa aagtatctat atatacttac atataattat attttatata 23760
taatatatac ttacacacac acttatatat agtgtgtgtg tatatgtata gctctgtgaa 23820
accatttgca aattttttga ataggttaag ttatagcaga tacaacaggt tgggcactgg 23880
atatataaag ataaacaaaa tatagtacag ttttctaaga ccataataat atatggtata 23940
ataatatttt tggttaaggg ttttcaagcc aaaagtaga attagttaca atgtaatctc 24000
tgcaatatat tagaaaaaca gcaaaaaaaa aaagatcaa tatacagata gtacctaacc 24060
cattctaggt gcttaagaaa ttggatgaat aaataaaaga ctgaataata gaattttgta 24120
ggtgaaaagg ggcttagatg ttatctagcc tgaacccata atttatagac atacatgccc 24180
aggaggttga attaatgtgg ccataatgga attgcttgcg ggttgacaga gctgaattcc 24240
tgggccagga atcctacttt ttagtctggt atcttttcca ctctgctagt ggttttcaca 24300
ctctgcctaa attggaagca aatgccagca cttggtcttt taaatctctt ggatgctgcc 24360 ctgtaagttt gatttttagt gtgtcaggaa ggtcaggcaa gaggcacagc tcttcaggcc 24420
agaggtatat aaacagctca cataaaaga ctgtagttcc tggaatcatc tctgtcacct 24480
gcagtaagcc caactacact gagtcacagc cttcagtttg ggccaatgcc atagggattt 24540
atcatttagg caaaagatcc aaattcccct tccctgcggt tgatggagct ttctgcacaa 24600
gaagagtagt ttattggttt attagagctt cttaaatggg tagaaaagct aaatgaaaaa 24660
aaaacggaag ttcctgggaa aaactgtgct aaataggtca gcagtgttac cctgtcacct 24720
gaagtttgag gcctaaaagt gaaaggattt cttacttgaa cagaaatccg atgatattag 24780
agaaagtggc tctcattctg ggtttatctt cagtgtctgg aatgatgcct ggcatgaagt 24840
ctctgtggaa gacagtttga aaaccacctt ttcaatataa tttcagttta ccaattaggg 24900
aattatcaat gcatcatcaa aggataagac tatactaatg tatccatgtt atatattttt 24960
tattcatact gttcttaagt gttcagttat taaagttact aaatgtactg aagtatttag 25020
gtcaaataaa aaatccctaa atggtacagt cttatccttt aaagtcttgc ttgggtttag 25080
aaaaggcttt tcttgccagg tagaaaggtc taatgtatct ctctttaatc ttaaaatatg 25140 tgagacactt gtatacaaag taggagaatg gttggatgga atattttgat aaatggtctg 25200
gcttgtgtag tagctctgtg aaaccactcg taaatttttt gaatatgcta agttatagca ..... 25260
gatacaacag gttgggcact ggatatataa agataaacta gatatagttc agttttctat ..... 25320
gacctcacac cagtgggagt gggaggaaga gataggtgta aatagtcata cttctatttt ..... 25380
tgtatgactt agtaagttga atcatctgtc aattactacg gagttctatt gcaagatgag ..... 25440
tcactctgct aggggtaaga tgatgtaata ctttagttta ttcatggagg aagagtatat ..... 25500
tccaggagaa aaattttgat aaagggcatt ccagttagag aaagcagaat ggataaagcc ..... 25560
ttgaaggaag aaaataggat agttagggaa cagtaaactc ttcggtgtgc ttacagccag ..... 25620
ggctacacag aagtgagcaa agctggtgaa gatggggatg ggggagtgga gtgagttgac ..... 25680
gctagaaagg gatgtagcaa atgtaactat tattccagaa tccaagtgtg caaaaatcat ..... 25740
aacttttatt ttgtataaca taatttctat tttgagtccc tcttcagtaa atgtggcaaa ..... 25800
aacccaatct cagttatttt ggggtagctc tcacctttcc ctggggtgtg tcttggaggc ..... 25860
caacatagtt tcagatgagt gagaaggcca agtgttgttc agcatcatca tctacatagg ..... 25920
cattccgtca agctctctcg tcctcctctc ttagaagctt tgatgtcatt ctgttccttc ..... 25980
cgtgtctctc taggtccaag gagattcttg gaagctgttc taaagccgtc ccttcaactc ..... 26040
actacatatt tccttttctg aggtcttgct ctagggtctg cttaggtgga aggaggcagg ..... 26100
cttcttttac ctggaccctc cagtgcctat gtagtcagca attctcaact acttattttc ..... 26160
ttctcagcat gtggaaaata ttttttcagt ctaaagggtg gagcttaatc tcatgaaatt ..... 26220
taaacaaaat attctgtcat gctcagtttc ctctccaacc ctcaaaggca taggctttgg ..... 26280
gaacatttta tatatatata tcgtgtaggt ttcttcttgg cttttattac atatataaat ..... 26340
ttttttaatg cagcagcagc tctctctgaa gacagtgact gcacagtgct ctagctactg ..... 26400
aaatgaggag ataggtgtgg gggtaaagga aatatctggt aaatcattgt tggtacttta ..... 26460
aacatatttt ctttcatgtg tatattgtgg caagattgca aaagaactaa tatgatgtat ..... 26520
catggatctt gatctttatc ctttatgcag tggactgtca ttgaaggtta tcaatggaag ..... 26580
cactatgatt atacttttgc tttcaaatat aaaatctggg gacaatatga agacctaact ..... 26640
ggggtagaaa gattagttac agtagtcttg ttacaggaga ggtaagaaat gatggggact ..... 26700
tgaaccagag cattaactgt gaagagggag aggatttgaa tagctctaga tatattttgg ..... 26760
agtcactgaa gctgtggaag tggctgaggt cagcctggag cctgcataga gttctgggaa ..... 26820
acatcagtgt gtaggaaaga gggagagcat aaggggtcag taactgagac taagggagaa ..... 26880
gcatcagagt ggtggaagta ggagaaccag gataatgagg tcctacaaac tggtgataac ..... 26940
cttcaatgac agtccactgc ataaaggatg atggttttat taatgggagt tcccctgcat ..... 27000
gtgccctctt gactgtcgtc atgtaagaca tgcctttgcc cctcctttgc cttctgccat ..... 27060
gattgtgagg cctccccagc catgtggaac tatgagtgca ttaaacctct ttttctttat ..... 27120
aaattaccca gtctcaggta tgtttttatc agcagcatac gaatagacta atatagttac ..... 27180
ataggtaaat gtgtgtcatg gggtttttgt gtagagatta tttcatcacc caggtattaa ..... 27240
gcctagtacc cactagttat ttttcctgat tctctccetc cccccaccot ccacactcca 27300
aaggtctcca gtgtgtgttg tttccctctg tgtgtccatg tgttctcatc atttagatcc 27360
tacttataag tgaaaacgtg gtatttggtt ttctcttcct gtgttagttt gctaaggata 27420
atggcctcca gccccctccc tgtccctgca aaggacatga tctcattctt ttttatggct 27480
gcatagtatt ccatggcatc tatgtaccac attttcttta tccagtctgt cattgattgg 27540
catttaggtt tattgctttt gtgaatagtg ctacaatgaa catatgcatg catgtgtctt 27600
tataatagaa ttatttatac ttctttgttg aatgcccaat ttttattttg actggtacac 27660
tttgaaaggt gaagcacaaa taaaaggaac ttggtgctag acttgttttc tatgtatttt 27720
gatggctgaa gctataggta tggtgtaatt ggacatacaa acataaagga ggcaattttg 27780
accttgaaca aaatttttta ctttttgagc tatttagaaa tgactgtaga cttagaagac 27840
atcattgctg aagaacattt atttaaatta tcttggtttc ttatcttaag aattgagaaa 27900
actggtagaa acaatgagta gaaaacagat atttagttag acttgaaaac tggtcttcat 27960
ttggtagtga catactggca acaattaata ctatatagag acatttcctg gtaaatacag 28020
aagatacata atgtaagctt tcccaagatt gatttggaga aaaatcctca cctgactatt 28080
caatgtcctg tttcattttt gaaaacgtgg agcagaagaa atctcacaac tgatttaaat 28140
aaaattcaa tctgaaaaag tttattgttg tattgacata aggcaaagca gtcaagaatc 28200
cctcttgcat gtttttctaa gcttctgcat ctctttgttt cttccttacc ttctgtctta 28260
ccctttaatg gagttttctc caggattctg tgcttggcct gctgttttct ccctatctgc 28320
tggcctgtcc ttatgggtaa ggagagatgg atttagctat ggcctctagc acatgaatcc 28380
taaatcttta gatctgaact gtctctaaaa ctggaaagtt acggtctgaa tacttattgg 28440
acatcttcaa ttgaatgtgc caggcactgc aaaacaaaaa tgttcaagac caaactaata 28500
atctctcctc tgaccagggt agctcttttt gctgacttct tatttcccac catgacacta 28560
gcattctcca aatgcttagg cttgatgttt tgatcatttt taaaattttt actctctcta 28620
cttatattct atcagttgct aagtcctgta acttcttcct ctataagtct cataagtatt 28680
tgctttcttt tctattctca tttagactct ctttccctac atagttatag tagctgaagt 28740
cccttagtct ctccacattc cagttaatcg tatacattgc tggtggctaa actctcgtag 28800
aacatacttc taattatatc ctattttgct caaaactgtc gaacaactgc tggtctgttt 28860
tcttgcacca tacctgcagg acccagatta tctttccagc tacataacca cctccctcca 28920
tccacacaac acacttccta cccctgtcaa aaacaggctt acccattctt caaacttgcc 28980
gttgactttt gtccatattg ttcactttgc tagaaatagt ttttcccctc ttttccatac 29040
ctgttgaaat cagactcaac tgaaacaggc cctttcattt taacataact ttttgcagtc 29100
tcgattgtga taatttatcc tagaaaatat ttactctttg tacatttata acatatgtgt 29160
ataagtatgt ttaacacagc attgcttgta acagtaaaat attggaaaca agtgaaatag 29220
ttgcaacatt gaaaaaatag atacacacat atatatactg acaggaaaaa atttggccac 29280
Page 15
taaatgtaaa ctatgactcc attgttaaaa aataagtgtg tgtgcttgtg tgtatataca 29340tgcgtacata aaatctgaaa atgtacacac aaaatgttg aattgtgctc gtgtgtgtat 29400atgtgtctga gtgtgtgtgt gtgtgtctgt gtgttagggg taaagttatg aaatactttt 29460ataatgtact ttttatccaa aaatgttagc atttgctaat attgggtgat gagctcactg 29520gtgtttatta cactgttctt tgtatttgat atttttaaat aagaataaaa ctggaaaaat 29580gctaagattt taaatatata ccttagtacc aagtactcaa tggcatttat caggacaaaa 29640gatttttctt tagtgggatc ttggagtgat attccaaaaa atattttaat tcaaggacac 29700tgatatttct gctgacagac tacctcttac caaatgtggg tttcttcttg caaataaaca 29760tcactgagtt ggtgtatatg tggttgcaca cagtcagcag agaagtattt gaatgaatgc 29820cataatgctt acacacaatt aaactgagt cagttcgacc tatttttatg taaatcatta 29880aatgaaatga gtttgattca tttttacatg tttattttta atggagacta aagagacata 29940aatggtatgt ttgttttgtg gtggtctagg tgatatcaat gatacagggt tcttcatgaa 30000tctggaggaa gacatgacca ggtaattaga cattctcctt actattgtta agtttttcta 30060tattcatcaa gttgtagaaa tgtttaaaac tttgcattat catcacagaa attttaagga 30120gaacaatact cttgatagtt tttagaagag tatatgtaga ctttttaaga aaagaattgg 30180ctgcataaag tatacaaaag ttagagttaa gcttaaaatg gacatatatg cattgatcaa 30240tgtagaatat catattaata tataaggaaa ttagaggagt ttaaggtagt cttttaaaat 30300gcagttgaat taagaatcat tattttctat aatatatatg tgccaggtgt ggtggcacat 30360agctacagtc ctagctactt gggaggctga tgtgagagga ttgcttgagc ccaggagttc 30420aaagctgcag tgagctatga tcacatcact gcactccagc ctgggggaaa gagcaggacc 30480ctgtccccca caccccccca caaaaaaga gagagacaga ataattattt tcaatggctt 30540ataattatta atttgtttcc tggatgtttt catacgacat gacaatgaaa gcagccaaaa 30600gaaatttttt tttaaactca acatctgatt ctgaacaata aaataaaaca tcccagaaaa 30660aatacagtct ctctattctt aactcacact ggggaagctt ttggttaatg atttacatat 30720cttgaagttt attctccaga acagcttata acagatgtcc aaggtacgt gtaggtgaat 30780cagaagaaaa tgattttaca ttactgtgta aatgatctgg gtacatttga aaaaaataga 30840tttttttttt tttttttttt tgagacagag tgtcactctg tcccccaggc tggagtgcag 30900tggtgcgatc tcggctcact gcgaccttcg attcctggct tcaagcgatt cttgtgtctc 30960agcttcccaa gtggctggga ttacaggcgt gtgtcactac acccagctaa tttttgtatt 31020
tttagtagag atggggtttc accatgttac ccaggctggc cttgaacttc tgacctcagg 31080
tgatccaccc accttggcet tccaaagtgc tgggattacg ggtgtgagcc gccatgccct 31140
gccgaaaatg gagacacttt atcatatttt gaaagctgag gaccaaatac ctccaaatat 31200
tcctgtggag aggcattgtt ttttacattg gggaaaatga aacagaatgt cttttaatca 31260
tcatgtacat gcttttgagt aaacatacac tgacagttta aaatacatta acttattctc 31320
aaagcatgag ttttccttgt gatgtggggt ttctattagc tatctcttgc tatataatga ..... 31380
acacttagat agaatactat ctcttgttgt atcatgaaca cttagatagt aattagttat ..... 31440
tgagttatct gttaactaaa acttagtgtt ctaattacat cttccttgat aaatgcattg ..... 31500
taagttacaa ggtagagttt ggggaaaact tgtttcccaa gattcctgat gggctgagtg ..... 31560
ctgaattgtg aacagcaatg tcagagactg tctgcttgtt acctagccat ctttatagct ..... 31620
tctgcatgat tcaaacacac agtttaattt gaagatggta agtagtgaag tttaaaaagt ..... 31680
gggaatattg gaatactgta ttttatttta gaatgtttgc ctttaaccta aaaattaatt ..... 31740
gtttagtgaa gcccaagaat tgtttgcatt tttttttttt gagatggagt tttactctgt ..... 31800
cacccaggtt ggagtgtagt ggcacgatct ctgtgcactg caacttctgc cttcaagcga ..... 31860
ttctcctgcc tcagcctcct gagtagctgg gattacaggt gtgcaccacc acacccggct ..... 31920
aatttttatt, attattatta ttattattat tgtcgtattt ttagtagaga cgggatttca ..... 31980
ccactttggt caggctggtc tcgaactcct gacttcatga tccgcccacc ttggcctcce ..... 32040
gaagtgctgg gattacaggt gtgaattgtt tgcttttctt tttttttttt tgagatggag ..... 32100
tccagctctg tcacccaggc tggagtgcag tggcacgatc tcggctcccg gcaacctccg ..... 32160
cttcctgggc tcaagcgatt cttctgcctc agcctcccaa gtaaccaagt agctgggact ..... 32220
acagtcgcac gccaccacgc ttggctaatt tttttgtatt ttttgtagag acagggtttc ..... 32280
atcatattgg ccagggtggt ctccaactcc tgaccttgtg atctgtctgc ctcggcctcc ..... 32340
caaagtgctg ggattatagg cgtgagccac cacgcctggc ctgtttgctt ttttaaatat ..... 32400
gaaaactcag agcaaggaaa cattaattca ggatttctag attgattaag ggtctgtttt ..... 32460
ttaagacaat actcgtgttt tttaatgcta gaatttaaat tggaagggcc actatttgat ..... 32520
tcgttaagca ttttaggata actcacgatg gccagttggg caaatgaaat aaaactactt ..... 32580
tttaaaaatt tcatttgtct cttggcatct aaactctact agattaatag cccaagggat ..... 32640
aatccctagc ccaaaaggcc aagggattat ttccatttct tgatcccact ttctgacaag ..... 32700
ttctgtcact tcttttcatc cccaatgtac agagaccaac ccaatgcact tatcaactga ..... 32760
tgctgatgta ataaaaatat tacagtggtg ggtcttacct gatgctgtga ttccccaata ..... 32820
taccacacca tctttcacag taactggcat tctctgagtg cttactatgt gctgggcact ..... 32880
ggtccaagta tatttcatgc atgagtattt ttatctctgt tttactaatt agaaaactga ..... 32940
ggcttagaaa agttaagcag cttgcctaag attacataac agtgaagtga agagctggga ..... 33000
tttgagccaa agcagaaatc attcttaacc attccaccat attggtgcaa acatagttgg ..... 33060
atggtgacta tcaactctgc ttgttacaac cccttcttat ttgttaagga tggcgagata ..... 33120
cccagataag tgaaaaaata ggaatgagtg gtctctttgg aaaaggtaca ttgattagga ..... 33180
aacagagagc atgtgcagtg ttcttgttgt ttttcaggta tgcctattat tacagtggaa ..... 33240
ttggtgctgg ggtgctggtt gctgcttaca ttcaggtttc attttggtgc ctggcagctg ..... 33300
gaagacaaat acacaaaatt agaaaacagt tttttcatgc tataatgcga caggagatag ..... 33360


| cattgcttca | catgccattt | atttagtaaa | aggagcceca | ggctggctgc | tagagaccta | 35460 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| agctttaatc | ctcactcagc | attacacttt | tggtatcatt | tcaaacaaa | ttcttaaatt | 35520 |
| ctttggatgt | caatcatatc | aaaaaagaaa | tgtagttcac | tctgtgtctg | ggaagcctct | 35580 |
| ttgtgtctgt | gagcctgagg | ctctaccttc | aaacctagag | acacagaagc | aaatagatat | 35640 |
| agcagaaagg | gacccaatca | cagcaagtga | cctggcccag. | gaagcctctt | ttgtactttg | 35700 |
| tgcccaagac | tgtcttccca | gcgcccagag | acactgggcc | actagagggc | attggtacag | 35760 |
| gaaaccacca | cactcttccc | gacctggcag | tctgatcttg | aaaaaccact | ttcactccct | 35820 |
| caagcagcag | aagcaggggc | cagtgggagc | cccagaagct | tcacataaat | caagccgacc | 35880 |
| aaaataatac | tacaamcact | ctgaatatta | aaccgctcct | ggaaccacag | ctcacaaaag | 35940 |
| tagaccaata | ctatcatgct | aaacttcaat | ggaacaactg | ccggctagag | taaggcaggt | 36000 |
| ttattttatt | ttaacttttt | aaaactttta | ttttaagttc | aggggtgtaa | gtgcaggtct | 36060 |
| gttccatagg | aaacttgtgt | catgggcatt | tgttatacag | attatttcat | tacccaagta | 36120 |
| ttaag | tacccattag | tta | tgactgtctc | cctcctccca | ctctccaccc | 36180 |
| tctaataggc | cccagtgtgt | gttgttcccc | tctatgtatc | catgtgttct | catcatttag | 36240 |
| ctcccactta | taagtgagaa | catgcagtat | ttagttttct | gttcctgagt | tagtttgcta | 36300 |
| agcataatgg | cctccagctc | catccatgtc | cctgcaaagg | acatgatctc | atttttttta | 36360 |
| tggctgcat | gta | gtgtctatgt | accacatttt | ttttttatcc | agtctatcat | 36420 |
| tgatgggcat | ttaggttgat | tccatccttg | ttttgtgcca | gttttcaagg | ggaatgcttc | 36480 |
| cagcttttgc | ccattcagta | taatgttagc | tgtaagtttg | tcatagatgg | cttttattat | 36540 |
| tttgaggtgt | gttcctttaa | taccttgttt | gttgagagtt | ttgatcatga | gtggatgttg | 36600 |
| cattttattg | aaagactttt | tgcatctatt | gagataagca | agtgcttttg | tttatgtgat | 36660 |
| gaatcacatt | tattgattta | catctgttga | gccaaacttg | catcctgagg | ataaagccta | 36720 |
| cttgatcatg | gtagataagc | tttttgatgt | gctgctgaat | tcagtttgcc | aggattttgc | 36780 |
| tgaggatatt | tgtatcatat | ttgttcatca | aggatactgg | cccaaagttt | ttttgttgtt | 36840 |
| gtatctctgc | caggttttgg | tatcaggatt | attctggcet | catagaatgg | gtcagagagg | 36900 |
| agtccctcca | cctaattttt | ttggaacatt | ttcagtagga | atggtatcag | ctcttctttt | 36960 |
| tacatttggc | agaattcagc | catgaattca | tctggtcctg | ggctttttta | gggggtggta | 37020 |
| agctattatt | actgcctcaa | tttcagaact | ctttattggt | ctgttcaggg | attttatttc | 37080 |
| ttcctggttc | agtcttggga | ggatgtatgt | gttcagtaat | ttatccattt | attctagatt | 37140 |
| ttctagttta | tgtgcacaga | agcattcata | atattctctg | atgggatcct | tgtgactgaa | 37200 |
| ttgttctgca | tcatgactgt | gttgtgtact | accatgtcca | gctagttttt | aatttttttg | 37260 |
| tagagatggg | gtttctatgt | gttgcccagg | ctggtcgtga | actcttggcc | tcaagtgatt | 37320 |
| ctcccacttg | tcaccccaaa | gtcctgggac | tataggtgtg | agtcatcaca | cctggccttt | 37380 |
| tcatatcctt g | gataatagga | taggttagag | ggtctttttt <br> Page 19 | tttttttttt | tcttttcaca | 37440 |

gacatgatct cctgaaaaac acaacactta atgaagattc aaccatactt ttgtctgctt ..... 37500
ataatgatgg aaggtcatag aaggaaaggg gagctagatt tgtagaattc attgtggtgt ..... 37560
ttgatatatt aataataaat taaaataaca tttttaaaca ctcctaactt taaagtattt ..... 37620
ctttaaagta tcactttgaa gtgatcttcc tatttttggt tgtctccatt gtggcaattg ..... 37680
ttgtggcacc atgaggtttc agaaggagca gaaaaacatt cctaaagcat gttgtatgag ..... 37740
cctcctctat tctacagatg tttgttaagg gatggcgctg gccaccctat ctcagacctc ..... 37800
ctaccgtgaa tggttctgca aatatggtgg gacagcccca ggaagattgc tttgagaagt ..... 37860
caggaatgtc tagtgaatca atggaatatt attttgggta ctaaagaggc taggtacgtg ..... 37920
cagcacaatg tctgcactta gaacataact aatcaatgtt catttgcctt catccaccca ..... 37980
ttcctctagc tgtcatagct ccttttactt tcacacacag gaaagggaga aatccactgt ..... 38040caaaacccaa acaactttga gctttttatg ccaagaaggg atagatcaat gttacttacataaatgagtc cacaagtgcc tatgaatttt ttattgaaag aatataattc tatttaatcagaatttttaa aataaattgg ttttgtgtaa ttttggggga ggggtgggga gaacagggtctcactctgtt gcccaggctg gactgcagtg gcacaaccac tgctcacggc agcctcgtcctcctaggctc aaacgatcct cccacctcag cctcctgagt agctgagaca ataggcatgtgccaccacac ctagctaatt tttaattttt ttgtagaggc ggggtcttgc catgctgcccaggctggtct agaattcctg ggctcaagca attattcatc tttggcatcc taaaatgttaggattacaag tgtgagccac cacacctggc ctgtcatgtg taatctttat ttaggtagttgaccacttca gcattctagg tacaataacg ttagcccttt tcccattgca attgatccagctttcataat aggaccctct gggacccaag ttcatgcatc agtggctggt ttcagggagaggtttacttc agtggctgtt aatacagacc aaaggcaagt aaaagacaga ttttgctctacacatgcatt aatgtataag cagcatttat gtatacattt atttttactt tacaaaagataaaattaacg tgtttttttc atgatccaaa attgtatttt taaacagata ctatcttcatttactgataa agaactctta gcgtatgcaa aagctggagc agtagctgaa gaggtcttggcagcaattag aactgtgatt gcatttggag gacaaaagaa agaacttgaa aggtttgagtttctttttta aatggataga tatgttaaat tctgtcttct caaatgccct tcagattgacagtgttattt ataagcattt tctcccatat atatgttttc ttaaatagct taatggatgtattacagcat aactaatatt ttggagaagt tgcaattcta attgtacttt tcttttatctgctaattcaa agagttttct agatgggcaa taacataaaa atagtttcca aagactgaaatgtattttat acctctgatt tttttcctct atcaagttaa ttagtgattc tctaatattgactgagtatt ttgtgtaaga tctatgggga cattttaatc cgtggttcat tttcctttctgcccctattc agctgacata tgctacatga tttggaagat aaaaatataa ccaaaaagcacttatgtgta caaaagttca aaatctttta tattatgtgt tgtttagtat tatgcttacaactgttccct tttattagat ttctgaaagc atagatggta atttattgaa atttagaata381003816038220382803834038400384603852038580386403870038760388203888038940390003906039120391803924039300393603942039480
taaatcagat tactgatttc taagcagcaa cagattttat gtaaaattga atataaatta
39540
taatattctg atgctataac acttcaatct tatgaatata aatatctttg aaggtatgaa aagtaaact gtattggtta agggggttat acattgtttg gtacagccat gaagaaaacc aagggaattt tcagaaatca gaataatggt tagctctagg gaagatgaga agttgtgtaa ggagagccga acacacagca ggggattctg tggttctggc agtgttctgt ttcttgatct caggtatact tacatgagta ttcattgcat aattttttta actgtgcata tgtgtttgga gtacttttgt tctatggtat attcacacat ataaatatat gtcttctgag aataatttat aatggttaag tgccatagtt caaaggaaag caatatctat gttccatttt gatttttgct aagtctttg cttttgcttt aaagatgtc tggctaataa aatgaggagt cctgcttaat gactgagcaa atattcaatg aaataaatcc ccggctctca tggagcttaa attctaatgg gggtggtaag gttacagaca accacaaata atcaaatagg taatacacca gttggggtaa atattatttg gggaaataaa ataagagaat catgaggata ctatttagat aagttattca gaaaatatt ttctgattag ttgatatttg agcagagaca taatgaagtt ttgagtttgg aattcttttc cagttttcct cagagaaaag gtggaccaga aaaataatat gcccaacagc ccaacagcat tggaagacag acaaaactta acagcttctt cctggtttct cagctggagg aaaacatctc cccatagaag ttagattctc aggcctgact gctatctgca gatggggtct gaatttatac tactctgtga agtaggaaac aatctctctc caaagctgag gatttcacat taaaaatgag tcctaatcca gtgataccag agtatctggt agaagaaata caaaattatt gaaaattaaa gaactaaaca tctaaatcaa gaagttaggg aaagaaaat aaaatattag aaggaaataa taagggtaag agcaaaatt aatttaatag aaaataaagg caaaagagtg gaagattgac atatccataa actaattctt taaaagact tataaataac tattttatat gtttaatcag gaacaaaaaa agaatgaata gggaaaactt gtaacagatt aaaaaatcta tttttaaatt gcagctttat accaataaat ttaaaaccc tcatgatata gataatatgc tagaaaaatg taggttaaaa ctctcactta agaaggagaa gatctgagta aacccataat cgttaaaata attgaattca aagttaaaaa tctccatttt ccactgcttc tctcccacaa gcagaaatca caaaacaagt ctggataatt ttgtaggcaa gttctaccac aacttcaaga aagagataac cctatcttac ataaattgtt gtagaaaatg gaaaaaaaaa gttgtgagtt cattttatgg ggctagtgta actttgatac tatattacca ggcagatata atatgagaaa ggaaaactaa aggacagttt cacctaaaac atagatgcag aaattataaa taaataaata attagctaaa taaatccatg tgaaaatac caaattgaaa acatcctgaa catgttggct ttatcctatg aatggaaaga tggttttaac attaggaaat ctattaatat aattaaccat attaatagac taaaagaaaa aatttgttgt gatgtaaggg gaaaagcatg aattatatat gattcatgac aaaaactcgt aaaactagga ataagaaaca accttcttaa tatattagag acccgcagga aacatttagt ggtgaaacat taaaaatatt ttaacactta aaataatgaa

39600
39660
39720
39780
39840
39900
39960
40020
40080
40140
40200
40260
40320
40380
40440
40500
40560
40620
40680
40740
40800
40860
40920
40980
41040
41100
41160
41220
41280
41340
41400
41460
41520
tgaaataaat acatactatt ttgtctgtca atttaaaaa taaaaaacaa ataatgaatg aggcaaatgt gtttactatc atttctatgt actgttgtac tagtactaga taataagacc aataaaacaa ataagagtaa gaatgaaaaa aagaaaacga aaagaaaaa cagtaagtat tggaaaatag gaagaaaact gtcgttattc acagacgata agattgtcta cacagaaatt ccagtaggtt ttataagcta acaaaataca ataagatctc tgggtaccta tcattattca gaataaatgg cattcctgta aaccaggaat gaacaattag aaaatatata gtatttcatc cattcagaat gaatccatca tctctgaaat ggaaataaac ctaacagtca gtgccatctt aaatttgcta tggatctgga ttgaagtcca attctatcag gcaggattta tgtttgcttc tgctattcac caagggcact gctaacctga gaccacctta cattaaatta tgtgtctagt tgtttttgga caacactggt agtgtgaatt catgctcaga cctgcatgag tcctactttg tggtccagat tctcagggag gtttttttcc ttccttcacc tggttggtgc caagactgag acatgcaagt tctgctgtgt gttgcttaat atactgaggt atagcccttt gagtcctagc tgcatatgga atgttctatt aaactcattt cttaatgctt tataaaatga cagaatacca gtacttgcta caattaaaga catcttaaat ttgaggaact gacaatattt ttaaggacat tattagcaac aaaaactata atgatggatg gaataaaacc tgagaaggag gaaatggaca ttgtgggagt gaagcccaac atttgtaact tagataaatc tgatatttta gttcactgaa agtttaaagg gattcagatt taatggagtg ctttaattag ttagatatga aaatattttt cctagcaaaa atgatcaaat tttaattaga agaataaaca aatggactta cacagctata tccagagatt cctatatttg ctctaagtag tattttacat gtggaatcat ttcttcagaa aatataagtt acttgaatgt atctttgaaa gggcagtctc cttcagagaa ctgaaactaa ttttgcttgg agatcaagaa ttgacgtaag actacaaact gatggcctaa tgtccttccc aagttggcca tttctttaaa gtttatttat ttgacaaatt tattcagcac ctacaacatg ccaggcactg tgctaggttc taggttttcc ttttcttttc ttttttattt tataaagttc taaacaaagc tagggctcta ggttttcagc tgtgaatcag aaacagaggt tattttgtga tcttcagtca aaatagtaaa atggataaaa atggaaggaa tggagaaagg gaggaagaaa tataaggaat cgtcaaagaa aaaaataaaa atattgattt aaaaatgtga tctgaaagaa tcaaagagtg tgaggatgat gaatggtcaa gttaacttta aatacggtgg ttggggcagg cctctttgag gaggtttcat ctcagctgag acctgaggga tgagaaaaag caagccatgg tgggaacaga gatgggatgg gtttgaagaa gagtcttcca agtttcctgg ggtggccaag aatttagtat gttctaggaa acgtcataag gcttgcatga ctgaagctca gttgaatgag gggagattgg ggtgggcagg ggccagaatg tggggatgat gggccatgat gaagagttag gattgtattc taagggcaca ggggagtggt tgagggtgtt aaaaagagga atgactatcc acgttatgtt ttcagattgt ttgtctgtaa tgggtctcat gaatggatgc gagaagtcaa agggtggaag cagggagacc agtggggagg ccggcagaat ggtccaggcc agagcatctc

41580
41640
41700
41760
41820
41880
41940
42000
42060
42120
42180
42240
42300
42360
42420
42480
42540
42600
42660
42720
42780
42840
42900
42960
43020
43080
43140
43200
43260
43320
43380
43440
43500
43560
cctgatctgg agcagaggca ggagaggggt gacaggtacc tagatttggg atataacttg
43620
aagtagaat tgatgactat tcttagacca gacataatca gttttctttc ctgttcttct 43680
ccttcccttc tctgcttcta cataaagtaa ttgaggtctt aatctggctt tctcttcaaa 43740
ttagcagtgc aggagggaag ctggctctct ttgaatggaa atttaaccag aagttaaaat 43800
aaattccatt caatcgtata gaatagtttt gttccttttc acttaaaaat atttttctct 43860
cttttatgtg ccttgaaat atatttatct attaatattt atttctgttg gcttaaccaa 43920
aattgttcct tattcttaaa gtagttactt gcattattga taaatatttt cacaatattt 43980
aagaagagtc atatttccta cataccaatg gttttgttat ataactttag tctcttttag 44040
tcattcctac tttcatgacc atttctgtac taagccttt gatgaatgga tgattttaat 44100
atctagtttt atgtatggca tttgtatgga aaaacaacaa agagcttatg gaaaaggagt 44160
attgtttata tctcatttcc ctttgccccc ttctttcctt tatttctcct ttctgatgtc 44220
taaatgattt tcttccttgt gcttctgtgc aaatgtttcc tctggctgca taggacaaca 44280
tcagactggt ttctacgtta gtcttggtaa gcgagaactg aaataaagga gataacagag 44340
tatagtacat atgtattaaa tgtccttaaa tatattctta tgtaagttat gcccagtggt 44400
agcttttacc tcagatcaac tcataaagtg agagatatgt atgtgtctgt gtgttgtgtg 44460
tgtgtgcata tgtgtatata aataattgaa acgtttctca aaataattt ttctgttcag 44520
atgataactg gcaagttgag gatttacatt ttggggggta gaaactaatg atgtgttgct
atgttgtcac tatgaagttc tgtttaaatt tcatctttag tgaacacggt agtgatgtaa
44580
atttcacttg gcatgctttt gtaactctgt aatgtaccta gttaaagtag tcatgaaagc
aagaaacaat gcatgttcat atctggacag tgcttgagct ctgctgtaca tgctaatgga
acaagttctc caaaggaaat taatataatg ctatataata tttatatagt tgtgaattta
tggatcaaag accttaaagt atgactttta tatggaagta agttcatgaa tttcaacaat tattagctcc tcctctatcc aaggcaatgc tctaggttct gcagggagtt gcgattaaga aaatacagac agatttaagg caatctgtcc tgtttctcta atccttttac cttggttcta catttgcaaa attccttatt tttctttaac tctgggctcc attagcttat ctccatagct atcctcttaa acagagtagt cttctctgtt cctttcaaat gaaggtcgag ttatagacat catctccttt aactttaccg tcaaaaattg accaccctct gagggtttaa aaaatatttg aattatcagt tgggaataac ctgtaagtca ttagaactgc agaaataaaa tctgtcagtg actccaacat tgtttttcga ttttttttca tttgctcatt tttgcccctt ttgtacaatt tcctatattt gctctaagaa gtattaatgg gtcctcacca ggtttggtta atgtagggta ggcctttgtg ctatgaatgg gttcagggct atagatatgc taataattta atttttctca gcccttgagt tagtcagaag ccgtggatgg cccaagccac caggtcagat cctgttagtt cagcagctgg cagctatggc atgacaagtg ttgggagcat taatagaggg aaggacacat agtcaatatg tagaaaggag aaggaagaga gggaatgctg atgaataatg atgtgctttt 45600
ataaacctgt aagccaaata atttttgtag gtcagtgatt tttaaaataa atgtctacgg ..... 45660
aagtaagata aattgtacag cttatttaaa tgtaaaaagt attacatggc aattcagaac ..... 45720
tctcaaaatg ctctcacaca tgtaacctct cagttggtca ttccgacagt cctagaggaa ..... 45780
ggcaggacag agagcaaact gtctccaaac ttccttaatg ttcagagatt taaatgactg ..... 45840
tccctttttt ttttttttag tactttctaa cattatctct ccattttatt aattaaagag ..... 45900
gcatgcagat tttgaaggaa agacacaggt tttttccatg gcttttactt tgtcttgtgtatgaaacaat gtaagtttta gaagtttgct aaatattgct tcacttattt gctattgctaaagccagaga aggaaatttt agaagactcc cccacaaata agttaatctc tataaatgtcagtttcttct tctgcaaaaa gggagtaaaa cagtactact atatagggtc aatgtatgagcagcaggaca aaatgcatgt atatcacagg actgaacaca tcctagccac tctgtaactatcatgctata tcgacatgtt ttcataaaat gtatatgtga tcatttttgt tctttttctcaggtacaaca aaaatttaga agaagctaaa agaattggga taaagaaagc tattacagccaatatttcta taggtgctgc tttcctgctg atctatgcat cttatgctct ggccttctggtatgggacca ccttggtcct ctcaggggaa tattctattg gacaagtact cactgtaagtggtttacatt gagaaatgaa ccattattat aaaatgccaa tgaaatccag taatgttggcttgactttag aagaatatat tttgaacaac tgttgaatac tacagacata tgcactactaggttaaaaag tggaaggcca ggcacggtgg ctcacacctg taatcccagc actttgggaggccgaggcgg gcagatcacg aggtcaggag atcgagacca tcctggctaa catggtgaaaccccgtctct acaaaaaata caaaaaatta gccgggcgtg gtggcgggcg cctgtagtcccagctactca ggaggctgag gcaagagaat ggcgttaacc cgggaggcgg agcttgcagtgacctgagat tgcgccactg caccgcagcc tgggtgacac agcaagactg tctcaaaaaaaaaaaaaaaa aaaaaaagtg gaaaaaaatt agatgcatcc gacaaccagt tttttcttttatttgcttgt ttgtttttgc agatatcttg cacaatttat cctgtagcct aaataatattaacgttagca atctcaatag catgaacttt tatttgttga gattatttgt gccaaacttgttgctaaaca cttcactttt attttctgat ttactattca caaaaatggc aaaatacacattttgacttc ttttcaaatt atttcgaatg gcagctaatc cacctcacaa attatggaagtataaattgt tcctcttcct aaaagattgt ttggaagttt tctgcattat agacctttgt47160gcaagcacag aagtcttgcc atttaatgtt atcctctttt attttaaaga caatgaaata47280
atttgcaaac caaactcaat tcatatctgg atctaataaa ataçactgtt agaagagtca ..... 47340
cttagttcct ttttcgtagg ccaagcagaa atagattgag gcaatgagag atgatatttt ..... 47400
gtagaatccc ttaaaacaat tccctttttc aggagccaga tttgtaatat ggattcaata ..... 47460
atcacttctt tttttttttt tttcaaattt gccttatcca acattgctca gtggggaatg ..... 47520
gcttttccca tgaagttttt tttttttagt atagtgatgc tgtgggtgat ccttggttgt ..... 47580
cataacagat gttttctgtt taactgtgat tgagataagc agtttaaatc cagtgttcac ..... 47640tgggcatgaa tatacttcac aaactggagg catccaaaat agggagtgtg gggtttgata47700tcggaccaac ttgattcaaa tggcagcttt gtcacttatt agctgtatga tttagggcag47760
ccagctaaac ttttcaagcc tctgctacct cttctgccaa atgatatgac aggaatgaac ..... 47820
cctaactttc agcgtgtttt gaggactaac tgagataatg aatatagaat cacttagcaa ..... 47880agttccttct gtgtaaatta ttgagcaata taattccatg tgatttagaa aaaatgaataagacaacctg ataaaggatc agtgtcttag gtgtgtgtgc tttataatta acaaaaaatg47940
atgacattaa ttctcatcaa ccccaaagag aaacacagtt aagtctttac caataagaaa ..... 48000
ctgatttgca caaaacaaaa aaacaaaagg cctaaagaga acacggcctt gtggcaaaaa ..... 4806048120
ttcatataca ctgttttatc atgatacaac atctctatga ttgctgggta gggagaacag ..... 48180
atagaatctt gcctgttttt gccaatgaag aaattgaagc caaagattct aagagaatgg ..... 48240
cctaacaaca tataactagt tattactaat ttaggattca gaactgggtc ttctgactcc ..... 48300
tctgctagca tttatttcca ctttacctta ctgatgatct aaataaaaaa aaaactagac ..... 48360
attgaagaaa tataaaaata tcctattata atattctagc cctgtgaaat ttaaatcatt ..... 48420
gtttattgag caagtaagac taattctgca ttctaaattg aaaagtattt atatttccaa ..... 48480
aatattttta gtatgggcaa catagtgaga tctcatctct ataatttttt tttttaatta ..... 48540
gccagatatg gtggcacttg tctgtagtcc cagctacttc ggagactgag gtgagagggt ..... 48600
tgtctgaacc cagaggttga agttgcagtg agccaagatt gcaccactgc agtccaggct ..... 48660
gggcaacaaa gcaagacctt gtcttaaaat ttttttaaaa aaattttatg ctatttttaa ..... 48720
ttgacaaatc ataattgcat tacacttacg gaatacaatg tgatgtttat gtatacacac ..... 48780
acatacacac aaaatgtgga gtgatcagag caagctaatt aacatatcta tcacctcact ..... 48840
tgacaccttt tgaggtgaca catttgaaat ttactcttgg ctgggtgtgg tggcccatgc ..... 48900
ctgtgatccc agcacttttg caggctgagg caggaggatt gcttgaggcc aggagaccag ..... 48960
gctgggacac gtagcgagac cctgtctcta gaaaaaataa aaaaaattag tcaggtgtga ..... 49020
tggtgcatgc ttgtagtcct agctacttgg gagactgagg caggaggatt gcttgagccc ..... 49080aggaggtaga ggttatgacg agctatgatt gcaccactgc actccatcct gggtgacaga
atgagaacct gtctcttaaa aaggaaaaaa aatgaaaaaa ggaattaaat gtactgttag ..... 4920049140
ttacaaactt ttaaaatatt tttaacacct agtatttata tgttgcctcg ccattttaaa ..... 49260
ttttaattac ctatgttaga aaaagacaga agaagtatga aaaaaaattg ctatcactat ..... 49320
ctcagtagcc tgatggtttt tcttcacatt cctcaggtat tcttttctgt attaattggg ..... 49380
gcttttagtg ttggacaggc atctccaagc attgaagcat ttgcaaatgc aagaggagca ..... 49440
gcttatgaaa tcttcaagat aattgataat gtaagtctga gttggccatc tatccaccta ..... 49500
tctaaaagat tgtccagtta agtcaatttc tttgtcactt tatccagctc tccacaaaat ..... 49560
atcactaaaa gtagttattg taacctagta atctcttaaa atttgattct gtttagaagc ..... 49620
caagtattga cagctattcg aagagtgggc acaaaccaga taatattaag ggaaatttgg ..... 49680
aattcagaaa tgttcacttc agttacccat ctcgaaaaga agttaaggta cagtgataaa 49740 tgattaatca acaattaatc tattgaatga agagtttctg atgttttctt gtagagatta 49800 taaaaagtg catgtatatt taaacctagt gaacagtcag ttcctatatc ctgtgtctgt 49860 gaattgcctt gaagtttttt tctcactcgt cctggtagat cttgaagggn ctgaacctga 49920 aggtgcagag tgggcagacg gtggccctgg ttggaaacag tggctgtggg aagagcacaa 49980 cagtccagct gatgcagagg ctctatgacc ccacagaggg gatggtgaga tgacccatgc 50040 gagctagacc ctgcggtgat cagcagtcac attgcacatc tttctgatgt tgccctttca 50100 attacaaatg tatgaaagtc acacttactt tttattccag gtcagtgttg atggacagga 50160 tattaggacc ataaatgtaa ggtttctacg ggaaatcatt ggtgtggtga gtcaggaacc 50220 tgtattgttt gccaccacga tagctgaaaa cattcgctat ggccgtgaaa atgtcaccat 50280 ggatgagatt gagaaagctg tcaaggaagc caatgcctat gactttatca tgaaactgcc 50340 tcatgtaagt tgtccttgcc ctttgccttt ctagaggtgc aaaaataaa atgcaggcct 50400 actatgcagg aagttaggaa actactataa atcggaagaa gggaaatcct aagaagggaa 50460 agtaagatta cttcagattt gaaagctcta gcagtatcaa ctggtcgtag atacattttt 50520 aaaaactgag gttggttatt gtgttaaata agatttaaag aactggacct gtattacttg 50580 tgagacttgg gctgtgtata ggattcctta ccaatttaaa atatgagctg agatagcttg 50640 tccttatgct aaatcattct gggttttctg tggtagaaat ttgacaccct ggttggagag 50700 agaggggcce agttgagtgg tgggcagaag cagaggatcg ccattgcacg tgccctggtt 50760 cgcaacccca agatcctcct gctggatgag gccacgtcag ccttggacac agaaagcgaa 50820 gcagtggttc aggtggctct ggataaggtc agtgaggctt agttcaaacc aacctgattt 50880 ataagcataa gaacattcta ctactaattc ttgttaatat tggtcttaga aaaggaaatt 50940 tctgatagct tctaggtgat tccttcagct attaaaataa aagcattggg cctctttgaa 51000 atctttttct atttgtttgt tttattgttc aatttctatt tatttctctg atcttatttt 51060 aaatgttgat gaatacattt tcatttgaag acacttgcta atcttttaaa ttaaaaaata 51120 gaaatataga cacatgtgaa agttcatctt cattgtgatc ttcaaaactt gactatgtgg 51180
ataaccctgt tatttaggtt ttgagagttt gtaatattgc caagaagaga aaaatacaac 51240
ctgaaggtcc atatataatt ttccaggtgt tgaatgccac ttgaagactc tatgcgaaat 51300
aagaaacctc ttatttccag gaaaggggca gatagcctgt gatactgaaa acctacctaa 51360
gccatgacag gttattgact atcaacagag tttgactgtc ctgcaattct ggagtccata 51420
tgactcattc accaaatagc atttgagtgt ttgccgtgtg ccgggcactg tgcctttgat 51480
ctccagcacg tgatagtaac ggggataatt ctgtgaggac cgagaatgtg gagatggaga 51540
cattataacc aaaggtgttc caagttgaga tgtcacagta gaattcaaag atgaactcat 51600
atttgtttca actctccctg tctctaataa aacaacactt gaatgttcct taacatcctg 51660
tcaatgtgct taataaattt ttgagagaga aaaaagcag cttactaaac attctgtgaa 51720
ccaaaataga ggccgatggg attctggtta ctatttttcc cctcattttg cttaatctgt ..... 51780
gatttcatct ctgtgttttt ctttttcttt ctttatttcc ttccttcctt ccttccetcc ..... 51840
ctccctcaat ccctccctct cttgctcttc ctcttccttt cctttctttc ctttcctttc ..... 51900
ctgaccttcc cttcctttca tttcctttcc cttcccttcc ctttctttcc cttcccttcc ..... 51960
cttcctttcc cttcccttcc cttcctttcc cttcccttcc cttcctttcc cttcccttcc ..... 52020
cttcctttcc ctccccttcc cttccctccc cttcccttcc ctccccttcc ctcccctccc ..... 52080
atcccctccc ctcccttttc ttttcttttt tctcttctct tctcttcctc tcctctcctg ..... 52140
tcttttcttt tcttatctta tcttatcttt tcttttcttg tttcttttct cactctgtca ..... 52200
ccaggctgaa gtgcagttgt gccatcatgg ccactacaac ctctgctgcc cagtctcaag ..... 52260
tgatcctccc acctcagcct cacaagtagc tgggaccaca ggtgtgtgcc accatgccaa ..... 52320
ggtttttttt tttttttttt ttttgagatg gagtctcgct ctgtcgccca ggctggagtg ..... 52380
cagtgggaca atcttggctc actgcaacct ctgcctcctg cctcagcctc ctgagtagct ..... 52440
aggattacag gcatgcaccg ccacacctgg ctaatttttg tatttttagt aaagacaggg ..... 52500
tttcgtcatg tggcccacgc tggtcttgaa ctcctgacct caggtgatcc acctgcctcg ..... 52560
gcctcccaaa gtgctgggat tacacgcgtg agctaccgtg cccagccccc agctattttt ..... 52620
tgatatattt gtagagatga ggtctcacta tcttgccccg actggtctca gactcctggg ..... 52680
ctcaagcaat cctcccgcct cagcctccca aagtgctgga attacaggag tgagccactg ..... 52740
cttactggtt tgcttatctg tgtttcctta ttaatctata gtgaaactat gtattaaatt ..... 52800
ataaataaaa acaattttaa aaggttatat tttaaaatac tttagggtgt aaattttgag ..... 52860
gggaaattcc acatacccct tttttcttaa aaagatacaa aaattgatct attttcttct ..... 52920
gtattttcta gtttctacca cctaattttt ccttgtgtat tttttctttt tgaagttttc ..... 52980
cacttctact tatcctatgg atcctgaaaa tgttgtgtgt tggttttgag aattgtattg ..... 53040ctagttatta gagagacata tagagtaaca aaaattatga gcattgggaa agttacaaag 53100gttagagaag tctcagacaa ggcctggata tctggctctg ttcctttatg aaaataaaag 53160agacttacgt gactcttcaa tttttccata attcttcaac ctaggaataa gtatcactaa 53220ctatggataa ggcacagtgt tgagtacttt atgtgcttta ttttatttag tcatcacaac 53280tactctaaga agtaaatact attattatcc ccattttaca tatgaataaa ttgagtctca 53340cacagtttcc ttggataaaa tattttattg gataaaataa attcataaat ttatttcagg 53400tcagtgtgac attgaggtct ggactttgct gcctcacatt tattgtctgt cttgttcatc 53460cagggggtca tgcgggatag gatattataa ttcctagggc tgattactag ccggtgtgta 53520tcagtacagc acaatggcct gtgtttgttt tgattggcca acgcctggtc tgtaggaatt 53580tgttggtttg tacaagcccc tgattattat tattttttta ttttttattt tttttttttg 53640agatggagtc tcactctgtc acccaggctg gagtgcagtg gtgcgatctt ggctcactgc 53700aagctccacc tcccaagtag cggggactac aggcacccgc caccatgccc agctaatttt 5376053760
ttgtattttt agtagagatg gggtttcact gtgttagcca ggatggtctt gatctcctga ..... 53820
ccttgtgatc cacccacctt ggcctcccaa agtgctggga ttacaggcgt gagccaccac ..... 53880
gcccggccea agcccctgat tattactgca aatttaggtt aaataaaata tttgggggct ..... 53940
tacataatat taatatgtga ctgttatatt tgtgtttgta tttattacaa ggaaacatca ..... 54000
tttttaacta ttatcaattg tctatacatt tattgaagtc agaggctatc ttatatagat ..... 54060
ttgatggttt tacaatgccc acagcattgg ttcagtaaat atatgttgaa tggttaagtt ..... 54120
tcttcaggta attgttaatg tattcaaaaa ccaaatttct ctctctttag gccagaaaag ..... 54180
gtcggaccac cattgtgata gctcatcgtt tgtctacagt tcgtaatgct gacgtcatcg ..... 54240ctggtttcga tgatggagtc attgtggaga aaggaaatca tgatgaactc atgaaagagaaaggcattta cttcaaactt gtcacaatgc aggtatagtt taacttcaga attttcctaagtcatctcag tgataaactg attttgcatt taatgctaaa aataaatatt atttgatttg543005436054420
attaccttac aagtaggaa acaacacctg ggggattcag gatgagacca gtgtttaaga ..... 54480ttttttttct ctcttgaaag aggggaaaat aaagaaggat aaacagataa aaaaattaaaaggtttcaag gtgagttatc cgttatagta gtcagtagcc acatgtgtct gtacagcatt54600
taaatggtag tgaatctgaa ttggaatgtt ttaagtagag tgtgcacatc ttgttctgaa ..... 54660
gacttagtac aaaaatgca aaatatctca atttttatat tgattacatg ttgaaacgat ..... 54720
atcatgggat atcaataata tattggatat attaggtgaa ataaaatata taactaaaat ..... 54780
taatttcact tctttctttt tacatttgtt aacatggctt ctagaaaatt taaaataaca ..... 54840
tggctcatag catggcttgt gccatatttc tgttgaaagc actgatccag agtaaacttg ..... 54900
attgagtcaa aacacccaca aaaaatggcg tgagaagata agatacatgg aatgaatgcc ..... 54960
agattatttc agtatatgca gtgggaattc ttctttggtg gttttcgttt tcaaatatca ..... 55020
cacacacaca cagacacaca cagacacaca taaaacaaca cagcagatta gctttatcct ..... 55080
tttctgcagt tactaaacaa attgctgttt tccttgatag ctgacattca gtgattagct ..... 55140
ttcattggtt aacacacagc ctaatgagct tttgcatatt tcttatttat tttagacagc ..... 55200
aggaaatgaa gttgaattag aaaatgcagc tgatgaatcc aaaagtgaaa ttgatgcctt ..... 55260
ggaaatgtct tcaaatgatt caagatccag tctaataaga aaaagatcaa ctcgtaggag ..... 55320
tgtccgtgga tcacaagccc aagacagaaa gcttagtacc aaagaggctc tggtatgaag ..... 55380
ggagatgcgg agtttgtttt aatctcacta actgtggttc cctagtttgg tgggctaggg ..... 55440
ctacggtagg agtgggaaca agagaggtta tccagaatcc tcctgtccta tcccccagaa ..... 55500
tgtcaacatt ttagaatcag gttagaattt aaaagtatta atttacacag cagaattttt ..... 55560agaattaaaa tttatagtgt aaagagacta tagcgggtct tcaaatatca aaatttccattctgtttact cctgcttata aatactcttg tcaggttctg agtaccaaat aaaagtaagt5562055680
gtttgtggaa acatcattta ttttttaaaa aataaaggag tattgtagca aaatttgtca ..... 55740
acattttttt gaagctaaat aataatcact atgacatttt ttaaagcaaa attgtcatca ..... 55800
cttattcatt gataaggaat aaggatagga tatattcctt tactaatttt tgtgcgtatg

55860
55920
55980
56040
56100
56160
56220
56280
56340
56400
56460
56520
56580
56640
56700
56760
56820
56880
56940
57000
57060
57120
57180
57240
57300
57360
57420
57480
57540
57600
57660
57720
57780
57840

actagggatg caggtacctg cctcttgatg attaaaagga tcatgtgacc aagtgctcag

59940
60000
60060
60120
60180
60240
60300
60360
60420
60480
60540
60600
60660
60720
60780
60840
60900
60960
61020
61080
61140
61200
61260
61320
61380
61440
61500
61560
61620
61680
61740
61800
61860
61920

| tataagtcag | attccactta | ggtctacact | gaccttgctc | cagggccaaa | ttcccatta | 61980 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| cccaaccagc | ctccaggcca | actgctgtgc | ccattatact | ttggcagctg | agctgatggt | 62040 |
| ttgtggaatg | tctcctccat | aaattgttaa | gtagggcaag | atttattaa | tgccttctg | 62100 |
| cttacaa | ctagtactta | gtattgtaag | gcattcaaac | ctgaatggtc | cctgtgttca | 62160 |
| agga | at | ggaga | t | aagaatattg | aaagcaagt | 62220 |
| tgat | tataa | tatagcacta | t | atatttttg | ggaaaaatga | 62280 |
| tagatact | tttaatgtaa | taaccag | aga | gggacc | cag | 62340 |
| gtaaaa | taaattcata | cttcaaattc | acaatcatgt | tagtatctgt | a | 62400 |
| atat | t | aaaa | tttatgttga | atgttatatt | taatctgga | 62460 |
| ttactttt | tgatttg | aa | tcacttatac | tgccctgagc | tacatttatc | 62520 |
| taactgctta | ttcaacctct | ctattggatg | tctaataaga | t | C | 62580 |
| caaaatggag | Ct | ttcccctccc | C | CC | cactcttccc | 62640 |
| ctata | aa | C | tagttgtttg | gctaaaaac | cttgaagcca | 62700 |
| tctttgac | ttctcttt | gttaacattg | caaaagcttc | ctaactgatc | tccctacctc | 62760 |
| catttttatc | ccaatcctgt | agattcacct | aaaaacagc | acagtcattt | tctaaaata | 62820 |
| gaaatcaaat | catatcccac | tg | t | C | acagggta | 62880 |
| aagt | tct | C | accataca | tgtggcccc | cactctctct | 62940 |
| caaatctcat | ctcctataac | cgccttt | caatacaccg | tctactccat | aaacagtggc | 63000 |
| tcactgccca | agagtagg | aggggaggat | cagaggagcc | aaatcagaac | acaaatctgt | 63060 |
| ccagaggaag | ggggtggctt | acaaattc | tacaaa | tctttttagg | ctgctggcag | 63120 |
| cctttaaggc | tttcttgctg | gtgtatgaac | agattaggta | tgcctttgcc | tcagggtctt | 63180 |
| tgcacatgct | gtttctgc | ccctgaggt | tctcatc | tccttcat | tcaggcctct | 63240 |
| gttcagatgt | ccccttagaa | aggccttctg | tttcccctcc | cagtctcta | aaatagcacc | 63300 |
| tcccctcact | tttctgttct | cttaccctg | ctcaattttt | ttttatttg | taccttccat | 63360 |
| ctctagaatc | taaacttcat | gaaagcagct | actttgtctt | ttgttcgt | caggtccaa | 63420 |
| cacttagaaa | tcatgcgtag | gagaaagtag | gcactcagaa | tttttctga | caaatgaaat | 63480 |
| gatctattta | tgtgttttta | tattaagttt | ctttcttgtg | attgaatgt | cacatcctga | 63540 |
| gtactaaatg | cagggggtat | aagtataaac | aaaactgacc | catcgctgc | cctcttggag | 63600 |
| ctgagagtct | cataaacagc | tttaaggtaa | taaaatcatt | ttctgtgcca | caggatgtga | 63660 |
| gttggtttga | tgaccetaaa | aacaccactg | gagcattgac | taccaggctc | gccaatgatg | 63720 |
| ctgctcaagt | taaaggggta | cgtgcctcct | ttctactggt | gtttgtctta | attggccatt | 63780 |
| ttggacccca | gcatgaaact | aattttctcc | ttacgggtgt | tagttatcat | cattaagaaa | 63840 |
| atgttgaata | aatatctaac | ctacgaatat | atcacatgct | ttttgtagca | acatgttaac | 63900 |
| tatttaaaca | ttatatactg | tagagcatat | agataactta | aaaccattt | gctattgctg | 63960 |

ttattcatgc tattaacaag atgcatgtag aatagttatt tagaaaagag agtataaagt gctcaatcaa cataaaacag taattgctac tgaagaaagg atgtatttaa ttgctgtaag aaagtttaga gtcactatgg ttacagaagg gagggaagac aatcctctaa aatataggtt gaaggaaatg aaaagcacat taaaaatta aggcaagaat agaataactt cagtctttat ctttaataac tttaaacttt aataatttta ataacttaaa ttttgctact gtatgaatct cttgatataa ctagatacta ttgaaccagc aggttttgat ttttggctga agtgacaatt tcttctacaa ctgtttatgg caaaagtcca caaaatgatg tagaatttga aaaaattcat gtaatctctg gtgtgtcttt tcccctcttg aaccttatcc atctttatct ttaaatcttt tctgtaagtt agtactatac taacatttct tctatctaat atatggggct tctttaagaa taaattaagc tataaatgag gaaatacata gagttataac gttgaaatat aaaccttagg agtccctctt tttctattgt ttggaatagt ttcagaaggc atagtaccag ctcctctttg taattgttgt tttaatacaa acttctttgc ctaaagcaaa caaaacaata aaaatcaagg tttagatcaa gttgtataga atgtaattac aggtgcacgc ctgtaatccc agctactcgg gaggctgagg tacgagaatt acttgaactc aggaggtgga ggttgcagag ctgagattgc accactgcac tccagcctgg gtgacaaagc aagactatgt ctcaaaaaaa aaaaaaatgc aaagaagaca gagtggctgg aataaagtga gtgaaaagaa gagtcataag tgtgttaagg tcagcattat atccagaagt agatggaaaa ccactgtagg gttttgaaca cagaagtgac 64920 64980 atgatctgaa attttgaaag gatcactata gaaactgtgt gaataggccg aagggggcaa 65040 gcatagaagg agtctgttgc agtaatccag gaggagatga tagtgtttta gactaatttg 65100 gttatacaaa aggctataag atataattaa ttctggatat attttagatg tacagccaac 65160 aatgtgttgg ttggataaga tgatggatat gagagaaagg gtatttaaag atgactccaa 65220 attcttttac ctgcacagtg gaaaaaaaa tggagttttt tttttttttt ttttttttaa 65280 agagacaggg tctctctttg tagcccaggc cagactgcag tggcatgatt acagctcact 65340 ataacctaga actcatgggc tcaagcaatc ctcccttctc agccttcctg gtagctggga ccacaggtgt gcaccaccat gtcgagctaa tttttaaatt ttttgtagag acagggtctc cctatgttgc ccagtctggt cttgaactcc tgggctcaag tgatcctccc tgcctcagcc tcccaaagtg ttgagattac aggtgtgagc tgccatgccc agctggagtt gatctttatg aattcagaa gcctgttgga gaagcaggct tgggggagaa tggagagttc tgtatgagac atggtaagtt tgtgacgtct gtattagaca tccatatgta gatgtcaaga aggctgaaat ttgcactgct aactttagta aacctttttt ataattggtt tactaaataa gttttactat gcttctccat tcattttggt cctcacaact ctatagactc ctactctgta ggaggaatat acatgagaca gtgagcatta gtctttggaa taaaggaaca gtaaccagtg caatgtgaca ttgcacaata tgacacagac cctgtggtat gggctcatgt gctttgatgg acaaatatcc 65400 65460 65520 65.580 65640 65700 cgcatcacat ttgacgtgga agacacaatc ctggaggcaa ctcagtcttt tgcagcagaa

| cccagagcag atgtattcag ggcattctgt attgcagttt ttcttcctgc ctctcaaatt | 66060 |  |
| :--- | :--- | :--- | :--- |
| cctctggtgt tatgacctgc cttaggccac agtgagttcc tatatttgta aaattggtgg | 66120 |  |
| tcactttttc cctccatagt gctgtttgtg aggccctttg ccatattcac tgtctctaat | 66180 |  |
| tgcctgctgg ggtcaacctt ctgcttttca cttgtttcct caaagacacc tcaaacttgg | 66240 |  |
| ccctgccaag atggctgcac tcaccttata tgagccactc aagcaaact gttttccaga | 66300 |  |
| acttaaaaca gtggctaaaa aggaaagacc agggaagagg aaatgagcag gttggcaaac | 66360 |  |
| attgtcaacc taggaaagcc caatgtagct gttatcacag attgactgag agaggactgt | 66420 |  |
| tggatctact ttacaccact agctgacctg tttttggctg acaggtttta gttcctcccc | 66480 |  |
| tcaaccctag tctttacaat gaaaattttc tggctactaa tctgtgctct tccattcttc | 66540 |  |
| ttagtcccca tatttctata gactcctact ctgtagaaat agatgtagga cagtgagtgt | 66600 |  |
| taatcttcag aataaaggaa tagtaatcag tgtaatgtct gaagtgctaa atcaaatgaa | 66660 |  |
| gccgaggggt taaatctcct tgcagtttaa ggtcttagac ttcttagtta ctctttttgg | 66720 |  |
| tacattaaag aatttgccat ctttgatcca tcttatttat ttctgtggcc cttcactact | 66780 |  |
| atccaaagct gcacatttat atgtctgaaa agttaaatag ccaggtaatt cctctgccat | 66840 |  |
| gaactcacac ctagaagtga ttctaggtga tattctaaga gcttctctca caagcccata | 66900 |  |
| agtcagtagc tgcagtggga aatgagctct gagtagaact tgtagatact gagaacagga | 66960 |  |
| gcaacttata tcatccctgt gccacagccg ttctccagct tgcatccttt gaccttcaaa | 67020 |  |
| ctaaagatgt gaaatgcaga gagctggctt tgaaaaccat tccagtctga tttatcccaa | 67080 |  |
| atttggcaaa tcatcccttt aatataagct caatgattgc atcaacaaaa tatacaggtg | 67140 |  |
| ttctattatt tatgatacct ctttcagccc tggtgcagct ccatggtatg gaatttaatg | 67200 |  |
| tatttaaagg ctactgatag ttatagcact cttttcctaa atactagtct ggtgtttata | 67260 |  |
| tcagacaggt gttaaatgat tttgtaggga ttaatgtagg gaattataaa ggattctatt | 67320 |  |
| gttactagaa aggggtcccg atgcagacct tgaaagaggg ttcttggatc ttgtacaaga | 67380 |  |
| aagaattcaa ggcgaatcca cagagtaaag tgaaagcaag tttattaaga aagtgaaaaa | 67440 |  |
| attaagaata gctacttcat aggcagaaca gcagcatggg atgctcagct gcttatactt | 67500 |  |
| attgttactt cttgattaca tgctaaacaa ggggtggatt attcatgagt ttcctaggaa | 67560 |  |
| aaggggtggg ctcttctcag aactgagge tcctcccatt tttagaccat atagggtaac | 67620 |  |
| gttgccatgg catttgtaaa ttgtcatggc gctggtgtgt ctttgagcat gctaatacat | 67680 |  |
| tataattagc gtataatgag cagtgagaac aaccagaggt cacttttgtc accatcttga | 67740 |  |
| ttttggtagg atttggctgg cttctttacc acagcttgtt ttatcagcaa ggtcttagcg | 67800 |  |
| acctgtatct tgtgctgacc tcctgtcgca tcctgtgact tagaatgcct aacctcctgg | 67860 |  |
| gaatgtagcc cagtagatct cagccttatt ttacccagcc ccttttcaga aaaggcaggt | 67920 |  |
| ctggggccag ggcaggcatt tggaatggct tgagggcaca gaatatttcc agggtagagg | 67980 |  |
| agatgtgctg gactaaatta tagagtgatg gactaaactg actttgtagc ttgacttttg | 68040 |  |
| Page 34 |  |  |

gtttcagagt tggaaatctg gtttacagtt ggacattata cagtggtgga ttaaattcac
68100 tttgcagttt gacttctggc tttagagctg gaactcagtg cacactagtg aaagtcatgc 68160 ttcagactcc ctctggaatt caaggggaag ataataatgc gcgcttacta tcttaaatta 68220 aattccataa ttttcagctc tgtatttttt ccaaattaaa cattatatct caaacagacc 68280 cagatatatt tgaatattat taatgacaaa cgttaggctt aaattacaaa taataatata 68340 cctaacattg gaaatttcca tcattcctag tttgtcagac tcctttatct tgctaatttg 68400 cagatattgc tttagtaatg ttgccgtgat taatgaaggt tttcttggta ttaaaagatc 68460 caagagata ggaatatgta attgaactct agattgttga tattctactt tcagcattct 68520 gaagtcatgg aaattcttac tgtagaaact caataaactt acaagtagac ctttactttt 68580 tagttcatta ctgataaaat aatgaatata gtctcatgaa ggtgagtttt cagajaaatag aagcatgagt tgtgaagata atatttttaa aatttctcta atttgttttg ttttgcaggc tataggttcc aggcttgctg taattaccca gaatatagca aatcttggga caggaataat tatatccttc atctatggtt ggcaactaac actgttactc ttagcaattg tacccatcat tgcaatagca ggagttgttg aaatgaaaat gttgtctgga caagcactga aagataagaa agaactagaa ggtnctggga aggtgagtca aactaaatat gattgattaa ttaagtagag taaagtattc taatcagtgt tattttgtta ctccctactg cttactatgc tctaagaatg tgtttataac cattcctcaa agcaatcttt ttcatgctta ttcagtaaat tagaaactta cagaaagtag caaagccagt tcttggactc aaaaactgat aattaacttt aacagacttt ttcagttttc aggccattgt cttcacactg ttcttccttc ctccccactt tcctccttcc cttagttatt ttcttctttc ttttctctca ctttcactct ctctccactc cttccttcct tctttccttc cttccttctt tccttctttt ctttcctttt ttccttcctt cccttcttcc ttcttttctt tttctttctt ttctttcttt ctttctttct ttctttcttt ctttcttgct ttcttgcttt cttgctttct tgctttcttg ctttcttgct ttctttcttt tctttcaagc ttaaatccat tcctttattg aggaaagtaa gcccatttta tttgtacatg tgagggggga gattaaatat ggaaaaatgc taggggtatt tattatatct gttttaaatt actaccactc tttctttttt tttatcatgc tcctcccttc atcctatttc tgtttatctt tacccttttt ttacttcttt tttttcccct gcatacttgt cttttttttc ccattattta acaaatgctt atgtggcatt tactgtgttt ccaggcaaat gtcttattcc ttatagcaac catatggggg tctattatct catttttcta gtggggctga ggcacaggca ggtcatggct ctggacttta gagctgatag gtcttggagc tgggattcaa gctcagacag ttgtgctcca aagttgtttc tctttctgtt ataaaacaaa gagttcctct gatggcagaa tgcagtctga tatcacatga tctgtatcat agtggaaatg agaggtcaga gcagggctga ctgccataac taacatttag gacagggata tatgtgtgat gaacattctg agattcccag gagttagggc agggactcac acagatcaga gtggctctgg ttgtcagtag gcccagctac ctcaccaagt gaatgatgaa68640



Page 37

| cggaagaaga | tataggagaa | gcagtgccag | aaaactaact | ttacgcaatc | tggcaaaaga | 74220 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gtttcggtta | ttcaggactg | cttttaactt | cttttatgac | atggaccctt | ctgtgataca | 74280 |
| ggcactgaaa | ctgaagcaaa | Cagtggagga | aggattggta | ccatattgaa | acatttttag | 74340 |
| agaaataaaa | aaggaaaaag | tcagacagag | accacaatgc | atgtcagtca | caccaggtgt | 74400 |
| gcctgccttt | cccgectccc | cttccacttc | ctctgcctct | tcccctctgc | cacccctgag | 74460 |
| acagcaagac | caatccetct | tcttcctcct | tttccttagc | ctattcaata | taaagatgat | 74520 |
| gaagataaag | acctttaaat | gatccacttc | cgcttaatga | atagtaaata | tattttctct | 74580 |
| tccttgtgat | tttcttaata | acattttctt | ccctctagct | tactttatta | taaaaataca | 74640 |
| aatatgtatt | acatataaca | tacaaaatac | atgttaattg | actgtttatg | ttattggtaa | 74700 |
| gacttctggt | tgacagcagg | ctattagtag | ttaagtttct | ggggagtcag | aagttatatg | 74760 |
| tgaattttca | actgcacgag | gggtcagtct | ccctaacccc | catgttgtcc | aaggctccat | 74820 |
| tgtagatgct | tttttacttt | tcagaaataa | aagccataaa | cttgtttatt | tgagggtagg | 74880 |
| aagagtaatg | tcaggaggct | attttttctt | tctagadaca | tacattttta | tttcttcaaa | 74940 |
| attatttagt | acacaacagt | ttccgaggga | aactcgatgc | tatttgttct | tggattagaa | 75000 |
| attctctctc | ctgatggtga | ttactgagtc | tccatttaaa | actcttggaa | taaacaaagc | 75060 |
| tgtgaggatg | cggggtgcta | attaagcctt | ttcctaattg | cttcattgtg | cgtcaagatg | 75120 |
| aagacagt | tgaag | atactgtttt | actcaccact | tttaggtctg | tgactcaaac | 75180 |
| tcccactttt | attcggctat | atacacacta | aaaagcaatg | acatttacaa | accaatctca | 75240 |
| gaccagacac | tcctgcctta | gaacatggtg | cacagaaaat | atttcttaaa | accattacac | 75300 |
| tgaaatatac | agtaaaatct | gtttttcagc | agacattgtt | atagtctgtt | gaattttctt | 75360 |
| actaatctag | aaaacctgtt | tgttagaatt | ctgataatta | gaaatatttc | tせtttttttt | 75420 |
| tgcttgtgaa | acttcagctt | ttattttatt | tttttatttt | tattatactt | taagttttag | 75480 |
| gatacatgta | cacaacgtgc | aggttagtta | catatgtata | catgtgccat | $g t t g g t g t g c$ | 75540 |
| tgtacccacc | aactcgtcat | ttaacattag | gtatatctcc | taatgctatc | cctcccccct | 75600 |
| tccctctacc | ccacaacagg | ccccagtgtg | tgatgttccg | cttcctgtgt | ccatgtgttc | 75660 |
| tcattgttca | gttcccacct | atgagtgaga | acatacggtg | tttggttttt | tgtccttgcg | 75720 |
| atagtttgct | gagaatgatg | gttctaatgt | gtctgacaaa | aggattgtct | ttgggaattt | 75780 |
| gaaggtgaat | tttttctcct | caccttttgc | tttctgcact | tttacgattt | tctaaagtga | 75840 |
| ctatatatca | tttttataat | gtgtaaaaga | agtttataca | atattttaaa | ataaacctgc | 75900 |
| cattttccta | attttctaag | tatcttgtgg | taaacataat | tcaatcttct | tggcctgtca | 75960 |
| gtgtaagaat | aatattttaa | attttatttt | taaataagtt | tttgtttcta | agaatgttac | 76020 |
| taattttttt | ttttacttct | gatagatttt | gatattagtc | ttcaaaactg | acttaatgtc | 76080 |
| ttatgaaatg | cttgctgtta | tgtttgaagt | taggtaattt | atgtaagatt | cagtgaagaa | 76140 |
| taagtggaat t | tccatgttta | tgattttaag | ctataaaaca <br> Page 38 | ctctaaatta | aatgtgtctt | 76200 |

tattagaatc tgttctgacc agtgcagagg ccaaagagag gaagaacatc ttaaacaata ..... 76260
aagagtcatt tctcttggtg acttataatt ctggaagtta tttctcttaa aatcatagca ..... 76320
ttaaaaggga ctttagagac cctctagtcc atcgtcctca ttttgcaaat gaggaaaatg ..... 76380
agacagcatg ttggttcaag gtggtgcggc tgatgtaggc tgaaatctca tcttgtacac ..... 76440
tggtgttctt tgctttttcc atatcccttt actcagactc cagaggtgat gaaggatgta ..... 76500
tgtttcctaa tcagattgcc ttgttggaag taacatttga ttacaacata attgaatgat ..... 76560
ggaaactttc tttttaagat ggagtctcac tctgttgccc aggctagagt gcagtgacgc ..... 76620
catcttgact cactgcaatc tctgtctcgc cagttcgagc gcttcctctg cctcccagta ..... 76680
gcatgggatt aaggcatgtg ccaccatgcc cagctaattt ttgtattttt agtagagatg ..... 76740
gggtttcacc atgttgatgg tggctggtct caaactcctt acttcaaatg atccacctgc ..... 76800
ctccatctcc caaactgctg ggattacagg catgagccac catacccagc ccaaaacttt ..... 76860
ctggaaaaca gattgatagt atgtgccaca ttccttaaaa aattaaaaaa attaattcaa ..... 76920
gccaggtgta gtgccatgtg cttgtagtct tagctacttg gcagactgag gcaggaggat ..... 76980
tgcttaccca ggtgtgtgag gctgcagtga gctatgatga tcacacctgt gaattgccac ..... 77040
tgcactgcag cctgggcaac agagcaagac cccgtttcta aaaaaaaag ttagttttct ..... 77100
ttgacttatt aatttcacct taagaaattt tcctaataaa ccaattcaat atggacaaat ..... 77160
gtttaggtac aaagatgttt atctcaccac tatttttaat aaaaaggaa ttgaaaacct ..... 77220
ggctcaacaa taaaggaata cttaattggt tatgatatta aaggactcat tacacatctc ..... 77280
attattaatg tgtatttaat gaacttggaa aatgcttttg atatgaaggt aaaaataatg ..... 77340
atatagagct aaatatagag tttcattcca atctttttaa atatatttat gcacttagga ..... 77400
aaaaaacaat atggaaatgt gtaaaatata cttttttttt aaaaaaagg acacatttat ..... 77460
tcagcattat gatcagacta ttacatttaa caatcaacag tatgggtgcc aaaaaaaatc ..... 77520
tacattaaaa ccctttgttg taatgcttta cactttccac agaacagaaa ctaaaagaat ..... 77580
ctgttacaca attagtcaca aatatagtcc tcgagttttt tacccataca catgagtatt ..... 77640
tgtctaaaac atgtcttctt gtagcactta ggccctgcca ccactgtgct tgtctgagtt ..... 77700
cacaaatctg ttgtaaactg tagcttccct gtcacttctc tggctcttat ctcctgctaa ..... 77760
gatttgtttc ctggcagtaa tttaaaatct tctgccactg ctgtagctac tgctgctact ..... 77820
ggaactgcca tagccacctt ggtttcatgg tttggcaaag tactggcctg taccagcata ..... 77880
ggggccagag cttctgcctc caaagtttcc tcccttcatg ggtccaaaat gtaaaactaa ..... 77940
ttgttgtaat tgccaaaatc attacaccac ctccaaaatt gcttccatga ttaccaaatc ..... 78000
cattatagcc atccccactg ccactatatc caccaccacc acagctgcca ccaaagccac ..... 78060
aatgaccact gaagtttcct ccacgaccaa agttgtcatt cccaccgaaa ctacctccac ..... 78120
gaccaccacc aaagttccca gagctgcttt gcctctttgg ctggatgaag cactcaccat ..... 78180
ctcttgcttt gacagggctt tcctaacttc acaagtgtgg ccattcacag tatgggtatt ..... 78240

ggtctcctaa gtcaagaaca tgcctgtttc ctctatgtca acccatcatg gcccctgcac ..... 80340
cttgtaggtg ttgggtaact gtgtgcagaa tgaatattta cgtagagtac ttccatactgtgtgtccaaa agtggggaga aagggagata acttttactt attgatccct aaccggctctttacagtcat tggattattt tctctttaca tcaacacttt gaggtgtagt taattctacattaaagataa agacactgag tctcagaggt taggcagttt cccaaatttg cacaactgttaagtgtaaag tgaggaccaa acttagttct ttggaaacga aacccatttt tgttggtcatgcctctgtgc cctactgcca acctatcaaa agttacattt taaggactga gaaatgaaagtggaatgagt tttcaaatgt cttcttttca gaaactcttt gaggaaagca cacatctttg 80760gaattacatt ttccttcacc caggcaatga tgtatttttc ctatgctgga tgtttccggt 80820ttggagccta cttggtggca cataaactca tgagctttga ggatgttctg ttgtaagtat 80880tgggctatta tttagttaag ctctaaaaat aaagctggga tgaacatgct tcatgtctag 80940ataggaaacc ctactgtgaa gcctcatgaa gagattctgg tgattcctaa atcggcattt 81000ctgcctctga gtcttcatgt gccaccattg aagcaccccc tttcatttgg aggagcagta 81060acttctttcc tcattgctgg ctcacacata gttgaccttt tcaaatctgt gactgagtgg 81120agtgattcca ttcggagatt ttgagaaggc cttggcattt gggaagaagc ctagccctga 81180gcaaggagtc tgactggctc cttttaaagg actttcttac agagcaagta aaatacagat 81240gtgttgtact aagttctgca agcctttggc aattccagga tatgtttact ttcccttgat 81300aagagaggaa ttggaaggta agagccaaat gaattcagaa atgacaaagg aaaagttata 81360ttgggatttt tctgctacat tgttctgaat gtagataatt gtacctcggt cagaggaaga 81420taagcctgaa gcaattatac tacaaaagta cccaaatatg caatttggtg gtcaaaagta 81480tgtgtaatct ttctgagctt ctagatttgg aggtgggtaa gattctgcct cttgatagca 81540taacataatt agtagcgata taattttata tttaaaccag aatcatataa gcctggcagt 81600ataatgtgtt agtatagtat ttctgtcctt tttaaacatt gagttgttca tgcattacag 81660tttgctcagg atgaccccca aaaagacatc taaatttcca ttaaagatgt acattggaca 81720aatgatatgc aggtgctatt tgtgattatg gcctagagat caaaaccaag tatcactggc 81780
attggggctt tgattagata attatttgat atattgcttt actccaaaaa attgaaittga ..... 81840
tgaaagttgc tgacattggg gcatttaggt ttgcaaaatc aaccttccaa gttaaggaaa ..... 81900
aggaagacct gtctgaagaa cagtgcatta gaaaaaggtc ccataggttt aaatgatcac ..... 81960
aagtccaaga ttaactatcg gtattttatt tagtgctagt taaacaaaca aacaaacaaa ..... 82020
aactaatgac acaatatgct acacaaacat aggcatagag tccagaggaa agaaagcgat ..... 82080
gaattccctg tattctgggc gggttcccaa gtattatgtc tgttccaggc ttgtgtttaa ..... 82140
aaaaagcata ttaataaatg tgtctagaga agggcagtga aaggagttat caaagaagca ..... 82200
aaggagtttt tagcaatatt tcaagacttg taggcctgtc ttatgtaaaa ggaagtaaat ..... 82260
tttttcttca gtttgaaaaa gaactgttta acaatagaca atcaaacatg ctacttccta ..... 82320
aaacagagga tggaagccaa tttaggggag gtgtccaggc acgaacatgg agagctggac ..... 82380
ttgatacctg taaggtcctt cccaacttta agttgctgtg attcccatgt catagataag ..... 82440
aacgtcaatg catcttaaga gcaacatgat atctggctct gtaagaaact ttcttttggt ..... 82500
tgcacaattc ataggttttt aagaatctga tgtaattcca acatcactga ctgtatccgt ..... 82560
tgttgattac cacaataatg ctgtgtaaca accagccact aaacctcaga gacatacata ..... 82620
tttttgccca caaacctatg ggttagctgt gaagttctac tgatctggat taggcatagt ..... 82680
ggatctcggc tggacttact catgtgtctg tagtggattg tgtctgtagc tggttggttg ..... 82740
ggattcagac agctctcctc catgtgtgcc tcatcctcca gcaggctatc ctgggtttga ..... 82800
gacagtggca gaattctgag agagagagag agagaaagag agagagaaca catgcatgca ..... 82860
tgctcacaaa gcatataagg ccccctgagt cataggcttg gaactggcaa aagtgatttc ..... 82920
caccatatcc tgttggccaa agcaagtcac aaggccagcc cagattcaaa gggtgctgca ..... 82980
aaggtcacag tgcaagaaga tgaggataca gagagaggta tacagaggga aaaaaatggg ..... 83040
ccattttgca atcaatctat cacatagaca tgaacttata aggaaatgtg tttgttttat ..... 83100
ttttaacatc tgttttataa ctattagagg caaagctctc tggttataga agtgtcaact ..... 83160
tttggccagg catggtggct cacacctata atcccggcac ttttggaggc tgaagcagga ..... 83220
ggttcacttc agcccggtag ttcgagacca gcctgggcaa tatagggaga cccccatctc ..... 83280
tacaaaaaat aaaacaatta gctggtgtgg tcatgcacag ttgtagtccc agctactaca ..... 83340
gaggctgagg tgggaagatc gcttaagccc tggagatcat atctttagtg agctctgatt ..... 83400
gcactactgc actccagcct gggcaacaca ggaagacccc acctcaaaaa aaaaaaaaaa ..... 83460
aaaaaaaagg agtgtcagct ttctagcatt gtgatggtaa tgctgtgcac atgttttgtg ..... 83520
tttgtgcttt ccagagtatt ttcagctgtt gtctttggtg ccatggccgt ggggcaagtc ..... 83580
agttcatttg ctcctgacta tgccaaagcc aaaatatcag cagcccacat catcatgatc ..... 83640
attgaaaaaa cccctttgat tgacagctac agcacggaag gcctaatgcc ggtgagtttg ..... 83700
atgtttcaac tgtttgatct actcctgact cctgaatgaa agtattttaa gtggaaactt ..... 83760
aataaaattt gtactttcaa atatgctgat gataaaataa aacttcctag atcatagatt ..... 83820
cctttcaatt actgctaata atatacatca acattcagta cttttacgta gcaaaggtta ..... 83880
tagggaaata ggaatactgc tcactttata agcaaaacct attaatcaga ttttttaaaa ..... 83940
acaatttttt tttagagaca gagtcttact ctgtcatcca ggctggagtg cagcagtatg ..... 84000
atcatagctc actgcagcct tgatcttctg ggctcaagcg atcctcctgc ctcagcctcc ..... 84060
caagtaactg ggactacaag cgtgtgccat catgcccagc taatttttta attatttgta ..... 84120
gagacagggt ctcgttatgt tgcccaggct ggttatcaga ttttattgta tgtaagttac ..... 84180
tgtattcctg aggaacagat ttgagttatt gtagctgtat tgcatattca tattgtctta ..... 84240
acaatacatg ctatgaaagc ttttactctt ttagatctca tttattaaat tctagcagtc ..... 84300
tgaggtcaag cacagtggct catgcctgta atcccagcac tttgggaggc caaggtgggt ..... 84360
ggatcacgtg aggtcaggag ttcgagacca gcttggccaa tatggtgaaa ctccatctca ..... 84420
aataaaaata aaaaaaaaaa gattagctgg gtgtggtggc acacgcctgt aatcccagct ..... 84480
acttgggagg ctgaggtacg agaattgctt gaacctgggg ggtggaggtt gcagtgagct ..... 84540
gagatcatgc cactgcactc cagcctgggt gacagagtga gactctgtct caaaaaata ..... 84600
aaaataaat aaataataaa attaaaataa aataaattct agcagtttga agtgaagcca ..... 84660
attgtaacac aaattaatta tcttctgaca cctggtaatc gagagagtta gctatacact ..... 84720
ttattttcag tattgcagca ttcaaattta ctgttattct tctcattgca gaacacattg ..... 84780
gaaggaaatg tcacatttgg tgaagttgta ttcaactatc ccacccgacc ggacatccca ..... 84840
gtgcttcagg gactgagcct ggaggtgaag aagggccaga cgctggctct ggtgggcagc ..... 84900
agtggctgtg ggaagagcac agtggtccag ctcctggagc ggttctacga ccccttggca ..... 84960
gggaaagtgg tgagcacact ttcacattta gctcagttca ggttttcatc atccaaatgt ..... 85020
ctgaatgtat ttaattctca actataagcc atgttttttc aaacctttaa acaacagtcc ..... 85080cacttggata aagtctgaga gcctaaatat ggtctccaag tggtgtcatc tgtcccagcc
aacttctcca ggctcccctc aacactaccc ctctaccctc ctttgcaagc accctttgta85140
ccacctgtct cccttgctgt acttaggttt cagctgactt gaaaagaacc aacaaaaatg ..... 85260
gaagtagcca gaaagataca ggacaggtgc taggagtgag atgaaagcca aggatggaga ..... 85320
ctgtttcaaa gatggtggtc agcaatgaca gatggtatag agagattggg taaggaggag ..... 85380
actgagaaga catgatagaa tctagcaatg aatgggctat ttctgacatt ttaaaatatt ..... 85440
cttagtgaag tagtgatggt aggaaacagt taagggtggc tgaagtatga ttaggggaag ..... 85500
aaatggagat gtaagtgagc atagattatc aagaagatgg aaagtaaaag aaaggcaaat ..... 85560
aaggacagca cttgagtggg gaggcagagt ccagggaaaa tgtttagggg ttgaactctt ..... 85620
gagcattttt atgactgggg atgagacaaa atcattgatg gagagagggc attgaaacat ..... 85680
catctgagga aaaaaaagta gaatcaacaa atttgggaaa tgtaaaaaga acatggaagt ..... 85740
atgccctttt gtcttagttg ctggggatga gggaagtgtc tggcagggag tagtgagttt ..... 85800
aaacgactat gatgggggat aggaaagaag acaaactagg aaggaataga attaagagca ..... 85860
agacattttc attatgtagg catgtaatta gttgtactga gaagtgatat ttaaaattct ..... 85920
taacaatggg aacagcatgg acactgacca atcagaatgg atgcctgcta gtacaagtca ..... 85980
gatggatgcc cacccgctgt gtatggacgc tggtagcatt taatgactca aatgaaaaga ..... 86040
gagttttcaa tctttagtct aggactcttt ttacttcctc acagcggtca ctcttactca ..... 86100
taaatgtttg ttattgtatt caaagacagt tatgccttta ttcagtttac tcttaaatca ..... 86160
ttacaagctt gttccctact tttactctct tgctttatat tttctgtgaa ctctgattgt ..... 86220
gtccttcaat tttgtggtga acttaaaaac aggactctaa ataaagacgt cgctgacagc ..... 86280
caaagcaatt agaggaggaa atgttagctc tcagacactg attatggccc tcaacagaca ..... 86340
gttattaagt ggtgtgttaa agattgtgct ataatgaatt gtagggcatg atatctgccc ..... 86400

aaatatcttt cagttgtgag gtattagtgt taccaaaacc caacctagag tgtttagatc ..... 88500
cagtcttggt tttcccccaa ttactaattg tgtgaccttg tataagttac ttaactttgc ..... 88560
tcatttcctt atcagtaaaa tgggaattaa catcacagtg accttataag tgttgctatg ..... 88620
aaggttaaac aagagactgt atgagacatg cttatcacag aacctacaac gttgtcagtg ..... 88680
cttgattttt aaaaatcaat tattatttca cagtaaatat gcatagaaga atacacattt ..... 88740
aacatttaca tttttgcgct aatgaaagca agcaatagac ttaacagtct acaatagaga ..... 88800
tagccaatca ttttaatctg gcctttcatt atttcttacc aataggttta ctttcctagt ..... 88860
tatattttta tttaggccag gattttttga gcacttctca tagattcccc agagttcctt ..... 88920
ggtctttcct cagaccaaca ccattcctca gccaaagaag ctctgcttct ttgatctgtt ..... 88980
ttacattcta ggcatctaaa cttgttttac ttaaagaaag aattcttcag tcaacacagg ..... 89040
ttttaaatag tttgcaattc tagggtatta tgcggggacg aggaggccca agagtgtatt ..... 89100
cagtgtaatg aagaaactca ttataacttg ctgagatcat tcagatttgc ctgtgattat ..... 89160
attagttagg cagtgtctgt tttcctccca ggggtataaa cctggaaaat tacaacaaaa ..... 89220
acaaaagcaa aaaccaaact cctcctttcc ttaaatttcc acttctcaag tacagtgttt ..... 89280
tatctaacaa gatctgctgc ttagccacat ccttgtttgg ctttcactgt ctctctccac ..... 89340
cctctacttt ccaccttcat ttattaacat tgtaaggaag cctggagcat acatgtgtgg ..... 89400
aatagtctcc atggcaactc agcttggaac taataaggtt atgggagagc ttccatccct ..... 89460
gcacctgcca gtgtcacaag cagaagccat gttcctgtag caaagattgt actctactgc ..... 89520
taggcagctg tccccttgag ccacccagcc aggcacatgg gatacagaga gtatatggct ..... 89580
cagcacgcac tagtcactac tattagaaaa gctcaaagct gagtcactgg tgccttcttc ..... 89640
agaagggatg aacagctctc tcacttgaat gccagaaaat tatcttgcaa agcagaccta ..... 89700
tctgatagac atatttgcat cagagtaggg cttgttatca gcaaggctgt aaggtgccct ..... 89760
ccccagtctt ctgcaggata atccagggag ccagattata gagaaagggc agccctctgt ..... 89820
tcctactcat ctggctcagt attgaggatc tccattcacc ctttcctacc cctgttcacc ..... 89880
tatccatccc ctgtagttcc tgacctgcaa agctattatg tggtcatagt tgttatttat ..... 89940
ttccatgcta atcctgggca ctgtttctct gaaaaatgga gatttaagaa taaggctgtc ..... 90000
aggatacatc tcagaagtat taggcgttgt attagtgtgg ctgctataac cacttgccac ..... 90060
aaattcagtg gcttaaaaca acaccaattt atgactttac acttctggag gtcagaagtc ..... 90120
taaaataggt ctcagtgggc tcatatcaag tgtcaggagg gctgtgttcc ttctagaggc ..... 90180
tccaaaggat gatctgtttt cttgcttgtg gccccttcct ccattttcaa aaccagcagt ..... 90240
ggctggttga gtctttctca cactgcattt gctgatactc ttctccctcc ttcttcttca ..... 90300
gttaagagcc cttatgatta cattagtttc accaagataa ttcaggataa tcttatttta ..... 90360
cagtcagctg attagcaacc ttacatctac tactttagtt tctttttgcc atgtaacata ..... 90420
acacattcac aggatccagg gattaggaca cagatgtctt gtgggagagg gaacattatt ..... 90480
ctgcctacca catgcataca tcagaaacca tggttgaaac acaggaaaca tgacagttcc ..... 90540
tcaaggcata caattatgac cttgttgggt taaccttcac tatccaaatt ttaatcacac ..... 90600
aaacttttcc ttaatctcac agtaacttgg cagtttcagt gtaagaaata atgatgttaa ..... 90660
ttgtgctaca ttcaaagtgt gctggtcctg aagttgatct gtgaactctt gttttcagct ..... 90720
gcttgatggc aaagaaataa agcgactgaa tgttcagtgg ctccgagcac acctgggcat ..... 90780cgtgtcccag gagcccatcc tgtttgactg cagcattgct gagaacattg cctatggagacaacagccgg gtggtgtcac aggaagagat ngtgagggca gcaaaggagg ccaacatacatgccttcatc gagtcactgc ctaatgtaag tctctcttca aataaacagc ctgggagcatgtggcagcct ctctggccta tagtttgatt tataaggggc tggtttccca gaagtgaagagaaattagca accaaatcac acccttacct gtatacaagc atctggccac acttcctgtttgggttagtt gttaccttta cctgatcacc tgaccctcct tgtgaggaag ggatgaaagtgttcgaccac ttcaggttta ggagagagga acatttctgg gataggagaa ctggaacaattgtcttgatc caaagctata ggcttgaggc tccacctttg tcagccttag gggtaagtacaatatctgga aagcctttca ctttaagtcc aagtacagag tctgggtccc cacctgcacatgctgcttcc ggcctgctga ggaagtaggc atgactgtct ctccccatgt ctctccccta90840909009096091020910809114091200912609132091380
tttctcttcc ttcttcctcc ttgcagctct ctcccaagcc tcaaaactca ctgtggagtc ..... 91440acaagtccagtcc agagggctga acccatgcca gaaagagtgc tctctctgaa taagggttga
accattaaag tgaccatggt atcatgtaat ctctcatcgt catacatttc atgggcccca ..... 91560
aactttgttt acagtgaatg gttgaaatct tttctagggc tccctgggct gtggtcctat ..... 91620
gggtgtcctc tgaattcctc ctaggaagtt ggataaatgg aacctctcat cctctctcaa ..... 91680gccctcttag gtcctgtgta ttagtttttg tgttttgttt tgttttggaa tacaagttga91740
ctattccttg tccaaaatac ttgggaccag aagtgttttg gattttggaa tatttgcata ..... 91800
tacataattc acttatattt catgtatacc ttatacacat agccagaagg caattttata ..... 91860
caaattttta ataactttga gcatgaaaca aagttttgac tgtgttttga ctgtgacctg ..... 91920
tcacatgaga tcatgtgtga aattttccac ttgtggcatc atgttggtgc tcaaaaagtt ..... 91980
tcagaatttc agattttgaa ttttttggat tatggatgct caacctgtat taaaagtcca ..... 92040
aaattagatt ttttccaacc tttattttag acttaggggg tacatgtaca ggtttattat ..... 92100
acctgggtat attgcatgat gctgaggttt cagggtacaa atgatcctgt cacctaggta ..... 92160
ctgagcacag aacccaatag gtagtttttc aacacttacc ccccttccct ccctacacta ..... 92220
atagtcctca gtttctattg ttgccacttt tcatctatga gtacctgatg tttagctgcc ..... 92280
acttataagt gagaacatgc agtatttggc tttctgttcc tgcattactt tgcttgggat ..... 92340
aatggcctcc agctgcatcc atgttgctgc aaaagttagg tcctatattt tgaagtgctt ..... 92400
ctcaacctag gtaatctacc tcacccatta tcaccagtta tgctctttat tctaaggaag ..... 92460
tgccccctaa aacaaagctc aggagcctca acccggcggg gaagacagtt tcctcacgag ..... 92520

| gcaggcaagc | aacaccaggt | ggctctcttt | cccaagattc | ccttcttcca | taggctcttc | 92580 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tttggataac | tgctgacacc | aatctcatta | agtcccaggg | aggtggcctt | gtctcccact | 92640 |
| ctcctttcac | agtctagccc | aactagataa | ttgaccettc | tccaggctca | aacatttaaa | 92700 |
| acagaaagct | ccaatcccct | ttactgatga | gaggatagtg | gttctcaagt | atagacacgg | 92760 |
| atcagcatca | cctgggaacc | tgttaggact | gcagatccag | caatccacgt | tttaagagcc | 92820 |
| caggtaatac | tgttgcaggc | tcaaattggg | gaaccactgg | gctaggggtt | aataaaacac | 92880 |
| aaacacagag | agtaactatt | tctatcttt | ttaaatatac | tttgcctctc | tggcatttgt | 92940 |
| ctaccacagg | acaaatcctg | gtctctaatc | tgtcacacag | aacacacctg | tagacaccat | 93000 |
| aattctcttc | ctcagt | agaccaggag | agaaaaggta | aaacccctaa | gtcctattac | 93060 |
| ctaaggatcc | acttctccta | agtgattaat | actagga | acctcaccct | gtaggtggct | 93120 |
| ggtttagcta | cctccttgta | agactgccgc | ac | tctgcctcaa | caggatcagt | 93180 |
| actatatcaa | gagtgcaact | agtaaaccct | ccactcacat | gagatacact | tcagtgctcc | 93240 |
| tgccttctga | gacagccaca | gctctggggt | agagccagtc | ccaagcatct | gccaaaacct | 93300 |
| catctggcgt | cttctccctc | tatttacagg | gggaaaccca | gggctgagtg | ccacaaaact | 93360 |
| agggttgcca | gatatagcaa | taaaaatag | gcaaccctgt | actaaaattt | cagcctggaa | 93420 |
| ataataaga | agt | tctaggttac | aagtgcagtc | acatcctgtt | tctgctctct | 93480 |
| cctgagta | Ct | tg | atcatagtga | tagttggtg | acaattatcc | 93540 |
| attgtgggaa | acaacacggg | aagaattttg | ggagc | aacaacaaaa | atgtttaaaa | 93600 |
| aggagatgtt | cattttcaaa | gttagatatc | aaaatggaat | ccattttgat | agcatccatt | 93660 |
| aaaattgttt | aagtgttgga | gatactttca | taaaaacaga | cactaaaagg | ctgatatccc | 93720 |
| aagaatgtct | ttcctcaatt | taaaatcatg | cagagagtgt | cacctgctgg | ttacaagtaa | 93780 |
| gatgagcttc | aattatttta | ttatgaatac | tttgttcaaa | tcaaatttcc | tttctcctgc | 93840 |
| caatctttgt | cccaaatata | atgaaatcag | aagatggaag | atactaacat | tgcaacaata | 93900 |
| ttcttcttca | ttaaaatgct | gacctgttta | cactaaaaca | tatcatctta | aggcccaaac | 93960 |
| atttattttc a | aaatagatag | taaacctgcc | ttttgggct | tttttactgt | tctttcctta | 94020 |
| attattcatg a | aaccattctt | gcttctgaa | cctaaatttt | tggagtatat | agaatcgtct | 94080 |
| atcctacttt | ctataagcaa | gctgttagaa | tttactttc | agttctactt | tcataacaac | 94140 |
| aaaaccttat | ttacagaaat | atagcactaa | agtaggagac | aaaggaactc | agctctctgg | 94200 |
| tggccagaaa | caacgcattg | ccatagctcg | tgcccttgtt | agacagcctc | atattttgct | 94260 |
| tttggatgaa | gccacgtcag | ctctggatac | agaaagtgaa | aaggtaagaa | tttaaattgg | 94320 |
| gttcatccat | acttgaatca | gtcatcctta | aaaatgctgt | attctccata | gtaaagaaca | 94380 |
| aaatgcagcc a | agattattga | cagttactta | catcaccttt | cagtaaaatg | taaacacttt | 94440 |
| ctgtagctgt ta | tattcaggag | agtaaagatg | acacaaggca | gattagtctt | agagcacata | 94500 |
| tgcttcacaa a | accatcacta | gatcagtttg | $\begin{array}{r} \text { cgttgaccta } \\ \text { Page } 47 \end{array}$ | agaagctgct | ttcaggaacc | 94560 |

atatttccta cttctactgg cagcatctct tcagcaacca taaggaaaag tggtagtgta ..... 94620
agacagtgtt aaaatctggc tgtacctttt aagataaata ttagcaatca tttctagtca ..... 94680
gagaaaacgt taaaaattca gattatctga aagacgaact cagtatgagg tattctctac ..... 94740
aataaaaaaa ttcgggagtc tgtatttggc ctttgctcag gcttcttgaa tttactgtga ..... 94800
gctggatttg aggtctgttt tcaaggatgt tgtaacatcc ttgaaactaa ggtaaacaca ..... 94860
ttagtaaatg tttttaaatg agtcagtgaa ctgtatttac tgccttaact ttaagagtag ..... 94920
atgactgctg gccggcacct actgccgtgg ctcacacctg taaccccaac actttggaag ..... 94980
gctgaggtgg gaggatcact tgagcccagg atttttgaga ccagcctggg taacacagag ..... 95040
agacattgtc tctacaaaaa ataaaataat tagctaggca tggtggtggg aacttgtagt ..... 95100
cccagctact cagaaggctg aggtggaaga atcgcttgag cccaggaggt tgaggctgca ..... 95160
gtaagccatg attgtaccac tacactccac cctggtgaca gagtgaggcc ctgtttcaaa ..... 95220
taaaaaaaaa aagttaggga tagataactg ctgcctcgga actaacccaa attagggttc ..... 95280
agtcatctga tatgacaaat cctttgtttt caaaaacctg ttctatttca ctgacacctt ..... 95340
gttaagaagg atctattgtt gtattgttgt ggagcttttt tttttttttt tctttgagat ..... 95400
ggagtctcgc tctgtcgccc aggctggagt gcagtggcgc aatcttggct ccctgcaagc ..... 95460
tctgcctccc gggttcacac cattctcctg cctcagcctc ccgactagct gggactacag ..... 95520
gcacccacca ccaagcccag ctaatttttt gtatttttga tagagatggg gtttcaccgt ..... 95580
gttagccagg atggtctcaa tctgaccttg tgatctgccc gccttggcct cccaaagtgc ..... 95640
tgggattaca ggcgtgagcc accatgcccg tcctactgtg gagcttttta tggaagagga ..... 95700
attagggaaa agaactatta tgagaattaa tctatgtgat tatggaatag gttgtccaag ..... 95760
aagccctgga caaagccaga gaaggccgca cctgcattgt gattgctcac cgcctgtcca ..... 95820
ccatccagaa tgcagactta atagtggtgt ttcagaatgg cagagtcaag gagcatggcacgcatcagca gctgctggca cagaaaggca tctatttttc aatggtcagt gtccaggctg 95940
gaacaaagcg ccagtga ..... 95957

```
<210> }
<211> 3843
<212> DNA
<213> Homo sapiens
<220>
<221>
<222> (1)..(3843)
<220>
<221> misc_feature
<222> (1236)..(1236)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (2677)..(2677)
<223> n may be any nucleotide
```

```
<220>
<221> misc_feature
<222> (3435)..(3435)
<223> n may be any nucleotide
<400> 2
atg gat ctt gaa ggg gac cgc aat gga gga gca aag aag aag aac ttt
Met Asp Leu Glu Gly Asp Arg Asn Gly Gly Ala Lys Lys Lys Asn Phe
1 5 10
ttt aaa ctg aac aat aaa agt gaa aaa gat aag aag gaa aag aaa cca
phe Lys Leu Asn Asn Lys ser Glu Lys Asp Lys Lys Glu Lys Lys pro
act gtc agt gta ttt tca atg ttt cgc tat tca aat tgg ctt gac aag
Thr val. Ser Val phe Ser Met phe Arg Tyr Ser Asn Trp Leu Asp Lys
    35 40 40
ttg tat atg gtg gtg gga act ttg gct gcc atc atc cat ggg gct gga
Leu Tyr met val val G7y Thr Leu Ala Ala Ile Ile His Gly Ala Gly
    50 55 60
ctt cct ctc atg atg ctg gtg ttt gga gaa atg aca gat atc ttt gca
Leu Pro Leu Met Met Leu val Phe Gly G7u Met Thr Asp Ile Phe Ala
65 70 70 70
aat gca gga aat tta gaa gat ctg atg tca aac atc act aat aga agt
Asn Ala gly Asn Leu Glu Asp Leu Met Ser Asn Ile Thr Asn Arg Ser
    85 90 95
gat atc aat gat aca ggg ttc ttc atg aat ctg gag gaa gac atg acc
Asp Ile Asn Asp Thr G7y Phe Phe Met Asn Leu Glu Glu Asp Met Thr
    100 105 110
agg tat gcc tat tat tac agt gga att ggt gct ggg gtg ctg gtt gct
Arg Tyr Ala Tyr Tyr Tyr Ser Gly Ile Gly Ala Gly Val Leu val Ala
    115 120 125
gct tac att cag gtt tca ttt tgg tgc ctg gca gct gga aga caa ata
Ala Tyr Ile Gln val ser Phe Trp Cys Leu Ala Ala G7y Arg Gln Ile
        130 135 140
cac aaa att aga aaa cag ttt ttt cat gct ata atg cga cag gag ata
His Lys Ile Arg Lys G7n Phe Phe His Ala Ile Met Arg Gln Glu Ile
145 150 155 160
ggc tgg ttt gat gtg cac gat gtt ggg gag ctt aac acc cga ctt aca
Gly Trp phe Asp val His Asp val Gly Glu Leu Asn Thr Arg Leu Thr
gat gat gtc tcc aag att aat gaa gga att ggt gac aaa att gga atg
Asp Asp val Ser Lys Ile Asn Glu Gly Ile G7y Asp Lys Ile Gly Met
Asp Asp val Ser Lys Ile Asn Glu Gly Ile Gly Asp Lys Ile Gly Met
ttc ttt cag tca atg gca aca ttt ttc act ggg ttt ata gta gga ttt
Phe Phe Gln Ser Met Ala Thr phe Phe Thr Gly Phe Ile val Gly Phe
        195 200 205
aca cgt ggt tgg aag cta acc ctt gtg att ttg gcc atc agt cct gtt
aca cgt ggt tgg aag cta acc ctt gtg att ttg gcc atc agt cct gtt
        210 215 220
ctt gga ctg tca gct gct gtc tgg gca aag ata cta tct tca ttt act
Leu G7y Leu Ser Ala Ala val Trp Ala Lys Ile Leu Ser Ser phe Thr
Leu G7y Leu Ser Ala Ala val Trp Ala Lys Ile Leu Ser Ser Phe Thr
gat aaa gaa ctc tta gcg tat gca aaa gct gga gca gta gct gaa gag130135140
                                    Page 49
```



| ITe |  | $\begin{aligned} & \text { Lys } \\ & 515 \end{aligned}$ |  | Pro | His |  | $\begin{aligned} & \text { Phe } \\ & 520 \end{aligned}$ | Asp | $\mathrm{hr}$ | eu | Va 1 | $\begin{aligned} & \text { G7y } \\ & 525 \end{aligned}$ | G7u | $\mathrm{Arg}$ | G7y |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gcc | cag | ttg | agt | ggt | ggg | cag | aag | cag |  | atc | gcc | att | gca |  | gcc | 1632 |
| Ala | $\begin{aligned} & \text { G7n } \\ & 530 \end{aligned}$ | Leu | Ser | G7y | Gly | $\begin{aligned} & \text { G7n } \\ & 535 \end{aligned}$ | Lys | G7n | Arg | ITe | A7a $540$ | ITe | Ala |  | A7a |  |
| ctg | gtt | cgC | aac | CCC | aag | atc | ctc | ctg | ctg | gat | gag | gcc | acg |  | gcc | 1680 |
| Leu | Val | Arg | Asn | Pro | Lys | Ile | Leu | Leu | Leu | Asp | G7u | Ala | Thr | Ser | Ala |  |
| 545 |  |  |  |  | 550 |  |  |  |  | 555 |  |  |  |  | 560 |  |
| ttg | gac | aca | gaa | agc | gaa | gca | gtg | gtt | cag | gtg | gct | ctg | gat | aag | gcc | 1728 |
| Leu | Asp | Thr | Glu | $\begin{aligned} & \text { Ser } \\ & 565 \end{aligned}$ | G7u | A7a | Val | Val | $\begin{aligned} & \text { 67n } \\ & 570 \end{aligned}$ | vai | A7a | Leu | Asp | $\begin{aligned} & \text { Lys } \\ & 575 \end{aligned}$ | Ala |  |
| aga | aaa | ggt | cgg | acc | acc | att | gtg | ata | gct | cat | cgt | ttg | tct | aca | $g t t$ | 1776 |
| Arg | Lys | G7y | $\begin{aligned} & \text { Arg } \\ & 580 \end{aligned}$ | Thr | Thr | Ile | Val | $\begin{aligned} & \text { Ile } \\ & 585 \end{aligned}$ | Ala | His | Arg | Leu | $\begin{aligned} & \text { ser } \\ & 590 \end{aligned}$ | Thr | Va1 |  |
| cgt | aat | gct | gac | gtc | atc | gct | ggt | ttc | gat | gat | gga | gtc | att | gtg | gag | 1824 |
| Arg | Asn | $\begin{aligned} & \text { A1a } \\ & 595 \end{aligned}$ | Asp | Val | ITe | Ala | $\begin{aligned} & \text { G7y } \\ & 600 \end{aligned}$ | Phe | Asp | Asp | Gly | $\begin{aligned} & \text { Va7 } \\ & 605 \end{aligned}$ | ITe | val | G7u |  |
| aaa | gga | aat | cat | gat | gaa | ctc | atg | aaa | gag | aaa | ggc | att | tac | ttc | aaa | 1872 |
| Lys | $\begin{aligned} & \text { G1y } \\ & 610 \end{aligned}$ | ASn | His | Asp | G7u | $\begin{aligned} & \text { Leu } \\ & 615 \end{aligned}$ | Met | Lys | Glu | Lys | $\begin{aligned} & \text { G7y } \\ & 620 \end{aligned}$ | I7e | Tyr | Phe | Lys |  |
| ctt | gtc | aca | atg | cag | aca | gca | gga | aat | gaa | gtt | gaa | tta | gaa | aat | gca | 1920 |
| $\begin{aligned} & \text { Leu } \\ & 625 \end{aligned}$ | Val | Thr | Met | G7n | $\begin{aligned} & \text { Thr } \\ & 630 \end{aligned}$ | Ala | GTy | Asn | Glu | $\begin{aligned} & \text { val } \\ & 635 \end{aligned}$ | G7u | Leu | G7u | Asn | $\begin{aligned} & \text { A7a } \\ & 640 \end{aligned}$ |  |
| gct | gat | gaa | tcc | aaa | agt | gaa | att | gat |  | ttg | gaa | atg | tct | tca | aat | 1968 |
| Ala | Asp | G7u | Ser | $\begin{aligned} & \text { Lys } \\ & 645 \end{aligned}$ | Ser | G7u | I7e | Asp | $\begin{aligned} & \text { A7a } \\ & 650 \end{aligned}$ | Leu | Glu | Met | ser | $\begin{aligned} & \text { Ser } \\ & 655 \end{aligned}$ | Asn |  |
| gat | tca | aga | tcc | agt | cta | ata | aga | aaa | aga | tca | act | cgt | agg | agt | gtc | 2016 |
| Asp | Ser | Arg | $\begin{aligned} & \text { Ser } \\ & 660 \end{aligned}$ | Ser | Leu | Ile | Arg | $\begin{aligned} & \text { Lys } \\ & 665 \end{aligned}$ | Arg | Ser | Thr | Arg | $\begin{aligned} & \text { Arg S } \\ & 670 \end{aligned}$ | ser | Val |  |
| cgt | gga | tca | caa | gcc c | caa | gac | aga | aag | ctt |  | acc | aaa | gag | gct | ctg | 2064 |
| Arg | G7y | $\begin{aligned} & \text { Ser } \\ & 675 \end{aligned}$ | G7n | Ala | G7n | Asp | Arg 680 | Lys | Leu | Ser | Thr | $\begin{aligned} & \text { Lys } \\ & 685 \end{aligned}$ | G7u A | A7a | Leu |  |
| gat | gaa | agt | ata | CCt | cca | gtt | tcc | ttt | tgg | agg | att | atg | aag | cta | aat | 2112 |
| Asp | $\begin{aligned} & \text { G7u } \\ & 690 \end{aligned}$ | Ser | I7e | Pro Pr | Pro | $\begin{aligned} & \text { Va1 } \\ & 695 \end{aligned}$ | Ser | Phe | Trp | Arg | $\begin{aligned} & \text { Ile } \\ & 700 \end{aligned}$ | Met | Lys | Leu | Asn |  |
| tta | act | gaa | tgg | cct ta | tat | ttt | gtt | gtt | ggt |  | ttt | tgt | gcc |  |  | 2160 |
| Leu 705 | Thr | G7u | Trp | Pro 7 | Tyr 710 | Phe | Val | Val | Gly | Val 715 | Phe | cys | Ala I | Ile | $\begin{aligned} & \text { Ile } \\ & 720 \end{aligned}$ |  |
| aat |  |  | ctg |  | cca | gca | ttt | gca | ata | ata | ttt | tca | aag | att | ata | 2208 |
| Asn | Gly | G7y | Leu | $\begin{aligned} & \mathrm{G} 7 \mathrm{n} \\ & 725 \end{aligned}$ | Pro | Ala | Phe | Ala | $\begin{aligned} & \text { Ile } \\ & 730 \end{aligned}$ | Ile | Phe | Ser L | Lys 7 | $\begin{aligned} & \text { I1e } \\ & 735 \end{aligned}$ | ITe |  |
| ggg | $g t t$ | ttt | aca | aga a | att | gat | gat | CCt | gaa | aca | aaa | cga | cag a | aat | agt | 2256 |
| Gly | val | Phe | $\begin{aligned} & \text { Thr } \\ & 740 \end{aligned}$ | Arg I | Ile | Asp | Asp | $\begin{aligned} & \text { Pro } \\ & 745 \end{aligned}$ | G1u | Thr | Lys A | Arg | $\begin{aligned} & \text { G7n } \\ & 750 \end{aligned}$ | Asn | ser |  |
| aac | ttg | ttt | tca | cta t | ttg | ttt | cta | gcc | ctt | gga | att | att t | tct t | ttt | att | 2304 |
| Asn | Leu | $\begin{aligned} & \text { Phe } \\ & 755 \end{aligned}$ | Ser | Leu L | Leu | Phe | $\begin{aligned} & \text { Leu } \\ & 760 \end{aligned}$ | Ala | Leu | GTy | ITe | $\begin{aligned} & \text { ITe } \\ & 765 \end{aligned}$ | $\text { Ser } P$ | Phe | ITe |  |
|  |  | ttc | ctt | cag g | ggt | ttc | aca | ttt | ggc | aaa |  | gga g | gag | atc | ctc | 2352 |
| Thr | $\begin{aligned} & \text { Phe } \\ & 770 \end{aligned}$ | Phe | Leu | Gln G | Gly | $\begin{aligned} & \text { Phe } \\ & 775 \end{aligned}$ | Thr | Phe | GTy | Lys 7 | $\begin{aligned} & \text { A1a } \\ & 780 \end{aligned}$ | Gly | G7u I | ITe | Leu |  |
| acc | aag | cgg | ctc | cga t | tac | atg |  | ttc | cga Pag | $\begin{gathered} \operatorname{tcc} \\ \text { ge } 51 \end{gathered}$ | $\operatorname{atg}$ | ctc | aga c | cag | gat | 2400 |



$<220>$
<221> misc feature
<223> The 'Xaa' at location 893 stands for Thr, Ala, Pro, or ser. $<220>$
<221> misc_feature
<222> (1145)..(1145)
<223> The 'Xaa' at location 11.45 stands for ITe, or Met.
<400> 3
Met Asp Leu Glu gly Asp Arg Asn Gly gly Ala Lys Lys Lys $\underset{10}{ } \underset{15}{\text { Asn }}$ Phe 1

Phe Lys Leu Asn Asn Lys Ser G7u $\underset{20}{\text { Lys Asp Lys Lys Glu Lys Lys Pro }}$

 Leu Pro Leu Met Met Leu Val Phe Gly G7u Met Thr Asp ITe Phe Ala $65 \quad 70 \quad 75 \quad 80$

Asn Ala Gly Asn $\underset{85}{\text { Leu G7u Asp Leu Met }} \underset{90}{\operatorname{Ser}}$ Asn Ile Thr Asn $\underset{95}{\text { Arg }}$ Ser Asp Ile Asn Asp Thr Gly Phe Phe Met Asn Leu Glu Glu Asp Met Thr Arg Tyr Ala Tyr Tyr Tyr Ser Gly Ile Gly Ala Gly Val $\begin{array}{r}120 \\ 125\end{array}$ Ala Tyr Ile Gln val ser Phe Trp Cys Leu Ala Ala Gly Arg Gin Ile 130135140 His Lys Ile Arg Lys G7n
145
150 Gly Trp Phe Asp val His Asp val G7y G7u Leu Asn Thr arg Leu Thr $\begin{aligned} & 170 \\ & 175\end{aligned}$ Asp Asp Val Ser Lys Ile Asn Glu Gly Ile Gly Asp Lys Ile Gly Met Phe Phe Gln Ser Met Ala Thr Phe Phe Thr Gly Phe Ile val Gly Phe 195200205

Thr $\underset{210}{\text { Arg Gly }}$ Trp Lys Leu Thr Leu Val Ile Leu Ala $\begin{gathered}220 \\ 215\end{gathered}$ Ile Ser Pro Val 210215

Leu Gly Leu Ser Ala Ala Val Trp Ala Lys ITe Leu Ser Ser Phe Thr
225
230

Asp Lys G7u Leu Leu Ala Tyr Ala Lys Ala gly Ala Val Ala glu glu


 Leu Ile Tyr Ala Ser Tyr Ala Leu Ala Phe
305
310 $\underset{315}{\operatorname{Trp}} \begin{aligned} & \text { Tyr G7y }\end{aligned}$
 Ser Val Leu ITe G7y Ala Phe Ser Val G7y G7n Ala Ser $\underset{3}{340}$ pro Ser Tle Glu Ala Phe Ala Asn Ala Arg G7y Ala Ala Tyr Glu
365
360 Ile Asp Asn Lys Pro Ser Ile Asp Ser Tyr Ser Lys ser gly His Lys Pro Asp Asn Ile Lys G7y Asn Leu G7u Phe $\begin{array}{r}\text { Arg } \\ 385 \\ 395\end{array} \quad$ Asn Val His Phe Ser
390 Tyr Pro Ser Arg lys G7u Val Lys Ile Leu Lys G7y Leu Asn Leu Lys Val Gln Ser G7y G7n Thr Val Ala $\underset{420}{ }{ }_{425}$ Val Gly Asn Ser G7y cys Gly
 G7y Met Val $\operatorname{ser}$ Val Asp G7y G7n Asp Ile Arg Thr
450
450 Phe Leu Arg Glu ITe ITe G7y val val Ser G7n G7u pro val Leu Phe
4705
470

ATa Thr Thr Ile Ala G7u Asn Ile Arg Tyr Gly Arg glu Asn Val $\begin{gathered}495 \\ 495\end{gathered}$ Thr

 Ala Gln Leu Ser G7y Gly $\underset{530}{\operatorname{G7n}}$ Lys G7n Arg I7e Ala $\underset{540}{ }$ Ile Ala Arg Ala Leu Val Arg Asn Pro Lys Ile Leu Leu Leu Asp G7u Ala Thr Ser Ala
545
550
 Arg Lys Gly $\underset{580}{\text { Arg }} \begin{array}{r}\text { Thr }\end{array}$ Thr Ile Val Ile Ala His Arg Leu Ser
585
590 Arg Asn Ala Asp Val Ile Ala G7y Phe Asp Asp gly Val Ile Val G7u 595 600 605

Lys G7y Asn His Asp G7u Leu Met Lys Glu Lys Gly Ile Tyr Phe Lys
610
620 Leu Val Thr Met G7n Thr Ala G7y Asn G7u val G7u Leu G7u Asn Ala
625
630 Ala Asp glu Ser $\underset{645}{\operatorname{Lys}}$ Ser glu Ile Asp Ala Leu glu Met Ser $\underset{650}{ } \underset{655}{ }$ Asn Asp Ser Arg $\underset{660}{\operatorname{Ser}}$ Ser Leu Ile Arg $\underset{665}{\operatorname{Ly}}$ Arg Ser Thr Arg $\underset{670}{\operatorname{Arg}}$ Ser Va7 Arg Gly $\underset{675}{\operatorname{Ser}}$ G7n Ala G7n Asp $\underset{680}{\text { Arg }}$ Lys Leu Ser Thr Lys G7u Ala Leu Asp G7u Ser Ile pro Pro Val Ser Phe Trp Arg $\begin{aligned} & \text { I7e Met Lys Leu Asn } \\ & 690\end{aligned}$ Leu Thr G7u Trp Pro Tyr Phe Val Val G1y Val
705 Asn Gly Gly Leu G7n Pro Ala Phe Ala $\underset{725}{ }{ }_{730}$ Ile Phe ser Lys $\underset{735}{ }$ Ile G7y Val Phe Thr Arg Ile Asp Asp Pro G7u Thr Lys Arg G7n Asn Ser Asn Leu Phe Ser Leu Leu Phe Leu Ala Leu Gly Ile $\underset{765}{760} \begin{aligned} & 765\end{aligned}$
 Thr Lys Arg Leu Arg
785 $\underset{790}{\text { Tyr Met val phe Arg }} \underset{795}{\operatorname{Ser}} \begin{aligned} & \text { Met Leu Arg Gin Asp } \\ & 800\end{aligned}$ Val Ser Trp Phe Asp Asp Pro Lys Asn Thr Thr G7y Ala Leu Thr Thr $\begin{gathered}815 \\ 810\end{gathered}$ Arg Leu Ala Asn Asp Ala Ala Gln Val
820 Lys Gly Ala Ile $\begin{gathered}\text { G7y } \\ 830\end{gathered}$
 Ile Ser Phe Sle Tyr Gly $\underset{850}{\operatorname{Trp}} \underset{855}{ }$ G7n Leu Thr Leu Leu Leu Leu Ala Ile Val
865

G7y G7n Ala Leu Lys $\underset{885}{ }$ Asp Lys Lys G7u Leu G7u G7y Xaa G7y $\underset{890}{\operatorname{Lys}}$ ITe



Arg Asn Ser Leu Arg Lys Ala His Ile Phe Gly Ile Thr Phe Ser Phe
930
935

Thr G7n Ala Met Met Tyr Phe Ser Tyr Ala G7y Cys Phe Arg Phe Gly
945
950




I7e Met ITe I7e G7u Lys Thr
1010
1015 Pro Leu I7e Asp Ser $\begin{aligned} & \text { Tyr Ser Thr } \\ & 1020\end{aligned}$
'G7u G7y
1025 Leu Met Pro Asn Thr Leu G7u Gly Asn Val $\begin{aligned} & 1030 \\ & 1035\end{aligned}$ Thr Phe Gly
Glu val
1040 G7n Gly $\underset{1055}{ }$ Leu Ser Leu G7u val $\begin{aligned} & 1060\end{aligned}$ Lys Lys G7y G7n $\underset{1065}{ }$ Lhr Leu A7a Leu

 Lys $\underset{1100}{\text { Glu }}$ Ile Lys Arg Leu Asn $\underset{1105}{ }$ Val Gln Trp Leu $\underset{1110}{\text { Arg Ala His Leu }}$
 Glu Asn $\underset{1130}{\text { I7e ATa Tyr Gly Asp }} \underset{1135}{ }$ Asn Ser Arg val $\underset{1140}{\text { Val }}$ Ser G7n Glu


 Ala Leu val Arg Gin Pro $\underset{1190}{\text { His }} \underset{1195}{ }$ Ile Leu Leu Leu Asp $\underset{1200}{ }$ G7u Ala Thr Ser Ala Leu Asp Thr G7u $\underset{1205}{\operatorname{ser}}$ G7u Lys Val val ${ }_{1210}^{\text {G7n }}$ GTu Ala Leu Asp Lys Ala Arg Glu Gly $\underset{12220}{\text { Arg }} \underset{1220}{ }$ Thr cys Ile val $\underset{1230}{\text { Ile }}$ Ala His Arg

Gly $\underset{1250}{\text { Arg }}$ Val Lys Glu His G7y $\underset{1255}{ }$ Thr His Gin Gln $\underset{1260}{\text { Leu }}$ Leu Ala Gln

Arg G7n 1280

```
<210> 4
<211> 20
<212> DNA
<213> Homo sapiens
<220>
<22I> misc_feature
<222> (7)..(7)
<223> n may be any nucleotide
<400> 4
gaagggnctg aacctgaagg 20
<210> 5
<211> 20
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (7)..(7)
<223> n may be any nucleotide
<400> 5
gaaggtnctg ggaaggtgag 20
<210> 6
<211> 57
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucTeotide
<400> 6
ngtgagggca gcaaaggagg ccaacataca tgccttcatc gagtcactgc ctaatgt . 57
<210> 7
<211> 18985
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (18985)..(18985)
<223> n may be any nucleotide
<400> 7
nctgaacctg aaggtgcaga gtgggcagac ggtggccctg gttggaaaca gtggctgtgg 60
gaagagcaca acagtccagc tgatgcagag gctctatgac cccacagagg ggatggtgag 120
atgacccatg cgagctagac cctgcggtga tcagcagtca cattgcacat ctttctgatg 180
                                    Page 59
```

| ttgccetttc | aattacaaat | gtatgaaagt | cacacttact | tttattcca | ggtcagtgtt | 240 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gatggacagg | atattaggac | cataaatgta | aggtttctac | gggaaatcat | tggtgtggtg | 300 |
| agtcaggaac | ctgtattgtt | tgccaccacg | atagctgaaa | acattcgcta | tggccgtgaa | 360 |
| aatgtcacca | tggatgagat | tgagaaagct | gtcaaggaag | ccaatgccta | tgactttatc | 420 |
| atgaaactgc | ctcatgtaag | ttgtccttgc | cctttgcctt | tctagaggtg | caaaaaataa | 480 |
| aatgcaggcc | tactatgcag | gaagttagga | aactactata | aatcggaaga | agggaaatcc | 540 |
| taagaaggga | aagtaagatt | acttcagatt | tgaaagctct | agcagtatca | actggtcgta | 600 |
| gatacatttt | taaaaactga | ggttggttat | tgtg | aagatttaaa | gaactggacc | 660 |
| tgtattactt | gtgagacttg | ggctgtgtat | agga | accaatttaa | aatatgagct | 720 |
| gagatagctt | gtccttatgc | taaatcattc | tgggttttct | gtggtagaaa | tttgacaccc | 780 |
| tggttggaga | gagaggggcc | cagttgagtg | gtgggcagaa | gcagaggatc | gccattgcac | 840 |
| gtgccetggt | tcgcaacccc | aagatcctcc | tgctggatga | ggccacgtca | gccttggaca | 900 |
| cagaaagcga | agcagtggtt | caggtggctc | ggataaggt | cagtgaggct | tagttcaaac | 960 |
| caacctgatt | tataagcata | agaacattct | actactaatt | cttgttaata | ttggtcttag | 1020 |
| aaaaggaaa | ttctga | ttctaggtga | cttcagc | tattaaaata | aaagcattgg | 1080 |
| gcctctttga | aatctttttc | tatttgtttg | ttttattgtt | caatttctat | ttatttctct | 1140 |
| gatcttattt | taaatgttga | tgaatacatt | ttcatttgaa | gacacttgct | aatcttttaa | 1200 |
| attaaaaaat | agaaatatag | acacatgtga | aagttcatct | tcattgtgat | cttcaaaact | 1260 |
| tgactatgtg | gataaccctg | ttatttaggt | tttgagagtt | tgtaatattg | ccaagaagag | 1320 |
| aaaaatacaa | cctgaaggtc | catatataat | tttccaggtg | ttgaatgcca | cttgaagact | 1380 |
| ctatgcgaaa | taagaaac | cttatttcca | ggaaaggggc | agatagcctg | tgatactgaa | 1440 |
| aacctaccta | agccatgaca | ggttattgac | tatcaacaga | gtttgactgt | cctgcaattc | 1500 |
| tggagtccat | atgactcatt | caccaaatag | catttgagtg | tttgccgtgt | gccgggcact | 1560 |
| gtgcctttga | tctccagcac | gtgatagtaa | cggggataat | tctgtgagga | ccgagaatgt | 1620 |
| ggagatggag | acattataac | caaaggtgtt | ccaagttgag | atgtcacagt | agaattcaaa | 1680 |
| gatgaactca | tatttgtttc | actctccct | ctctaata | aaacaacact | tgaatgttcc | 1740 |
| ttaacatcct | gtcaatgtgc | taataaatt | tttgagagag | aaaaaaagca | gcttactaaa | 1800 |
| cattctgtga | accaaaatag | aggccgatgg | gattctggtt | actatttttc | ccctcatttt | 1860 |
| gcttaatctg | tgatttcatc | tctgtgtttt | tctttttctt | tctttatttc | cttccttcct | 1920 |
| tccttccctc | cctccctcaa | tccctccctc | tcttgctctt | cctcttcctt | tcctttcttt | 1980 |
| cctttccttt | cctgaccttc | ccttcctttc | atttcctttc | ccttcccttc | cctttctttc | 2040 |
| ccttcccttc | ccttcctttc | ccttcccttc | ccttcctttc | ccttcccttc | ccttcctttc | 2100 |
| ccttcccttc | ccttcctttc | cctccecttc | ccttccctcc | ccttcccttc | cctccccttc | 2160 |
| cctcccctcc | catcecctcc | cctccctttt | cttttctttt Page 60 | ttctcttctc | ttctcttcct | 2220 |

ctcctctcct gtcttttctt ttcttatctt atcttatctt ttcttttctt gtttcttttc ..... 2280
tcactctgtc accaggctga agtgcagttg tgccatcatg gccactacaa cctctgctgc ..... 2340
ccagtctcaa gtgatcctcc cacctcagcc tcacaagtag ctgggaccac aggtgtgtgc ..... 2400
caccatgcca aggttttttt tttttttttt tttttgagat ggagtctcgc tctgtcgccc ..... 2460
aggctggagt gcagtgggac aatcttggct cactgcaacc tctgcctcct gcctcagcct ..... 2520
cctgagtagc taggattaca ggcatgcacc gccacacctg gctaattttt gtatttttag ..... 2580
taaagacagg gtttcgtcat gtggcccacg ctggtcttga actcctgacc tcaggtgatc ..... 2640
cacctgcctc ggcctcccaa agtgctggga ttacacgcgt gagctaccgt gcccagcccc ..... 2700
cagctatttt ttgatatatt tgtagagatg aggtctcact atcttgcccc gactggtctc ..... 2760
agactcctgg gctcaagcaa tcctcccgcc tcagcctccc aaagtgctgg aattacagga ..... 2820
gtgagccact gcttactggt ttgcttatct gtgtttcctt attaatctat agtgaaacta ..... 2880
tgtattaaat tataaataaa aacaatttta aaaggttata ttttaaaata ctttagggtg ..... 2940
taaattttga ggggaaattc cacatacccc ttttttctta aaagataca aaaattgatc ..... 3000
tattttcttc tgtattttct agtttctacc acctaatttt tccttgtgta ttttttcttt ..... 3060
ttgaagtttt ccacttctac ttatcctatg gatcctgaaa atgttgtgtg ttggttttga ..... 3120
gaattgtatt gctagttatt agagagacat atagagtaac aaaaattatg agcattggga ..... 3180
aagttacaaa ggttagagaa gtctcagaca aggcctggat atctggctct gttcctttat ..... 3240
3300
gaaaataaaa gagacttacg tgactcttca atttttccat aattcttcaa cctaggaata
3360
agtatcacta actatggata aggcacagtg ttgagtactt tatgtgcttt attttattta
gtcatcacaa ctactctagg aagtaaatac tattattatc cccattttac atatgaataa ..... 3420
attgagtctc acacagtttc cttggataaa atattttatt ggataaaata aattcataaa ..... 3480
tttatttcag gtcagtgtga cattgaggtc tggactttgc tgcctcacat ttattgtctg ..... 3540
tcttgttcat ccagggggtc atgcgggata ggatattata attcctaggg ctgattacta ..... 3600
gccggtgtgt atcagtacag cacaatggcc tgtgtttgtt ttgattggcc aacgcctggt ..... 3660
ctgtaggaat ttgttggttt gtacaagccc ctgattatta ttattttttt attttttatt ..... 3720
tttttttttt gagatggagt ctcactctgt cacccaggct ggagtgcagt ggtgcgatct ..... 3780
tggctcactg caagctccac ctcccaagta gcggggacta caggcacccg ccaccatgcc ..... 3840
cagctaattt tttgtatttt tagtagagat ggggtttcac tgtgttagcc aggatggtct ..... 3900
tgatctcctg accttgtgat ccacccacct tggcctccca aagtgctggg attacaggcg ..... 3960
tgagccacca cgcccggccc aagcccctga ttattactgc aaatttaggt taaataaaat ..... 4020
atttgggggc ttacataata ttaatatgtg actgttatat ttgtgtttgt atttattaca ..... 4080
aggaaacatc atttttaact attatcaatt gtctatacat ttattgaagt cagaggctat ..... 4140
cttatataga tttgatggtt ttacaatgcc cacagcattg gttcagtaaa tatatgttga ..... 4200
atggttaagt ttcttcaggt aattgttaat gtattcaaaa accaaatttc tctctcttta ..... 4260



gggctttgag ctggtccctt aaatgttgac gatttgaatt ggtgaagcaa aagaagttag 10440 aggactacct tcataacatg gtgcttgggg ggactggaac ctagcttgcc tagaaaacag 10500 gtgagaacaa agtctatgag ttataggacc ttagaagtct agagatagca aaatgcctaa 10560 agatgttaga ccacccaact cctcatctta gtaacatggg gagggagtcc agtgggtaat 10620 taaaagctca ggctctggat tcccattctg gttaggattc tcatgcatct ggattaagat 10680 tccaactctg tgaccttata actgtgggct atggatctta tcaggctctc taagcctcag 10740 tttcttctgt aaagtgggct tttctgtcta caccacccgt acagccttgt gccatggcac 10800 acagtcacag aaacatagca agcccttgaa atcaggcttt ctgactttgt ctaatctcct 10860 gctttagcaa agacatcaat tctccctcct tttatttaaa tggtggctgg gtccctgaca 10920 aggtatgttc ctgcccacag ggtttcacat ttggcaaagc tggagagatc ctcaccaagc 10980 ggctccgata catggttttc cgatccatgc tcagacaggt atgtctatcg agggctgtgc 11040 cctgggatgt gtagaactcc ccatgtgtgc cccttggact cagacagtgg gagctctgtc 11100 atgtttccca ggcctagctc ctatattggt tgtccctcag tctcacgtca gagatgcggt 11160 tgaaccgccc actaggcaca gatgaggctc tactgtctgt caggtctggg tgggcacaag 11220 gaactggacc agtttgagac agagcctcca gcgtggattc acctccagcc tgtgtctcag 11280 cagtggtggg cagtgcaggc agaaggcatc tttgataccc ctcatccctt tcctctcttt 11340 ccttctttct cctggcaaac ccagggcacg agtgtttctt cccaccaaga gatgccctct 11400 gtactctttc tttccgcaaa tcaaaactca tccctgtgtc ctgttcttca ctgcccattt 11460 tcttttctga ggctagtctg aaatactagt tcaagctgca gtgttactgg ctgataatgg 11520 gtttaatgga atgcttctca cattttgcag cttcaaaacc ctaaccattg acacgtgtga 11580 atgttttcct ggggaaatgg ggaaggaaat tagaggaatg taactcagag cagcctggtt 11640 caaaggggaa agttccttta acaatcatga aaattttgta tgtgacctaa taactttccg 11700 ttttaaaaat cactaatcac ttgccattga gtaaaatgat gctttagaag tctgccccag 11760 atgtgccagg ggtactcgaa ccctggctaa gaggcatcag tttggtgtgt taggctttct 11820 agagggcatt cacattttac cagctgtctc tggttcctca gttcttcccc attcctccca 11880 tcaccattta gaaagaaaat gtatttatgg gaattgctag ctagtgactg acagagccag 11940 gactgaatct aagtgagaga gcacagtttg atgggaaacc ctgtccttgg actgtcaggt 12000 cgaactgtat ttataagtca gattccactt aggtctacac tgaccttgct ccagggccaa 12060 atttcccatt acccaaccag cctccaggcc aactgctgtg cccattatac tttggcagct 12120 gagctgatgg tttgtggaat gtctcctcca taaattgtta agtagggcaa gcatttatta 12180 agtgccttct gcttacaagg tctagtactt agtattgtaa ggcattcaaa cctgaatggt 12240 ccctgtgttc aaggaagtta cattcctgtg tggagacact tttaaccact taagaatatt 12300 gaaaagcaag ttgatacata ctataataaa ttatagcact attttttctt aaatattttt 12360 gggaaaaatg atagatactt ctttaatgta aataaccagt ttagataact tcttagggac 12420 Page 65
cacagttatt tgtaaaacat ataaattcat acttcaaatt cacaatcatg ttagtatctg ..... 12480
tgtatttaaa aatataaatg gttttacaaa gaaaatctta ttttatgttg aatgttatat ..... 12540
tttaatctgg attacttttg ctgatttgtt tttgactcaa atcacttata ctgccctgag ..... 12600
ctacatttat ctaactgctt attcaacctc tctattggat gtctaataag agtttcactg ..... 12660
tttaacattt ccaaaatgga gcttttgact cttcccctcc ccaccagcct tattccttac ..... 12720
ccactcttcc cctatatcag taagcgacag ctccgttcta ctagttgttt gggctaaaaa ..... 12780
ccttgaagcc atctttgact cttctctttc tgttaacatt gcaaaagctt cctaactgat ..... 12840
ctccctacct ccatttttat cccaatcctg tagattcacc taaaacagc cacagtcatt ..... 12900
tttctaaaat agaaatcaaa tcatatccca ctgcccaaac ctgctgaagg ctgcctgtag ..... 12960
ctcacagggt aaagtctgga atcttgacag tcgcctgtaa gaccatacac tatgtggccc ..... 13020
ccactctctc tcaaatctca tctcctataa ccgccttttt ccaatacacc gtctactcca ..... 13080
taaacagtgg ctcactgccc aagagtaggg gaggggagga tcagaggagc caaatcagaa ..... 13140
cacaaatctg tccagaggaa gggggtggct tttacaaatt catacaaagg ttctttttag ..... 13200
gctgctggca gcctttaagg ctttcttgct ggtgtatgaa cagattaggt atgcctttgc ..... 13260
ctcagggtct ttgcacatgc tgtttctgct ttccctgagg tattctcatc ctttccttca ..... 13320
ttcaggcctc tgttcagatg tccccttaga aaggccttct gtttcccctc cccagtctct ..... 13380
aaaatagcac ctcccctcac ttttctgttc ttcttaccct gctcaatttt tcttttattt ..... 13440
gtaccttcca tctctagaat ctaaacttca tgaaagcagc tactttgtct tttttgttcg ..... 13500
tacaggtcca acacttagaa atcatgcgta ggagaaagta ggcactcaga aatttttctg ..... 13560
acaaatgaaa tgatctattt atgtgttttt atattaagtt tctttcttgt gtattgaatg ..... 13620
tcacatcctg agtactaaat gcagggggta taagtataaa caaaactgac cccatcgctg ..... 13680
ccctcttgga gctgagagtc tcataaacag ctttaaggta ataaaatcat tttctgtgcc ..... 13740
acaggatgtg agttggtttg atgaccctaa aaacaccact ggagcattga ctaccaggct ..... 13800
cgccaatgat gctgctcaag ttaaaggggt acgtgcctcc tttctactgg tgtttgtctt ..... 13860
aattggccat tttggacccc agcatgaaac taattttctc cttacgggtg ttagttatca ..... 13920
tcattaagaa aatgttgaat aaatatctaa cctacgaata tatcacatgc tttttgtagc ..... 13980
aacatgttaa ctatttaaac attatatact gtagagcata tagataactt ataaaccatt ..... 14040
tgctattgct gttattcatg ctattaacaa gatgcatgta gaatagttat ttagaaaaga ..... 14100
gagtataaag tgctcaatca acataaaaca gtaattgcta ctgaagaaag gatgtattta ..... 14160
attgctgtaa gaaagtttag agtcactatg gttacagagg ggagggaaga caatcctcta ..... 14220
aaatataggt tgaaggaaat gaaaagcaca ttaaaaaatt aaggcaagaa tagaataact ..... 14280
tcagtcttta tctttaataa ctttaaactt taataatttt aataacttaa attttgctac ..... 14340
tgtatgaatc tcttgatata actagatact attgaaccag caggttttga tttttggctg ..... 14400
aagtgacaat ttcttctaca actgtttatg gcaaaagtcc acaaaatgat gtagaatttg ..... 14460
aaaaaattca tgtaatctct ggtgtgtctt ttcccctctt gaaccttatc catctttatc 14520 tttaaatctt ttctgtaagt tagtactata ctaacatttc ttctatctaa tatatggggc 14580 ttctttaaga ataaattaag ctataaatga ggaaatacat agagttataa cgttgaaata 14640 taaaccttag gagtccctct ttttctattg tttggaatag tttcagaagg catagtacca 14700 gctcctcttt gtaattgttg ttttaataca aacttctttg cctaaagcaa acaaaacaat 14760 aaaatcaag gtttagatca agttgtatag aatgtaatta caggtgcacg cctgtaatcc 14820 cagctactcg ggaggctgag gtacgagaat tacttgaact caggaggtgg aggttgcaga 14880 gctgagattg caccactgca ctccagcctg ggtgacaaag caagactatg tctcaaaaaa 14940 aaaaaaaatg caaagaagac agagtggctg gaataaagtg agtgaaaaga agagtcataa 15000 gtgtgttaag gtcagcatta tatccagaag tagatggaaa accactgtag ggttttgaac 15060 acagaagtga catgatctga aattttgaaa ggatcactat agaaactgtg tgaataggcc 15120 gaagggggca agcatagaag gagtctgttg cagtaatcca ggaggagatg atagtgtttt 15180 agactaattt ggttatacaa aaggctataa gatataatta attctggata tattttagat 15240 gtacagccaa caatgtgttg gttggataag atgatggata tgagagaaag ggtatttaaa 15300 gatgactcca aattctttta cctgcacagt ggaaaaaaaa atggagtttt tttttttttt 15360 ttttttttta aagagacagg gtctctcttt gtagcccagg ccagactgca gtggcatgat 15420 tacagctcac tataacctag aactcatggg ctcaagcaat cctcccttct cagccttcct 15480 ggtagctggg accacaggtg tgcaccacca tgtcgagcta atttttaaat tttttgtaga 15540 gacagggtct ccctatgttg cccagtctgg tcttgaactc ctgggctcaa gtgatcctcc 15600 ctgcctcagc ctcccaaagt gttgagatta caggtgtgag ctgccatgcc cagctggagt 15660 tgatctttat gaaattcaga agcctgttgg agaagcaggc ttgggggaga atggagagtt 15720 ctgtatgaga catggtaagt ttgtgacgtc tgtattagac atccatatgt agatgtcaag 15780 aaggctgaaa tttgcactgc taactttagt aaaccttttt tataattggt ttactaaata 15840 agttttacta tgcttctcca ttcattttgg tcctcacaac tctatagact cctactctgt 15900 aggaggaata tacatgagac agtgagcatt agtctttgga ataaaggaac agtaaccagt 15960 gcaatgtgac attgcacaat atgacacaga ccctgtggta tgggctcatg tgctttgatg 16020 gacaaatatc ccgcatcaca tttgacgtgg aagacacaat cctggaggca actcagtctt 16080 ttgcagcaga acccagagca gatgtattca gggcattctg tattgcagtt tttcttcctg 16140 cctctcaaat tcctctggtg ttatgacctg ccttaggcca cagtgagttc ctatatttgt 16200 aaattggtg gtcacttttt ccctccatag tgctgtttgt gaggcccttt gccatattca 16260 ctgtctctaa ttgcctgctg gggtcaacct tctgcttttc acttgtttcc tcaaagacac 16320 ctcaaacttg gccctgccaa gatggctgca ctcaccttat atgagccact caagcaaaac 16380 tgttttccag aacttaaaac agtggctaaa aaggaaagac cagggaagag gaaatgagca 16440 ggttggcaaa cattgtcaac ctaggaaagc ccaatgtagc tgttatcaca gattgactga 16500 Page 67
gagaggactg ttggatctac tttacaccac tagctgacct gtttttggct gacaggtttt ..... 16560
agttcctccc ctcaacccta gtctttacaa tgaaaatttt ctggctacta atctgtgctc ..... 16620
ttccattctt cttagtcccc atatttctat agactcctac tctgtagaaa tagatgtagg ..... 16680
acagtgagtg ttaatcttca gaataaagga atagtaatca gtgtaatgtc tgaagtgcta ..... 16740
aatcaaatga agccgagggg ttaaatctcc ttgcagttta aggtcttaga cttcttagtt ..... 16800
actctttttg gtacattaaa gaatttgcca tctttgatcc atcttattta tttctgtggc ..... 16860
ccttcactac tatccaaagc tgcacattta tatgtctgaa aagttaaata gccaggtaat ..... 16920
tcctctgcca tgaactcaca cctagaagtg attctaggtg atattctaag agcttctctc ..... 16980
acaagcccat aagtcagtag ctgcagtggg aaatgagctc tgagtagaac ttgtagatac ..... 17040
tgagaacagg agcaacttat atcatccctg tgccacagcc gttctccagc ttgcatcctt ..... 17100
tgaccttcaa actaaagatg tgaaatgcag agagctggct ttgaaaacca ttccagtctg ..... 17160
atttatccca aatttggcaa atcatccctt taatataagc tcaatgattg catcaacaaa ..... 17220
atatacaggt gttctattat ttatgatacc tctttcagcc ctggtgcagc tccatggtat ..... 17280
ggaatttaat gtatttaaag gctactgata gttatagcac tcttttccta aatactagtc ..... 17340
tggtgtttat atcagacagg tgttaaatga ttttgtaggg attaatgtag ggaattataa ..... 17400
aggattctat tgttactaga aaggggtccc gatgcagacc ttgaaagagg gttcttggat ..... 17460
cttgtacaag aaagaattca aggcgaatcc acagagtaaa gtgaaagcaa gtttattaag ..... 17520
aagtgaaaa aattaagaat agctacttca taggcagaac agcagcatgg gatgctcagc ..... 17580
tgcttatact tattgttact tcttgattac atgctaaaca aggggtggat tattcatgagtttcctagga aaaggggtgg gctcttctca gaactgaggg ttcctcccat ttttagacca 17700tatagggtaa cgttgccatg gcatttgtaa attgtcatgg cgctggtgtg tctttgagca 17760tgctaataca ttataattag cgtataatga gcagtgagaa caaccagagg tcacttttgt 17820caccatcttg attttggtag gatttggctg gcttctttac cacagcttgt tttatcagca 17880aggtcttagc gacctgtatc ttgtgctgac ctcctgtcgc atcctgtgac ttagaatgcc 17940taacctcctg ggaatgtagc ccagtagatc tcagccttat tttacccagc cccttttcag 18000
aaaaggcagg tctggggcca gggcaggcat ttggaatggc ttgagggcac agaatatttc ..... 18060
cagggtagag gagatgtgct ggactaaatt atagagtgat ggactaaact gactttgtag ..... 18120
cttgactttt ggtttcagag ttggaaatct ggtttacagt tggacattat acagtggtgg ..... 18180
attaaattca ctttgcagtt tgacttctgg ctttagagct ggaactcagt gcacactagt ..... 18240
gaaagtcatg cttcagactc cctctggaat tcaaggggaa gataataatg cgcgcttact ..... 18300
atcttaaatt aaattccata attttcagct ctgtattttt tccaaattaa acattatatc ..... 18360
tcaaacagac ccagatatat ttgaatatta ttaatgacaa acgttaggct taaattacaa ..... 18420
ataataatat acctaacatt ggaaatttcc atcattccta gtttgtcaga ctcctttatc ..... 18480
ttgctaattt gcagatattg ctttagtaat gttgccgtga ttaatgaagg ttttcttggt ..... 18540
attaaaagat ccaaagagat aggaatatgt aattgaactc tagattgttg atattctact ..... 18600
ttcagcattc tgaagtcatg gaaattctta ctgtagaaac tcaataaact tacaagtaga ..... 18660
cctttacttt ttagttcatt actgataaaa taatgaatat agtctcatga aggtgagttt ..... 18720
tcagaaaata gaagcatgag ttgtgaagat aatattttta aaatttctct aatttgtttt ..... 18780
gttttgcagg ctataggttc caggcttgct gtaattaccc agaatatagc aaatcttggg ..... 18840
acaggaataa ttatatcctt catctatggt tggcaactaa cactgttact cttagcaatt ..... 18900
gtacccatca ttgcaatagc aggagttgtt gaaatgaaaa tgttgtctgg acaagcactg ..... 18960
aaagataaga aagaactaga aggtn ..... 18985
<210> 8

$<211>21978$

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)..(1)

$<223>\mathrm{n}$ may be any nucleotide

<220>

<221> misc_feature

<222> (21978)..(21978)

<223> n may be any nucleotide

<400> 8
nctgggaagg tgagtcaaac taaatatgat tgattaatta agtagagtaa agtattctaa 60
tcagtgttat tttgttactc cctactgctt actatgctct aagaatgtgt ttataaccat 120
tcctcaaagc aatctttttc atgcttattc agtaaattag aaacttacag aaagtagcaa 180
agccagttct tggactcaaa aactgataat taactttaac agactttttc agttttcagg 240
ccattgtctt cacactgttc ttccttcctc cccactttcc tccttccctt agttattttc 300
ttctttcttt tctctcactt tcactctctc tccactcctt ccttccttct ttccttcctt 360
ccttctttcc ttcttttctt tccttttttc cttccttccc ttcttccte ttttcttttt 420
ctttcttttc tttctttctt tctttcttcc tttctttctt tcttgctttc ttgctttctt 480
gctttcttgc tttcttgctt tcttgctttc tttcttttct ttcaagctta aatccattcc 540
tttattgagg aaagtaagcc cattttattt gtacatgtga ggggggagat taaatatgga 600
aaatgctag gggtatttat tatatctgtt ttaaattact accactcttt cttttttttt 660
atcatgctcc tcccttcatc ctatttctgt ttatctttac ccttttttta cttctttttt 720
ttcccctgca tacttgtctt ttttttccca ttatttaaca aatgcttatg tggcatttac 780
tgtgtttcca ggcaaatgtc ttattcctta tagcaaccat atgggggtct attatctcat 840
ttttctagtg gggctgaggc acaggcaggt catggctctg gactttagag ctgataggtc 900
ttggagctgg gattcaagct cagacagttg tgctccaaag ttgtttctct ttctgttata 960
aaacaaagag ttcctctgat ggcagaatgc agtctgatat cacatgatct gtatcatagt 1020
Page 69
ggaaatgaga ggtcagagca gggctgactg ccataactaa catttaggac agggatatat ..... 1080
gtgtgatgaa cattctgaga ttcccaggag ttagggcagg gactcacaca gatcagagtg ..... 1140
gctctggttg tcagtaggcc cagctacctc accaagtgaa tgatgaagaa ggccccagat ..... 1200
gttggtcatt gccactaatg ctttgtcctt tacctctctg ctatctcttc agactcttta ..... 1260
tcccattttt ggggggttcc ctctggaaaa tcttttgggc cagctgtagt acctccccag ..... 1320
ccagtgtctg taggggacaa acctcatgcg aggcatccta tcaggctccc ggaggcctat ..... 1380
ttctggaggt acatcaggat ggtgctgacc cacagggttc tattgtggtc tggctcaagt ..... 1440
aacgttggcc cctgccaggg aatacattct gtgccaatga cttcctttaa tattgtactt ..... 1500
tgggggtctc ccccatggtg ttgactcttc ccctgagaag atggaaaaat ctaggacaaa ..... 1560
atataaaaaa aggagaaagc attctagtta acagtgtcat tatttaatga ggtatttgaa ..... 1620
atgtcccaga agggagtctt agattgcttt taggaagaag atgatacaat ctgatcctaa ..... 1680
gctaattcat gcaaaccaca ggttagcacc tttgaagtga cacatgagat ttaagactct ..... 1740
tggagttctg tcagaggggc caaaggaagg ggagcagatg gaaatgaatt atgtgtggat ..... 1800
atggcagatt ttctgagcaa aggtcaggga tgtgatacat tttctaattc aggggtggct ..... 1860
catgagaggg acagaagtag gtagggaaag tagtggatct gcaccaacca gtccacataa ..... 1920
tattgaaaat caacttccac ctgaacacac ttctcaagct gctatttagt acttcttata ..... 1980
taacaagtat cattctaaca gctctaaata cttgatctaa tttaatgctg tcaataaccc ..... 2040
tgttgaatta tccctgttgg atagataatg aaacccaaga ttcagagaga tcaagtaact ..... 2100
tgctcacagt cacacagcaa cagcctccta accaatctcc ctgcttcaac ttcttccatt ..... 2160
cttatccaac tatagtctac attcttacag aagtcattct gacctctcat aaataaaaat ..... 2220
ctaataatgc caacccttac ttaaaactcc caaatggctt ctgatcatac ttggaacaaa ..... 2280
atctaatctt ttatcatgga ttaccaaact ctgtatgatc taactcctga tgacctttat ..... 2340
ggccttatct ggcatcacag tctcttatgc ctgtgtgccc tggcccagtg ggcattcttgcaattcctgg aatgttttaa gctgattgcc tctgatcttc tcttggctga ctttacgttattcatatccc aaaattatta cctcttcaga gaggccattc ctgacacctg aattaattcatgatacttcg cagcccttga tgtttctttc ctatttcttt gtttattggt taatggtgtaccatttgett attgattaat ggtgtatgtg acattagagc ccacacttat tctgactgcatgatgtcaaa acctgctcat aacctctacc ctgcctgtga tctgtgctct gcagtttttcagcacattat tgtggactag tgagcatttt tataaaatac aggtgaccct tgaacagcacgggtttgaac tgagtgggtc cacttatatg taggtttttt tcaataaatg tattggaaaattttttgaag acttgcaaca atttgaaaaa actcacaagc catgtagcct agaaatatcaaaaattaag aaaaagttaa gtatgttata aatgcataaa atatatgtag ttactcgtcttattatttac taccataaga tatacacaaa cctataatta aaaaagtta aaagtgatcaaattttatgc acacaaacac ttacagacca tataagtcac cattcacagt caagacacat


tacttctgat agattttgat attagtcttc aaaactgact taatgtctta tgaaatgctt ..... 7200
gctgttatgt ttgaagttag gtaatttatg taagattcag tgaagaataa gtggaattcc ..... 7260
atgtttatga ttttaagcta taaaacactc taaattaaat gtgtctttat tagaatctgt ..... 7320
tctgaccagt gcagaggcca aagagaggaa gaacatctta aacaataaag agtcatttct ..... 7380
cttggtgact tataattctg gaagttattt ctcttaaaat catagcatta aaagggactt ..... 7440
tagagaccet ctagtccatc gtcctcattt tgcaaatgag gaaaatgaga cagcatgttg ..... 7500
gttcaaggtg gtgcggctga tgtaggctga aatctcatct tgtacactgg tgttctttgc ..... 7560
tttttccata tccctttact cagactccag aggtgatgaa ggatgtatgt ttcctaatca ..... 7620
gattgccttg ttggaagtaa catttgatta caacataatt gaatgatgga aactttcttt ..... 7680
ttaagatgga gtctcactct gttgcccagg ctagagtgca gtgacgccat cttgactcac ..... 7740
tgcaatctct gtctcgccag ttcgagcgct tcctctgcct cccagtagca tgggattaag ..... 7800gcatgtgcca ccatgcccag ctaatttttg tatttttagt agagatgggg tttcaccatgttgatggtgg ctggtctcaa actccttact tcaaatgatc cacctgcctc catctcccaa78607920
actgctggga ttacaggcat gagccaccat acccagccca aaactttctg gaaaacagat ..... 7980
tgatagtatg tgccacattc cttaaaaaat taaaaaaatt aattcaagcc aggtgtagtg ..... 8040
ccatgtgctt gtagtcttag ctacttggca gactgaggca ggaggattgc ttacccaggt ..... 8100
gtgtgaggct gcagtgagct atgatgatca cacctgtgaa ttgccactgc actgcagcct ..... 8160
gggcaacaga gcaagacccc gtttctaaaa aaaaagtta gttttctttg acttattaat ..... 8220
ttcaccttaa gaaattttcc taataaacca attcaatatg gacaaatgtt taggtacaaa ..... 8280
gatgtttatc tcaccactat ttttaataaa aaaggaattg aaaacctggc tcaacaataa ..... 8340
aggaatactt aattggttat gatattaag gactcattac acatctcatt attaatgtgt ..... 8400
atttaatgaa cttggaaaat gcttttgata tgaaggtaaa aataatgata tagagctaaa ..... 8460
tatagagttt cattccaatc tttttaaata tatttatgca cttaggaaaa aaacaatatg ..... 8520
gaaatgtgta aaatatactt tttttttaaa aaaaggaca catttattca gcattatgat ..... 8580
cagactatta catttaacaa tcaacagtat gggtgccaaa aaaaatctac attaaaaccc ..... 8640
tttgttgtaa tgctttacac tttccacaga acagaaacta aagaatctg ttacacaatt ..... 8700
agtcacaaat atagtcctcg agttttttac ccatacacat gagtatttgt ctaaaacatg ..... 8760
tcttcttgta gcacttaggc cctgccacca ctgtgcttgt ctgagttcac aaatctgttg ..... 8820
taaactgtag cttccctgtc acttctctgg ctcttatctc ctgctaagat ttgtttcctg ..... 8880
gcagtaattt aaaatcttct gccactgctg tagctactgc tgctactgga actgccatag ..... 8940
ccaccttggt ttcatggttt ggcaaagtac tggcctgtac cagcataggg gccagagctt ..... 9000
ctgcctccaa agtttcctcc cttcatgggt ccaaaatgta aaactaattg ttgtaattgc ..... 9060
caaaatcatt acaccacctc caaaattgct tccatgatta ccaaatccat tatagccatc ..... 9120
cccactgcca ctatatccac caccaccaca gctgccacca aagccacaat gaccactgaaPage 73
gtttcctcca cgaccaaagt tgtcattccc accgaaacta cctccacgac caccaccaaa ..... 9240
gttcccagag ctgctttgcc tctttggctg gatgaagcac tcaccatctc ttgctttgac ..... 9300
agggctttcc taacttcaca agtgtggcca ttcacagtat gggtatttct cagtgacagt ..... 9360
cttacccatg gagtcatggt cgtcaaaagt tactaaagca aagccccttt tcttatcact ..... 9420
gccttggtca gtcatgattt caatcacttc aatttttcca tactattcaa aataatctcc ..... 9480
taggtgatgt tcttcagtgt ctttaatgcc accaacaaat atctttttca cagttgggtg ..... 9540
ggcatctggt ctttgagaat cttctcttga gacagctctc tttggttcca caactcttcc ..... 9600
atccaccttg tatggccttg cattcatggc tgcatccacc tcctccacag tggcatatgt ..... 9660
gacaaaccca aagtcccttg agcgcttggt atttggatct ctcattacca cgcaatccat ..... 9720
gagcattccc cattgctcaa aatggctcct caggctctca ttggtcaacc aatgtagagc ..... 9780
tcaacctcca atgaagaatt tcctcagctg tttgggctca ttaggagact ctgacttaga ..... 9840
catgacggca ggggaaagag agactttaag gatgcttcct tggtggcgtc cacgggcaga ..... 9900
aaggggtaag cgtccacagg cagagaggag taagcctttg aatgtatcta aaatttactt ..... 9960
tttattgcct gcattctttc actattttcc aaacattctt caattgtcac aaatggcaat ..... 10020
gataaagaga aaaatataaa catcacattt taaaaataag tgtaaaataa ctgtgaactt ..... 10080
aaaatgtgat catcatagaa gaaagagcac tgcaatagaa gtactgtgag ttctctggtt ..... 10140
aattttgcga agccacgtaa agctgtgtgg ctttaagcaa tgccataatc catttaagtc ..... 10200
ttagcttcta tttctgcaaa ggagaagttt gaattagtca atcttccttc cagatccata ..... 10260
actcagtttg ataaattact tagtatcttt ttctacagag aaaatgctca tacataataa ..... 10320
taaatatgta atcataaaat tattttcatt agtctgtttt atagaattca aattaatcac ..... 10380
cactatttac tcttgtgcct cttggtgatc ggtgctgtct gttacagatc gctactgaag ..... 10440
caatagaaaa cttccgaacc gttgtttctt tgactcagga gcagaagttt gaacatatgt ..... 10500
atgctcagag tttgcaggta ccatacaggt aataaccgct gaagagtggg aggagagtgt ..... 10560
gaataatttt tcaatcatca tatttgtttt cagagggatt actttggcta gaaggtaggg ..... 10620
agcaagtgga gaaagtgctc gaaggtaaac cattgagaaa cagttgtaat tatgcaggag ..... 10680
agaaagtaca agaccctgaa ctaaggcagg gacatctctg aggtagaacc tgtaagaatg ..... 10740
ggtcactgat gagaagggag ggagacatga tgctgagaat gactatctga tgtccaggta ..... 10800
ggatatgacc ctataatttg ctctagttga aaatgagtta tttatggaac ctgaaatttg ..... 10860
aggtacctgt gggacacaga gaggatggtg cttcctgggt ttacctgtct ttctgctttt ..... 10920
taccccttct ccagtgcaga ccttcttcct ttaatacagt catcccagtg agggctctaa ..... 10980
taagtggtgt gaacataagc aagcccaacc tttctgagtc tgtttcttat aataaaagga ..... 11040
aagctggact ttattgatgg atttttaagc ttttatttaa aaaaaaatga tggtagagcc ..... 11100
tttttgtcta aaaaaaaatt atttgaaatc ttcatatgtg atgaaggtaa aagcagagtt ..... 11160
gatctggtag caggaggggt tggaagccca gggctgccca cttgctaacc tgcccccacc ..... 11220
cacacctcca tatcactgag ctggatccat atcatgattc tagaatgtca accccttttt 11280 aagccttct ctgagacccc cgaagaattt agtgcttctc ctttcttcct ataacagatt 11340 attcataagg cacctctaat gcataaatag taattcaact caagtcaggt ctcctaagtc 11400 aagaacatgc ctgtttcctc tatgtcaacc catcatggcc cctgcacctt gtaggtgttg 11460 ggtaactgtg tgcagaatga atatttacgt agagtacttc catactgtgt gtccaaaagt 11520 ggggagaaag ggagataact tttacttatt gatccctaac cggctcttta cagtcattgg 11580 attattttct ctttacatca acactttgag gtgtagttaa ttctacatta aagataaga 11640 cactgagtct cagaggttag gcagtttccc aaatttgcac aactgttaag tgtaaagtga 11700 ggaccaaact tagttctttg gaaacgaaac ccatttttgt tggtcatgcc tctgtgccct 11760 actgccaacc tatcaaaagt tacattttaa ggactgagaa atgaaagtgg aatgagtttt 11820 caaatgtctt cttttcagaa actctttgag gaaagcacac atctttggaa ttacattttc 11880 cttcacccag gcaatgatgt atttttccta tgctggatgt ttccggtttg gagcctactt 11940 ggtggcacat aaactcatga gctttgagga tgttctgttg taagtattgg gctattattt 12000 agttaagctc taaaaataaa gctgggatga acatgcttca tgtctagata ggaaacccta 12060 ctgtgaagcc tcatgaagag attctggtga ttcctaaatc ggcatttctg cctctgagtc 12120 ttcatgtgcc accattgaag cacccccttt catttggagg agcagtaact tctttcctca 12180 ttgctggctc acacatagtt gaccttttca aatctgtgac tgagtggagt gattccattc 12240 ggagattttg agaaggcctt ggcatttggg aagaagccta gccctgagca aggagtctga 12300 ctggctcctt ttaaaggact ttcttacaga gcaagtaaaa tacagatgtg ttgtactaag 12360 ttctgcaagc ctttggcaat tccaggatat gtttactttc ccttgataag agaggaattg 12420 gaaggtaaga gccaaatgaa ttcagaaatg acaaaggaaa agttatattg ggatttttct 12480 gctacattgt tctgaatgta gataattgta cctcggtcag aggaagataa gcctgaagca 12540 attatactac aaaagtaccc aaatatgcaa tttggtggtc aaaggtatgt gtaatctttc 12600 tgagcttcta gatttggagg tgggtaagat tctgcctctt gatagcataa cataattagt 12660 agcgatataa ttttatattt aaaccagaat catataagcc tggcagtata atgtgttagt 12720 atagtatttc tgtccttttt aaacattgag ttgttcatgc attacagttt gctcaggatg 12780 acccccaaaa agacatctaa atttccatta aagatgtaca ttggacaaat gatatgcagg 12840 tgctatttgt gattatggcc tagagatcaa aaccaagtat cactggcatt ggggctttga 12900 ttagataatt atttgatata ttgctttact ccaaaaatt gaattgatga aagttgctga 12960 cattggggca tttaggtttg caaaatcaac cttccaagtt aaggaaaagg aagacctgtc 13020
tgaagaacag tgcattagaa aaaggtccca taggtttaaa tgatcacaag tccaagatta 13080
actatcggta ttttatttag tgctagttaa acaaacaaac aaacaaaaac taatgacaca 13140
atatgctaca caaacatagg catagagtcc agaggaaaga aagcgatgaa ttccctgtat 13200

[^1]| ataaatgtgt ctagagaagg gcagtgaag gagttatcaa agaagcaaag gagtttttag | 13320 |
| :--- | :--- | :--- |
| caatatttca agacttgtag gcctgtctta tgtaaagga agtaaatttt ttcttcagtt | 13380 |
| tgaaaagaa ctgtttaaca atagacaatc aacatgcta cttcctaaaa cagaggatgg | 13440 |
| aagccaattt aggggaggtg tccaggcacg aacatggaga gctggacttg atacctgtaa | 13500 |
| ggtccttccc aactttaggt tgctgtgatt cccatgtcat agataagaac gtcaatgcat | 13560 |
| cttaagagca acatgatatc tggctctgta agaaactttc ttttggttgc acaattcata | 13620 |
| ggtttttaag aatctgatgt aattccaaca tcactgactg tatccgttgt tgattaccac | 13680 |
| aataatgctg tgtaacaacc agccactaaa cctcagagac atacatattt ttgcccacaa | 13740 |
| acctatgggt tagctgtgaa gttctactga tctggattag gcatagtgga tctcggctgg | 13800 |
| acttactcat gtgtctgtag tggattgtgt ctgtagctgg ttggttggga ttcagacagc | 13860 |
| tctcctccat gtgtgcctca tcctccagca ggctatcctg ggtttgagac agtggcagaa | 13920 |
| ttctgagaga gagagagaga gaagagaga gagaacacat gcatgcatgc tcacaaagca | 13980 |
| tataaggccc cctgagtcat aggcttggaa ctggcaaag tgatttccac catatcctgt | 14040 |
| tggccaaagc aagtcacaag gccagcccag attcaaaggg tgctgcaaag gtcacagtgc | 14100 |
| aagaagatga ggatacagag agaggtatac agagggaaaa aaatgggcca ttttgcaatc | 14160 |
| gttatgttgc ccaggctggt tatcagattt tattgtatgt aagttactgt attcctgagg | 15300 |
| aatctatcac atagacatga acttataagg aaatgtgttt gttttatttt taacatctgt | 14220 |
| tttataacta ttagaggcaa agctctctgg ttatagaagt gtcaactttt ggccaggcat | 14280 |
| atactgctca ctttataagc aaaacctatt aatcagattt tttaaaaaca attttttttt | 15060 |
| agagacagag tcttactctg tcatccaggc tggagtgcag cagtatgatc atagctcact | 15120 |
| ggtggctcac acctataatc ccggcacttt tggaggctga agcaggaggt tcacttcagc | 14340 |

aacagatttg agttattgta gctgtattgc atattcatat tgtcttaaca atacatgcta ..... 15360
tgaaagcttt tactctttta gatctcattt attaaattct agcagtctga ggtcaagcac ..... 15420
agtggctcat gcctgtaatc ccagcacttt gggaggccaa ggtgggtgga tcacgtgagg ..... 15480
tcaggagttc gagaccagct tggccaatat ggtgaaactc catctcaaat aaaaataaaa ..... 15540
aaaaaaagat tagctgggtg tggtggcaca cgcctgtaat cccagctact tgggaggctg ..... 15600
aggtacgaga attgcttgaa cctggggggt ggaggttgca gtgagctgag atcatgccac ..... 15660
tgcactccag cctgggtgac agagtgagac tctgtctcaa aaaaataaaa aataaataaa ..... 15720
taataaaatt aaaataaaat aaattctagc agtttgaagt gaagccaatt gtaacacaaa ..... 15780
ttaattatct tctgacacct ggtaatcgag agagttagct atacacttta ttttcagtat ..... 15840
tgcagcattc aaatttactg ttattcttct cattgcagaa cacattggaa ggaaatgtca ..... 15900
catttggtga agttgtattc aactatccca cccgaccgga catcccagtg cttcagggac ..... 15960
tgagcctgga ggtgaagaag ggccagacgc tggctctggt gggcagcagt ggctgtggga ..... 16020
agagcacagt ggtccagctc ctggagcggt tctacgaccc cttggcaggg aaagtggtga ..... 16080
gcacactttc acatttagct cagttcaggt tttcatcatc caaatgtctg aatgtattta ..... 16140
attctcaact ataagccatg ttttttcaaa cctttaaaca acagtcccac ttggataaag ..... 16200
tctgagagcc taaatatggt ctccaagtgg tgtcatctgt cccagccaac ttctccaggc ..... 16260
tcccctcaac actacccotc taccctcctt tgcaagcacc ctttgtacca cctgtctccc ..... 16320
ttgctgtact taggtttcag ctgacttgaa aagaaccaac aaaaatggaa gtagccagaa ..... 16380
agatacagga caggtgctag gagtgagatg aaagccaagg atggagactg tttcaaagat ..... 16440
ggtggtcagc aatgacagat ggtatagaga gattgggtaa ggaggagact gagaagacat ..... 16500
gatagaatct agcaatgaat gggctatttc tgacatttta aaatattctt agtgaagtag ..... 16560
tgatggtagg aaacagttaa gggtggctga agtatgatta ggggaagaaa tggagatgta ..... 16620
agtgagcata gattatcaag aagatggaaa gtaaaagaaa ggcaaataag gacagcactt ..... 16680
gagtggggag gcagagtcca gggaaaatgt ttaggggttg aactcttgag catttttatg ..... 16740
aćtggggatg agacaaaatc attgatggag agagggcatt gaaacatcat ctgaggaaaa ..... 16800
aaaagtagaa tcaacaaatt tgggaaatgt aaaaagaaca tggaagtatg cccttttgtc ..... 16860
ttagttgctg gggatgaggg aagtgtctgg cagggagtag tgagtttaaa cgactatgat ..... 16920
gggggatagg aaagaagaca aactaggaag gaatagaatt aagagcaaga cattttcatt ..... 16980
atgtaggcat gtaattagtt gtactgagaa gtgatattta aaattcttaa caatgggaac ..... 17040
agcatggaca ctgaccaatc agaatggatg cctgctagta caagtcagat ggatgcccac ..... 17100
ccgctgtgta tggacgctgg tagcatttaa tgactcaaat gaaaagagag ttttcaatct ..... 17160
ttagtctagg actcttttta cttcctcaca gcggtcactc ttactcataa atgtttgtta ..... 17220
ttgtattcaa agacagttat gcctttattc agtttactct taaatcatta caagcttgtt ..... 17280
ccctactttt actctcttgc tttatatttt ctgtgaactc tgattgtgtc cttcaatttt ..... 17340

| gtggtgaact | taaaaacagg | actctaaata | aagacgtcgc | tgacagccaa | agcaattaga | 17400 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ggaggaaatg | ttagctctca | gacactgatt | atggccctca | acagacagtt | attaagtggt | 17460 |
| gtgttaaaga | ttgtgctata | atgaattgt | gggcatgata | tctgccotca | aatactttaa | 17520 |
| ctgaggagat | gggg | gttttgt | ttaaaggaga | taggcctaga | gctagttctt | 17580 |
| aagagataag | gatttgcata | tggg | gtagggatga | cattcaatgt | aagagaaaat | 17640 |
| acaagctgcg | atcagtgatt | tg | gc | ggacactttg | gctgaaatgg | 17700 |
| gagtttagt | tgagtagtgg | tc | a | gaactcatgt | aggaaagatt | 17760 |
| taaagatgtc | accatcca | aatgtactca | tagtcacttt | ccagtgggaa | aaacacttgg | 17820 |
| tgataacaaa | at | g | atacaatac | agcacaataa | tgatacaata | 17880 |
| aggaataatt | aatgattatt | Ct | ataaacacaa | gcagaacata | ataaaacatg | 17940 |
| gacagcatca | gttacagttc | agagtagga | acagaaactg | ccctaggtat | ttcaagtaga | 18000 |
| taagta | ata | tta | acaaaatctt | tggaagggct | ggaagagtga | 18060 |
| aaggaggcta | tttgctccca | ga | tcacagcact | gcagctgtga | ttctgtaacc | 18120 |
| aagaagctgc | agaaaattat | $g$ | cagctgttcc | aaatttaaag | acacgagact | 18180 |
| agaatctggc | tgttgcag | ttcttgtct | gccaaagatg | agttaagccg | ttagcttacc | 18240 |
| acagctg | aag | $g \mathrm{gt}$ | tcccagattt | tacactgtgc | atctccctcg | 18300 |
| tagaccotga | cttacttcca | gaaccaaggg | aaactgggaa | atagttttta | gccttctagt | 18360 |
| cccttgatgt | ataaaaggtc | agacacagaa | tatatgaaa | atgaatgctg | agtgcatagg | 18420 |
| acagtaagc | tcccaagaca | ttattgataa | tgcactgaaa | acaaagtgtc | ctcatttatt | 18480 |
| gcaggaatat | ccattatctt | ctcctgtgca | tttgcactag | ggcttggaaa | tactgccaga | 18540 |
| tggtgcactg | cagccagact | tctcagcaga | aagtgtccac | aataaactct | agtaaatgta | 18600 |
| caaagttttg | tttatggaca | tcaatgagta | gccetgaata | actcagtgga | ttactactaa | 18660 |
| ttagttttta | ttgccatgta | aaaatgagtc | tgtggatttg | agcataagaa | ttaaagaaag | 18720 |
| gattggaccc | taaggacccc | ctgaagtcag | gaaattctcg | tgaccatcag | tggaccttga | 18780 |
| atcccatgtt | gaaaaacatt | gacctgaaat | ggtggttcta | aagcttcggt | gaatattaga | 18840 |
| atggcctcaa | gagctagtaa | aaaacacagc | cggcctggat | tattcaagta | ggctagggtt | 18900 |
| tggcctttta | tttttataat | attccgaggt | gattctgatg | ccaaccatca | ttcgagagcc | 18960 |
| actgaactgg | taaatttgaa | gaaagattta | ggttaaattg | tggagagtct | tgaataatta | 19020 |
| agctatgaaa | ttgtgaagat | gggaggatcc | tccaagtttt | gagggacctg | aaggttatac | 19080 |
| aattgtaggg | attcaactta | aagtaaaaaa | aaaaatttt | ttttttaatg | caaaattaga | 19140 |
| tacaaaggtg | gatatttact | tcaaatgaaa | aaaaaatcat | aaccaaatac | tggaagctta | 19200 |
| aagattctgg | tcccttttca | tttttaggta | atttatcatc | aatacttaca | gggaagagct | 19260 |
| tcctgactac a | agccttgcct | ccctcccacc | cagaaaactg | ttcgacttcc | cagaatttac | 19320 |
| atgtaattga | gggctcttga a | agcttaacct | tcatcagtgt Page 78 | cactgttaaa | tcacccotgc | 19380 |



```
atgattacat tagtttcacc aagataattc aggataatct tattttacag tcagctgatt 21480
agcaacctta catctactac tttagtttct ttttgccatg taacataaca cattcacagg 21540
atccagggat taggacacag atgtcttgtg ggagagggaa cattattctg cctaccacat 21600
gcatacatca gaaaccatgg ttgaaacaca ggaaacatga cagttcctca aggcatacaa 21660
ttatgacctt gttgggttaa ccttcactat ccaaatttta atcacacaaa cttttcctta 21720
atctcacagt aacttggcag tttcagtgta agaaataatg atgttaattg tgctacattc 21780
aaagtgtgct ggtcctgaag ttgatctgtg aactcttgtt ttcagctgct tgatggcaaa 21840
gaaataaagc gactgaatgt tcagtggctc cgagcacacc tgggcatcgt gtcccaggag 21900
cccatcctgt ttgactgcag cattgctgag aacattgcct atggagacaa cagccgggtg 21960
gtgtcacagg aagagatn
```

<210> 9

```
<210> 9
<211> 40962
<211> 40962
<212> DNA
<212> DNA
<213> Homo sapiens
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (18985)..(18985)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (40962)..(40962)
<223> n may be any nucleotide
<400> 9
nctgaacctg aaggtgcaga gtgggcagac ggtggccctg gttggaaaca gtggctgtgg }6
gaagagcaca acagtccagc tgatgcagag gctctatgac cccacagagg ggatggtgag 120
atgacccatg cgagctagac cctgcggtga tcagcagtca cattgcacat ctttctgatg 180
ttgccctttc aattacaaat gtatgaaagt cacacttact ttttattcca ggtcagtgtt 240
gatggacagg atattaggac cataaatgta aggtttctac gggaaatcat tggtgtggtg 300
agtcaggaac ctgtattgtt tgccaccacg atagctgaaa acattcgcta tggccgtgaa 360
aatgtcacca tggatgagat tgagaaagct gtcaaggaag ccaatgccta tgactttatc 420
atgaaactgc ctcatgtaag ttgtccttgc cctttgcctt tctagaggtg caaaaaataa 480
aatgcaggcc tactatgcag gaagttagga aactactata aatcggaaga agggaaatcc 540
taagaaggga aagtaagatt acttcagatt tgaaagctct agcagtatca actggtcgta 600
gatacatttt taaaaactga ggttggttat tgtgttaaat aagatttaaa gaactggacc }66
tgtattactt gtgagacttg ggctgtgtat aggattcctt accaatttaa aatatgagct }\mathbf{720
Page 80
```


agactcctgg gctcaagcaa tcctcccgcc tcagcctccc aaagtgctgg aattacagga ..... 2820
gtgagccact gcttactggt ttgcttatct gtgtttcctt attaatctat agtgaaacta ..... 2880
tgtattaaat tataaataaa aacaatttta aaaggttata ttttaaaata ctttagggtg ..... 2940
taaattttga ggggaaattc cacatacccc ttttttctta aaaagataca aaaattgatc ..... 3000
tattttcttc tgtattttct agtttctacc acctaatttt tccttgtgta ttttttcttt ..... 3060
ttgaagtttt ccacttctac ttatcctatg gatcctgaaa atgttgtgtg ttggttttga ..... 3120
gaattgtatt gctagttatt agagagacat atagagtaac aaaaattatg agcattggga ..... 3180
aagttacaaa ggttagagaa gtctcagaca aggcctggat atctggctct gttcctttat ..... 3240
gaaaataaaa gagacttacg tgactcttca atttttccat aattcttcaa cctaggaata ..... 3300
agtatcacta actatggata aggcacagtg ttgagtactt tatgtgcttt attttattta ..... 3360
gtcatcacaa ctactctaag aagtaaatac tattattatc cccattttac atatgaataa ..... 3420
attgagtctc acacagtttc cttggataaa atattttatt ggataaaata aattcataaa ..... 3480
tttatttcag gtcagtgtga cattgaggtc tggactttgc tgcctcacat ttattgtctg ..... 3540
tcttgttcat ccagggggtc atgcgggata ggatattata attcctaggg ctgattacta ..... 3600
gccggtgtgt atcagtacag cacaatggcc tgtgtttgtt ttgattggcc aacgcctggt ..... 3660
ctgtaggaat ttgttggttt gtacaagccc ctgattatta ttattttttt attttttatt ..... 3720
tttttttttt gagatggagt ctcactctgt cacccaggct ggagtgcagt ggtgcgatct ..... 3780
tggctcactg caagctccac ctcccaagta gcggggacta caggcacccg ccaccatgcc ..... 3840
cagctaattt tttgtatttt tagtagagat ggggtttcac tgtgttagcc aggatggtct ..... 3900
tgatctcctg accttgtgat ccacccacct tggcctccca aagtgctggg attacaggcg ..... 3960
tgagccacca cgcccggccc aagcccctga ttattactgc aaatttaggt taaataaaat ..... 4020
atttgggggc ttacataata ttaatatgtg actgttatat ttgtgtttgt atttattaca ..... 4080
aggaaacatc atttttaact attatcaatt gtctatacat ttattgaagt cagaggctat ..... 4140
cttatataga tttgatggtt ttacaatgcc cacagcattg gttcagtaaa tatatgttga ..... 4200
atggttaagt ttcttcaggt aattgttaat gtattcaaaa accaaatttc tctctcttta ..... 4260
ggccagaaaa ggtcggacca ccattgtgat agctcatcgt ttgtctacag ttcgtaatgc ..... 4320
tgacgtcatc gctggtttcg atgatggagt cattgtggag aaaggaaatc atgatgaact ..... 4380
catgaaagag aaaggcattt acttcaaact tgtcacaatg caggtatagt ttaacttcag ..... 4440
aattttccta agtcatctca gtgataaact gattttgcat ttaatgctaa aaataaatat ..... 4500
tatttgattt gattacctta caaagtagga aacaacacct gggggattca ggatgagacc ..... 4560
agtgtttaag attttttttc tctcttgaaa gaggggaaaa taaagaagga taaacagata ..... 4620
aaaaaattaa aaggtttcaa ggtgagttat ccgttatagt agtcagtagc cacatgtgtc ..... 4680
tgtacagcat ttaaatggta gtgaatctga attggaatgt tttaagtaga gtgtgcacat ..... 4740
cttgttctga agacttagta caaaaaatgc aaaatatctc aatttttata ttgattacat ..... 4800

cactgacttt tgattctaca aaatatatat gactattcag gtctactcat gcctggttag
attgtttcta gcccttggaa ttatttcttt tattacattt ttccttcagg taaatgtttc 8940 cattttcact atattcattt tgagaaatgc atcattattc aactgggggg atttataaac 9000 atcagtgtaa ttggcctttt agtagaactt ctctattaac atggctacta acagccaagt 9060 ttttctgaca tagtgaaaaa agatgtttgc ctcctctggc tcccttgctt accttctcct 9120 ctccaccctt acctcccgca atgaaaccag ggggaacaga gtatttggcc tgatatgatg 9180 attggaggtg aaaggcaggg acttcaaaat ggggtgtggg ggagccctgg atgtaagttt 9240 ggatataagt attgccagta aatgtaaata ctcaaagaaa tgtcttcccg tgttctaaaa 9300 gcaacaacaa acaaacaaaa ccccataagc gggtcacatg ggtttaaaaa gggtttgaag 9360 agtttatcgt tcaacttttg ttcaaggaga aagaatcatt tcagtgatgg aagaatgtca 9420 atcctccaac aaaaatgta ggaaacattt gtaaagagtc agatttttac agcttgcaaa 9480 tcatttggct gaaatgaaat ggcaaggaat taagtactct aaaagtttag tatgagggcc 9540 aggtgtggtg gcttatgcct gtaatcccag cactttggga ggccgaggtg agtggatcac 9600 ctgaggtcag gagtttgaga ccagactggc caacatggcg gaactctgtc tccgctaaaa 9660 atacaaaat taacagggcg tggtgatcca cacctgtagt ctcagctact cagtaggctg 9720 aagcaggaaa tgtctcagga acccaggagg ccaggaggag gaggttgctg tgagccatga 9780 tcgtgccact gcactccagc ctgggcaaca gagtgagtcc ctgtctctaa ataaataaat 9840 aaataaataa atactgtttg gtatggcaag acagtattgg ttttggttca agtgctcctt 9900 gtacctgttt ctgattttgt gtccttgggc acataagtaa actgtctaag cctctgtttc 9960 cttagcggta aactagggat gcaggtacct gcctcttgat gattaaaagg atcatgtgac 10020 caagtgctca ggcctgcatg aggcagtagt aggtaatcac ttattaattg aatgattagc 10080 cacctgtaat ttttaaagat aaaactgtga ctagatctct ttatttaaca tgtaagcatg 10140 tattacattt ttaaaaaata gaatattttt ctggacatta taagaggtgc aaaagaaaaa 10200 acagactgaa ttctttcttt actgagtact tacagcctca ctggggaatt tgaccttact 10260 agaataaatg tgacactcta actacttata ggattgccat accaactatt aataccacag 10320 gtgtttggag gaggtaaaga ttgctcaatg taaatttggt taagaaagtt ggagtgaggt 10380 gggctttgag ctggtccctt aaatgttgac gatttgaatt ggtgaagcaa aagaagttag 10440 aggactacct tcataacatg gtgcttgggg ggactggaac ctagcttgcc tagaaaacag 10500 gtgagaacaa agtctatgag ttataggacc ttagaagtct agagatagca aaatgcctaa 10560 agatgttaga ccacccaact cctcatctta gtaacatggg gagggagtcc agtgggtaat 10620 taaaagctca ggctctggat tcccattctg gttaggattc tcatgcatct ggattaagat 10680 tccaactctg tgaccttata actgtgggct atggatctta tcaggctctc taagcctcag 10740 tttcttctgt aaagtgggct tttctgtcta caccacccgt acagccttgt gccatggcac 10800 acagtcacag aaacatagca agcccttgaa atcaggcttt ctgactttgt ctaatctcct 10860 gctttagcaa agacatcaat tctccctcct tttatttaaa tggtggctgg gtccctgaca 10920
aggtatgttc ctgcccacag ggtttcacat ttggcaaagc tggagagatc ctcaccaagc ..... 10980
ggctccgata catggttttc cgatccatgc tcagacaggt atgtctatcg agggctgtgc ..... 11040
cctgggatgt gtagaactcc ccatgtgtgc cccttggact cagacagtgg gagctctgtc ..... 11100
atgtttccca ggcctagctc ctatattggt tgtccctcag tctcacgtca gagatgcggt ..... 11160
tgaaccgccc actaggcaca gatgaggctc tactgtctgt caggtctggg tgggcacaag ..... 11220
gaactggacc agtttgagac agagcctcca gcgtggattc acctccagcc tgtgtctcag ..... 11280
cagtggtggg cagtgcaggc agaaggcatc tttgataccc ctcatccctt tcctctcttt ..... 11340
ccttctttct cctggcaaac ccagggcacg agtgtttctt cccaccaaga gatgccctct ..... 11400
gtactctttc tttccgcaaa tcaaaactca tccctgtgtc ctgttcttca ctgcccattt ..... 11460
tcttttctga ggctagtctg aaatactagt tcaagctgca gtgttactgg ctgataatgg ..... 11520
gtttaatgga atgcttctca cattttgcag cttcaaaacc ctaaccattg acacgtgtga ..... 11580
atgttttcct ggggaaatgg ggaaggaaat tagaggaatg taactcagag cagcctggtt ..... 11640
caaaggggaa agttccttta acaatcatga aaattttgta tgtgacctaa taactttccg ..... 11700ttttaaaaat cactaatcac ttgccattga gtaaaatgat gctttagaag tctgccccag11760
atgtgccagg ggtactcgaa ccctggctaa gaggcatcag tttggtgtgt taggctttct ..... 11820
agagggcatt cacattttac cagctgtctc tggttcctca gttcttcccc attcctccca ..... 11880tcaccattta gaaagaaaat gtatttatgg gaattgctag ctagtgactg acagagccaggactgaatct aagtgagaga gcacagtttg atgggaaacc ctgtccttgg actgtcaggtcgaactgtat ttataagtca gattccactt aggtctacac tgaccttgct ccagggccaaatttcccatt acccaaccag cctccaggcc aactgctgtg cccattatac tttggcagctgagctgatgg tttgtggaat gtctcctcca taaattgtta agtagggcaa gcatttatta1194012000120601212012180agtgccttct gcttacaagg tctagtactt agtattgtaa ggcattcaaa cctgaatggt 12240ccctgtgttc aaggaagtta cattcctgtg tggagacact tttaaccact taagaatatt 12300gaaaagcaag ttgatacata ctataataaa ttatagcact attttttctt aaatattttt 12360gggaaaaatg atagatactt ctttaatgta aataaccagt ttagataact tcttagggac 12420cacagttatt tgtaaaacat ataaattcat acttcaaatt cacaatcatg ttagtatctg 12480tgtatttaaa aatataaatg gttttacaaa gaaaatctta ttttatgttg aatgttatat 12540tttaatctgg attacttttg ctgatttgtt tttgactcaa atcacttata ctgccctgag 12600ctacatttat ctaactgctt attcaacctc tctattggat gtctaataag agtttcactg 12660tttaacattt ccaaaatgga gcttttgact cttcccctcc ccaccagcct tattccttac 12720ccactcttcc cctatatcag taagcgacag ctccgttcta ctagttgttt gggctaaaaa 12780ccttgaagcc atctttgact cttctctttc tgttaacatt gcaaaagctt cctaactgat 12840ctccctacct ccatttttat cccaatcctg tagattcacc taaaaacagc cacagtcatt 12900tttctaaaat agaaatcaaa tcatatccca ctgcccaaac ctgctgaagg ctgcctgtag 12960
ctcacagggt aaagtctgga atcttgacag tcgcctgtaa gaccatacac tatgtggccc ..... 13020
ccactctctc tcaaatctca tctcctataa ccgccttttt ccaatacacc gtctactcca ..... 13080
taaacagtgg ctcactgccc aagagtaggg gaggggagga tcagaggagc caaatcagaa ..... 13140
cacaaatctg tccagaggaa gggggtggct tttacaaatt catacaaagg ttctttttag ..... 13200
gctgctggca gcctttaagg ctttcttgct ggtgtatgaa cagattaggt atgcctttgc ..... 13260
ctcagggtct ttgcacatgc tgtttctgct ttccctgagg tattctcatc ctttccttca ..... 13320
ttcaggcctc tgttcagatg tccccttaga aaggccttct gtttcccctc cccagtctct ..... 13380
aaaatagcac ctcccctcac ttttctgttc ttcttaccct gctcaatttt tcttttattt ..... 13440
gtaccttcca tctctagaat ctaaacttca tgaaagcagc tactttgtct tttttgttcg ..... 13500
tacaggtcca acacttagaa atcatgcgta ggagaaagta ggcactcaga aatttttctg ..... 13560
acaaatgaaa tgatctattt atgtgttttt atattaagtt tctttcttgt gtattgaatg ..... 13620
tcacatcctg agtactaaat gcagggggta taagtataaa caaaactgac cccatcgctg ..... 13680
ccctcttgga gctgagagtc tcataaacag ctttaaggta ataaaatcat tttctgtgcc ..... 13740
acaggatgtg agttggtttg atgaccctaa aaacaccact ggagcattga ctaccaggct ..... 13800
cgccaatgat gctgctcaag ttaaaggggt acgtgcctcc tttctactgg tgtttgtctt ..... 13860
aattggccat tttggacccc agcatgaaac taattttctc cttacgggtg ttagttatca ..... 13920
tcattaagaa aatgttgaat aaatatctaa cctacgaata tatcacatgc tttttgtagc ..... 13980
aacatgttaa ctatttaaac attatatact gtagagcata tagataactt ataaaccatt ..... 14040
tgctattgct gttattcatg ctattaacaa gatgcatgta gaatagttat ttagaaaaga ..... 14100
gagtataaag tgctcaatca acataaaaca gtaattgcta ctgaagaaag gatgtattta ..... 14160
attgctgtaa gaaagtttag agtcactatg gttacagaag ggagggaaga caatcctcta ..... 14220
aaatataggt tgaaggaaat gaaaagcaca ttaaaaaatt aaggcaagaa tagaataact ..... 14280
tcagtcttta tctttaataa ctttaaactt taataatttt aataacttaa attttgctac ..... 14340
tgtatgaatc tcttgatata actagatact attgaaccag caggttttga tttttggctg ..... 14400
aagtgacaat ttcttctaca actgtttatg gcaaaagtcc acaaaatgat gtagaatttg ..... 14460
aaaaattca tgtaatctct ggtgtgtctt ttcccctctt gaaccttatc catctttatc ..... 14520
tttaaatctt ttctgtaagt tagtactata ctaacatttc ttctatctaa tatatggggc ..... 14580
ttctttaaga ataaattaag ctataaatga ggaaatacat agagttataa cgttgaaata ..... 14640
taaaccttag gagtccctct ttttctattg tttggaatag tttcagaagg catagtacca ..... 14700gctcctcttt gtaattgttg ttttaataca aacttctttg cctaaagcaa acaaaacaataaaaatcaag gtttagatca agttgtatag aatgtaatta caggtgcacg cctgtaatcccagctactcg ggaggctgag gtacgagaat tacttgaact caggaggtgg aggttgcagagctgagattg caccactgca ctccagcctg ggtgacaaag caagactatg tctcaaaaaa14760148201488014940
aaaaaaaatg caaagaagac agagtggctg gaataaagtg agtgaaaaga agagtcataa ..... 15000
gtgtgttaag gtcagcatta tatccagaag tagatggaaa accactgtag ggttttgaac ..... 15060
acagaagtga catgatctga aattttgaaalggatcactat agaaactgtg tgaataggcc ..... 15120
gaagggggca agcatagaag gagtctgttg cagtaatcca ggaggagatg atagtgtttt ..... 15180
agactaattt ggttatacaa aaggctataa gatataatta attctggata tattttagat ..... 15240
gtacagccaa caatgtgttg gttggataag atgatggata tgagagaaag ggtatttaaa ..... 15300
gatgactcca aattctttta cctgcacagt ggaaaaaaa atggagtttt tttttttttt ..... 15360
ttttttttta aagagacagg gtctctcttt gtagcccagg ccagactgca gtggcatgat ..... 15420
tacagctcac tataacctag aactcatggg ctcaagcaat cctcccttct cagccttcct ..... 15480
ggtagctggg accacaggtg tgcaccacca tgtcgagcta atttttaaat tttttgtaga ..... 15540
gacagggtct ccctatgttg cccagtctgg tcttgaactc ctgggctcaa gtgatcctcc ..... 15600
ctgcctcagc ctcccaaagt gttgagatta caggtgtgag ctgccatgcc cagctggagt ..... 15660
tgatctttat gaaattcaga agcctgttgg agaagcaggc ttgggggaga atggagagtt ..... 15720
ctgtatgaga catggtaagt ttgtgacgtc tgtattagac atccatatgt agatgtcaag ..... 15780
aaggctgaaa tttgcactgc taactttagt aaaccttttt tataattggt ttactaaata ..... 15840
agttttacta tgcttctcca ttcattttgg tcctcacaac tctatagact cctactctgt ..... 15900
aggaggaata tacatgagac agtgagcatt agtctttgga ataaaggaac agtaaccagt ..... 15960
gcaatgtgac attgcacaat atgacacaga ccctgtggta tgggctcatg tgctttgatg ..... 16020
gacaaatatc ccgcatcaca tttgacgtgg aagacacaat cctggaggca actcagtctt ..... 16080ttgcagcaga acccagagca gatgtattca gggcattctg tattgcagtt tttcttcctg16140
cctctcaaat tcctctggtg ttatgacctg ccttaggcca cagtgagttc ctatatttgt ..... 16200
aaaattggtg gtcacttttt ccctccatag tgctgtttgt gaggcccttt gccatattca ..... 16260
ctgtctctaa ttgcctgctg gggtcaacct tctgcttttc acttgtttcc tcaaagacac ..... 16320
ctcaaacttg gccctgccaa gatggctgca ctcaccttat atgagccact caagcaaaac ..... 16380
tgttttccag aacttaaaac agtggctaaa aaggaaagac cagggaagag gaaatgagca ..... 16440ggttggcaaa cattgtcaac ctaggaaagc ccaatgtagc tgttatcaca gattgactgagagaggactg ttggatctac tttacaccac tagctgacct gtttttggct gacaggtttt1650016560
agttcctccc ctcaacccta gtctttacaa tgaaaatttt ctggctacta atctgtgctc ..... 16620
ttccattctt cttagtcccc atatttctat agactcctac tctgtagaaa tagatgtagg ..... 16680
acagtgagtg ttaatcttca gaataaagga atagtaatca gtgtaatgtc tgaagtgcta ..... 16740
aatcaaatga agccgagggg ttaaatctcc ttgcagttta aggtcttaga cttcttagtt ..... 16800
actctttttg gtacattaaa gaatttgcca tctttgatcc atcttattta tttctgtggc ..... 16860
ccttcactac tatccaaagc tgcacattta tatgtctgaa aagttaaata gccaggtaat ..... 16920
tcctctgcca tgaactcaca cctagaagtg attctaggtg atattctaag agcttctctc ..... 16980
acaagcccat aagtcagtag ctgcagtggg aaatgagctc tgagtagaac ttgtagatac ..... 17040


| ctctaagaat | gtgtttataa | cattcctca | agcaa | tttcatgctt | a | 19140 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ttagaaactt | acagaaagta | gcaaagccag | ttcttggact | caaaaactga | taattaactt | 19200 |
| taacagactt | tttcagt | caggccattg | tcttcacact | gttcttcctt | cctccccact | 19260 |
| ttcctccttc | cttagttat | ttcttcttt | ttttct | actttcact | tctctccact | 19320 |
| ccttccttcc | ttctttcctt | cttccttct | C | ctttc | ttccttcct | 19380 |
| せcccttcttc | cttctt | t | tttctttctt | ctttctttc | ttctttctt | 19440 |
| tctttctt | tttcttgctt | tgctttc | ttgctttctt | cttectgc | せtctttctt | 19500 |
| ttctt | cttaaa | t | gagg | gcccatttt | tttgtacat | 19560 |
| gtgagggggg | agattaaata | tggaaaaatg | ctagggg | ttattatatc | tgttttaaat | 19620 |
| tactacca | ctttcttttt | ttttatcatg | c | atcctattt | ctgtttatct | 19680 |
| ttaccctt | t | ttttttcccc | tgcatactt | ctttttttt | cccattattt | 19740 |
| aacaaatgct | tatgt | ttactgtgtt | tccaggcaaa | tgtcttattc | ttatagcaa | 19800 |
| ccatatggg | gtctattatc | cattttt | agtggggctg | aggcacaggc | aggtcatggc | 19860 |
| tctgga | agagctga | ggtcttggag | ctggga | agctcagaca | $g t t g t g c t c c$ | 19920 |
| aaagtt | ctctttctgt | a | agag | tgatggcaga | atgcagtctg | 19980 |
| atatcacatg | atctgta | tagtggaaat | gagaggtca | agcagggctg | actgccataa | 20040 |
| ctaacattta | ggacagggat | tatgtgtga | tgaacattc | gagattccca | ggagttaggg | 20100 |
| cagggactca | cacagatcag | agtggctctg | gttgtcagt | ggcccagcta | cctcaccaag | 20160 |
| tgaatg | agaaggcccc | agatg | cattgccact | atgctttgt | cctttacctc | 20220 |
| tctgctatct | cttcagac | tatcccat | ttttgggggg | ttccctctgg | azaatctttt | 20280 |
| gggccagctg | tagtacctcc | ccagccagtg | tctgtagggg | caaacctca | tgcgaggcat | 20340 |
| cctatcaggc | tcccggaggc | tatttctgg | aggtacat | ggatggtgct | gacccacagg | 20400 |
| gttctattgt | ggtctggctc | gtaac | ggcccctgc | agggaataca | ttctgtgcca | 20460 |
| atgacttcct | ttaatattgt | actttggggg | tctcccccat | ggtgttgact | cttcccctga | 20520 |
| gaagatggaa | aaatctagga | caaaatataa | aaaaaggaga | aagcattcta | gttaacagtg | 20580 |
| tcattattta | atgaggtatt | gaaatgtcc | agaagggag | tcttagattg | cttttaggaa | 20640 |
| gaagatgata | caatctgatc | ctaagctaat | tcatgcaaac | cacaggttag | cacctttgaa | 20700 |
| gtgacacatg | agatttaaga | ctcttggagt | tctgtcagag | gggccaaagg | aaggggagca | 20760 |
| gatggaaatg | aattatgtgt | ggatatggca | gattttctga | gcaaaggtca | gggatgtgat | 20820 |
| acattttcta | attcaggggt | ggctcatgag | agggacagaa | gtaggtaggg | aaggtagtgg | 20880 |
| atctgcacca | accagtccac | ataatattga | aaatcaactt | ccacctgaac | acacttctca | 20940 |
| agctgctatt | tagtacttct | tatataacaa | gtatcattct | aacagctcta | aatacttgat | 21000 |
| ctaatttaat | gctgtcaata | accctgttga | attatccctg | ttggatagat | aatgaaaccc | 21060 |
| aagattcaga | gagatcaagt | aacttgctca | cagtcacaca | gcaacagcct | cctaaccaat | 21120 |

ctccctgctt caacttcttc cattcttatc caactatagt ctacattctt acagaagtca 21.180 ttctgacctc tcataaataa aaatctaata atgccaaccc ttacttaaaa ctcccaaatg 21240 gcttctgatc atacttggaa caaaatctaa tcttttatca tggattacca aactctgtat 21300 gatctaactc ctgatgacct ttatggcctt atctggcatc acagtctctt atgcctgtgt 21360 gccctggcce agtgggcatt cttgcaattc ctggaatgtt ttaagctgat tgcctctgat 21420 cttctcttgg ctgactttac gttattcata tcccaaaatt attacctctt cagagaggcc 21480 attcctgaca cctgaattaa ttcatgatac ttcgcagccc ttgatgtttc tttcctattt 21540 ctttgtttat tggttaatgg tgtaccattt gtttattgat taatggtgta tgtgacatta 21600 gagcccacac ttattctgac tgcatgatgt caaaacctgc tcataacctc taccctgcct 21660 gtgatctgtg ctctgcagtt tttcagcaca ttattgtgga ctagtgagca tttttataaa 21720 atacaggtga cccttgaaca gcacgggttt gaactgagtg ggtccactta tatgtaggtt 21780 tttttcaata aatgtattgg aaaatttttt gaagacttgc aacaatttga aaaaactcac 21840 aagccatgta gcctagaaat atcaaaaaat taagaaaaag ttaagtatgt tataaatgca 21900 taaatatat gtagttactc gtcttattat ttactaccat aagatataca caaacctata 21960 attaaaaaaa gttaaaagtg atcaaatttt atgcacacaa acacttacag accatataag 22020 tcaccattca cagtcaagac acatgtcaac aactgtaaag atacagtgtt acatgttttg 22080 gctctgtgtc ctcatccaaa tctcaccttg aatggtaata atcttcacat gttatgggag 22140 ggtcccagtg ggaggtaaat gaatcctggg agtgggtttt tcccatgcta ttctcatgat 22200 agtgaataag tctcatgaga tctgatgatt ttataaaggg gagttcccct gcatatgctg 22260 tcttgcctgc tgccatgtga gatgtccctc tgctcttcct tcatcttccg ccatgattgt 22320 gaggcctccc cagctatgtg gaactgtgag tccattaaac ctctttcctt tataaattac 22380 ccaatcttgg gcatgtcttt attagcagca tgagaacaga ctaatacata gtgttaaatc 22440 ataaaataaa accatagtat gtactgtact accccagtaa ttttgtagcc acttcctgtt 22500 gctactgcag tgagttcaag tgttgtacct gcttaaaatg cagtgacact aacaatctca 22560 gcaggagcag ttcatctctc cagtaaattg catattgcaa taaaaagtga tctctcatga 22620 ttcttgcgta ttttaaaat catgtttagt gcaataccat aaacattaaa taacaccagg 22680 gaacccatat caagtgacac tagtgttgct ggaagtgctc ccaacaagca gagaaaagtc 22740 atgacattac aggaaaagt tgaactgctt gatacatact atagatgtga ggtctgcagc 22800 tgcagttagt tgcctgccat ttcaagataa atgaatctag cataagaacc attgtaaaag 22860 aaaacaaaac aaaacaaaaa caaaattcgt aaaagccatc actgcagcta caacagcagg 22920 tgcaaaaacc ttgcactttt tgtacaatac atttatctca tattgaaaat gcagctttca 22980 catgggtgca ggattgctgt gagaaaggca tacctataga ctctaatatg attagaggaa 23040 aagcaaagtc aattcaaatt aaaggtaaga ttaaggatca gagctgggtc atttaatgcc 23100 agcaaaggat ggtttgataa ttttagaaag atatttggct ttaaaaaaat gtcaagatag 23160

| caggagaagg | agcttctgcc | aatcaagagg | cagcagaaat | gttcctagat | gctattaaaa | 23220 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| aaaatcattg | agaagaaagg | tatctgcct | gaaaaggttt | taatgcagt | cgaaagtgcc | 23280 |
| ctattttaga | aaaaaaaaat | ccacaaagg | acttttatta | agg | gaagcaagtc | 23340 |
| caagga | tg | gagtaggc | , | gtata | $g c a a t c c g g$ | 23400 |
| tttaggatga | gtccatatct | - | tcttatgtgt | aagct | $c$ | 23460 |
| cttgaaggga | aaagataagc | ctcagctgcc | agtattttgg | ttgtacaaca | aggcatg | 23520 |
| gacaatgaga | gca | t | gatgc |  | a | 23580 |
| gaaagt | agacagt | gact | ag | ttgatatta | acaatgctc | 23640 |
| C | ag | agtttaacat | caaaggtatt | aag | C | 23700 |
| taaa | tctaattcag | cctttagatg | t | agaatcttt | a | 23760 |
| taaatttt | agacat | g | tgttgcctgg | ggtctcg | aactcctgag | 23820 |
| ctcaag | C | g | agtgctagg | tacagg | ggagccactg | 23880 |
| t | C | tttttaaggc | a | cacatggta | tctacagaa | 23940 |
| aggattg | atgctatggt | agagagccac | g | aacatcatgc | agtctgtaa | 24000 |
| tcctaaaacg | aaa | ctggaga | tgtg | atgtacatg | tttcacagg | 24060 |
| atttat | gagc | ggaaa | aga | ggagatat | aaaaaaaaaa | 20 |
| aaagg | g | ttcaaggtat | agattttgga | gaa | agcaaataaa | 80 |
| catcacac | ga | a | $g$ | tgag | caaaccact | 24240 |
| gccaaaca | gcggaaga | ataggaga | gcagtgc | aaaactaac | tttacgcaat | 24300 |
| ctggcaaaag | ag | attcaggact | t | tcttttatga | catggaccct | 24360 |
| tctgtgatac | ggc | tgaag | cagtggagg | aggattggt | accatattga | 24420 |
| aacattttta | gagaaat | agga | agacaga | accacaatg | catgtcagtc | 24480 |
| acaccaggtg | tgcctgc | gc | tccactt | cctctgcctc | ccccoteg | 24540 |
| ccacccotga | gacagcaaga | aatccctc | tcctcc | tttccttag | cctattcaat | 24600 |
| ataaagatga | tgaagata | accttta | atccac | cgcttaat | aatagtaaat | 24660 |
| atattttctc | ccttgtga | ttcttaa | acattttc | ccctctag | tactttatt | 24720 |
| ataaaaatac | aatatgta | atata | a | atgttaat | gactgtttat | 24780 |
| gttattggta | agacttctgg | gacagcag | ctattagta | ttaagtttc | tggggagtca | 24840 |
| gaagttatat | tgaattttc | ctgcacga | gggtcagtc | ccctaaccc | catgttgtc | 24900 |
| caaggctcca | tgtagatgc | tttttactt | cagaaata | aagccataa | acttgtttat | 24960 |
| ttgagggtag | gaagagtaat | gtcaggaggc | attttttct | ttctagaaac | atacattttt | 25020 |
| atttcttcaa | aattatttag | tacacaacag | ttccgaggg | aaactcgatg | ctatttgttc | 25080 |
| ttggattaga | aattctctct | cctgatggtg | attactgagt | tccatttaa | aacticttgga | 25140 |
| ataaacaaag | ctgtgaggat | gcggggtgct | aattaagcct | tttcctaatt | gcttcattgt | 25200 |


| gcgtcaagat | gaagacagtc | ctgaagttat | atactgttt | actcaccac | ttaggtct | 25260 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| gtgactcaaa | ctcccacttt | tattcggcta | tatacacact | aaaaagcaat | gacatttaca | 25320 |
| aaccaatctc | agaccagaca | ctcctgcctt | agaacatggt | acagaaaa | tatttcttaa | 25380 |
| aaccattaca | ctga | cagtaaaatc | tgtttttcag | agacattgt | tatagtctgt | 25440 |
| tgaat | tactaa | g | ttgttagaat | tgataatt | agaaatattt | 25500 |
| ctt | t! | a | tattttat | tttttattt | tattatact | 25560 |
| t | gg | acacaacgtg | agg | ca | atgtgcca | 25620 |
| tgt | C | C | ttaacatta | C | aatgctat | 25680 |
| ccctcccccc | ttccctctac | cccacaacag | t | C | $g$ | 25740 |
| tccatgtg | ctcattg | agttcccacc | atgagtgag | acatacgg | tttggtttt | 25800 |
| ttgtccttgc | gatagtttgc | tgagaa | g | tgtctga | aaggattgtc | 25860 |
| t | t | ttttttctcc | caccttttg | C | t | 25920 |
| ttctaa | actatatatc | atttttataa | $g$ | aagtttatac | aatattttaa | 25980 |
| aataaacctg | ccattttcct | aattttctaa | gtatcttgtg | gtaaacataa | ttcaatcttc | 26040 |
| ttggc | agtg | taatatttta | t | gt | tgtttct | 26100 |
| a | C | tttttacttc | tgatagattt | tgatattagt | ttcaaaact | 26160 |
| gacttaa | cttatg | gcttgctg | tgtttgaag | ttaggtaatt | tatgtaagat | 6220 |
| tcagtgaag | aagtggaa | tg | tgatttta | gctataaaac | actctaaatt | 26280 |
| aaatgtgtct | tt | C | cagtgcagag | gccaaagaga | ggaagaacat | 26340 |
| cttaaacaat | aaagag | ttctcttggt | actt | ctggaag | tttctctta | 26400 |
| aaatcatagc | attaaaaggg | actttagaga | cctctagtc | atcgtcctc | attttgcaaa | 26460 |
| tgaggaaaat | gagacagcat | gttggttcaa | ggtggtgcgg | tgatgtagg | ctgaaatctc | 26520 |
| atcttgtaca | ctggt | C | ta | actcaga | ccagaggtga | 26580 |
| tgaaggatgt | gtttccta | cagattgc | tgttgga | aacatttg | attacaacat | 26640 |
| aattgaatga | tggaaacttt | ctttttaaga | tggagtctca | ctctgttgcc | caggctagag | 26700 |
| tgcagtgacg | ccatcttgac | tcactgcaat | ctctgtctcg | ccagttcgag | cgcttcctct | 26760 |
| gcctcccagt | agcatgggat | aggcatgt | ccaccatgc | cagctaa | ttgtatttt | 26820 |
| tagtagagat | ggggtttcac | catgttgatg | gtggctggtc | caaactcct | tacttcaaat | 26880 |
| gatccacctg | cctccatctc | ccaaactgct | gggattacag | gcatgagcca | ccatacccag | 26940 |
| cccaaaactt | tctggaaaac | agattgatag | tatgtgccac | ttccttaaa | aaattaaaaa | 27000 |
| aattaattca | agccaggtgt | agtgccatgt | cttgtagtc | tagctactt | ggcagactga | 27060 |
| ggcaggagga | ttgcttaccc | aggtgtgtga | ggctgcagtg | agctatgatg | atcacacctg | 27120 |
| tgaattgcca | ctgcactgca | gcctgggcaa | cagagcaaga | ccccgtttct | aaaaaaaaaa | 27180 |
| gttagttttc | tttgacttat | taatttcacc | ttaagaaatt | ttcctaataa | accaattcaa | 27240 |

tatggacaaa tgtttaggta caaagatgtt tatctcacca ctatttttaa taaaaaagga ..... 27300
attgaaaacc tggctcaaca ataaaggaat acttaattgg ttatgatatt aaaggactca ..... 27360
ttacacatct cattattaat gtgtatttaa tgaacttgga aaatgctttt gatatgaagg ..... 27420
taaaaataat gatatagagc taaatataga gtttcattcc aatcttttta aatatattta ..... 27480
tgcacttagg aaaaaaacaa tatggaaatg tgtaaaatat actttttttt taaaaaaaag ..... 27540
gacacattta ttcagcatta tgatcagact attacattta acaatcaaca gtatgggtgc ..... 27600
caaaaaaat ctacattaaa accctttgtt gtaatgcttt acactttcca cagaacagaa ..... 27660
actaaaagaa tctgttacac aattagtcac aaatatagtc ctcgagtttt ttacccatac ..... 27720
acatgagtat ttgtctaaaa catgtcttct tgtagcactt aggccctgcc accactgtgc ..... 27780
ttgtctgagt tcacaaatct gttgtaaact gtagcttccc tgtcacttct ctggctctta ..... 27840
tctcctgcta agatttgttt cctggcagta atttaaaatc ttctgccact gctgtagctactgctgctac tggaactgcc atagccacct tggtttcatg gtttggcaaa gtactggcet27960
gtaccagcat aggggccaga gcttctgcct ccaaagtttc ctcccttcat gggtccaaaa ..... 28020tgtaaaacta attgttgtaa ttgccaaaat cattacacca cctccaaaat tgcttccatgattaccaaat ccattatagc catccccact gccactatat ccaccaccac cacagctgccaccaaagcca caatgaccac tgaagtttcc tccacgacca aagttgtcat tcccaccgaaactacctcca cgaccaccac caaagttccc agagctgctt tgcctctttg gctggatgaagcactcacca tctcttgctt tgacagggct ttcctaactt cacaagtgtg gccattcacagtatgggtat ttctcagtga cagtcttacc catggagtca tggtcgtcaa aagttactaaagcaaagcce cttttcttat cactgccttg gtcagtcatg atttcaatca cttcaatttttccatactat tcaaaataat ctcctaggtg atgttcttca gtgtctttaa tgccaccaacaaatatcttt ttcacagttg ggtgggcatc tggtctttga gaatcttctc ttgagacagctctctttggt tccacaactc ttccatccac cttgtatggc cttgcattca tggctgcatccacctcctcc acagtggcat atgtgacaaa cccaaagtcc cttgagcget tggtatttggatctctcatt accacgcaat ccatgagcat tccccattgc tcaaaatggc tcctcaggctctcattggtc aaccaatgta gagctcaacc tccaatgaag aatttcctca gctgtttgggctcattagga gactctgact tagacatgac ggcaggggaa agagagactt taaggatgcttccttggtgg cgtccacggg cagaaagggg taagcgtcca caggcagaga ggagtaagcctttgaatgta tctaaaattt actttttatt gcctgcattc tttcactatt ttccaaacattcttcaattg tcacaaatgg caatgataaa gagaaaaata taaacatcac attttaaaaataagtgtaaa ataactgtga acttaaaatg tgatcatcat agaagaaaga gcactgcaatagaagtactg tgagttctct ggttaatttt gcgaagccac gtaaagctgt gtggctttaagcaatgccat aatccattta agtcttagct tctatttctg caaaggagaa gtttgaattagtcaatcttc cttccagatc cataactcag tttgataaat tacttagtat ctttttctac280802814028200282602832028380284402850028560286202868028740288002886028920

aaaatacaga tgtgttgtac taagttctgc aagcctttgg caattccagg atatgtttac ..... 31380
tttcccttga taagagagga attggaaggt aagagccaaa tgaattcaga aatgacaaag ..... 31440
gaaaagttat attgggattt ttctgctaca ttgttctgaa tgtagataat tgtacctcgg ..... 31500
tcagaggaag ataagcctga agcaattata ctacaaaagt acccaaatat gcaatttggt ..... 31560
ggtcaaaagt atgtgtaatc tttctgagct tctagatttg gaggtgggta agattctgcc ..... 31620
tcttgatagc ataacataat tagtagcgat ataattttat atttaaacca gaatcatata ..... 31680
agcctggcag tataatgtgt tagtatagta tttctgtcct ttttaaacat tgagttgttc ..... 31740
atgcattaca gtttgctcag gatgaccccc aaaaagacat ctaaatttcc attaaagatg ..... 31800
tacattggac aaatgatatg caggtgctat ttgtgattat ggcctagaga tcaaaaccaa ..... 31860
gtatcactgg cattggggct ttgattagat aattatttga tatattgctt tactccaaaa ..... 31920aattgaattg atgaaagttg ctgacattgg ggcatttagg tttgcaaaat caaccttccaagttaaggaa aaggaagacc tgtctgaaga acagtgcatt agaaaaaggt cccataggtttaaatgatca caagtccaag attaactatc ggtattttat ttagtgctag ttaaacaaacaaacaaacaa aaactaatga cacaatatgc tacacaaaca taggcataga gtccagaggaaagaaagcga tgaattccct gtattctggg cgggttccca agtattatgt ctgttccagg3198032040321003216032220
cttgtgttta aaaaaagcat attaataaat gtgtctagag aagggcagtg aaaggagtta ..... 32280
tcaaagaagc aaaggagttt ttagcaatat ttcaagactt gtaggcctgt cttatgtaaa ..... 32340
aggaagtaaa ttttttcttc agtttgaaaa agaactgttt aacaatagac aatcaaacat ..... 32400gctacttcct aaaacagagg atggaagcca atttagggga ggtgtccagg cacgaacatggagagctgga cttgatacct gtaaggtcct tcccaacttt aagttgctgt gattcccatgtcatagataa gaacgtcaat gcatcttaag agcaacatga tatctggctc tgtaagaaactttcttttgg ttgcacaatt cataggtttt taagaatctg atgtaattcc aacatcactgactgtatccg ttgttgatta ccacaataat gctgtgtaac aaccagccac taaacctcagagacatacat atttttgccc acaaacctat gggttagctg tgaagttcta ctgatctggattaggcatag tggatctcgg ctggacttac tcatgtgtct gtagtggatt gtgtctgtagctggttggtt gggattcaga cagctctcct ccatgtgtgc ctcatcctcc agcaggctatcctgggtttg agacagtggc agaattctga gagagagaga gagagaaaga gagagagaacacatgcatgc atgctcacaa agcatataag gccccctgag tcataggctt ggaactggcaaaggtgattt ccaccatatc ctgttggcca aagcaagtca caaggccagc ccagattcaaagggtgctgc aaaggtcaca gtgcaagaag atgaggatac agagagaggt atacagagggaaaaaaatgg gccattttgc aatcaatcta tcacatagac atgaacttat aaggaaatgtgtttgtttta tttttaacat ctgttttata actattagag gcaaagctct ctggttatagaagtgtcaac ttttggccag gcatggtggc tcacacctat aatcccggca cttttggaggctgaagcagg aggttcactt cagcccggta gttcgagacc agcctgggca atatagggag 3336032460325203258032640327003276032820328803294033000330603312033180332403330033360
acccccatct ctacaaaaaa taaaacaatt agctggtgtg gtcatgcaca gttgtagtcc ..... 33420
cagctactac agaggctgag gtgggaagat cgcttaagcc ctggagatca tatctttagt ..... 33480
gagctctgat tgcactactg cactccagcc tgggcaacac aggaagaccc cacctcaaaa ..... 33540
aaaaaaaaaa aaaaaaaagg gagtgtcagc tttctagcat tgtgatggta atgctgtgca ..... 33600
catgttttgt gtttgtgctt tccagagtat tttcagctgt tgtctttggt gccatggccg ..... 33660
tggggcaagt cagttcattt gctcctgact atgccaaagc caaaatatca gcagcccaca ..... 33720
tcatcatgat cattgaaaaa acccctttga ttgacagcta cagcacggaa ggcctaatgc ..... 33780
cggtgagttt gatgtttcaa ctgtttgatc tactcctgac tcctgaatga aagtattta ..... 33840
agtggaaact taataaaatt tgtactttca aatatgctga tgataaaata aaacttccta ..... 33900
gatcatagat tcctttcaat tactgctaat aatatacatc aacattcagt acttttacgt ..... 33960
agcaaaggtt atagggaaat aggaatactg ctcactttat aagcaaaacc tattaatcag ..... 34020
attttttaaa aacaattttt ttttagagac agagtcttac tctgtcatcc aggctggagt ..... 34080gcagcagtat gatcatagct cactgcagcc ttgatcttct gggctcaagc gatcctcctg 34140cctcagcctc ccaagtaact gggactacaa gcgtgtgcca tcatgcccag ctaatttttt 34200aattatttgt agagacaggg tctcgttatg ttgcccaggc tggttatcag attttattgt 34260atgtaagtta ctgtattcct gaggaacaga tttgagttat tgtagctgta ttgcatattc 34320atattgtctt aacaatacat gctatgaaag cttttactct tttagatctc atttattaaa 34380ttctagcagt ctgaggtcaa gcacagtggc tcatgectgt aatcccagca ctttgggagg 34440ccaaggtggg tggatcacgt gaggtcagga gttcgagacc agcttggcca atatggtgaa 34500actccatctc aaataaaaat aaaaaaaaa agattagctg ggtgtggtgg cacacgcctg 34560taatcccagc tacttgggag gctgaggtac gagaattgct tgaacctggg gggtggaggt 34620tgcagtgagc tgagatcatg ccactgcact ccagcctggg tgacagagtg agactctgtc 34680tcaaaaaaat aaaaaataaa taaataataa aattaaaata aaataaattc tagcagtttg 34740aagtgaagcc aattgtaaca caaattaatt atcttctgac acctggtaat cgagagagtt 34800agctatacac tttattttca gtattgcagc attcaaattt actgttattc ttctcattgc 34860agaacacatt ggaaggaaat gtcacatttg gtgaagttgt attcaactat cccacccgac 34920cggacatccc agtgcttcag ggactgagcc tggaggtgaa gaagggccag acgctggctc 34980tggtgggcag cagtggctgt gggaagagca cagtggtcca gctcctggag cggttctacg 35040accccttggc agggaaagtg gtgagcacac tttcacattt agctcagttc aggttttcat 35100catccaaatg tctgaatgta tttaattctc aactataagc catgtttttt caaaccttta 35160aacaacagtc ccacttggat aaagtctgag agcctaaata tggtctccaa gtggtgtcat 35220ctgtcccagc caacttctcc aggctcccct caacactacc cctctaccct cctttgcaag 35280caccctttgt accacctgtc tcccttgctg tacttaggtt tcagctgact tgaaaagaac 35340caacaaaaat ggaagtagcc agaaagatac aggacaggtg ctaggagtga gatgaaagcc 35400
aaggatggag actgtttcaa agatggtggt cagcaatgac agatggtata gagagattgg
35460
gtaaggagga gactgagaag acatgataga atctagcaat gaatgggcta tttctgacat tttaaaatat tcttagtgaa gtagtgatgg taggaaacag ttaagggtgg ctgaagtatg attaggggaa gaaatggaga tgtaagtgag catagattat caagaagatg gaaagtaaaa gaaaggcaaa taaggacagc acttgagtgg ggaggcagag tccagggaaa atgtttaggg gttgaactct tgagcatttt tatgactggg gatgagacaa aatcattgat ggagagaggg cattgaaaca tcatctgagg aaaaaaagt agaatcaaca aatttgggaa atgtaaaaag aacatggaag tatgcccttt tgtcttagtt gctggggatg agggaagtgt ctggcaggga gtagtgagtt taaacgacta tgatggggga taggaaagaa gacaaactag gaaggaatag aattaagagc aagacatttt cattatgtag gcatgtaatt agttgtactg agaagtgata tttaaaattc ttaacaatgg gaacagcatg gacactgacc aatcagaatg gatgcctgct agtacaagtc agatggatgc ccacccgctg tgtatggacg ctggtagcat ttaatgactc aaatgaaaag agagttttca atctttagtc taggactctt tttacttcct cacagcggtc actcttactc ataaatgttt gttattgtat tcaaagacag ttatgccttt attcagttta ctcttaaatc attacaagct tgttccctac ttttactctc ttgctttata ttttctgtga actctgattg tgtccttcaa ttttgtggtg aacttaaaaa caggactcta aataaagacg tcgctgacag ccaaagcaat tagaggagga aatgttagct ctcagacact gattatggcc ctcaacagac agttattaag tggtgtgtta aagattgtgc tataatgaat tgtagggcat gatatctgcc ctcaaatact ttaactgagg agatggggca tatagttttg tgagttaaag gagataggcc tagagctagt tcttaagaga taaggatttg catatgggtg aacagtaggg atgacattca atgtaagaga aaatacaagc tgcgatcagt gatttgttcc caatgcagca aataggacac tttggctgaa atgggagttt agtttgagta gtggttctca aacttctctg atgagaactc atgtaggaaa gatttaaaga tgtcaccatc cactaatgta ctcatagtca ctttccagtg ggaaaaacac ttggtgataa caaaatgttg gagttgaaaa caacaataca atacagcaca ataatgatac aataaggaat aattaatgat tattctgcaa gtatataaac acaagcagaa cataataaaa catggacagc atcagttaca gttcagagta ggaaacagaa actgccctag gtatttcaag tagataagta tttaataaag ggaattacat gcttacaaaa tctttggaag ggctggaaga gtgaaaggag gctatttgct cccagacttc caaatcacag cactgcagct gtgattctgt aaccaagaag ctgcagaaaa ttatgccgat atcacagctg ttccaaattt aaagacacga gactagaatc tggctgttgc agaaattctt gtctgccaaa gatgagttaa gccgttagct taccacagct gctgaaggag agctggtgtc tacctcccag attttacact gtgcatctcc ctcgtagacc ctgacttact tccagaacca agggaaactg ggaaatagtt tttagccttc tagtcccttg atgtataaa ggtcagacac agaagtatat gaaaatgaat gctgagtgca taggacagta agcatcccaa gacattattg ataatgcact

35520
35580
35640
35700
35760
35820
35880
35940
36000
36060
36120
36180
36240
36300
36360
36420
36480
36540
36600
36660
36720
36780
36840
36900
36960
37020
37080
37140
37200
37260
37320
37380
37440

tacatgtgtg gaatagtctc catggcaact cagcttggaa ctaataaggt tatgggagag 39540
cttccatccc tgcacctgcc agtgtcacaa gcagaagcca tgttcctgta gcaaagattg 39600
tactctactg ctaggcagct gtccccttga gccacccagc caggcacatg ggatacagag 39660
agtatatggc tcagcacgca ctagtcacta ctattagaaa agctcaaagc tgagtcactg 39720
gtgccttctt cagaagggat gaacagctct ctcacttgaa tgccagaaaa ttatcttgca 39780
aagcagacct atctgataga catatttgca tcagagtagg gcttgttatc agcaaggctg 39840
taaggtgccc tccccagtct tctgcaggat aatccaggga gccagattat agagaaaggg 39900
cagccctctg ttcctactca tctggctcag tattgaggat ctccattcac cctttcctac 39960
ccctgttcac ctatccatcc cctgtagttc ctgacctgca aagctattat gtggtcatag 40020
ttgttattta tttccatgct aatcctgggc actgtttctc tgaaaaatgg agatttaaga 40080
ataaggctgt caggatacat ctcagaagta ttaggcgttg tattagtgtg gctgctataa 40140
ccacttgcca caaattcagt ggcttaaaac aacaccaatt tatgacttta cacttctgga 40200
ggtcagaagt ctaaaatagg tctcagtggg ctcatatcaa gtgtcaggag ggctgtgttc 40260
cttctagagg ctccaaagga tgatctgttt tcttgcttgt ggccccttcc tccattttca 40320
aaaccagcag tggctggttg agtctttctc acactgcatt tgctgatact cttctccctc 40380
cttcttcttc agttaagagc ccttatgatt acattagttt caccaagata attcaggata 40440
atcttatttt acagtcagct gattagcaac cttacatcta ctactttagt ttctttttgc 40500
catgtaacat aacacattca caggatccag ggattaggac acagatgtct tgtgggagag 40560
ggaacattat tctgcctacc acatgcatac atcagaaacc atggttgaaa cacaggaaac 40620
atgacagttc ctcaaggcat acaattatga ccttgttggg ttaaccttca ctatccaaat 40680
tttaatcaca caaacttttc cttaatctca cagtaacttg gcagtttcag tgtaagaaat 40740
aatgatgtta attgtgctac attcaaagtg tgctggtcct gaagttgatc tgtgaactct 40800
tgttttcagc tgcttgatgg caaagaaata aagcgactga atgttcagtg gctccgagca 40860
cacctgggca tcgtgtccea ggagcccatc ctgtttgact gcagcattgc tgagaacatt 40920
gcctatggag acaacagccg ggtggtgtca caggaagaga tn 40962

```
<210> 10
<211> 20
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (7)..(7)
<223> n may be any nucleotide
<400> }1
```

gaagggnctg aacctgaagg 20
<210> 11
<211> 18

```
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (7)..(7)
<223> n may be any nucleotide
<400> }1
gaaggtnctg ggaagatc
<210> 12
<211> 507
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (507)..(507)
<223> n may be any nucleotide
<400> 12
gtattttcag ctgttgtctt tggtgccatg gccgtggggc aagtcagttc atttgctcct 60
gactatgcca aagccaaaat atcagcagcc cacatcatca tgatcattga aaaaacccct 120
ttgattgaca gctacagcac ggaaggccta atgccgaaca cattggaagg aaatgtcaca }18
tttggtgaag ttgtattcaa ctatcccacc cgaccggaca tcccagtgct tcagggactg 240
agcctggagg tgaagaaggg ccagacgctg gctctggtgg gcagcagtgg ctgtgggaag 300
agcacagtgg tccagctcct ggagcggttc tacgacccct tggcagggaa agtgctgctt 360
gatggcaaag aaataaagcg actgaatgtt cagtggctcc gagcacacct gggcatcgtg 420
tcccaggagc ccatcctgtt tgactgcagc attgctgaga acattgccta tggagacaac 480
agccgggtgg tgtcacagga agagatn 507
<210> 13
<211> 1442
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (1442)..(1442)
<223> n may be any nucleotide
<400> 13
nctgaacctg aaggtgcaga gtgggcagac ggtggccctg gttggaaaca gtggctgtgg
tgaaaatgtc accatggatg agattgagaa agctgtcaag gaagccaatg cctatgactt ..... 300
tatcatgaaa ctgcctcata aatttgacac cctggttgga gagagagggg cccagttgag ..... 360
tggtgggcag aagcagagga tcgccattgc acgtgccctg gttcgcaacc ccaagatcct ..... 420
cctgctggat gaggccacgt cagccttgga cacagaaagc gaagcagtgg ttcaggtggc ..... 480
tctggataag gccagaaaag gtcggaccac cattgtgata gctcatcgtt tgtctacagt ..... 540
tcgtaatgct gacgtcatcg ctggtttcga tgatggagtc attgtggaga aaggaaatca ..... 600
tgatgaactc atgaaagaga aaggcattta cttcaaactt gtcacaatgc agacagcagg ..... 660
aaatgaagtt gaattagaaa atgcagctga tgaatccaaa agtgaaattg atgccttgga ..... 720
aatgtcttca aatgattcaa gatccagtct aataagaaaa agatcaactc gtaggagtgt ..... 780
ccgtggatca caagcccaag acagaaagct tagtaccaaa gaggctctgg atgaaagtat ..... 840
acctccagtt tccttttgga ggattatgaa gctaaattta actgaatggc cttattttgt ..... 900
tgttggtgta ttttgtgcca ttataaatgg aggcctgcaa ccagcatttg caataatatt ..... 960
ttcaaagatt ataggggttt ttacaagaat tgatgatcct gaaacaaaac gacagaatag ..... 1020
taacttgttt tcactattgt ttctagccct tggaattatt tcttttatta catttttcct ..... 1080
tcagggtttc acatttggca aagctggaga gatcctcacc aagcggctcc gatacatggt ..... 1140
tttccgatcc atgctcagac aggatgtgag ttggtttgat gaccctaaaa acaccactgg ..... 1200
agcattgact accaggctcg ccaatgatgc tgctcaagtt aaaggggcta taggttccag ..... 1260
gcttgctgta attacccaga atatagcaaa tcttgggaca ggaataatta tatccttcat ..... 1320
ctatggttgg caactaacac tgttactctt agcaattgta cccatcattg caatagcagg ..... 1380
agttgttgaa atgaaaatgt tgtctggaca agcactgaaa gataagaaag aactagaagg ..... 1440
tn ..... 1442
```

<210> 14
<211> 759
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (759)..(759)
<223> n may be any nucleotide
<400> }1
nctgggaaga tcgctactga agcaatagaa aacttccgaa ccgttgtttc tttgactcag
6 0
gagcagaagt ttgaacatat gtatgctcag agtttgcagg taccatacag aaactctttg120
aggaaagcac acatctttgg aattacattt tccttcaccc aggcaatgat gtatttttcc 180
tatgctggat gtttccggtt tggagcctac ttggtggcac ataaactcat gagctttgag 240
Page 102

```
```

gatgttctgt tagtattttc agctgttgtc tttggtgcca tggccgtggg gcaagtcagt 300
tcatttgctc ctgactatgc caaagccaaa atatcagcag cccacatcat catgatcatt 360
gaaaaaaccc ctttgattga cagctacagc acggaaggcc taatgccgaa cacattggaa 420
ggaaatgtca catttggtga agttgtattc aactatccca cccgaccgga catcccagtg 480
cttcagggac tgagcctgga ggtgaagaag ggccagacgc tggctctggt gggcagcagt 540
ggctgtggga agagcacagt ggtccagctc ctggagcggt tctacgaccc cttggcaggg 600
aaagtgctgc ttgatggcaa agaaataaag cgactgaatg ttcagtggct ccgagcacac }66
ctgggcatcg tgtcccagga gcccatcctg tttgactgca gcattgctga gaacattgcc }72
tatggagaca acagccgggt ggtgtcacag gaagagatn 759
<210> }1
<211> 2200
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)..(1)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (1442)..(1442)
<223> n may be any nucleotide
<220>
<221> misc_feature
<222> (2200)..(2200)
<223> n may be any nucleotide
<400> 15
nctgaacctg aaggtgcaga gtgggcagac ggtggccctg gttggaaaca gtggctgtgg
gaagagcaca acagtccagc tgatgcagag gctctatgac cccacagagg ggatggtcag 120
tgttgatgga caggatatta ggaccataaa tgtaaggttt ctacgggaaa tcattggtgt 180
ggtgagtcag gaacctgtat tgtttgccac cacgatagct gaaaacattc gctatggccg 240
tgaaaatgtc accatggatg agattgagaa agctgtcaag gaagccaatg cctatgactt 300
tatcatgaaa ctgcctcata aatttgacac cctggttgga gagagagggg cccagttgag 360
tggtgggcag aagcagagga tcgccattgc acgtgccctg gttcgcaacc ccaagatcct 420
cctgctggat gaggccacgt cagccttgga cacagaaagc gaagcagtgg ttcaggtggc 480
tctggataag gccagaaaag gtcggaccac cattgtgata gctcatcgtt tgtctacagt 540
tcgtaatgct gacgtcatcg ctggtttcga tgatggagtc attgtggaga aaggaaatca 600
tgatgaactc atgaaagaga aaggcattta cttcaaactt gtcacaatgc agacagcagg 660
aaatgaagtt gaattagaaa atgcagctga tgaatccaaa agtgaaattg atgccttgga }72
aatgtcttca aatgattcaa gatccagtct aataagaaaa agatcaactc gtaggagtgt 780

```

Thr val \(\underset{35}{ } \quad \underset{40}{ } \quad\) Val phe ser Met phe Arg Tyr Ser Asn Trp Leu Asp Lys Leu Tyr Met val val Gly \(\underset{50}{\operatorname{Thr}} \underset{55}{\text { Thr }}\) Leu Ala Ala ITe ITe His Gly Ala Gly Leu Pro Leu Met Met Leu Val Phe G7y Giu Met Thr Asp Ile Phe Ala Asn Ala G7y Asn Leu Glu Asp Leu Met ser Asn Ile Thr Asn Arg Ser 859095
Asp Ile Asn Asp Thr Gly Phe Phe Met Asn Leu Glu glu Asp Met Thr \(\begin{array}{r}105 \\ 100\end{array}\)
Arg Tyr Ala Tyr Tyr Tyr Ser G7y Ile Gly Ala Gly Val Leu Val Ala
Ala Tyr Ile G7n Val Ser Phe Trp Cys Leu Ala ATa G7y Arg G7n ITe
135
130


Asp Asp val Ser Lys I7e Asn G7u G7y Ile G7y Asp Lys I7e G7y Met
Phe phe G7n Ser Met Ala Thr Phe Phe Thr Gly Phe Ile val Gly Phe Thr Arg G7y Trp Lys Leu Thr Leu val Ile Leu Ala Ile Ser pro val 21021520
\(\underset{225}{L_{2}}\) Gly Leu Ser Ala Ala val Trp Ala Lys \(\begin{array}{r}\text { I7e } \\ 230\end{array}\) Leu Ser Ser Phe Thr Asp Lys Glu Leu \(\underset{245}{\text { Leu Ala Tyr Ala Lys Ala Gly Ala val Ala glu glu }} \underset{250}{255}\) Val Leu Ala Ala Ile Arg Thr val \(\underset{260}{ } \frac{17}{265}\) Ala Phe Gly Gly Gln Lys Lys Glu Leu G7u Arg Tyr asn lys Asn Leu Glu glu Ala lys arg Ile gly 275280285
Ile Lys Lys Ala Ile Thr Ala Asn Ile Ser Ile gly Ala Ala Phe Leu 290 295 300 Page 105

 G7u Ala Phe Ala Asn Ala Arg gly Ala Ala Tyr G7u \(\begin{gathered}\text { Ile } \\ 365\end{gathered}\)
 Pro Asp Asn Ile Lys Gly Asn Leu G7u Phe Arg Asn Va7 His Phe Ser
385
395 Tyr Pro Ser Arg \(\underset{405}{\operatorname{Lys}}\) G7u Val Lys ITe Leu Lys G7y Leu Asn \(\underset{410}{415}\) Leus val Gin Ser G7y G7n Thr val Ala \(\underset{420}{ }{ }_{425}\) val G7y Asn Ser G7y cys G7y Lys Ser Thr
435 Thr val G7n Leu Met Gin Arg Leu Tyr Asp Pro Thr G7u G7y Met Val ser Val Asp G7y Gln Asp ITe Arg Thr ITe Asn Val Árg Phe Leu Arg G7u Ile Ile G7y val Val Ser Gln G7u Pro val Leu Phe
470
 Met Asp G7u \(\underset{500}{\text { ITe G7u Lys Ala val Lys G7u Ala Asn Ala }} \underset{505}{\text { Tyr }} \begin{array}{r}510\end{array}\) Ile Met Lys Leu Pro His Lys Phe Asp Thr Leu val Gly G7u Arg G7y
515
520 Ala G7n Leu Ser Gly Gly \(\underset{530}{\operatorname{Gln}}\) Lys G7n Arg Ile Ala \(\underset{540}{ }\) Ile Ala Arg Ala Leu Val Arg Asn Pro Lys Ile Leu Leu Leu Asp G7u Ala Thr Ser Ala
545
550 Leu Asp Thr G7u Ser Glu Ala val Val \(\begin{gathered}\text { G7n Val Ala Leu Asp } \\ 565 \\ 570 \\ \text { Page } 106\end{gathered}\)

ITe Ser Phe ITe Tyr Gly Trp G7n Leu Thr Leu Leu Leu Leu Ala Ile
850
850
Val \(\begin{aligned} & \text { Ser G7n G7u Pro Ile Leu } \\ & 11150\end{aligned}\) Phe Asp Cys Ser \(\underset{1125}{ }\) Ala G7u Asn
Ile Ala \(\underset{1130}{ }\) Tyr Gly Asp Asn \(\operatorname{Ser}_{1135}\) Arg Val Val ser \({ }_{1140}\) G7u Glu Ile
Val \(\begin{array}{r}\text { Arg Ala ATa Lys Glu Ala } \\ 1145 \\ 1150\end{array}\) Asn Ile His Ala phe \(\underset{1155}{ }\) ITe G7u Ser
Leu Pro Asn Lys Tyr Ser Thr
1160
1165
Leu \(\underset{1175}{\operatorname{Ser}}\) Gly G7y G7n Lys Gln \(\underset{1180}{ }\) Arg Ile Ala Ile Ala \(\underset{1185}{ }\) Arg Ala Leu

Leu Asp Thr G7u Ser G7u Lys val val gin glu Ala \(\underset{1215}{1205}\) Leu Asp Lys

Thr \(\underset{1235}{\text { Ile }}\) G7n Asn Ala Asp Leu \(\underset{1240}{\text { Lle Val Val Phe G7n }} \underset{1245}{ }\) Asn G7y Arg

Ile Tyr \(\underset{1265}{ }\) Phe Ser Met val \(\underset{1270}{ } \quad\) Val Gin Ala gly \(\underset{1275}{\text { Thr }}\) Lys Arg G7n

\(<210>17\)

<211> 1103

<212> DNA

<213> Homo sapiens
<220>
<221> misc_feature
\(<222>\) (810). (810)
<223> \(n\) may be any nucleotide
<400> 17
ctgcagtgac cactgcccca tcattgctgg ctgaggtggt tggggtccat ctggctatct 60
gggcagctgt tctcttctct cctttctctc ctgtttccag acatgcagta tttccagaga 120
gaaggggcca ctctttggca aagaacctgt ctaacttgct atctatggca ggacctttga 180
agggttcaca ggaagcagca caaattgata ctattccacc aagccatcag ctccatctca 240
tccatgccct gtctctcctt taggggtccc cttgccaaca gaatcacaga ggaccagcct 300
gaaagtgcag agacagcagc tgaggcacag ccaagagctc tggctgtatt aatgacctaa ..... 360420480540600
tcataatgac aggaaaatag aaaatataaa cctcatttta attctttcac agaaaggtta
ccaccagatc tcataaaaac tcattgtcac gaggacaaca cctaaggcgg gatggtgtga ..... 3120 ..... 318032403300
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline tgatgctctg & cccagggtct & gggccctgga & tggaaaatta & ggagcccatg & tccacatggc & 5160 \\
\hline cagccgcagc & agtcagctct & ttggcttaca & tcatctttcc & cacat & aggcttgttt & 5220 \\
\hline ctcagattct & aattctctca & ggtgagggct & ttgttgtcta & attactatcc & aggaatttca & 5280 \\
\hline tattttttcc & ctgtgcaaaa & gcaataattt & cccogccacc & ttttccaggt & caactcttta & 5340 \\
\hline gtagatgtac & ccccaagatg & cacattcctg & ggacctttgt & ttgcacagtt & aaaatgtcac & 5400 \\
\hline ccctgaaatg & tcgatacagg & aaggtttgtt & tttaagtttc & agtgaaaact & ctgagcaagt & 5460 \\
\hline gttgtaattt & gctgtgtccg & atgtgtagag & gggacatttt & ctcagaactt & ccatgttaag & 5520 \\
\hline ctggaaaact & ggaaagtgag & ttcactttgt & cattctgtca & ctcgttcatt & ttctcactca & 5580 \\
\hline acaacatgcc & tcatacttac & ctaaatctgc & tagactaaag & gagttccctg & gtgtctgtaa & 5640 \\
\hline ctttccaatt & ctgctagaac & tctagagcga & gctcatgaaa & taaatgaaaa & ggatgacaaa & 5700 \\
\hline gagataaaac & actgtgcatt & ctcttctgat & gctaattcac & tttcccttgg & cctcagtttc & 5760 \\
\hline cccatgtgcc & cctggag & atcattcagg & gattcatgag & attttcaaga & caacacatga & 5820 \\
\hline aamagcaaaa & agacatcaga & aagacaaaga & ggtacttagt & atttatacac & aaggataagt & \[
5880
\] \\
\hline cattcagtat & ccacaacact & tggagagaat & tcaagagtga & ttttaaattt & cccttttcaa & 5940 \\
\hline atacctcctc & tgttttctct & tatttccttt & tgacgtctc & caaataagct & tcctctaact & 6000 \\
\hline gccagcaagt & ctgat & tggcttcgac & tgttttcatc & ccaattagag & gcagggttaa & 6060 \\
\hline gtacatta & aataataatc & aaatattatt & ttgtttctcc & tcccagggc & tttgtatgtt & 6120 \\
\hline tgacatggaa & tgtcataaaa & agtatggaa & agtgtggggg & tgagtattct & ggaaacttcc & 6180 \\
\hline attggataga & cttgtttcta & tgatgagttt & accccactgc & acagaggaca & gtctcagccc & 6240 \\
\hline aaagcctctt & gggatgaagc & ac & taactacaa & acagagagaa & gttctctgaa & 6300 \\
\hline agaagaagat & atttattt9g & gtgtagagta & ttgcaatggg & aatctgcatg & cctttataaa & 6360 \\
\hline ctatgtgcaa & attcagggaa & gtaaagcaag & acaaagaggc & tccaaggaaa & atatgaggag & 6420 \\
\hline gatttcttat & cagttttgaa & ataattatcc & ttcgctacaa & agatcagtaa & caagggtgac & 6480 \\
\hline gcctcaccaa & ggttggacag & gcagttgctg & ggcaggtgtc & cttgcagaaa & tatttttttt & 6540 \\
\hline aatgttggga & tggcctttgt & gcaagcttgt & agttttgcgg & agtcttttgt & gatagttttg & 6600 \\
\hline ttatcaggca & cacaagcatg & agaatcctct & cttcatagcc & ttctttgatt & tatttgtcag & 6660 \\
\hline ggtttttaca & cacacacaca & cacacacaca & cacacacaca & actagtgaca & tcattttggt & 6720 \\
\hline tctaacaaca & ttcacactgg & ttattgtaaa & acttttcgaa & ggttgtccta & ccaaggatcc & 6780 \\
\hline catgtgtcac & caggtgtcaa & gttctacagt & ctgaactagg & ctgggagcat & tgtgattact & 6840 \\
\hline tttctccaga & ctttggtggc & ccagggactc & acagcatcat & gctctgtcca & gtgtctgcct & 6900 \\
\hline attcccctct & tctttttttt & ttccttagg & tgccctttta & ttacatgtgt & tgtctcagac & 6960 \\
\hline ccttctaata & tgtgctcata & aatacatcat & atcatctcct & tcccacatca & attcactttc & 7020 \\
\hline aattaaaagc & caaaactctt & tcatttagac & tttggattta & aagtgctttt & gaatgaaggg & 7080 \\
\hline ttgagagata & atagagaaat & agattggcaa & accatttata Page 11 & ctctgctgtt & gttgtttttt & 7140 \\
\hline
\end{tabular}
aattttatct gcaagtgtgg aacttttcat tctgttttgt tattaaattt aagccaagac ..... 7200
tttttaatag aagggtatat aagcatttct ttgtctatac cttcctgctg aatttgaaga ..... 7260
aatgctgaat attcttaacc actggcgggc tgatggactg tgattttatt ttatttttta ..... 7320
tttttagttt tttaaattat actttaagtt ctgggttaca tgcatagaat gtgtagtttt ..... 7380
gttacatagg tatacacgtg ccatggtggt ttgctgcacc catcaacctg tcacctacat ..... 7440
taggtatttc tcctaatgct atctctcccc tagcccccca cccaacaaca ggccccagtg ..... 7500
tgtgatgttt ccctccccgt gtccatgtgt tctcattgtt caactcctac ttaggagtga ..... 7560
gaacatgtgg tgttgagttt tctgatcttg tgatagtttg ctgagaatga tggtttccag ..... 7620
cttcatcctt gtccctgcaa aggacatgaa ctcattgttt ttttatggct gcatagtatt ..... 7680
ccatggggta tacgtggcac attttcttta tccagtctgt cactgatgga catttgggtt ..... 7740
ggttccaagt ctttggtatt gtgaatagtt ctgcaataaa catatgtgtg catgtgtctt ..... 7800
tatcatagaa tgatttatgc tttgggtata tgcccagtaa tgggattgct gggtcaaatg ..... 7860
gtatttctag ttctagatcc ttgaggaatc accacactgt cttccacaat ggttgaacta ..... 7920
atttacactc ccaccaacag tgtaaaagtg ttcctatttc tccacatcct ctccagcatc ..... 7980
tgttgtttcc tgacttttta atgatcacca tactacctgg catgagatgg tatctcattg ..... 8040
tggttttgat ttgcatttct ctaatgacca gtgatgatga gcattttttc acatgtctgt ..... 8100
tggctgcata gatgtcttct tttgagaagt gtctgttcat atcctttgcc tattttttga ..... 8160
tggggttgtt tgcttttttt cttgtaaatt tgtttaagtt ctttgtagat tctggatgtt ..... 8220
agcccttcgt cagatggata gattgcaaaa attttctccc attctgtagg ttgcctgttt ..... 8280
gctctgatga tagtttcttt tgctgtgtag aagctcttta gtttaatcat atcccatttg ..... 8340
tcaattttgg cttttgttgc cattgctttt ggtgttatat ttatgaagcc tttgcccatg ..... 8400
cctgtgtcct gaatggtatt gcccaggttt tcttctagga tttttatggt cctaggtctt ..... 8460
acatttaagt ctttaatcca tcttgagtta atttttgtat aaggtgtaag gaaggggtcc ..... 8520
agtttcaatt ttctgcatat ggctaggcag tttcaccaac accatttatt aaataggaaa ..... 8580
tcttttcccc attgcttttg tgtgtcaggt ttgtcaaaca tcagatggta gtagatgcat ..... 8640
ggtgttattt ctgaggcctc tgttctgttc cattgatcta tatttctgtt ttggtacctg ..... 8700
taccatgctg ttttggttac tgtagccttt tagtataatt tgaagtcagg tagcgtgatg ..... 8760
cctccagttt tgttcttttt gcttaggatt gtcttgtcta tgtgggctct tttttggttc ..... 8820
catatgaact ttaaagtagt tttttccaat tctatgagga aagtcagtgg tagcttgatg ..... 8880
gaaatagcat tgaatctata aattaccttg ggcagtatgg ccattttcat gatatggagt ..... 8940
cttcctaccc atgagcatgg aatgttcttc catttgtttg tgtcctcttt tatttcattg ..... 9000
agcagcggtt tgtagttctc cttgaagagg agaacttcac atcccttgta agctggattc ..... 9060
ctagatattt tattctcttt gtagtaattg tgaatgggag ttcactcatg atttggctct ..... 9120
ctgtttgtct attattggtg tgtaggaatg cttgtgattt ttgtaaattg attttgtatc ..... 9180

attatccttc atcatatgaa gacttgggtg gctcctgtgt cagactcttg ctgtgtgtca ..... 11340
11280
aaccatggtc tctttgaaga gctcttttgt ctttcaatgt ctcttccttg tttggcccac
tacctaatga actagaacct aagattactg tgtattgtac aactaaggga ttatgtaagg ..... 11400
tcaggatcaa agtctggctt cctgggttgg gctccagctg tagaataagg ctgttgatgt ..... 11460
ttaatcaact ctgttttttt cacacagctt ttatgatggt caacagcctg tgctggctat ..... 11520
cacagatcct gacatgatca aaacagtgct agtgaaagaa tgttattctg tcttcacaaa ..... 11580
ccggagggta agcattcatg tgttgaaatt aaaatactga ttgattaaat ttatattttg ..... 11640
aaattcttat atattcatag acagttgcct aaaaaatgtc caggaaggtt ccacgtccac ..... 11700
ttcatcctgt cccccccgaa tggtaacatc ttgcaatctt gcataactat aaatacagta ..... 11760
tattcatgtt actatatagt acaggaaatt gacattgata cagttcacag agcttattca ..... 11820
gatttcacca gctttatgtg ccctcatttc tgttctatgc aaatttatca caatcataga ..... 11880
tttgtgcaat gaccagcatg atcaaaacgc agaacccatt tgtgttctat acaaatttat ..... 11940
cacatcatta gtagatttgt gcaatgacca gcatgatcaa gatgcagatc cactctgtct ..... 12000
ccatgtggct ccctcctgct atcctacagt cacaagtctt tttctctgac agggtccata ..... 12060
12120
tcagagcaaa ctatttttta ttttgaggtg atcaatgtat taatatttcc ttttctggat12180
12240
tttttctaaa agtcataag tttaatgctt tactaaatgt agctctatta tcaattttgt
12300
gttatttttt gtataaggta tgagatttag atcaaggttc attttttttg tggcttatgg
12360
ctgtccaatt cctccaacgc catttttgga aagatgggta tattgttaaa aatcaactgg
12420
gcatacttct ttgcatctct cacgtactac tgtgtcccat tgatcactgt gtttattatt ..... 12420
ccaccaatac cacactgtgt tgaccctagt agctgtacac taactcttaa cctcgtgtag ..... 12480
agtgattctt cccactttta ttgacttatt ttttcagatt tgttttaact cttcctggtc ..... 12540
ctctgctttt ccataaaaat tcagaatgag tttctcagtg tctacaaata aacgtgctgt ..... 12600
tattttgaca agaattgaat taaatatata gaccagttta tggaaaatgt gtatctttac ..... 12660
tgtttgatct ttcaattcat gaacatagta tgcctctcca tttccttaga ttttctttga ..... 12720
ttgcttttaa catcttgtag ttttcagcat agagatcctg tacatgtttt gttagattta ..... 12780
cagctaaatt atttcatttt tgttggaatg actgtaaatg atattatgtt tttcttgtat ..... 12840
tttcagatat tgattgttgt tacatagaaa tgtgctttga ttttgtgtgt tgatctggta ..... 12900
tcctagaacc ttgcagaact gacatagtag ttctagaagt ctctttgtat atgccttgag ..... 12960
attttacaca ttgtcagtta tgtcacttga aaatagagac aattctttta tttcctttcc ..... 13020
aatctgtatg ccttttattt cttttctttt gtctattgca ctaggacttc ccagtataat ..... 13080
gttgagtaaa tgtggtgaat ttttgaattg ttcctaatct taggagggaa gccttcagtt ..... 13140
ctttcatcat taagcatcat ttagatgaag tgggctgttt ttttgtagat tctctttatc ..... 13200
aaattgagga attttctctc cctagtttgc tgagttttta tcataaatga atgctggaag ..... 13260
\begin{tabular}{llll} 
tactcattat aaaaaaatg gctatggaag atgaaagaaa gtttcaagca tgttggcttg & 13320 \\
ataggccagt tccaagttgg caaaataat tatctctttt ttctttctat ccatgaaata & 13380 \\
aaaaattaag agacaagaat gtttatggaa ttgcattatt tcttcaaaat atgttcctag & 13440 \\
ttttaaaggt attacctact atttttttta aaaccatcac attgaggcac ttctttttca & 13500 \\
cgttgcccat gctgcaggag aacataaaga cagcttgtct gaggcaacat acaatccacc & 13560 \\
aaagtcacct gcttgtctgt ctccactccg tctctacact gcagaagtgc taggtcttga & 13620 \\
ttctgtttat tgtactggaa gaacacattc tctaccacgt ggataatttg catgtaaaag & 13680 \\
aagactggga tatagaggct ggagacgaca tcagggtctc ctgaacactg ccgccacccc & 13740 \\
cccgccccgc cccacacaaa tacatcccag gacactgggc atctgggata aatctctatt & 13800 \\
gagcatctag catagggccc atcacccagt agacagtcac taaatagttg ttgaataagt & 13860 \\
gttcctgttt aacacatttt ctacaaccat ggagacctcc acaactgatg taggacaaaa & 13920 \\
tgtttctgct ttgaactcta gccttttggt ccagtgggat ttatgaaag tgccatctct & 13980 \\
atagctgagg atgaagaatg gaagagatta cgatcattgc tgtctccaac cttcaccagt & 14040 \\
ggaaaactca aggaggtatg aaaataacat gagttttaat aagaaactta aagaatgaat & 14100 \\
ctggtgggga caggtataaa ataagatcac agtccctttc caaggggtag tccactgaat & 14160 \\
ttgagctgcc taaaaatggt cttttatctt tatgtacaga aaacacatca caaaattcat & 14220 \\
tataaaatgt cacttactgc tccatgctgg ggaaagccat gtccttctgg gactagagtc & 14280 \\
tgcacattta actatgggtg gtgttgtgtt ttgtgcttag atggtcccta tcattgccca & 14340 \\
gtatggagat gtgttggtga gaaatctgag gcgggaagca gagacaggca agcctgtcac & 14400 \\
cttgaaagag taagtagaag cgcagccatg gggttctgag ctgtcatgaa cccctccagc & 14460 \\
tgcctgccat ggagctgata ttcctgctgt tgggttattc cagtgaccag acaaaaggag & 14520 \\
ggctgtggta atgcaacttc aatgggtctc ccaagatggg gcagctccga tgaggaggtg & 14580 \\
gggcagctgg aggaaaagga tcttctcccc tgtgcacagg ggccagggtt tacatatcca & 14640 \\
ttaaattgtc accttggata ttctagaaga ctaaatatat cctttagggg gaaaaagtgt & 14700 \\
gattgtacca aagttttaag catggagtgt atgggatggt ggaagggaa ggcacttggt & 14760 \\
atctgttggt tggcagtgag taggttggaa gagttataat ggagaactta gaataacttt & 14820 \\
gatcatttca tgtttttttc tgaggatatc agtagaatac taaatattaa aattcctacc & 14880 \\
atttcttttt cctccagtct caaagagaga gggtggtaaa aacactatag gtagggcaag & 14940 \\
cctattattt gctatctaca cttatgcagt aaaaacaggt gtaatctgag tttgtcctgg & 15000 \\
gcagaccagg gatatgtggt cactcactat agaaatttcc aaatcaaatt ttgagagatt & 15060 \\
tttttttaac caggacatta ttggtcatta tattttacaa aaataattct gctgtcaggg & 15120 \\
caacctcagc tcaccacagc tggggatagt ggaattttcc aaagcttgag cagggagtat & 15180 \\
agagaataag gatgatattt ctaggagctc agaacagggt actgttgctt tgtaaagtgc & 15240 \\
tgaagaggaa tcggctctgg gcatagagtc tgcagtcagg caatatcacc tgtcttgagc & 15300 \\
\hline
\end{tabular}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline cccttaggaa & gagttaatta & ttctactctt & gttctgctga & agcacagtgc & ttacccatct & 15360 \\
\hline tgtatcatcc & acaatcaata & catgctact & tagttgtct & atagtgggtc & tctgtcttcc & 15420 \\
\hline tatgatgggc & tccttgatct & cagaggtag & tctaattcag & ttcagtgtc & ccatcacacc & 15480 \\
\hline cagcgtaggg & ccagctgcat & cac & tga & tgat & gtgtgataga & 15540 \\
\hline aggtgatcta & gtagatctga & aagt & ct & gtcttgactg & gacatgtggg & 15600 \\
\hline tttcctgttg & catgcataga & ggaaggatgg & taaaaaggtg & ctgattttaa & ttttccacat & 15660 \\
\hline ctttctccac & tcagcgtctt & tggggcctac & agcatggatg & tgatcactag & cacatcattt & 15720 \\
\hline ggagtgaaca & tcgactctct & caa & caagacccct & ttgtggaaaa & caccaagaag & 15780 \\
\hline cttttaagat & tgatttttt & g & a & taagtatgtg & gactactatt & 15840 \\
\hline tccttttatt & tatct & tcttaaaaat & aactgcttta & tgagatata & aatcaccatg & 15900 \\
\hline taattcatcc & acttaaaata & tacagttcag & tgatttgtag & tacatttgaa & gatatgtgtg & 15960 \\
\hline accatcatca & ttttaaactt & taa & tt & tagagacctc & atacattttt & 16020 \\
\hline agctatcagc & cccctgtcac & aaa & a & accactaatc & aactttctgc & 16080 \\
\hline ctc & t & tggac & tagaaatga & tattaattca & tcagggtttt & 16140 \\
\hline ttattctcta & gtt & t & tctgtatca & tttctttct & tctgctggct & 16200 \\
\hline tcaggcttag & tttgccctt & ttcgtttact & tgttgtggc & tgaacatag & attactgatt & 16260 \\
\hline tgtgattttt & ttgttcctct & aaatttagac & ttacagctg & taactttccc & tctgagcact & 16320 \\
\hline tcctttgcta & aat & attgtggce & atcacatct & tagtttttgt & tcacctcaaa & 16380 \\
\hline acagt & tttg & ggtttctac & ttgactcat & tggttactta & aatgtttatt & 16440 \\
\hline atttaacttc & cacatatgtg & tgagtttctc & t & cccttattga & tttatcttt & 16500 \\
\hline attccatgat & aggtgacaga & gatatgctgt & ttatttcta & tcttgactac & ctactatttc & 16560 \\
\hline ttgaacagca & agattaattt & tgagcttcag & ttatgattt & gggttattct & aggagactgt & 16620 \\
\hline agtccaatag & ataaaggcaa & agagattagg & gcattgaatt & ttgttccttt & tatccttcaa & 16680 \\
\hline aagatgcaca & aggggctgct & gatctcactg & ctgtagcggt & gctccttatg & catagacctg & 16740 \\
\hline cccttgctca & gccactggcc & tgaaagaggg & gcaaaagtca & tagaaggaat & ggcttccagt & 16800 \\
\hline tgagaacctt & gatgtctttt & actcttctgg & tggtagaga & aactagaat & gctccaggt & 16860 \\
\hline aaattttgca & cattcacaat & gaatttcttt & tctgttttt & ttttgtttt & cctacagca & 16920 \\
\hline gtctttccat & tcctcatccc & aattcttgaa & gtattaaata & tctgtgtgtt & tccaagagaa & 16980 \\
\hline gttacaaatt t & ttttaagaaa & atctgtaaaa & aggatgaaag & aaagtcgcct & cgaagataca & 17040 \\
\hline caaaaggtaa a & aatgtggtgg & tagttatagg & aggatgttta & gtttttcata & attttttaga & 17100 \\
\hline taatatacat a & atgatcagtg & cagttacctg & tatgttttta & aagaatgctt & ttaacatgaa & 17160 \\
\hline gactgctcat g & gtttagagca & agagaattca & tttggtagaa & atgcagaaag & tggggtttgg & 17220 \\
\hline ggaaggagat a & atgagaatga & gtcagagaca & gcacatgaaa & tttgatatca & gacacaacaa & 17280 \\
\hline ttagtatgcc a & acggcataaa & ttttattgag & ataaaacttt Page 118 & accactttac 8 & ttcccttcaa & 17340 \\
\hline
\end{tabular}
taaattgtca gaggataaac attactgttt ggaaatatat tttactgata ttatgctttc cccagatcat aaattggtaa agactcattt caagtacaaa cctgttattt tacacattct

17400 caaatgaaag tccctatcag gccacctgct gaagtgtagc gtgtgttaaa tttgagccat ctcacatgat agccagattt gtttcaggaa aacatcctgc tttccaagga tttagaaggg tatgtttttc actggtgatt caggcaacat gcatcagatt tctggtcttc aaggagccat attctcagaa gggagatcaa ggaccacgct tgtgatttac ttctgacttc aggagccact ttctgtcagt gaaatttctc tttttgcttc tagcaccgag tggatttcct tcagctgatg attgactctc agaattcaaa agaaactgag tcccacaaag gtaaccagag tgtttctgag ggctacttgt ggggcactca gagggaaggc cttgttctga aaatgtgcag gaagtattcc aggatgatga gaatttctgc cacatagcag aacgacacat gtttgaatgt tataagtggt agttggaggc actttctaga ggcatgcagg catagatagc catgttctaa gagtaaaggg caaccctaag caaacctggc atgctagaaa gtcagtctgc ggtctgtgga tcacctacat cagatcaaat gccaattctc agcctccttc agatccactg aataaaaatt tctgcctaga aatttattag gttgctgcaa aattaattgt gggtttttcc attactttta attgcaaaaa aatgcaatta aaagtaatgg caaaaaccac aataactttt gcagctacct aatacatcta acatccaatc aggagccacg ctgttccaaa aggggtgatt ttaactggca gtacttgttg aaagtgtgtt caccaggtta atctactgca aagttatttt ccactttgga atggattagt aacttgtggg gaaaacctct gagaccgtgt aaatatcctc tttgtcttca atgtttcacc tacgagtttt acaacccatg tatgtttttt actgaagtca ctattactat gacagtggca aaatgatgac catggtcaat tacaagccac caagacttgg caaccatctc acaaaattcc tgaatattta actattggtt ctagagagca ggactgggct tactccagca tactgcttta aatatatcca tgtctacatc cacttttgtc tgtatgtcta tgtatctatc tatgtatcta tctagctatg tatctatcta tctatctatc atctatctat ctatcatcta tccatctatc atctatcatt tatccatcta tcatctatct acttatatat tatctatcta tcatctaact attcatctat ctatcccaat ataacttgct gtgataaagg aaatagtcta ttgttttact gtttcatata gaaatcacta gacacatatg gctattgagc actggatatg tggctagtgc cattgaagat caattttaaa tggtatttca gtttaattta ataaaatttg attttaaata gccactagtg ggtagtggct accatattgg acagcagagc tctaaacttt gagattatag ttcaatttca catcagtatc atcaggttca tgataactga atactatgat taggtagttg aatttactta taactgcatc acagaagtct tcactggtaa atcacagctc tgtcgacctt tctcacactc ctttcatatt ggttttggtt gtgaattaca tggttggagc aggcattata tttatttcta tgttccaggt ctctaaaggt cctaatccag tcctgatcaa acagaccagt gatggaccat cctgagcttc tctcaggaga aaaatcaaga ggggcccaac ttgtaatcat aggagcttat gctattttaa tgccatccat cagactacaa tcaattacca ctcatctagc 19380 Page 119

\begin{tabular}{|c|c|c|c|c|c|c|}
\hline aatcgtgttc & tctggttaca & tccaaccaga & accttctttg & tcatttttgg & gatagaaagg & 21480 \\
\hline gactagttta & tcctcaaatt & atttatggag & attttat & atatagtgtt & tctctccaca & 21540 \\
\hline ttttctgtat & ataacaaaag & tcct & agtgtgtgta & acatata & tatacacata & 21600 \\
\hline tatatgt & tg & tgtgtgtata & tatatataca & catatataca & ggttatgct & 21660 \\
\hline aagga & gt & tg & C & accttc & taactcctc & 21720 \\
\hline cacacat & agt & t & aatgggcatg & acagttatta & aacacttta & 21780 \\
\hline taaatgc & gat & cagt & tagtctctgg & gctcctaat & \(g\) & 21840 \\
\hline tac & a & aagttaattc & aaaatctcaa & a & t & 21900 \\
\hline t & ca & ctatgatact & gtgctacaga & t & gacatggtg & 21960 \\
\hline gtgaatg & cgctcaga & attcccaatt & C & ttgagagggt & tgcaamaaa & 22020 \\
\hline gatgttga & tcaatgg & g & aggggtgg & ggtgatgat & ccaagctat & 22080 \\
\hline g & g & g & g & C & C & 22140 \\
\hline aag & g & C & ccagcctggt & agcatat & ctgcctctc & 22200 \\
\hline ttaatctac & ggacag & g & aattatttgc & g & tatttagag & 22260 \\
\hline atttttttaa & tcatcada & attattgt & cact & aaccatagac & tagaaaaaag & 22320 \\
\hline adaa & tc & t & ttacgatgaa & t & a & 22380 \\
\hline tg & g & t & せta & acatgat & aatgctatt & 22440 \\
\hline attatgtc & ccaaatga & atttttttca & taaatataat & caaatattta & gaataacat & 22500 \\
\hline tttaataaaa & aacatgc & taatcttcaa & gaa & gcttactgg & aacagataga & 22560 \\
\hline caaagaaagc & agtgatgata & ctgcattaca & tcggtacagt & at & cctgtgtaa & 22620 \\
\hline attatctgac & ttcaacta & tatggagg & tggggga & agagggaga & atggagaga & 22680 \\
\hline agaagaagga & ggagaaggag & ggagaaga & aggtaagg & agaaggagg & agaattagaa & 22740 \\
\hline aaacaagaga & ggagatgaga & agga & aaataa & tttg & gtacaagaca & 22800 \\
\hline atttcccctt & ccttcctc & tgaccaatg & aagtgtg & tgaggcagg & aacctacttt & 22860 \\
\hline tccatcagtc & agtcccatca & cttatgtgcc & tttatagt & ggacacat & accaccctga & 22920 \\
\hline atataatttc & agtgtttaga & aataagtatt & tttgcaac & ctatttatct & catctcaaca & 22980 \\
\hline agactgaaag & ctcctatagt & gtcaggagag & tagaaaggat & tgtagctta & caattctcat & 23040 \\
\hline agcaaaataa & gcatagcagg & atttcaatga & cagc & aggtat & gtgtactact & 23100 \\
\hline agttgagggg & tggcccctaa & gtaagaaacc & ctaacatgta & actcttaggg & gtattatgtc & 23160 \\
\hline attaactttt & taaaaatcta & ccaacgtgga & accagattca & gcaagaagaa & caaggacaac & 23220 \\
\hline atagatcctt & acatatacac & accctttgga & agtggaccca & gaaactgcat & tggcatgagg & 23280 \\
\hline tttgctctca & tgaacatgaa & acttgctcta & atcagagtcc & ttcagaactt & ctccttcaaa & 23340 \\
\hline ccttgtaaag & aaacacaggt & tagtcaattt & tctataaaaa & taatgttgta & ttaataattc & 23400 \\
\hline ttttaactga & tggtctgta & tttttaaaa & agaatatgct Page 12 & tgtttaatct & tttactaatt & 23460 \\
\hline
\end{tabular}


```

<210> 19
<211> 2059
<212> DNA
<213> Homo sapiens
<220>
<221> CDS

```



ITe Asp \(\underset{195}{\text { Ser }}\) Leu Asn Asn Pro G7n Asp Pro Phe Val \(\underset{200}{\text { G7u }} \underset{205}{ }\) Asn Thr Lys

Lys Leu Leu Arg Phe Asp Phe Leu Asp Pro Phe Phe Leu Ser Ile Thr
210
215

Va7
225 Phe Pro Phe Leu Ile Pro Ile Leu G7u Va7 \(\begin{aligned} & 230 \\ & 235\end{aligned}\) Leu Asn Ile Cys Va7



Gln Leu Met ITe Asp Ser Gln Asn Ser Lys G7u Thr \(\underset{280}{275} \underset{285}{\text { Glu }}\) Ser His Lys

Ala \(\underset{290}{\text { Leu }}\) Ser Asp Leu Glu \(\underset{295}{\text { Leu Val Ala Gln Ser }} \underset{300}{\text { Ile }}\) Ile Phe Ile Phe

Ala Gly Tyr G7u Thr Thr Ser Ser Val Leu Ser Phe Ile Met Tyr G7u
305
310


ATa Val Leu Pro Asn Lys Ala Pro Pro Thr Tyr Asp Thr Val Leu Gln \(\begin{aligned} & 345 \\ & 340\end{aligned}\)

 G7y
385 Met Phe Ile Pro Lys G7y Val Val Val \(\begin{array}{r}\text { Met } \\ 390\end{array} \quad\) Ile pro ser Tyr Ala Leu His Arg Asp Pro Lys Tyr Trp Thr \(\underset{405}{ } \underset{410}{ } \quad\) Pro Glu Lys Phe Leu \(\underset{415}{ }\) Pro

G7u Arg Phe Ser Lys Lys Asn Lys Asp Asn Ile Asp Pro Tyr ITe Tyr

Thr Pro Phe Gly Ser Gly Pro Arg Asn Cys Ile Gly Met Arg Phe Ala

Leu Met Asn Met Lys Leu Ala Leu Ile Arg Val Leu Gin Asn Phe Ser
450

ITe Asp Ser Leu Asn Asn Pro G7n Asp Pro Phe Val \begin{tabular}{r} 
G7u Asn \\
200
\end{tabular}


Ala \(\underset{290}{\text { Leu }}\) Ser Asp Leu G7u \(\underset{295}{\text { Leu Val Ala Gln Ser }} \underset{300}{ }\) Ile Tle Phe Ile Phe
Ala Gly Tyr G7u Thr Thr Ser Ser val Leu \begin{tabular}{l} 
Ser \\
305 \\
310
\end{tabular} Phe Ile Met Tyr G7u
320


Ala val Leu pro Asn Lys Ala Pro Pro Thr Tyr Asp Thr Val Leu Gin \(\begin{gathered}345 \\ \\ 340\end{gathered}\)


ITe Ala Met Arg Leu Glu \(\underset{370}{\operatorname{Arg}} \underset{375}{\text { Val }}\) Cys Lys Lys \(\underset{380}{\text { Asp }}\) Val G7u ITe Asn

G7y Met Phe Ile Pro Lys g7y val val val Met Ile pro Ser Tyr Ala
385
390 \(\quad \begin{array}{r}400\end{array}\)

Leu His Arg Asp Pro Lys Tyr Trp Thr G7u Pro G7u Lys Phe Leu Pro

G7u Arg Phe Ser Lys Lys Asn Lys Asp Asn ITe Asp Pro Tyr ITe Tyr

Thr Pro Phe Gly Ser Gly Pro Arg Asn Cys Ile Gly Met
435 \(\quad \begin{aligned} & 445 \\ & 435\end{aligned}\)

Leu Met Asn Met Lys Leu Ala Leu Ile Arg Val \(\underset{450}{450} \underset{460}{\text { Leu }}\) G7n Asn Phe Ser
\begin{tabular}{|c|c|}
\hline Phe Lys Pro Cys Lys G7u Thr G7n Ile Pro Leu Lys Leu Ser Leu G7y
465 & \\
\hline G7y Leu Leu G7n Pro Glu Lys pro val val Leu Lys val Glu \begin{tabular}{r}
490 \\
495
\end{tabular} & \\
\hline \[
\begin{gathered}
\text { Asp G7y Thr val Ser G7y Ala } \\
500
\end{gathered}
\] & \\
\hline \[
\begin{array}{ll}
<210> & 22 \\
<211> & 35769 \\
<212> & \text { DNA } \\
<213> & \text { Homo sapiens }
\end{array}
\] & \\
\hline ```
<220>
<221> misc_feature
<222> (7087)..(7087)
<223> n may be any nucleotide
``` & \\
\hline <400> 22 & \\
\hline ctgctgtgca gggcagggaa gctccaggca aacagcceag caaacagcag cactcagcta & 60 \\
\hline aaaggaagac tcacagaaca cagttgaaga aggaaagtgg cgatggacct catcccaaat & 120 \\
\hline ttggcggtgg aaacctggct tctcctggct gtcagcctgg tgctcctcta tctgtgagta & 180 \\
\hline actgtccaaa ctcctctctt tgtttccttg gacttggggt gctaatcggg ccccttttcc & 240 \\
\hline cttatctgtt ttgaagatca aaagagatgt tcaaggagaa gtagctgaag tgttggacgc & 300 \\
\hline tacaaacgca tagaagttat tattatctta tgcagatcta tgaatgaata aataagcatt & 360 \\
\hline tctcccatcc accttctaat tttggtgact aggagggttt agggacagca tttggtagtg & 420 \\
\hline ggaatgattt gattagctta gatctgacga agactaatca atgaaaacat ggcagcggca & 480 \\
\hline gattacaaac tgctgatcat gatggacagt gtgatcctca tccccttccc aggctctggg & 540 \\
\hline gattctgggt acaggaagga gtggcttgca tttttgtctc attaattcgc tttctgggtt & 600 \\
\hline ctgtgtctgc tggaagggat gtgtagctgt attgcccotg tagacctggt tcctgctccc & 660 \\
\hline ccgccttcca acccaggata tcatttacat aacgcaccag gggacaccaa gacttcatgg & 720 \\
\hline gaagctgtcc cctggctctt ccctctttcc tgtgccatgc ccctgaaaat cccctccotc & 780 \\
\hline ctatgagtca ctcctccacc ctgtcataca caggatggtt tatcttgcaa tgattaacct & 840 \\
\hline ctagagcaaa ggagacctgg aggaagtttc gaggatttat tctttgcttt aatctttttc & 900 \\
\hline ctcccgtctc tgggaggcta ggattaatat agagctttgt ttctcaccta atgggaatct & 960 \\
\hline actagcagcc tgaaaaggca ggagccatga aagccaattt ggattttaca tatttttcce & 1020 \\
\hline ctttatgtta cagtacagga gggcaaaccc tctcactggt gggattcctg gcatcctaga & 1080 \\
\hline gcaggtggag agaagagtta ctttccactg tgggtagtgg aggctccacc tgtcccatta & 1140 \\
\hline acttctacct caatttgact tttattaaga gcagggaacc acaatgacat gaaaatagac & 1200 \\
\hline actataaacc tcattttaat tctttcacag aaagcttagg aattcagtga gttgtggcaa & 1260 \\
\hline catggtttcc attgtctaac atttttaaat gaattgatat ggtttaaatt cattcatttt Page 130 & 1320 \\
\hline
\end{tabular}
taaaccagaa ttttttggag atagactatt tccagcatgt tccttctgga tggtaaaaca ..... 1380
gggctgttag ttcagtattt gtgacaataa gtgtgtgtaa aataatgtca cctttcctga ..... 1440
atgtcaggaa tatgagtcta atgcacaaat gtatacctct aagacaagac tgcacgtctt ..... 1500
ttcaaatata cctgtccggc catttatttt aataactcct tttcgaatat acctgcttag ..... 1560
cagattgtct taaactctca ggacagggga gtaagcaaga ctgtgagcca gtgacgatag ..... 1620
caaaggcttc caggtaggat ccatatgaag tgagaaaata ttcctcagct ctcagggtag ..... 1680
aactccaaag agatattcat gggtcctggc cccaccgtgg aggtcactca aagggcaaac ..... 1740
aggttggcat ctcatctgct tcaagcctgg acacaggggc accatctgtg tcactctgtg ..... 1800
tgtggtctgc catgttgtgg gccggtcact acagactcgg gcagccaggc agacaatgcc ..... 1860
ttagccttag acaatgctgg tgcagcccag gagtcagaaa atgcagtgta gaccaggccc ..... 1920
tccttaggcc aacacaatta catgcaatag atgactggct tttctgttag tctcttcact ..... 1980
ggacccaaag gctgcattac tctaccagag gggagctgga aagaaactaa agagttcgcc ..... 2040
cagcacagca tctgccttga catggtacca tgtgaatcta gacactcacc aagatctttc ..... 2100
cttgggggcc aatgctgctg acacattaac tcaatagctt gtcctcacct gagaggtcag ..... 2160
gtaatgtgtt taaagttcag gagcagagat tagtgtcatt gatttgacat ggctgtgaca ..... 2220
acaaaggagg gaactgaagt gggaataccc aaggccaccc tggctttggc aggtggtgca ..... 2280
cgcacttcca ctaactgttc tggggcaggg aaccaaatgt atgactgggc ctgctcatgc ..... 2340
tgcccctgct gagtcctcca aaccctgccc ttcatgtaat ttctcagttt tattttatca ..... 2400
cattttataa gtcactggat gtttacaaaa tgtttggaac ctatactgcc ttgaaggcta ..... 2460
acctctaaag aggagtaaac aaggtcttaa tacaactctc cgggacgttt tatcattact ..... 2520
tatcttatat gccatactgc accatttgct atcaacagga aagtacctgg actttggaag ..... 2580
gtccctctgt gtcttttagc tgaaagtaca tatgaggcat gtggattctt ttatgcacat ..... 2640
catctttttc agccacattt ttgtagtttg cctctctgga gccaactgtg tggggctagc ..... 2700
agcttcacag ctgaatcagt gtctggcaac ctcttccttc agcctctctt cttcctccag ..... 2760
ttttccatcc ctcagtcaca ccggaggggg aaggtctgca aggatccaga accatcagtt ..... 2820
ggaggagttt gcacatgact catgaaagat gagttccagg caggcctgcc atagtgaaca ..... 2880
ccaggcttaa tgggtttttc ctcagagata cttcacgtac agaggcagtg aactgactgc ..... 2940
tttctggttg accaccttga aaaagatgag tgtgcctggc actgtgcttc tcaggtgagt ..... 3000
atgacctgag aagtattagt tgctggttct tctgcacaca atcattcaag gacatatgga ..... 3060
tcaaccatcc tcctcaacag ctcaaatcaa ccagatcatc tgaccacaga gactgaggtg ..... 3120
tacctgaaag ctgcccacat ttctataagg ccaatagaag ccatgaacac agttgtcaat ..... 3180
ctgtagaaat aaggactcca tgactcctcc aaggcctctc tgtgaatgaa cgtttaagaa ..... 3240
gggctagatc ctaaaacagg gtcagagctt agagggaaga aaaagcataa acatttctga ..... 3300
gcaaattgta agggcagtgt caccataggc tcccagtgac cctctgtgat tgagtgcata3360
cagtgatgca aaatctcatc atcagtgcaa aagacaaaaa aaatcttact ctttctacct ..... 3420
aggatgagag tccccaaatc agcgaagagt ccacttacta aacagacata aggaaatgaa ..... 3480
gtgtcctgga agaattcctg cctgaacctc tcaggagcat ttgaggacat ttatcaagta ..... 3540
ttcactccag gattgggact atgaagactt cagctgcttt cagctaatca ttgagacttt ..... 3600
tcaggggtct cagaatagtc aggaaaggac ctgatgagtg aatgcaatta ctgatgttgg ..... 3660
agttgctgtt attatttatc gtgtacatat tacctccctc tcttgaccat tccagttcct ..... 3720
gagtaactca ccagccctct gatctataaa gtcacaatcc ctgtgacctg atttctgttt ..... 3780
cactttgtag atatgggacc cgtacacatg gactttttaa gagactggga attccagggc ..... 3840
ccacacctct gcctttgttg ggaaatgttt tgtcctatcg tcaggtgagt tgcttgagct ..... 3900
tcctcttttg cttcttatgg ttgcaaacat cagcttagtt ccatcagtaa aaatgcccct ..... 3960
ccttgggagg gagttctgag gtttcacatt ttcagaaatg gtgggactgg gtgcagtgga ..... 4020
tcatgcctgt aatctcagcc tctgtgaggc caagactggc aaattgcttg agcccaggag ..... 4080
tttgagaaca gcctgggcaa cacagtgaga cacctgtctc tagaaagaaa aaattacctg
tgcatgatat ggtagcccat gcctgtagtc ccagctactc tgaatgttaa ggtgggagga
ttgtatgaac ccaggaagtc aaggctgtat tgagctgtga tcgcaccact gcactccagc
ttggtcaaca gaacaagaca gaaaggaaga aagaaagaga gagagagaaa gaaagagaga
ggaaggagag gggaggggag gggagaggag tggggaggag aggagaggag aggagagaaa
aggagaggag agaggagagg agaggaaaag gtgtgtaggc tccacccaaa gcatggccag
gtttacccct ggagggaaag tcacaagctc atgtccagaa ggccagtagc agcaagctgc
tctccagccc agatttccta tcctgtgtac ctggagcttg tttctcagat tctaactctc41404200
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline ttttgtttcc & tctcccaggg & tctctggaaa & tttgacacag & agtgctataa & aaagtatgga & 5460 \\
\hline aaaatgtggg & ggtgagtatt & ctgaaaacct & ccattggata & gacctgctac & tgtgaggagg & 5520 \\
\hline ttaccccact & gcaggatagt & ctctgcccag & gtcttcatgg & gatgaagctc & ttgtcaacct & 5580 \\
\hline aaatacaaac & agagagaggt & tctctgaaag & aagaggataa & ttacttggga & gtagaatatt & 5640 \\
\hline gcaatgggaa & tctgc & gttataaact & atgtgcaaa & cagggaggt & aaacaagaca & 5700 \\
\hline aagatgctcc & atagaaaata & tgagaagaat & ctcat & tttgagata & attattgtta & 5760 \\
\hline gctacaaaga & tcaataacaa & gggtgatgcc & acaccaagg & tggacaggca & gttgctggac & 5820 \\
\hline aggtgt & gcagaaatat & ttttgtgtaa & agttgaaa & gcctttgtgc & aaagttgtgg & 5880 \\
\hline tttttgtaga & C & atagttttgt & t & acaagcataa & gaatcctctc & 5940 \\
\hline ttcatagcct & tcttgg & tatttgtcag & ggttaaaaaa & caattagtga & catcactttg & 6000 \\
\hline gttctgataa & agttcacact & cgctattgta & aaacttttcg & aggcttgtcc & taccaaggat & 6060 \\
\hline cccatgtgtc & accaggt & gaggtcttca & gtctgaacta & ggctaggagc & attgtggtta & 6120 \\
\hline ccacttttct & gcaggttttg & gtggcccagg & gactcccagc & atcgccttct & gtccagtgtc & 6180 \\
\hline tgcctattcc & ctcttct & ttttcttcct & taggtgccet & ttatcacat & gcattgtctc & 6240 \\
\hline agacccttc & aatatgtgc & cataaatgca & tggcatcatc & tccttcccac & attgattcac & 6300 \\
\hline tttcaa & aagccaaaac & tccttcattt & agactgaat & taacatgtgc & ttttgaaaga & 6360 \\
\hline agggttgaga & gataatagag & aaacagattg & ggaaaccact & tatgctccac & ttttttaaac & 6420 \\
\hline tttctctgca & agtatggaat & tttttgttct & gctttgttgt & taaatttaa & gccaaaactt & 6480 \\
\hline cttaatagaa & ggatatacaa & atatttattg & gtttatacca & ttgcacttac & tttgaagaag & 6540 \\
\hline agatgctgaa & tattattaaa & ccattgtgtt & ccctggtggg & ctgatggact & gtgattttat & 6600 \\
\hline aaggtggtct & cagccaattg & cagcagctgt & tccotgtcag & aggggctaga & \(g g t t t g g t g a\) & 6660 \\
\hline gagcagtgga & tgaggtgcag & tggtgtgttt & gttcactaga & agcaagtggg & agaaagcttt & 6720 \\
\hline gcctctttgt & acttcttcat & cttctcccct & caagtcctca & gaatccacag & cgctgactgt & 6780 \\
\hline ggagtgctgt & ggagctggca & tggcccatac & aggcaacatg & acttagtaga & cagatgacac & 6840 \\
\hline gctctagatg & tccatgggcc & ccacaccaac & tgccottgca & gcatttagtc & cttgtgagca & 6900 \\
\hline cttgatgatt & tacctgcctt & caatttttca & ctgacctaa & ttctttttg & ataatgaagt & 6960 \\
\hline attttaaaca & tataaaacat & tatggagagt & ggcataggag & atacccacgt & atgtaccacc & 7020 \\
\hline cagcttaacg & aatgctctac & tgtcatttct & aaccataatc & tctttaaaga & gctcttttgt & 7080 \\
\hline ctttcantat & ctcttccctg & tttggaccac & attacccttc & atcatatgaa & gccttgggtg & 7140 \\
\hline gctcctgtgt & gagactcttg & ctgtgtgtca & caccctaatg & aactagaacc & taaggttgct & 7200 \\
\hline gtgtgtcgta & caactagggg & tatggattac & ataacataat & gatcaaagtc & tggcttcctg & 7260 \\
\hline ggtgtggctc & cagctgcaga & atcgggctag & tgaagtttaa & tcagctccgt & tgtccccaca & 7320 \\
\hline cagaacgtat & gaaggtcaac & tccetgtgct & ggccatcaca & gatcccgacg & tgatcagaac & 7380 \\
\hline agtgctagtg & aaagaatgtt & attctgtctt & \begin{tabular}{l}
cacaaatcga \\
Page 13
\end{tabular} & \(\qquad\) & tccatttttt & 7440 \\
\hline
\end{tabular}
gaaatttaaa taatgattga tccactgatt aaatttttat tttgaaaaaa acatatattc ..... 7500
acagaaggtt acctaaaaaa tgtacaggaa ggttccatgt actcttcatc ctgtcccgcc ..... 7560
cagtggtaac atcttgcaat cttgtatatt gcaatatata tctagtatat tcatattatc ..... 7620
aggttggcac aaaagttaaa atggcaaact acaggctggg cataatggct catgcctgta ..... 7680
atcccagcac tttgggaggc cgaggcaggt ggatcacgag gtcaggagtt cgagatcagc ..... 7740
ctgaccaaca tggtgaaacc ccatctctac taaaaataca aaaattagct gcgtgtggtg ..... 7800
gcatgcgcct gtagtcccag ctactcagta gtctgagaca ggagaatcgc ttgaacctgg ..... 7860
gaggcggagg ttgcagtgag ccgagatcac gccattatac tccagtctgg gcaacccaat ..... 7920
gagactccat ctcaaacaac aacaacaaca acaacaacaa aaaccggcaa actgcaataa ..... 7980
cttttgcacc aacctaatac tatagtacag gaaattgact ttgatatagt ttacagagct ..... 8040
tttcagattt caccagtttt acatgccctt gtttgtgtgt gtttatgtgt gtgggtagtt ..... 8100
ctaagcaatt tttcacattc gtagatttgt gcaacgacca gcaccatcaa gatgcagacc ..... 8160
cattccgtca ccatgtggct ccctcctgct gtcctacagt cacaacatgg agtttgtctt ..... 8220
tttctctgac aggttctata tcagagcaaa cttttattta tttgaggagg ccaatgtatt ..... 8280
aatatttcct tttatggatt gttcttttgg tgttaagtct gaaaatcctt tgcttagccc ..... 8340
tccttcctac attgcttttt ctaagagtta tatagtttaa cactttacaa aatgtaactc ..... 8400
tattacccat tttgtgttaa tatttgcata agttatgaga tttagatcaa ggttcatttt ..... 8460
ctgtggacta tggctgtcca aatgttccaa caccattttg gaaaggtagg catattgtca ..... 8520
aaactcagct gagtatattt tgtgaatcta tttcttattg tttactcctc cactaatacc ..... 8580
acactgtggt gactctagta gctgtacagt aactcttaac atcatatagg gcaattcttt ..... 8640
ccactttatt gatttatatt ttcagaatgg ctttagcttt tcttgtccct tgcctttcca ..... 8700
taaaaattca gaataagctt gtaagtgtct acaaacaaac ctgccataat tttgataaga ..... 8760
attaaagcag aggtgtccaa tcttttggct tccctgggcc acagtggaag aagaagtgtc ..... 8820
gtgggccaca cataaaatac acacacacac acacacacac acacacacac acacacacac ..... 8880
acacaaatgg tctgtgtata gttttcatta tatatctacc accacagata agcaaaaatg ..... 8940
tccttgcata ataatcctaa ttatgcactg ccccattcag agggtctttc aaaatcattg ..... 9000
aacaggttcc aagtttgcaa tcactgatac agaaaatgta catatctagc taaacttcac ..... 9060
tacttttttg atatttttta ttataaaga aaagagaaca acataaaact agtggggtac ..... 9120
ttgacattgt ttttgagaaa ctaatccatc agtatctggc ttgatggaag tagttgcaat ..... 9180
tctcagtgag ttctcaaggt gctcatcaga tattttggtt ctaattttac tcttcgtgtt ..... 9240
cttcatcctt gaaaatagta gctcacaaat gtaagtgctg ccaaaaagca atgacatgaa ..... 9300
caaggtgtga ttgtgaagca agggatattt gtcattggga agacaggtct tacaaaagtc ..... 9360
cagtaaagag gcaaaatcaa atttttctat aagttgaaca tcagattgca gctctaggca ..... 9420
ttccatttca aaattgccag gtaacatata tatgtcgact gaaaatggag ttgcaaatat ..... 9480
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline caaaatat & tgatgatttt & & & & & 9540 \\
\hline aaattgaaaa & gcaaggctgc & gtatttttgg & ctgttcacag & gaccatgttt & agccaacatg & 9600 \\
\hline tcgaaatgca & taaaattgtt & tgccttaat & gagcttgcc & taatttcag & tttcatatgg & 9660 \\
\hline aatgctgtta & tggtttgaaa & cattgtatt & ttaagt & ttcaacttg & aagacacagg & 9720 \\
\hline tttaactc & ttaaatgggc & cgtcaaaccc & actaaaaatg & taaatctgt & aagccagttt & 9780 \\
\hline tcattgtcaa & gttctggcac & aattttg & gataccata & aacagcttga & tttcacatca & 9840 \\
\hline caaagcataa & aatcttta & g & cttaaccat & ttacttcta & aaaagtgaat & 9900 \\
\hline gacttgctag & agtcagc & atac & ggaat & aaactagcg & atgattcaat & 9960 \\
\hline tcctgggcce & ttgtgaaatt & tacagccttg & atgacaattt & catgacgt & atctactttt & 10020 \\
\hline aaagcttgtg & cacatggat & ttcttgatgt & ttatgcaat & aatacttcat & caaatgtgag & 10080 \\
\hline ttttgtgtgg & caactgcatc & CLa & gtacaagtc & cctctctttt & acctaccatc & 10140 \\
\hline gccagggcag & catctgtagc & tatc & tgt & ggacaaaga & aaattgcttt & 10200 \\
\hline aacatatttt & t & a & cttgatttag & -9t9 & aatagcatg & 10260 \\
\hline gtgacatttc & gattt & ga & attcatcatc & atac & ataaaaatag & 10320 \\
\hline caagttgtgc & cgtatctgt & gtgtcagtg & tcatccat & caccaaagca & taaaatttta & 10380 \\
\hline aattagcagt & t & aaa & aatagattt & ccaatttct & ccaattctcc & 10440 \\
\hline tggctatag & ctggtgagac & aaactga & aga & agtttctcaa & g9caaataat & 10500 \\
\hline atctaccaca & tcttccagac & tgct & aactcac & tcagtaaatg & gttttgattt & 10560 \\
\hline ttttgctatt & aaatttg & cacataa & ggttttacc & tacgattga & gtccgagttg & 10620 \\
\hline taacttttta & aaaaatc & tttgttgaaa & agacagactt & ttttcagtt & ctgctatttt & 10680 \\
\hline gtccttac & cacataca & ccaaatttg & cagcacgt & tgcatataa & tgcctcttca & 10740 \\
\hline aattgtagtc & tttgaaaa & cacaaa & cgtggaaat & taagcagagt & gcttcgctat & 10800 \\
\hline ttgcctcaac & aagaaaaagt & tttgt & tttcattg & acaatcttc & cttcatccat & 10860 \\
\hline aatttttgtt & ttttagggtt & cttttt & acattgtgg & aagccattct & ggaattaaaa & 10920 \\
\hline gcattataat & agataagc & atattt & tttattatg & gaaattaaca & gataggaaaa & 10980 \\
\hline tagaacagaa & agcaaggttt & taatca & aagaatact & tacatgtctt & taaataata & 11040 \\
\hline ttaaacacct & catcta & ggtagg & aatatta & gataattgc & tgggttttac & 11100 \\
\hline ttgccaaatt & ccacaaaca & cctaatac & tgacagtg & aattcaact & gtccgtgatt & 11160 \\
\hline agaagataac & acactggaag & cgcacacca & ccataaaact & gaagccacac & atgcgtacaa & 11220 \\
\hline atggcgacag & tgtctggtgt & cagcagcgc & tctgccttgt & cagaataca & cacttgaatt & 11280 \\
\hline ctttgtcaca & attcacttca & cgtggcactg & caatagcgtc & ctctcgctct & ttgttagtta & 11340 \\
\hline attttaatgg & cttttaattt & cttcttgctg & aactgtttgc & aattataatg & caaattatgg & 11400 \\
\hline atactagtcc & attatttgtg & gatgtgacat & actctgatta & cccctttcca & ttccattgtt & 11460 \\
\hline gtctacgaag & ttcacacttg & agaatcacat & agtcaaatta & caaaattaca & aaaaaaattg & 11520 \\
\hline
\end{tabular}
caaaaaact caaaatgttt taagaaagtt tccacatttg tattgggata cattcaaagc

11580
11640
11700
11760
11820
11880
11940
12000
12060
12120
12180
12240
12300
12360
12420
12480
12540
12600
12660
12720
12780
12840
12900
12960
13020
13080
13140
13200
13260
13320
13380
13440
13500
13560
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline \(g g c g g c a\) & agtcagcaag & ggaaa & ctc & & & \\
\hline acttatctgt & taacttgtca & ccataaatat & tctgggagat & taaatacata & ctttagaaat & 13680 \\
\hline taaaaaaaca & tgattgtatc & aaagttttga & gtgtagtgga & atggaactg & tgggtaagca & 13740 \\
\hline agcatttggt & acttgttgcc & tgcattggg & taagatggga & aagttacaat & ggggaacttg & 13800 \\
\hline gaacaatttc & aatcccttca & tgg & gagaat & gcaaactatg & aactattaaa & 13860 \\
\hline ccttcccact & acttcctttt & cctccaatct & caaadaagaa & agggtgctag & aaatgctatg & 13920 \\
\hline tgtagagcaa & gcctattatt & tgctgtctac & aatggtatgt & gcttcaatta & tgcaggaacg & 13980 \\
\hline acaggtgtaa & tctgagcctg & tcctgttcag & acttgggaca & tgtggtcact & cagttttggg & 14040 \\
\hline ttctccaaat & caatgttgga & aga & C & agaacattct & tgattgtcac & 14100 \\
\hline atcttacaaa & aatgactctg & C & aacttcaggt & cagaggagct & ggggatagtg & 14160 \\
\hline gggttttcca & gagcattagc & agggagtgta & gagaataaag & gatgatattt & ctaggaactc & 14220 \\
\hline agaacagggt & gttactgttt & tgtaaagtgt & gaagaggaa & ttggctctgg & gcatagagtc & 14280 \\
\hline tgtagtcaga & caacgc & t & cactaggaa & gagttaatta & ttctactctt & 14340 \\
\hline gttctgctga & agcacagagc & tacatatct & atatcatcc & cactcaaca & catgctactg & 14400 \\
\hline tagttgtctg & ataatgggtc & ctgtcttcc & atgactggg & ctccttgacc & tcagaggtga & 14460 \\
\hline gtctaactc & gcttggtgtc & tccatcaccc & cagcatagg & gccagctcca & tcactggcac & 14520 \\
\hline cagataa & C & gagtagatgg & aagatgattc & agcagatagt & tctgaaagtc & 14580 \\
\hline tgtggctctt & tatgtg & actggatat & ct & ca & tagtggaag & 14640 \\
\hline gacggtaaga & ggtgctgatt & aattttcc & tatctttct & ccactcagca & tctttggggc & 14700 \\
\hline ctacagcatg & gatgtgatta & tggcacatc & ttggagtg & acatcgact & ctctcaacaa & 14760 \\
\hline tccacaaga & ccctttgtgg & agagcactaa & aagttccta & aatttggt & tcttagatcc & 14820 \\
\hline attatttctc & tcaataagta & tgtgggctat & atttctttc & ctcttttta & aaaataactg & 14880 \\
\hline ctttcttgac & atataattca & catatcgtat & attcatcca & cttaaajggt & acaattccat & 14940 \\
\hline tgtttttaag & ataatcaaaa & tatgtatga & cattactat & tgtaaactaa & aatgtttttg & 15000 \\
\hline tcaatctaga & gccctcacac & ctttagctg & caacacccc & ccacaaacc & ccactgccct & 15060 \\
\hline aagcatccaa & taatcaactt & tctgcctcta & agatttgcc & attctggac & acttcataga & 15120 \\
\hline aataatatca & ttgatttttc & ctgttgttt & ttattctct & tttcatgag & tttattttag & 15180 \\
\hline tctgttattt & ctttctttt & gctggcttta & ggtttcattt & gctcttcttc & ttttagtgtt & 15240 \\
\hline ttgtggtgta & aataattata & atcaatttga & atattttct & cttttaaat & ttagatatta & 15300 \\
\hline cagctataaa & tttccctctg & agcactggtt & tggctacatc & ctgtgttttg & gtacatcatg & 15360 \\
\hline ccttcttttt & gttcatctca & aaacaatttc & ttgttgccct & tttgatttct & gctttgactc & 15420 \\
\hline actggtcact & taaaactgta & ttgtttaact & tccacaaatg & tatgagtttc & ccaaatttct & 15480 \\
\hline ttcccttatt & gatttctagt & tttattccat & ggaagttgat & gtacatatgc & tgtgttaatt & 15540 \\
\hline ctatcttgac & tatcatttcc & tgaacagcat & gattaagtta Page 13 & agcagcagat & tatggtctac & 15600 \\
\hline
\end{tabular}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline aatccaa & aaactctagt & ccaatagata & aggctaaga & gtcagggaa & ttaattcta & 15660 \\
\hline ttactttggt & cactccaaag & actcagaagg & tgccattga & ctcactgctg & tagtggtgtt & 15720 \\
\hline tcctatgtat & agacctgc & ttgctcagtc & ccggcctg & agaagggca & aacatgataa & 15780 \\
\hline aaggaatggg & tt & a & gttcttattc & tattactgg & agagaaaat & 15840 \\
\hline tat & ccaggtaaag & - & caatgatttc & g & t & 15900 \\
\hline cccacag & C & ccttacccca & & t & t & 15960 \\
\hline ccaaaaga & ccataaa & tttaagtaaa & & gaatgaagaa & agtcgcctc & 16020 \\
\hline aacga & aa & atctgatggt & ttaaatga & gatgtttag & a & 16080 \\
\hline atttag & tatacacatg & atagagcatg & tctgtatt & ttaaaaata & agacagaga & 16140 \\
\hline acttatg & agaacaagag & aagccatttg & tagaaataa & agaaggagat & ggggaagga & 16200 \\
\hline gatgaga & agtcagag & atagcattta & aacttgaaa & tcaggcacaa & aattagtat & 16260 \\
\hline g & aamcagta & ga & taccactt & tcttc & taataaattg & 16320 \\
\hline tcaaag & aag & tttgaaaata & attttactg & tattgt & tcctcatat & 16380 \\
\hline cacaga & taadg & ttttaagtcc & agactctta & ttttacatat & ctgcaatta & 16440 \\
\hline aagg & gagg & ga & gacatgtagt & gtgtggtaaa & gtgagtgtt & 16500 \\
\hline tc & g & gggt & gag & gttgcaagg & ctggctaact & 16560 \\
\hline cagctt & ttcacgagcc & ctagaggcca & gccgaaggat & tctgcag & cagggagaca & 16620 \\
\hline ggaccagg & acccagctgt & tgaa & tatatagagt & ag & ggaatattt & 16680 \\
\hline gaaaatgc & C & tg & agttctggaa & tgtcaggag & attaatctat & 16740 \\
\hline acggacac & Ct & g & aga & agtacttct & cctgggtaa & 16800 \\
\hline gcagttat & agag & Cg & tctttggcca & aajgctgta & ccaaaagac & 16860 \\
\hline agagaagatg & agaa & g & cgaaaaagca & tggacatg & tagctagat & 16920 \\
\hline ttgtttcagg & aa & gctttccaag & atttaga & atgtttttg & tcactggtg & 16980 \\
\hline actcaggtaa & cacgtctt & gaag & ggaggtt & ggagggaag & tcaagaaggg & 17040 \\
\hline aggttgagga & gc & ttt & gacttcacg & gtcactttc & gccaaagaa & 17100 \\
\hline atctctcctt & tgcttctag & ccgactag & tttccttca & ctgatgatt & actcccaga & 17160 \\
\hline attcgaaaga & gagt & aaaggta & ggagt & cttctgagg & gctactggcg & 17220 \\
\hline gggacactaa & gagggagggc & tgt & aatgtgcag & gaagtattcc & ggaagatga & 17280 \\
\hline gaatttttgc & acatagcag & acaacacac & ttagatgt & ataaatggt & gctggaggc & 17340 \\
\hline actttccaga & gcccacagg & atagccatg & tccaggctg & aagggcaac & cctaagcaaa & 17400 \\
\hline cctagaatgc & ttggaggaca & tcagtggtt & tgtggatcac & ctacatgaga & tcaaatgcca & 17460 \\
\hline \(g t t c t c a g c c\) & tcctccagat & caccaagtg & agaacctcta & cttggaaatt & tatatcaaac & 17520 \\
\hline ataccgatca & ggaagcacac & tatcccagta & agggtgattt & taactggcag & tacttgaaag & 17580 \\
\hline tgtgttcgca & aggttaatct & actgcaaagt & tttatttttc & cctttgaaat & gcataagtaa & 17640 \\
\hline
\end{tabular}
ctaatggggg acacctctga taccatgtaa atctacttca atcttcagtc ttgtatctac 17700 tagttttatg acccatggat ggttttaacc aaaaccatta ttactaagac agtggcaaaa 17760 tgataaccat ggtcaatttc aagctaccaa gatttggcaa ccatctcaca aaatttttga 17820 atatttaaca attggttcta gagagcagga ctcagcagac tccagtatac cactttaaac 17880 atgtccatgt ctacatctac ttctgtctgt ctatctatct gtcaatcatc tatctgccta 17940 taatttatca attaatcatc tatctatctc aacaaaactt gctgtgataa agaaaatagt 18000 ctatcatttc actgtttcat atagaaatca ctagacacat atggctattg agtactggac 18060 atgtggccaa tgccactgaa gaacaatttt taagagtatt tatttttaat tgaataaaat 18120 ttgaatttaa atagccacat gtggatagtg gctaccagat tggacagcag agctcccaac 18180 tttaaaatta cagttcaatt tcaactcagt ataatggggt tcaatgtaac tgagtaaaat 18240 aattggatgg ttgaatttac ccacagcagc atacagaaat attcactgat aaatcagaac 18300 tctgtagacc tttctcacac tcattttata ttgtgtttgg ttgtgagtta catgattgct 18360 gcaggcacca tatttatttc tgtgctccag gtctctaagg gtcctaatcc agtcctgacc 18420 aaacagacta gtgatggacc atcgtgagct tctctcagga gaaatatcaa gagggaggcc 18480 aacctgtaat cataagaact tctgctattt taatgccatt catcagacta cagtcaatca 18540 ccatgcttct ggctttttgt ctatctctgc tgtcttgtac atcctgagat agtccattct 18600 gagaactgta ccctagatct tgtattgcct gatgcctgtc aaagatgtaa tccatgctgc 18660 ttaagtgagg ttgtgcacac aaatcaccat atctcctgca agtttggatt ttgattcagt 18720 agttcgatgg tggggtttga gattctgcat ttctaataag ctcccagatg tggctggtgc 18780 tgctggtcca tgaaacacac tttgagtagc aagaggtgat ctgtagctca gtattggtcc 18840 tttaagttcc ctcaaacata tatagagaaa aggtcctaaa tattgcaaat tctctcaaag 18900 tttgtcaagc tatattggaa ttctctcaaa gtctgtcaag ctctattgta gaaaatcaaa 18960 tttttattgg gaaaaagcct accccatatt tacttacaga taaagtactt ttaggatcat 19020 tcaaggcaca cacccataac actgagtatg taagacagaa atgctctctc tggaaattac 19080 agcagtgctg gtgctgggat gccatgatga ggagtgtgtg gcccacaatc atgtagacct 19140 tgggaaaacc tggattaaaa tgattttgcg tcatcctggc cctgtataag atacatatca 19200 gaatgaaaac cactcccagt gtgactttga attgcttttc cattttttct tcttgggatt 19260 agagagcttc acttagattt catctaagct gtgatgttgt acgttgacct gatttaccta 19320 aatgtcttt cctctccttt cagctctgtc tgatctggag ctcgcagccc agtcaataat 19380 cttcattttt gctggctatg aaaccaccag cagtgttctt tccttcactt tatatgaact 19440 ggccactcac cctgatgtcc agcagaaact gcaaaaggag attgatgcag ttttgcccaa 19500
taaggtgagg ggatgacccc tggagatgaa gggaagaggt gaagccttag caaaaatgcc 19560 tcctcaccac tccccaggag aatttttata aaaagcataa tcactgattc cttcactgac 19620 ataatgtagg aagcctctga ggagaaaaac aaagggagaa acatagagaa cggttgctac 19680
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline tggcagaagc & ataagatctt & & & \(g \mathrm{ttcacctg}\) & a & 19740 \\
\hline tcacaataat & gctaagtaaa & aaaaaaaaaa & aaaaaaaaaa & aaaaaaaagg & agtgtggcga & 19800 \\
\hline gaagatggcc & aaacaggaac & agctccagtc & C & agcgtgagca & acacagaaga & 19860 \\
\hline cgaatgatt & ctgcatttcc & aactgaggta & ccgggtgcat & ctcaatgggg & g & 19920 \\
\hline agtgggtg & ggacagtggg & tgcagtgcac & ccagcctgag & ccaaagcagg & gcgaggcatc & 19980 \\
\hline acct & g & g9ggtcaggg & aattcccttt & cctaggggtg & cggacagca & 20040 \\
\hline cctggaaaat & cag & ccaccctaat & actgcgcttt & tctgatggtc & ttagcaaacg & 20100 \\
\hline gcacac & agattatatc & ccgcgcatgg & ctcggagggt & cctacgccca & tggagcctcg & 20160 \\
\hline Ct & gcacagc & tgaga & actgcaaggc & agcagcaagg & ctgggggagg & 20220 \\
\hline g & a & g & taaacaaagc & ggaag & ctcaaactgg & 20280 \\
\hline gtga & C & a & cctgcctctg & tagactccac & C & 20340 \\
\hline agagcata & caaccaaaag & gcagcagaaa & C & C & tgtctgac & 20400 \\
\hline agc & agagtagt & t & acacagctgg & agatctgaga & acagacagac & 20460 \\
\hline t & a & tgacceccga & C & C & cccccagtag & 20520 \\
\hline g & acacctcaca & cggccgggta & g & acaaaacttc & gaaggaatg & 20580 \\
\hline atcaggcagc & agcat & ggtcaccaat & accgctg & tgcagcctcc & actcctgata & 20640 \\
\hline CC & caggg & g & ggcaaactc & aacagacct & gcagctgagg & 20700 \\
\hline a & tca & aactaacaaa & cagaaaggac & atccacacca & aaacccatct & 20760 \\
\hline gtacat & atcatcaaag & atcaaa & gat & caaagatggg & ggaaaaacag & 20820 \\
\hline cagaaaaact & gaaaaatc & aaatcagag & cacctctcct & ctccaaagg & aacgcagctc & 20880 \\
\hline cgcacc & ac & ggatggagaa & tgactttgac & gagttgagag & aagaaggctt & 20940 \\
\hline cagacgatca & aactact & agctaaagga & a & ccca & aagaagttaa & 21000 \\
\hline aaaccttgaa & aaaagattag & acaaatgg & actaga & tcaatgcag & agaagtcctt & 21060 \\
\hline aaaggacctg & atggagctg & gaccatggc & cgaga & gtgatgaa & gcacaagcct & 21120 \\
\hline cagtagcca & ttcaatcaac & ggaagaaag & ggtatcagtg & tggaagatc & aaatgaatga & 21180 \\
\hline aatgaagaaa & gaagagaagt & tagaaga & agaataaaa & agaaaggaac & aaggectcca & 21240 \\
\hline agaaatatgg & gactatgtga & aagaccaaa & ctacgtctg & ttggtgtac & ctgaaagtga & 21300 \\
\hline cggggagaat & agaacgaagt & ggaaaacac & gcag & ttatccagg & agaacttccc & 21360 \\
\hline caatctagca & aggcaggcca & acattcaaat & tcaggaaata & cagagaacgc & cacaaagata & 21420 \\
\hline ctcctcgaga & agagcaactc & aagacacat & attgtcaga & tcaccaaag & ttgaaatgaa & 21.480 \\
\hline ggaaaaaatg & ttaagggcag & ccagagagaa & aggtcgggtt & acccacaaac & acaaacccat & 21540 \\
\hline cagactaaca & gtggatctct & cggcagaaac & tctacaagcc & agtagagagt & gggggccaat & 21600 \\
\hline attcaacatt & cttaaggaaa & agaattttca & acccagaatt & tcatttccag & Ccaaactaag & 21660 \\
\hline cttcataagt & gaaggagaaa & taaaatactt & tacagacaag & caaatgctga & gagattttgt & 21720 \\
\hline
\end{tabular}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline caccaccagg & cctgccctaa & aagagctctt & gaaggaagca & ctaaacatgg & aaaggaacaa & 21780 \\
\hline ctggtaccag & ccactgcaaa & aacatgccaa & attgtaaaga & ccatcgaggc & taaggagaaa & 21840 \\
\hline ctgcatcaac & taacgagcaa & aataatcagc & taacatcata & atgacaggat & caaattcaca & 21900 \\
\hline tataaaaata & ttaaccttaa & atgtaaacgg & gctaaatgct & ccaattaaaa & gacacagact & 21960 \\
\hline ggcaaactgg & atagagtcaa & gacccatcgg & tgtgctgtat & tcaggaaacc & catctcacgt & 22020 \\
\hline gcaaagtaac & acataggctc & aaaataaagg & gatggaggaa & gatctaccaa & gcaaatggac & 22080 \\
\hline aacaaaaaaa & ggcaggggtt & gca & tctctgataa & aacaggcttt & aaaccaacaa & 22140 \\
\hline agatcaaaag & agacaaagaa & ggccattaca & taatggtaaa & gggatcaatt & caacaagaag & 22200 \\
\hline agctaactat & cctaaatata & tatgcaccca & atacaggagc & acccagattc & atgaagcaag & 22260 \\
\hline tctttagaga & cttacaaaga & gagttagact & cccacacaat & aataatggaa & gactttaaca & 22320 \\
\hline ccacactgtc & aacactagac & ag & gacagaaagt & taagaaggat & atccaggaat & 22380 \\
\hline tgaactcagc & tctgcacaaa & gtggacataa & tagacatcta & cagaactctc & caccccaaat & 22440 \\
\hline caacagaata & tacattcttt & tcagcaccac & accacaccta & ttccaaaatt & aaccacatag & 22500 \\
\hline ttggaagtaa & ag & agcaaatgta & aaagaacaga & cattataaca & aactgtctct & 22560 \\
\hline cagaccacag & tgcaatcaaa & ctagaactca & ggattcagaa & actcactcaa & aaccgctcaa & 22620 \\
\hline ctacatggaa & actgaacaac & ctgctcctga & atgactactg & ggtacataac & gaaatgaagg & 22680 \\
\hline cagaaataaa & gatgttcttt & gaaaccaaca & agaacaaaga & cacaacatac & cagaatctct & 22740 \\
\hline gggccacatt & caaagcaatg & tgtagaggga & aatttatagc & actaaatgcc & tacaagagaa & 22800 \\
\hline agcaggaaag & atctaacatt & gacaccctaa & catcacaatg & aaaagaacta & gagaagcagg & 22860 \\
\hline agcaaacaca & ttcaaaagat & agcagaaggc & aagaaataac & taagatcaga & gcagaactga & 22920 \\
\hline aggaaacaga & gacacaaaaa & aacccttcaa & aaaaatcaat & gaatccagga & gctggttttt & 22980 \\
\hline tgaaaagatc & aacaaaattg & atagaatgct & agcaagacta & ataaagaaga & aaagagagaa & 23040 \\
\hline gaatcaaata & gatgcaataa & aaatgataaa & ggggatatca & ccacccatcc & cacagaaata & 23100 \\
\hline caaactacca & tcagagaata & ctataadcac & ctctatgcaa & ataaactaga & aaatctagaa & 23160 \\
\hline gaaatggata & aattcctcga & cacatacact & ctcccaagac & taaaccagga & agaagttgaa & 23220 \\
\hline actctgaata & gaccaataac & aggttctgaa & attgaggcaa & taattaatag & cttaccaacc & 23280 \\
\hline aaaaaaagtc & caggaccaga & tggattcacc & gccgaattct & accagaggta & caaggaggac & 23340 \\
\hline ctggtaccat & tctttctgaa & actattccaa & tcaatagaaa & aagagggaat & cotccotaac & 23400 \\
\hline tcattttatg & aggccagcat & catcctgata & ccaaagcctg & gcagagacac & aaccaaaaaa & 23460 \\
\hline gagaatttta & gaccaatatc & cctgatgaac & agtgatacaa & aaatcctcaa & taaaatactg & 23520 \\
\hline gcaaaccgaa & tccagcagca & catcaaaaag & cttatccacc & atgatcaagt & gggcttcatc & 23580 \\
\hline cctgggatgc & aaggctggtt & caacatacgc & aaatcaataa & acataatcca & gcatataaac & 23640 \\
\hline agaaccaacg & acaaaaccca & catgattatc & tcaatagatg & cagaaaaggc & ctttaacaad & 23700 \\
\hline attcaacagc & ccttcatgct & aaaaactctg & aataaattag Page 141 & gtattgatgg 1 & aacctatctc & 23760 \\
\hline
\end{tabular}
aaaataataa gagcaaattt atgacaaacc cacagccaat atcatactga atggacaaaa
23820 actggaatca ttccctttga aaactggcac aagacaggga tgccctctct caccactcct attcaacata gtgttggaag ttctggccag ggcaatcagg caagagaaag aaataaaggg tattcaatta ggaaaagagg aagtcaaatt gtccctgttt gcagatgaca tgattgtata tctagaaaac cccatcgtct cagcccaaaa tctccttaag ctgataaaca acttcagcaa agtatcagga tacaaaatca atgtgcaaaa atcacaaata ttcttataca ccaataacag acaaacagag agccaaatca tgagtgaact cccattcaca attgcttcaa agacaataaa atacctagga attcaactta caagggatgt gaaggacctc ttcaaggaga attacaaacc actgctcaat gaaataaag aagatacaaa caaatggaac aacattccat gctcatgggt aggaagaatc aatatcatga aaatggccat actgcccaag gtaatttata gattcagtgc catcgccatc aagctaccaa tgactttctt cacagaactg gaaaaaacta ctttaaagtt catatggaac caaaaagag cccgcattgc caagtcaatc ctaagccaaa agaacaaagc cggaggcatc atgctacctg acttcaaact atactacaag gctacagtaa ccaaaacagc atggtactgg taccaaaaca gagatattga tcaatggagc agaacagagc cctgagaaag aatgccacat atctacaacc atctgatctt tgacaaacct gacaaaaaca agcagtgggg aaaggattcc ctatttaata aatggtgctg ggaaaactgg ctagccatat atagaaagct gaaactggat cccttcctta caccttatac aaaaattaat tcaagatgga ttaaagactt acatgttaga cctaaaacca taaaaaccct agaagaaaac ctaggcaata tcattcaata cagaggcatg ggcaaggact tcatgtctaa aacaccaaaa gcaatggcaa caaaagccaa aattgacaaa tgggatctaa tgaaactaaa gagcttctgc acagcaaaag aaactaccat cagagtgaac aggcaaccga cagaatggga gaaaattttt gcaacctact catctgacaa agggctaata tccagaatct acaatgatct caaacaaatt tacaagaaaa aaacacaacc ccatcaacaa gtgggggaag gatatgaaca gacacttctc aaaagacatt tatgcagcca atagacacat gaaaaaatgt tcatcatcac tggccatcaa agaaatgcaa atcaaaacca caatgagata ccatctcacg ccagttagaa tggcgatcat taaaaagtca ggaaacaaca ggtgctggag aggatgtgga gaaaacagga acacttttac actgttggtg ggactgtaaa ctagttcaac cattgtggaa gtcagtgtgg tgattcctca gggatctaga actagaaata ccatttgacc cagccatccc attactgggt atatacccaa aggattataa atcatcctgc tataacaca catgcacact tatgtttatt gcagcactat tcacaatagc aaagacttgg aaccaaccca aatgtccaat aatgatagac tggattaaga aaatgtggca catatacacc atggaatgct atgcagccat aaaaatgat gagttcatgt cctttgtaga gacatggatg aagctggaaa ccatcattct cagcaaacta tggcaaggac aaaaaccaa acactgtatg ttctcactcg taggtgggaa ttgaacaatg agaacacatg gacacaggaa ggggaatatc acacactggg gcctgttttg gggtgggagg agtggggagg gatagcatta ggagatatac 25800

\begin{tabular}{|c|c|c|c|c|c|c|}
\hline atattgca & tatgccataa & accatttaac & catgatgtta & ag & aggctttt & 27900 \\
\hline tattcctttc & tgttttttct & atgctgtgcc & ctttagctct & ctgaatttaa & cagaadettt & 27960 \\
\hline aaaacatgct & tccacattcc & tgctttc & aacgttactt & gctatttcct & ctgtagtaat & 28020 \\
\hline tataagagtg & caggctgagg & tcctgagaag & tcctcatccc & taatggttta & agccacttca & 28080 \\
\hline ctgaagacac & aagacagcac & aggtcctcct & ggtcctat & gtggctgcag & tcctgtgcca & 28140 \\
\hline gctcccttat & actctcagta & gacatctcac & acactcctcc & ttggaggtgt & ttgagcatg & 28200 \\
\hline ctcttctggg & aattcaggga & caaggtcagg & ccttaggcac & agttcgc & ctggatatag & 28260 \\
\hline ttggtgtttt & cccattactg & tattattaag & caaaatttag & aatgaa & ttagggtact & 28320 \\
\hline ggctggtgat & tcaggatgct & tgggatctag & actttcatta & gcccctacct & gcaagtttgc & 28380 \\
\hline tgatgggagg & aaccttgtct & tgttggtcat & ggtgtcccta & gtgctag & ggagtctgca & 28440 \\
\hline cataatactt & \(g\) & taagtcagag & ctgaccaagt & t & ctggagtaga & 28500 \\
\hline ggacttctat & gtttcctgc & agctcagcac & tccacctcc & tgtggctgca & ctaatacgaa & 28560 \\
\hline atcagagacc & actcgctgta & cttcactttg & aatcactcag & tcaccaaaaa & gatagtgctt & 28620 \\
\hline gccatgtgtc & aggaacttgg & ctaggcaggg & agaaattcat & atgatttata & taaatccata & 28680 \\
\hline aatccatatg & atttacataa & atccataaat & t & tatacgtata & tgtgtgtgta & 28740 \\
\hline tatatatatt & agagaatgt & tgacatatac & acaagtacat & gttaccgaca & ccagcctata & 28800 \\
\hline gaatagtttt & cgtgcatctc & catatatcta & tcactggttc & caacagccat & caatccatgt & 28860 \\
\hline tagctgccec & atccaaatgc & caccatcacc & ctcctcctga & ctatcatgt & attttgaagc & 28920 \\
\hline aatagcctgt & aaatatttca & gaatgctctc & caaaatataa & agactcctgt & aaaaacatat & 28980 \\
\hline gacaacaatg & ccattattac & ttctttgaa & tcaacatttt & ttccttaata & taatcaaata & 29040 \\
\hline tttagaaatc & aaatttgaat & aaaacatggg & tcaatcttca & aagaatttat & agcttaatgg & 29100 \\
\hline aacagatcaa & ggaaagcagg & gatgacacta & cagtagggta & gcatcatatg & cccatgtaac & 29160 \\
\hline ttatgtgact & taaactatcc & tgtaagggtg & tgggggagaa & agagaggaag & agatggagag & 29220 \\
\hline aagaaaaagg & aagagaagga & ggaggagaag & gaggcagagg & agaaggtgga & cggggaaggt & 29280 \\
\hline agagaggagg & aggaggggaa & ttagaaaaaa & agagatgaca & ggagaaggaa & agggaaaaat & 29340 \\
\hline aacaacttga & aatagcacaa & gacgttttct & ccttctcctt & tctcaatgag & catgtgacca & 29400 \\
\hline acacaagtgt & gagttgaggc & aggaatccac & tttccatcc & atcagtctta & tcatttatgt & 29460 \\
\hline gccttttata & gtgtgaacac & atcaccaccc & tgaatataat & tttagtgttt & agagataaat & 29520 \\
\hline attatttgca & acaatattca & tctcatctca & agaaacgctc & ctatagggta & tggagaattt & 29580 \\
\hline aamggacctg & taggttatga & tgattataac & gaaataacca & aagcaggatt & tcaatgacca & 29640 \\
\hline gcccacaaaa & gtatcctgtg & tactactggt & tgggaggtgg & aggggggttg & ttcttaagta & 29700 \\
\hline agaaccceta & acatgtaact & ctgtggtttt & tatgtttcat & taactattta & atctaccaat & 29760 \\
\hline atggaactag & gttcagtaag & aagaaggaca & gcatagatcc & ttacatatac & acaccctttg & 29820 \\
\hline gaactggacc & cagaaactgc & attggcatga & \[
\begin{aligned}
& \text { ggtttgctct } \\
& \text { Page } 144
\end{aligned}
\] & \begin{tabular}{l}
catgaacatg \\
4
\end{tabular} & aaacttgctc & 29880 \\
\hline
\end{tabular}
taatcagagt ccttcagaac ttctccttca aaccttgtaa agaaacacag gtcagtacac ..... 29940
tttctgtatg ttttattaag aattttttta actgaagggt atatattttt taaaagaata ..... 30000
tgcatgttta tcttttaata attcattcta tgggccaaag aacctacttg gatccatctt ..... 30060
tgatcattaa ggatgcttca gttctggact tcaaaacctg tagcattaag aacatcatgt ..... 30120
aaagtccaca cagattagca tgacatgatt atgtgtagtc tctttgaacc tgagtaagtt ..... 30180
taaattcagt ttcaagtcaa ttggaaagaa gtgttttgca caatcatgaa gtgcaatgat ..... 30240
tacctggctg tgacttaaat ggtgttctcc atcaccagaa cctgcagaag ctctctcatg ..... 30300
acagtggttc tcaaccacta gctgtatatt ggaatcacca gggagcttca aaaattcatg ..... 30360
atgcctgtga catctcagaa attctaaact aattaaccca gagcgtgact aggttctgtc ..... 30420
atgctgtcgg gtgaacccct gattagttct cacgtgaagc caaggtggag aatgactaat ..... 30480
ttcaggcatt tctggtggat atgaaggact accatagagc agggctatcc ttactccttg ..... 30540
accttatgtt ccaggtgata catttaaga aagatttaga atcttttctc tgaagaagtt ..... 30600
aaagaacaga tgtcattgat tcatattaag caatagccta taagtcttat ttccaggacc ..... 30660
ggtgtattta atatgcaact ctacccctta agtacacttt gtgcttggga gaggaggagg ..... 30720
atggagatgg ttgccatctt atctatggct tcagggcagc tgtgtagctt tcctatgtgt ..... 30780
gtattcaggc agggggctca gccctgagag aaagtgggcc tctggcacac ctgggacagg ..... 30840
gaagatattc cctggcaagc tctcaggcat ctcaggctgg cacttctttg tatccatggc ..... 30900
aatttgcttt cccctcactg aactgagatc agaatgttac tctgttggtg gctcccccaa ..... 30960
cagtgaaggg gtgactcagt gacaatagtg ctagaagtat gagtcaaaac actgtacaac ..... 31020
ttgagaaatt ccccgtttgc actacgcttg gaagccaaga ggagatgtta aaaagaaaag ..... 31080
aataattctt tctgaagaca tttcccatca ttgcacttga tgggttcaac tgggaagggt ..... 31140
tactagactc tggaagttga aaactgccca cataattaaa ctgtacaaca gctactcagg ..... 31200
attaccttgc aagttttaac ctataaaaat ttaactttat atagcacttc caaaatagtt ..... 31260
tgccataata cctactaatc tggatttaat ttttaaaact catcctttta acttaagatt ..... 31320
taaataaaaa aaaaaaaaca cgagtccaca agaatttgtc tcaggcctgg cacagagtca ..... 31380
gtgctccata aatattttgt taaacgatgg atggtgagtg cttttactat ccagtattta ..... 31440
cccagcttat agattaagta tgaagagttc aagatacatg gtgttaagag tcgtttttat ..... 31500
atgcttgcaa agcatttttg tcatattttt tctactttgc ttccatcttt tcttctttca ..... 31560
cttcatttat taattctcca tatgcttgtt taactattgt agatcccctt gaaattagac ..... 31.620
acgcaaggac ttcttcaacc agaaaaaccc attgttctaa aggtggattc aagagatgga ..... 31680
accctaagtg gagaatgagt tattctaagg atttctactt tggtcttcaa gaaagctgtg ..... 31740
ccccagaaca ccagagattt caacttagtc aataaaacct tgaaataaag atgggcttaa ..... 31800
tctaatgtac tgcatgagta gttggtgatt ttgtacattc attgagctct cccagagtct ..... 31860
gtgtagagtg ttgtgcatta tgtagtataa aggaggtgac caggtaagtg acagataggt ..... 31920
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline agactcagct & t & tcataggact & acctctaccc & & agcattatca & 31980 \\
\hline actcctcctg & agctctcatc & agagaataaa & tatttctcaa & C & cataactttt & 32040 \\
\hline aagaaaaata & agaattatca & tgatgactct & aatagtg & tttatatcac & gttttatttg & 32100 \\
\hline taatattcta & taagttttat & attaagcgaa & gtgataaaat & cccctttaca & aaaatattat & 32160 \\
\hline ctgatgccat & CC & aaagagaaat & ctatag & gaatgactga & aaaccagcaa & 32220 \\
\hline ataaacattt & tttat & gt & 292999 & cctttgtcag & aattccaatt & 32280 \\
\hline tgattattaa & cataggtg & agttaatctg & ctgtg & gcccattg & tggagaaaat & 32340 \\
\hline attcatagt & tcattctg & tttgaag & aacatatt & tg & caacgazgca & 32400 \\
\hline cttatcat & ta & g & ttttaccaca & ctcccetga & catttctgga & 32460 \\
\hline acacaggaaa & catgt & tatacgtctt & g & tcacctccc & attgtctta & 32520 \\
\hline atgcaatgaa & cactgaataa & aaattgtca & tcgtcag & tgattgggca & gcatgtctaa & 32580 \\
\hline aagcacta & catt & t & tcattttccc & tccttttctg & aatactaaag & 32640 \\
\hline ccatta & g & C & cacacattaa & ggtggacaag & agagtcatgg & 32700 \\
\hline tggctcca & tcagattc & agtgtgctgg & ggaaggcatc & cacatggagg & ggcagcctga & 32760 \\
\hline cctggaagcg & ggagcccaag & caatcagaga & aggggtccac & acagaggtgt & ggccttcaag & 32820 \\
\hline agcagccaga & gcctaaa & ggcetggaga & acccacgtga & ggtgaggagg & gtatccctga & 32880 \\
\hline gtg & at & \(g t\) & t & t & agtaaataaa & 32940 \\
\hline gtatactgg & agctaggt & gtcacttttg & cagaaaagag & tcatagattc & agaaagggag & 33000 \\
\hline aaagctagca & ttaatcctat & ggtgttagat & tggaatggat & gtatcagtgt & acattcatac & 33060 \\
\hline ttttctagat & agatag & tagatagac & agatgataga & tagataatag & atagttgata & 33120 \\
\hline gataattaga & tgtaaatata & tgtgtttgta & tgtgtgagca & tgcatgtgtg & tgtgtgtgtg & 33180 \\
\hline tgtgtgtgtg & tgtgtgtgtg & tgtgtataaa & atatattc & cctacttca & tgatagggct & 33240 \\
\hline agataacaat & gacatttcaa & tagcaatgag & tatacttagt & gcctagatct & tggcttatga & 33300 \\
\hline atatcatttt & ccactgaagg & gaaccagagc & gttagaga & aatagctgat & gccagggcta & 33360 \\
\hline ggactaaaaa & tgttcaagat & gagcccagga & actttttgt & gccaggaagt & aagaaaacac & 33420 \\
\hline tcaaatggtt & tctaaatggt & ttggaaa & atttctaaa & tgatatttcc & atatgacttc & 33480 \\
\hline caaatgatat & tttaaaaacc & aaggactcc & accaaagaac & tattagaact & gataaacaaa & 33540 \\
\hline ttcagtaaat & ttgcaggata & aaaatcaac & tacaaaatt & ggtagcactt & ctatatgcca & 33600 \\
\hline gcaaggaaca & atctgaaata & aaatcaatag & ccacaaataa & aattaaatac & ctaggaatta & 33660 \\
\hline acttaatcaa & agaagtgaaa & gatctctaca & atgaaaacta & taaaacaccg & atgaaagaaa & 33720 \\
\hline ttgaagagga & cacaaaaaca & tagaaagata & ttccatgttt & atatatttca & agaatcaata & 33780 \\
\hline ttgttaaaaa & tgcccacaat & acccaaagca & atatatggat & tcaatgcaat & ccctatcaaa & 33840 \\
\hline ataccaatga & cattcttcac & agaaatattt & ttttaatcct & aaaatttgta & tggattcaca & 33900 \\
\hline aaagaccaag & aagagccaaa & gctaacatca & gcaaaaagaa & caaaactgga & ggaaacacat & 33960 \\
\hline
\end{tabular}

```

<210> 23
<211> 1760
<212> DNA
<213> Homo sapiens

```


```

gttattctaa ggacttctac tttggtcttc aagaaagctg tgccccagaa caccagagat1658

```
ttcaacttag tcaataaaac cttgaaataa agatgggctt aatctaatgt aaaaaaaaaa ..... 1718
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa ..... 1760
```

<210> 24
<211> 502
<212> PRT
<213> Homo sapiens
<400> 24
Met Asp Leu ITe pro Asn Leu Ala Val Glu Thr Trp Leu Leu Leu Ala

```

Phe Lys \(\underset{35}{\text { Arg Leu Gly Ile Pro }} \underset{40}{\text { Gly }}\) Pro Thr Pro Leu \(\underset{45}{\text { Pro }}\) Leu Leu Gly

Tyr Lys Lys Tyr G7y
65
70



ITe \(\underset{130}{\text { Arg }}\) Ser Leu Leu Ser \(\underset{135}{\text { Pro }} \begin{gathered}\text { Thr Phe Thr Ser G7y Lys Leu Lys G7u } \\ 140\end{gathered}\)

Arg Arg Glu Ala \(\underset{165}{\text { G7u }}\) Lys Gly Lys Pro Va才 \(\underset{170}{\text { Va }}\) Thr Leu Lys Asp \(\underset{175}{\text { I7e }}\) Phe
Gly Ala Tyr Ser met Asp val Ile \(\begin{gathered}\text { Thr } \\ 185 \\ 180\end{gathered}\) Gly Thr Ser Phe G7y val Asn \(\begin{gathered}190\end{gathered}\)
Ile Asp Ser Leu Asn Asn Pro Gln Asp Pro Phe Val Glu Ser Thr Lys
    195200205
Lys Phe Leu Lys Phe Gly phe Leu Asp Pro Leu 210 Phe Leu Ser Ile Ile
                                    Page 150
Leu Phe Pro Phe Leu Thr Pro Val Phe Glu Ala Leu Asn Val Ser Leu
225
230
G7y Thr Leu Ser G7y G7u ..... 500
<210> . 25
<211> 3612
<212> DNA<213> Homo sapiens
<220>
<221> misc_feature
<222> (1923).. (1923)
\(<223>\mathrm{n}\) may be any nucleotide<400> 25cagggaagct ccaggcaaag agcccagcaa acagcagcac tcagctaaaa ggaagactca60
cagaacacag ttgaagaagg aaagtggcga tggacctcat cccaaatttg gcggtggaaa ..... 120
cctggcttct cctggctgtc agcctggtgc tcctctatct atatgggace cgtacacatg ..... 180
gactttttaa gagactggga attccagggc ccacacctct gcctttgttg ggaaatgttt ..... 240
tgtcctatcg tcagggtctc tggaaatttg acacagagtg ctataaaaag tatggaaaaa ..... 300
tgtgggggtg agtattctga aaacctccat tggatagacc tgctactgtg aggaggttac ..... 360
cccactgcag gatagtctct gcccaggtct tcatgggatg aagctcttgt caacctaaat ..... 420
acaaacagag agaggttctc tgaaagaaga ggataattac ttgggagtag aatattgcaa ..... 480
tgggaatctg cttgccgtta taaactatgt gcaaattcag ggaggtaaac aagacaaaga ..... 540
tgctccatag aaaatatgag aagaatctca taactgtttt gagataatta ttgttagcta ..... 600
caaagatcaa taacaagggt gatgccacac caaggttgga caggcagttg ctggacaggt ..... 660
gtccttgcag aaatattttt gtgtaaagtt gaaatagcct ttgtgcaaag ttgtggtttt ..... 720
tgtagacact tttgtaatag ttttgtttcc aggaacacaa gcataagaat cctctcttca ..... 780
tagccttctt gggatttatt tgtcagggtt aaaaaacaat tagtgacatc actttggttc ..... 840
tgataaagtt cacactcgct attgtaaaac ttttcgaggc ttgtcctacc aaggatccca ..... 900
tgtgtcacca ggtatcgagg tcttcagtct gaactaggct aggagcattg tggttaccac ..... 960
ttttctgcag gttttggtgg cccagggact cccagcatcg ccttctgtcc agtgtctgcc ..... 1020
tattcccctc ttcttttttt cttccttagg tgccctttta tcacatgcat tgtctcagac ..... 1080
ccttctaata tgtgctcata aatgcatggc atcatctcct tcccacattg attcactttc ..... 1140
aattaaaagc caaaactcct tcatttagac tgaatttaac atgtgctttt gaaagaaggg ..... 1200
ttgagagata atagagaaac agattgggaa accacttatg ctccactttt ttaaactttc ..... 1260
tctgcaagta tggaattttt tgttctgctt tgttgtttaa atttaagcca aaacttctta ..... 1320
atagaaggat atacaaatat ttattggttt ataccattgc acttactttg aagaagagat ..... 1380
gctgaatatt attaaaccat tgtgttccct ggtgggctga tggactgtga ttttataagg ..... 1440
tggtctcagc caattgcagc agctgttcce tgtcagaggg gctagaggtt tggtgagagc ..... 1500
agtggatgag gtgcagtggt gtgtttgttc actagaagca agtgggagaa agctttgcct 1560 ctttgtactt cttcatcttc tcccctcaag tcctcagaat ccacagcgct gactgtggag 1620 tgctgtggag ctggcatggc ccatacaggc aacatgactt agtagacaga tgacacgctc 1680 tagatgtcca tgggccccac accaactgce cttgcagcat ttagtccttg tgagcacttg 1740 atgatttacc tgccttcaat ttttcactga cctaatattc tttttgataa tgaagtattt 1800 taaacatata aaacattatg gagagtggca taggagatac ccacgtatgt accacccagc 1860 ttaacgaatg ctctactgtc atttctaacc ataatctctt taaagagctc ttttgtcttt 1920 cantatctct tccctgtttg gaccacatta cccttcatca tatgaagcct tgggtggctc 1980 ctgtgtgaga ctcttgctgt gtgtcacacc ctaatgaact agaacctaag gttgctgtgt 2040 gtcgtacaac taggggtatg gattacataa cataatgatc aaagtctggc ttcctgggtg 2100 tggctccagc tgcagaatcg ggctagtgaa gtttaatcag ctccgttgtc cccacacaga 2160 acgtatgaag gtcaactccc tgtgctggcc atcacagatc ccgacgtgat cagaacagtg 2220 ctagtgaaag aatgttattc tgtcttcaca aatcgaaggt ctttaggccc agtgggattt 2280 atgaaaagtg ccatctcttt agctgaggat gaagaatgga agagaatacg gtcattgctg 2340 tctccaacct tcaccagcgg aaaactcaag gagatgttcc ccatcattgc ccagtatgga 2400 gatgtattgg tgagaaactt gaggcgggaa gcagagaaag gcaagcctgt caccttgaaa 2460 gacatctttg gggcctacag catggatgtg attactggca catcatttgg agtgaacatc 2520 gactctctca acaatccaca agaccccttt gtggagagca ctaagaagtt cctaaaattt 2580 ggtttcttag atccattatt tctctcaata atactctttc cattccttac cccagttttt 2640 gaagcattaa atgtctctct gtttccaaaa gataccataa attttttaag taaatctgta 2700 aacagaatga agaaaagtcg cctcaacgac aaacaaaagc accgactaga tttccttcag 2760 ctgatgattg actcccagaa ttcgaaagaa actgagtccc acaaagctct gtctgatctg 2820 gagctcgcag cccagtcaat aatcttcatt tttgctggct atgaaaccac cagcagtgtt 2880 ctttccttca ctttatatga actggccact caccctgatg tccagcagaa actgcaaaag 2940 gagattgatg cagttttgcc caataaggca ccacctacct atgatgccgt ggtacagatg 3000 gagtaccttg acatggtggt gaatgaaaca ctcagattat tcccagttgc tattagactt 3060 gagaggactt gcaagaaaga tgttgaaatc aatggggtat tcattcccaa agggtcaatg 3120 gtggtgattc caacttatgc tcttcaccat gacccaaagt actggacaga gcctgaggag 3180 ttccgccctg aaaggttcag taagaagaag gacagcatag atccttacat atacacaccc 3240 tttggaactg gacccagaaa ctgcattggc atgaggtttg ctctcatgaa catgaaactt 3300 gctctaatca gagtccttca gaacttctcc ttcaaacctt gtaaagaaac acagatcccc 3360 ttgaaattag acacgcaagg acttcttcaa ccagaaaaac ccattgttct aaaggtggat 3420 tcaagagatg gaaccctaag tggagaatga gttattctaa ggacttctac tttggtcttc 3480 aagaaagctg tgccccagaa caccagagat ttcaacttag tcaataaaac cttgaaataa 3540 Page 153
agatgggctt aatctaatgt aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3600
aaaaaaaaaa aa
```

<210> 26
<211> 502
<212> PRT
<213> Homo sapiens
<400> 26

```
Met Asp Leu Ile pro Asn Leu Ala val Glu Thr Trp Leu Leu Leu Ala
\(15010 \quad 15\)

Phe Lys \(\underset{35}{\text { Arg Leu G7y Ile Pro G7y pro Thr Pro Leu Pro Leu Leu Gly }} \underset{45}{ }\)
Asn val Leu Ser Tyr Arg \(\underset{50}{\text { G7n }} \underset{50}{ }\) Gly Leu Trp Lys \(\underset{60}{\text { Phe }}\) Asp Thr G7u Cys
\begin{tabular}{l} 
Tyr Lys Lys Tyr Gly \\
65 \\
\hline 70
\end{tabular}
Val Leu Ala Ile \(\underset{85}{\operatorname{Thr}}\) Asp Pro Asp Val \(\underset{90}{\text { Ile }}\) Arg Thr val Leu Val \(\underset{95}{ }\) Lys
Glu Cys Tyr Ser val Phe Thr Asn
100
105 Arg Ser Leu G7y Pro Val Gly
Phe met Lys Ser Ala Ile Ser Leu Ala glu Asp glu glu trp Lys Arg
Ile \(\underset{130}{\text { Arg Ser Leu Leu Ser Pro }} \underset{135}{ }\) Thr Phe Thr Ser G7y \(\underset{140}{ }\) Lys Leu Lys G7u
Met Phe pro Ile Ile Ala G1n Tyr Gly Asp val Leu Val Arg Asn leu
\(145150 \quad 155160\)
Arg Arg Glu Ala Glu Lys gly Lys Pro Val Thr Leu Lys Asp \({ }_{170}{ }_{175}\) Phe

Ile Asp Ser Leu Asn Asn pro G7n Asp Pro Phe Val G7u ser Thr Lys
195200205
Lys phe Leu Lys Phe G7y phe Leu Asp Pro Leu Phe Leu ser Ile Ile
        \(210 \quad 215 \quad 220\)
Leu Phe Pro Phe Leu Thr Pro Val Phe Glu Ala Leu Asn Val Ser Leu
225
230

G7u Arg Phe Ser Lys Lys Lys Asp Ser Tle Asp Pro Tyr ITe Tyr Thr Pro Phe G7y Thr G7y Pro Arg Asn Cys Ile Gly Met Arg Phe Ala Leu Met Asn
450 Lys Pro Cys Lys G7u Thr G7n ITe Pro Leu Lys Leu Asp Thr G7n G7y
465
470 Leu Leu G7n pro G7u Lys Pro ITe Val \(\underset{485}{\text { Leu }}\) Lys Val Asp Ser \(\underset{49}{\text { Arg }}\) Asp

Gly Thr Leu \(\begin{gathered}\text { Ser G7y G7u } \\ 500\end{gathered}\)
```

<210> 27
<211> 102
<212> PRT
<213> Homo sapiens
<400> 27
Met Asp Leu Ile Pro Asn Leu Ala val glu Thr Trp Leu Leu Leu Ala
Val Ser Leu val Leu Leu Tyr Leu Tyr Gly Thr Arg Thr tis H0
Phe Lys
Asn val Leu Ser Tyr Arg Gln Gly Leu Trp Lys Phe Asp Thr Glu Cys
Tyr Lys Lys Tyr G7y Lys Met Trp G7y Thr Tyr G7u G7y G7n Leu Pro
Val Leu Ala Ile Thr Asp Pro Asp Val \

```
G7u cys Tyr Ser val Phe
        100
<210> 28
<211> 23
<212> DNA
<213> Artificial
<220>
<223> PCR Primer
<400> 28
gttcacttca gttacccatc tcg 23
<210> 29
<211> 23
<212> DNA
<213> Artificial
<220>
<223> PCT Primer
<400> 29
tatcctgtcc atcaacactg acc
<210> 30
\(<211>22\)
<212> DNA
<213> Artificial
```

<220>
<223> PCR Primer
<400> 30
aggctatagg ttccaggctt gc
<210> 31
<211> 23
<212> DNA
<213> Artificial
<220>
<223> PCT primer
<400> 31
agaacagtgt gaagacaatg gcc
<210> 32
<211> 24
<212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 32
atctcacagt aacttggcag tttc 24
<210> 33
<211> 22
<212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 33
aacccaaaca ggaagtgtgg cc
<210> 34
<211> }2
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 34
gtcagttcct atatcctgtg tctg 24
<210> 35
<211> 23
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 35
tcctgtccat caacactgac ctg

```
<210> 36
```

<211> 22
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 36
cccatcattg caatagcagg ag 22
<210> 37
<211> 23
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 37
gaacagtgtg aagacaatgg cct 23
<210> 38
<211> 23
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 38
gctggtcctg aagttgatct gtg 23
<210> 39
<211> 23
<212> DNA
<213> Artificial
<220>
<223> Sequencing primer
<400> 39
aaacaggaag tgtggccaga tgc


[^0]:    ${ }^{\text {a }}$ All patients are diagnosed with cutaneous T-cell lymphoma except for 12 patients in Group 2 who are diagnosed with various refractory cancers;
    ${ }^{\mathrm{b}}$ Data are presented as a median and range.

[^1]:    tctgggcggg ttcccaagta ttatgtctgt tccaggcttg $\underset{\text { Page }}{75} \operatorname{tgtttaaaa}$ aagcatatta 13260

