Title: HAPLOTYPES OF THE IL1B GENE

Abstract: Novel genetic variants of the Interleukin 1, Beta (IL1B) gene are described. Various genotypes, haplotypes, and haplotype pairs that exist in the general United States population are disclosed for the IL1B gene. Compositions and methods for haplotyping and/or genotyping the IL1B gene in an individual are also disclosed. Polynucleotides defined by the sequence of the haplotypes disclosed herein are also described.
HAPLOTYPES OF THE IL1B GENE

RELATED APPLICATIONS

This application claims the benefit of International Application No. PCT/US00/25794 filed September 20, 2000.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically-important proteins. In particular, this invention provides genetic variants of the human interleukin 1, beta (IL1B) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a lead compound that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by activity at non-intended targets. The lead compound identified in this screening process then undergoes further in vitro and in vivo testing to determine its absorption, disposition, metabolism and toxicological profiles. Typically, this testing involves use of cell lines and animal models with limited, if any, genetic diversity.

What this approach fails to consider, however, is that natural genetic variability exists between individuals in any and every population with respect to pharmaceutically-important proteins, including the protein targets of candidate drugs, the enzymes that metabolize these drugs and the proteins whose activity is modulated by such drug targets. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a pharmaceutically-important protein may be manifested as significant variation in expression, structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in the treatment of individuals with a drug whose design is based upon a single representative example of the target or enzyme(s) involved in metabolizing the drug. For example, it is well-established that some drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. Also, there is significant variation in how well people metabolize drugs and other exogenous chemicals, resulting in substantial interindivial variation in the toxicity and/or efficacy of such exogenous substances (Evans et al., 1999, Science 286:487-491). This variability in efficacy or toxicity of a drug in genetically-diverse patients makes many drugs ineffective or even dangerous in certain groups of the population, leading to the failure of such drugs in clinical trials or their early withdrawal from the market even though they could be highly beneficial for other groups in the
population. This problem significantly increases the time and cost of drug discovery and development, which is a matter of great public concern.


The standard for measuring genetic variation among individuals is the haplotype, which is the ordered combination of polymorphisms in the sequence of each form of a gene that exists in the population. Because haplotypes represent the variation across each form of a gene, they provide a more accurate and reliable measurement of genetic variation than individual polymorphisms. For example, while specific variations in gene sequences have been associated with a particular phenotype such as disease susceptibility (Roses AD supra; Ulbrecht M et al. 2000 Am J Respir Crit Care Med 161: 469-74) and drug response (Wolfe CR et al. 2000 BMJ 320:987-90; Dahl BS 1997 Acta Psychiatr Scand 96 (Suppl 391): 14-21), in many other cases an individual polymorphism may be found in a variety of genomic backgrounds, i.e., different haplotypes, and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 Am J Hum Genet 63:595-612; Ulbrecht M et al. 2000 supra; Drysdale et al. 2000 PNAS 97:10483-10488). Thus, there is an unmet need in the pharmaceutical industry for information on what haplotypes exist in the population for pharmaceutically-important genes. Such haplotype information would be useful in improving the efficiency and output of several steps in the drug discovery and development process, including target validation, identifying lead compounds, and early phase clinical trials (Marshall et al., supra).

One possible drug target for the treatment of inflammatory and immune disorders is the interleukin-1 beta (IL1B) gene or its encoded product. IL1B is a proinflammatory cytokine produced primarily by mononuclear phagocytes in response to infectious agents. IL1B is a major fever-inducing cytokine, which can lead to excessive loss of body weight and a negative nitrogen balance, as well as seizures in children. In combination with other cytokines, IL1B is believed to play a significant role in latter stages of the septic response and in propagating the inflammatory process that causes many of the pathologic and clinical manifestations of rheumatoid arthritis. IL1B may also be a causative factor in
postmenopausal osteoporosis and is believed to be involved in the pathogenesis of inflammatory bowel disease (IBD), insulin-dependent diabetes mellitus (IDDM), multiple sclerosis, and local osteolytic hypercalcemia (LOH), which is a condition suffered by cancer patients in which cancers have spread to the bone or bone marrow.

The gene encoding IL1B, which is also expressed by endothelial cells, B lymphocytes, natural killer cells, fibroblasts, smooth-muscle cells, keratinocytes and glial cells, is part of an IL-1 gene cluster located on the long arm of human chromosome 2 that includes genes encoding interleukin 1α (IL-1α), IL1B and the IL-1 receptor antagonist (IL-1Ra) (Loughrey et al., Cytokine 10:984-988, 1998). The interleukin-1, beta gene is located on chromosome 2q14 and contains 7 exons, the last six of which encode a precursor IL1B polypeptide of 269 amino acids, which is believed to be processed into a mature form consisting of the carboxyl-terminal 153 amino acids (Bensi et al., Gene 52:95-101, 1987; March et al., Nature 315:641-647, 1985). Reference sequences for the IL1B gene (GenBank Reference No. 4081185 or GenBank Accession No: M15840.1; SEQ ID NO: 1), coding sequence (GenBank Accession No:NM_000576.1), and protein are shown in Figures 1, 2 and 3, respectively.

Three single-nucleotide polymorphisms (SNPs) that affect expression or activity of the IL1B gene have been reported. One is a polymorphism of C or T located at -511 in the promoter region. This polymorphism corresponds to nucleotide 343 in Figure 1. The 343T allele has recently been shown to be associated with an increased risk of both hypochlorhydria induced by H. pylori and gastric cancer (El-Omar et al., Nature 2000 Mar 23;404(6776):398-402). H. pylori infection is associated with a variety of clinical outcomes including gastric cancer and duodenal ulcer disease. Individuals with gastritis predominantly localized to the antrum retain normal (or even high) acid secretion, whereas individuals with extensive corpus gastritis develop hypochlorhydria and gastric atrophy, which are presumptive precursors of gastric cancer. The association of the 343T allele with disease may be explained by the biological properties of IL1B, which is a powerful inhibitor of gastric acid secretion. Host genetic factors that affect IL1B may determine why some individuals infected with H. pylori develop gastric cancer while others do not (El-Omar et al., supra).

The second SNP, originally identified as a TaqI RFLP (Pociot et al., Eur. J. Clin. Invest. 22:396-402, 1992) and now referred to as the +3953 polymorphism, has been reported to constitute a T or C at codon 105 in exon 5 (Korman et al., J. Clin. Periodontol. 24:72-77, 1997; Loughrey et al., Cytokine 10:984-988, 1998; Santilla et al., Scand. J. Immunol. 47:195-198, 1998). This information would indicate to the skilled artisan that the location of the +3953 SNP corresponds to nucleotide 4336 in Fig. 1. Additionally, a SNP constituting a guanine or adenine variation at a position corresponding to nucleotide 4259 in Fig. 1 was originally identified as a BsoF I dimorphism (Gusach JF et al. 1996 Cytokine 8:598-602).

Several groups have reported that certain alleles at one or both of positions 343 and 4336 are associated with increased susceptibility to or increased severity of various inflammatory-related diseases, including inflammatory bowel disease (Nemetz et al., Immunogenetics 49:527-531, 1999); early-onset

IL1B haplotypes for the two SNPs at positions corresponding to nucleotides 4259 and 4336 of Fig. 1 were identified and characterized with respect to frequency of occurrence in 12 European Caucasians. Identifying which of these haplotypes is possessed by an individual has been suggested to have utility in predicting susceptibility to chronic obstructive pulmonary disorders such as asthma (Duff GW et al. U.S. Patent No. 6,140,047 issued Oct. 31, 2000). These two IL1B SNPs have also been included in haplotypes of the IL1 gene cluster, which additionally contain 3 SNPs in the IL1A gene, one SNP in the IL1R gene and two SNPs in the region between IL1B and IL1R. One of the IL1 gene cluster haplotypes, designated haplotype 44112332 and containing IL1B alleles of thymine at position 343 and cytosine at position 4336 of Fig. 1, was significantly overrepresented in a patient group with non-insulin dependent diabetes mellitus (NIDDM) relative to a healthy control group and was concluded to be associated with inflammatory disease (WO 98/54359).

In addition, expanded haplotypes for six SNPs spanning the IL1B gene were derived by Applicants for a population of 93 humans (WO 01/21639). The locations of these SNPs in Figure 1 are: 343, 346, 4259, 4336, 6421, and 6883.

Because of the potential for individual polymorphisms and haplotypes in the IL1B gene to affect susceptibility to a number of diseases, it would be useful to know whether additional polymorphisms exist in the IL1B gene, as well as how such polymorphisms are combined in different copies of the gene (haplotypes), and whether the frequencies of such haplotypes vary among different ethnic groups. Such information could be applied for studying the biological function of IL1B, as well as in identifying drugs targeting IL1B for the treatment of disorders related to its abnormal expression or function.

**SUMMARY OF THE INVENTION**

Accordingly, the inventors herein have discovered 3 novel polymorphic sites in the IL1B gene in addition to the three novel polymorphic sites discovered and reported previously by Applicants in WO 01/21639. These six polymorphic sites (PS) correspond to the following nucleotide positions in Figure 1: 346 (PS2), 3453 (PS3), 3467 (PS4), 3614 (PS5), 6421 (PS8) and 6883 (PS9). The polymorphisms at these sites are adenine or thymine at PS2, thymine or cytosine at PS3, guanine or cytosine at PS4, cytosine or thymine at PS5, adenine or guanine at PS8 and cytosine or adenine at PS9. In addition, the inventors have determined the identity of the alleles at these sites, as well as at the previously identified sites at nucleotide positions 343 (PS1), 4259 (PS6) and 4336 (PS7), in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African
descent, Asian, Caucasian and Hispanic/Latino. From this information, the inventors deduced a set of haplotypes and haplotype pairs for PS1-PS9 in the IL1B gene, which are shown below in Tables 4 and 3, respectively. Each of these IL1B haplotypes defines a naturally-occurring isoform (also referred to herein as an “isogene”) of the IL1B gene that exists in the human population. The frequency with which each haplotype and haplotype pair occurs within the total reference population and within each of the four major population groups included in the reference population was also determined.

Thus, in one embodiment, the invention provides a method, composition and kit for genotyping the IL1B gene in an individual. The genotyping method comprises identifying the nucleotide pair that is present at one or more polymorphic sites selected from the group consisting of PS2, PS3, PS4, PS5, PS8 and PS9 in both copies of the IL1B gene from the individual. A genotyping composition of the invention comprises an oligonucleotide probe or primer which is designed to specifically hybridize to a target region containing, or adjacent to, one of these novel IL1B polymorphic sites. A genotyping kit of the invention comprises a set of oligonucleotides designed to genotype each of these novel IL1B polymorphic sites. In a preferred embodiment, the genotyping kit comprises a set of oligonucleotides designed to genotype each of PS1-PS9. The genotyping method, composition, and kit are useful in determining whether an individual has one of the haplotypes in Table 4 below or has one of the haplotype pairs in Table 3 below.

The invention also provides a method for haplotyping the IL1B gene in an individual. In one embodiment, the haplotyping method comprises determining, for one copy of the IL1B gene, the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS2, PS3, PS4, PS5, PS8 and PS9. In another embodiment, the haplotyping method comprises determining whether one copy of the individual’s IL1B gene is defined by one of the IL1B haplotypes shown in Table 4, below, or a sub-haplotype thereof. In a preferred embodiment, the haplotyping method comprises determining whether both copies of the individual’s IL1B gene are defined by one of the IL1B haplotype pairs shown in Table 3, below, or a sub-haplotype pair thereof. The method for establishing the IL1B haplotype or haplotype pair of an individual is useful for improving the efficiency and reliability of several steps in the discovery and development of drugs for treating diseases associated with IL1B activity, e.g., inflammatory and immune disorders.

For example, the haplotyping method can be used by the pharmaceutical research scientist to validate IL1B as a candidate target for treating a specific condition or disease predicted to be associated with IL1B activity. Determining for a particular population the frequency of one or more of the individual IL1B haplotypes or haplotype pairs described herein will facilitate a decision on whether to pursue IL1B as a target for treating the specific disease of interest. In particular, if variable IL1B activity is associated with the disease, then one or more IL1B haplotypes or haplotype pairs will be found at a higher frequency in disease cohorts than in appropriately genetically matched controls. Conversely, if each of the observed IL1B haplotypes are of similar frequencies in the disease and control groups, then it may be inferred that variable IL1B activity has little, if any, involvement with that disease. In either case,
the pharmaceutical research scientist can, without a priori knowledge as to the phenotypic effect of any IL1B haplotype or haplotype pair, apply the information derived from detecting IL1B haplotypes in an individual to decide whether modulating IL1B activity would be useful in treating the disease.

The claimed invention is also useful in screening for compounds targeting IL1B to treat a specific condition or disease predicted to be associated with IL1B activity. For example, detecting which of the IL1B haplotypes or haplotype pairs disclosed herein are present in individual members of a population with the specific disease of interest enables the pharmaceutical scientist to screen for a compound(s) that displays the highest desired agonist or antagonist activity for each of the most frequent IL1B isoforms present in the disease population. Thus, without requiring any a priori knowledge of the phenotypic effect of any particular IL1B haplotype or haplotype pair, the claimed haplotyping method provides the scientist with a tool to identify lead compounds that are more likely to show efficacy in clinical trials.

The method for haplotyping the IL1B gene in an individual is also useful in the design of clinical trials of candidate drugs for treating a specific condition or disease predicted to be associated with IL1B activity. For example, instead of randomly assigning patients with the disease of interest to the treatment or control group as is typically done now, determining which of the IL1B haplotype(s) disclosed herein are present in individual patients enables the pharmaceutical scientist to distribute IL1B haplotypes and/or haplotype pairs evenly to treatment and control groups, thereby reducing the potential for bias in the results that could be introduced by a larger frequency of a IL1B haplotype or haplotype pair that had a previously unknown association with response to the drug being studied in the trial. Thus, by practicing the claimed invention, the scientist can more confidently rely on the information learned from the trial, without first determining the phenotypic effect of any IL1B haplotype or haplotype pair.

In another embodiment, the invention provides a method for identifying an association between a trait and a IL1B genotype, haplotype, or haplotype pair for one or more of the novel polymorphic sites described herein. The method comprises comparing the frequency of the IL1B genotype, haplotype, or haplotype pair in a population exhibiting the trait with the frequency of the IL1B genotype or haplotype in a reference population. A higher frequency of the IL1B genotype, haplotype, or haplotype pair in the trait population than in the reference population indicates the trait is associated with the IL1B genotype, haplotype, or haplotype pair. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. In a particularly preferred embodiment, the IL1B haplotype is selected from the haplotypes shown in Table 4, or a sub-haplotype thereof. Such methods have applicability in developing diagnostic tests and therapeutic treatments for inflammatory and immune disorders.

In yet another embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IL1B gene or a fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at PS2, cytosine at PS3, cytosine at PS4, thymine at PS5, guanine at PS8 and adenine at PS9. In a preferred embodiment,
the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of thymine at PS1, adenine at PS6 and thymine at PS7.

A particularly preferred polymorphic variant is an isogene of the IL1B gene. A IL1B isogene of the invention comprises cytosine or thymine at PS1, adenine or thymine at PS2, thymine or cytosine at PS3, guanine or cytosine at PS4, cytosine or thymine at PS5, guanine or adenine at PS6, cytosine or thymine at PS7, adenine or guanine at PS8 and cytosine or adenine at PS9. The invention also provides a collection of IL1B isogenes, referred to herein as a IL1B genome anthology.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for a IL1B cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig.2) and the polymorphic cDNA comprises thymine at a position corresponding to nucleotide 138. In a preferred embodiment, the polymorphic variant comprises an additional polymorphism of thymine at a position corresponding to nucleotide 315. A particularly preferred polymorphic cDNA variant comprises the coding sequence of a IL1B isogene defined by haplotypes 1-18.

Polynucleotides complementary to these IL1B genomic and cDNA variants are also provided by the invention. It is believed that polymorphic variants of the IL1B gene will be useful in studying the expression and function of IL1B, and in expressing IL1B protein for use in screening for candidate drugs to treat diseases related to IL1B activity.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express IL1B for protein structure analysis and drug binding studies.

The present invention also provides nonhuman transgenic animals comprising one of the IL1B polymorphic genomic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the IL1B isogenes in vivo, for in vivo screening and testing of drugs targeted against IL1B protein, and for testing the efficacy of therapeutic agents and compounds for inflammatory and immune disorders in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the IL1B gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the IL1B gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing IL1B haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the IL1B gene (Genenaghence Reference No. 4081185; contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence
indicated with a bracket ([ or ] ) and the numerical position below the sequence and the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence. SEQ ID NO:36 is equivalent to Figure 1, with the two alternative allelic variants of each polymorphic site indicated by the appropriate nucleotide symbol (R= G or A, Y= T or C, M= A or C, K= G or T, S= G or C, and W= A or T; WIPO standard ST.25).

Figure 2 illustrates a reference sequence for the IL1B coding sequence (contiguous lines; SEQ ID NO:2), with the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3 illustrates a reference sequence for the IL1B protein (contiguous lines; SEQ ID NO:3).

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the IL1B gene. As described in more detail below, the inventors herein discovered 18 isogenes of the IL1B gene by characterizing the IL1B gene found in genomic DNAs isolated from an Index Repository that contains immortalized cell lines from one chimpanzee and 93 human individuals. The human individuals included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (21 individuals), African descent (20 individuals), Asian (20 individuals), or Hispanic/Latino (18 individuals). To the extent possible, the members of this reference population were organized into population subgroups by their self-identified ethnogeographic origin as shown in Table 1 below.
Table 1. Population Groups in the Index Repository

<table>
<thead>
<tr>
<th>Population Group</th>
<th>Population Subgroup</th>
<th>No. of Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>African descent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sierra Leone</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Burma</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>China</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Korea</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Philippines</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Vietnam</td>
<td>4</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>British Isles</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>British Isles/Central</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>British Isles/Eastern</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Central/Eastern</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Eastern</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Central/Mediterranean</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Mediterranean</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Scandinavian</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caribbean</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Caribbean (Spanish Descent)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Central American (Spanish Descent)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Mexican American</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>South American (Spanish Descent)</td>
<td>3</td>
</tr>
</tbody>
</table>

In addition, the Index Repository contains three unrelated indigenous American Indians (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family.

The IL1B isogenes present in the human reference population are defined by haplotypes for 9 polymorphic sites in the IL1B gene, 6 of which are believed to be novel. The IL1B polymorphic sites identified by the inventors are referred to as PS1-PS9 to designate the order in which they are located in the gene (see Table 2 below), with the novel polymorphic sites referred to as PS2, PS3, PS4, PS5, PS8 and PS9. Using the genotypes identified in the Index Repository for PS1-PS9 and the methodology described in the Examples below, the inventors herein also determined the pair of haplotypes for the IL1B gene present in individual human members of this repository. The human genotypes and haplotypes found in the repository for the IL1B gene include those shown in Tables 3 and 4, respectively. The polymorphism and haplotype data disclosed herein are useful for validating whether IL1B is a suitable target for drugs to treat inflammatory and immune disorders, screening for such drugs and reducing bias in clinical trials of such drugs.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

**Allele** - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

**Candidate Gene** – A gene which is hypothesized to be responsible for a disease, condition, or the
response to a treatment, or to be correlated with one of these.

**Gene** - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

**Genotype** - An unphased 5’ to 3’ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

- **Full-genotype** - The unphased 5’ to 3’ sequence of nucleotide pairs found at all polymorphic sites examined herein in a locus on a pair of homologous chromosomes in a single individual.
- **Sub-genotype** - The unphased 5’ to 3’ sequence of nucleotides seen at a subset of the polymorphic sites examined herein in a locus on a pair of homologous chromosomes in a single individual.

**Genotyping** - A process for determining a genotype of an individual.

**Haplotype** - A 5’ to 3’ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

- **Full-haplotype** - The 5’ to 3’ sequence of nucleotides found at all polymorphic sites examined herein in a locus on a single chromosome from a single individual.
- **Sub-haplotype** - The 5’ to 3’ sequence of nucleotides seen at a subset of the polymorphic sites examined herein in a locus on a single chromosome from a single individual.

**Haplotype pair** - The two haplotypes found for a locus in a single individual.

**Haplotyping** - A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

**Haplotype data** - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

**Isoform** - A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

**Isogene** - One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

**Isolated** - As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

**Locus** - A location on a chromosome or DNA molecule corresponding to a gene or a physical or
phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment – A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single
copy of the locus is not known.

As discussed above, information on the identity of genotypes and haplotypes for the IL1B gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is expected to be useful for a variety of drug discovery and development applications. Thus, the invention also provides compositions and methods for detecting the novel IL1B polymorphisms and haplotypes identified herein.

The compositions comprise at least one IL1B genotyping oligonucleotide. In one embodiment, an IL1B genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term “oligonucleotide” refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of a IL1B polynucleotide, i.e., a IL1B isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-IL1B polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the IL1B gene using the polymorphism information provided herein in conjunction with the known sequence information for the IL1B gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a “perfect” or “complete” complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is “substantially complementary” to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions.
Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5’ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., “Genetic Prediction of Hemophilia A” in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruaño et al., 87 Proc. Natl. Acad. Sci. USA 6296-6300, 1990. Typically, an ASO will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotides of the invention include ASO probes and ASO primers. ASO probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15mer, the 8th or 9th position in a 16mer, and the 10th or 11th position in a 20mer). An ASO primer of the invention has a 3’ terminal nucleotide, or preferably a 3’ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. ASO probes and primers hybridizing to either the coding or noncoding strand are contemplated by the invention.

ASO probes and primers listed below use the appropriate nucleotide symbol (R= G or A, Y= T or C, M= A or C, K= G or T, S= G or C, and W= A or T; WIPO standard ST.25) at the position of the polymorphic site to represent the two alternative allelic variants observed at that polymorphic site.

A preferred ASO probe for detecting IL1B gene polymorphisms comprises a nucleotide sequence, listed 5’ to 3’, selected from the group consisting of:

AAGCCATWAAAACAG (SEQ ID NO:4) and its complement,
GGGATGAYCTTTAGA (SEQ ID NO:5) and its complement,
AAGCTCGGATTTCTT (SEQ ID NO:6) and its complement,
TGGATGGGAGCTCC (SEQ ID NO:7) and its complement,
TTCGACGAGAAGTT (SEQ ID NO:8) and its complement, and
GCGTGCTTTTTG (SEQ ID NO:9) and its complement.

A preferred ASO primer for detecting IL1B gene polymorphisms comprises a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

TTTGAAGCCATWA (SEQ ID NO:10); CCCTCGCTGGTTTTGW (SEQ ID NO:1611)
CAGAGGGGATGAYC (SEQ ID NO:12); CGAGCTTCTAAAGRT (SEQ ID NO:13);
CTTGAAGCGCTGCGS (SEQ ID NO:14); GCTTGAAAGAATCSC (SEQ ID NO:15);
GCCCTCGGATGGYG (SEQ ID NO:16); GTAGCTGATGCGRC (SEQ ID NO:17);
AATTAAGCCACACRG (SEQ ID NO:18); GGGGCAACTCTTCYG (SEQ ID NO:19);
GCTATAGCCTGGAMT (SEQ ID NO:20); and AGACAACAGGAAAKT (SEQ ID NO:21).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as “primer-extension oligonucleotides”. In a preferred embodiment, the 3'-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site.

A particularly preferred oligonucleotide primer for detecting IL1B gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

TGAAAGCCAT (SEQ ID NO:22); TCGCTGTGTTT (SEQ ID NO:23);
AGAGGGATGAG (SEQ ID NO:24); GCTTCTAAAG (SEQ ID NO:25);
TAGAGCTCCTG (SEQ ID NO:26); TGAAAGAATC (SEQ ID NO:27);
CTCTGTGATAG (SEQ ID NO:28); GCTTGATGCC (SEQ ID NO:29);
AAATTCCACAC (SEQ ID NO:30); CCCAACCTTTC (SEQ ID NO:31);
ATAGCCTGGAG (SEQ ID NO:32); and CAACAGGAAA (SEQ ID NO:33).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

IL1B genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized IL1B genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping
oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the IL1B gene in an individual. As used herein, the terms “IL1B genotype” and “IL1B haplotype” mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the IL1B gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid sample comprising the two copies of the IL1B gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more polymorphic sites selected from the group consisting of PS2, PS3, PS4, PS5, PS8 and PS9 in the two copies to assign a IL1B genotype to the individual. As will be readily understood by the skilled artisan, the two “copies” of a gene in an individual may be the same allele or may be different alleles. In a preferred embodiment of the genotyping method, the identity of the nucleotide pair at one or more of the polymorphic sites selected from the group consisting of PS1, PS6 and PS7 is also determined. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-PS9.

Typically, the nucleic acid sample is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid sample may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from a tissue in which the IL1B gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5’ and 3’ untranslated regions. If a IL1B gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid sample containing only one of the two copies of the IL1B gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS2, PS3, PS4, PS5, PS8 and PS9 in that copy to assign a IL1B haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the IL1B gene or fragment such as one of the methods described above for preparing IL1B isogenes, with targeted in vivo cloning being the preferred approach. As will be readily
appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two IL1B gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional IL1B clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the IL1B gene in an individual. In some embodiments, the haplotyping method also comprises identifying the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS6 and PS7. In a particularly preferred embodiment, the nucleotide at each of PS1-PS9 is identified.

In another embodiment, the haplotyping method comprises determining whether an individual has one or more of the IL1B haplotypes shown in Table 4. This can be accomplished by identifying, for one or both copies of the individual’s IL1B gene, the phased sequence of nucleotides present at each of PS1-PS9. The present invention also contemplates that typically only a subset of PS1-PS9 will need to be directly examined to assign to an individual one or more of the haplotypes shown in Table 4. This is because at least one polymorphic site in a gene is frequently in strong linkage disequilibrium with one or more other polymorphic sites in that gene (Drysdale, CM et al. 2000 PNAS 97:10483-10488; Rieder MJ et al. 1999 Nature Genetics 22:59-62). Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stephens, JC 1999, Mol. Diag. 4:309-317). Techniques for determining whether any two polymorphic sites are in linkage disequilibrium are well-known in the art (Weir B.S. 1996 Genetic Data Analysis II, Sinauer Associates, Inc. Publishers, Sunderland, MA).

In a preferred embodiment, an IL1B haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more polymorphic sites selected from the group consisting of PS2, PS3, PS4, PS5, PS8 and PS9 in each copy of the IL1B gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-PS9 in each copy of the IL1B gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the IL1B gene, or a fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled
artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., Proc. Natl. Acad. Sci. USA 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., Science 241:1077-1080, 1988).

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., Proc. Natl. Acad. Sci. USA 89:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the IL1B gene of an individual may also be determined by
hybridization of a nucleic acid sample containing one or both copies of the gene, or fragment(s) thereof, to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.


A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the “Genetic Bit Analysis” method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., *Nucl. Acids Res.* 17:8392, 1989; Ruaño et al., *Nucl. Acids Res.* 19, 6877-6882, 1991; WO 93/22456; Turki et al., *J. Clin. Invest.* 95:1635-1641, 1995).

In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping another polymorphic site that is in linkage disequilibrium with the polymorphic site that is of interest. Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

In another aspect of the invention, an individual's IL1B haplotype pair is predicted from its IL1B genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying a IL1B genotype for the individual at two or more IL1B polymorphic sites described herein, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing IL1B haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one
embodiment, the reference haplotype pairs include the IL1B haplotype pairs shown in Table 3.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African-descent, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by \(2n = \log(1-q)/\log(1-p)\) where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair \(H_1 / H_2\) is equal to

\[
p_{H_1H_2}(H_1 / H_2) = 2p(H_1)p(H_2) \quad \text{if} \quad H_1 \neq H_2 \quad \text{and} \quad p_{H_1H_2}(H_1 / H_2) = p(H_1)p(H_2) \quad \text{if} \quad H_1 = H_2.
\]

A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), single molecule dilution, or allele-specific long-range PCR (Michalotos-Beloïn et al., Nucleic Acids Res. 24:4841-4843, 1996).

In one embodiment of this method for predicting a IL1B haplotype pair for an individual, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct
molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., supra). A preferred process for predicting IL1B haplotype pairs from IL1B genotypes is described in U.S. Provisional Application Serial No. 60/198,340 and the corresponding International Application filed April 18, 2001.

The invention also provides a method for determining the frequency of a IL1B genotype, haplotype, or haplotype pair in a population. The method comprises, for each member of the population, determining the genotype or the haplotype pair for the novel IL1B polymorphic sites described herein, and calculating the frequency any particular genotype, haplotype, or haplotype pair is found in the population. The population may be a reference population, a family population, a same sex population, a population group, or a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for IL1B genotypes, haplotypes, and/or haplotype pairs are determined in a reference population and used in a method for identifying an association between a trait and a IL1B genotype, haplotype, or haplotype pair. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s), haplotype(s), or haplotype pair(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s), haplotype(s), or haplotype pair(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes, haplotypes, and/or haplotype pairs observed in the populations are compared. If a particular IL1B genotype, haplotype, or haplotype pair is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that IL1B genotype, haplotype or haplotype pair. Preferably, the IL1B genotype, haplotype, or haplotype pair being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 3 and 4, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting IL1B or response to a therapeutic treatment for a medical condition. As used herein, “medical condition” includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other
disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and a IL1B genotype, haplotype, or haplotype pair, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the IL1B gene for each individual in the trial population is genotyped and/or haplotype, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and IL1B genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their IL1B genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the IL1B gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in PCT Application Serial No.
PCT/US00/17540, entitled “Methods for Obtaining and Using Haplotype Data”.


Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the IL1B gene. As described in PCT Application Serial No. PCT/US00/17540, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, supra, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of IL1B genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the IL1B gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the IL1B gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying IL1B genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

In another embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the IL1B gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant IL1B gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS2, PS3, PS4, PS5, PS8 and PS9, and may also comprise one or more additional polymorphisms selected from the group consisting of thymine at PS1, adenine at PS6 and thymine at PS7. Similarly, the nucleotide sequence of a variant fragment of the IL1B gene is identical to the corresponding portion of
the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence of the IL1B gene (or other reported IL1B sequences) or to portions of the reference sequence (or other reported IL1B sequences), except for genotyping oligonucleotides as described above.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of thymine at PS2, cytosine at PS3, cytosine at PS4, thymine at PS5, guanine at PS8 and adenine at PS9. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the IL1B gene which is defined by any one of haplotypes 1-18 shown in Table 4 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the IL1B gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the novel polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing in vitro mutagenesis using procedures well-known in the art.

IL1B isogenes may be isolated using any method that allows separation of the two "copies" of the IL1B gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted in vivo cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad. Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 1989, supra; Ruaño et al., 1991, supra; Michalatos-Beloin et al., supra).

The invention also provides IL1B genome anthologies, which are collections of IL1B isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. A IL1B genome anthology may comprise individual IL1B isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like.

Alternatively, two or more groups of the IL1B isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred IL1B genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 4 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded IL1B protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but
are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant IL1B sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the IL1B gene will produce IL1B mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of a IL1B cDNA comprising a nucleotide sequence which is a polymorphic variant of the IL1B reference coding sequence shown in Figure 2. Thus, the invention also provides IL1B mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2), or its corresponding RNA sequence, except for having thymine at a position corresponding to nucleotide 138, and may also comprise an additional polymorphism of thymine at a position corresponding to nucleotide 315. A particularly preferred polymorphic cDNA variant comprises the coding sequence of a IL1B isogene defined by haplotypes 1-18. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphism described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized IL1B cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.
As used herein, a polymorphic variant of a IL1B gene fragment comprises at least one novel polymorphism identified herein and has a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, such fragments are between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the IL1B polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the IL1B gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the IL1B genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular IL1B protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the IL1B isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular IL1B isogene. Expression of a IL1B isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco; N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of IL1B mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of IL1B mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue in vivo or ex vivo. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which
are not as easily recognized by endogenous nucleases.

Effect(s) of the polymorphisms identified herein on expression of IL1B may be investigated by preparing recombinant cells and/or nonhuman recombinant organisms, preferably recombinant animals, containing a polymorphic variant of the IL1B gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into IL1B protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired IL1B isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the IL1B isogene is introduced into a cell in such a way that it recombines with the endogenous IL1B gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired IL1B gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the IL1B isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the IL1B isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant nonhuman organisms, i.e., transgenic animals, expressing a variant IL1B gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053.
Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the IL1B isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human IL1B isogene and producing human IL1B protein can be used as biological models for studying diseases related to abnormal IL1B expression and/or activity, and for
screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel IL1B isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel IL1B isogenes; an antisense oligonucleotide directed against one of the novel IL1B isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel IL1B isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel IL1B isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington’s Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the IL1B gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The IL1B polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle
database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD-ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

EXAMPLE 1

This example illustrates examination of various regions of the IL1B gene for polymorphic sites.

Amplification of Target Regions

The following target regions were amplified using either the PCR primers represented below or 'tailed' PCR primers, each of which includes a universal sequence forming a noncomplementary 'tail' attached to the 5' end of each unique sequence in the PCR primer pairs. The universal 'tail' sequence for the forward PCR primers comprises the sequence 5'-TGTTAAACGACGGCCAGT-3' (SEQ ID NO:34) and the universal 'tail' sequence for the reverse PCR primers comprises the sequence 5'-AGGAAACAGCTATGACC-3' (SEQ ID NO:35). The nucleotide positions of the first and last nucleotide of the forward and reverse primers for each region amplified are presented below and correspond to positions in Figure 1.

<table>
<thead>
<tr>
<th>Fragment No.</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>PCR Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment 1</td>
<td>628-649</td>
<td>complement of 1144-1122</td>
<td>517 nt</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>4148-4171</td>
<td>complement of 4713-4690</td>
<td>566 nt</td>
</tr>
<tr>
<td>Fragment 3</td>
<td>8-28</td>
<td>complement of 577-557</td>
<td>570 nt</td>
</tr>
<tr>
<td>Fragment 4</td>
<td>97-118</td>
<td>complement of 749-727</td>
<td>653 nt</td>
</tr>
<tr>
<td>Fragment 5</td>
<td>6343-6368</td>
<td>complement of 6929-6907</td>
<td>587 nt</td>
</tr>
<tr>
<td>Fragment 6</td>
<td>3397-3418</td>
<td>complement of 3916-3894</td>
<td>520 nt</td>
</tr>
</tbody>
</table>
These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

- **Reaction volume** = 10 µl
- 10 x Advantage 2 Polymerase reaction buffer (Clontech) = 1 µl
- 100 ng of human genomic DNA = 1 µl
- 10 mM dNTP = 0.4 µl
- Advantage 2 Polymerase enzyme mix (Clontech) = 0.2 µl
- Forward Primer (10 µM) = 0.4 µl
- Reverse Primer (10 µM) = 0.4 µl
- Water = 6.6 µl

**Amplification profile:**
- 97°C - 2 min. 1 cycle
- 97°C - 15 sec.
- 70°C - 45 sec. 10 cycles
- 72°C - 45 sec.
- 97°C - 15 sec.
- 64°C - 45 sec. 35 cycles
- 72°C - 45 sec.

**Sequencing of PCR Products**

The PCR products were purified using a Whatman/Polyfiltronics 100 µl 384 well unifilter plate essentially according to the manufacturers protocol. The purified DNA was eluted in 50 µl of distilled water. Sequencing reactions were set up using Applied Biosystems Big Dye Terminator chemistry essentially according to the manufacturers protocol. The purified PCR products were sequenced in both directions using either the primer sets represented below with the positions of their first and last nucleotide corresponding to positions in Figure 1, or the appropriate universal 'tail' sequence as a primer.

Reaction products were purified by isopropanol precipitation, and run on an Applied Biosystems 3700 DNA Analyzer.

**Sequencing Primer Pairs**

<table>
<thead>
<tr>
<th>Fragment No.</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment 1</td>
<td>693-712</td>
<td>complement of 1092-1073</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>4172-4191</td>
<td>complement of 4665-4646</td>
</tr>
<tr>
<td>Fragment 3</td>
<td>27-47</td>
<td>complement of 503-492</td>
</tr>
<tr>
<td>Fragment 4</td>
<td>188-207</td>
<td>complement of 707-688</td>
</tr>
<tr>
<td>Fragment 5: Tailed Seq.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fragment 6: Tailed Seq.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Analysis of Sequences for Polymorphic Sites**

Sequence information for a minimum of 80 humans was analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., *Nucleic Acids Res.* 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations...
in the IL1B gene are listed in Table 2 below.

<table>
<thead>
<tr>
<th>Polymorphic Site Number</th>
<th>PolyIda</th>
<th>Nucleotide Reference</th>
<th>Variant CDS Variant</th>
<th>AA Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS1^a</td>
<td>5077411</td>
<td>343</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>PS2</td>
<td>5077413</td>
<td>346</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>PS3</td>
<td>9117201</td>
<td>3453</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>PS4</td>
<td>9117106</td>
<td>3467</td>
<td>G</td>
<td>C</td>
</tr>
<tr>
<td>PS5</td>
<td>9116916</td>
<td>3614</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>PS6^a</td>
<td>5077421</td>
<td>4259</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>PS7^b</td>
<td>5077423</td>
<td>4336</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>PS8</td>
<td>5077427</td>
<td>6421</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>PS9</td>
<td>5077439</td>
<td>6883</td>
<td>C</td>
<td>A</td>
</tr>
</tbody>
</table>

aPolyId is a unique identifier assigned to each PS by Genaissance Pharmaceuticals, Inc.

bReported previously in the literature.

EXAMPLE 2

This example illustrates analysis of the IL1B polymorphisms identified in the Index Repository for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 3 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 3, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 3 were inferred based on linkage disequilibrium and/or Mendelian inheritance.
Table 3. Genotypes and Haplotype Pairs Observed for IL1B Gene

<table>
<thead>
<tr>
<th>Genotype Number</th>
<th>PS1</th>
<th>PS2</th>
<th>PS3</th>
<th>PS4</th>
<th>PS5</th>
<th>PS6</th>
<th>PS7</th>
<th>PS8</th>
<th>PS9</th>
<th>HAP Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G</td>
<td>C/A</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G/C</td>
<td>C</td>
<td>G</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>C</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>T</td>
<td>A</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A/G</td>
<td>C</td>
<td>A/G</td>
<td>C/A</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>C/T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>C</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>C/T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>C/T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>C/T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>18</td>
<td>C/T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>C/T</td>
<td>A/T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>A/G</td>
<td>C/T</td>
<td>A/G</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>22</td>
<td>T/C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C/T</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>23</td>
<td>T/C</td>
<td>A</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>24</td>
<td>T/C</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>25</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A/G</td>
<td>C</td>
<td>A/G</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>26</td>
<td>T</td>
<td>A</td>
<td>T/C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>14</td>
</tr>
<tr>
<td>27</td>
<td>T/C</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A/G</td>
<td>C</td>
<td>A/G</td>
<td>C</td>
<td>10</td>
</tr>
<tr>
<td>28</td>
<td>C/T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>G/A</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>29</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C/T</td>
<td>G</td>
<td>C</td>
<td>7</td>
</tr>
</tbody>
</table>

The haplotype pairs shown in Table 3 were estimated from the unphased genotypes using a computer-implemented extension of Clark's algorithm (Clark, A.G. 1990 Mol Bio Evol 7, 111-122) for assigning haplotypes to unrelated individuals in a population sample, as described in U.S. Provisional Application Serial No. 60/198,340 entitled “A Method and System for Determining Haplotypes from a Collection of Polymorphisms” and the corresponding International Application filed April 18, 2001. In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is augmented with haplotypes obtained from two families (one three-generation Caucasian family and one two-generation African-American family) and then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 18 human IL1B haplotypes shown in Table 4 below.
Table 4. Haplotypes Identified in the IL1B Gene

<table>
<thead>
<tr>
<th>HAP No.</th>
<th>HAP ID</th>
<th>Polymorphic Sites</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>91200012</td>
<td></td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>2</td>
<td>91200008</td>
<td></td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>3</td>
<td>9120016</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
</tr>
<tr>
<td>4</td>
<td>9120010</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>5</td>
<td>9120014</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>6</td>
<td>9120013</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>7</td>
<td>9120002</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>8</td>
<td>9120009</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
</tr>
<tr>
<td>9</td>
<td>9120015</td>
<td></td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>10</td>
<td>9120004</td>
<td></td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>11</td>
<td>9120019</td>
<td></td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>12</td>
<td>9120011</td>
<td></td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>13</td>
<td>9120005</td>
<td></td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
</tr>
<tr>
<td>14</td>
<td>9120003</td>
<td></td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>15</td>
<td>9120017</td>
<td></td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>16</td>
<td>9120006</td>
<td></td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>C</td>
</tr>
<tr>
<td>17</td>
<td>9120007</td>
<td></td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
</tr>
<tr>
<td>18</td>
<td>9120018</td>
<td></td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
</tbody>
</table>

Table 5 below shows the percent of chromosomes characterized by a given IL1B haplotype for all unrelated individuals in the Index Repository for which haplotype data was obtained. The percent of these unrelated individuals who have a given IL1B haplotype pair is shown in Table 6. In Tables 5 and 6, the "Total" column shows this frequency data for all of these unrelated individuals, while the other columns show the frequency data for these unrelated individuals categorized according to their self-identified ethnogeographic origin. Abbreviations used in Tables 5 and 6 are AF = African Descent, AS = Asian, CA = Caucasian, HL = Hispanic-Latino, and NA = Native American.
Table 5. Frequency of Observed IL1B Haplotypes In Unrelated Individuals

<table>
<thead>
<tr>
<th>HAP No.</th>
<th>HAP ID</th>
<th>Total</th>
<th>CA</th>
<th>AF</th>
<th>AS</th>
<th>HL</th>
<th>AM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9120012</td>
<td>1.22</td>
<td>0.0</td>
<td>2.5</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>9120008</td>
<td>1.83</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>9120016</td>
<td>0.61</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>9120010</td>
<td>1.83</td>
<td>2.38</td>
<td>2.5</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>9120014</td>
<td>0.61</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>6</td>
<td>9120013</td>
<td>0.61</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>9120002</td>
<td>33.54</td>
<td>21.43</td>
<td>50.0</td>
<td>27.5</td>
<td>36.11</td>
<td>33.33</td>
</tr>
<tr>
<td>8</td>
<td>9120009</td>
<td>1.83</td>
<td>0.0</td>
<td>7.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>9</td>
<td>9120015</td>
<td>0.61</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>9120004</td>
<td>10.37</td>
<td>2.38</td>
<td>5.0</td>
<td>30.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>11</td>
<td>9120019</td>
<td>0.61</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>12</td>
<td>9120011</td>
<td>1.83</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>13</td>
<td>9120005</td>
<td>10.98</td>
<td>16.67</td>
<td>5.0</td>
<td>0.0</td>
<td>22.22</td>
<td>16.67</td>
</tr>
<tr>
<td>14</td>
<td>9120003</td>
<td>21.34</td>
<td>40.48</td>
<td>7.5</td>
<td>20.0</td>
<td>13.39</td>
<td>33.33</td>
</tr>
<tr>
<td>15</td>
<td>9120017</td>
<td>0.61</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>16</td>
<td>9120006</td>
<td>8.54</td>
<td>16.67</td>
<td>5.0</td>
<td>5.0</td>
<td>8.33</td>
<td>0.0</td>
</tr>
<tr>
<td>17</td>
<td>9120007</td>
<td>2.44</td>
<td>0.0</td>
<td>2.5</td>
<td>2.5</td>
<td>2.78</td>
<td>16.67</td>
</tr>
<tr>
<td>18</td>
<td>9120018</td>
<td>0.61</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 6. Frequency of Observed IL1B Haplotype Pairs In Unrelated Individuals

<table>
<thead>
<tr>
<th>HAP1</th>
<th>HAP2</th>
<th>Total</th>
<th>CA</th>
<th>AF</th>
<th>AS</th>
<th>HL</th>
<th>AM</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7</td>
<td>9.76</td>
<td>4.76</td>
<td>10.0</td>
<td>10.0</td>
<td>16.67</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>4.88</td>
<td>0.0</td>
<td>0.0</td>
<td>20.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>4.88</td>
<td>14.29</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>1.22</td>
<td>4.76</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>2.44</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>11.11</td>
<td>0.0</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>2.44</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>13</td>
<td>16</td>
<td>2.44</td>
<td>4.76</td>
<td>0.0</td>
<td>0.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>1.22</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>6.1</td>
<td>4.76</td>
<td>10.0</td>
<td>5.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>4.88</td>
<td>9.52</td>
<td>5.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>6.1</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
<td>16.67</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>4.88</td>
<td>0.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.56</td>
<td>33.33</td>
</tr>
<tr>
<td>7</td>
<td>18</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>2.44</td>
<td>4.76</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>8.54</td>
<td>19.05</td>
<td>0.0</td>
<td>0.0</td>
<td>11.11</td>
<td>33.33</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>8.54</td>
<td>19.05</td>
<td>5.0</td>
<td>0.0</td>
<td>11.11</td>
<td>0.0</td>
</tr>
<tr>
<td>14</td>
<td>10</td>
<td>2.44</td>
<td>0.0</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>1.22</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>5.56</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>9.76</td>
<td>14.29</td>
<td>5.0</td>
<td>10.0</td>
<td>5.56</td>
<td>33.33</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>3.66</td>
<td>0.0</td>
<td>15.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
The size and composition of the Index Repository were chosen to represent the genetic diversity across and within four major population groups comprising the general United States population. For example, as described in Table 1 above, this repository contains approximately equal sample sizes of African-descent, Asian-American, European-American, and Hispanic-Latino population groups. Almost all individuals representing each group had all four grandparents with the same ethnogeographic background. The number of unrelated individuals in the Index Repository provides a sample size that is sufficient to detect SNPs and haplotypes that occur in the general population with high statistical certainty. For instance, a haplotype that occurs with a frequency of 5% in the general population has a probability higher than 99.9% of being observed in a sample of 80 individuals from the general population. Similarly, a haplotype that occurs with a frequency of 10% in a specific population group has a 99% probability of being observed in a sample of 20 individuals from that population group. In addition, the size and composition of the Index Repository means that the relative frequencies determined therein for the haplotypes and haplotype pairs of the IL1B gene are likely to be similar to the relative frequencies of these IL1B haplotypes and haplotype pairs in the general U.S. population and in the four population groups represented in the Index Repository. The genetic diversity observed for the three Native Americans is presented because it is of scientific interest, but due to the small sample size it lacks statistical significance.

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.
What is Claimed is:

1. A method for haplotyping the interleukin 1, beta (IL1B) gene of an individual, which comprises determining which of the IL1B haplotypes shown in the table immediately below defines one copy of the individual’s IL1B gene, wherein each of the IL1B haplotypes comprises a set of polymorphisms whose locations and identities are set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Haplotype Number</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17</td>
<td>1 2</td>
<td>343</td>
<td>36</td>
<td>27-1092</td>
</tr>
<tr>
<td>C C C C C C T C T T T T T T</td>
<td>A A A A A A A A A A A A T</td>
<td></td>
<td>36</td>
<td>27-1092</td>
</tr>
<tr>
<td>C T T T T T T C C C C T T T T T</td>
<td></td>
<td>3453</td>
<td>36</td>
<td>3397-3916</td>
</tr>
<tr>
<td>G G G G G G G G G G G G G G</td>
<td></td>
<td>3467</td>
<td>36</td>
<td>3397-3917</td>
</tr>
<tr>
<td>C C C C C C T C C C C C C</td>
<td></td>
<td>3614</td>
<td>36</td>
<td>3397-3918</td>
</tr>
<tr>
<td>A G A A G G A A A A A A A</td>
<td></td>
<td>4259</td>
<td>36</td>
<td>4172-4665</td>
</tr>
<tr>
<td>C T C C C C C C C C C C C</td>
<td></td>
<td>4336</td>
<td>36</td>
<td>4172-4666</td>
</tr>
<tr>
<td>G G G G G G G G G G G G G</td>
<td></td>
<td>6421</td>
<td>36</td>
<td>6343-6929</td>
</tr>
<tr>
<td>C C C C A C C C C C C A C C C</td>
<td></td>
<td>6883</td>
<td>36</td>
<td>6343-6930</td>
</tr>
</tbody>
</table>

*a: Alleles for haplotypes are presented 5’ to 3’ in each column
*PS = polymorphic site;
*Position of PS within the indicated SEQ ID NO;
*Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO. of the sequenced region.

2. The method of claim 1, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS9 on the one copy of the individual’s IL1B gene.

3. A method for haplotyping the interleukin 1, beta (IL1B) gene of an individual, which comprises determining which of the IL1B haplotype pairs shown in the table immediately below defines both copies of the individual’s IL1B gene, wherein each of the IL1B haplotype pairs consists of first and second haplotypes which comprise first and second sets of polymorphisms whose locations and identities are set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Haplotype Pairs</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/3 13/12 16/11 17/2 13/16 14/15 17/12 7/6 7/10</td>
<td>1</td>
<td>343</td>
<td>36</td>
<td>27-1092</td>
</tr>
<tr>
<td>C C T T T T C C T C C C C T</td>
<td>A A A A A A A A A A A A</td>
<td></td>
<td>36</td>
<td>27-1092</td>
</tr>
<tr>
<td></td>
<td>T T C C T C T C T T C T C T</td>
<td></td>
<td>3453</td>
<td>36</td>
</tr>
<tr>
<td>G G G G G G G G G G G G G</td>
<td></td>
<td>3467</td>
<td>36</td>
<td>3397-3917</td>
</tr>
<tr>
<td>C C C C C C C C C C C</td>
<td></td>
<td>3614</td>
<td>36</td>
<td>3397-3918</td>
</tr>
<tr>
<td>G G G G G A G G G G G G A</td>
<td></td>
<td>4259</td>
<td>36</td>
<td>4172-4665</td>
</tr>
<tr>
<td>C T C C C C C C C C C C</td>
<td></td>
<td>4336</td>
<td>36</td>
<td>4172-4666</td>
</tr>
<tr>
<td>G G G G G G G G G G G G</td>
<td></td>
<td>6421</td>
<td>36</td>
<td>6343-6929</td>
</tr>
<tr>
<td>C C C C C C C C C C C</td>
<td></td>
<td>6883</td>
<td>36</td>
<td>6343-6930</td>
</tr>
<tr>
<td>Haplotype Pairs</td>
<td>PS&lt;sup&gt;b&lt;/sup&gt; No.</td>
<td>PS Position&lt;sup&gt;c&lt;/sup&gt;</td>
<td>SEQ ID NO.</td>
<td>Region Examined&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>7/7</td>
<td>C</td>
<td>1</td>
<td>343</td>
<td>3627-1092</td>
</tr>
<tr>
<td>10/10</td>
<td>A</td>
<td>2</td>
<td>346</td>
<td>3627-1092</td>
</tr>
<tr>
<td>14/14</td>
<td>T</td>
<td>3</td>
<td>3453</td>
<td>363397-3916</td>
</tr>
<tr>
<td>13/13</td>
<td>G</td>
<td>4</td>
<td>3467</td>
<td>363397-3917</td>
</tr>
<tr>
<td>7/16</td>
<td>C</td>
<td>5</td>
<td>3614</td>
<td>363397-3918</td>
</tr>
<tr>
<td>7/17</td>
<td>G</td>
<td>6</td>
<td>4259</td>
<td>364172-4665</td>
</tr>
<tr>
<td>7/18</td>
<td>C</td>
<td>7</td>
<td>4336</td>
<td>364172-4666</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>6421</td>
<td>366343-6929</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>6883</td>
<td>366343-6930</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Haplotype Pairs</th>
<th>PS&lt;sup&gt;b&lt;/sup&gt; No.</th>
<th>PS Position&lt;sup&gt;c&lt;/sup&gt;</th>
<th>SEQ ID NO.</th>
<th>Region Examined&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/4</td>
<td>C</td>
<td>1</td>
<td>343</td>
<td>3627-1092</td>
</tr>
<tr>
<td>14/13</td>
<td>A</td>
<td>2</td>
<td>346</td>
<td>3627-1092</td>
</tr>
<tr>
<td>14/9</td>
<td>T</td>
<td>3</td>
<td>3453</td>
<td>363397-3916</td>
</tr>
<tr>
<td>10/4</td>
<td>G</td>
<td>4</td>
<td>3467</td>
<td>363397-3917</td>
</tr>
<tr>
<td>14/1</td>
<td>C</td>
<td>5</td>
<td>3614</td>
<td>363397-3918</td>
</tr>
<tr>
<td>14/16</td>
<td>G</td>
<td>6</td>
<td>4259</td>
<td>364172-4665</td>
</tr>
<tr>
<td>14/10</td>
<td>C</td>
<td>7</td>
<td>4336</td>
<td>364172-4666</td>
</tr>
<tr>
<td>10/2</td>
<td>G</td>
<td>8</td>
<td>6421</td>
<td>366343-6929</td>
</tr>
<tr>
<td>7/14</td>
<td>C</td>
<td>9</td>
<td>6883</td>
<td>366343-6930</td>
</tr>
</tbody>
</table>

*Haplotype pairs are represented as 1<sup>st</sup> haplotype/2<sup>nd</sup> haplotype; with alleles of each haplotype shown 5' to 3' as 1<sup>st</sup> polymorphism/2<sup>nd</sup> polymorphism in each column;

<sup>b</sup>PS = polymorphic site;

<sup>c</sup>Position of PS within the indicated SEQ ID NO;

<sup>d</sup>Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO of the sequenced region.

4. The method of claim 3, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS9 on both copies of the individual's IL1B gene.

5. A method for genotyping the interleukin 1, beta (IL1B) gene of an individual, comprising determining for the two copies of the IL1B gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS) selected from the group consisting of PS3, PS4, and PS5, wherein the one or more PS have the location and alternative alleles shown in SEQ ID NO: 36.

6. The method of claim 5, wherein the determining step comprises:
   (a) isolating from the individual a nucleic acid mixture comprising both copies of the IL1B gene, or a fragment thereof, that are present in the individual;
   (b) amplifying from the nucleic acid mixture a target region containing the selected polymorphic site;
   (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
   (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the
reaction, wherein said terminators are complementary to the alternative nucleotides present at the selected polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

7. The method of claim 5, further comprising determining the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS2, PS8 and PS9.

8. The method of claim 5, which comprises determining for the two copies of the IL1B gene present in the individual the identity of the nucleotide pair at each of PS1-PS9.

9. A method for haplotyping the interleukin 1, beta (IL1B) gene of an individual which comprises determining, for one copy of the IL1B gene present in the individual, the identity of the nucleotide at two or more polymorphic sites (PS) selected from the group consisting of PS3, PS4, and PS5, wherein the selected PS have the location and alternative alleles shown in SEQ ID NO:36.

10. The method of claim 9, further comprising determining the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS2, PS8 and PS9, wherein the selected PS have the locations and alternative alleles shown in SEQ ID NO:36.

11. The method of claim 9, further comprising determining the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS6 and PS7, wherein the selected PS have the locations and alternative alleles shown in SEQ ID NO:36.

12. The method of claim 9, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid sample containing only one of the two copies of the IL1B gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing the selected polymorphic site;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the selected polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

13. A method for predicting a haplotype pair for the interleukin 1, beta (IL1B) gene of an individual comprising:

(a) identifying a IL1B genotype for the individual, wherein the genotype comprises the nucleotide pair at two or more polymorphic sites (PS) selected from the group consisting of PS3, PS4, and PS5, wherein the selected PS have the location and alternative alleles shown in SEQ ID NO:36;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;
(c) comparing the possible haplotype pairs to the haplotype pair data set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Haplotype Pairs</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/3: 13/12 16/11 7/2 13/16 14/15 7/12 7/1 7/6 7/10</td>
<td>C C T T T T T C T C C C C C T</td>
<td>1</td>
<td>343</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>A A A A A A A A A A A A A A A A A A A</td>
<td>2</td>
<td>346</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>T T C C T C T T T T T C T T T C T</td>
<td>3</td>
<td>3453</td>
<td>36-3397-3916</td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C C C C C C C C</td>
<td>5</td>
<td>3614</td>
<td>36-3397-3918</td>
</tr>
<tr>
<td></td>
<td>C T T C C C C C T C C C C C C C C C C</td>
<td>7</td>
<td>4336</td>
<td>36-4172-4666</td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C C C</td>
<td>9</td>
<td>6883</td>
<td>36-6343-6930</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Haplotype Pairs</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/7: 10/10 14/14 13/13 7/5 7/6 7/13 7/17 7/18</td>
<td>C C T T T T T T C C C C C T C T T C T T</td>
<td>1</td>
<td>343</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>A A A A A A A A A A A A A A A A A A A</td>
<td>2</td>
<td>346</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>T T C C T C T T T T T C T T T C T</td>
<td>3</td>
<td>3453</td>
<td>36-3397-3916</td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C C C C C C C</td>
<td>5</td>
<td>3614</td>
<td>36-3397-3918</td>
</tr>
<tr>
<td></td>
<td>G G A A A A G G G G G G G G G G G G G A</td>
<td>6</td>
<td>4259</td>
<td>36-4172-4665</td>
</tr>
<tr>
<td></td>
<td>C C C C C C T C T C C C C C C C C C</td>
<td>7</td>
<td>4336</td>
<td>36-4172-4666</td>
</tr>
<tr>
<td></td>
<td>G G A A A A G G G G G G G G G G G G G A</td>
<td>8</td>
<td>6421</td>
<td>36-6343-6929</td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C C</td>
<td>9</td>
<td>6883</td>
<td>36-6343-6930</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Haplotype Pairs</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/4 14/13 14/9 10/4 14/1 14/16 14/10 10/2 7/14 7/8</td>
<td>C C T T T T T T T T T T C C C C C C C</td>
<td>1</td>
<td>343</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>A A A A A A A A A A A A A A A A A A A</td>
<td>2</td>
<td>346</td>
<td>36-27-1092</td>
</tr>
<tr>
<td></td>
<td>T T T C T T T T T T T C C C C C C C T T</td>
<td>3</td>
<td>3453</td>
<td>36-3397-3916</td>
</tr>
<tr>
<td></td>
<td>C C C C C C T C C C C C C C C C C C C</td>
<td>5</td>
<td>3614</td>
<td>36-3397-3918</td>
</tr>
<tr>
<td></td>
<td>G A A G A A A A A A A A A A A A A A A A A</td>
<td>6</td>
<td>4259</td>
<td>36-4172-4665</td>
</tr>
<tr>
<td></td>
<td>C C T C C C C C C C C C C C C C C C C C T</td>
<td>7</td>
<td>4336</td>
<td>36-4172-4666</td>
</tr>
<tr>
<td></td>
<td>G A A G A A A A A A A A A A A A A A A A</td>
<td>8</td>
<td>6421</td>
<td>36-6343-6929</td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C C</td>
<td>9</td>
<td>6883</td>
<td>36-6343-6930</td>
</tr>
</tbody>
</table>

*a Haplotype pairs are represented as 1st haplotype/2nd haplotype; with alleles of each haplotype shown 5' to 3' as 1st polymorphism/2nd polymorphism in each column;*  
*PS = polymorphic site;*  
*Position of PS within the indicated SEQ ID NO;*  
*Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO of the sequenced region;*  

and  

(d) assigning a haplotype pair to the individual that is consistent with the data.  

14. The method of claim 13, further comprising determining the identity of the nucleotide at one or
more polymorphic sites selected from the group consisting of PS2, PS8 and PS9.

15. The method of claim 13, wherein the identified genotype of the individual comprises the nucleotide pair at each of PS1-PS9, which have the location and alternative alleles shown in SEQ ID NO:36.

16. A method for identifying an association between a trait and at least one haplotype or haplotype pair of the interleukin 1, beta (IL1B) gene which comprises comparing the frequency of the haplotype or haplotype pair in a population exhibiting the trait with the frequency of the haplotype or haplotype pair in a reference population, wherein the haplotype is selected from haplotypes 1-18 shown in the table presented immediately below, wherein each of the haplotypes comprises a set of polymorphisms whose locations and identities are set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Haplotype Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PS&lt;sup&gt;b&lt;/sup&gt; No.</th>
<th>PS Position&lt;sup&gt;c&lt;/sup&gt;</th>
<th>SEQ ID NO.</th>
<th>Region Examined&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 C C C C C C T T T T T T T</td>
<td>1 343</td>
<td>36 27-1092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 A A A A A A A A A A A A</td>
<td>2 346</td>
<td>36 27-1092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 C T T T T T T C C C T T T T</td>
<td>3 3453</td>
<td>36 3397-3916</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 G C C C G G G G G G G G G</td>
<td>4 3467</td>
<td>36 3397-3917</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 C C C C C C C C C C C</td>
<td>5 3614</td>
<td>36 3397-3918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 A G G A A A G G A A G G A</td>
<td>6 4259</td>
<td>36 4172-4665</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 C T C C C C C C C C C C C</td>
<td>7 4336</td>
<td>36 4172-4666</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 G G G G G G G G A G G G A</td>
<td>8 6421</td>
<td>36 6343-6929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 C C C A C C C C C C A C C</td>
<td>9 6883</td>
<td>36 6343-6930</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Alleles for haplotypes are presented 5' to 3' in each column
<sup>b</sup> PS = polymorphic site;
<sup>c</sup> Position of PS within the indicated SEQ ID NO;
<sup>d</sup> Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO. of the sequenced region,

and wherein the haplotype pair is selected from the haplotype pairs shown in the table immediately below, wherein each of the IL1B haplotype pairs consists of first and second haplotypes which comprise first and second sets of polymorphisms whose locations and identities are set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Haplotype Pairs&lt;sup&gt;e&lt;/sup&gt;</th>
<th>PS&lt;sup&gt;f&lt;/sup&gt; No.</th>
<th>PS Position&lt;sup&gt;g&lt;/sup&gt;</th>
<th>SEQ ID NO.</th>
<th>Region Examined&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/3</td>
<td>C T T T T T C C C C T T T T C T C C C C C C C C C C C C</td>
<td>1 343</td>
<td>36 27-1092</td>
<td></td>
</tr>
<tr>
<td>16/11</td>
<td>A A A A A A A A A A A A A A A A A A A A A A</td>
<td>2 346</td>
<td>36 27-1092</td>
<td></td>
</tr>
<tr>
<td>13/12</td>
<td>T T C C T C T C C T T A A A A A A A A A A A A A A A A A A</td>
<td>3 3453</td>
<td>36 3397-3916</td>
<td></td>
</tr>
<tr>
<td>14/15</td>
<td>C C C C C C C C C C C C C C C C C C C C C C C C C C C C C</td>
<td>5 3614</td>
<td>36 3397-3918</td>
<td></td>
</tr>
<tr>
<td>7/10</td>
<td>C T T C C C C C C C C C C C C C C C C C C C C C C C C C C C</td>
<td>7 4336</td>
<td>36 4172-4666</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C C C C C C C C C C C C C C A C C C C C C C C C C C C C C C</td>
<td>9 6883</td>
<td>36 6343-6930</td>
<td></td>
</tr>
<tr>
<td>Haplotype Pairs&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PS&lt;sup&gt;b&lt;/sup&gt; No.</td>
<td>PS Position&lt;sup&gt;c&lt;/sup&gt;</td>
<td>SEQ ID NO.</td>
<td>Region Examined&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>7/7</td>
<td>10/10</td>
<td>14/14</td>
<td>13/13</td>
<td>7/5</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>343</td>
<td>36</td>
<td>27-1092</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>346</td>
<td>36</td>
<td>27-1092</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>3</td>
<td>3453</td>
<td>36</td>
<td>3937-3916</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>4</td>
<td>3647</td>
<td>36</td>
<td>3937-3917</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>3614</td>
<td>36</td>
<td>3937-3918</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>4259</td>
<td>36</td>
<td>4172-4665</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>7</td>
<td>4336</td>
<td>36</td>
<td>4172-4666</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>8</td>
<td>6421</td>
<td>36</td>
<td>6343-6929</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>9</td>
<td>6883</td>
<td>36</td>
<td>6343-6930</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Haplotype pairs are represented as 1<sup>st</sup> haplotype/2<sup>nd</sup> haplotype; with alleles of each haplotype shown 5' to 3' as 1<sup>st</sup> polymorphism/2<sup>nd</sup> polymorphism in each column;

<sup>b</sup>PS = polymorphic site;

<sup>c</sup>Position of PS within the indicated SEQ ID NO;

<sup>d</sup>Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO of the sequenced region,

wherein a higher frequency of the haplotype or haplotype pair in the trait population than in the reference population indicates the trait is associated with the haplotype or haplotype pair.

17. The method of claim 16, wherein the trait is a clinical response to a drug targeting IL1B.

18. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the interleukin 1 alpha (IL1B) gene at a polymorphic site (PS) selected from the group consisting of PS3, PS4, and PS5, wherein the selected PS have the location and alternative alleles shown in SEQ ID NO: 36.

19. The composition of claim 18, further comprising at least one genotyping oligonucleotide for detecting a polymorphism in the interleukin 1 alpha (IL1B) gene at a polymorphic site (PS) selected from the group consisting of PS2, PS8, and PS9.

20. The composition of claim 18, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IL1B gene at a region containing the polymorphic site.

21. The composition of claim 20, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-9, the complements of SEQ ID
22. The composition of claim 18, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.

23. The composition of claim 22, wherein the primer extension oligonucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS:22-33.

24. A kit for genotyping the interleukin 1, beta (IL1B) gene of an individual, which comprises a set of oligonucleotides designed to genotype each of polymorphic sites (PS) PS3, PS4, and PS5, wherein the selected PS have the location and alternative alleles shown in SEQ ID NO:36.

25. The kit of claim 24, which further comprises oligonucleotides designed to genotype each of PS2, PS8 and PS9.

26. The kit of claim 24, which further comprises oligonucleotides designed to genotype each of PS1, PS6 and PS7.

27. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which comprises a interleukin 1, beta (IL1B) isogene, wherein the IL1B isogene is selected from the group consisting of isogenes 1-18 shown in the table immediately below and wherein each of the isogenes comprises the regions of the SEQ ID NOS shown in the table immediately below and wherein each of the isogenes 1-18 is further defined by the corresponding set of polymorphisms whose locations and polymorphisms are set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Isogene Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PS&lt;sup&gt;b&lt;/sup&gt; No.</th>
<th>PS Position&lt;sup&gt;c&lt;/sup&gt;</th>
<th>SEQ ID NO.</th>
<th>Region Examined&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>343</td>
<td>3627-1092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>346</td>
<td>3627-1092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3453</td>
<td>363397-3916</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3467</td>
<td>363397-3917</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3614</td>
<td>363397-3918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4259</td>
<td>364172-4665</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4336</td>
<td>364172-4666</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>6421</td>
<td>366343-6929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6883</td>
<td>366343-6930</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Alleles for isogenes are presented 5’ to 3’ in each column  
<sup>b</sup>PS = polymorphic site;  
<sup>c</sup>Position of PS within the indicated SEQ ID NO;  
<sup>d</sup>Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO of the sequenced region;

(b) a second nucleotide sequence which comprises a fragment of the first nucleotide sequence, wherein the fragment comprises one or more polymorphisms selected from the group consisting of cytosine at PS3, cytosine at PS4, and thymine at PS5, wherein the selected polymorphism has the location set forth in the table immediately above; and

(c) a third nucleotide sequence which is complementary to the first or second nucleotide sequence.
28. The isolated polynucleotide of claim 27, which is a DNA molecule and comprises both the first and third nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

29. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 27, wherein the organism expresses a IL1B protein encoded by the first nucleotide sequence.

30. The recombinant organism of claim 29, which is a nonhuman transgenic animal.

31. The isolated polynucleotide of claim 27 which consists of the second nucleotide sequence.

32. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a coding sequence for a interleukin 1, beta (IL1B) isogene selected from the group consisting of coding sequences 3c, 8c, 13c, 17c, and 9c shown in the table immediately below and wherein each of the coding sequences 3c, 8c, 13c, 17c, and 9c comprises SEQ ID NO:2 except at each of the polymorphic sites (PS) which have the locations and polymorphisms set forth in the table immediately below:

<table>
<thead>
<tr>
<th>Coding Sequence Subhaplotype Number</th>
<th>PS(^a) No.</th>
<th>PS Position(^b)</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1c, 2c, 4c-7c, 10c-12c, 14c-16c, 18c</td>
<td>3c, 8c, 13c, 17c, 9c</td>
<td>5</td>
<td>138</td>
</tr>
<tr>
<td>C C T T</td>
<td>7</td>
<td>315</td>
<td>2</td>
</tr>
</tbody>
</table>

\(^a\) Alleles for coding sequence haplotypes are presented 5' to 3' in each column.

\(^b\) PS = polymorphic site;

\(^c\) Position of PS within the indicated SEQ ID NO.

and

(b) a fragment of the coding sequence, wherein the fragment comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 138.

33. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 32, wherein the organism expresses a interleukin 1, beta (IL1B) protein encoded by the polymorphic variant sequence.

34. The recombinant organism of claim 33, which is a nonhuman transgenic animal.

35. A computer system for storing and analyzing polymorphism data for the interleukin 1, beta gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the haplotypes set forth in the table immediately below:
Haplotype Number

<table>
<thead>
<tr>
<th>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C C C C C C C C C C C C C C C C C C C T T T T T T T T T T T T</td>
</tr>
<tr>
<td>A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A T</td>
</tr>
<tr>
<td>C C T T T T T C C C C C C C C C C C C C C A A A A A</td>
</tr>
<tr>
<td>C C C C C C C C C C C C C C C C C C C C C C C C</td>
</tr>
<tr>
<td>C C T C C C C C C C C C C C C C C C C C C C C C C C C</td>
</tr>
<tr>
<td>C C C C C C C C C C C C C C C C C C C C C C C C C C C C</td>
</tr>
</tbody>
</table>

Notes:
- Alleles for haplotypes are presented 5' to 3' in each column.
- PS = polymorphic site.
- Position of PS within the indicated SEQ ID NO.
- Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO. of the sequenced region.

...and the haplotype pairs set forth in the table immediately below:

Haplotype Pairs

<table>
<thead>
<tr>
<th>7/3 13/12 16/11</th>
<th>7/2 13/16 14/15</th>
<th>7/12 7/1 7/6</th>
<th>7/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>C C T T T T T C C T T T T T C C C C C C T T T T T T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T T C C T C T C C T T T T T C T T T T T T T T T T T T T T T T T T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C C C C C C C C C C C C C C C C C C C C C C C C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C C T C C C C C C C C C C C C C C C C C C C C C C C C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C C C C C C C C C C C C C C C C C C C C C C C C C C C C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- Haplotype pairs are represented as 1st Haplotype/2nd Haplotype; with alleles of each haplotype in shown 5' to 3' as 1st polymorphism/2nd polymorphism in each adjacent column.
- Location of PS within the indicated SEQ ID NO.
- Region examined represents the nucleotide positions defining the start and stop positions of the sequenced region.
### Haplotype Pairs

<table>
<thead>
<tr>
<th></th>
<th>7/4</th>
<th>14/13</th>
<th>14/9</th>
<th>10/4</th>
<th>14/1</th>
<th>14/16</th>
<th>14/10</th>
<th>10/2</th>
<th>7/14</th>
<th>7/8</th>
<th>PS No.</th>
<th>PS Position</th>
<th>SEQ ID NO.</th>
<th>Region Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>5</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>9</td>
</tr>
</tbody>
</table>

*Haplotype pairs are represented as 1st Haplotype/2nd Haplotype; with alleles of each haplotype in shown 5' to 3' as 1st polymorphism/2nd polymorphism in each adjacent column;*  
*Location of PS within the indicated SEQ ID NO;*  
*Region examined represents the nucleotide positions defining the start and stop positions of the sequenced region.*

36. A genome anthology for the interleukin 1, beta (IL1B) gene which comprises IL1B isogenes defined by any one of haplotypes 1-18 set forth in the table shown below:

### Haplotype Number

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

*Alleles for haplotypes are presented 5' to 3' in each column*  
*PS = polymorphic site;*  
*Position of PS within the indicated SEQ ID NO;*  
*Region examined represents the nucleotide positions defining the start and stop positions within the indicated SEQ ID NO. of the sequenced region.*
POLYMORPHISMS IN THE IL1B GENE

AAAGTATGTG CAGTGTATATA TCTGCTGTTGC TTTCCACCTTG TCCACATA
ACTAAATTTA AACATTTTCC TAACTGTGGGA AAATCCAGTA TTTTAATGTG 100
GACATCAACT GACACACGAT TCTGAGAAAA ACAATGCTAT TTTGCACTGG
TGATACATTG GCAAAATCTG TCTAGGTGGT TCTACCTCTGT CCCCCATAT 200
AACCCGAGAA TTCATTCCAGT TTATATGGCC CCTCCTCTCA AAGCAGTCCA
CCTAACACTT TTTTCTCCTTT CTTTTAACTT GAATTGGAA TCCAGATATTC 300
AACACGAGAAA TTTCTCAGCC TCTCAGCTCT GCTTTTGAAA GCCATAAAAA
TT TT
CAGCGAGGGG AAGAATCGCAA GATAAACCACT CTTCTGAGGG CACAGGCCAC 400
AAGAGGCTGC TCTGAGATTC TCTTCAGCCA ATCTCAATGG CTTCAAGTAT
ACTTTTATCT CTTCCCAACAA TTGTGCTGAA GGGAGTGCTTT TCTCAGCTCT 500
GCCCTCTGAT TGGTAGGGCAT ATTCCTCCTCT TCTCCCTCTT TCTTTCTCTG
TCTCCCTCTC TTCGCCCCCT CTCGAGGTCT TCTTACCAA A
AGCCAGGTGT AATAATATGC TTTAGACTCG GAGAATATTC TGGGAGATGGA 600
TACTGCTTAT CTACAGCTGT ACRCCTTAAA GGTATGCTC AAGGCCTCTG
CTCAGCAGCT TCTAGCCTAA TACATTGTGA GTGTTGGGTT GTTTTACCAA
ATGCTTTTCT TCTAGCCCAA AGAAGTACCTC TTCCACATAT TCTACTTAT
AAGCTGACGTTAT TCTTCGTGCAG CTAGTGGGGA TGCTGAGAAGA
ACTACACCT GACAGTATAG CGTCATGGGA AGTCTATCTA TTTCTCCTCTT
TTACAGCGGT GCTCAGAGCA GCCATCGGCAG AGATCTCTGA GCTGCCAGT

[exon 2: 924..970]
GAAATGATGG GTCATATTAC A GTCAGTGGAG AGCTAGAGAC CAGTAACTAG 1000

AGCACGCTTCC CTCTTCTCAGA AGTAGACTGT TATCTGTGCT TGGAGACCCAG
ATTTTTTCCTCA TAATGGTTCA CTTTCTAGTG CGAAAGCTGA GCCAATGGA A 1100
TCTGATTTAA AGACCTACTT CCCATTACAA GTCTCACTGC CTTTGAGGAC
TGGAAGCTAT CAGATGTGTT GTGTTGGGAA GCCTCCCTCA GAGCCTAATG
GGGAGGAAGA GCTCACAAGGC CACAAACTAA GAATTTCTTC GAGAAGTTG
GGCTTGAGGG GAGGGGAGGA GCTCATAGTT TAGCTGACCT GCTGAGCTGA 1300
GAGTAGTCAT GCTGGCCACT GACAGCTGCTT GCAGCATATC AGCTGAACAT
TGCCAGGCTT TCTCGCCCAC TTGGAGAGCC AGCCTTTGTT GAGTTCTTCA
TGAACAGCAG AGCTTCTGATC AAGGGGAGAG AAAATAAAAC CTGTTAGGAGA 1500
GACATAATGT AGACAGTCTCC CTTCTTCTAT TACAGTGCC AGTGGAGTAGA

[exon 3: 1536..1587]
CTTGTTCATT TGAAGCTGATG GCCCTAAACA GATGAAAGTA AGACTATGG 1600

TTTAACCTCC AACCCAGGGG AGGGGCTTAA CACAGCAGAA GCTCAAGGAA
GGAGTTTCTG GGCACCTTTG ATGCCATGGT ATTTTTGTAT TTGAGACCTT 1700
TACCTCCTCT CGACTGACAA AGCTGCTCAG CCTCTGCTGA CTTGCCACCT
TGGTCACTAT ATTCAACAAG TCTCCTCACA AGGTGAGGT TGTTGGCCAC
ACTTTGGGTG GACAGGGGAG GAGTAGTGAT AATGTTTCCC ATTTCTGATG
AGGAAGACAA CGAGGCTTTC TAACAATATTT GATTATCTCT TTTATAGAT
GGATCTCCAG TATCGGACCA CTGGGATGTA AACCTGGAA CCAAAGGATT
ATCTTGAGAA CTAACATATG TCCTACCCCT TTGGAGTGAAG TTATTTTGG
CTACCTGAGGA TAAAAGATAG AAAACACAGA CACTTCTGTT TGGATCATGA 2000
AGATCATTGA AGACTGAAAA GAAAACCATC GACTGCGTGG TGGACCTTAG
CAAGCTCATT TACCCGCTTT GACCTGCAAC TGGAAATATA AGGCTGAGCC
TGGTGATGCT CTTGAGATCC AGCAGTCTGC AGCTCGGATT CTGGAAATAT
TTCTCGGATG AGCTAGAGGG ATTTGCGGCA CAAATTGGTC TTTTCCAGAG
TCTACCTTAC AAAGAGCCTGT GTTATATCCC TGCTGCTCTT TCTGTTTAT 2300
ATGATGGCTC AGTGAGCCCT TCTGCTCGCC AGCAATGACC CTAGCTGAGT
CGTTGTGCA GTTGTCGAGG AGCCACTTTT CTCTGGCTTT ATTATTATCC
AGTTGTGTGT AGCCTCCCTAA GCTCAGTCTC ATCTTTTGGT

FIGURE 1A
AAACATATT TCTTTTAGAG TTTTAAAGACT GGCATAACTTT GTGGGCTGCA 2500
GCTGTGGGAG GAGGCCATTG GCTGGTCTGC CTTGCCCTTTG CCCCCATTCG 2600
CTCTTCCAGC AGCTTGGCCG TGCTCCAGGC AGGAAATCTT CTCCGTCTCA 2700
ACTTTCTTTT GTGCACCTTA AGGTCTCTTT AACTGCTTCTT CAAGCCTTTG 2700
AACAATTATC AGCTCCCATA GCCAACCTTC GAAGGAGCTTT ATGAGGAGGC 2800
TTCTCTGATA GGTTGAAA GAGGCTACCT TGACAGCTGG TAAAGTGGAA 2900
AACTTTGCTG CATCAGCTGG AAATAGAAA CTTGAGCCTG TGAGACCTAG 3000
AGTATATTTG TGGATGAAT AGGTCACACA CTGAAATGAA AGTGGGGGCA 3100
TGCTGAAA AGGTAGAAAC AGTGGAAAGT GATTTGATCG TAAAGAATAG 3200
CCCTGAAAGT GCAGAAGAGC CCACATTCTA GTCCATGCTC TGTTTCTAATC 3300
CTGTATGCTG CCCCCATGTA CACACATACG CTCTTGAGCT TCAAGAGGC 3400
ATAACTTTTT ATATTATTAT TATTTTATTT TGTTTTTGTG TGGTTTTTCTT 3500
TGTTTTATGA AACAGAGTCT CACCTCTGTG CCCAGCTGGA AGTGCCATCG 3600
TACAAATCCTG GCTTACTCGA TCTCCACACT CTTGAGTTCA AGCGTATCTC 3700
CTCCCTCAGT CTCTCGAATA GCTAGGATTAA CAGGTGCACC CACACACACC 3800
CAGCTAAATTT TTGTTATTTTT AGTAGAAGAG GGGTTTCGCC ATGTGGGCCA 3900
GGCTGTTTGT GAAGCTCTGA CTTAAATGAT GTACCCACCT CGGCTTCCCA 4000
AAAGGCTTGG TGAGAGAGCA CCGAGCCAAA AGAGATTCAGT TGAAGGCAAG 4100
GATCTTTTAGA AGCTCGGGAT TCTTTTCAAGC CTTTTTCTCC TCTTGTGACT 4200

T

A

CAACATGTCG TCCCATATTC AGAACCATATC TTCTTGGACA CATGGGATAA 4300

T

[exon 5: 4323..]
CGAGGCTTAT TGGCAGCATG CACCTGTACG ATCAGTGAAC TGCAACGCTCC 4400
GGAGCGTACA GCAGAAAGC TTTAGTGATTG CTGGCTCCATA TAAGCTGGAA 4500
GCTCTCCACC TCCAGGGACA GGAATGGAG CAAAGAGGTA AATGGGAACA 4600

T

[exon 7: 4487..]
TACATGTGGC CTTGGCCCGG CTTCCCTGGC CTTGGCTGGG CTTACGAAAT 4700
AAAGATGTAA AGAGAGATAT ACCAGGCTTG GCCTTGAAGA TGCACCCAG 4800

T

FIGURE IB
TGGAGCGATA ATGGTGTCCAT CCCCTCCCCA GTCTTCCCCC TTGGCCCAA 4900
CAGTCCGTTT CACCAATGTC AGTGGTTCTCT TGTAGGAAAT TTTTACGCC 5000
CAGCAGGATCAT TATTATCTCT CCGCTGTAAC GGGCAAAGGT TTCAAGTGCT 5100
GTGAAGCATCAT TGAACTAGTG GTGATGCTGC TTGGGATCGT CAGCACAGTG 5200
CAAGAACCTCT TGACCTGGAG TGTTGGCCACG TAAAGCTGCT GACTTTGAG 5300
GACCGTCTCA CTGCGGGTTA AGTCATATGC TCGGGCTGCT CAATCTGAAC 5400
TCTGCTTCTG GACACCCAG AGTGGTGCAG TCTCTCCCA AGTGGTGCAG 5500
ACCTGCTTCT TCCGCTTCTT CATTTAGGGA CAGTGGTCAG TCTCTTACCT 5600
TTGTTGTCTG CATGCAGCCT CATACCCGAG TCTCTCCTCTG CATACCCGAG 5700
AGCTTATTTA TCTGGTGCCG TGTTGGGCAG CAGAAACCC TATATATATAT 5800
CTCTGAGGAA ACCATCAGCC TCTGTGGTGA CAGAAACCC TATATATATAT 5900
[exon 6: 5723...]
AGAGAAGACCT TATGACAATA TACCTGTTGC CTTGGGCTCCA AGGAAAGG 6000
ATCCTGACTCT GTCTGCCTGG TTTGAAATG TAAAGCCCAT TCTACGCTG 6100
GAGGTAAGTG AATGCTATAG GTTGACAACC TCTCTGCTCC ACTCTAGGAGG 6200
CTTCACTGAT CAGAACGGTT CTGCCAGATT ACCAAATCA AAAAAAAAT 6300
TCTACTGCTC TCTGTCTCTT AGTTTTCTG AATAGGTTG GGGCAAGG 6400
TTGTTTTTCT TCTGGTCCAGA GTGTAAGATCC CAAAGATTTAC CCAGAAGAA 6500
[exon 7: 6570...]
AGATGGAAAA CGATTGGTCT TCAACAAGAA TCAACAGAA TACACAGCTG 6600
GAATTGAGAG CTGGCTGATT CCCCCATGG TACATGACGA CCTTCTCAGA 6700
AGAAAGACCT TGCTGCTTCC TGGGGAGGAC CAAAGGGGGA CGAGATATAA 6800
CTGATCTTCC CATGACAGTT GTGTTCTCCT CAAAGAGGCT GTACCCAGA 6900
AGTCCCTGATCT GAAATGTGGA CTAATACCTT AGGGCTGGCA GAAAGGGAAC
AGAAAGATTG CTAGTAGCCGA GTTCTTCTCAGG TCTCTTACCA

A
CAAAACCGAAG CGTGGCTGATT TGGGGATCGT CTAGGAGATG CTCCCGCTCA
TCAGCCCCAC GGCTGCTGATT CTGGGGGAC CAGGCCCTTT
TGCTGACGGA CGGGCGACG TCTTCTCGTCA TCTTGTGCTC CAGGGCTG
CAAAGACCT GCAGCAGGTG GCTCCGACTT GTTCCTAAGA AACCTCGCTC
CAATTCCTGA TGGGCAGACC TCTTCTCATG TATTATTATA CTGTTGTGTTT
TGTATTACAT CTTGTCTCAA TTTATTCAAA GGGGCGAAGA AGTACCGAGT
TCTGTAAGAA AGCCTGATTT TAAATAGCTA TGGATGCAAT TCTTTGCTG
CTGGGGCTGT CTCTTAAAAA CAAGTCTTTT ATTAAAGCAT GAAGATATAA
AAGCTCAGAT TATTAAATAT GGAATATTAA TAAATAGGCA TAAATCATAC

FIGURE 1C
TGTTCAATGG TTCTGAAATA AACTTCCTCTG AAGAAAAAAA AAGGGTCTTT 7400
CCTGATCATT GACTTGCTTT GGATTTGACA CTGAACAGTA AAGACAACA 7500
GGGCTGTGAG AGTTCCTTGGG GGACTAAAGC CCACCTCTTC ATGCTGATG 7500
GCTGCAAGTG CACCTAGAAA TATCCCTTGG CCACCGAGA CTATCTCTCT 7600
CACCCATCCC CTTATTCTC CTTGTCCACAG AGAAGGATAT TCAGTGCAACAT 7700
TCTGGAAACAG GATCAGCTGA AGCACTGCGA GGAACAGGCA CTGCTAGTAA
CAGCTACCAG TGATTTATCT ATCAATGACAC CAAACATCTG TTGAGCAAGC
GCTATGTACG AGGAGCTGGG AGTACAGAGA TGAGAACAGT CACCAAGTCC
TCCTCAGATA GGAGAGGCAG CTAAGTATAAA GCAGAAACAA GGTAACTAGA
CAAGTAGAGT AAGATAAAGA ACA 7824

FIGURE 1D
POLYMORPHISMS IN THE CODING SEQUENCE OF IL1B

ATGGCAGAAG TACCTGAGCT CGCCAGTGAA ATGATGGCTT ATTACAGTGG
CAATGAGGAT GACTTGTTCCT TGGAGCTGAA TGGCCCTAAA CAGATGAAGT
GCTCCTTCCA GGACCTGGCC CTCTGCCCTC TGGAGGGCGG CATCCAGCTA

CGAATCTCCG ACCACCACTA CAGCAAGGAG TCTCAGCAGG CCGGTCAGT
TGTCTGGCC ATGGACAAGC TGAGAAGAT GCTGTTCCTC TGCCACAGA
CCTCCAGGA GAAFGACCTG AGCACCCTCT TTCCCTTCAT CTGTTGAAGA
GAACCTATCT TCTTCGACAC ATGGGATAAC GAGGCTTTAG TGCCAGATGC

T
ACCTGTAGCA TCACTGAACT GCACGCTCCG GGAGTCACAG CAAAAAGCT
TGTTGATGTC TGGTCCATAT GAACGTGAAAG CTCTCCACCT CGAGGACAG
GATATTGGGC ACAAAGTGTT GTTTCTCATG TTCTTGATAC AAGGAGAAGA
RAGTAATGAG AAAATACCTG TGTCCTGGGG CCTCAGGAAA AAGAATCTGT
ACCTGTCTGCG GTGTGATGGA GATGATAAGC CCACTCTACA GCTGGAGAGT
GTTGATTTCCA AAAATTACCC AAGGGAGAAG ATGGAAAAGG GATTTGCTTT
CAACAGATA GAAATCGATA AAGAGCGGAG ATTTGAGTCT GCCCAGTTCC
CCAACCTGGA TATCAGCACC TCTCAAGCAG AAAACATGGC CTGCTTCTCG
GGAGGGGACCA AAGGCGGCAAG CAGGATACCT GACTTCACCA TGCAAATTGT
GTTTCCCTAA

FIGURE 2A
AMINO ACID SEQUENCE OF THE IL1B PROTEIN

MAEVPELASE MMAYYSGNED DLFFEADGPK QMKCSFOQDLQ LCPLDGGQQL
RISDHHYSKG FRQAASVVVA MDKLKLMLVP CPQTFQENDL STFFPFIQEE 100
EPIFFDWDN EAYVHDAPVR SINCJLRRDSQ QKSLVMGIQPY ELKALHQLQGQ
DMEQQVVFSTM SFVQGEESND KIPVALGLKE KNYLYLSCVLK DDKPTLQLES 200
VDPKNYPKKK MEKRFVFNHK IINNKLFEES AQFPWYIST SQAENMPVFL
GGTGGQDIT DFTMQFVSS 269
<110> Genaissance Pharmaceuticals 
   Bentivegna, Steven C 
   Chew, Anne 
   Choi, Julie 
   Denton, R. Rex 
   Lee, Helen H. 
   Nandabalan, Krishnan 
   Stephens, J. Claiborne

<120> Haplotypes of the IL1B Gene

<130> MWH-0010PCT2

<140> TBA
<141> 2001-05-22

<150> PCT/US00/25794
<151> 2000-09-20

<160> 36

<170> PatentIn Ver. 2.1

<210> 1
<211> 7824
<212> DNA
<213> Homo sapien

<400> 1
aaagtatgtg catgtataaa tctgtgtgtc ttccaccttg tcccatatat actaaattta 60 
aacattcttc taacgtggga aataccagta ttttaatgtg gacatcaact gcacaacgt 120 
tgctaggaa acaatgcata ttgcaacttg tgatacattt gcacaattct gcataagttg 180 
ctactccctg ccctccatag aacaagagaa ttatctcagt ttattagccc cctccoccta 240 
gaagctccaa ccaataacct ttccoccttt cctttaacct gatttggaat tcaggtatcc 300 
aacagagaaaa ttctctcagcc tcctacttct gcctttggaa gccataaaaa aacgctgggaa 360 
gaaactcgaca gataaacaac ctcttgagag cacaagggaca aacagctgct gcctgggattc 420 
tcctcagcc caatctcattg ctcagaagctg accttataatct tccttacacc taggggtcct 480 
gggagtctct ctctctctct gcctctctctgt gtgtagctat atctctctct tcctctcttt 540 
tctctctctg tctctctctct cctctctctct cctctctctct tctctctctct gttgcaaaa 600 
agccaggtgt attataatgc ttatgtactg ggaatatattc tggaatggaa tactgtttat 660 
tcaaccagctg acacccctaa gggttaggttc aagacctcttg cctccagctct cctagcacaa 720 
tacacctgca gttgggtggtt gtttagcga aatgtttttct ctgagaaccca aagagctctctc 780 
ttcacaccat tcattcatt catccagagat caattttgct catgactgcc atgcactgga 840 
tgctggaga aatccacacc gaaagctgga actcatggag ggtccacagtt atctttctctt 900 
ttacacacct gttcgaagcga gcacagcagg aagtactctga gctcgccagt gaaatgatgg 960 
cttattacag gtcagctggag agctgtagac cagaatacatg agcaggtctct cttcttcagag 1020
aaacaaagaa ggctggaaac caaagcaact atcttttagg tggaaactat tcttaaagaa 3960
gatctttgag tcctagtcga ctttgcacatcctcctactt ttcaggggct cttttcacttaa 4020
cattgctcacc aaggtggcgt acatcctcttg aggctctattg gaaatgacat ccaaccttttaa 4080
cctaaagtaa tcctgtgtgct ccggccagttg aggagtataa gagcttgagaac aagcaacat 4140
gctcaaatag gcagtggttac cccagctcactc ccagcctctcct ccctactgtgt 4200
gattggtcatca aagtttgggc gtattagttgc agttatcctc cagagatgata aagaaatcct 4260
tctgcctcacc gaggctgcctc cttctgtattt tataatccaa caaagactgc cccaccttac 4320
agaaaacctc ttccttcgcact catggtggat cagaggtttat gtgcagcctag cccctgtcag 4380
atcacatgac tcgcagccgca caaagaaagc ttggtgtagtt cttggtcctata 4440
tgaaactgaa gctcttcgacc ttcagggcca ggtatgtagc caaagagata aatggaacca 4500
tcctgtggttccc ccctgcgtgct ccctggtgccc cttcggctatg ttaaacaagt 4560
agaagaagtta aaaaaagccaa atgatcaaca caagttgaaa aaaaattataa aaaggaat 4620
acaacatttt tttctttagaaa tggcactctt gattgagctgt gcaggctgtag ttttaaacag 4680
gagtgtgaccc tattgaaatgc agagttgaag gcgtttgact cccacttctcc aaaaaatggc 4740
tctgggctttgc tggcatcactgc aaattttccc tataatcctcc catctagttt gacactgtgcct 4800
aactgcgactt ttcctggtctc ttcctgtcttg agagttgcagtt aaagagtgac ttgggagcata 4860
atgatgctcat cctccctttccc gtttcctgcct cagttggccaa cagttgctgc ccaccaagtg 4920
aggtgggttttt tttcttccaggc ccttcaagagc attttgatgaa tattagttgttt 4980
ttcagagcagag ggtgacgtcct cggccagccaa agcgcaagta tccctgttgcct 5040
ccagcagactcc gaaagggcttc tgagcagggag tggccgacactgt 5100
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5160
tggacagctgat ctaagctattg tggctcagcct ggtctcttccag cctcttttccag 5220
agatcaaaaat gatgatctcttc cggccagctg cctcttggtctct cctctgtcttt tggacgctcct 5280
ccccctttcc ctgttgcactc tctggtgtcctt aacagttttc tttctgtgtag ttttgggagc 5340
agtacagcttt cctttgtcttt cccctctgtcc atagtgctccat aagttttgct cttgggtggaa 5400
ttttttttcct ttcctctgtcct cttgattcgct ttcctgtgcctc aagtttttggg aagttttttaa 5460
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5520
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5580
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5640
tttgggcttttt ctgtactacccct cccctctgtcc atagtgctccat aagttttgct cttgggtggaa 5700
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5760
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5820
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5880
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 5940
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6000
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6060
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6120
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6180
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6240
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6300
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6360
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6420
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6480
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6540
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6600
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6660
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6720
agatatcactt ctctctgggggc aggattttgc ctctcagcgc gcacattttg tttttgccct 6780
aaagagcagt gatccagag ctgctctgtgc tgaatggtgaa ctcataatcct gaggctggca 6840
ggaagggac aagaaaggttt ttagatcagg stataagcctg ccattcctgctt ttgctcaac 6900
catgcccaaa cctgcctgctt taggtagagt ctaagagatt ctcctgtcga ccagcagcagga 6960
cagctcgctc tctctttaaa gcggccatccc ccaagcctttt tgggagccta ggcttctcctc 7020
acccctcctta ctcaactttaaa gcggccatctg cagaaccaac cgcacatatt ggcttctaaga 7080
aaccctccttgc cattcgtctcc cacatctctga tggaccaatgc cttctctattt tatattttata 7140
ttttttttttt tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Met Ala Glu Val Pro Glu Leu Ala Ser Glu Met Met Ala Tyr Tyr Ser  
1  5  10  15  
Gly Asn Glu Asp Asp Leu Phe Phe Glu Ala Asp Gly Pro Lys Gln Met  
20  25  30  
Lys Cys Ser Phe Glu Asp Leu Asp Leu Cys Pro Leu Asp Gly Gly Ile  
35  40  45  
Gln Leu Arg Ile Ser Asp His His Tyr Ser Lys Gly Phe Arg Gln Ala  
50  55  60  
Ala Ser Val Val Val Ala Met Asp Lys Leu Arg Lys Met Leu Val Pro  
65  70  75  80  
Cys Pro Gln Thr Phe Gln Glu Asn Asp Leu Ser Thr Phe Phe Pro Phe  
85  90  95  
Ile Phe Glu Glu Glu Pro Ile Phe Phe Asp Thr Trp Asp Asn Glu Ala  
100 105 110  
Tyr Val His Asp Ala Pro Val Arg Ser Leu Asn Cys Thr Leu Arg Asp  
115 120 125  
Ser Gln Glu Ser Leu Val Met Ser Gly Pro Tyr Glu Leu Lys Ala  
130 135 140  
Leu His Leu Gln Gly Glu Met Asp Met Glu Gln Val Val Phe Ser Met  
145 150 155 160  
Ser Phe Val Gln Gly Glu Ser Asp Lys Ile Pro Val Ala Leu  
165 170 175  
Gly Leu Lys Glu Lys Asn Leu Tyr Leu Ser Cys Val Leu Lys Asp Asp  
180 185 190  
Lys Pro Thr Leu Gln Leu Glu Ser Val Asp Pro Lys Asn Tyr Pro Lys  
195 200 205  
Lys Lys Met Glu Lys Arg Phe Val Phe Asn Lys Ile Glu Ile Asn Asn  
210 215 220  
Lys Leu Glu Phe Glu Ser Ala Gln Phe Pro Asn Trp Tyr Ile Ser Thr  
225 230 235 240  
Ser Gln Ala Glu Asn Met Pro Val Phe Leu Gly Gly Thr Lys Gly Gly  
245 250 255
Gln Asp Ile Thr Asp Phe Thr Met Gln Phe Val Ser Ser
   260      265

<210> 4
<211> 15
<212> DNA
<213> Homo sapien

<400> 4
aagccatwaa aacag  15

<210> 5
<211> 15
<212> DNA
<213> Homo sapien

<400> 5
gggatgayct ttaga  15

<210> 6
<211> 15
<212> DNA
<213> Homo sapien

<400> 6
aagctcgsga ttctt  15

<210> 7
<211> 15
<212> DNA
<213> Homo sapien

<400> 7
tggatgyygc catscc  15

<210> 8
<211> 15
<212> DNA
<213> Homo sapien

<400> 8
ttcacacrga aagtt  15
<210> 9
<211> 15
<212> DNA
<213> Homo sapien

<400> 9
gcctggamtt tcttg

<210> 10
<211> 15
<212> DNA
<213> Homo sapien

<400> 10
ttttgaagac catwa

<210> 11
<211> 15
<212> DNA
<213> Homo sapien

<400> 11
ccctcgcgtgt tttwa

<210> 12
<211> 15
<212> DNA
<213> Homo sapien

<400> 12
cagagaggga tgayc

<210> 13
<211> 15
<212> DNA
<213> Homo sapien

<400> 13
cagctttcta aagrt

<210> 14
<211> 15
<212> DNA
<213> Homo sapien

<400> 14
ccttagaagc tcgsg

<210> 15
<211> 15
<212> DNA
<213> Homo sapien

<400> 15
gcttagaaga atcsc

<210> 16
<211> 15
<212> DNA
<213> Homo sapien

<400> 16
gccctctgga tgyyg

<210> 17
<211> 15
<212> DNA
<213> Homo sapien

<400> 17
gtagctggtgcrc

<210> 18
<211> 15
<212> DNA
<213> Homo sapien

<400> 18
aataaattca cacrg

<210> 19
<211> 15
<212> DNA
<213> Homo sapien
<400> 19
gggcacaact ttcyg
15

<210> 20
<211> 15
<212> DNA
<213> Homo sapien

<400> 20
gctatagcct ggamt
15

<210> 21
<211> 15
<212> DNA
<213> Homo sapien

<400> 21
agacaacagg aaakt
15

<210> 22
<211> 10
<212> DNA
<213> Homo sapien

<400> 22
tgaaagcccat
10

<210> 23
<211> 10
<212> DNA
<213> Homo sapien

<400> 23
tcgctgttttt
10

<210> 24
<211> 10
<212> DNA
<213> Homo sapien

<400> 24
agagggatga
10
<210> 25
<211> 10
<212> DNA
<213> Homo sapien

<400> 25
gcttctaaag

<210> 26
<211> 10
<212> DNA
<213> Homo sapien

<400> 26
tagaagctcg

<210> 27
<211> 10
<212> DNA
<213> Homo sapien

<400> 27
tgaaagaatc

<210> 28
<211> 10
<212> DNA
<213> Homo sapien

<400> 28
cctctgatgga

<210> 29
<211> 10
<212> DNA
<213> Homo sapien

<400> 29
gctggatgccc

<210> 30
<211> 10
DNA
Homo sapien

aaattcac

cccaccttttc

atagcctgga

caacaggaaa

tgtaaaacga cggccagt

16
<400> 35
aggaacagc tatkccat

<210> 36
<211> 7824
<212> DNA
<213> Homo sapien

<220>
<221> allele
<222> (343)
<223> Polymorphic Site PS1: c or t

<220>
<221> allele
<222> (346)
<223> Polymorphic Site PS2: a or t

<220>
<221> allele
<222> (3453)
<223> Polymorphic Site PS3: t or c

<220>
<221> allele
<222> (3467)
<223> Polymorphic Site PS4: g or c

<220>
<221> 5'UTR
<222> (3614)
<223> Polymorphic Site PS5: c or t

<220>
<221> allele
<222> (4259)
<223> Polymorphic Site PS 6: g or a

<220>
<221> allele
<222> (4336)
<223> Polymorphic Site PS7: c or t

<220>
<221> allele
<222> (6421)
<223> Polymorphic Site PS8: a or g
<220>
<221> allele
<222> (6883)
<223> Polymorphic Site PS9: c or a

<400> 36
aaagatagtgt cagtataaa tctgtgtgtc ttccactttg tcccacatat actaaatatta 60
aacatccctc taaaggtggaa aataccagta atttaatgtg gacatcactc gcacaagcagat 120
tgctaggata aacaatagta ttgctcattg tatacattttgc caaataatctg tataagtgttg 180
tcatctcttg ccctctcagtg aacagagaga ttatctctgt gtttacctgc 240
gaaagtccca ccataacctct ttctccccttct otctatctactc gatttgtgataa tcaaggtatc 300
aacagagaa attttcctcagg tctttctctct ctttgcattgt gcacagtaga 360
agcaagttgt tacataaatgc ttaagacattg gcagaatatc attgctatgg gaggatatgg 420
cttacagtct acacacttaa cacttgtattgc aatctatcactt gcttactatgc actttatcaac 480
tacatgccata gttggggttttt gtttactgttaa actttatctctc tctttactttt 540
ccctcctctc tctctttcttt ctctctctctct cttttactgaaat tttttgcaaa 600
tgcttacgaga atacactcatc gcacaagcagct cattcacttc attttccttctt gcacaagcagct 720
ttacacagtgt tattagttattc gttgcttactt gttttacttattc gttttacttattc 840
tgtgctgattc tggagacacttc attcacttactt actttacttttttt acttcacttactt 900
tacacagtgt ggttgattttt gttgcttactt gttttacttattc gttttacttattc 1020
tttcttttttt tttttttttttt tttttttttttt tttttttttttt 1120
gaggtggtta accagcttact accagctttact accagctttact 1220
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 1320
cccacttactc cctacttactc acatgacttc tattttttttt tattttttttt 1420
cagctttttta cagctttttta cagctttttta cagctttttta 1520
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 1620
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 1720
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 1820
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 1920
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2020
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2120
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2220
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2320
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2420
agcttacttactcc atcactttactacc gagaataatag ccacatgatgtt 2520
agtacagctt caccccttctc ctttccttccc acattgatcata gttggtcgcct ctctctggtgga 5400
tgggcaacatt gccagcgcgt gcacaaatggt cttcctctcct ctctctctctc agcattaaaaa 5460
atgtagccct cttctctccct tccgcttccta ctgtctatgag gctcttgagaa acctccaggc 5520
ttggagggga aacccttaaat caacaaatggt accctgctaat gttctgtgag aagtcgataat 5580
atcctgtgctc tatgctgccag tgcgctctccct gcagtaaggcct gagcgagctgacc 5640
tgtatctcact ctatgggtggtt gcctgaacccct cactgctgtgctt tcccattaaccat 5700
caagccccccc cttttctcttc agtggtgttgct ctatgtctctt gtgcattcagaa aagaggaaatgc 5760
aagagccaaata ctcctggtggct cctggctccct caagagagctt cttggaagtttgg tttttctacgc 5820
tgagagctactt gagacacccgta gctggttaaat cggaactttgct cttgacagctt 5880
cttctgccacct cccatcachac ccaacacaggc cttctcccgcac cccatccctct 5940
aagaaagagct cggcagctcgat tttctggctaa cttttgtgtaa aattttacacac 6000
atttttcac ctcgctagccag cctcaacctctc gcctcagacaattgatc 6060
ttattagcctt tacccattact taggagcgtttct tacaattacat cagacctcctcagc 6120
cttggattgct ttcttagttgtttc cccatcactctt cttggtttttt ccaacagctat 6180
tttgacacagctt cttccttcattttcataaa ctttcatcagaag ttttccataacat 6240
rgaaagtgcct cccggcagctgcc agtctccgcatggtccttcct ctgttctggtgct 6300
gatgctcctt gttgagcagcct cggcctgcagcct cttctggtcctct tcctgctgtgct 6360
tctccgcatc ttaagggctttc ttcacactgatta gcctcaagcctt gagtttttctc ttaatctcttt 6420
ctcgtgatatg ccttcagcaattt cctgctgctgctt cccatttctttt gcagcgtggac 6480
aagagagctct tgcctcgctcact ccgtaagagct cctgctctctgt ctaccaatctcttc 6540
ctcaattactcat cttttctcttcagactcagttt cccctgctgctt cccatcactctt tttttctacgc 6600
agatcgattag ctcgctagccag cctcagccactt cttggtttttt ccaacagctat 6660
ctcctcgaagctt cagaggccgcgtc tatcagatactgct cttcctcctcta cagactcctctcctt 6720
ttttctctcttcctt ctggttttctct ctggttttctct ctggttttctct ctggttttctct 6780
cttccggagct ccttccttcctt cctcctctcctt cttcctctctcctt cttcctctctcctt 6840
gagctcctgact cctcctctcctt cctcctctcctt cttcctctctcctt cttcctctctcctt 6900
cctcctgact cctcctctcctt cctcctctcctt cctcctctcctt cttcctctctcctt 6960
cctcctgact cctcctctcctt cctcctctcctt cctcctctcctt cttcctctctcctt 7020
acacctgccttc tctacctctcctt ggacccactcc cccctgctgctt cccatcactctt tttttctacgc 7080
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7140
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7200
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7260
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7320
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7380
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7440
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7500
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7560
tttttctctctt cttcctctcctt cctcctctcctt cctcctctcctt cttcctctcctt 7620
tttttctctctt cttcctctcctt cctcctctcctt cttcctctcctt cttcctctcctt 7680
tttttctctctt cttcctctcctt cctcctctcctt cctcctctcctt cttcctctcctt 7740
tttttctctctt cttcctctcctt cctcctctcctt cctcctctcctt cttcctctcctt 7800
tttttctctctt cttcctctcctt cctcctctcctt cctcctctcctt cttcctctcctt 7860
tttttctctctt cttcctctcctt cctcctctcctt cctcctctcctt cttcctctcctt 7924
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12Q 1/68; C12P 19/34; C07H 21/02, 21/04
US CL.: 585/6, 91.2; 586/22.1, 23.1, 24.1, 24.3, 24.31, 24.33
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 485/6, 91.2; 520/22.1, 23.1, 24.1, 24.3, 24.31, 24.33

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>US 5,521,381 A (WALLACE et al) 28 May 1996, see entire document.</td>
<td>1-4, 9-12N</td>
</tr>
<tr>
<td>Y</td>
<td>US 5,851,762 A (SIMONS) 22 December 1998, see entire document.</td>
<td>1-4, 9-12</td>
</tr>
<tr>
<td>Y</td>
<td>WO 00/47619 A1 (INTERLEUKIN GENETICS, INC.) 17 August 2000, see entire document.</td>
<td>1-4, 9-12</td>
</tr>
<tr>
<td>Y</td>
<td>WO 00/37679 A2 (INTERLEUKIN GENETICS, INC.) 29 June 2000, see entire document.</td>
<td>1-4, 9-12</td>
</tr>
<tr>
<td>Y</td>
<td>WO 99/54707 A2 (MEDICAL SCIENCE SYSTEMS, INC.) 28 October 1999, see entire document.</td>
<td>1-4, 9-12</td>
</tr>
</tbody>
</table>

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
*"A" document defining the general state of the art which is not considered to be of particular relevance
*"E" earlier document published on or after the international filing date
*"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
*"O" document referring to an oral disclosure, use, exhibition or other means
*"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"W" document member of the same patent family

Date of the actual completion of the international search: 17 OCTOBER 2001
Date of mailing of the international search report: 1 DEC 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 205-9250

Authorized officer
JEFFREY FREDMAN
Telephone No. (703) 205-3014

Form PCT/ISA/210 (second sheet) (July 1998)
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
</table>
## Box I  Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(3)(a) for the following reasons:

1. □ Claims Nos.:  
   because they relate to subject matter not required to be searched by this Authority, namely:

2. □ Claims Nos.:  
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. □ Claims Nos.:  
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II  Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

   Please See Extra Sheet.

1. □ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. □ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. X As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:  
   1-4, 9-12

4. □ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  

**Remark on Protest**

X The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.
B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

EAST, CAPLUS, MEDLINE, BIOSIS
search terms: haplotype, genotype, mutation, polymorphism, variant, alteration, allele, interleukin, beta, gene, IL1B,
SNP, sequence, nucleic, DNA, RNA

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Groups 1-18, claim(s) 1-4, 9-19, in part, drawn to methods for haplotyping IL1B comprising determining whether the
individual has one of the IL1B haplotypes shown in the tables.

Groups 19-27, claim(s) 5-8, in part drawn to a method for genotyping the IL1B gene. It is noted that Groups
orrespond to polymorphic sites PS1-9, respectively and pairs thereof.

Groups 28-46, claim(s) 13-15, in part drawn to a method for predicting a haplotype pair for the IL1B gene by
identifying a IL1B genotype for the individual at two or more polymorphic sites PS1-9.

Groups 47-65, claim(s) 16, 17, in part drawn to a method for identifying an association between a trait and a haplotype
between one of the 18 haplotypes and haplotype pairs of IL1B gene.

Groups 66-81, claim(s) 18-23, 27, 28, 81, in part, drawn to a composition comprising at least one genotyping
oligonucleotide for detecting a polymorphism in the IL1B gene.

Group 82, claims 24-26, drawn to a kit comprising a set of oligonucleotides designed to genotype each of the
polymorphic sites.

Groups 83-101, claims 32, in part, drawn to a polymorphic variant of a reference sequence for
IL1B gene or a fragment thereof.

Group 102-120, claim(s) 29, 30, 35, 34, in part drawn to a recombinant nonhuman organisms comprising one of the 18
haplotypes or haplotype pairs respectively.

Group 121-138, claim(s) 35, in part drawn to a computer system comprising polymorphism data wherein the data
comprises the haplotypes and haplotype pairs shown in the tables.

Groups 139-150, claim(s) 36, in part, drawn to a genome anthologies comprising IL1B isogenes having any one of the
haplotypes or haplotype pairs of the tables.

The products claimed in Claims 18-23, 31 and 32, include fragments of variant sequences, and the claims simply
require the presence of a single polymorphic site in the coding sequence, so that a random hexamer mixture,
commercially available, would anticipate the claims. Accordingly, the claims are sufficiently broad so as to
encompass nucleic acid fragments taught in the art. Teaching the sequence of the human IL1B gene and
fragments thereof. As the nucleic acid products do not represent a contribution over the prior art, the claims
lack a special technical feature that is the same as or that corresponds to a special technical feature of the other
claimed inventions. Thus, there is no special technical feature linking the recited Groups, as would be necessary
to fulfill the requirement for unity of invention.

It is also noted that each of the present claims has been presented in improper Markush format, as distinct
products and distinct methods are improperly joined in the claims. With respect to the claims, each polymorphic
site and each molecule containing said polymorphic site is structurally and functionally distinct from and has
a different special technical feature than each other polymorphic site and molecules containing said site. The
chemical structure of each polymorphism and of each molecule containing the same differ from each other.
For example, a polymucleotide comprising PS1 is chemically, structurally, and functionally different from a molecule
comprising PS5. As the products and methods encompassed by the claims do not share a special technical feature,
the distinct products and methods may not properly be presented in the alternative. Accordingly, the claims have
been separated into a number of groups corresponding to the number of different inventions encompassed by the
claims, and the claims will be examined only as they read upon the invention of the elected group. For the same
reasons, the remainder of the claims have been separated in a number of groups corresponding to the number of
different inventions encompassed thereby.
With particular respect to the haplotype and genotyping claims, it is noted that the haplotypes and genotypes encompassed by these claims are also distinct from each other and from the single polymorphisms recited. For example, a molecule of haplotype 1, comprising a particular combination of polymorphisms, differs chemically, structurally, and functionally from a molecule of haplotype 2 and from a molecule comprising a single polymorphism (e.g., PSI). The special technical feature of each haplotype or genotype is the combination of polymorphisms contained therein, which feature is lacking from and not shared with each other haplotype or genotype or with, e.g., a molecule comprising any single polymorphism set forth in the claims. Similarly, with respect to the pairs of polymorphism, each combination of polymorphisms differs from each other combination and from each of the other combinations discussed above (i.e., haplotypes, genotypes, and single polymorphic sites).

Thus, the claims have been separated into a number of groups corresponding to the number of different inventions encompassed thereby, and the claims will be examined only as they read upon the invention of the elected group.

Further polynucleotides, kits, and various compositions, recombinant organisms, computer system, and genome anthologies are additionally drawn to multiple, distinct products lacking the same or corresponding special technical features. The nucleic acids are composed of nucleotides and function in, e.g., methods of nucleic acid hybridization or amplification. These groups are directed to different combinations of nucleic acids which are different from one another and may be employed in different methods. Further the computer systems are composed of, e.g., a CPU, a display device, an input device, etc., and function in, e.g., methods of electronic sequence comparison. Accordingly, the products of each of these Groups differ structurally and functionally from each other. As products of different sets of Groups differ from each other in structure, function, and effect, they do not belong to a recognized class of chemical compound, or have both a "common property or activity" and a common structure as would be required to show that the inventions are "of a similar nature".

Further, the methods of each of the method Groups have different objectives and require different process steps. The haplotyping methods require steps of identifying haplotypes and haplotype pairs to achieve the objectives of haplotyping. The genotyping methods require steps of identifying a single nucleotide on one gene copy to achieve the objective of genotyping. The predictive methods require steps of identifying two polymorphisms in a gene to achieve the objective of "predicting a haplotype pair". The association methods require steps of comparing frequencies of haplotypes in a population to achieve the objective of "identifying an association between a trait" and a haplotype. In addition to differences in objectives, effects, and method steps, it is again noted that the claims of the present Groups are not directed to the detection or identification of molecules having the same or common special technical feature, for the reasons discussed above.