United States

Patent Application Publication

Stanton, JR.

Pub. No.: US 2007/0037164 A1
Pub. Date: Feb. 15, 2007

TUMOR NECROSIS FACTOR ALPHA

Inventor: Vincent P. Stanton JR., Belmont, MA (US)

Correspondence Address:
FISH & RICHARDSON PC
225 FRANKLIN ST
BOSTON, MA 02110 (US)

Assignee: Variagenics, Inc., a Delaware corporation

Appl. No.: 10/902,585
Filed: Jul. 29, 2004

Related U.S. Application Data

Continuation of application No. 09/969,188, filed on Oct. 1, 2001, now abandoned, which is a division of application No. 09/649,035, filed on Aug. 25, 2000, now abandoned, which is a continuation-in-part of application No. 09/590,749, filed on Jun. 8, 2000, now abandoned, which is a continuation-in-part of application No. 09/495,780, filed on Feb. 1, 2000, now abandoned, which is a continuation-in-part of application No. 09/492,712, filed on Jan. 27, 2000, now abandoned, which is a continuation-in-part of application No. PC1/US00/01392, filed on Jan. 20, 2000.

Said application No. 09/969,188 is a continuation-in-part of application No. 09/451,252, filed on Nov. 29, 1999, now abandoned, which is a continuation-in-part of application No. 09/427,835, filed on Oct. 26, 1999, now abandoned, and which is a continuation-in-part of application No. 09/414,330, filed on Oct. 6, 1999, now abandoned, and which is a continuation-in-part of application No. 09/389,993, filed on Sep. 3, 1999, now abandoned, and which is a continuation-in-part of application No. 09/370,841, filed on Aug. 9, 1999, now abandoned, and which is a continuation-in-part of application No. 09/300,747, filed on Apr. 26, 1999, now abandoned.

Provisional application No. 60/131,334, filed on Apr. 26, 1999. Provisional application No. 60/131,191, filed on Apr. 26, 1999. Provisional application No. 60/121,047, filed on Feb. 22, 1999.

Publication Classification

Int. Cl.
C12Q 1/68 (2006.01)
C07H 21/04 (2006.01)
C07K 14/47 (2006.01)

U.S. Cl. 435/6; 536/24.3; 530/350

ABSTRACT

The present disclosure describes the use of genetic variance information for genes involved in inflammatory or immunologic disease, disorder, or dysfunction. The variance information is indicative of the expected response of a patient to a method of treatment. Methods of determining relevant variance information and additional methods of using such variance information are also described.
TUMOR NECROSIS FACTOR ALPHA

RELATED APPLICATIONS

[0001] This application is a Stanton, U.S. application Ser. No. 09/950,749, filed Jun. 8, 2000, which is a continuation in part of Stanton, U.S. application Ser. No. 09/495,780, filed Feb. 1, 2000, entitled GENE SEQUENCE VARIATIONS WITH UTILITY IN DETERMINING THE TREATMENT OF INFLAMMATORY OR IMMUNOLOGIC DISEASE, which is a continuation in part of Stanton, International Application No. PCT/US00/01392, filed Jan. 20, 2000, entitled GENE SEQUENCE VARIATIONS WITH UTILITY IN DETERMINING THE TREATMENT OF INFLAMMATORY OR IMMUNOLOGIC DISEASE, which is a continuation in part of Stanton, U.S. application Ser. No. 09/427,835, filed Oct. 26, 1999, entitled GENE SEQUENCE VARIATIONS WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE, and Stanton and Adams, U.S. application Ser. No. 60/300,747, filed Apr. 26, 1999, entitled GENE SEQUENCE VARIATIONS WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE, and claims the benefit of U.S. Provisional Patent Application, Stanton & Adams, Ser. No. 60/131,334, filed Apr. 26, 1999, entitled GENE SEQUENCE VARIATIONS WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE, which are hereby incorporated by reference in their entireties, including drawings and tables.

BACKGROUND OF THE INVENTION

[0002] This application concerns the field of mammalian therapeutics and the selection of therapeutic regimens utilizing host genetic information, including gene sequence variances within the human genome in human populations.

[0003] The information provided in the Background of the Invention is not admitted to be prior art to the present invention but is provided solely to assist the understanding of the reader.

[0004] Many drugs or treatments have highly variable safety and efficacy effects on different individuals. Due to such variability, a given drug or treatment may be effective in one individual but ineffective or ill tolerated in another individual. Thus, administering such drugs to individuals, who would not benefit from such a drug in another individual, would result in wasted cost and time. It could also directly worsen the patient’s condition and even cause the patient’s death.

[0005] For some drugs, variances in individual reaction, measured in selected pharmacokinetic parameters, has been shown to be inheritable by over 50%. For a limited number of drugs, variances in DNA sequences of specific genes involved in drug action and metabolism have been identified. These variances have been shown to vary significantly from individual to individual. The determination of genetic variance in pharmacogenetics is an important issue for the selection of drugs in the clinic. Methods for identifying genetic variances and determining their utility in the selection of optimal therapy for specific patients are described.

[0006] In this application, we address the difficulties in treating inflammatory diseases and other diseases in which modulation of immunologic function provides the basis for therapeutic intervention, including for example, diseases treated with antinflammatory, analgesic or antipyretic drugs as well as autacoids, eicosanoids, interleukins, cytokines or their agonists or antagonists. Diseases or conditions involving the inflammatory response or the immune system constitute a complex and heterogeneous group of diseases, involving all organ systems from the central nervous system and the circulatory system to the viscera and skin. These diseases may be acute or chronic, may progress from an acute stage to a chronic condition, or may exhibit a waxing and waning pattern of flare ups and remissions. Due to their wide anatomical distribution, this group of diseases can (collectively) lead to impairment of a wide range of essential physiological functions. The unifying theme in the treatment of these diseases is the modulation of inflammatory mediators or immune function. The evaluation of long term response to therapy is, for many of these diseases, the most important index of treatment efficacy, due to the progressive nature of inflammatory or immunological diseases. Since it is often difficult to assess the long-term effects of treatment over a short observation period (particularly for diseases with a waxing and waning pattern) there is considerable utility in a genetic test that can predict long term outcomes. Many treatments for diseases with significant inflammatory or immunological components are quite costly. Again, the development of a genetic test that can be used to identify a high responder group may change the economics of treating a population with an expensive therapy. Therapeutics that modulate immunologic or inflammatory responses also frequently pose significant risks for patients. Thus, a test that would allow more judicious use of potentially harmful compounds or biologicals on patients likely to suffer side effects, and/or those patients unlikely to benefit from treatment, would have considerable use both in drug development and in effective use of approved treatments. As healthcare becomes increasingly costly, the ability to allocate healthcare resources effectively becomes more urgent, and methods that lead to safer and more economical use of medicines will contribute to more effective use of healthcare resources.

SUMMARY OF THE INVENTION

[0007] The present invention is concerned generally with the field of identifying an appropriate treatment regimen for an inflammatory disease (or a disease in which modulation of the inflammatory response or the immune system is being tested for therapeutic effect) based upon genotype in mammals, particularly in humans. It is further concerned with the genetic basis of inter-patient variation in response to therapy, including drug therapy. Specifically, this invention describes the identification of gene sequence variances useful in the field of therapeutics for optimizing efficacy and safety of drug therapy. This will be accomplished by establishing diagnostic tests for variances and demonstrating their value in the development, marketing, and use of pharmaceutical products in the clinic. Methods for identifying genetic variances and determining their utility in the selection of optimal therapy for specific patients are described. In general, the invention relates to methods for identifying genetically defined patient subsets that respond to drug therapy differently from other subsets or controls.

[0008] The inventors have determined that the identification of gene sequence variances in genes that may affect
response to therapeutic interventions directed against diseases in which there is abnormal immune/inflammatory response can be exploited to improve therapeutic outcomes for such diseases. Such variances can be used, for example, to identify patients in whom specific therapeutic interventions are likely to be efficacious, well tolerated, and safe. Methods are described in this application for determining whether genetic variances account for variable drug efficacy and safety and for determining whether a given drug or other therapy may be safe and effective in a class of patients with a particular genotype. Methods are also described for developing diagnostic tests so that pharmacogenetic tests can be performed in the care of individual patients. Also provided in this invention are identifications of genes and sequence variances which can be useful in connection with predicting differences in response to treatment and selection of appropriate treatment of a disease or condition. A target gene and variances have utility in pharmacogenetic association studies and diagnostic tests to improve the use of certain drugs or other therapies including, but not limited to, the drug classes and specific drugs identified in the 1999 Physicians' Desk Reference (53rd edition), Medical Economics Data, 1998, or the 1995 United States Pharmacopeia XVIII National Formulary XVIII, Interpharm Press, 1994, Examples 1 through 3 or other sources as described below. [0009] The terms “disease” or “condition” are commonly recognized in the art and designate the presence of signs and/or symptoms in an individual or patient that are generally recognized as abnormal. Diseases or conditions may be diagnosed and categorized based on pathological changes. Signs may include any objective evidence of a disease such as changes that are evident by physical examination of a patient or the results of diagnostic tests which may include, among others, laboratory tests to determine the presence of DNA sequence variances or variant forms of certain genes in a patient. Symptoms are subjective evidence of disease or a patients condition, i.e. the patients perception of an abnormal condition that differs from normal function, sensation, or appearance, which may include, without limitations, physical disabilities, morbidity, pain, and other changes from the normal condition experienced by an individual. Various diseases or conditions include, but are not limited to; those categorized in standard textbooks of medicine including, without limitation, textbooks of nutrition, allopathic, homeopathic, and osteopathic medicine. In certain aspects of this invention, the disease or condition is selected from the group consisting of the types of diseases listed in standard texts such as Harrison’s Principles of Internal Medicine (14th Ed) by Anthony S. Fauci, Eugene Braunwald, Kurt J. Isselbacher, et al. (Editors), McGraw Hill, 1997, or Robbins Pathologic Basis of Disease (6th edition) by Ramzi S. Cotran, Vinay Kumar, Tucke Collins & Stanley L. Robbins, W B Saunders Co., 1998, or other texts described below. This application is directed particularly to diseases in which abnormal function of the immune system or the inflammatory response is part of the disease process, or in which modulation of immune or inflammatory function is being tested as a therapeutic intervention. [0010] In connection with the methods of this invention, unless otherwise indicated, the term “suffering from a disease or condition” means that a person is either presently subject to the signs and symptoms, or is more likely to develop such signs and symptoms than a normal person in the population. Thus, for example, a person suffering from a condition can include a developing fetus, a person subject to a treatment or environmental condition which enhances the likelihood of developing the signs or symptoms of a condition, or a person who is being given or will be given a treatment which increase the likelihood of the person developing a particular condition. For example, tardive dyskinesia is associated with long-term use of anti-psychotics; dyskinesias, paranoid ideation, psychotic episodes and depression have been associated with use of L-dopa in Parkinson’s disease; and dizziness, diplopia, ataxia, sedation, impaired mentation, weight gain, and other undesired effects have been described for various anticonvulsant therapies. Thus, methods of the present invention which relate to treatments of patients (e.g., methods for selecting a treatment, selecting a patient for a treatment, and methods of treating a disease or condition in a patient) can include primary treatments directed to a presently active disease or condition, secondary treatments which are intended to cause a biological effect relevant to a primary treatment, and prophylactic treatments intended to delay, reduce, or prevent the development of a disease or condition, as well as treatments intended to cause the development of a condition different from that which would have been likely to develop in the absence of the treatment. [0011] The term “therapy” refers to a process which is intended to produce a beneficial change in the condition of a mammal, e.g., a human, often referred to as a patient. A beneficial change can, for example, include one or more of: restoration of function, reduction or symptoms, limitation or retardation of progression of a disease, disorder, or condition or prevention, limitation or retardation of deterioration of a patient’s condition, disease or disorder. Such therapy can involve, for example, nutritional modifications, administration of radiation, administration of a drug, behavioral modifications, and combinations of these, among others. [0012] The term “drug” as used herein refers to a chemical entity or biological product, or combination of chemical entities or biological products, administered to a person to treat or prevent or control a disease or condition. The chemical entity or biological product is preferably, but not necessarily a low molecular weight compound, but may also be a larger compound, for example, an oligomer of nucleic acids, amino acids, or carbohydrates including without limitation proteins, oligonucleotides, ribozymes, DNAzymes, glycoproteins, lipoproteins, and modifications and combinations thereof. A biological product is preferably a monoclonal or polyclonal antibody or fragment thereof such as a variable chain fragment; cells; or an agent or product arising from recombinant technology, such as, without limitation, a recombinant protein, recombinant vaccine, or DNA construct developed for therapeutic, e.g., human therapeutic, use. The term “drug” may include, without limitation, compounds that are approved for sale as pharmaceutical products by government regulatory agencies (e.g., U.S. Food and Drug Administration (USFDA or FDA), European Medicines Evaluation Agency (EMEA), and a world regulatory body governing the International Conference of Harmonization (ICH) rules and guidelines), compounds that do not require approval by government regulatory agencies, food additives or supplements including compounds commonly characterized as vitamins, natural products, and completely or incompletely characterized mixtures of chemical entities including natural compounds or purified or partially purified natural products. The term “drug” as used herein is synonym-
mous with the terms “medicine”, “pharmaceutical product”, or “product”. Most preferably the drug is approved by a government agency for treatment of a specific disease or condition.

[0013] A “low molecular weight compound” has a molecular weight <5,000 Da, more preferably <2500 Da, still more preferably <1000 Da, and most preferably <700 Da.

[0014] Those familiar with drug use in medical practice will recognize that regulatory approval for drug use is commonly limited to approved indications, such as to those patients afflicted with a disease or condition for which the drug has been shown to be likely to produce a beneficial effect in a controlled clinical trial. Unfortunately, it has generally not been possible with current knowledge to predict which patients will have a beneficial response, with the exception of certain diseases such as bacterial infections where suitable laboratory methods have been developed. Likewise, it has generally not been possible to determine in advance whether a drug will be safe in a given patient. Regulatory approval for the use of most drugs is limited to the treatment of selected diseases and conditions. The descriptions of approved drug usage, including the suggested diagnostic studies or monitoring studies, and the allowable parameters of such studies, are commonly described in the “label” or “insert” which is distributed with the drug. Such labels or inserts are preferably required by government agencies as a condition for marketing the drug and are listed in common references such as the Physicians Desk Reference (PDR). These and other limitations or considerations on the use of a drug are also found in medical journals, publications such as pharmacology, pharmacy or medical textbooks including, without limitation, textbooks of nutrition, allopathic, homeopathic, and osteopathic medicine.

[0015] Many widely used drugs are effective in a minority of patients receiving the drug, particularly when one controls for the placebo effect. For example, the PDR shows that about 45% of patients receiving Cognex (tacrine hydrochloride) for Alzheimer’s disease show no change or minimal worsening of their disease, as do about 68% of controls (including about 5% of controls who were much worse). About 58% of Alzheimer’s patients receiving Cognex were minimally improved, compared to about 33% of controls, while about 2% of patients receiving Cognex were much improved compared to about 1% of controls. Thus a tiny fraction of patients had a significant benefit. Response to many cancer chemotherapy drugs is even worse. For example, 5-fluorouracil is standard therapy for advanced colorectal cancer, but only about 20-40% of patients have an objective response to the drug, and, of these, only 1-5% of patients have a complete response (complete tumor disappearance; the remaining patients have only partial tumor shrinkage). Conversely, up to 20-30% of patients receiving 5-FU suffer serious gastrointestinal or hematopoietic toxicity, depending on the regimen.

[0016] Thus, in a first aspect, the invention provides a method for selecting a treatment for a patient suffering from a disease or condition by determining whether or not a gene or genes in cells of the patient (in some cases including both normal and disease cells, such as cancer cells) contain at least one sequence variance which is indicative of the effectiveness of the treatment of the disease or condition. The gene or genes are specified herein, in Tables 1, and 3. Preferably the at least one variance includes a plurality of variances which may provide a haplotype or haplotypes. Preferably the joint presence of the plurality of variances is indicative of the potential effectiveness or safety of the treatment in a patient having such plurality of variances. The plurality of variances may each be indicative of the potential effectiveness of the treatment, and the effects of the individual variances may be independent or additive, or the plurality of variances may be indicative of the potential effectiveness if at least 2, 3, 4, or more appear jointly. The plurality of variances may also be combinations of these relationships. The plurality of variances may include variances from one, two, three or more gene loci.

[0017] In preferred embodiments of aspects of the invention involving genes relating to inflammatory or immunological conditions the gene product is involved in a function as described in the Background of the Invention or otherwise described herein.

[0018] In some cases, the selection of a method of treatment, i.e., a therapeutic regimen, may incorporate selection of one or more from a plurality of medical therapies. Thus, the selection may be the selection of a method or methods which is/are more effective or less effective than certain other therapeutic regimens (with either having varying safety parameters). Likewise or in combination with the preceding selection, the selection may be the selection of a method or methods, which is safer than certain other methods of treatment in the patient.

[0019] The selection may involve either positive selection or negative selection or both, meaning that the selection can involve a choice that a particular method would be an appropriate method to use and/or a choice that a particular method would be an inappropriate method to use. Thus, in certain embodiments, the presence of the at least one variance is indicative that the treatment will be effective or otherwise beneficial (or more likely to be beneficial) in the patient. Stating that the treatment will be effective means that the probability of beneficial therapeutic effect is greater than in a person not having the appropriate presence or absence of particular variances. In other embodiments the presence of the at least one variance is indicative that the treatment will be ineffective or contra-indicated for the patient. For example, a treatment may be contra-indicated if the treatment results, or is more likely to result, in undesirable side effects, or an excessive level of undesirable side effects. A determination of what constitutes excessive side effects will vary, for example, depending on the disease or condition being treated, the availability of alternatives, the expected or experienced efficacy of the treatment, and the tolerance of the patient. As for an effective treatment, this means that it is more likely that desired effect will result from the treatment administration in a patient with a particular variance or variances than in a patient who has a different variance or variances. Also in preferred embodiments, the presence of the at least one variance is indicative that the treatment is both effective and unlikely to result in undesirable effects or outcomes, or vice versa (is likely to have undesirable side effects but unlikely to produce desired therapeutic effects).

[0020] In reference to response to a treatment, the term “tolerance” refers to the ability of a patient to accept a
treatment, based, e.g., on deleterious effects and/or effects on lifestyle. Frequently, the term principally concerns the patients perceived magnitude of deleterious effects such as nausea, weakness, dizziness, and diarrhea, among others. Such experienced effects can, for example, be due to general or cell-specific toxicity, activity on non-target cells, cross-reactivity on non-target cellular constituents (non-mechanism based), and/or side effects of activity on the target cellular substituents (mechanism based), or the cause of toxicity may not be understood. In any of these circumstances one may identify an association between the undesirable effects and variances in specific genes.

[0021] Adverse responses to drugs constitute a major medical problem, as shown in two recent meta-analyses (Lazaro et al., Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies, JAMA 279:1200-1205, 1998; Bonn, Adverse drug reactions remain a major cause of death, Lancet 351:1183, 1998). An estimated 2.2 million hospitalized patients in the United States had serious adverse drug reactions in 1994, with an estimated 106,000 deaths (Lazaro et al.). To the extent that some of these adverse events are due to genetically encoded biochemical diversity among patients in pathways that effect drug action, the identification of variances that are predictive of such effects will allow for more effective and safer drug use.

[0022] In embodiments of this invention, the variance or variant form or forms of a gene is/are associated with a specific response to a drug. The frequency of a specific variance or variant form of the gene may correspond to the frequency of an efficacious response to administration of a drug. Alternatively, the frequency of a specific variance or variant form of the gene may correspond to the frequency of an adverse event resulting from administration of a drug. Alternatively the frequency of a specific variance or variant form of a gene may correspond closely with the frequency of a beneficial or adverse response, yet the variance may still be useful for identifying a patient subset with high response or toxicity incidence because the variance may account for only a fraction of the patients with high response or toxicity. In such a case the preferred course of action is identification of a second or third or additional variances that permit identification of the patient groups not usefully identified by the first variance. Preferably, the drug will be effective in more than 20% of individuals with one or more specific variances or variant forms of the gene, more preferably in 40% and most preferably in >60%. In other embodiments, the drug will be toxic or create clinically unacceptable side effects in more than 10% of individuals with one or more variances or variant forms of the gene, more preferably in >30%, more preferably in >50%, and most preferably in >70% or in more than 90%.

[0023] Also in other embodiments, the method of selecting a treatment includes eliminating a treatment, where the presence or absence of the at least one variance is indicative that the treatment will be ineffective or contra-indicated, e.g., would result in excessive weight gain. In other preferred embodiments, in cases in which undesirable side-effects may occur or are expected to occur from a particular therapeutic treatment, the selection of a method of treatment can include identifying both a first and second treatment, where the first treatment is effective to treat the disease or condition, and the second treatment reduces a deleterious effect of the first treatment.

[0024] The phrase “eliminating a treatment” refers to removing a possible treatment from consideration, e.g., for use with a particular patient based on the presence or absence of a particular variance(s) in one or more genes in cells of that patient, or to stopping the administration of a treatment which was in the course of administration.

[0025] Usually, the treatment will involve the administration of a compound preferentially active or safe in patients with a form or forms of a gene, where the gene is one identified herein. The administration may involve a combination of compounds. Thus, in preferred embodiments, the method involves identifying such an active compound or combination of compounds, where the compound is less active or is less safe or both when administered to a patient having a different form of the gene.

[0026] Also in preferred embodiments, the method of selecting a treatment involves selecting a method of administration of a compound, combination of compounds, or pharmaceutical composition, for example, selecting a suitable dosage level and/or frequency of administration, and/or mode of administration of a compound. The method of administration can be selected to provide better, preferably maximum therapeutic benefit. In this context, “maximum” refers to an approximate local maximum based on the parameters being considered, not an absolute maximum.

[0027] Also in this context, a “suitable dosage level” refers to a dosage level which provides a therapeutically reasonable balance between pharmacological effectiveness and deleterious effects. Often this dosage level is related to the peak or average serum levels resulting from administration of a drug at the particular dosage level.

[0028] Similarly, a “frequency of administration” refers to how often in a specified time period a treatment is administered, e.g., once, twice, or three times per day, every other day, once per week, etc. For a drug or drugs, the frequency of administration is generally selected to achieve a pharmacologically effective average or peak serum level without excessive deleterious effects (and preferably while still being able to have reasonable patient compliance for self-administered drugs). Thus, it is desirable to maintain the serum level of the drug within a therapeutic window of concentrations for the greatest percentage of time possible without such deleterious effects as would cause a prudent physician to reduce the frequency of administration for a particular dosage level.

[0029] A particular gene or genes can be relevant to the treatment of more than one disease or condition, for example, the gene or genes can have a role in the initiation, development, course, treatment, treatment outcomes, or health-related quality of life outcomes of a number of different diseases, disorders, or conditions. Thus, in preferred embodiments, the disease or condition or treatment of the disease or condition is any which involves a gene from the gene list described herein as Tables 1 and 3.

[0030] Determining the presence of a particular variance or plurality of variances in a particular gene in a patient can be performed in a variety of ways. In preferred embodiments, the detection of the presence or absence of at least
one variance involves amplifying a segment of nucleic acid including at least one of the at least one variances. Preferably a segment of nucleic acid to be amplified is 500 nucleotides or less in length, more preferably 100 nucleotides or less, and most preferably 45 nucleotides or less. Also, preferably the amplified segment or segments includes a plurality of variances, or a plurality of segments of a gene or of a plurality of genes.

[0031] In another aspect determining the presence of a set of variances in a specific gene related to treatment of inflammation or immune disease disease or other related genes, or genes listed in Tables 1 and 3, may entail a haplotyping test that requires allele specific amplification of a large DNA segment of no greater than 25,000 nucleotides, preferably no greater than 10,000 nucleotides and most preferably no greater than 5,000 nucleotides. Alternatively one allele may be enriched by methods other than amplification prior to determining genotypes at specific variant positions on the enriched allele as a way of determining haplotypes. Preferably the determination of the presence or absence of a haplotype involves determining the sequence of the variant sites by methods such as chain terminating DNA sequencing or minisequencing, or by oligonucleotide hybridization or by mass spectrometry.

[0032] The term “genotype” in the context of this invention refers to the alleles present in DNA from a subject or patient, where an allele can be defined by the particular nucleotide(s) present in a nucleic acid sequence at a particular site(s). Often a genotype is the nucleotide(s) present at a single polymorphic site known to vary in the human population.

[0033] In preferred embodiments, the detection of the presence or absence of the at least one variance involves contacting a nucleic acid sequence corresponding to one of the genes identified above or a product of such a gene with a probe. The probe is able to distinguish a particular form of the gene or gene product or the presence or a particular variance or variances, e.g., by differential binding or hybridization. Thus, exemplary probes include nucleic acid hybridization probes, peptide nucleic acid probes, nucleotide-containing probes which also contain at least one nucleotide analog, and antibodies, e.g., monoclonal antibodies, and other probes as discussed herein. Those skilled in the art are familiar with the preparation of probes with particular specificities. Those skilled in the art will recognize that a variety of variables can be adjusted to optimize the discrimination between two variant forms of a gene, including changes in salt concentration, temperature, pH and addition of various compounds that affect the differential affinity of GC vs. AT base pairs, such as tetramethyl ammonium chloride. (See Current Protocols in Molecular Biology by F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. D. Seidman, K. Struhl, and V. B. Chanda (editors, John Wiley & Sons.)

[0034] In other preferred embodiments, determining the presence or absence of the at least one variance involves sequencing at least one nucleic acid sample. The sequencing involves sequencing of a portion or portions of a gene and/or portions of a plurality of genes which includes at least one variance site, and may include a plurality of such sites. Preferably, the portion is 500 nucleotides or less in length, more preferably 100 nucleotides or less, and most preferably 45 nucleotides or less in length. Such sequencing can be carried out by various methods recognized by those skilled in the art, including use of dye terminations methods (e.g., using dye-labeled dye nucleotides) and the use of mass spectrometric methods. In addition, mass spectrometric methods may be used to determine the nucleotide present at a variance site. In preferred embodiments in which a plurality of variances is determined, the plurality of variances can constitute a haplotype or collection of haplotypes. Preferably the methods for determining genotypes or haplotypes are designed to be sensitive to all the common genotypes or haplotypes present in the population being studied (for example, a clinical trial population).

[0035] The terms “variant form of a gene”, “form of a gene”, or “allele” refer to one specific form of a gene in a population, the specific form differing from other forms of the same gene in the sequence of at least one, and preferably more than one, variant sites within the sequence of the gene. The sequences at these variant sites that differ between different alleles of the gene are termed “gene sequence variances” or “variances” or “variants”. The term “alternative form” refers to an allele that can be distinguished from other alleles by having distinct variances at least one, and preferably more than one, variant sites within the gene sequence. Other terms known in the art to be equivalent include mutation and polymorphism, although mutation is often used to refer to an allele associated with a deleterious phenotype. In preferred aspects of this invention, the variances are selected from the group consisting of the variances listed in the variance tables herein or in a patent or patent application referenced and incorporated by reference in this disclosure. In the methods utilizing variance presence or absence, reference to the presence of a variance or variances means particular variances, i.e., particular nucleotides at particular polymorphic sites, rather than just the presence of any variance in the gene.

[0036] Variances occur in the human genome at approximately one in every 500-1,000 bases within the human genome when two alleles are compared. When multiple alleles from unrelated individuals are compared the density of variant sites increases as different individuals, when compared to a reference sequence, will often have sequence variances at different sites. At most variant sites there are only two alternative nucleotides involving the substitution of one base for another or the insertion/deletion of one or more nucleotides. Within a gene there may be several variant sites. Variant forms of the gene or alternative alleles can be distinguished by the presence of alternative variances at a single variant site, or a combination of several different variances at different sites (haplotypes).

[0037] It is estimated that there are 3,300,000,000 bases in the sequence of a single haploid human genome. All human cells except germ cells are normally diploid. Each gene in the genome may span 100-1,000,000 bases of DNA sequence or 100-20,000 bases of mRNA. It is estimated that there are between 60,000 and 150,000 genes in the human genome. The “identification” of genetic variances or variant forms of a gene involves the discovery of variances that are present in a population. The identification of variances is needed for development of a diagnostic test to determine whether a patient has a variant form of a gene that is known to be associated with a disease, condition, or predisposition or with the efficacy or safety of the drug. Identification of
previously undiscovered genetic variances is distinct from the process of “determining” the status of known variances by a diagnostic test (often referred to as genotyping). The present invention provides exemplary variances in genes listed in the gene tables, as well as methods for discovering additional variances in those genes and a comprehensive written description of such additional possible variances. Also described are methods for DNA diagnostic tests to determine the DNA sequence at a particular variant site or sites.

[0038] The process of “identifying” or discovering new variances involves comparing the sequence of at least two alleles of a gene, more preferably at least 10 alleles and most preferably at least 50 alleles (keeping in mind that each somatic cell has two alleles). The analysis of large numbers of individuals to discover variances in the gene sequence between individuals in a population will result in detection of a greater fraction of all the variances in the population. Preferably the process of identifying reveals whether there is a variance within the gene; more preferably identifying reveals the location of the variance within the gene; more preferably identifying provides knowledge of the sequence of the nucleic acid sequence of the variance, and most preferably identifying provides knowledge of the combination of different variances that comprise specific variant forms of the gene or alleles (referred to as alleles). In identifying new variances it is often useful to screen different population groups based on racial, ethnic, gender, and/or geographic origin because particular variances may differ in frequency between such groups. It may also be useful to screen DNA from individuals with a particular disease or condition of interest because they may have a higher frequency of certain variances than the general population.

[0039] The process of genotyping involves using diagnostic tests for specific variances that have already been identified. It will be apparent that such diagnostic tests can only be performed after variances and variant forms of the gene have been identified. Identification of new variances can be accomplished by a variety of methods, alone or in combination, including, for example, DNA sequencing, SSCP, heteroduplex analysis, denaturing gradient gel electrophoresis (DGGE), heteroduplex cleavage (either enzymatic as with T4 Endonuclease 7, or chemical as with osmium tetroxide and hydroxylamine), computational methods (described herein), and other methods described herein as well as others known to those skilled in the art. (See, for example: Cotton, R. G. H., Slowly but surely towards better scanning for mutations, Trends in Genetics 13(2): 43-6, 1997 or Current Protocols in Human Genetics by N. C. Dracoli, J. L. Haines, B. R. Korf, D. T. Moir, C. C. Morton, C. E. Seidman, D. R. Smith, and A. Boyle (editors), John Wiley & Sons.) In the context of this invention, the term “analyzing a sequence” refers to determining at least some sequence information about the sequence, e.g., determining the nucleotides present at a particular site or sites in the sequence, particularly sites that are known to vary in a population, or determining the base sequence of all or of a portion of the particular sequence.

[0040] In the context of this invention, the term “haplotype” refers to a cis arrangement of two or more polymorphic nucleotides, i.e., variances, on a particular chromosome, e.g., in a particular gene. The haplotype preserves information about the phase of the polymorphic nucleotides—that is, which set of variances were inherited from one parent, and which from the other. A genotyping test does not provide information about phase. For example, an individual heterozygous at nucleotide 25 of a gene (both A and C are present) and also at nucleotide 100 (both G and T are present) could have haplotypes 25A-100G and 25C-100T, or alternatively 25A-100T and 25C-100G. Only a haplotyping test can discriminate these two cases definitively.

[0041] The terms “variances,” “variants” and “polymorphisms,” as used herein, may also refer to a set of variances, haplotypes or a mixture of the two. Further, the term variance, variant or polymorphism (singular), as used herein, also encompasses a haplotype. This usage is intended to minimize the need for cumbersome phrases such as: “. . . measure correlation between drug response and a variance, variances, haplotype, haplotypes or a combination of variances and haplotypes . . . .”, throughout the application. Instead, the italicized text in the foregoing sentence can be represented by the word “variance”, “variant” or “polymorphism”. Similarly, the term genotype, as used herein, means a procedure for determining the status of one or more variances in a gene, including a set of variances comprising a haplotype. Thus phrases such as “. . . genotype a patient . . .” refer to determining the status of one or more variances, including a set of variances for which phase is known (i.e. a haplotype).

[0042] In preferred embodiments of this invention, the frequency of the variance or variant form of the gene in a population is known. Measures of frequency known in the art include “allele frequency”, namely the fraction of genes in a population that have one specific variance or set of variances. The allele frequencies for any gene should sum to 1. Another measure of frequency known in the art is the “heterozygote frequency” namely, the fraction of individuals in a population who carry two alleles, or two forms of a particular variance or variant form of a gene, one inherited from each parent. Alternatively, the number of individuals who are homozygous for a particular form of a gene may be a useful measure. The relationship between allele frequency, heterozygote frequency, and homozygote frequency is described for many genes by the Hardy-Weinberg equation, which provides the relationship between allele frequency, heterozygote frequency and homozygote frequency in a freely breeding population at equilibrium. Most human variances are substantially in Hardy-Weinberg equilibrium. In a preferred aspect of this invention, the allele frequency, heterozygote frequency, and homozygote frequencies are determined experimentally. Preferably a variance has an allele frequency of at least 0.01, more preferably at least 0.05, still more preferably at least 0.10. However, the allele may have a frequency as low as 0.001 if the associated phenotype is, for example, a rare form of toxic reaction to a treatment or drug. Beneficial responses may also be rare.

[0043] In this regard, “population” refers to a defined group of individuals or a group of individuals with a particular disease or condition or individuals that may be treated with a specific drug identified by, but not limited to geographic, ethnic, race, gender, and/or cultural indices. In most cases a population will preferably encompass at least ten thousand, one hundred thousand, one million, ten million, or more individuals, with the larger numbers being more preferable. In a preferred aspect of this invention, the population refers to individuals with a specific disease or
condition that may be treated with a specific drug. In an aspect of this invention, the allele frequency, heterozygote frequency, or homozygote frequency of a specific variance or variant form of a gene is known. In preferred embodiments of this invention, the frequency of one or more variances that may predict response to a treatment is determined in one or more populations using a diagnostic test.

[0044] It should be emphasized that it is currently not generally practical to study an entire population to establish the association between a specific disease or condition or response to a treatment and a specific variance or variant form of a gene. Such studies are preferably performed in controlled clinical trials using a limited number of patients that are considered to be representative of the population with the disease. Since drug development programs are generally targeted at the largest possible population, the study population will generally comprise men and women, as well as members of various racial and ethnic groups, depending on where the clinical trial is being performed. This is important to establish the efficacy of the treatment in all segments of the population.

[0045] In the context of this invention, the term “probe” refers to a molecule which detectably distinguishes between target molecules differing in structure. Detection can be accomplished in a variety of different ways depending on the type of probe used and the type of target molecule. Thus, for example, detection may be based on discrimination of activity levels of the target molecule, but preferably is based on detection of specific binding. Examples of such specific binding include antibody binding and nucleic acid probe hybridization. Thus, for example, probes can include enzyme substrates, antibodies and antibody fragments, and nucleic acid hybridization probes. Thus, in preferred embodiments, the detection of the presence or absence of the at least one variance involves contacting a nucleic acid sequence which includes a variance site with a probe, preferably a nucleic acid probe, where the probe preferentially hybridizes with a form of the nucleic acid sequence containing a complementary base at the variance site as compared to hybridization to a form of the nucleic acid sequence having a non-complementary base at the variance site, where the hybridization is carried out under selective hybridization conditions. Such a nucleic acid hybridization probe may span two or more variance sites. Unless otherwise specified, a nucleic acid probe can include one or more nucleic acid analogs, labels or other substituents or moieties so long as the base-pairing function is retained.

[0046] As is generally understood, administration of a particular treatment, e.g., administration of a therapeutic compound or combination of compounds, is chosen depending on the disease or condition which is to be treated. Thus, in certain preferred embodiments, the disease or condition is one for which administration of a treatment is expected to provide a therapeutic benefit; in certain embodiments, the compound is a compound identified herein, e.g., in a drug table (Tables 5-13, 16, 17, 18).

[0047] As used herein, the terms “effective” and “effectiveness” includes both pharmacological effectiveness and physiological safety. Pharmacological effectiveness refers to the ability of the treatment to result in a desired biological effect in the patient. Physiological safety refers to the level of toxicity, or other adverse physiological effects at the cellular, organ and/or organism level (often referred to as side-effects) resulting from administration of the treatment. On the other hand, the term “ineffective” indicates that a treatment does not provide sufficient pharmacological effect to be therapeutically useful, even in the absence of deleterious effects, at least in the unstratified population. (Such a treatment may be ineffective in a subgroup that can be identified by the presence of one or more sequence variances or alleles.) “Less effective” means that the treatment results in a therapeutically significant lower level of pharmacological effectiveness and/or a therapeutically greater level of adverse physiological effects, e.g., greater liver toxicity.

[0048] Thus, in connection with the administration of a drug, a drug which is “effective against” a disease or condition indicates that administration in a clinically appropriate manner results in a beneficial effect for at least a statistically significant fraction of patients, such as an improvement of symptoms, a cure, a reduction of disease load, reduction in tumor mass or cell numbers, extension of life, improvement in quality of life, or other effect generally recognized as positive by medical doctors familiar with treating the particular type of disease or condition.

[0049] Effectiveness is measured in a particular population. In conventional drug development the population is generally every subject who meets the enrollment criteria (i.e. has the particular form of the disease or condition being treated). It is an aspect of the present invention that segmentation of a study population by genetic criteria can provide the basis for identifying a subpopulation in which a drug is effective against the disease or condition being treated.

[0050] The term “deleterious effects” refers to physical effects in a patient caused by administration of a treatment which are regarded as medically undesirable. Thus, for example, deleterious effects can include a wide spectrum of toxic effects injurious to health such as death of normally functioning cells when only death of diseased cells is desired, nausea, fever, inability to retain food, dehydration, damage to critical organs such as arrhythmias, renal tubular necrosis, fatty liver, or pulmonary fibrosis leading to coronary, renal, hepatic, or pulmonary insufficiency among many others. In this regard, the term “contra-indicated” means that a treatment results in deleterious effects such that a prudent medical doctor treating such a patient would regard the treatment as unsuitable for administration. Major factors in such a determination can include, for example, availability and relative advantages of alternative treatments, consequences of non-treatment, and permanency of deleterious effects of the treatment.

[0051] It is recognized that many treatment methods, e.g., administration of certain compounds or combinations of compounds, may produce side-effects or other deleterious effects in patients. Such effects can limit or even preclude use of the treatment method in particular patients, or may even result in irreversible injury, dysfunction, or death of the patient. Thus, in certain embodiments, the variance information is used to select both a first method of treatment and a second method of treatment. Usually the first treatment is a primary treatment which provides a physiological effect directed against the disease or condition or its symptoms. The second method is directed to reducing or eliminating one or more deleterious effects of the first treatment, e.g., to
reduce a general toxicity or to reduce a side effect of the primary treatment. Thus, for example, the second method can be used to allow use of a greater dose or duration of the first treatment, or to allow use of the first treatment in patients for whom the first treatment would not be tolerated or would be contra-indicated in the absence of a second method to reduce deleterious effects or to potentiate the effectiveness of the first treatment.

[0052] In a related aspect, the invention provides a method for selecting a method of treatment for a patient suffering from a disease or condition by comparing at least one variance in at least one gene in the patient, with a list of variances in the gene from Tables 1 and 3, or other gene related to inflammatory or immunological disease, which are indicative of the effectiveness of at least one method of treatment. Preferably the comparison involves a plurality of variances or a haplotype indicative of the effectiveness of at least one method of treatment. Also, preferably the list of variances includes a plurality of variances.

[0053] Similar to the above aspect, in preferred embodiments the at least one method of treatment involves the administration of a compound effective in at least some patients with a disease or condition; the presence or absence of the at least one variance is indicative that the treatment will be effective in the patient; and/or the presence or absence of the at least one variance is indicative that the treatment will be ineffective or contra-indicated in the patient; and/or the treatment is a first treatment and the presence or absence of the at least one variance is indicative that a second treatment will be beneficial to reduce a deleterious effect or to potentiate the effectiveness of the first treatment; and/or the at least one treatment is a plurality of methods of treatment. For a plurality of treatments, preferably the selecting involves determining whether any of the methods of treatment will be more effective than at least one other of the plurality of methods of treatment. Yet other embodiments are provided as described for the preceding aspect in connection with methods of treatment using administration of a compound; treatment of various diseases, and variances in particular genes.

[0054] In the context of variance information in the methods of this invention, the term “list” refers to one or more variances which have been identified for a gene of potential importance in accounting for inter-individual variation in treatment response. Preferably there is a plurality of variances for the gene, preferably a plurality of variances for the particular gene. Preferably, the list is recorded in written or electronic form. For example, identified variances of identified genes are recorded for some of the genes in Tables 1 and 3, additional variances for genes in Table 1 are provided in Table 1 of Stanton & Adams, U.S. application Ser. No. 09/300,747, supra, and additional gene variance identification tables are provided in a form which allows comparison with other variance information. The possible additional variances in the identified genes are provided in Table 3 in Stanton & Adams, U.S. application Ser. No. 09/300,747, supra.

[0055] In addition to the basic method of treatment, often the mode of administration of a given compound as a treatment for a disease or condition in a patient is significant in determining the course and/or outcome of the treatment for the patient. Thus, the invention also provides a method for selecting a method of administration of a compound to a patient suffering from a disease or condition, by determining the presence or absence of at least one variance in cells of the patient at least one identified gene from Tables 1 and 3, where such presence or absence is indicative of an appropriate method of administration of the compound. Preferably, the selection of a method of treatment (a treatment regimen) involves selecting a dosage level or frequency of administration or route of administration of the compound or combinations of those parameters. In preferred embodiments, two or more compounds are to be administered, and the selecting involves selecting a method of administration for one, two, or more than two of the compounds, jointly, concurrently, or separately. As understood by those skilled in the art, such plurality of compounds may be used in combination therapy, and thus may be formulated in a single drug, or may be separate drugs administered concurrently, serially, or separately. Other embodiments are as indicated above for selection of second treatment methods, methods of identifying variances, and methods of treatment as described for aspects above.

[0056] In another aspect, the invention provides a method for selecting a patient for administration of a method of treatment for a disease or condition, or of selecting a patient for a method of administration of a treatment, by comparing the presence or absence of at least one variance in a gene as identified above in cells of a patient, with a list of variances in the gene, where the presence or absence of the at least one variance is indicative that the treatment or method of administration will be effective in the patient. If the at least one variance is present in the patient’s cells, then the patient is selected for administration of the treatment.

[0057] In preferred embodiments, the disease or the method of treatment is as described in aspects above, specifically including, for example, those described for selecting a method of treatment.

[0058] In another aspect, the invention provides a method for identifying a subset of patients with enhanced or diminished response or tolerance to a treatment method or a method of administration of a treatment where the treatment is for a disease or condition in the patient. The method involves correlating one or more variances in one or more genes as identified in aspects above in a plurality of patients with response to a treatment or a method of administration of a treatment. The correlation may be performed by determining the one or more variances in the one or more genes in the plurality of patients and correlating the presence or absence of each of the variances (alone or in various combinations) with the patient’s response to treatment. The variances may be previously known to exist or may also be determined in the present method or combinations of prior information and newly determined information may be used. The enhanced or diminished response should be statistically significant, preferably such that p=0.10 or less, more preferably 0.05 or less, and most preferably 0.02 or less. A positive correlation between the presence of one or more variances and an enhanced response to treatment is indicative that the treatment is particularly effective in the group of patients having those variances. A positive correlation of the presence of the one or more variances with a diminished response to the treatment is indicative that the treatment will be less effective in the group of patients having those variances. Such information is useful, for example, for
selecting or de-selecting patients for a particular treatment or method of administration of a treatment, or for demonstrating that a group of patients exists for which the treatment or method of treatment would be particularly beneficial or contra-indicated. Such demonstration can be beneficial, for example, for obtaining government regulatory approval for a new drug or a new use of a drug.

In preferred embodiments, the variances are in at least one of the identified genes listed on Tables 1 and 3, or are particular variances described herein. Also, preferred embodiments include drugs, treatments, variance identification or determination, determination of effectiveness, and/or diseases as described for aspects above or otherwise described herein.

In preferred embodiments, the correlation of patient responses to therapy according to patient genotype is carried out in a clinical trial, e.g., as described herein according to any of the variations described. Detailed description of methods for associating variances with clinical outcomes using clinical trials are provided below. Further, in preferred embodiments the correlation of pharmacological effect (positive or negative) to treatment response according to genotype or haplotype in such a clinical trial is part of a regulatory submission to a government agency leading to approval of the drug. Most preferably the compound or compounds would not be approachable in the absence of the genetic information allowing identification of an optimal responder population.

As indicated above, in aspects of this invention involving selection of a patient for a treatment, selection of a method or mode of administration of a treatment, and selection of a patient for a treatment or a method of treatment, the selection may be positive selection or negative selection. Thus, the methods can include eliminating a treatment for a patient, eliminating a method or mode of administration of a treatment to a patient, or elimination of a patient for a treatment or method of treatment.

Also, in methods involving identification and/or comparison of variances present in a gene of a patient, the methods can involve such identification or comparison for a plurality of genes. Preferably, the genes are functionally related to the same disease or condition, or to the aspect of disease pathophysiology that is being subjected to pharmacological manipulation by the treatment (e.g., a drug), or to the activation or inactivation or elimination of the drug, and more preferably the genes are involved in the same biochemical process or pathway.

In another aspect, the invention provides a method for identifying the forms of a gene in an individual, where the gene is one specified as for aspects above, by determining the presence or absence of at least one variance in the gene. In preferred embodiments, the at least one variance includes at least one variance selected from the group of variances identified in variance tables herein. Preferably, the presence or absence of the at least one variance is indicative of the effectiveness of a therapeutic treatment in a patient suffering from a disease or condition and having cells containing the at least one variance.

The presence or absence of the variances can be determined in any of a variety of ways as recognized by those skilled in the art. For example, the nucleotide sequence of at least one nucleic acid sequence which includes at least one variance site (or a complementary sequence) can be determined, such as by chain termination methods, hybridization methods or by mass spectrometric methods. Likewise, in preferred embodiments, the determining involves contacting a nucleic acid sequence or a gene product of one of the sequences with a probe which specifically identifies the presence or absence of a form of the gene. For example, a probe, e.g., a nucleic acid probe, can be used which specifically binds, e.g., hybridizes, to a nucleic acid sequence corresponding to a portion of the gene and which includes at least one variance site under selective binding conditions. As described for other aspects, determining the presence or absence of at least two variances and their relationship on the two gene copies present in a patient can constitute determining a haplotype or haplotypes.

Other preferred embodiments involve variances related to types of treatment, drug responses, diseases, nucleic acid sequences, and other items related to variances and variance determination as described for aspects above.

In yet another aspect, the invention provides a pharmaceutical composition which includes a compound which has a differential effect in patients having at least one copy, or alternatively, two copies of a form of a gene as identified for aspects above and a pharmaceutically acceptable carrier, excipient, or diluent. The composition is adapted to be preferentially effective to treat a patient with cells containing the one, two, or more copies of the form of the gene.

In preferred embodiments of aspects involving pharmaceutical compositions, active compounds, or drugs, the material is subject to a regulatory limitation or restriction on approved uses or indications or is subject to a recommendation, e.g., by the U.S. Food and Drug Administration (FDA), limiting approved use of the composition to patients having at least one copy of the particular form of the gene which contains at least one variance. Alternatively, the composition is subject to a regulatory limitation or restriction on approved uses, or a recommendation to limit use, indicating that the composition is not approved for use or should not be used in patients having at least one copy of a form of the gene including at least one variance. Also in preferred embodiments, the composition is packaged, and the packaging includes a label or insert indicating or suggesting beneficial therapeutic approved use of the composition in patients having one or two copies of a form of the gene including at least one variance. Alternatively, the label or insert limits approved use, or recommends limiting use, of the composition to patients having zero or one or two copies of a form of the gene including at least one variance. The latter embodiment would be likely where the presence of the at least one variance in one or two copies in cells of a patient means that the composition would be ineffective or deleterious to the patient. Also in preferred embodiments, the composition is indicated for use in treatment of a disease or condition which is one of those identified for aspects above. Also in preferred embodiments, the at least one variance includes at least one variance from those identified herein.

The term "packaged" means that the drug, compound, or composition is prepared in a manner suitable for distribution or shipping with a box, vial, pouch, bubble pack, or other protective container, which may also be used in
combination. The packaging may have printing on it and/or printed material may be included in the packaging.

In preferred embodiments, the drug is selected from the drug classes or specific exemplary drugs identified in an example, in a table herein, and is subject to a regulatory limitation or suggestion or warning as described above that limits or suggests limiting approved use to patients having specific variances or variant forms of a gene identified in Examples or in the gene list provided below in order to achieve maximal benefit and avoid toxicity or other deleterious effect.

A pharmaceutical composition can be adapted to be preferentially effective in a variety of ways. In some cases, an active compound is selected which was not previously known to be differentially active, or which was not previously recognized as a therapeutic compound. Alternatively, the compound was previously known as a therapeutic compound, but the composition is formulated in a manner appropriate for administration for treatment of a disease or condition for which a gene of this invention is involved in treatment response, and the active compound had not been formulated appropriately for such use before. For example, a compound may previously have been formulated for topical treatment of a skin condition, but is found to be effective in IV or other internal treatment of a disease identified for this invention. For compounds that are differentially effective on the gene, such alternative formulations are adapted to be preferentially effective. In some cases, the concentration of an active compound which has differential activity can be adjusted such that the composition is appropriate for administration to a patient with the specified variances. For example, the presence of a specified variance may allow or require the administration of a much larger dose, which would not be practical with a previously utilized composition. Conversely, a patient may require a much lower dose, such that administration of such a dose with a prior composition would be impractical or inaccurate. Thus, the composition may be prepared in a higher or lower unit dose form, or prepared in a higher or lower concentration of the active compound or compounds. In yet other cases, the composition can include additional compounds needed to enable administration of a particular active compound in a patient with the specified variances, which was not in previous compositions, e.g., because the majority of patients did not require or benefit from the added component.

The term “differential” or “differentially” generally refers to a statistically significant difference in a property or effect. Preferably, the difference is also functionally significant. Thus, “differential binding or hybridization” is sufficient difference in binding or hybridization to allow discrimination using an appropriate detection technique. Likewise, “differential effect” or “differentially active” in connection with a therapeutic treatment or drug refers to a difference in the level of the effect or activity which is distinguishable using relevant parameters and techniques for measuring the effect or activity being considered. Preferably the difference in effect or activity is also sufficient to be clinically significant, such that a corresponding difference in the course of treatment or treatment outcome would be expected, at least on a statistical basis.

Also usefully provided in the present invention are probes which specifically recognize a nucleic acid sequence corresponding to a variance or variances in a gene as identified in aspects above or a product expressed from the gene, and are able to distinguish a variant form of the sequence or gene or product from one or more other variant forms of that sequence, gene, or gene product under selective conditions. Those skilled in the art recognize and understand the identification or determination of selective conditions for particular probes or types of probes. An exemplary type of probe is a nucleic acid hybridization probe, which will selectively bind under selective binding conditions to a nucleic acid sequence or a gene product corresponding to one or the genes identified for aspects above. Another type of probe is a peptide or protein, e.g., an antibody or antibody fragment which specifically or preferentially binds to a polypeptide expressed from a particular form of a gene as characterized by the presence or absence of at least one variance. Thus, in another aspect, the invention concerns such probes. In the context of this invention, a “probe” is a molecule, commonly a nucleic acid, though also potentially a protein, carbohydrate, polymer, or small molecule, that is capable of binding to one variance or variant form of the gene to a greater extent than to a form of the gene having a different base at one or more variance sites, such that the presence of the variance or variant form of the gene can be determined. Preferably the probe distinguishes at least one variance identified in Examples, tables or lists below or in Tables 1 or 3 of Stuwan & Adams, U.S. application Ser. No. 09/300,747, supra.

In preferred embodiments, the probe is a nucleic acid probe at least 15, preferably at least 17 nucleotides in length, more preferably at least 20 or 22 or 25, preferably 50 or fewer nucleotides in length, more preferably 200 or 100 or fewer, still more preferably 50 or fewer, and most preferably 30 or fewer. In preferred embodiments, the probe has a length in a range between from any one of the above lengths to any other of the above lengths (including endpoints). The probe specifically hybridizes under selective hybridization conditions to a nucleic acid sequence corresponding to a portion of one of the genes identified in connection with above aspects. The nucleic acid sequence includes at least one and preferably two or more variance sites. Also in preferred embodiments, the probe has a detectable label, preferably a fluorescent label. A variety of other detectable labels are known to those skilled in the art. Such a nucleic acid probe can also include one or more nucleic acid analogs.

In preferred embodiments, the probe is an antibody or antibody fragment which specifically binds to a gene product expressed from a form of one of the above genes, where the form of the gene has at least one specific variance with a particular base at the variance site, and preferably a plurality of such variances.

In connection with nucleic acid probe hybridization, the term “specifically hybridizes” indicates that the probe hybridizes to a sufficiently greater degree to the target sequence than to a sequence having a mismatched base at least one variance site to allow distinguishing such hybridization. The term “specifically hybridizes” thus means that the probe hybridizes to the target sequence, and not to non-target sequences, at a level which allows ready identification of probe-target sequence hybridization under selective hybridization conditions. Thus, “selective hybridization conditions” refer to conditions which allow such differential
binding. Similarly, the terms "specifically binds" and "selective binding conditions" refer to such differential binding of any type of probe, e.g., antibody probes, and to the conditions which allow such differential binding. Typically hybridization reactions to determine the status of variant sites in patient samples are carried out with two different probes, one specific for each of the (usually two) possible variant nucleotides. The complementary information derived from the two separate hybridization reactions is useful in corroborating the results.

Likewise, the invention provides an isolated, purified or enriched nucleic acid sequence of 15 to 500 nucleotides in length, preferably 15 to 100 nucleotides in length, more preferably 15 to 50 nucleotides in length, and most preferably 15 to 30 nucleotides in length, which has a sequence which corresponds to a portion of one of the genes identified for aspects above. Preferably the lower limit for the preceding ranges is 17, 20, 22, or 25 nucleotides in length. In other embodiments, the nucleic acid sequence is 30 to 300 nucleotides in length, or 45 to 200 nucleotides in length, or 45 to 100 nucleotides in length. The nucleic acid sequence includes at least one variance site. Such sequences can, for example, be amplification products of a sequence which spans or includes a variance site in a gene identified herein. Likewise, such a sequence can be a primer, or amplification nucleotide which is able to bind to or extend through a variance site in such a gene. Yet another example is a nucleic acid hybridization probe comprised of such a sequence. In such probes, primers, and amplification products, the nucleotide sequence can contain a sequence or site corresponding to a variance site or sites, for example, a variance site identified herein. Preferably the presence or absence of a particular variant form in the heterozygous or homozygous state is indicative of the effectiveness of a method of treatment in a patient.

Likewise, the invention provides a set of primers or amplification oligonucleotides (e.g., 2, 3, 4, 6, 8, 10 or even more) adapted for binding to or extending through at least one gene identified herein. In preferred embodiments the set includes primers or amplifications oligonucleotides adapted to bind to or extend through a plurality of sequence variants in a gene(s) identified herein. The plurality of variants preferably provides a haplotype. Those skilled in the art are familiar with the use of amplification oligonucleotides (e.g., PCR primers) and the appropriate location, testing and use of such oligonucleotides. In certain embodiments, the oligonucleotides are designed and selected to provide variance-specific amplification.

In reference to nucleic acid sequences which correspond to a gene, the term "correspond" refers to a nucleotide sequence relationship, such that the nucleotide sequence has a nucleotide sequence which is the same as the reference gene or an indicated portion thereof, or has a nucleotide sequence which is exactly complementary in normal Watson-Crick base pairing, or is an RNA equivalent of such a sequence, e.g., an mRNA, or is a cDNA derived from an mRNA of the gene.

In another aspect, the invention provides a method for determining a genotype of an individual in relation to one or more variances in one or more of the genes identified in above aspects by using mass spectrometric determination of a nucleic acid sequence which is a portion of a gene identified for other aspects of this invention or a complementary sequence. Such mass spectrometric methods are known to those skilled in the art. In preferred embodiments, the method involves determining the presence or absence of a variance in a gene; determining the nucleotide sequence of the nucleic acid sequence; the nucleotide sequence is 100 nucleotides or less in length, preferably 50 or less, more preferably 30 or less, and still more preferably 20 nucleotides or less. In general, such a nucleotide sequence includes at least one variance site, preferably a variance site which is informative with respect to the expected response of a patient to a treatment as described for above aspects.

As indicated above, many therapeutic compounds or combinations of compounds or pharmaceutical compositions show variable efficacy and/or safety in various patients in whom the compound or compounds is administered. Thus, it is beneficial to identify variances in relevant genes, e.g., genes related to the action or toxicity of the compound or compounds. Thus, in a further aspect, the invention provides a method for determining whether a compound has a differential effect due to the presence or absence of at least one variance in a gene or a variant form of a gene, where the gene is a gene identified for aspects above.

The method involves identifying a first patient or set of patients suffering from a disease or condition whose response to a treatment differs from the response (to the same treatment) of a second patient or set of patients suffering from the same disease or condition, and then determining whether the occurrence or frequency of occurrence of at least one variance in at least one gene differs between the first patient or set of patients and the second patient or set of patients. A correlation between the presence or absence of the variance or variances and the response of the patient or patients to the treatment indicates that the variance provides information about variable patient response. In general, the method will involve identifying at least one variance in at least one gene. An alternative approach is to identify a first patient or set of patients suffering from a disease or condition and having a particular genotype, haplotype or combination of genotypes or haplotypes, and a second patient or set of patients suffering from the same disease or condition that have a genotype or haplotype or sets of genotypes or haplotypes that differ in a specific way from those of the first set of patients. Subsequently the extent and magnitude of clinical response can be compared between the first patient or set of patients and the second patient or set of patients. A correlation between the presence or absence of a variance or variances or haplotypes and the response of the patient or patients to the treatment indicates that the variance provides information about variable patient response and is useful for the present invention.

The method can utilize a variety of different informative comparisons to identify correlations. For example a plurality of pairwise comparisons of treatment response and the presence or absence of at least one variance can be performed for a plurality of patients. Likewise, the method can involve comparing the response of at least one patient homozygous for at least one variance with at least one patient homozygous for the alternative form of that variance or variances. The method can also involve comparing the response of at least one patient heterozygous for at least one variance with the response of at least one patient homozygous for the at least one variance. Preferably the heterozy-
gous patient response is compared to both alternative homozygous forms, or the response of heterozygous patients is grouped with the response of one class of homozygous patients and said group is compared to the response of the alternative homozygous group.

[0083] Such methods can utilize either retrospective or prospective information concerning treatment response variability. Thus, in a preferred embodiment, it is previously known that patient response to the method of treatment is variable.

[0084] Also in preferred embodiments, the disease or condition is as for other aspects of this invention, for example, the treatment involves administration of a compound or pharmaceutical composition.

[0085] In preferred embodiments, the method involves a clinical trial, e.g., as described herein. Such a trial can be arranged, for example, in any of the ways described herein, e.g., in the Detailed Description.

[0086] The present invention also provides methods of treatment of a disease or condition, preferably a disease or condition related to an inflammatory or immunological disease or other inflammatory or immunological clinical symptomology. Such methods combine identification of the presence or absence of particular variances, preferably in a gene or genes from Tables 1 and 3, with the administration of a compound; identification of the presence of particular variances with selection of a method of treatment and administration of the treatment, and identification of the presence or absence of particular variances with elimination of a method of treatment based on the variance information indicating that the treatment is likely to be ineffective or contra-indicated, and thus selecting and administering an alternative treatment effective against the disease or condition. Thus, preferred embodiments of these methods incorporate preferred embodiments of such methods as described for such sub-aspects.

[0087] As used herein, a “gene” is a sequence of DNA present in a cell that directs the expression of a “biologically active” molecule or “gene product”, most commonly by transcription to produce RNA and translation to produce protein. The “gene product” is most commonly a RNA molecule or protein or a RNA or protein that is subsequently modified by reacting with, or combining with, other constituents of the cell. Such modifications may include, without limitation, modification of proteins to form glycoproteins, lipoproteins, and phosphoproteins, or other modifications known in the art. RNA may be modified without limitation by polyadenylation, splicing, capping or export from the nucleus or by covalent or noncovalent interactions with proteins. The term “gene product” refers to any product directly resulting from transcription of a gene. In particular this includes partial, precursor, and mature transcription products (i.e., pre-mRNA and mRNA), and translation products with or without further processing including, without limitation, lipidation, phosphorylation, glycosylation, or combinations of such processing.

[0088] The term “gene involved in the origin or pathogenesis of a disease or condition” refers to a gene that harbors mutations or polymorphisms that contribute to the cause of disease, or variances that affect the progression of the disease or expression of specific characteristics of the disease. The term also applies to genes involved in the synthesis, accumulation, or elimination of products that are involved in the origin or pathogenesis of a disease or condition including, without limitation, proteins, lipids, carbohydrates, hormones, or small molecules.

[0089] By “pathway” or “gene pathway” is meant the group of biologically relevant genes involved in a pharmacodynamic or pharmacokinetic mechanism of drug, agent, or candidate therapeutic intervention. These mechanisms may further include any physiologic effect the drug or candidate therapeutic intervention renders.

[0090] The term “gene involved in the action of a drug” refers to any gene whose product affects the efficacy or safety of the drug or affects the disease process being treated by the drug, and includes, without limitation, genes that encode gene products that are targets for drug action, gene products that are involved in the metabolism, activation or degradation of the drug, gene products that are involved in the bioavailability or elimination of the drug to the target, gene products that affect biological pathways that, in turn, affect the action of the drug such as the synthesis or degradation of competitive substrates or allosteric effectors or rate-limiting reaction, or, alternatively, gene products that affect the pathophysiology of the disease process via pathways related or unrelated to those altered by the presence of the drug compound. (Particular variances in the latter category of genes may be associated with patient groups in whom disease etiology is more or less susceptible to amelioration by the drug. For example, there are several pathophysiological mechanisms in hypertension, and depending on the dominant mechanism in a given patient, that patient may be more or less likely than the average hypertensive patient to respond to a drug that primarily targets one pathophysiological mechanism. The relative importance of different pathophysiological mechanisms in individual patients is likely to be affected by variances in genes associated with the disease pathophysiology.) The “action” of a drug refers to its effect on biological products within the body. The action of a drug also refers to its effects on the signs or symptoms of a disease or condition, or effects of the drug that are unrelated to the disease or condition leading to unanticipated effects on other processes. Such unanticipated processes often lead to adverse events or toxic effects. The terms “adverse event” or “toxic event” are known in the art and include, without limitation, those listed in the FDA reference system for adverse events.

[0091] In accordance with the aspects above and the Detailed Description below, there is also described for this invention an approach for developing drugs that are explicitly indicated for, and/or for which approved use is restricted to individuals in the population with specific variances or combinations of variances, as determined by diagnostic tests for variances or variant forms of certain genes involved in the disease or condition involved in the action or metabolism or transport of the drug. Such drugs may provide more effective treatment for a disease or condition in a population identified or characterized with the use of a diagnostic test for a specific variance or variant form of the gene if the gene is involved in the action of the drug or in determining a characteristic of the disease or condition. Such drugs may be developed using the diagnostic tests for specific variances or variant forms of a gene to determine the inclusion of patients in a clinical trial.
[0092] Thus, the invention also provides a method for producing a pharmaceutical composition by identifying a compound which has differential activity or effectiveness against a disease or condition in patients having at least one variance in a gene, preferably in a gene from Tables 1 and 3, compounding the pharmaceutical composition by combining the compound with a pharmaceutically acceptable carrier, excipient, or diluent such that the composition is preferentially effective in patients who have at least one copy of an effective variance. In some cases, the present has two copies of the variance or variances. In preferred embodiments, the active component of the pharmaceutical composition is a compound listed in the compound tables below (Tables 5-13, 16, 17, 18), or a compound chemically related to one of the listed compounds.

[0093] Similarly, the invention provides a method for preparing a pharmaceutical agent by identifying a compound which has differential activity against a disease or condition in patients having at least one copy of a form of a gene, preferably a gene from Tables 1 and 3, having at least one variance and synthesizing the compound in an amount sufficient to provide a pharmaceutical effect in a patient suffering from the disease or condition. The compound can be identified by conventional screening methods and its activity confirmed. For example, compound libraries can be screened to identify compounds which differentially bind to products of variant forms of a particular gene product, or which differentially affect expression of variant forms of the particular gene, or which differentially affect the activity of a product expressed from such gene. Alternatively, the design of a compound can exploit knowledge of the variances provided herein to avoid significant allele specific effects, in order to reduce the likelihood of significant pharmacogenetic effects during the clinical development process. Preferred embodiments are as for the preceding aspect.

[0094] In another aspect, the invention provides a method of treating a disease or condition in a patient by selecting a patient whose cells have an allele of an identified gene, preferably a gene selected from the genes listed in Tables 1 and 5. The allele contains at least one variance correlated with an effective response to a treatment of said disease or condition. The method also includes altering the level of activity in cells of the patient of a product of the allele, where the altering provides a therapeutic effect.

[0095] Preferably the allele contains a variance as shown in Tables 1 and 3 or other variance table herein, or in Table 1 or 3 of Stanton & Adams, U.S. application Ser. No. 09/300,747. Also preferably, the altering involves administering to the patient a compound preferentially active on at least one but less than all alleles of the gene.

[0096] Preferred embodiments include those as described above for other aspects of treating a disease or condition.

[0097] As recognized by those skilled in the art, all the methods of treating described herein include administration of the treatment to a patient.

[0098] In a further aspect, the invention provides a method for determining a method of treatment effective to treat a disease or condition by altering the level of activity of a product of an allele of a gene selected from the genes listed in Tables 1 and 3, and determining whether that alteration provides a differential effect (with respect to reducing or alleviating a disease or condition, or with respect to variation in toxicity or tolerance to a treatment) in patients with at least one copy of at least one allele of the gene as compared to patients with at least one copy of one alternative allele. The presence of such a differential effect indicates that altering the level or activity of the gene provides at least part of an effective treatment for the disease or condition.

[0099] Preferably the method for determining a method of treatment is carried out in a clinical trial, e.g., as described above and/or in the Detailed Description below.

[0100] In still another aspect, the invention provides a method for performing a clinical trial or study, which includes selecting or stratifying subjects in the trial or study using a variance or variances or haplotypes from one or more genes specified in Tables 1 and 3. Preferably the differential efficacy, tolerance, or safety of a treatment in a subset of patients who have a particular variance, variances, or haplotype in a gene or genes from Tables 1 and 3 is determined by conducting a clinical trial and using a statistical test to assess whether a relationship exists between efficacy, tolerance, or safety and the presence or absence of any of the variances or haplotype in one or more of the genes. Results of the clinical trial or study are indicative of whether a higher or lower efficacy, tolerance, or safety of the treatment in a subset of patients is associated with any of the variance or variances or haplotype in one or more of the genes. In preferred embodiments, the clinical trial or study is a Phase I, II, III, or IV trial or study. Preferred embodiments include the stratifications and/or statistical analyses as described below in the Detailed Description.

[0101] In preferred embodiments, normal subjects or patients are prospectively stratified by genotype in different genotype-defined groups, including the use of genotype as an enrollment criterion, using a variance, variances or haplotypes from Tables 1 and 3, and subsequently a biological or clinical response variable is compared between the different genotype-defined groups. In preferred embodiments, normal subjects or patients in a clinical trial or study are stratified by a biological or clinical response variable in different biologically or clinically-defined groups, and subsequently the frequency of a variance, variances or haplotypes from Tables 1 and 3 is measured in the different biologically or clinically defined groups.

[0102] In preferred embodiments, e.g., of the above two analyses (and in other aspects of this invention involving patient or normal subject stratification), the normal subjects or patients in a clinical trial or study are stratified by at least one demographic characteristic selected from the group consisting of sex, age, racial origin, ethnic origin, or geographic origin.

[0103] Generally the method will involve assigning patients or subjects to a group to receive the method of treatment or to a control group.

[0104] In yet another aspect, the invention provides experimental methods for finding additional variances in a gene provided in Tables 1 and 3. A number of experimental methods can also beneficially be used to identify variances. Thus, the invention provides methods for producing cDNA (Example 4) and detecting additional variances in the genes provided in Tables 1 or 3 using the single strand conforma-
tion polymorphism (SSCP) method (Example 5), the T4 Endonuclease VII method (Example 6) or DNA sequencing (Example 7) or other methods pointed out below. The application of these methods to the identified genes will provide identification of additional variances that can affect inter-individual variation in drug or other treatment response. One skilled in the art will recognize that many methods for experimental variance detection have been described (in addition to the exemplary methods of examples 5, 6, and 7) and can be utilized. These additional methods include chemical cleavage of mismatches (see, e.g., Ellis T P et al., Chemical cleavage of mismatch: a new look at an established method. Human Mutation 11(5):345-53, 1998), denaturing gradient gel electrophoresis (see, e.g., Van Oorsouw N J, et al., Design and application of 2-D DGGE-based gene mutational scanning tests. Genet Anal. 14(5-6):205-13, 1999) and heteroduplex analysis (see, e.g., Ganguly A, et al., Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced kinks in DNA heteroduplexes. Proc Natl Acad Sci USA. 90 (21):10325-9, 1993). Table 3 of Stanton & Adams, U.S. application Ser. No. 09/300,747, supra, provides a description of the additional possible variances that could be detected by one skilled in the art by testing an identified gene in Tables 1 and 3 using the variance detection methods described or other methods which are known or are developed.

[0105] The present invention provides a method for treating a patient at risk for an inflammatory or immunologic disease or condition (for example to prevent or delay the onset of frank disease) or a patient already diagnosed with inflammatory or immunologic disease or a disease associated with said disease. The methods include identifying such a patient and determining the patient’s genotype or haplotype for an identified gene or genes. The patient identification can, for example, be based on clinical evaluation using conventional clinical metrics and/or evaluation of a genetic variance or variances in one or more genes, preferably a gene or genes from Table 1. The invention provides a method for using the patient’s genotype status to determine a treatment protocol which includes a prediction of the efficacy and/or safety of a therapy.

[0106] In an another related aspect, the invention provides a method for identifying a patient for participation in a clinical trial of a therapy for the treatment of a disease inflammatory or immunologic disease or an associated inflammatory or immunopathological condition. The method involves determining the genotype or haplotype of a patient with (or at risk for) an inflammatory or immunological disease. Preferably the genotype is for a variance in a gene from Table 1. Patients with eligible genotypes are then assigned to a treatment or placebo group, preferably by a blinded randomization procedure. In preferred embodiments, the selected patients have no copies, one copy or two copies of a specific allele of a gene or genes identified in Table 1. Alternatively, patients selected for the clinical trial may have zero, one or two copies of an allele belonging to a set of alleles, where the set of alleles comprise a group of related alleles. One procedure for rigorously defining a set of alleles is by applying phylogenetic methods to the analysis of haplotypes. (See, for example: Templeton A. R., Crandall K. A. and C. F. Sing A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 1992 October; 132(2):619-33.) Regardless of the specific tools used to group alleles, the trial would then test the hypothesis that a statistically significant difference in response to a treatment can be demonstrated between two groups of patients each defined by the presence of zero, one or two alleles (or allele groups) at a gene or genes. Said response may be a desired or an undesired response. In a preferred embodiment, the treatment protocol involves a comparison of placebo vs. treatment response rates in two or more genotype-defined groups. For example a group with no copies of an allele may be compared to a group with two copies, or a group with no copies may be compared to a group consisting of those with one or two copies. In this manner different genetic models (dominant, co-dominant, recessive) for the transmission of a treatment response trait can be tested. Alternatively, statistical methods that do not posit a specific genetic model, such as contingency tables, can be used to measure the effects of an allele on treatment response.

[0107] In another preferred embodiment, patients in a clinical trial can be grouped (at the end of the trial) according to treatment response, and statistical methods can be used to compare allele (or genotype or haplotype) frequencies in two groups. For example responders can be compared to nonresponders, or patients suffering adverse events can be compared to those not experiencing such effects. Alternatively, response data can be treated as a continuous variable and the ability of genotype to predict response can be measured. In a preferred embodiments patients who exhibit extreme phenotypes are compared with all other patients or with a group of patients who exhibit a divergent extreme phenotype. For example if there is a continuous or semi-continuous measure of treatment response (for example the Alzheimer’s Disease Assessment Scale, the Mini-Mental State Examination or the Hamilton Depression Rating Scale) then the 10% of patients with the most favorable responses could be compared to the 10% with the least favorable, or the patients one standard deviation above the mean score could be compared to the remainder, or to those one standard deviation below the mean score. One useful way to select the threshold for defining a response is to examine the distribution of responses in a placebo group. If the upper end of the range of placebo responses is used as a lower threshold for an “outlier response” then the outlier response group should be almost free of placebo responders. This is a useful threshold because the inclusion of placebo responders in a “true” response group decreases the ability of statistical methods to detect a genetic difference between responders and nonresponders.

[0108] In an another related aspect, the invention provides a method for identifying a patient for participation in a clinical trial of a therapy for the treatment of inflammatory or immunological disease or an associated dysfunction or disorder. The method involves characterizing a patient with a disease risk and determining the patient’s genotype, allele status, for an identified gene, preferably a gene from Tables 1 and 3. In preferred embodiments, the method further involves determining the patient’s allele status and selecting those patients having at least one wild type allele, preferably having two wild type alleles for an identified gene, as candidates likely to develop inflammatory or immunological disease or associated pathological disease, disorder or dysfunction.
In a related aspect, the invention provides a method for developing a disease management protocol that entails diagnosing a patient with a disease or a disease susceptibility, determining the genotype of the patient at a gene or genes correlated with treatment response and then selecting an optimal treatment based on the disease and the genotype (or genotypes or haplotypes). The disease management protocol may be useful in an education program for physicians, other caregivers or pharmacists; may constitute part of a drug label; or may be useful in a marketing campaign.

By “disease management protocol” or “treatment protocol” is meant a means for devising a therapeutic plan for a patient using laboratory, clinical and genetic data, including the patient’s diagnosis and genotype. The protocol clarifies therapeutic options and provides information about probable prognoses with different treatments. The treatment protocol may provide an estimate of the likelihood that a patient will respond positively or negatively to a therapeutic intervention. The treatment protocol may also provide guidance regarding optimal drug dose and administration and likely timing of recovery or rehabilitation. A “disease management protocol” or “treatment protocol” may also be formulated for asymptomatic and healthy subjects in order to forecast future disease risks based on laboratory, clinical and genetic variables. In this setting the protocol specifies optimal preventive or prophylactic interventions, including use of compounds, changes in diet or behavior, or other measures. The treatment protocol may include the use of a computer program.

In a preferred embodiment, the treatment protocol involves a comparison of the allele status of a patient with a control population and a responder population. This comparison allows for a statistical calculation of a patient’s likelihood of responding to a therapy, e.g., a calculation of the correlation between a particular allele status and treatment response. In the context of this aspect, the term “wild-type allele” refers to an allele of a gene which produces a product having a level of activity which is most common in the general population. Two different alleles may both be wild-type alleles for this purpose if both have essentially the same level of activity (e.g., specific activity and numbers of active molecules).

In preferred embodiments of above aspects involving prediction of drug efficacy, the prediction of drug efficacy involves candidate therapeutic interventions that are known or have been identified to be metabolized in hepatic tissues. Preferably the candidate therapeutic intervention will be effective in patients with the genotype of at least one allele, and preferably two alleles from Tables 1 and 3, but have a risk to develop inflammatory or immunological disease.

In particular applications of the invention, wherein the above aspects involving a gene variance evaluation or treatment selection or patient selection or method of treatment, the method includes a determination of the genotypic allele status of the patient. Determination of the patient’s allele status as being heterozygous or homozygous is predictive of the patient having a poor response to a candidate therapeutic intervention and development of inflammatory or immunological disease or similar inflammatory or immunological dysfunction. In preferred embodiments, the above methods are used for or include identification of inflammatory or immunological disease and/or the likelihood of occurrence and/or severity of such disease. In preferred embodiments, the invention is suitable for identifying a patient with inflammatory or immunological disease but with inflammatory or immunological dysfunction. The method preferably involves determination of the allele status or variance presence or absence determination for at least one gene from Tables 1 and 3.

In another aspect, the invention provides a method for treating a patient at risk for an inflammatory or immunological disease or condition by a) identifying a patient with such a risk, b) determining the genotypic allele status of the patient, and c) converting the data obtained in step b) into a treatment protocol that includes a comparison of the genotypic allele status determination with the allele frequency of a control population. This comparison allows for a statistical calculation of the patient’s risk for having an inflammatory or immunological disease or condition, e.g., based on correlation of the allele frequencies for a population with response or disease occurrence and/or severity. In preferred embodiments, the method provides a treatment protocol that predicts a patient being heterozygous or homozygous for an identified allele to exhibit signs and or symptoms of inflammatory or immunological disease or condition, whereas a patient who is wild-type homozygous for the said allele, as responding favorably to these therapies.

In a related aspect, the invention provides a method for treating a patient at risk for or diagnosed with an inflammatory or immunological disease or condition using the methods of the above aspect and conducting a step c) which involves determining the gene allele load status of the patient. This method further involves converting the data obtained in steps b) and c) into a treatment protocol that includes a comparison of the allele status determinations of these steps with the allele frequency of a control population. This affords a statistical calculation of the patient’s risk for having inflammatory or immunological disease or inflammatory or immunological dysfunction. In a preferred embodiment, the method is useful for identifying inflammatory or immunological disease or dysfunction. In addition, in related embodiments, the methods provide a treatment protocol that predicts a patient to be at high risk for an inflammatory or immunological disease responding by exhibiting signs and symptoms of inflammatory or immunological dysfunction if the patient is determined as having a genotype or allelic difference in the identified gene or genes. Such patients are preferably given alternative therapies.

The invention also provides a method for improving the safety of candidate therapies for the identification of an inflammatory or immunological disease or dysfunction. The method includes the step of comparing the relative safety of the candidate therapeutic intervention in patients having different alleles in one or more than one of the genes listed in Tables 1 and 3. Preferably, administration of the drug is preferentially provided to those patients with an allele type associated with increased efficacy. In a preferred embodiment, the alleles of identified gene or genes used are wild-type and those associated with reduced biological activity.

As used herein, by “therapy associated with inflammatory or immunological disease” is meant any therapy
resulting in dysfunction or signs and symptoms of an inflammatory or immunologic condition or dysfunction, or those associated with the pathophysiological manifestations of a clinically diagnosed inflammatory or immunologic disorder or syndrome. A suitable therapy can be a pharmacological agent or drug that may enhance or inhibit metabolic pathways identified to affect the molecular structure or function of the parent candidate therapeutic intervention thereby affecting inflammatory or immunologic disease progression or any of these inflammatory or immunologic dysfunctions.

[0118] By “inflammatory or immunological dysfunction” is meant a disease or syndrome in which symptomology is similar to an inflammatory or immunologic disease. Specifically included are: arthritis, asthma, chronic obstructive pulmonary disease, autoimmune disease, inflammatory bowel disease, immunosuppression related to transplantation, pain associated with inflammation, psoriasis, atherosclerosis, and hepatitis.

[0119] By “drug efficacy” is meant the determination of an appropriate drug, drug dosage, administration schedule, and prediction of therapeutic utility.

[0120] By “allele load” is meant the relative ratio of identified gene alleles in the patient’s chromosomal DNA.

[0121] By “identified allele” is meant a particular gene isoform that can be distinguished from other identified gene isoforms using the methods of the invention.

[0122] By “PCR, RT-PCR, or ligase chain reaction amplification” is meant subjecting a DNA sample to a Polymerase Chain Reaction step or ligase-mediated chain reaction step, or RNA to a RT-PCR step, such that, in the presence of appropriately designed primers, a nucleic acid fragment is synthesized or fails to be synthesized and thereby reveals the allele status of a patient. The nucleic acid may be further analyzed by DNA sequencing using techniques known in the art.

[0123] By “gene allele status” is meant a determination of the relative ratio of wild type identified alleles compared to an allelic variant that may encode a gene product of reduced catalytic activity. This may be accomplished by nucleic acid sequencing, RT-PCR, PCR, examination of the identified gene translated protein, a determination of the identified protein activity, or by other methods available to those skilled in the art.

[0124] By “treatment protocol” is meant a therapy plan for a patient using genetic and diagnostic data, including the patient’s diagnosis and genotype. The protocol enhances therapeutic options and clarifies prognosis. The treatment protocol may include an indication of whether or not the patient is likely to respond positively to a candidate therapeutic intervention that is known to affect inflammatory or immunological function. The treatment protocol may also include an indication of appropriate drug dose, recovery time, age of disease onset, rehabilitation time, symptomology of attacks, and risk for future disease. A treatment protocol, including any of the above aspects, may also be formulated for asymptomatic and healthy subjects in order to forecast future disease risks and determine what preventive therapies should be considered or invoked in order to lessen these disease risks. The treatment protocol may include the use of a computer software program to analyze patient data.

[0125] By “patient at risk for a disease” or “patient with high risk for a disease” is meant a patient identified or diagnosed as having inflammatory or immunological disease or inflammatory or immunological dysfunction or having a genetic predisposition or risk for acquiring inflammatory or immunological disease or inflammatory or immunological dysfunction, where the predisposition or risk is higher than average for the general population or is sufficiently higher than for other individuals as to be clinically relevant. Such risk can be evaluated, for example, using the methods of the invention and techniques available to those skilled in the art.

[0126] By “converting” is meant compiling genotype determinations to predict either prognosis, drug efficacy, or suitability of the patient for participating in clinical trials of a candidate therapeutic intervention with known propensity of inflammatory or immunological disease or dysfunction. For example, the genotype may be compiled with other patient parameters such as age, sex, disease diagnosis, and known allelic frequency of a representative control population. The converting step may provide a determination of the statistical probability of the patient having a particular disease risk, drug response, or patient outcome.

[0127] By “prediction of patient outcome” is meant a forecast of the patient’s likely health status. This may include a prediction of the patient’s response to therapy, rehabilitation time, recovery time, cure rate, rate of disease progression, predisposition for future disease, or risk of having relapse.

[0128] By “therapy for the treatment of a disease” is meant any pharmacological agent or drug with the property of healing, curing, or ameliorating any symptom or disease mechanism associated with inflammatory or immunological disease or inflammatory or immunological conditions, or dysfunction.

[0129] By “responder population” is meant a patient or patients that respond favorably to a given therapy.

[0130] In another aspect, the invention provides a kit containing at least one probe or at least one primer (or other amplification oligonucleotide) or both (e.g., as described above) corresponding to a gene or genes listed in Tables 1 and 3 or other gene related to an immunological or inflammatory disease or condition. The kit is preferably adapted and configured to be suitable for identification of the presence or absence of a particular variance or variances, which can include or comprise a nucleic acid sequence corresponding to a portion of a gene. A plurality of variances may comprise a haplotype of haplotypes. The kit may also contain a plurality of either or both of such probes and/or primers, e.g., 2, 3, 4, 5, 6, or more of such probes and/or primers. Preferably the plurality of probes and/or primers are adapted to provide detection of a plurality of different sequence variances in a gene or plurality of genes, e.g., in 2, 3, 4, 5, or more genes or to amplify and/or sequence a nucleic acid sequence including at least one variance site in a gene or genes. Preferably one or more of the variance or variances to be detected are correlated with variability in a treatment response or tolerance, and are preferably indicative of an effective response to a treatment. In preferred embodiments, the kit contains components (e.g., probes and/or primers) adapted or useful for detection of a plurality of variances (which may be in one or more genes) indicative of the effectiveness of at least one treatment, preferably of a
plurality of different treatments for a particular disease or condition. It may also be desirable to provide a kit containing components adapted or useful to allow detection of a plurality of variences indicative of the effectiveness of a treatment or treatment against a plurality of diseases. The kit may also optionally contain other components, preferably other components adapted for identifying the presence of a particular variance or variences. Such additional components can, for example, independently include a buffer or buffers, e.g., amplification buffers and hybridization buffers, which may be in liquid or dry form, a DNA polymerase, e.g., a polymerase suitable for carrying out PCR (e.g., a thermostable DNA polymerase), and deoxy nucleotide triphosphates (dNTPs). Preferably a probe includes a detectable label, e.g., a fluorescent label, enzyme label, light scattering label, or other label. Preferably the kit includes a nucleic acid or polypeptide array on a solid phase substrate. The array may, for example, include a plurality of different antibodies, and/or a plurality of different nucleic acid sequences. Sites in the array can allow capture and/or detection of nucleic acid sequences or gene products corresponding to different variences in one or more different genes. Preferably the array is arranged to provide variance detection for a plurality of variences in one or more genes which correlate with the effectiveness of one or more treatments of one or more diseases, which is preferably a variance as described herein.

[0131] The kit may also optionally contain instructions for use, which can include a listing of the variences correlating with a particular treatment or treatments for a disease or diseases and/or a statement or listing of the diseases for which a particular variance or variences correlates with a treatment efficacy and/or safety.

[0132] Preferably the kit components are selected to allow detection of a variance described herein, and/or detection of a variance indicative of a treatment, e.g., administration of a drug, pointed out herein.

[0133] Additional configurations for kits of this invention will be apparent to those skilled in the art.

[0134] The invention also includes the use of such a kit to determine the genotype(s) of one or more individuals with respect to one or more variance sites in one or more genes identified herein. Such use can include providing a result or report indicating the presence and/or absence of one or more variant forms or a gene or genes which are indicative of the effectiveness of a treatment or treatments.

[0135] In another aspect, the invention provides a method for determining whether there is a genetic component to intersubject variation in a surrogate treatment response. The method involves administering the treatment to a group of related (preferably normal) subjects and a group of unrelated (preferably normal) subjects, measuring a surrogate pharmacodynamic or pharmokinetic drug response variable in the subjects, performing a statistical test measuring the variation in response in the group of related subjects and, separately in the group of unrelated subjects, comparing the magnitude or pattern of variation in response or both between the groups to determine if the responses of the groups are different, using a predetermined statistical measure of difference. A difference in response between the groups is indicative that there is a genetic component to intersubject variation in the surrogate treatment response.

[0136] In preferred embodiments, the size of the related and unrelated groups is set in order to achieve a predetermined degree of statistical power.

[0137] In another aspect, the invention provides a method for evaluating the combined contribution of two or more variences to a surrogate drug response phenotype in subjects (preferably normal subjects) by genotyping a set of unrelated subjects participating in a clinical trial or study, e.g., a Phase I trial, of a compound. The genotyping is for two or more variences (which can be a haplotype), thereby identifying subjects with specific variences, where the two or more specific variences define two or more genotype-defined groups. A drug is administered to subjects with two or more of said specific variences, and a surrogate pharmacokinetic or pharmokinetic drug response variable is measured in the subjects. A statistical test or tests is performed to measure response in the groups separately, where the statistical tests provide a measurement of variance in response with each group. The magnitude or pattern of variance in response or both is compared between the groups to determine if the groups are different using a predetermined statistical measure of difference.

[0138] In preferred embodiments, the specific variences are homozygous variences for two variences. In preferred embodiments, the comparison is between groups of subjects differing in three or more variences, e.g., 3, 4, 5, 6, or even more variences.

[0139] In another aspect, the invention provides a method for providing contract research services to clients (preferably in the pharmaceutical and biotechnology industries), by enrolling subjects (e.g., normal and/or patient subjects) in a clinical drug trial or study unit (preferably a Phase I drug trial or study unit) for the purpose of genotyping the subjects in order to assess the contribution of genetic variation to variation in drug response, genotyping the subjects to determine the status of one or more variences in the subjects, administering a compound to the subjects and measuring a surrogate drug response variable, comparing responses between two or more genotype-defined groups of subjects to determine whether there is a genetic component to the interperson variability in response to said compound; and reporting the results of the Phase I drug trial to a contracting entity. Clearly, intermediate results, e.g., response data and/or statistical analysis of response or variation in response can also be reported.

[0140] In preferred embodiments, at least some of the subjects have disclosed that they are related to each other and the genetic analysis includes comparison of groups of related individuals. To encourage participation of sufficient numbers of related individuals, it can be advantageous to offer or provide compensation to one or more of the related individuals based on the number of subjects related to them who participate in the clinical trial, or on whether at least a minimum number of related subjects participate, e.g., at least 3, 5, 10, 20, or more.

[0141] In a related aspect, the invention provides a method for recruiting a clinical trial population for studies of the influence of genetic variation on drug response, by soliciting subjects to participate in the clinical trial, obtaining consent of each of a set of subjects for participation in the clinical trial, obtaining additional related subjects for participation in the clinical trial by compensating one or more of the related
subjects for participation of their related subjects at a level based on the number of related subjects participating or based on participation of at least a minimum specified number of related subjects, e.g., at minimum levels as specified in the preceding aspect.

[0142] In the various aspects of this invention, a gene is a gene as identified in Tables 1 or 3, or is a gene in a pathway as identified in Table 1, preferably a gene having a function as identified in Table 1. Some such additional genes, and exemplary variances are provided in a parent application.

[0143] In the various aspects of this invention, the drug is preferably a drug as identified in a drug table herein, or is a drug in the same chemical class.

[0144] By “pathway” or “gene pathway” is meant the group of biologically relevant genes involved in a pharmacodynamic or pharmacokinetic mechanism of drug, agent, or candidate therapeutic intervention. These mechanisms may further include any physiologic effect the drug or candidate therapeutic intervention renders. Included in this are “biochemical pathways” which is used in its usual sense to refer to a series of related biochemical processes (and the corresponding genes and gene products) involved in carrying out a reaction or series of reactions. Generally in a cell, a pathway performs a significant process in the cell.

[0145] By “pharmacological activity” used herein is meant a biochemical or physiological effect of drugs, compounds, agents, or candidate therapeutic interventions upon administration and the mechanism of action of that effect.

[0146] The pharmacological activity is then determined by interactions of drugs, compounds, agents, or candidate therapeutic interventions, or their mechanism of action, on their target proteins or macromolecular components. By “agonist” or “mimetic” or “activators” is meant a drug, agent, or compound that activate physiologic components and mimic the effects of endogenous regulatory compounds. By “antagonists,” “blockers” or “inhibitors” is meant drugs, agents, or compounds that bind to physiologic components and do not mimic endogenous regulatory compounds, or interfere with the action of endogenous regulatory compounds at physiologic components. These inhibitory compounds do not have intrinsic regulatory activity, but prevent the action of agonists. By “partial agonist” or “partial antagonist” is meant an agonist or antagonist, respectively, with limited or partial activity. By “negative agonist” or “inverse antagonists” is meant that a drug, compound, or agent that can interact with a physiologic target protein or macromolecular component and stabilizes the protein or component such that agonist-dependent conformational changes of the component do not occur and agonist mediated mechanism of physiological action is prevented. By “modulators” or “factors” is meant a drug, agent, or compound that interacts with a target protein or macromolecular component and modifies the physiological effect of an agonist.

[0147] As used herein the term “chemical class” refers to a group of compounds that share a common chemical scaffold but which differ in respect to the substituent groups linked to the scaffold. Examples of chemical classes of drugs include, for example, phenothiazines, piperidines, benzodiazepines and aminoglycosides. Members of the phenothiazine class include, for example, compounds such as chlorpromazine hydrochloride, mesoridazine besylate, thioridazine hydrochloride, acetophenazine maleate trifluperazine hydrochloride and others, all of which share a phenothiazine backbone. Members of the piperidine class include, for example, compounds such as meperidine, diphenoxylate and loperamide, as well as phenylpiperidines such as fentanyl, sufentanil and alfentanil, all of which share the piperidine backbone. Chemical classes and their members are recognized by those skilled in the art of medicinal chemistry.

[0148] As used herein the term “surrogate marker” refers to a biological or clinical parameter that is measured in place of the biologically definitive or clinically most meaningful parameter. In comparison to definitive markers, surrogate markers are generally either more convenient, less expensive, provide earlier information or provide pharmacological or physiological information not directly obtainable with definitive markers. Examples of surrogate biological parameters: (i) testing erythrocyte membrane acetylcholinesterase levels in subjects treated with an acetylcholinesterase inhibitor intended for use in Alzheimer’s disease patients (where inhibition of brain acetylcholinesterase would be the definitive biological parameter); (ii) measuring levels of CD4 positive lymphocytes as a surrogate marker for response to a treatment for acquired immune deficiency syndrome (AIDS). Examples of surrogate clinical parameters: (i) performing a psychometric test on normal subjects treated for a short period of time with a candidate Alzheimer’s compound in order to determine if there is a measurable effect on cognitive function. The definitive clinical test would entail measuring cognitive function in a clinical trial in Alzheimer’s disease patients. (ii) Measuring blood pressure as a surrogate marker for myocardial infarction. The measurement of a surrogate marker or parameter may be an endpoint in a clinical study or clinical trial, hence “surrogate endpoint”.

[0149] As used herein the term “related” when used with respect to human subjects indicates that the subjects are known to share a common line of descent; that is, the subjects have a known ancestor in common. Examples of preferred related subjects include sibs (brothers and sisters), parents, grandparents, children, grandchildren, aunts, uncles, cousins, second cousins and third cousins. Subjects less closely related than third cousins are not sufficiently related to be useful as “related” subjects for the methods of this invention, even if they share a known ancestor, unless some related individuals that lie between the distantly related subjects are also included. Thus, for a group of related individuals, each subject shares a known ancestor within three generations or less with at least one other subject in the group, and preferably with all other subjects in the group or has at least that degree of consanguinity due to multiple known common ancestors. More preferably, subjects share a common ancestor within two generations or less, or otherwise have equivalent level of consanguinity. Conversely, as used herein the term “unrelated”, when used in respect to human subjects, refers to subjects who do not share a known ancestor within 3 generations or less, or otherwise have known relatedness at that degree.

[0150] As used herein the term “pedigree” refers to a group of related individuals, usually comprising at least two generations, such as parents and their children, but often comprising three generations (that is, including grandparents
or grandchildren as well). The relation between all the subjects in the pedigree is known and can be represented in a genealogical chart.

[0151] As used herein the term “hybridization”, when used with respect to DNA fragments or polynucleotides encompasses methods including both natural polynucleotides, non-natural polynucleotides or a combination of both. Natural polynucleotides are those that are polymers of the four natural deoxyribonucleotides (deoxyadenosine triphosphate [dA], deoxycytosine triphosphate [dC], deoxyguanine triphosphate [dG] or deoxythymidine triphosphate [dT]), usually designated simply thymidine triphosphate [T]) or polymers of the four natural ribonucleotides (adenosine triphosphate [A], cytosine triphosphate [C], guanine triphosphate [G] or uridine triphosphate [U]). Non-natural polynucleotides are made up in part or entirely of nucleotides that are not natural nucleotides; that is, they have one or more modifications. Also included among non-natural polynucleotides are molecules related to nucleic acids, such as peptide nucleic acid (PNA). Non-natural polynucleotides may be polymers of non-natural nucleotides, polymers of natural and non-natural nucleotides (in which there is at least one non-natural nucleotide), or otherwise modified polynucleotides. Non-natural polynucleotides may be useful because their hybridization properties differ from those of natural polynucleotides. As used herein the term “complementary”, when used in respect to DNA fragments, refers to the base pairing rules established by Watson and Crick: A pairs with T or U; G pairs with C. Complementary DNA fragments have sequences that, when aligned in antiparallel orientation, conform to the Watson-Crick base pairing rules at all positions or at all positions except one. As used herein, complementary DNA fragments may be natural polynucleotides, non-natural polynucleotides, or a mixture of natural and non-natural polynucleotides.

[0152] As used herein “amplify” when used with respect to DNA refers to a family of methods for increasing the number of copies of a starting DNA fragment. Amplification of DNA is often performed to simplify subsequent determination of DNA sequence, including genotyping or haplotyping. Amplification methods include the polymerase chain reaction (PCR), the ligase chain reaction (LCR) and methods using Q beta replicase, as well as transcription-based amplification systems such as the isothermal amplification procedure known as self-sustained sequence replication (3SR, developed by T. R. Gingeras and colleagues), strand displacement amplification (SDA, developed by G. T. Walker and colleagues) and the rolling circle amplification method (developed by P. Lizardi and D. T. Ward).

[0153] As used herein “contract research services for a client” refers to a business arrangement wherein a client entity pays for services comprising in part or in whole of work performed using the methods described herein. The client entity may include a commercial or non-profit organization whose primary business is in the pharmaceutical, biotechnology, diagnostics, medical device or contract research organization (CRO) sector, or any combination of those sectors. Services provided to such a client may include any of the methods described herein, particularly including clinical trial services, and especially the services described in the Detailed Description relating to a Pharmacogenetic Phase I Unit. Such services are intended to allow the earliest possible assessment of the contribution of a variance or variances or haplotypes, from one or more genes, to variation in a surrogate marker in humans. The surrogate marker is generally selected to provide information on a biological or clinical response, as defined above.

[0154] As used herein, “comparing the magnitude or pattern of variation in response” between two or more groups refers to the use of a statistical procedure or procedures to measure the difference between two different distributions. For example, consider two genotype-defined groups, AA and aa, each homozygous for a different variance or haplotype in a gene believed likely to affect response to a drug. The subjects in each group are subjected to treatment with the drug and a treatment response is measured in each subject (for example a surrogate treatment response). One can then construct two distributions: the distribution of responses in the AA group and the distribution of responses in the aa group. These distributions may be compared in many ways, and the significance of any difference qualified as to its significance (often expressed as a p value), using methods known to those skilled in the art. For example, one can compare the means, medians or modes of the two distributions, or one can compare the variance or standard deviation. Or, if the form of the distributions is not known, one can use nonparametric statistical tests to test whether the distributions are different, and whether the difference is significant at a specified level (for example, the p<0.05 level, meaning that, by chance, the distributions would differ to the degree measured less than one in 20 similar experiments). The types of comparisons described are similar to the analysis of heritability in quantitative genetics, and would draw on standard methods from quantitative genetics to measure heritability by comparing data from related subjects.

[0155] Another type of comparison that can be usefully made is between related and unrelated groups of subjects. That is, the comparison of two or more distributions is of particular interest when one distribution is drawn from a population of related subjects and the other distribution is drawn from a group of unrelated subjects, both subjected to the same treatment. (The related subjects may comprise small groups of related subjects, each compared only to their relatives.) A comparison of the distribution of a drug response variable (e.g. a surrogate marker) between two such groups may provide information on whether the drug response variable is under genetic control. For example, a narrow distribution in the group(s) of related subjects (compared to the unrelated subjects) would tend to indicate that the measured variable is under genetic control (i.e. the related subjects, on account of their genetic homogeneity, are more similar than the unrelated individuals). The degree to which the distribution was narrower in the related individuals (compared to the unrelated individuals) would be proportionate to the degree of genetic control. The narrowness of the distribution could be quantified by, for example, computing variance or standard deviation. In other cases the shape of the distribution may not be known and nonparametric tests may be preferable. Nonparametric tests include methods for comparing medians such as the sign test, the slippage test, or the rank correlation coefficient (the non-parametric equivalent of the ordinary correlation coefficient). Pearson’s Chi square test for comparing an observed set of frequencies with an expected set of frequencies can also be useful.
The present invention provides a number of advantages. For example, the methods described herein allow for use of a determination of a patient’s genotype for the timely administration of the most suitable therapy for that particular patient. The methods of this invention provide a basis for successfully developing and obtaining regulatory approval for a compound even though efficacy or safety of the compound in an unstratified population is not adequate to justify approval. From the point of view of a pharmaceutical or biotechnology company, the information obtained in pharmacogenetic studies of the type described herein could be the basis of a marketing campaign for a drug. For example, a marketing campaign that emphasized the superior efficacy or safety of a compound in a genotype or haplotype restricted patient population, compared to a similar or competing compound used in an undifferentiated population of all patients with the disease. In this respect a marketing campaign could promote the use of a compound in a genetically defined subpopulation, even though the compound was not intrinsically superior to competing compounds when used in the undifferentiated population with the target disease. In fact even a compound with an inferior profile of action in the undifferentiated disease population could become superior when coupled with the appropriate pharmacogenetic test.

By “comprising” is meant including, but not limited to, whatever follows the word “comprising”. Thus, use of the term “comprising” indicates that the listed elements are optional or mandatory, but that other elements are optional and may or may not be present. By “consisting of” is meant including, and limited to, whatever follows the phrase “consisting of”. Thus, the phrase “consisting of” indicates that the listed elements are required or mandatory, and that no other elements may be present. By “consisting essentially of” is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase “consisting essentially of” indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

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

DETAILED DESCRIPTION OF THE INVENTION

The tables are first briefly described.

Table 1, the Inflammation and Immunology Gene Table, lists genes that may be involved in pharmacological response to inflammation or immunologic therapeutics, or that may define disease subsets with different prognosis and consequent implications for treatment. The table has seven columns. Column 1, headed “Class” provides broad groupings of genes relevant to the pharmacology of anti-inflammation drugs or drugs affecting the immune system. Column 2, headed “Pathway”, provides a more detailed categorization of the different classes of genes by indicating the overall purpose of large groups of genes. These pathways contain genes implicated in the etiology or treatment response of the various diseases detailed in Table 2. Column 3, headed “Function”, further categorizes the pathways listed in column 2. Some categories in column 2 are not further categorized in column 3.

Column 4, headed “Name”, lists the genes belonging to the class, pathway and function shown to the left (in columns 1-3). The gene names given are generally those used in the OMIM database or in GenBank, however one skilled in the art will recognize that many genes have more than one name, and that it is a straightforward task to identify synonymous names. For example, many alternate gene names are provided in the OMIM record for a gene.

In column 5, headed “OMIM”, the Online Mendelian Inheritance in Man (OMIM) record number is listed for each gene in column 4. This record number can be entered next to the words: “Enter one or more search keywords:” at the OMIM world wide web site. The url is: http://www3.ncbi.nlm.nih.gov/Omim/searchomim.html. An OMIM record exists for most characterized human genes. The record often has useful information on the chromosome location, function, alleles, and human diseases or disorders associated with each gene.

Column 6, headed “GID”, provides the GenBank identification number (hence GID) of a genomic, cDNA, or partial sequence of the gene named in column 4. Usually the GID provides the record of a cDNA sequence. Many genes have multiple GenBank accession numbers, representing different versions of a sequence obtained by different research groups, or corrected or updated versions of a sequence. As with the gene name, one skilled in the art will recognize that alternative GenBank records related to the named record can be obtained easily. All other GenBank records listing sequences that are alternate versions of the sequences named in the table are equally suitable for the inventions described in this application. (One straightforward way to obtain additional GenBank records for a gene is on the internet. General instructions can be found at the NCBI web site at: http://www3.ncbi.nlm.nih.gov. More specifically, the GenBank record number in column 6 can be entered at the url: http://www3.ncbi.nlm.nih.gov/Entrez/ nucleotide.html. Once the GenBank record has been retrieved one can click on the “nucleotide neighbors” link and additional GenBank records from the same gene will be listed.

Column 7, headed “locus”, provides the chromosome location of the gene listed on the same row. The chromosome location helps confirm the identity of the named gene if there is any ambiguity.

Table 2 is a matrix showing the intersection of genes and therapeutic indications—that is, which categories of genes are most likely to account for interpatient variation in response to treatments for which diseases. The Table is displayed on two pages. On each page the first two columns provide a framework for organizing the genes listed in Table 1. Column 1 is similar to the ‘Class’ column in Table 1, while column 2 is a combination of the ‘Pathway’ and ‘Function’ columns in Table 1. It is intended that the summary terms listed in columns 1 and 2 be read as referring to all the genes in the corresponding sections of Table 1. The remaining columns in Table 2 list ten inflammatory or immunologic diseases. The information in the Table lies in the shaded boxes at the intersection of various “Pathways” (the rows) and treatment indications (the 13 columns) An
intersection box is shaded when a row corresponding to a particular pathway (and by extension all the genes listed in that pathway in Table 1) intersects a column for a specific inflammatory disease or immune disorders such that the pathway and genes are of possible use in explaining inter-patient differences in response to treatments for the inflammation or immunologic indication. Thus, the Table enables one skilled in the art to identify therapeutically relevant genes in patients with one of the 13 indications for the purposes of stratification of these patients based upon genotype and subsequent correlation of genotype with drug response. The shaded intersections indicate preferred sets of genes for understanding the basis of inter-patient variation in response to therapy of the indicated disease indication, and in that respect are exemplary. Any of the genes in the table may account for inter-patient variation in response to treatments for any of the diseases listed. Thus, the shaded boxes indicate the gene pathways that one skilled in the art would first investigate in trying to understand inter-patient variation in response to therapy for the listed indications.

Table 3 is a partial list of DNA sequence variances in genes relevant to the methods described in the present invention. These variances were identified by the inventors in studies of selected genes listed in Table 1, and are provided here as useful for the methods of the present invention. The variances in Table 3 were discovered by one or more of the methods described below in the Detailed Description or Examples. Table 3 has eight columns. Column 1, the “Name” column, contains the Human Genome Organization (HUGO) identifier for the gene. Column 2, the “GID” column provides the GenBank accession number of a genomic, cDNA, or partial sequence of a particular gene. Column 3, the “OMIM_ID” column contains the record number corresponding to the Online Mendelian Inheritance in Man database for the gene provided in columns 1 and 2. This record number can be entered at the world wide web site http://www3.ncbi.nlm.nih.gov/Omim/searchomim.html to search the OMIM record on the gene. Column 4, the VGX_Symbol column, provides an internal identifier for the gene. Column 5, the “Description” column provides a descriptive name for the gene, when available. Column 6, the “Variance_Start” column provides the nucleotide location of a variance with respect to the first listed nucleotide in the GenBank accession number provided in column 2. That is, the first nucleotide of the GenBank accession is counted as nucleotide 1 and the variant nucleotide is numbered accordingly. Column 7, the “Variance” column provides the nucleotide location of a variance with respect to an ATG codon believed to be the authentic ATG start codon of the gene, where the A of ATG is numbered as one (1) and the immediately preceding nucleotide is numbered as minus one (-1). This reading frame is important because it allows the potential consequence of the variant nucleotide to be interpreted in the context of the gene anatomy (5' untranslated region, protein coding sequence, 3' untranslated region). Column 7 also provides the identity of the two variant nucleotides at the indicated position. For example, in the first entry in Table 3, AB0000221, the variance is 364C>T, indicating the presence of a T or C at nucleotide 427 of GenBank sequence AB0000221. Column 8, the “CD5S_Context” column indicates whether the variance is in a coding region but silent (S); in a coding region and results in an amino acid change (e.g., R347C, where the letters are one letter amino acid abbreviations and the number is the amino acid residue in the encoded amino acid sequence which is changed); in a sequence 5' to the coding region (5); or in a sequence 3' to the coding region (3). As indicated above, interpreting the location of the variance in the gene depends on the correct assignment of the initial ATG of the encoded protein (the translation start site). It should be recognized that assignment of the correct ATG may occasionally be incorrect in GenBank, but that one skilled in the art will know how to carry out experiments to definitively identify the correct translation initiation codon (which is not always an ATG). In the event of any potential question concerning the proper identification of a gene or part of a gene, due for example, to an error in recording an identifier or the absence of one or more of the identifiers, the priority for use to resolve the ambiguity is GenBank accession number, OMIM identification number, HUGO identifier, common name identifier.

Tables 5-13, 16, 17, 18 provide lists of exemplary compounds in clinical development for various inflammatory or immunologic diseases. The compounds listed in the tables are exemplary; that is, the methods of the invention will apply to other compounds as well. Each table has four columns. The first column is titled “Product Name”, the second column is titled “Chemical Name”, the third “Action” and the fourth “Indication”. Under these headings are listed rows of compounds. For each compound there is a brief summary of information about the product name, its pharmacological action and potential clinical uses. The first column, “Product Name”, provides the generic name and/or alphanumeric designation of the compound, as well as its trade name in some cases (in capital letters). The second column, “Chemical Name” provides the full chemical name of the compound. The listed compounds, or compounds chemically related to those listed (e.g. by modification of one or more chemical moieties of the listed compounds), are suitable for the methods of this invention. The third column, “Action”, summarizes in a word or phrase an important pharmacological action of the compound, or what is currently believed to be an important pharmacological action—in most cases additional pharmacological actions are known but not listed to conserve space; alternatively, subsequent studies may reveal additional or alternative pharmacological actions. (Sources listed in the detailed description will help clarify whether additional pharmacological actions have been discovered.) The fourth column, “Indication”, provides an exemplary disease or condition for which the compound is currently being, or has already been, developed. In many cases the compound is being, has already been, or will likely be developed for other indications. Again, one skilled in the art will know how to identify additional drug development programs for these compounds. For example, a compound in development for one inflammatory disease is likely to be evaluated in the treatment of other inflammatory diseases.

Detailed Description of the Preferred Embodiments

1. Inflammatory or Immunological Disease, Disorders, or Dysfunctions

Inflammatory or immunological diseases and clinical symptoms include diseases and processes such as: arthritis (including rheumatoid arthritis, osteoarthritis, and other degenerative syndromes of the joints), asthma, chronic obstructive pulmonary disease (including bronchitis, bronchiectasis, emphysema and other pulmonary diseases asso-
Associated with obstruction to air flow, interstitial or restrictive lung diseases, autoimmune disease (including systemic lupus erythematosus, scleroderma and other diseases characterized by autoantibodies), transplantation (often treated with long term immunosuppressive therapy), pain associated with inflammation, psoriasis and other inflammatory skin diseases, atherosclerosis (for which there is strong data supporting the role of inflammatory pathogenetic mechanisms), and hepatitis, among other diseases. One skilled in the art will recognize that there may be overlap between some of the conditions listed.

Challenges in treating diseases with a significant inflammatory or immunological component include: (i) limited understanding of the pathophysiology of these diseases and conditions, (ii) a complex mix of immune/inflammatory mediators operating simultaneously, with the primary (initiating) events often unclear and the relative importance of different mediators unknown, and (iii) medical interventions that rarely produce specific effects or address the underlying pathophysiological basis of the disease or condition. Thus, medical management of inflammatory or immunologic disorders is empirical in nature. This is associated with multiple undesirable side effects, and thus, disease progression is common. Based upon these clinical realities and the difficulties drug developers face in developing new treatments for diseases with an inflammatory or immunologic component, the use of genotype-based stratification to identify populations enriched for responders, non-responders, and/or those likely to develop undesirable side effects will provide clear commercial and medical benefits. Ultimately, medical practitioners and patients will also benefit from better choices of medicines with superior safety and efficacy characteristics when used in conjunction with genetic diagnostic tests.

Inflammation is a complex process that comprises different cellular and physiologic events that can be initiated by tissue injury, by abnormal immune function. They can also be initiated by a wide variety of other endogenous or exogenous factors, not all of which are understood. The inflammatory process can also escape normal regulatory control and become part of the disease process.

Autoimmunity is one aspect of some diseases associated with abnormal immunologic function. Such diseases are characterized by the presence of autoantibodies and oligoclonal B cell populations. Immunological reactions associated with loss of self tolerance may be localized to a specific tissue or may be systemic. Ultimately, in severe cases, the immune system produces life threatening damage to tissues and result in compromised physiological functions. Autoimmunity can be initiated by a variety of endogenous (genetic predisposition and others) and exogenous (chemicals, drugs, microorganisms, and others) factors.

Exemplary diseases characterized by abnormal inflammatory or immunologic responses (also referred to herein as inflammatory or immune diseases or disorders) are described below. These diseases are suitable for application of methods described in this invention for identification of variances in a gene or genes involved in therapeutic response, e.g. efficacy, tolerability or toxicity.

A. Arthritis

Description of Arthritis

Arthritis comprises a variety of diseases characterized by pain, swelling, and limited movement in joints and connective tissues. Arthritis is usually chronic and there are three prevalent forms of the disease: rheumatoid arthritis (RA), osteoarthritis (OA), and fibromyalgia. In RA, the synovial joint lining becomes inflamed as a result of hyperactive immune response. There are an estimated 2.1 million Americans with RA, two thirds being women. In OA, the cartilage that covers the ends of the bones within joints deteriorates and causing pain and loss of movement as bone begins to rub against bone. There are an estimated 20.7 million Americans with OA, the majority being over the age of 45. In fibromyalgia, widespread pain affects muscles, attachments of muscles to bone, and the connective tissues, i.e., the ligaments and tendons. There are an estimated 3.7 million individuals diagnosed with fibromyalgia syndrome. Other serious and common forms of arthritis or related disorders include, among others, gout, systemic lupus erythematosus, scleroderma, ankylosing spondylitis, and juvenile arthritis.

Rheumatoid arthritis involves the disarrangefl joints and can affect a variety of other organs. The clinical hallmarks of RA include: morning stiffness; swelling of three or more joints; swelling of hand joints (proximal interphalangeal, metacarpophalangeal, or wrist); symmetric swelling; subcutaneous nodules; serum rheumatoid factor; and erosions and or periarticular osteopenia in hand or wrist joints as often observed on radiograph.

Osteoarthritis is a degenerative process in joint tissues that may occur in response to aging, genetic, or environmental factors. Its symptoms include progressive degeneration of cartilage, bone remodeling, and overgrowth of bone. Clinical hallmarks of OA include deep aching pain in the afflicted joints (hands, knees, spine, and hips), morning stiffness of short duration, variable joint thickening, and effusion. Pathologically OA is characterized by breakdown of cartilage. Destruction of joint cartilage involves direct physical injury, enzymatic degradation as a result of the injury to chondrocytes, and subchondral bone stiffening as a result of the bone remodeling.

Current Therapies for Arthritis

Agents used to treat RA fall into one of the following four categories: analgesics (NSAIDs, salicylates), disease modifying antirheumatic agents (gold compounds, cytotoxic), hormones (glucocorticoids), and skin and mucosal membrane preparations. Therapies for the treatment of OA focus on decreasing pain (analgesics) and physical therapies to increase joint mobility.

Analgesics: Typically, pain associated with arthritis can be controlled with NSAIDs including but not limited to, salicylates, para-aminophenol derivatives, indole and indene derivatives, heteroarayl acetic acids, aryloproprionic acids, anthranilic acids, enolic acids, or alkanones. Antiinflammatory agents such as cyclooxygenase inhibitors, and lipooxygenase inhibitors, among others, can be used to block the inflammation physiological pathways which mediate pain and progression of the disease. However, because these drugs are limited in their efficacy in advanced or more severe stages of arthritis, these agents are usually add-on therapies.
NSAIDs principally act by inhibiting prostaglandin and leukotriene synthesis. These compounds inhibit key enzymes in the biosynthetic pathway, i.e. cyclooxygenase. There are drugs that selectively inhibit isoforms of cyclooxygenase 1 and 2 (COX-1, COX-2). These drugs enhance patient tolerance due to the prevalence of COX-2 induction occurring in inflammation mediated by cytokines and others.

Further, pyrimidine synthesis inhibitors can be used as an antiinflammatory agent in arthritis, e.g. lefunomide.

Disease-Modifying Antiinflammatory Drugs or agents: Agents involved in the modification of clinical disease manifestation, reduction in inflammation, or slowing the disease progression are referred to as disease-modifying antiinflammatory drugs (DMARDs). They include gold salts (auranthioglycine, aurothiomalate, auranofin), hypoglycemics (angiotension converting enzyme inhibitors), allopurinol, immunosuppressives (azathioprine, cyclosporine), agents to treat metallic poison (penicillamine), depen, naprosen, isumun, antimalarial (chloroquine, hydroxychloroquine), alkylating agents (cyclophosphamide), absorbable sulfonanides (sulfasalazine), irriants and counter-irritants (capsaicin), antimicrobial agents (tetracyclines), and antinflammatory (methylprednisolone).

Hormones and Growth Factors: Agents acting at hormone receptors or growth factor receptors include steroids (glucocorticoids), adrenocorticotrophic hormone (corticotropic), and tumor necrosis factor inhibitors (soluble TNF receptors (enbrel) and TNF monoclonal antibody (remicade). The autoinmunity component of the disease is driven primarily by activated T-cells and give rise to cytokines IL-1 and TNF at the rheumatoid synovium. These agents are known to interfere with the actions of these cytokines.

Corticosteroids affect the inflammation within the joints by: decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-alpha, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

Skin and mucosal membrane preparations: Irritants and counter-irritants can be used to treat arthritic joints, such as but not limited to, Capaicin.

Chlorambucil, cyclosporine, cyclophosphamide are agents that are available for use in the treatment of refractory RA or with severe extraarticular complications such as vasculitis, corneal perforation or other severe systemic maladies associated with RA.

Limitations of Current Therapies for Arthritis

Low Efficacy Limitations of Therapies for Arthritis

The therapies discussed above are limited to the slowing or retarding the progression of arthritis. As joint degeneration progresses and result in irreversible damage, options become limited. Thus, therapies for arthritis are aimed at reducing manifestation of symptoms by controlling the clinical manifestations of inflammation.

Reduction in clinical symptoms of arthritis following DMARDs therapy is only evident after several weeks to months of therapy. Slow clinical relevance of these therapies limits determination of optimal therapy for individuals with arthritis and makes it difficult to select optimal therapy for any given stage of the disease.

Toxicity or Undesired Side Effects as Therapeutic Limitations of Arthritis

Toxicity and undesired side effects associated with above current therapies for arthritis require monitoring. Drugs used to treat arthritis may cause death, disability, disease, and pose risks to an unborn child. The undesired side effects or toxicities are listed for each drug category as described above.

Analgesics associated side effects include dyspepsia, gastric or small bowel bleeding, ulceration, renal insufficiency, confusion, rash, headache, hepatic toxicity. NSAIDs also reversibly inhibit platelet aggregation and prolong bleeding time.

Antirheumatic agents (DMARDs) associated side effects include: antimalarials—retinal or macular damage; sulfonamides—hematologic toxicities (leukopenia, thrombocytopenia, hemolysis in patients with glucose 6-phosphate dehydrogenase (G6PD) deficiency); antinmometabolites—hepatic compromise including hepatic fibrosis, ascites, esophageal varices, cirrhosis, pneumonitis, myelosuppression; immunosuppressives—myelosuppression, (cyclophosphorine; renal insufficiency anemia, hypertension); agents to treat metallic poison—rash, stomatitis, dysgeusia or metallic taste, myelosuppression (thrombocytopenia), proteinuria, nephritic syndrome or renal failure, and induction of autoimmune syndromes (systemic lupus erythematosus, myasthenia gravis, polymyositis, Goodpasture’s syndrome), gold preparations: hematologic, renal, pulmonary, and proteinuria; and chlorambucil—myelosuppression, myeloproliferative disorders, malignancy, and hemorrhagic cystitis.

Soluble TNF receptors agents have been shown to induce sepsis and predispose patients to serious infections. Further this product was associated with site of injection reactions, infections, and headache.

Glucocorticoid associated side effects include increased appetite, weight gain, fluid retention, acne, ecchymosis, development of cushion facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, psedo-tumor cerebri, psychosis, glaucoma, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamic-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

Since the majority of RA patients are women in their reproductive years, the level and extent the agents used to treat RA affects or has a potential to affect mothers during pregnancy or the developing fetus, either across the placenta or through breast milk. Clinical medical therapeutic decisions should weigh the use of all of the above current therapies for RA against known capacity of these agents to affect both the mother and the child.
Description of Mechanism of Action Hypotheses for Future Drug Development for Arthritis

[0193] Rheumatoid arthritis has been thought to be the result of host genetic factors, immunoregulatory abnormalities and autoimmunity, and triggering or persistent microbial infection.

[0194] Host genetic factors. The HLA-DR4 antigen (HLA, human leukocyte antigen) is significantly increased in RA patients. Recent studies show that a subtype of the HLA-DR4 share similar epitope among several MHC class II molecules and predispose to RA.

[0195] Autoimmune component. In over 80% of RA patients autoantibodies to the Fc portion of IgG (rheumatoid factors, RF) are present and can be used to determine diagnosis. The higher the titer of RFs, joint disease and extrarticular manifestations are more severe.

[0196] Related to the autoimmune component of the disease, ICAM-1 inhibitors, or other agents to reduce adhesion have been developed.

[0197] Microbial Infections. Of all the examined pathogens, only the Epstein-Barr virus (EBV) has remained unproven as a cause of RA. EBV has been shown to share a similar epitope as the HLA-DR4 epitopes, but EBV is ubiquitous and has yet to be a proven cause of RA.

[0198] A gene, genes, or gene pathway involved in the etiology of arthritis or associated disorders or potential sites for targeted drug therapy of arthritis are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of arthritis are listed in Table 5.

B. Chronic Obstructive Pulmonary Disease

Description of Chronic Obstructive Pulmonary Disease

[0199] Chronic obstructive pulmonary disease (COPD) is an imperfect term that refers to four pulmonary disorders: simple chronic bronchitis, chronic bronchitis, chronic obstructive bronchitis, and emphysema. A common characteristic of the disease is airway obstruction. Airways obstruction denotes the slowing of forced expiration. A decrease in the forced expiratory volume in 1 second (FEV1) to forced vital capacity (FVC) indicates that airflow is impaired. Forced expiration is determined primarily by intrinsic resistance of the airways, compressibility of the airways, and lung elastic recoil. Reduced maximal expiratory flow results from high airway resistance, reduced lung recoil, or excessive airways collapsibility. The overall cost of these illnesses to society is enormous due to the extent of the number of individuals afflicted with COPD. Approximately 15 million Americans are afflicted, and COPD is currently the fourth-leading cause of mortality. The high morbidity and mortality rates associated with COPD are linked to the failure to identify at-risk patients for intervention. Large reserves of pulmonary function in lungs and the slow progressive nature of the disease can often delay the clinical diagnosis and therapeutic intervention.

[0200] Simple chronic bronchitis is predominantly characterized by chronic productive cough and is usually the result of low-grade exposure to bronchial irritants. This syndrome is associated with enhanced mucous secretion, reduced ciliary activity, and impaired resistance to bronchial infection. Bronchitis patients range from those who experience sporadic cough producing mucous to those with a severe, disabling condition manifested by one or more of the following symptoms of increased resistance to airflow, hypoxia, hypercapnia, and irreversible narrowing of the small airways such as bronchiolitis and bronchiectasis.

[0201] Repeated exposure to bronchial irritants in individuals with hyperactive or sensitive airways can lead to bronchospasm, i.e. bronchial smooth muscle constriction, that is frequently accompanied by excess mucous production and edema of the bronchial walls. Episodic bronchospasm in individuals with chronic bronchitis is termed asthmatic bronchitis and is applied to those individuals with chronic airway constriction, chronic productive cough, and episodic bronchospasm.

[0202] Emphysema is characterized by abnormal, excessive, permanent enlargement of airway spaces distal to the terminal bronchioles, and is accompanied by destruction of their walls and may or may not be associated with fibrotic tissue. These changes result in a reduction of elastic recoil permitting excessive airway collapse upon expiration and leads to irreversible airway flow obstruction. Emphysema is strongly related to and correlated to inhalation of tobacco smoke such as cigarette or cigar smoking.

[0203] In emphysema, there is a loss of elastic recoil leading to pulmonary hyperinflation. The hyperinflation reaches a limit when the diaphragm is pushed flat and no longer functions effectively. The chest wall is expanded to the point that it pushes inward rather than exerting its normal outward force. These anatomical changes alter inspiration to the point that exertion is nearly impossible.

[0204] A deficiency in alpha 1-antitrypsin can predispose individuals to signs and symptoms of COPD. In these individuals there is a marked alveolar wall destruction with a non-uniform pattern of air space enlargement. In these patients there may be excessive formation of thick mucous and is often accompanied by persistent cough.

[0205] Complications of COPD include hypoxemia, cor pulmonale, hypercapnia, and dyspnea. Sustained chronic hypoxemia is a condition that leads to pulmonary vasoconstriction that with time becomes irreversible and leads to cor pulmonale.

Current Therapies for COPD

[0206] Current therapies for treatment of COPD are aimed at: reducing the airway obstruction reversibly, controlling the persistent cough and sputum production, reducing or eliminating airway infections, increasing exercise tolerance to the maximum allowable at the individual's level of physiological deficit, controlling the remedial disease complications, i.e. cardiovascular dysfunction and arterial hypoxemia, and relieving anxiety and depression or other psychiatric symptoms that accompany patients attempts to cope with the debilitating clinical manifestations. Lastly, all treatment regimens include education and supportive therapy to encourage patients with COPD to cease behaviors that may exacerbate symptoms such as inhalation of pulmonary irritants, i.e. smoking and others, and substance abuse, i.e. narcotics and sedatives.
Bronchodilators can be inhaled or administered by oral, subcutaneous, or intravenous routes.

Beta-adrenergic agonists or other sympathomimetic agents are used to produce rapid acute bronchodilation.

Anticholinergics agents are used to produce sustained bronchodilation. Nebulized atropine has been supplemented with the advent of a quaternary ammonium salt, ipratropium bromide, which undergoes minimal systemic absorption and thus has limited anticholinergic toxicity. Ipratropium has been shown to be effective in patients that have not responded to beta-adrenergic agonists and can reduce sputum volume without altering viscosity.

Anticholinergics and beta-adrenergic agonist combinations have been used with some success. Such combinations reduce the need to administer high doses, due to additive effects, and therefore reduce the likelihood for adverse effects or toxic side effects.

Theophylline is a methylxanthine bronchodilator. Theophylline improves airflow flow, decreases dyspnea, reduces pulmonary arterial pressure, increases arterial oxygen tension, improves diaphragmatic strength and endurance, increases right ventricular function (pulmonary vasodilator and cardiac inotropic effects), and may produce antiinflammatory effects.

Expectorants can be used to increase secretion clearance in patients with COPD. Although this therapy has not been demonstrated to render clinical benefit, it is an add-on therapy that enables the patient to experience an enhanced productive cough.

Anti-Inflammatory agents

Prolonged use of corticosteroids have been used to retard the rate of decline in FEV1 in COPD subjects. However, it has been determined that systemic corticosteroids are beneficial for acute exacerbations of COPD but are not used for long-term treatment and have not been proven to retard the progression of the disease. Corticosteroids affect the decline of FEV1 in the airways by: decreasing growth and development of mast cells, inducing apoptosis; suppressing lymphocyte generation of IL-5 and other cytokines; inhibiting some mediator release; inhibiting cytokine production; inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF); and inhibiting nitric oxide synthesis.

Alpha1-protease inhibitor deficiency as a cause of early development of emphysema has increased the awareness of the role of protease-antiprotease and oxidant-antioxidant imbalances in COPD. Intravenous delivery of alpha1-protease inhibitor can provide the appropriate levels for individuals with a genetic deficiency and individuals with acquired deficiency.

Antiproteases and Antioxidants

Alpha1-protease inhibitor deficiency as a cause of early development of emphysema has increased the awareness of the role of protease-antiprotease and oxidant-antioxidant imbalances in COPD. Intravenous delivery of alpha1-protease inhibitor can provide the appropriate levels for individuals with a genetic deficiency and individuals with acquired deficiency.

Mucolytics and secretion clearance agents can be used to assist in the removal of secretions during productive cough. These agents can thin secretions in patients with chronic bronchitis.

Supplemental oxygen therapy is used to treat detrimental effects of sustained chronic hypoxemia and hypercapnia. Correction of this condition is one of the treatments shown to have a positive effect on the survival rate in patients with COPD.

Treatment of cases of cor pulmonale includes the use of diuretics and positive inotropic agents such as digitalis. Careful monitoring is needed in these patients due to a development of marked right ventricular hypertrophy.

Dyspnea may be severely disabling despite aggressive therapy. Judicious use of opiates to control dyspnea and increase exercise tolerance have been proven beneficial. Unfortunately, opiates can have a respiratory depressive effect and care should be taken to deliver the appropriate therapeutic dose.

Many patients with COPD find themselves anxious or depressed or both. Appropriate use of psychostimulants can be used to control the signs and symptoms of anxiety and depression.

Surgical procedures can be performed to attempt to restore pulmonary capacity and function. Lung volume reduction surgery is useful to remove a portion of emphysematous lung tissue so that the diaphragm can return to its normal dome shape and the chest wall can reassume its normal configuration, mechanics, and physiology. Bullectomy is a procedure in which large bullae and surrounding lung tissue are removed. This allows for the remaining tissue to expand and once again function normally. Another procedure is lung transplantation. This expensive and aggressive approach is usually reserved for younger patients, particularly those who are alpha 1-antitrypsin deficient.

Limitations of Current Therapies for COPD

The most common limitations for the use of bronchodilators is the mistaken use of inhalants and inadequate patient education.

Beta adrenergic therapy is limited by three factors: 1) the density of beta2 receptors in the airways decreases with age, 2) despite the selectivity of the beta2 receptor agonists, there is cross reactivity to P1 receptors and may affect the myocardium and other peripheral tissues, and 3) there is beta-adrenergic receptor desensitization. Most of the recommended doses of beta adrenergic agonists provide less than maximal bronchodilation. Beta-adrenergic agonists can cause tremor, reflex tachycardia, tachyphylaxis, cardiomyopathy, and other cardiac toxic effects. Tachycardia is particularly problematic in the elderly or for those individuals who are at cardiac risk. Further, beta-adrenergic agonists have been shown to cause hyperkalemia. The majority of patients with COPD are current or former smokers, all of whom are at risk of coexisting coronary artery disease, thus in the compendium of therapies it is desirable to have alternatives to beta-adrenergic agonists.

Anticholinergics as bronchodilators have been associated with systemic side effects. In particular, systemic anticholinergic side effects include bradycardia (if pronounced, includes compensatory tachycardia), dry mouth, inhibition of sweating, dilatation of the pupils, and visual blurring. Ipratropium has a slow onset of action and a longer duration of action than beta-adrenergic agonists which can be
Theophylline continues to be a controversial treatment due to misconceptions of its role as a bronchodilator, drug delivery problems, and conflicting results of comparative studies during acute exacerbations. Further, theophylline has a limited therapeutic window, i.e. the dose needed to achieve bronchodilation is close to the dose associated with undesirable or adverse side effects including convulsions, cardiac arrhythmias, tachycardia, vasodilation, and diuresis. Further complicating therapy with theophylline is the intra-patient variability in efficacious response.

Long-term use of corticosteroids can be useful for patients in which continued symptoms or severe airflow limitations exist despite therapy with other agents. Only 20-30% of these patients experience therapeutic benefit for long-term use and indiscriminate use often leads to adverse effects without benefits. Unfortunately there have not been identified predictors of responders or nonresponders to long term steroid use in patients with COPD. Therefore, only those patients that attempt long-term steroid use and have documented clinical improvement should continue steroid therapy. Unfortunately, those patients in which long-term steroid use results in no benefit are subjected to potential adverse effects or toxicities. Glucocorticoid associated side effects include: increased appetite, weight gain, fluid retention, acne, eczymnosis, development of cushoid facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, pseudotumor cerebri, psychosis, glaucoma, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamic-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

Mucolytic and secretion clearance agents have been shown to improve thinning secretions. So far, there is little evidence to suggest that these agents render clinical improvement. Further cough suppressants may impair secretion clearance and possibly increase risk of pulmonary infection.

Description of Mechanism of Action Hypotheses for Future Drug Development of Candidate Therapeutic Interventions of COPD

Since the predominant smokers, smoking cessation programs and agents used to help patients quit smoking will be a valuable addition to therapeutic regimens. Nicotine replacement therapies such as nicotine patches (transdermal), gum, and transnasal formulations as well as bupropion (an antidepressant or other in this category) should be considered.

Other therapies to be considered are novel bronchodilators for inhalation therapy without the use of chofluorohydrocarbons (CFCs), next generation anticholinergic therapies, alpha 1 antitrypsinase augmentation therapies, and refinement of surgical procedures.

A gene, genes, or gene pathway involved in the etiology of COPD or associated disorders or potential sites for targeted drug therapy of COPD are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of COPD are listed in Table 6.

C. Autoimmune Disease

Description of Autoimmune Disease

An immune response to “self” antigens, or autoimmunity, can vary from minimal to severe depending on the extent of the loss of self tolerance and to localization of antigens. There is a distinction between autoimmune response, which may or may not be pathologic, and autoimmune disease, which does lead to pathologic conditions. In autoimmune disease there is a combination of the following types of evidence: 1) identification of the target antigens, 2) identification and isolation of self-reactive autoantibodies or self-reactive lymphocytes, 3) identification of clinical evidence, i.e. familial hereditary data, lymphocyte infiltration, MHC association and clinical symptomatic improvement with immunosuppressive agents. Initiation of autoimmune disease is thought to require one or more of the following: genetic predisposition to loss of tolerance, environmental factors that stimulate aberrant immune response, or loss or dysfunction of cellular or organ physiological processes leading to pathological immune response. Since many auto-reactive clones of T and B cells exist and are normally regulated by homeostatic mechanisms, loss or breakdown of this system of checks and balances can lead to activation or enhancement of these auto-reactive clones and ultimately lead to autoimmune disease.

There are few autoimmune disease indications wherein inflammation and immune response gene pathways should be considered in the stratification or therapeutic choice of patient groups based upon genotype. There are multiple examples of autoimmune diseases or diseases that have an autoimmune component. They include: amyotrophic lateral sclerosis, anti-phospholipid syndrome, aplastic anemia, autoimmune hemolytic anemia, diabetes mellitus type 1, Guillain-Barre syndrome, idiopathic thrombocytopenic purpura, Grave’s disease, myasthenia gravis, polyarthritis, rheumatoid arthritis, Hashimoto’s thyroiditis, uveitis, Wegener granulomatosis, periarteritis nodosa, ocular pemphigoid, pemphigus vulgaris, psoriasis, Goodpasture’s syndrome, Churg-Strauss vasculitis, poly-dermatomyositis, Cogan syndrome-autoimmune inner ear disease, hemolytic uremic syndrome, idiopathic glomerulonephritis, inflammatory bowel disease, Crohn’s disease, microscopic polyarteritis, and multifocal motoneuron neuropathy. Here we discuss four specific diseases that represent larger patient populations and are representative of diseases in which therapy can be aimed at suppressing the hyperactivity of the immune system. These include multiple sclerosis, systemic lupus erythematosus, scleroderma, diabetes mellitus type I, sarcoidosis, and nephritis.

Multiple Sclerosis

Multiple sclerosis (MS) is a disorder of multifocal sites of myelin sheath destruction, perivascular-lymphocytic cuffing and variable degree of oligodendroglial loss. In profound cases, there is glial, axonal transection, and neuronal and axonal loss. There are an estimated 300,000 Americans diagnosed with MS. The estimated cost of MS is $5 billion dollars.

Clinically, MS begins with a relapsing illness with episodes of neurological dysfunction lasting several weeks,
followed by substantial or complete improvement. This is identified as the relapsing-remitting stage of the disease found to be predominantly in females (1.6:1). There are some patients that remain in this stage of the disease for decades; others may rapidly progress to the next stage. As time progresses, and repeated relapses occur, recovery becomes less and less complete or as substantial. In these cases, a gradual relapse independent clinical progression develops and is termed secondary progressive MS. Further, a nonrelapsing-nonremitting form is characterized by a gradual progression and steady worsening of neurological function without any recovery or improvement. A steady but gradual neurological decline, predominately identified in males, characterizes the primary progressive form of MS. Clarity in understanding the significance of these varying disease patterns and diagnosis is dependent on quality neurological examination overtime.

[0239] Systemic Lupus Erythematous

[0240] Systemic lupus erythmatosus (SLE) is a disease characterized by inflammation in many different organ systems associated with the production of antibodies to reactive to nuclear, cytoplasmic, and cell membrane antigens. Clinical manifestations of the disease include reddish rash on the cheeks, fatigue, anemia, rashes, sun sensitivity, alopecia, arthritis, pericarditis, pleurisy, vasculitis, nephritis, and central nervous system disease. The immune hyperreactivity appears to derive from immune hyperreactivity and loss of self-tolerance. In these patients antibodies are produced against several nuclear components, notably antinuclear antibodies to native double stranded DNA, single stranded DNA, or nucleohistones.

[0241] Scleroderma

[0242] Scleroderma is a chronic disease marked by increases of fibrotic tissue involving the circulatory system, connective tissue (in particular the skin), visceral organs, and the immune system. There are approximately 500-700, 000 Americans diagnosed with scleroderma. There are two types of scleroderma, localized and systemic. In localized scleroderma (linear and morphea) the disorder of the connective tissue is limited to the skin, the tissues just beneath the skin, and muscle. Internal organs are not affected. In systemic scleroderma (sclerosis) vascular, digestive, pulmonary, renal, muscle and joints may be affected. Raynaud’s syndrome (frequent spasms of small arteries induced by temperature changes and emotion resulting in deprivation of blood supply to peripheral tissues), CREST syndrome (calcium deposits, Raynaud’s syndrome, loss of muscular control of the esophagus, sclerodactyly, and telangiectasia), and Sjogren’s syndrome (inflammation of the conductive, cornea, tear, and salivary glands with progressive destruction by lymphocytes and plasma cells) are both subcategories of scleroderma.

[0243] The clinical manifestations of scleroderma include the following symptoms: fatigue, swelling and numbness of the hands and feet, shiny skin and disappearance of skin folds, ulcers on the fingers, calcium deposits on the fingers, joint inflammation, joints tightening into bend position, muscle weakness, itchy skin, difficulty in swallowing, shortness of breath, fatty diarrhea or constipation, and loss of body hair. Although ultimately renal impairment and failure is a common endpoint, therapy affecting the hypertensive phase or renal involvement has changed the mortality rate.

[0244] Diabetes Mellitus Type I

[0245] This form of diabetes involves the chronic inflammatory destruction of the insulin-producing islet cells of the pancreas. Although this form of diabetes is treated similarly to the type II form (which is not linked to autoimmunity), i.e. insulin replacement therapy, early identification of type I versus type II individuals may be useful to thwart the autoimmune destruction of the β-cells. There are an estimated 500,000 to 1 million Americans that have type 1 diabetes, it is the seventh leading cause of death, and the following is a list of the progressive complications that are associated with the unregulated carbohydrate balance in tissues: retinopathy leading to blindness, nephropathy (diabetic nephropathy is the leading cause of end-stage renal disease), coronary and cardiovascular disease, neuropathy (severe forms can lead to amputation), impotence (diabetic neuropathy and cardiovascular disease can lead to impotence), and stroke.

[0246] Sarcoïdosis

[0247] Sarcoïdosis is a granulomatous disorder characterized by enhanced cellular immune response at one or more involved sites. The prevalence of sarcoïdosis is 5 in 100,000, so approximately 13,000 patients have been diagnosed. Between 80-90% of patients with sarcoïdosis have pulmonary involvement, however, any organ can be affected. Pulmonary involvement includes dyspnea with or without exertion, persistent dry cough, and atypical chest pain. Cor pulmonale can develop as a complication of pulmonary dysfunction and further progress to right atria dilatation and right ventricular hypertrophy. Ocular involvement includes disturbance in visual acuity, and in chronic cases may lead to glaucoma, cataract formation and retinal neovascularization. In 80% of the cases, sarcoïdosis is self-limiting and results in minimal symptomology, discomfort, or debilitation. However in the remaining 20%, sarcoïdosis patients face potentially serious debilitation, disfigurement, and can be life threatening. Misdagnosis is frequent and can limit appropriate therapeutic intervention.

[0248] Nephritis

[0249] Inflammation of the kidneys results in impaired renal function. Nephritis can be either interstitial or glomerular. In either case, mononuclear cells infiltrate in the interstitium of the renal cortex. Eosinophils, and in some cases, polymorphonuclear leukocytes are found in a similar compartment. The infiltrate may be diffuse or patchy and may be accompanied by fibrinotic tissue. Membranous nephropathy may develop and lead to impairment of glomerular filtration rate. There is evidence to suggest both cytotoxic T cells and T-cell mediated delayed hypersensitivity are involved. Nephritis is a component of the clinical manifestation of systemic lupus erythematosis, scleroderma, and other autoimmune diseases and disorders.

Current therapy for Autoimmune Diseases and Disorders

[0250] Agents used to treat autoimmune disease fall into one of the following four categories: analgesics (NSAIDs, salicylates), immunosuppressive agents, hormones (glucocorticoids), and skin and mucosal membrane preparations

[0251] Analgesics: Typically, pain associated with autoimmune disease can be controlled with NSAIDs including but not excluded to, salicylates, para-aminophenol derivatives,
indole and indene derivatives, heteroaroyl acetic acids, arylpropionic acids, anthranilic acids, enolic acids, or alkanones. Antinflammatory agents such as cyclooxygenase inhibitors, lipoxigenase inhibitors, and others can be used to block the inflammation physiological pathway which mediates pain. However, because these drugs are limited in their efficacy in advanced or more severe stages of autoimmune disease, these agents are add-on therapies.

[0252] NSAIDs derive their principle mechanism of action by the inhibition of prostaglandin and leukotriene synthesis. These compounds inhibit key enzymes in the biosynthetic pathway, i.e. cyclooxygenase. There are drugs that selectively inhibit isoforms of cyclooxygenase 1 and 2 (COX-1, COX-2) which enhances patient tolerance due to the prevalence of COX-2 induction occurs in inflammation mediated by cytokines and others.

[0253] Immunosuppressive drugs or agents: Agents involved in the modification of the immune system for the treatment of autoimmune disease are immunosuppressive agents. Immunosuppressives include azathioprine, cyclosporine, penicillin, antimalarials (chloroquine, hydroxychloroquine), alkyllating agents (cyclophosphamide), and antimetabolites (methotrexate).

[0254] Hormones and Growth Factors: Agents acting at hormone receptors or growth factor receptors include steroids (glucocorticoids), adrenocorticotrophic hormone (corticotropic), and tumor necrosis factor inhibitors ( soluble TNF receptors (entrel) and TNF monoclonal antibody (remicade). Since the autoimmunity component of the disease is driven primarily by activated T-cells, which gives rise to cytokines IL-1 and TNF at the affected areas. These agents are known to interfere with the actions of these cytokines.

[0255] Corticosteroids affect the immune response by: decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

[0256] Plasma Exchange: A useful technique for the removal of autoantibodies is a process called plasmapheresis or plasma exchange. In this process, antibodies are removed that mediate humoral immune response to the autoantigen.

[0257] Antioxidants: Many of the therapies in use for these autoimmune diseases are aimed at reducing the level and extent of tissue damage mediated by T-cell immune response. For example, dimethyl sulfoxide, dimethyl sulfone, para-aminobenzoic acid, and vitamin E are included in this category.

Limitations of Current Therapies for Autoimmune Disease

Limitations Current Therapies for Autoimmune Disease Based upon Low Efficacy

[0258] The therapies discussed above are related to the slowing or retarding the progression of autoimmune disease. As immune response tissue damage occurs, degeneration of the function progresses, irreversible damage occurs, and therapeutic options become limited. Thus, therapies for autoimmune disease are aimed at reduction of manifestation of symptoms by controlling the clinical manifestations of inflammation and the hypersensitive immune response.

[0259] The reduction of clinical symptoms of autoimmune disease following immunosuppressive therapy by one of the agents listed above is only evident after several weeks to months after therapy. The slow clinical relevance of these therapies limits the clinician to determine optimal therapy for individuals with autoimmune disease, and provides a risk for selection of optimal therapy for any given stage of the disease. Furthermore, there may be delays in identifying those patients that have an autoimmune hyperreactivity, and this can delay therapeutic intervention.

Limitations Current Therapies for Autoimmune Disease Based Upon Toxicity or Undesired Side Effects

[0260] There are toxicities and undesired side effects associated with the above current therapies for autoimmune disease that require monitoring. Drugs used to treat autoimmune disease may cause death, disability, disease, and place an unborn child at risk. The undesired side effects or toxicities are listed for each drug category as described above.

[0261] Analgesics associated side effects include dyspepsia, gastric or small bowel bleeding, ulceration, renal insufficiency, confusion, rash, headache, hepatic toxicity. NSAIDs also reversibly inhibit platelet aggregation and prolong bleeding time.

[0262] Immunosuppressive therapies have associated side effects including—antimalarials: retinal or macular damage; sulfonamides: hematologic toxicities (leukopenia, thrombocytopenia, hemolysis in patients with glucose 6-phosphate dehydrogenase (G6PD) deficiency); antimetabolites: hepatic compromise including hepatic fibrosis, ascites, esophageal varices, cirrhosis, pneumonitis, myelosuppression; immunosuppressives: myelosuppression, (cyclosporine: renal insufficiency anemia, hypertension); penicillin: rash, stomatitis, dysgeusia or metallic taste, myelosuppression (thrombocytopenia), proteinuria, nephrotic syndrome or renal failure, and induction of autoimmune syndromes (systemic lupus erythematosus, myasthenia gravis, polymyocytis, Goodpasture’s syndrome).

[0263] Glucocorticoid associated side effects include: increased appetite, weight gain, fluid retention, acne, ecchymosis, development of cushion facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, pseudotumor cerebri, psychosis, glaucoma, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamus-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

[0264] Since the majority of autoimmune disease patients are women in their reproductive years, the level and extent the agents used to treat autoimmune disease affects or has a potential to affect the mother during pregnancy, cross the placenta, affect the developing fetus, or be excreted in breast milk during lactation are important issues facing the skilled practitioner. Clinical medical therapeutic decisions should weigh the use of all of the above current therapies for
autoimmune disease against known capacity of these agents to affect both the mother and the child.

Description of Mechanism of Action Hypotheses for Future Drug Development for the Treatment of Autoimmune Disease

[0265] Autoimmune disease has been thought to be the result of host genetic factors, immunoregulatory abnormalities and autoimmunity, and triggering or persistent microbial infection.

[0266] A gene, genes, or gene pathway involved in the etiology of autoimmune diseases or disorders or associated disorders or potential sites for targeted drug therapy of autoimmunity are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development are listed for the treatment of autoimmune disease or disorder, Tables 7 and 9, and for systemic lupus erythematosus, Table 8.

D. Immunosuppression-Transplantation

Description of Transplantation

[0267] There are many different conditions in which medical or surgical therapy is unable to halt, retard, or treat the underlying disease, disorder, or dysfunction. Although many refractory diseases, disorders, or dysfunctions do not lead to severe cases, there are some in which the progression leads to conditions in which the remaining therapeutic alternative is replacement of the diseased tissue with normal donated tissue by transplantation. These end stage conditions include both primary disease or complications from a disease. For example whole organ transplantation is an end-stage therapeutic alternative in the following indications, end-stage cardiomyopathy, end-stage renal disease, pulmonary disease, cirrhosis of the liver, as well as other end-stage diseases affecting whole organ function.

[0268] Besides whole, or partial organ transplantation there are programs aimed at replacing cells in specific tissues to enable or restore physiologic function. For example, cellular transplantation includes, but are not excluded to, grafting bone marrow cells in patients with hematopoietic or lymphocytic cancers, dopaminergic producing cells in brains of patients with Parkinson’s disease, striated muscle cells in patient’s with Duchenne’s muscular dystrophy, myocytes or cardiomyocytes in patient’s with ischemic heart disease or cardiomyopathy, and replacement of neurons or astrocytes or glial cells in neurodegenerative disease including but not excluded to Alzheimer’s disease, amyotrophic lateral sclerosis, multiple sclerosis, Huntington’s disease, refractory pain, epilepsy, and stroke.

[0269] In this way, transplantation includes autografts, isografts, allografts or xenografts and can involve whole organ or cellular grafts. With the exception of autologous transplantation, all other transplantation procedures include pre- and post-surgical immunosuppression to blunt graft rejection or graft versus host disease. Successful immunosuppression in this setting includes an appropriate balance between the need to prevent the process of graft rejection and the risk of suppressing the recipient’s immune system to the extent that they become vulnerable to infection or other complications.

[0270] Transplantation is immunologically mediated. Both T cells and circulating antibodies are induced against allografts or xenografts. While the antibodies are responsible for rejection of erythrocytes, T-cells are mainly responsible for the rejection of most other type of tissue. The antigens found on grafted tissue which initiate the rapid rejection of an allograft are found on most cell membranes and are encoded by genes in the major histocompatibility complex (MHC) which are called the HLA. The structures encoded in these genes, MHC class I and class II molecules, are involved in the determining the discrimination between self and non-self. The degree of the histocompatibility between donor and recipient can be determined serologically, by genotyping, or by a mixed lymphocyte reaction. Survival of HLA mismatched allografts is prolonged by anti-inflammatory agents, cytotoxic agents, antimetabolites, and other modalities aimed at immunosuppressing the recipient. These approaches have proven clinical success in terms of graft survival and clinical symptomology.

[0271] Rejection can occur at any time, and is either hyperacute, acute or delayed. The rate, extent, and underlying mechanism of transplantation rejection varies dramatically from individual to individual. Physiological factors include patency of blood circulation, lymphatic drainage, expression of antigens on the graft, and others that can influence the rejection rate.

[0272] In hyperacute rejection, preexisting host antibodies to antigens found on the grafted tissue mount an immune response. These antibodies activate complement, followed by platelet activation and deposition causing swelling and interstitial hemorrhage in a whole organ graft, or specific cell targeting in a cellular transplant. Cell mediated immunity is not activated in the hyperacute response.

[0273] In acute rejection, infiltration of lymphocytes and macrophages recognize the foreign antigen on the graft cells, and initiate a cascade of intragraft events that ultimately leads to host cellular and humoral mediated destruction of the grafted tissue and if unchecked will result in irreversible loss of the graft. This acute process occurs rapidly and does not in the first stages affect the vital structures of a whole organ graft, which allows for identification of the process and implementation of therapy. In many cases, an acute rejection episode can be reversed, and approximately 30-50% of whole organ graft recipients undergo one or more of these episodes in the early transplant period.

[0274] Delayed or chronic rejection occurs in a slower process than acute rejection and ultimately leads to a gradual loss of function in the grafted tissues. In chronic or delayed rejection, both cell mediated immunity and humoral immunity is activated. Chronic rejection is characterized by arteriosclerosis, in which the smooth muscle cells lining the arteries in the graft organ proliferate to create lesions and lead to fibrosis, with a result of constricting blood flow. As a result of the chronic immune rejection, there is slow and progressive destruction of the grafted organ or cells. If damage to the tissue is extensive, very little can be done to save the graft.

Current Immunosuppressive Therapies

[0275] The goal of clinical immunosuppression in the transplantation setting is to control allograft rejection. Clinical immunosuppression involves the non-specific suppression of both cell-mediated and humoral immune reactivity to the grafted tissue. Although a number of methods have been
proposed, successful prolongation of graft survival has been attained through the use of a combination of therapies that suppress both the lymphocytic interaction and proliferation and therapies that deplete the pool of available lymphocytes.

Antiproliferative Agents

These agents are useful to blunt the proliferative phase of lymphocyte activation of the immune response.

Purine Analogs

Azathioprine acts to inhibit the proliferation of T cells. Azathioprine is cleaved to 6-mercaptopurine and it is this active compound that serves to suppress the T-cell mediated antigenic determination and engraftment. Azathioprine is a relatively non-selective immunosuppressive agent. Other agents in the same class as azathioprine, i.e. antimetabolites, include but are not excluded to, mercaptopurine, chlorambucil, and cyclophosphamide.

Pyrimidine Analogs

These agents (cytosine arabinoside) inhibit DNA synthesis and therefore have their greatest effect on the immune response during the proliferative phase of lymphocyte activation. These agents inhibit primary antibody response and have minimal effects on the cell-mediated immunity.

Folic Acid Analogs

These agents (methotrexate, aminopterin) inhibit dihydrofolate reductase preventing the conversion of folic acid to tetrahydrofolic acid. This conversion is necessary for the production of DNA and RNA.

Alkylationing Agents

These agents (nitrogen mustard, phenylalanine mustard, busulfan, cyclophosphamide) alter the structure of the DNA and RNA. These agents have reactive ring structures which combine with electron rich groups such as tertiary nitrogen in purines or pyrimidines, or —NH2, —COOH, —SH, —PO3H2 groups. These reactions alter the composition of the DNA, and if not repaired, chromosomal replication will be altered in activated proliferating cells. The use of alkylating agents in the setting of transplantation is time dependent and is effective just before or during the activation of the immune system by antigen. Cyclophosphamide has been shown to have a greater effect on B-cells rather than T-cells, thereby inhibiting the humoral response to a greater degree.

Antibiotics

These agents (actinomycin D, mitomycin C, puramycin, chloramphenicol) inhibit either nucleic acid or protein synthesis.

Cyclosporin acts by inhibiting the production of IL-2, which results in an inhibition of the proliferation of T and B lymphocytes. Cyclosporin is widely prescribed for transplantation patients due to the clinical advantage of potent immunosuppression with limited myelosuppression.

FK-506 (Tacrolimus) is an agent that acts by inhibiting the production of IL-2 which prevents the proliferation of T and B lymphocytes.

Mycophenolate mofetil is rapidly converted to mycophenolic acid which selectively inhibits T and B cell proliferation. Mycophenolate mofetil has an advantage over azathiprine because it does not damage chromosomes.

Lymphocyte Depletion agents

Antilymphocytic globulin (ALG) is an agent that binds to circulating T-lymphocytes and the cells coated with the ALG are lysed and cleared by the reticuloendothelial system. ALG is more commonly used for renal transplantation, showing little to no benefit for liver or bone marrow transplantation.

Radiation

Total lymphoid irradiation or total body irradiation is based upon the immunosuppression observed after this procedure was used in patients with Hodgkin's lymphoma. The radiation causes breakdown in the nucleic acid structure, and the effect is time dependent since there are systems within all cells for the repair of DNA. Since the radiation affects those cells in M or G2 phase, those cells in G1 or S phase are resistant.

Monoclonal Antibodies

A murine monoclonal antibody is available to deplete the circulating CD3 lymphocytes. This antibody reacts with the T3 recognition site of the T-lymphocytes and blocks the recognition of the Class I and II antigens. This leads to prevention of the activation of the effector lymphocytes. This antibody has been useful in the treatment of rejection of renal, pancreatic, hepatic, cardiac, and pulmonary whole organ transplantations.

Steroids—such as the glucocorticoids are widely used in transplantation in combination with other drugs. As well as providing antiinflammatory therapy, corticosteroids suppress immune function by inhibiting the activation of T cells. Corticosteroids affect the inflammation within the airways by: decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis. Steroids are highly effective in the early induction and maintenance regimens and are first line therapy in acute allograft rejection.

Blood transfusions can be used to cause allo sensitization if the recipient is exposed to donor antigens in the presence of azathiprine. In this way, induction of a specific degree of hyporeactivity against graft antigens can result by a potential suppressor cell phenomena.

Limitations of Immunosuppressive Therapies

Limitations of Immunosuppressive Therapies Due to Lack of Efficacy

As suggested, the efficacy of immunosuppression is a balance between prevention of graft rejection or graft versus host disease and subjecting a patient unnecessarily to blunted immune defenses to ward off infections. All too often, this balance is not achieved and on one end the patient succumbs to infections or on the other the graft is rejected. It has been estimated that 30% of the transplantation patients are in this category.
Limitations of Immunosuppressive Therapies Due to Toxicities or Undesired Side Effects

Antiproliferative Agents
Azathioprine is associated with suppression of bone marrow production, and blood disorders including anemia, thrombocytopenia, and leukopenia. Hepatotoxicity occurs in a dose-independent manner, and is irreversible. Azathioprine is associated with chromosome damage and therefore is mutagenic.

Methotrexate and aminopterin are associated with bone marrow suppression, mucosal breakdown, gastrointestinal bleeding, megaloblastic hematopoiesis.

Alkylating Agents are associated with stomatitis, nausea, vomiting, diarrhea, skin rash, anemia, and alopecia. Specifically, cyclophosphamide has been associated with fluid retention, hemorrhagic cystitis, and cardiac toxicity.

Cyclosporin is associated with gingival hyperplasia, hirsutism, tremor, hypertension, hyperkalemia, hepatotoxicity, hyperglycemia, hypomagnesemia, hypercholesterolemia, hypertriglyceridemia, and hyperuricemia, nausea and gastrointestinal irregularities, and renal dysfunction. Nephrotoxicity associated with cyclosporin manifests as tubular necrosis, interstitial fibrosis, and tubular atrophy.

FK506 is associated with neurotoxicity, nephrotoxicity, and disturbances of glucose metabolism. The major neurotoxic symptoms are reversible and dose dependent and include headache, tremors, parasthesias, insomnia, increased sensitivity to light, mood changes, aphasia, and seizures. There has been a suggested association of FK-506 with cardiomyopathy and it is contraindicated in pregnancy.

Lymphocyte Depletion Agents
AI.Gs are associated with anemia, thrombocytopenia, and allergic reactions including urticaria, anaphylactoid reactions, serum sickness, joint pain, fever, and malaise.

Radiation is associated with higher incidence of infections and chromosomal breakage and mutations.

Monoclonal antibody therapy has been associated with the production of human anti-mouse antibodies (HAMA) in 80% of the treated patients and the sensitization rate is 15-40% thus limiting retreatment rates. Side effects are fever, chills, nausea, vomiting, headache, dyspnea, wheezing, pulmonary edema, tachycardia, hypotension, aseptic meningitis, seizures, and coma. These symptoms are related to the inordinate release of cytokines TNF, IL-1, and interferon-gamma. Although these symptoms can be reduced by pretreatment with steroids, acetaminophen, or diphenhydramine the HAMA response precludes repeated use.

Steroids—Glucocorticoid associated side effects include: increased appetite, weight gain, fluid retention, acne, ecchymosis, development of cushion facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperinsular state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, psuedotumor cerebri, psychosis, glaucoma, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamic-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

Complications of Immunosuppression

In addition to the above listed toxicities and undesirable side effects, potent immunosuppression as required in the transplantation setting leads to prolonged immune compromise and predisposes the patient to infections (80% of the patients) and cancer (ranging between 10-40% of the patients). This risk has been proposed to result from impaired immune surveillance mechanisms, chronic antigenic stimulation, reactivation of latent oncogenic viruses and the direct oncogenic effects of the immunosuppressive agents.

Moreover, 40% of the deaths of transplant patients are attributable to the complications of infections or a combination of infection and graft rejection. The infections experienced by transplant patients are 50% bacterial, 30% viral, 15% fungal. Some of the common bacterial infections are Staphylococcus aureus, Staphylococcus epidermidis, and gram-negative rods in line sepsis. Urinary tract infections, pneumonias, wound infections, and surgical infections (including cholecystitis, appendicitis, diverticular disease, ulcer, etc.). Common viral infections include cytomegalovirus, Epstein-Barr virus, Herpes Simplex Virus, and variella zoster virus. Further, common fungal or protozoan infections include Candida albicans, Aspergillus flavus, Cryptococcus neoformans, Coccidiodes immitis, Histoplasma capsulatum, Nocardia asteroides, and Pneumocystis carinii.

Description of Mechanism of Action Hypotheses for Future Immunosuppressive Drug Development

The majority of the hypotheses for future therapeutic interventions for graft rejection and graft immunoreactivity are based upon the understanding the immunologic mechanisms that cause and perpetuate the rejection within the graft.

A gene, genes, or gene pathway involved in the etiology of transplantation or immunosuppression or associated disorders or potential sites for targeted drug therapy of transplantation are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of anemia are listed in Tables 9 and 10.

E. Pain Associated with Inflammation

Description of Pain Associated with Inflammation

Pain associated with inflammation can be caused by pathologic processes in somatic structures or viscera, or by prolonged dysfunction of parts the peripheral nervous system. Pain associated with inflammation may be the result of recurrent injuries, trauma, headache, arthritis, chronic obstructive pulmonary disease, psoriasis, or other pathologies. Pain associated with inflammation may be acute or chronic depending on the level and extent of the inflammation.

Current Therapies for Pain Associated with Inflammation

Therapeutic management of pain resulting from inflammation includes a three step ladder approach: non-opioid analgesics are stepwise prescribed in combination with moderate to potent opiates. The guidelines call for a determination by the patient and the physician of pain relief. Broadly speaking, the guidelines are as follows: mild pain is
treated with non-opioid analgesics, moderate or persisting pain is treated with a weak opioid plus non-opioid analgesics, and severe pain that persists or increases is treated with a potent opioid plus non-opioid analgesics.

[0318] Analgesics: Typically, pain associated with inflammation can be controlled with NSAIDs including but not excluded to, salicylates, para-aminophenol derivatives, indole and indene derivatives, heteroaryl acetic acids, aryloproionic acids, anthranilic acids, enolic acids, or alkamones. Antiinflammatory agents such as cyclooxygenase inhibitors, lipooxygenase inhibitors, and others can be used to block the inflammation physiological pathway which mediate pain and the progression of the disease. However, because these drugs are limited in their efficacy in advanced or more severe stages of arthritis, these agents are add-on therapies.

[0319] NSAIDs derive their principle mechanism of action by the inhibition of prostaglandin and leukotriene synthesis. These compounds inhibit key enzymes in the biosynthetic pathway, i.e. cyclooxygenase. There are drugs that selectively inhibit isoforms of cyclooxygenase 1 and 2 (COX-1, COX-2) which enhances patient tolerance due to the prevalence of COX-2 induction occurs in inflammation mediated by cytokines and others.

[0320] Further, pyrimidine synthesis inhibitors can be used as an antinflammatory agent in arthritis, e.g. leflunomide.

Limitations of Current Therapies for Pain Associated with Inflammation

Limitation of Therapies for Pain Associated with Inflammation Due to Low Efficacy

[0321] The therapies discussed above are limited to the slowing or retarding the progression of arthritis. As degeneration of the joints progresses, and irreveresible damage occurs, the options become limited. Thus, therapies for arthritis are aimed at reduction ofmanifestation of symptoms by controlling the clinical manifestations of inflammation.

Limitations of Therapies of Pain Associated with Inflammation due too Toxicity or Undesired Side Effects

[0322] Analgesics associated side effects include dyspepsia, gastric or small bowel bleeding, ulceration, renal insufficiency, confusion, rash, headache, hepatic toxicity. NSAIDs also reversibly inhibit platelet aggregation and prolong bleeding time.

Description of Mechanism of Action Hypotheses for Future Pain Associated with Inflammation Drug Development

[0323] The persistence of pain most likely involves a cascade of pathological neurochemical events that lead to abnormal sensory hyperexcitability and excitotoxicity. The genes listed in FIG. 1 are part of a pathway that are all involved in producing prostaglandins or leukotrienes, which are two potent mediators of inflammation. Inordinate levels of prostaglandins have been implicated in pain associated with inflammation, and several drugs target this branch of the pathway, to inhibit the action of leukotrienes. When a cell receives a pro-inflammatory stimulus, such as tumor necrosis factor, membrane phospholipids, or interleukin-1, as shown in the figure, membrane phospholipases are activated, and arachidonic acid is released from membrane phospholipids into the cell. The liberated arachidonic acid is then metabolized either by the cyclooxygenase enzymes, which leads to the production of prostaglandins, or the lipooxygenase family of enzymes, which leads to the production of leukotrienes. There are several types of prostaglandins and leukotrienes, and many of the enzymes listed here function to convert one form into another.

[0324] The presence of leukotrienes and prostaglandins can lead to a persistence of neural hyperexcitability involving a sequence of neuroplastic events.

[0325] A gene, genes, or gene pathway involved in the etiology of pain or associated disorders or potential sites for targeted drug therapy of pain are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of pain associated with inflammation are listed in Table 11.

F. Psoriasis

Description of Psoriasis

[0326] Papulosquamous skin disorders have diverse etiologies and include psoriasis, Reiter’s syndrome, pityriasis rosea, lichen planus, ointyriasis rubra pilaris, secondary syphilis, mycosis fungoides, and ichthyosis vulgaris.

[0327] Psoriasis is a genetically determined, chronic epidermal proliferative disease with an unpredictable course. Psoriasis appears as erythematous plaques with silvery, mica-like scales, and is usually nonpruritic. The plaques appear anywhere on the body and almost never involves the mucous membranes. There are variations of psoriasis including guttate psoriasis, inverse psoriasis, pustular psoriasis, erythroderma, and psoriatic arthritis. There is an increased prevalence of psoriasis in subjects with the HLA antigen BW17, B13, and BW37. Further, 30% of cases have a family history of psoriasis. The Koebner phenaomena is a hallmark characteristic of psoriasis, e.g. intense trauma (scratches or surgical incisions) to the skin induces new linear papulosquamous lesions.

[0328] This multifactorial disease is characterized by an accelerated cell cycle in an increased number of dividing cells that results in rapid epidermal cell proliferation. It is estimated that 4-5 million Americans have psoriasis, 100,000 have severe cases, and 1 in 20 have psoriatic arthritis.

Current Therapies for Psoriasis

[0329] The goals of the therapeutic regimens is to limit the epidermal proliferation underlying the dermal inflammation. There are both topical and systemic treatments available, however in either category the treatment suppresses the condition for only as long as is administered. The treatment of psoriasis entails a stepwise increase of extent of the therapy ranging from topical applications to phototherapy to systemic interventions to prevent the epidermal proliferation.

[0330] In the first step topical treatments include corticosteroid ointments, vitamin D containing ointments, preparations containing coal tar or anthralin, salicylic acid containing ointments, and various other moisturizers and bath solutions. These steps are aimed at reducing the itching, scaling, and progression of the lesions.
In the second step, phototherapy other than natural sunlight can be used to thwart the epidermal cell proliferation. In these cases, ultraviolet light is administered to affected areas or uniformly to the body. In phototherapy, light delivered to the skin activates porphyrin molecules. These activated molecules transfer their energy to form cytoxic singlet oxygen leading to lethal alteration of cellular membranes and subsequent tissue destruction. In UVB therapy, UVB light is administered alone or with ointments containing coal tar, anthralin, or salicylic acid. UVA light is administered with psoralen.

In the third step of therapeutic regimens for psoriasis, systemic agents are administered to those cases refractory to the previously described first two steps. These compounds include retinoids, methotrexate, hydroxyurea, cyclosporin, azathioprine, 5-fluorouracil, cyclophosphamide, vinblastine, dapsone, and sulfasalazine.

Limitations of Current Therapies for Psoriasis

The main limitation of the current therapies for psoriasis is that the drugs are only efficacious during the administration. Further, periods of remission and outbreaks are difficult to impossible to predict. It has been shown that patients should rotate their treatments to retain efficacy. This can lead to missed schedules and requires patient education. Lastly, for all the listed therapies there is unreliable efficacy in their ability to stop proliferation and inflammation of the lesions.

Toxicities of the current therapies include the following: phototherapy can lead to other skin lesions and sunburn. Cytotoxic agents used as immunosuppressive agents including methotrexate, 5-fluorouracil, cyclophosphamide, and vinblastine have associated side effects including hepatic compromise including hepatic fibrosis, ascites, esophageal varices, cirrhosis, pneumonitis, myelosuppression, cyclosporine: renal insufficiency anemia, hypertension.

A gene, genes, or gene pathway involved in the etiology of psoriasis or associated disorders or potential sites for targeted drug therapy of psoriasis are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of psoriasis are listed in Table 12.

G. Atherosclerosis

Atherosclerosis is a complex combination of hyperlipidemia, injury to the endothelium, and inflammation. The interaction of these multiple processes in association with local genetic and hemodynamic influences may promote the formation of atheromatous plaques as a reparative response of the arterial wall. Atherosclerotic plaques are composed of thrombogenic lipid-rich core protected by a fibrous cap comprising smooth muscle cells and inflammatory cells. The inflammatory cells are predominantly macrophages. As atherosclerotic plaques build blood flow is reduced creating ischemia in tissues down stream from the area of the plaque.

In another model, the stenosis created by the plaques may be a part of the resulting ischemic event. Frequently, less obstructive but more vulnerable plaques occur which are characterized by a thinned fibrous cap, marked lipid accumulation, a large number of macrophages, and a smaller amount of smooth muscle cells. It has been proposed that since these plaques are more prone to rupture creating contact with the highly thrombogenic materials of the lipid-rich nucleus of these lesions, thrombosis is stimulated.

Advanced atherosclerotic lesions are caused by a series of cellular and molecular events involving replication of smooth muscle cells and macrophages on the vessel wall. The interaction of these cells with the T lymphocytes can lead to a fibroproliferative response. Large amounts of connective tissue produced by these smooth muscle cells comprise macrophages, T lymphocytes, smooth muscle cells, connective tissue, necrotic residues, and varying amounts of lipids and lipoproteins.

Endothelial cells maintain the vessel surface in a non-thrombogenic state, preventing platelet and leukocyte adhesion, and in maintaining the vascular tonus by releasing nitric oxide, prostaglandin, and endothelin. These cells also produce growth factors, cytokines, and chemokines to maintain the integrity of the collagen- and proteoglycan-rich basement membrane. Changes in some of these functions may trigger cell interactions with monocytes, platelets, smooth muscle cells, and lymphocytes. Hyperlipidemia and hypercholesterolemia are sufficient to induce dysfunction of the endothelial modulation of the vasoactive reactions and arteriolar tone.

The inflammatory mechanisms involved in the initial events or atherosclerosis are classic components of a specialized type of chronic inflammatory response that precedes the migration and proliferation of smooth muscle cells of the vessel wall. The formation and accumulation of foam cells in the intima leads to the first stage of the atherosclerotic lesion. In this stage, the accumulation of fatty strain comprise a mixture of macrophages, lipids, and T lymphocytes representing a purely inflammatory response. If the stimulating agent is maintained, i.e. hyperlipidemia, hypercholesterolemia, or other risk factor, then the protective inflammatory response will also persist and they may become deleterious to the cells lining the arterial wall. This condition may lead to an intermediate lesion that may contain multiple smooth muscle cell layers, macrophages, and T lymphocytes. A fibrous capsule is formed covering the contents of the lesion.

There is evidence to suggest that the inflammatory process and specific immune mechanisms are involved in atherogenesis. At sites close to the plaque rupture, inflammatory processes are observed resulting from T cell-dependent autoimmune response. This may lead to an inflammatory reactions participating in the destabilization of the fibrous cap. Immunoglobulins, T lymphocytes, and macrophages are found in the plaques. B lymphocytes and plasmocytes may also be detected in the adventitia adjacent to the plaques. Autoimmune reactions against the oxidized lipoproteins have been observed. The macrophages are transformed into foam cells and in the presence of LDL, form immuno-complexes with the LDL by Fe fragments of the immuno-globulins. These LDL immuno-complexes can induce numerous metabolic and functional changes which can directly or indirectly damage the endothelial cells leading to the progression of the atherosclerotic lesion.
Despite the evidence of the involvement of the immune system in atherogenesis, the complexity of the immune reactions and response impairs the clarification of the involvement of these mechanisms at the various stages of atherosclerosis. The sequence of immune response event suggests an initial mechanism to respond to injury. However, this protective inflammatory response in the presence of persistent stimulus and the formation of a fibroproliferative response can be amplified.

Attempts to modify the specific cell interactions with growth factor mediators or intracellular signalling molecules has provided a window to the potential prevention or regression of the lesions.

A gene, genes, or gene pathway involved in the etiology of atherosclerosis or associated disorders or potential sites for targeted drug therapy of atherosclerosis are depicted in Table 2 with the specific gene list in Table 1. Current candidate therapeutic interventions in development for the treatment of atherosclerosis are listed in Table 13.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents for use in Inflammation and Inflammation-Mediated Diseases

The advantages of a clinical research and drug development program that includes the use of polymorphic genotyping for the stratification of patients for the appropriate selection of candidate therapeutic intervention includes: 1) identification of patients that may respond earlier and show signs and symptoms of clinical improvement or side effects and toxicities, 2) identification of the primary gene and relevant polymorphic variance that directly affects manifestation of efficacy, safety, toxicity, or undesired side effects, 3) identification of pathophysiologic relevant variance or variances and potential therapies affecting those allelic genotypes or haplotypes, and 4) identification of allelic variances or haplotypes in genes that indirectly affects efficacy, safety or both.

Based upon these advantages, designing and performing a clinical trial, either prospective or retrospective, which includes a genotype stratification arm will incorporate analysis of clinical outcomes and genetic variation associated with those outcomes, and hypothesis testing of the statistically relevant correlation of the genotypic stratification and therapeutic benefits. If statistical relevance is detectable, these studies will be incorporated into regulatory filings. Ultimately, these clinical trial data will be considered during the approval for marketing process, as well as, incorporated into accepted medical management of inflammatory disease.

Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions for Inflammation Disease

There is evidence to suggest that there are safety response differences to drug therapy in arthritis which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability as described in the Examples. The Detailed Description below demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmaceutical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Table 1, and pathway matrix Table 2 and discussed below are candidates for the genetic analysis and product development using the methods described above.

A sample of therapies approved or in development for preventing or treating the progression of the above described indications currently known in the art are listed in Tables 5 through 13. In these tables, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

Pharmacogenomics studies for these drugs, as well as other agents, drugs, compounds or candidate therapeutic interventions, could be performed by identifying genes that are involved in the function of a drug including, but not limited to is absorption, distribution metabolism, or elimination, the interaction of the drug with its target as well as potential alternative targets, the response of the cell to the binding of a drug to a target, the metabolism (including synthesis, biodistribution or elimination) of natural compounds which may alter the activity of the drug by complementary, competitive or allosteric mechanisms that potentiate or limit the effect of the drug, and genes involved in the etiology of the disease that alter its response to a particular class of therapeutic agents. It will be recognized to those skilled in the art that this broadly includes proteins involved in pharmacokinetics as well as genes involved in pharmacodynamics. This also includes genes that encode proteins homologous to the proteins believed to carry out the above functions, which are also worth evaluation as they may carry out similar functions. Together the foregoing proteins constitute the candidate genes for affecting response of a patient to the therapeutic intervention. Using the methods described above, variances in these genes can be identified, and research and clinical studies can be performed to establish an association between a drug response or toxicity and specific variances.

Identification of Interpatient Variation in Response; Identification of Genes and Variances Relevant to Drug Action; Development of Diagnostic Tests; and Use of Variance Status to Determine Treatment

Development of therapeutics in man follows a course from compound discovery and analysis in a laboratory (preclinical development) to testing the candidate therapeutic intervention in human subjects (clinical development). The preclinical development of candidate therapeutic interventions for use in the treatment of human diseases, disorders, or conditions begins at the discovery stage whereby a candidate therapy is tested in vitro to achieve a desired biochemical alteration of a biochemical or physiological event. If successful, the candidate is generally tested in animals to determine toxicity, adsorption, distribution, metabolism and excretion in a living species. Occasionally, there are available animal models that mimic human diseases, disorders, and conditions in which testing the candidate therapeutic intervention can provide support-
ive data to warrant proceeding to test the compound in humans. It is widely recognized that preclinical data is imperfect in predicting response to a compound in man. Both safety and efficacy have to ultimately be demonstrated in humans. Therefore, given economic constraints, and considering the complexities of human clinical trials, any technical advance that increases the likelihood of successfully developing and registering a compound, or getting new indications for a compound, or marketing a compound successfully against competing compounds or treatment regimens, will find immediate use. Indeed, there has been much written about the potential of pharmacogenetics to change the practice of medicine. In this application we provide descriptions of the methods one skilled in the art would use to advance compounds through clinical trials using genetic stratification as a tool to circumvent some of the difficulties typically encountered in clinical development, such as poor efficacy or toxicity. We also provide specific genes, variation in which may account for interpatient variation in treatment response, and further we provide specific exemplary variances in those genes that may account for variation in treatment response.

[0352] The study of sequence variation in genes that mediate and modulate the action of drugs may provide advances at virtually all stages of drug development. For example, identification of amino acid variances in a drug target during preclinical development would allow development of non-allele selective agents. During early clinical development, knowledge of variation in a gene related to drug action could be used to design a clinical trial in which the variances are taken account of by, for example, including secondary endpoints that incorporate an analysis of response rates in genetic subgroups. In later stages of clinical development the goal might be to first establish retrospectively whether a particular problem, such as liver toxicity, can be understood in terms of genetic subgroups, and thereby controlled using a genetic test to screen patients. Thus genetic analysis of drug response can aid successful development of therapeutic products at any stage of clinical development. Even after a compound has achieved regulatory approval its commercialization can be aided by the methods of this invention, for example by allowing identification of genetically defined responder subgroups in new indications (for which approval in the entire disease population could not be achieved) or by providing the basis for a marketing campaign that highlights the superior efficacy and/or safety of a compound coupled with a genetic test to identify preferential responders. Thus the methods of this invention will provide medical, economic and marketing advantages for products, and over the longer term increase therapeutic alternatives for patients.

[0353] As indicated in the Summary above, certain aspects of the present invention typically involve the following process, which need not occur separately or in the order stated. Not all of these described processes should be present in a particular method, or need be performed by a single entity or organization or person. Additionally, if certain of the information is available from other sources, that information can be utilized in the present invention. The processes are as follows: a) variability between patients in the response to a particular treatment is observed; b) at least a portion of the variable response is correlated with the presence or absence of at least one variance in at least one gene; c) an analytical or diagnostic test is provided to determine the presence or absence of the at least one variance in individual patients; d) the presence or absence of the variance or variances is used to select a patient for a treatment or to select a treatment for a patient, or the variance information is used in other methods described herein.

A. Identification of Intercpatient Variability in Response to a Treatment

[0354] Intercpatient variability is the rule, not the exception, in clinical therapeutics. One of the best sources of information on interpatient variability is the nurses and physicians supervising the clinical trial who accumulate a body of first hand observations of physiological responses to the drug in different normal subjects or patients. Evidence of interpatient variation in response can also be measured statistically, and may be best assessed by descriptive statistical measures that examine variation in response (beneficial or adverse) across a large number of subjects, including in different patient subgroups (men vs. women; whites vs. blacks; Northern Europeans vs. Southern Europeans, etc.). In accord with the other portions of this description, the present invention concerns DNA sequence variances that can affect one or more of:

i. The susceptibility of individuals to a disease;

ii. The course or natural history of a disease;

[0355] iii. The response of a patient with a disease to a medical intervention, such as, for example, a drug, a biologic substance, physical energy such as radiation therapy, or a specific dietary regimen. (The terms 'drug', 'compound' or 'treatment' as used herein may refer to any of the foregoing medical interventions.) The ability to predict either beneficial or detrimental responses is medically useful.

[0356] Thus variation in any of these three parameters may constitute the basis for initiating a pharmacogenetic study directed to the identification of the genetic sources of interpatient variation. The effect of a DNA sequence variance or variances on disease susceptibility or natural history (i and ii, above) are of particular interest as the variances can be used to define patient subgroups which behave differently in response to medical interventions such as those described in (iii). The methods of this invention are also useful in a clinical development program where there is not yet evidence of interpatient variation (perhaps because the compound is just entering clinical trials) but such variation in response can be reliably anticipated. It is more economical to design pharmacogenetic studies from the beginning of a clinical development program than to start at a later stage when the costs of any delay are likely to be high given the resources typically committed to such a program.

[0357] In other words, a variance can be useful for customizing medical therapy at least for either of two reasons. First, the variance may be associated with a specific disease subset that behaves differently with respect to one or more therapeutic interventions (i and ii above); second, the variance may affect response to a specific therapeutic intervention (iii above). Consider for exemplary purposes pharmacological therapeutic interventions. In the first case, there may be no effect of a particular gene sequence variance on the observable pharmacological action of a drug, yet the
disease subsets defined by the variance or variances differ in their response to the drug because, for example, the drug acts on a pathway that is more relevant to disease pathophysiology in one variance-defined patient subset than in another variance-defined patient subset. The second type of useful gene sequence variance affects the pharmacological action of a drug or other treatment. Effects on pharmacological responses fall generally into two categories: pharmacokinetic and pharmacodynamic effects. These effects have been defined as follows in Goodman and Gilman’s Pharmacologic Basis of Therapeutics (ninth edition, McGraw Hill, New York, 1986): “Pharmacokinetics” deals with the absorption, distribution, biotransformations and excretion of drugs. The study of the biochemical and physiological effects of drugs and their mechanisms of action is termed “pharmacodynamics.”

Useful gene sequence variances for this invention can be described as variances which partition patients into two or more groups that respond differently to a therapy or that correlate with differences in disease susceptibility or progression, regardless of the reason for the difference, and regardless of whether the reason for the difference is known. The latter is true because it is possible, with genetic methods, to establish reliable associations even in the absence of a pathophysiological hypothesis linking a gene to a phenotype, such as a pharmacological response, disease susceptibility or disease prognosis.

B. Identification of Specific Genes and Correlation of Variances in Those Genes with Response to Treatment of Diseases or Conditions

It is useful to identify particular genes which do or are likely to mediate the efficacy or safety of a treatment method for a disease or condition, particularly in view of the large number of genes which have been identified and which continue to be identified in humans. As is further discussed in section C below, this correlation can proceed by different paths. One exemplary method utilizes prior information on the pharmacology or pharmacokinetics or pharmacodynamics of a treatment method, e.g., the action of a drug, which indicates that a particular gene is, or is likely to be, involved in the action of the treatment method, and further suggests that variances in the gene may contribute to variable response to the treatment method. For example, if a compound is known to be glucuronidated then a glucuronyltransferase is likely involved. If the compound is a phenol, the likely glucuronyltransferase is UGT1 (either the UGT1*1 or UGT1*6 transcripts, both of which catalyze the conjugation of planar phenols with glucuronic acid). Similar inferences can be made for many other biotransformation reactions.

Alternatively, if such information is not known, variances in a gene can be correlated empirically with treatment response. In this method, variances in a gene which exist in a population can be identified. The presence of the different variances or haplotypes in individuals of a study group, which is preferably representative of a population or populations of known geographic, ethnic and/or racial background, is determined. This variance information is then correlated with treatment response of the various individuals as an indication that genetic variability in the gene is at least partially responsible for differential treatment response. It may be useful to independently analyze variances in the different geographic, ethnic and/or racial groups as the presence of different genetic variances in these groups (i.e., different genetic background) may influence the effect of a specific variance. That is, there may be a gene-gene interaction involving one unstudied gene, however the indicated demographic variables may act as a surrogate for the unstudied allele. Statistical measures known to those skilled in the art are preferably used to measure the fraction of interpatient variation attributable to any one variance, or to measure the response rate in different subgroups defined genetically or defined by some combination of genetic, demographic and clinical criteria.

C. Development of a Diagnostic Test to Determine Variance Status

In accordance with the description in the Summary above, the present invention generally concerns the identification of variances in genes which are indicative of the effectiveness of a treatment in a patient. The identification of specific variances, in effect, can be used as a diagnostic or prognostic test. Correlation of treatment efficacy and/or toxicity with particular genes and gene families or pathways is provided in Stanton et al., U.S. Provisional Application 60/093,484, filed Jul. 20, 1998, entitled GENE SEQUENCE VARIANCES WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE (concerns the safety and efficacy of compounds active on folate or pyrimidine metabolism or action) and Stanton, U.S. Provisional Application No. 60/121,047, filed Feb. 22, 1999, entitled GENE SEQUENCE VARIANCES WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE (concerning Alzheimer’s disease and other dementias and cognitive disorders), which are hereby incorporated by reference in their entireties including drawings.

Genes identified in the examples below and in the Tables and Figures can be used in the methods of the present invention. A variety of genes which the inventors realize may account for interpatient variation in response to treatments for inflammatory or immune diseases, conditions, disorders, and/or the development of same are listed in Tables 1 and 3. Gene sequence variances in said genes are particularly useful for aspects of the present invention.

Methods for diagnostic tests are well known in the art. Generally in this invention, the diagnostic test involves determining whether an individual has a variance or variant form of a gene that is involved in the disease or condition or the action of the drug or other treatment or effects of such treatment. Such a variance or variant form of the gene is preferably one of several different variances or forms of the gene that have been identified within the population and are known to be present at a certain frequency. In an exemplary
method, the diagnostic test involves determining the sequence of at least one variance in at least one gene after amplifying a segment of said gene using a DNA amplification method such as the polymerase chain reaction (PCR). In this method, DNA for analysis is obtained by amplifying a segment of DNA or RNA (generally after converting the RNA to cDNA) spanning one or more variances in the gene sequence. Preferably, the amplified segment is >500 bases in length, in an alternative embodiment the amplified segment is >100 bases in length, most preferably >65 bases in length.

In some cases it will be desirable to determine a haplotype instead of a genotype. In such a case the diagnostic test is performed by amplifying a segment of DNA or RNA (cDNA) spanning more than one variance in the gene sequence and preferably maintaining the phase of the variance on each allele. The term “phase” refers to the relationship of variances on a single chromosomeal copy of the gene, such as the copy transmitted from the mother (maternal copy or maternal allele) or the father (paternal copy or paternal allele). The haplotyping test may take part in two phases, where first genotyping tests at two or more variant sites reveal which sites are heterozygous in each patient or normal subject. Subsequently the phase of the two or more variant sites can be determined. In performing a haplotyping test preferably the amplified segment is >500 bases in length, more preferably it is >1,000 bases in length, and most preferably it is >2,500 bases in length. One way of preserving phase is to amplify one strand in the PCR reaction. This can be done using one or a pair of oligonucleotide primers that terminate (i.e. have a 3' end that stops) opposite the variant site, such that one primer is perfectly complementary to one variant form and the other primer is perfectly complementary to the other variant form. Other than the difference in the 3' most nucleotide the two primers are identical (forming an allelic primer pair). Only one of the allelic primers is used in any PCR reaction, depending on which strand is being amplified. The primer for the opposite strand may also be an allelic primer, or it may prime from a non-polymeric region of the template. This method exploits the requirement of most polymerases for perfect complementarity at the 3' terminus of the primer in a primer-template complex. See, for example: Lo Y M, Patel P, Newton C R, Markham A F, Fleming K A and J S Wainscoat. (1991) Direct haplotype determination by double ARMS: specificity, sensitivity and genetic applications. Nucleic Acids Res July 11; 19(13):3561-7.

It is apparent that such diagnostic tests are performed after initial identification of variances within the gene, which allows selection of appropriate allele specific primers.

Diagnostic genetic tests useful for practicing this invention belong to two types: genotyping tests and haplotyping tests. A genotyping test simply provides the status of a variance or variances in a subject or patient. For example suppose nucleotide 150 of hypothetical gene X on an autosomal chromosome is an adenine (A) or a guanine (G) base. The possible genotypes in any individual are AA, AG or GG at nucleotide 150 of gene X.

In a haplotyping test there is at least one additional variance in gene X, say at nucleotide 810, which varies in the population as cytosine (C) or thymine (T). Thus a particular copy of gene X may have any of the following combinations of nucleotides at positions 150 and 810: 150A-810C, 150A-810T, 150G-810C or 150G-810T. Each of the four possibilities is a unique haplotype. If the two nucleotides interact in either RNA or protein, then knowing the haplotype can be important. The point of a haplotyping test is to determine the haplotypes present in a DNA or cDNA sample (e.g. from a patient). In the example provided there are only four possible haplotypes, but, depending on the number of variances in the gene and their distribution in human populations there may be three, four, five, six or more haplotypes at a given gene. The most useful haplotypes for this invention are those which occur commonly in the population being treated for a disease or condition. Preferably such haplotypes occur in at least 5% of the population, more preferably in at least 10%, still more preferably in at least 20% of the population and most preferably in at least 30% or more of the population. Conversely, when the goal of a pharmacogenetic program is to identify a relatively rare population that has an adverse reaction to a treatment, the most useful haplotypes may be rare haplotypes, which may occur in less than 5%, less than 2%, or even in less than 1% of the population. One skilled in the art will recognize that the frequency of the adverse reaction provides a useful guide to the likely frequency of salient causative haplotypes.

Based on the identification of variances or variant forms of a gene, a diagnostic test utilizing methods known in the art can be used to determine whether a particular form of the gene, containing specific variances or haplotypes, or combinations of variances and haplotypes, is present in at least one copy, one copy, or more than one copy in an individual. Such tests are commonly performed using DNA or RNA collected from blood, cells, tissue scrapings or other cellular materials, and can be performed by a variety of methods including, but not limited to, PCR based methods, hybridization with allele-specific probes, enzymatic mutation detection, chemical cleavage of mismatches, mass spectrometry or DNA sequencing, including minisequencing. Methods for haplotyping are described above. In particular embodiments, hybridization with allele specific probes can be conducted in two formats: (1) allele specific oligonucleotides bound to a solid phase (glass, silicon, nylon membranes) and the labelled sample in solution, as in many DNA chip applications, or (2) bound sample (often cloned DNA or PCR amplified DNA) and labelled oligonucleotides in solution (either allele specific or short — e.g. 7mers or 8mers — so as to allow sequencing by hybridization). Preferred methods for diagnosing testing of variances are described in four patent applications Stanton et al, entitled A METHOD FOR ANALYZING POLYNUCLEOTIDES, Ser. Nos. 09/394, 467; 09/394,457; 09/394,774; and 09/394,387; all filed Sep. 10, 1999. The application of such diagnostic tests is possible after identification of variances that occur in the population. Diagnostic tests may involve a panel of variances from one or more genes, often on a solid support, which enables the simultaneous determination of more than one variance in one or more genes.

D. Use of Variance Status to Determine Treatment

The present disclosure describes exemplary gene sequence variances in genes identified in a gene table herein (e.g., Table 3), and variant forms of these gene that may be determined using diagnostic tests. As indicated in the Summary, such a variance-based diagnostic test can be used to determine whether or not to administer a specific drug or
In such cases, the procedure for using the drug is restricted or limited on the basis of a diagnostic test for determining the presence of a variance or variant form of a gene. Alternatively, the use of a genetic test may be advised as best medical practice, but not absolutely required, or it may be required in a subset of patients, e.g., those using one or more other drugs, or those with impaired liver or kidney function. The procedure that is dictated or recommended based on genotype may include the route of administration of the drug, the dosage form, dosage, schedule of administration or use with other drugs; any or all of these may require selecting or determination consistent with the results of the diagnostic test or a plurality of such tests. Preferably the use of such diagnostic tests to determine the procedure for administration of a drug is incorporated in a text such as those listed above, or medical textbooks, for example, textbooks of medicine, laboratory medicine, therapeutics, pharmacy, pharmacology, nutrition, allopathic, homeopathic, and osteopathic medicine. As previously stated, preferably such a diagnostic test or tests are required by regulatory authorities and are incorporated in the label or insert as well as in the Physicians Desk Reference.

Variances and variant forms of genes useful in conjunction with treatment methods may be associated with the origin or the pathogenesis of a disease condition. In many useful cases, the variant form of the gene is associated with a specific characteristic of the disease condition that is the target of a treatment, most preferably response to specific drugs or other treatments. Examples of diseases or conditions ameliorable by the methods of this invention are identified in the Examples and tables below; in general treatment of diseases with current methods, particularly drug treatment, always involves some unknown element (invoking efficacy or toxicity or both) that can be reduced by appropriate diagnostic methods.

Alternatively, the gene is involved in drug action, and the variant forms of the gene are associated with variability in the action of the drug. For example, in some cases, one variant form of the gene is associated with the action of the drug such that the drug will be effective in an individual who inherits one or two copies of that form of the gene. Alternatively, a variant form of the gene is associated with the action of the drug such that the drug will be toxic or otherwise contra-indicated in an individual who inherits one or two copies of that form of the gene.

In accord with this invention, diagnostic tests for variances and variant forms of genes as described above can be used in clinical trials to demonstrate the safety and efficacy of a drug in a specific population. As a result, in the case of drugs which show variability in patient response correlated with the presence or absence of a variance or variants, it is preferable that such drug is approved for sale or use by regulatory agencies with the recommendation or requirement that a diagnostic test be performed for a specific variance or variant form of a gene which identifies specific populations in which the drug will be safe and/or effective. For example, the drug may be approved for sale or use by regulatory agencies with the specification that a diagnostic test be performed for a specific variance or variant form of a gene which identifies specific populations in which the drug will be toxic. Thus, approved use of the drug, or the procedure for use of the drug, can be limited by a diagnostic test for such variances or variant forms of a gene; or such a diagnostic test may be considered good medical practice, but not absolutely required for use of the drug.

As indicated, diagnostic tests for variances as described in this invention may be used in clinical trials to establish the safety and efficacy of a drug. Methods for such clinical trials are described below and/or are known in the art and are described in standard textbooks. For example, diagnostic tests for a specific variance or variant form of a gene may be incorporated in the clinical trial protocol as inclusion or exclusion criteria for enrollment in the trial, to allocate certain patients to treatment or control groups within the clinical trial or to assign patients to different treatment cohorts. Alternatively, diagnostic tests for specific variances may be performed on all patients within a clinical trial, and statistical analysis performed comparing and contrasting the efficacy or safety of a drug between individuals with different variances or variant forms of the gene or genes. Preferred embodiments involving clinical trials include the genetic stratification strategies, phases, statistical analyses, sizes, and other parameters as described herein.

Similarly, diagnostic tests for variances can be performed on groups of patients known to have efficacious responses to the drug to identify differences in the frequency of variances between responders and non-responders. Likewise, in other cases, diagnostic tests for variance are performed on groups of patients known to have toxic responses to the drug to identify differences in the frequency of the variance between those having adverse events and those not having adverse events. Such outlier analyses may be particularly useful if a limited number of patient samples are available for analysis. It is apparent that such clinical trials can be or are performed after identifying specific variances or variant forms of the gene in the population. In defining outliers it is useful to examine the distribution of responses in the placebo group; outliers should preferably have responses that exceed in magnitude the extreme responses in the placebo group.

The identification and confirmation of genetic variances is described in certain patents and patent applications. The description therein is useful in the identification of variances in the present invention. For example, a strategy for the development of anticancer agents having a high therapeutic index is described in Housman, International Application PCT/US/94 08473 and Housman, INHIBITORS OF ALTERNATIVE ALLELES OF GENES.
ENCODING PROTEINS VITAL FOR CELL VIABILITY OR CELL GROWTH AS A BASIS FOR CANCER THERAPEUTIC AGENTS, U.S. Pat. No. 5,702,890, issued Dec. 30, 1997, which are hereby incorporated by reference in their entireties. Also, a number of gene targets and associated variances are identified in Housman et al., U.S. patent application Ser. No. 09/045,053, entitled TARGET ALLELES FOR ALLELE-SPECIFIC DRUGS, filed Mar. 19, 1998, which is hereby incorporated by reference in its entirety, including drawings.

[0378] The described approach and techniques are applicable to a variety of other diseases, conditions, and/or treatments and to genes associated with the etiology and pathogenesis of such other diseases and conditions and the efficacy and safety of such other treatments.

[0379] Useful variances for this invention can be described generally as variances which partition patients into two or more groups that respond differently to a therapy (a therapeutic intervention), regardless of the reason for the difference, and regardless of whether the reason for the difference is known.

III. From Variance List to Clinical Trial: Identifying Genes and Gene Variances that Account for Variable Responses to Treatment

[0380] There are a variety of useful methods for identifying a subset of genes from a large set of candidate genes that should be prioritized for further investigation with respect to their influence on inter-individual variation in disease predisposition or response to a particular drug. These methods include for example, (1) searching the biomedical literature to identify genes relevant to a disease or the action of a drug, (2) screening the genes identified in step 1 for variances. A large set of exemplary variances are provided in Table 3. Other methods include (3) using computational tools to predict the functional effects of variances in specific genes, (4) using in vitro or in vivo experiments to identify genes which may participate in the response to a drug or treatment, and to determine the variances which affect gene, RNA or protein function, and may therefore be important genetic variables affecting disease manifestations or drug response, and (5) retrospective or prospective clinical trials. Computational tools are described in U.S. Patent Application, Stanton et al., Ser. No. 09/330,747, filed Apr. 26, 1999, entitled GENE SEQUENCE VARIANCES WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE, and in Stanton et al., Ser. No. 09/419,705, filed Oct. 14, 1999, entitled VARIANCE SCANNING METHOD FOR IDENTIFYING GENE SEQUENCE VARIANCES, which are hereby incorporated by reference in their entireties, including drawings. Other methods are considered below in some detail.

[0381] (1) To begin, one preferably identifies, for a given treatment, a set of candidate genes that are likely to affect disease phenotype or drug response. This can be accomplished most efficiently by first assembling the relevant medical, pharmacological and biological data from available sources (e.g., public databases and publications). One skilled in the art can review the literature (textbooks, monographs, journal articles) and online sources (databases) to identify genes most relevant to the action of a specific drug or other treatment, particularly with respect to its utility for treating a specific disease, as this beneficially allows the set of genes to be analyzed ultimately in clinical trials to be reduced from an initial large set. Specific strategies for conducting such searches are described below. In some instances the literature may provide adequate information to select genes to be studied in a clinical trial, but in other cases additional experimental investigations of the sort described below will be preferable to maximize the likelihood that the salient genes and variances are moved forward into clinical studies. Specific genes relevant to understanding interpatient variation in response to treatments for major inflammatory and immunologic diseases are listed in Table 1. In Table 2 preferred sets of genes for analysis of variable therapeutic response in specific diseases are highlighted. These genes are exemplary; they do not constitute a complete set of genes that may account for variation in clinical response. Experimental data are also useful in establishing a list of candidate genes, as described below.

[0382] (2) Having assembled a list of candidate genes generally the second step is to screen for variances in each candidate gene. Experimental and computational methods for variance detection are described in this invention, and tables of exemplary variances are provided: (Table 3) as well as methods for identifying additional variances and a written description of such possible additional variances in the cDNAs of genes that may affect drug action (see Stanton & Adams, application Ser. No. 09/300,747, filed Apr. 26, 1999, entitled GENE SEQUENCE VARIANCES WITH UTILITY IN DETERMINING THE TREATMENT OF DISEASE, incorporated in its entirety.

[0383] (3) Having identified variances in candidate genes the next step is to assess their likely contribution to clinical variation in patient response to therapy, preferably by using informatics-based approaches such as DNA and protein sequence analysis and protein modeling. The literature and informatics-based approaches provide the basis for prioritization of candidate genes, however it may in some cases be desirable to further narrow the list of candidate genes, or to measure experimentally the phenotype associated with specific variances or sets of variances (e.g. haplotypes).

[0384] (4) Thus, as a third step in candidate gene analysis, one skilled in the art may elect to perform in vitro or in vivo experiments to assess the functional importance of gene variances, using either biochemical or genetic tests. (Certain kinds of experiments—for example gene expression profiling and proteome analysis—may not only allow refinement of a candidate gene list but may also lead to identification of additional candidate genes.) Combination of two or all of the three above methods will provide sufficient information to narrow and prioritize the set of candidate genes and variances to a number that can be studied in a clinical trial with adequate statistical power.

[0385] (5) The fourth step is to design retrospective or prospective human clinical trials to test whether the identified allelic variance, variances, or haplotypes or combination thereof influence the efficacy or toxicity profiles for a given drug or other therapeutic intervention. It should be recognized that this fourth step is the crucial step in producing the type of data that would justify introducing a diagnostic test for at least one variance into clinical use. Thus while each of the above four steps are...
useful in particular instances of the invention, this final step is indispensable. Further guidance and examples of how to perform these five steps are provided below.

[0386] (6) A fifth (optional) step entails methods for using a genotyping test in the promotion and marketing of a treatment method. It is widely appreciated that there is a tendency in the pharmaceutical industry to develop many compounds for well established therapeutic targets. Examples include beta adrenergic blockers, hydroxyethylglutaryl (HMG) CoA reductase inhibitors (statins), dopamine D2 receptor antagonists and serotonin transporter inhibitors. Frequently the pharmacology of these compounds is quite similar in terms of efficacy and side effects. Therefore the marketing of one compound vs. other members of the class is a challenging problem for drug companies, and is reflected in the lesser success that late products typically achieve compared to the first and second approved products. It occurred to the inventors that genetic stratification can provide the basis for identifying a patient population with a superior response rate or improved safety to one member of a class of drugs, and that this information can be the basis for commercialization of that compound. Such a commercialization campaign can be directed at caregivers, particularly physicians, or at patients and their families, or both.

1. Identification of Candidate Genes Relevant to the Action of a Drug

[0387] Practice of this invention will often begin with identification of a specific pharmaceutical product, for example a drug, that would benefit from improved efficacy or reduced toxicity or both, and the recognition that pharmacogenetic investigations as described herein provide a basis for achieving such improved characteristics. The question then becomes which genes and variances, such as those provided in this application in Tables 1 and 3, would be most relevant to interpatient variation in response to the drug. As discussed above, the set of relevant genes includes both genes involved in the disease process and genes involved in the interaction of the patient and the treatment—for example genes involved in pharmacokinetic and pharmacodynamic action of a drug. The biological and biomedical literature and online databases provide useful guidance in selecting such genes. Specific guidance in the use of these resources is provided below.

Review the Literature and Online Sources

[0388] One way to find genes that affect response to a drug in a particular disease setting is to review the published literature and available online databases regarding the pathophysiology of the disease and the pharmacology of the drug. Literature or online sources can provide specific genes involved in the disease process or drug response, or describe biochemical pathways involving multiple genes, each of which may affect the disease process or drug response.

[0389] Alternatively, biochemical or pathological changes characteristic of the disease may be described; such information can be used by one skilled in the art to infer a set of genes that can account for the biochemical or pathologic changes. For example, to understand variation in response to a drug that modulates interleukin-6 (IL-6) levels in a immune disorder associated with altered levels of IL-6 one would preferably study, at a minimum, variances in genes responsible for IL-6 biosynthesis, release from the cell, receptor binding, presynaptic reuptake, and degradation or metabolism. Genes responsible for each of these functions should be examined for variation that may account for interpatient differences in drug response or disease manifestations. As recognized by those skilled in the art, a comprehensive list of such genes can be obtained from textbooks, monographs and the literature.

[0390] There are several types of scientific information, described in some detail below, that are valuable for identifying a set of candidate genes to be investigated with respect to a specific disease and therapeutic intervention. First there is the medical literature, which provides basic information on disease pathophysiology and therapeutic interventions. A subset of this literature is devoted to specific description of pathologic conditions. Second there is the pharmacology literature, which will provide additional information on the mechanism of action of a drug (pharmacodynamics) as well as its principal routes of metabolic transformation (pharmacokinetics) and the responsible proteins. Third there is the biomedical literature (principally genetics, physiology, biochemistry and molecular biology), which provides more detailed information on metabolic pathways, protein structure and function and gene structure. Fourth, there are a variety of online databases that provide additional information on metabolic pathways, gene families, protein function and other subjects relevant to selecting a set of genes that are likely to affect the response to a treatment.

Medical Literature


[0392] In addition to these general texts, there are a variety of more specialized medical texts that provide greater detail about specific disorders which can be utilized in developing a list of candidate genes and variances relevant to interpatient variation in response to a treatment. For example, within the field of medicine there are standard textbooks for each of the subspecialties. Some examples include:
Heart Disease: A Textbook of Cardiovascular Medicine (2 Volume set) by E. Braunwald (Editor), W B Saunders Co., 1996.


Wintrone’s Clinical Hematology (10th Ed) by G. R. Lee, J. Foerster (Editor) and J. Lukens (Editors) (2 Volumes) Lippincott, Williams & Wilkins, 1998.


The Management of Pain by J. J. Bonica, Lea and Febiger, 1992


In addition to these subspecialty texts there are many textbooks and monographs that concern more restricted disease areas, or specific diseases. Such books provide more extensive coverage of pathophysiological mechanisms and therapeutic options. The number of such books is too great to provide examples for all but a few diseases, however one skilled in the art will be able to readily identify relevant texts. One simple way to search for relevant titles is to use the search engine of an online bookseller such as http://www.amazon.com or http://www.barnesandnoble.com using the disease or drug (or the group of diseases or drugs to which they belong) as search terms. For example a search for asthma would turn up titles such as Asthma: Basic Mechanisms and Clinical Management (3rd edition) by P. J. Barnes, I. W. Rodger and N. C. Thomson (Editors), Academic Press, 1998 and Airways and Vascular Remodelling in Asthma and Cardiovascular Disease: Implications for Therapeutic Intervention, by C. Page & J. Black (Editors), Academic Press, 1994.

Pathology Literature

In addition to medical texts there are texts that specifically address disease etiology and pathologic changes associated with disease. A good general pathology text is Robbins Pathologic Basis of Disease (6th edition) by R. S. Cotran, V. Kumar, T. Collins and S. L. Robbins, W B Saunders Co., 1998. Specialized pathology texts exist for each organ system and for specific diseases, similar to medical texts. These texts are useful sources of information for one skilled in the art for developing lists of genes that may account for some of the known pathologic changes in disease tissue. Exemplary texts are as follows:

Bone Marrow Pathology. 2nd edition, by B. J. Buijn, I. Lampert. & D. Clark, Blackwell Science, 1996


Pulmonary Pathology by M. A. Grippi, Lippincott-Raven, 1995.


Pharmacology, Pharmacogenetics and Pharmacy Literature

There are also both general and specialized texts and monographs on pharmacology that provide data on pharmacokinetics and pharmacodynamics of drugs. The discussion of pharmacodynamics (mechanism of action of the drug) in such texts is often supported by a review of the biochemical pathway or pathways that are affected by the drug. Also, proteins related to the target protein are often listed; it is important to account for variation in such proteins as the related proteins may be involved in drug pharmacology. For example, there are 14 known serotonin receptors. Various pharmacological serotonin agonists or antagonists have different affinities for these different receptors. Variation in a specific receptor may affect the pharmacology not only of drugs targeted to that receptor, but also drugs that are principally agonists or antagonists of different receptors. Such compounds may produce different effects on two allelic forms of a non-targeted receptor, for example on variant form may bind the compound with higher affinity than the other, or a compound that is principally an antagonist for one allele may be a partial agonist for another allele. Thus genes encoding proteins structurally related to the target protein should be screened for variance in order to successfully realize the methods of the present invention. A good general pharmacology text is Goodman & Gilman’s the Pharmacological Basis of Therapeutics (9th Ed) by J. G. Hardman, L. E. Limbird, P. B. Molinoff, R. W. Ruddon and A. G. Gilman (Editors) McGraw Hill, 1996. There are also
texts that focus on the pharmacology of drugs for specific disease areas, or specific classes of drugs (e.g. natural products) or adverse drug interactions, among other subjects. Specific examples include:


[0398] There are also texts on pharmacogenetics which are particularly useful for identifying genes which may contribute to variable pharmacokinetic response. In addition there are texts on some of the major xenobiotic metabolizing proteins, such as the cytochrome P450 genes.


Cytochrome P450: Structure, Mechanism, and Biochemistry by P. R. Ortiz de Montellano (Editor), Plenum Publishing Corp, 1995.


Genetics, Biochemistry and Molecular Biology Literature

[0399] In addition to the medical, pathology, and pharmacology texts listed above there are several information sources that one skilled in the art will turn to for information on the genetic, physiologic, biochemical, and molecular biological aspects of the disease, disorder or condition or the effect of the therapeutic intervention on specific physiologic processes. The biomedical literature may include information on nonhuman organisms that is relevant to understanding the likely disease or pharmacological pathways in man.

[0400] Also provided below are illustrative texts which will aid in the identification of a pathway or pathways, and a gene or genes that may be relevant to interindividual variation in response to a therapy. Textbooks of biochemistry, genetics and physiology are often useful sources for such pathway information. In order to ascertain the appropriate methods to analyze the effects of an allelic variance, variances, or haplotypes in vitro, one skilled in the art will review existing information on molecular biology, cell biology, genetics, biochemistry; and physiology. Such texts are useful sources for general and specific information on the genetic and biochemical processes involved in disease and in drug action, as well as experimental procedures that may be useful in performing in vitro research on an allelic variance, variances, or haplotype.

[0401] Texts on gene structure and function and RNA biochemistry will be useful in evaluating the consequences of variances that do not change the coding sequence (silent variances). Such variances may alter the interaction of RNA with proteins or other regulatory molecules affecting RNA processing, polyadenylation, or export.

Molecular and Cellular Biology


Essentials of Molecular Biology, D. Freifelder and Malachinski Jones and Bartlett, 1993.


Molecular Genetics


General Biochemistry


Biochemistry, G. Zubay, 1998. Wm. C. Brown Communications


Transcription


RNA


mRNA Metabolism and Post-transcriptional Gene Regulation, Harford and Morris (eds.), 1997. Wiley-Liss
Translation


General Physiology


Online Databases

[0402] Those skilled in the art are familiar with how to search the biomedical literature, such as, e.g., libraries, online PubMed, abstract listings, and online mutation databases. One particularly useful resource is maintained at the web site of the National Center for Biotechnology Information (ncbi): http://www.ncbi.nlm.nih.gov/. From the ncbi site one can access Online Mendelian Inheritance in Man (OMIM). OMIM can be found at: http://www3.ncbi.nlm.nih.gov/Omim/searchomim.html. OMIM is a medically oriented database of genetic information with entries for thousands of genes. The OMIM record number is provided for many of the genes in tables 1 and 3 (see column 3), and constitutes an excellent entry point for identification of references that point to the broader literature. Another useful site at NCBI is the Entrez browser, located at http://www3.ncbi.nlm.nih.gov/Entrez/. One can search genomes, nucleotide sequences, proteins, 3D structures, taxonomy or the biomedical literature (PubMed) via the Entrez site. More generally links to a number of useful sites with biomedical or genetic data are maintained at sites such as MedWeb at the Emory University Health Sciences Center Library: http://WWW.MedWeb.Emory.Edu/MedWeb/; Riken, a Japanese web site at: http://www.rtc.riken.go.jp/; othersite.html with links to DNA sequence, structural, molecular biology, bioinformatics, and other databases; at the Oak Ridge National Laboratory web site: http://www.ornl.gov/bgmis/links.html; at the Yahoo website of Diseases and Conditions: http://dir.yahoo.com/health/diseases and conditions/index.html. Each of the indicated web sites has additional useful links to other sites.

[0403] Another type of database with utility in selecting the genes on a biochemical pathway that may affect the response to a drug is databases that provide information on biochemical pathways. Examples of such databases include the Kyoto Encyclopaedia of Genes and Genomes (KEGG), which can be found at: http://www.genome.ad.jp/kegg/kegg.html. This site has pictures of many biochemical pathways, as well as links to other metabolic databases such as the well known Boehringer Mannheim biochemical pathways charts: http://www.expasy.ch/cgi-bin/search-biochem-index. The metabolic charts at the latter site are comprehensive, and excellent starting points for working out the salient enzymes on any given pathway.

Each of the web sites mentioned above has links to other useful web sites, which in turn can lead to additional sites with useful information. Research Libraries

[0404] Those skilled in the art will often require information found only at large libraries. The National Library of Medicine (http://www.nlm.nih.gov/) is the largest medical library in the world and its catalogs can be searched online. Other libraries, such as university or medical school libraries are also useful to conduct searches. Biomedical books such as those referred to above can often be obtained from online bookstores as described above.

Biomedical Literature

[0405] To obtain up to date information on drugs and their mechanism of action and biotransformation; disease pathophysiology; biochemical pathways relevant to drug action and disease pathophysiology; and genes that encode proteins relevant to drug action and disease one skilled in the art will consult the biomedical literature. A widely used, publically accessible web site for searching published journal articles is PubMed (http://www.ncbi.nlm.nih.gov/PubMed/). At this site, one can search for the most recent articles (within the last 1-2 months) or older literature (back to 1966). Many Journals also have their own sites on the world wide web and can be searched online. For example see the IDEAL web site at: http://www.apnet.com/www/gp/aboutid.html. This site is an online library, featuring full text journals from Academic Press and selected journals from W. B. Saunders and Churchill Livingstone. The site provides access (for a fee) to nearly 2000 scientific, technical, and medical journals.

Experimental Methods for Identification of Genes Involved in the Action of a Drug

[0406] There are a number of experimental methods for identifying genes and gene products that mediate or modulate the effects of a drug or other treatment. They encompass analyses of RNA and protein expression as well as methods for detecting protein—protein interactions and protein—ligand interactions. Two preferred experimental methods for identification of genes that may be involved in the action of a drug are (1) methods for measuring the expression levels of many mRNA transcripts in cells or organisms treated with the drug (2) methods for measuring the expression levels of many proteins in cells or organisms treated with the drug.

[0407] RNA transcripts or proteins that are substantially increased or decreased in drug treated cells or tissues relative to control cells or tissues are candidates for mediating the action of the drug. Preferably the level of an mRNA is at least 30% higher or lower in drug treated cells, more preferably at least 50% higher or lower, and most preferably two fold higher or lower than levels in non-drug treated control cells. The analysis of RNA levels can be performed on total RNA or on polyadenylated RNA selected by oligodeoxyT affinity. Further, RNA from different cell compartments can be analyzed independently—for example nuclear vs. cytoplasmic RNA. In addition to RNA levels, RNA kinetics can be examined, or the pool of RNAs currently being translated can be analyzed by isolation of RNA from polysomes. Other useful experimental methods include protein interaction methods such as the yeast two hybrid system and variants thereof which facilitate the detection of protein—protein interactions. Preferably one of the interacting proteins is the drug target or another protein strongly implicated in the action of the compound being assessed.

[0408] The pool of RNAs expressed in a cell is sometimes referred to as the transcriptome. Methods for measuring the transcriptome, or some part of it, are known in the art. A recent collection of articles summarizing some current methods appeared as a supplement to the journal Nature Genetics. (The Chipping Forecast. Nature Genetics supplement, vol-
A preferred method for measuring expression levels of mRNAs is to spot PCR products corresponding to a large number of specific genes on a nylon membrane such as Hybond N Plus (Amersham-Pharmacia). Total cellular mRNA is then isolated, labelled by random oligonucleotide priming in the presence of a detectable label (e.g., alpha 33P labelled radiolabeled nucleotide) and hybridized with the filter containing the PCR products. The resulting signals can be analyzed by commercially available software, such as can be obtained from Clontech/Molecular Dynamics or Research Genetics, Inc.

Experiments have been described in model systems that demonstrate the utility of measuring changes in the transcriptome before and after changing the growth conditions of cells, for example by changing the nutrient environment. The changes in gene expression help reveal the network of genes that mediate physiological responses to the altered growth condition. Similarly, the addition of a drug to the cellular or in vivo environment, followed by monitoring these changes in gene expression can aid in identification of gene networks that mediate pharmacological responses.

The pool of proteins expressed in a cell is sometimes referred to as the proteome. Studies of the proteome may include not only protein abundance but also protein subcellular localization and protein-protein interaction. Methods for measuring the proteome, or some part of it, are known in the art. One widely used method is to extract total cellular protein and separate it in two dimensions, for example first by size and then by isoelectric point. The resulting protein spots can be stained and quantitated, and individual spots can be excised and analyzed by mass spectrometry to provide definitive identification. The results can be compared from two or more cell lines or tissues, at least one of which has been treated with a drug. The differential up or down modulation of specific proteins in response to drug treatment may indicate their role in mediating the pharmacologic actions of the drug. Another way to identify the network of proteins that mediate the actions of a drug is to exploit methods for identifying interacting proteins. By starting with a protein known to be involved in the action of a drug—for example the drug target—one can use systems such as the yeast two hybrid system and variants thereof (known to those skilled in the art; see Ausubel et al., Current Protocols in Molecular Biology, op. cit.) to identify additional proteins in the network of proteins that mediate drug action. The genes encoding such proteins would be useful for screening for DNA sequence variations, which in turn may be useful for analysis of interpatient variation in response to treatments. For example, the protein 5-lipoxygenase (5LO) is an enzyme which is at the beginning of the leukotriene biosynthetic pathway and is a target for anti-inflammatory drugs used to treat asthma and other diseases. In order to detect proteins that interact with 5-lipoxygenase the two-hybrid system was recently used to isolate three different proteins, none previously known to interact with 5LO. (Provost et al., Interaction of 5-lipoxygenase with cellular proteins. Proc. Natl. Acad. Sci. U.S.A. 96: 1881-1885, 1999.) A recent collection of articles summarizing some current methods in proteomics appeared in the August 1998 issue of the journal Electrophoresis (volume 19, number 11). Other useful articles include: Blackstock WP, et al. Proteomics: quantitative and physical mapping of cellular proteins. Trends Biotechnol. 17 (3): p. 121-7, 1999, and Patton W. F., Proteome analysis II. Protein subcellular redistribution: linking physiology to genomics via the proteome and separation technologies involved. J Chromatogr B Biomed Sci App. 722(1-2):203-23. 1999.

Since many of these methods can also be used to assess whether specific polymorphisms are likely to have biological effects, they are also relevant in section 3, below, concerning methods for assessing the likely contribution of variances in candidate genes to clinical variation in patient responses to therapy.

2. Screen for Variances in Genes That May Be Related to Therapeutic Response

Having identified a set of genes that may affect response to a drug the next step is to screen the genes for variances that may account for interindividual variation in response to the drug. There are a variety of levels at which a gene can be screened for variances, and a variety of methods for variance screening. The two main levels of variance screening are genomic DNA screening and cDNA screening. Genomic variance detection may include screening the entire genomic segment spanning the gene from 2 kb to 10 kb upstream of the transcription start site to the polyadenylation site, or 2 to 10 kb beyond the polyadenylation site. Alternatively genetic variance detection may (for intron containing genes) include the exons and some region around them containing the splicing signals, for example, but not all of the intronic sequences. In addition to screening introns and exons for variances it is generally desirable to screen regulatory DNA sequences for variances. Promoter, enhancer, silencer and other regulatory elements have been described in human genes. The promoter is generally proximal to the transcription start site, although there may be several promoters and several transcription start sites. Enhancer, silencer and other regulatory elements may be intragenic or may lie outside the introns and exons, possibly at a considerable distance, such as 100 kb away. Variances in such sequences may affect basal gene expression or regulation of gene expression. In either case such variation may affect the response of an individual patient to a therapeutic intervention, for example a drug, as described in the examples. Thus in practicing the present invention it is useful to screen regulatory sequences as well as transcribed sequences, in order to identify variances that may affect gene transcription. Frequently the genomic sequence of a gene can be found in the sources above, particularly by searching GenBank or Medline (PubMed). The name of the gene can be entered at a site such as Entrez: http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html. Using the genomic sequence and information from the biomedical literature one skilled in the art can perform a variance detection procedure such as those described in examples 15, 16 and 17.

Variance detection is often first performed on the cDNA of a gene for several reasons. First, available data on functional sequence variances suggests that variances in the transcribed portion of a gene may be most likely to have functional consequences as they can affect the interaction of the transcript with a wide variety of cellular factors during the complex processes of RNA transcription, processing and translation, with consequent effects on RNA splicing, stability, translational efficiency or other processes. Second, as a practical matter the cDNA sequence of a gene is often
available before the genomic structure is known, although the reverse will be true in the future as the sequence of the human genome is determined. Third, the cDNA is often compact compared to the genomic locus, and can be screened for variances with much less effort. If the genomic structure is not known then only the cDNA sequence can be scanned for variances. Methods for preparing cDNA are described in Example 14. Methods for variance detection on cDNA are described below and in the examples.

[0414] In general it is preferable to catalog genetic variation at the genomic DNA level because there are an increasing number of well documented instances of functionally important variances that lie outside of transcribed sequence. Also, to properly use optimal genetic methods to assess the contribution of a candidate gene to variation in a phenotype of interest it is desirable to understand the character of sequence variation in the candidate gene: what is the nature of linkage disequilibrium between different variances in the gene; are there sites of recombination within the gene; what is the extent of homoplasy in the gene (i.e. occurrence of two variant sites that are identical by state but not identical by descent because the same variance arose at least twice in human evolutionary history on two different haplotypes); what are the different haplotypes and how can they be grouped to increase the power of genetic analysis?

[0415] Methods for variance screening have been described, including DNA sequencing. See for example: U.S. Pat. No. 5,698,400; Detection of mutation by resolavase cleavage; U.S. Pat. No. 5,217,863: Detection of mutations in nucleic acids; and U.S. Pat. No. 5,750,335: Screening for genetic variation, as well as the examples and references cited therein for examples of useful variance detection procedures. Detailed variance detection procedures are also described in examples 15, 16 and 17. One skilled in the art will recognize that depending on the specific aims of a variance detection project (number of genes being screened, number of individuals being screened, total length of DNA being screened) one of the above cited methods may be preferable to the others, or yet another procedure may be optimal. A preferred method of variance detection is chain terminating DNA sequencing using dye labeled primers, cycle sequencing and software for assessing the quality of the DNA sequence as well as specialized software for calling heterozygotes. The use of such procedures has been described by Nickerson and colleagues. See for example: Rieder M. J., et al. Automating the identification of DNA variations using quality-based fluorescence resequencing: analysis of the human mitochondrial genome. Nucleic Acids Res. 26 (4):967-73, 1998, and: Nickerson D. A., et al. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. 25 (14):2745-51, 1997. Although the variances provided in Table 3 comprise principally of CDNA variances, it is an aspect of this invention that detection of genomic variances is also a useful method for identification of variances that may account for interpatient variation in response to a therapy.

[0416] Another important aspect of variance detection is the use of DNA from a panel of human subjects that represents a known population. For example, if the subjects are being screened for variances relevant to a specific drug development program it is desirable to include both subjects with the target disease and healthy subjects in the panel, because certain variances may occur at different frequencies in the healthy and disease populations and can only be reliably detected by screening both populations. Also, for example, if the drug development program is taking place in Japan, it is important to include Japanese individuals in the screening population. In general, it is always desirable to include subjects of known geographic, racial or ethnic identity in a variance screening experiment so the results can be interpreted appropriately for different patient populations, if necessary. Also, in order to select optimal sets of variances for genetic analysis of a gene locus it is desirable to know which variances have occurred recently—perhaps on multiple different chromosomes—and which are ancient. Inclusion of one or more apes or monkeys in the variance screening panel is one way of gaining insight into the evolutionary history of variances. Chimpanzees are preferred subjects for inclusion in a variance screening panel.

3. Assess the Likely Contribution of Variances in Candidate Genes to Clinical Variation in Patient Responses to Therapy

[0417] Once a set of genes likely to affect disease pathophysiology or drug action has been identified, and those genes have been screened for variances, said variances (e.g., provided in Table 3) can be assessed for their contribution to variation in the pharmacological or toxicological phenotypes of interest. Such studies are useful for reducing a large number of candidate variances to a smaller number of variances to be tested in clinical trials. There are several methods which can be used in the present invention for assessing the medical and pharmaceutical implications of a DNA sequence variance. They range from computational methods to in vitro and/or in vivo experimental methods, to prospective human clinical trials, and also include a variety of other laboratory and clinical measures that can provide evidence of the medical consequences of a variance. In general, human clinical trials constitute the highest standard of proof that a variance or set of variances is useful for selecting a method of treatment, however, computational and in vitro data, or retrospective analysis of human clinical data may provide strong evidence that a particular variance will affect response to a given therapy, often at lower cost and in less time than a prospective clinical trial. Moreover, at an early stage in the analysis when there are many possible hypotheses to explain interpatient variation in treatment response, the use of informatics-based approaches to evaluate the likely functional effects of specific variances is an efficient way to proceed.

[0418] Informatics-based approaches to the prediction of the likely functional effects of variances include DNA and protein sequence analysis (phylogenetic approaches and motif searching) and protein modeling (based on coordinates in the protein database, or pdb; see http://www.rcsb.org/pdb/). See, for example: Kawabata et al. The Protein Mutant Database. Nucleic Acids Research 27: 355-357, 1999; also available at: http://pdb.ddbj.nig.ac.jp. Such analyses can be performed quickly and inexpensively, and the results may allow selection of certain genes for more extensive in vitro or in vivo studies or for more variance detection or both.

[0419] The three dimensional structure of many medically and pharmaceutically important proteins, or homologs of such proteins in other species, or examples of domains present in such proteins, is known as a result of x-ray crystallography studies and, increasingly, nuclear magnetic
resonance studies. Further, there are increasingly powerful tools for modeling the structure of proteins with unsolved structure, particularly if there is a related (homologous) protein with known structure. (For reviews see: Rost et al., Protein fold recognition by prediction-based threading, J. Mol. Biol. 270:471-480, 1997; Firestone et al., Threading your way to protein function, Chem. Biol. 3:779-783, 1996)

There are also powerful methods for identifying conserved domains and vital amino acid residues of proteins of unknown structure by analysis of phylogenetic relationships. (Delenge et al., Protein structure prediction: Implications for the biologist, Biochimie 79:681-686, 1997; Taylor et al., Multiple protein structure alignment, Protein Sci. 3:1858-1870, 1994) These methods can permit the prediction of functionally important variations, either on the basis of structure or evolutionary conservation. For example, a crystal structure can reveal which amino acids comprise a small molecule binding site. The identification of a polymorphic amino acid residue in the topological neighborhood of such a site, and, in particular, the demonstration that at least one variant form of the protein has a variant amino acid which impinges on (or which may otherwise affect) the chemical environment around the small molecule binding pocket differently from another variant form, provides strong evidence that the variance may affect the function of the protein. From this it follows that the interaction of the protein with a treatment method, such an administered compound, will likely be variable between different patients. One skilled in the art will recognize that the application of computational tools to the identification of functionally consequential variances involves applying the knowledge and tools of medicinal chemistry and physiology to the analysis.

Phylogenetic approaches to understanding sequence variation are also useful. Thus if a sequence variance occurs at a nucleotide or encoded amino acid residue where there is usually little or no variation in homologs of the protein of interest from non-human species, particularly evolutionarily remote species, then the variance is more likely to affect function of the RNA or protein. Computational methods for phylogenetic analysis are known in the art, (see below for citations of some methods).

Computational methods are also useful for analyzing DNA polymorphisms in transcriptional regulatory sequences, including promoters and enhancers. One useful approach is to compare variances in potential or proven transcriptional regulatory sequences to a catalog of all known transcriptional regulatory sequences, including conserved binding domains for all transcription factor binding domains. See, for example, the databases cited in: Burks, C. Molecular Biology Database List. Nucleic Acids Research 27: 1-9, 1999, and links to useful databases on the internet at: http://www.eup.co.uk/nar/Volume_27/issue_01/summary/gkc105_gml.html. In particular see the Transcription Factor Database (Heinemeyer, T., et. al. (1999) Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. Nucleic Acids Res. 27: 318-322, or on the internet at: http://转入.175.44.40/TRANSFAC/index.html). Any sequence variances in transcriptional regulatory sequences can be assessed for their effects on mRNA levels using standard methods, either by making plasmid constructs with the different allelic forms of the sequence, transfecting them into cells and measuring the output of a reporter transcript, or by assays of cells with different endogenous alleles of variances. One example of a polymorphism in a transcriptional regulatory element that has a pharmacogenetic effect is described by Drazen et al. (1999) Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. Nature Genetics 22: 168-170. Drazen and co-workers found that a polymorphism in an Sp1-transcription factor binding domain, which varied among subjects from 3-6 tandem copies, accounted for varied expression levels of the 5-Lipoxygenase gene when assayed in vitro in reporter construct assays. This effect would have been flagged by an informatics analysis that surveyed the 5-lipoxygenase candidate promoter region for transcriptional regulatory sequences (resulting in discovery of polymorphism in the Sp1 motif).

4. Perform In Vitro or In Vivo Experiments to Assess the Functional Importance of Gene Variances

There are two broad types of studies useful for assessing the likely importance of variances: analysis of RNA or protein abundance (as described above in the context of methods for identifying candidate genes for explaining interpatient variation in treatment response) or analysis of functional differences in different variant forms of a gene, mRNA or protein. Studies of functional differences may involve direct measurements of biochemical activity of different variant forms of an mRNA or protein, or may involve assaying the influence of a variance or variances on various cell properties, including both tissue culture and in vivo studies.

The selection of an appropriate experimental program for testing the medical consequences of a variance may differ depending on the nature of the variance, the gene, and the disease. For example if there is already evidence that a protein is involved in the pharmacologic action of a drug, then the in vitro or in vivo demonstration that an amino acid variance in the protein affects its biochemical activity is strong evidence that the variance will have an effect on the pharmacology of the drug in patients, and therefore that patients with different variant forms of the gene may have different responses to the same dose of drug. If the variance is silent with respect to protein coding information, or if it lies in a noncoding portion of the gene (e.g., a promoter, an intron, or a 5'- or 3'-untranslated region) then the appropriate biochemical assay may be to assess mRNA abundance, half life, or translational efficiency. If, on the other hand, there is no substantial evidence that the protein encoded by a particular gene is relevant to drug pharmacology, but instead is a candidate gene on account of its involvement in disease pathophysiology, then the optimal test may be a clinical study addressing whether two patient groups distinguished on the basis of the variance respond differently to a therapeutic intervention. This approach reflects the current reality that biologists do not sufficiently understand gene regulation, gene expression and gene function to consistently make accurate inferences about the consequences of DNA sequence variances for pharmacological responses.

In summary, if there is a plausible hypothesis regarding the effect of a protein on the action of a drug, then in vitro and in vivo approaches, including those described below, will be useful to predict whether a given variance is therapeutically consequential. If, on the other hand, there is no evidence of such an effect, then the preferred test is an empirical clinical measure of the impact to the variance on
efficacy or toxicity in vivo (which requires no evidence or assumptions regarding the mechanism by which the variance may exert an effect on a therapeutic response). However, given the expense and statistical constraints of clinical trials, it is preferable to limit clinical testing to variances for which there is at least some experimental or computational evidence of a functional effect.

Experimental Methods: Genomic DNA Analysis

[0425] Variances in DNA may affect the basal transcription or regulated transcription of a gene locus. Such variances may be located in any part of the gene but are most likely to be located in the promoter region, the first intron, or in 5' or 3' flanking DNA, where enhancer or silencer elements may be located. Methods for analyzing transcription are well known to those skilled in the art and exemplary methods are briefly described above and in some of the texts cited elsewhere in this application. Transcriptional run off assay is one useful method. Detailed protocols can be found in texts such as: Current Protocols in Molecular Biology edited by: F. M. Ausubel, et al. John Wiley & Sons, Inc. 1999, or, Molecular Cloning: A Laboratory Manual by J. Sambrook, E. F. Fritsch and T Maniatis. 1989. 3 vols, 2nd edition, Cold Spring Harbor Laboratory Press

Experimental Methods: RNA Analysis

[0426] RNA variances may affect a wide range of processes including RNA splicing, polyadenylation, capping, export from the nucleus, interaction with translation initiation, elongation or termination factors, or the ribosome, or interaction with cellular factors including regulatory proteins, or factors that may affect mRNA half life. However, the effect of most RNA sequence variances on RNA function, if any, should ultimately be measurable as an effect on RNA or protein levels—either basal levels or regulated levels or levels in some abnormal cell state, such as cells from patients with a disease. Therefore, one preferred method for assessing the effect of RNA variances on RNA function is to measure the levels of RNA produced by different alleles in one or more conditions of cell or tissue growth. Said measuring can be done by conventional methods such as Northern blots or RNAase protection assays (kits available from Ambion, Inc.), or by methods such as the Taqman assay (developed by the Applied Biosystems Division of the Perkin Elmer Corporation), or by using arrays of oligonucleotides or arrays of cDNAs attached to solid surfaces. Systems for arraying cDNAs are available commercially from companies such as Nanogen and General Scanning. Complete systems for gene expression analysis are available from companies such as Molecular Dynamics. For recent reviews of systems for high throughput RNA expression analysis see the supplement to volume 21 of Nature Genetics entitled “The Chipping Forecast”, especially articles beginning on pages 9, 15, 20 and 25.

[0427] Additional methods for analyzing the effect of variances on RNA include secondary structure probing, and direct measurement of half life or turnover. Secondary structure can be determined by techniques such as enzymatic probing (using enzymes such as T1, T2 and S1 nucleases), chemical probing or RNAase H probing using oligonucleotides. Most RNA structural assays are performed in vitro, however some techniques can be performed on cell extracts or even in living cells, using fluorescence resonance energy transfer to monitor the state of RNA probe molecules.

Experimental Methods: Protein Analysis

[0428] There are a variety of experimental methods for investigating the effect of an amino acid variance on response of a patient to a treatment. The preferred method will depend on the availability of cells expressing a particular protein, and the feasibility of a cell-based assay vs. assays on cell extracts, on proteins produced in a foreign host, or on proteins prepared by in vitro translation.

[0429] For example, the methods and systems listed below can be utilized to demonstrate differential expression, stability and/or activity of different variant forms of a protein, or in phenotype/genotype correlations in a model system.

[0430] For the determination of protein levels or protein activity a variety of techniques are available. The in vitro protein activity can be determined by transcription or translation in bacteria, yeast, baculoivirus, COS cells (transient), Chinese Hamster Ovary (CHO) cells, or studied directly in human cells, or other cell systems can be used. Further, one can perform pulse chase experiments to determine if there are changes in protein stability (half-life).

[0431] One skilled in the art can construct cell based assays of protein function, and then perform the assays in cells with different genotypes or haplotypes. For example, identification of cells with different genotypes, e.g. cell lines established from families and subsequent determination of relevant protein phenotypes (e.g. expression levels, post translational modifications, activity assays) may be performed using standard methods.

[0432] Assays of protein levels or function can also be performed on cell lines (or extracts from cell lines) derived from pedigrees in order to determine whether there is a genetic component to variation in protein levels or function. The experimental analysis is as above for RNAs, except the assays are different. Experiments can be performed on naive cells or on cells subjected to various treatments, including pharmacological treatments.

[0433] In another approach to the study of amino acid variances one can express genes corresponding to different alleles in experimental organisms and examine effects on disease phenotype (if relevant in the animal model), or on response to the presence of a compound. Such experiments may be performed in animals that have disrupted copies of the homologous gene (e.g. gene knock out animals engineered to be deficient in a target gene), or variant forms of the human gene may be introduced into germ cells by transgenic methods, or a combination of approaches may be used. To create animal strains with targeted gene disruptions a DNA construct is created (using DNA sequence information from the host animal) that will undergo homologous recombination when inserted into the nucleus of an embryonic stem cell. The targeted gene is effectively inactivated due to the insertion of non-natural sequence—for example a translation stop codon or a marker gene sequence that interrupts the reading frame. Well known PCR based methods are then used to screen for those cells in which the desired homologous recombination event has occurred. Gene knockouts can be accomplished in worms, drosophila, mice or other organisms. Once the knockout cells are created (in whatever species) the candidate therapeutic intervention can be administered to the animal and pharmacological or biological responses measured, including gene expression
levels. If variant forms of the gene are useful in explaining interpatient variation in response to the compound in man, then complete absence of the gene in an experimental organism should have a major effect on drug response. As a next step various human forms of the gene can be introduced into the knockout organism (a technique sometimes referred to as a knock-in). Again, pharmacological studies can be performed to assess the impact of different human variances on drug response. Methods relevant to the experimental approaches described above can be found in the following exemplary texts:

General Molecular Biology Methods


Short Protocols in Molecular Biology, Ausubel et al. October 1995. 3rd edition, John Wiley and Sons


Polymerase Chain Reaction (PCR)


The Polymerase Chain Reaction, K. B. Mullis et al. (eds.), 1994. Birkhauser


General Procedures for Discipline Specific Studies


IV. Clinical Trials

[0434] A clinical trial is the definitive test of the utility of a variance or variances for the selection of optimal therapy. A clinical trial in which an interaction of gene variances and clinical outcomes (desired or undesired) is explored will be referred to herein as a “pharmacogenetic clinical trial”. Pharmacogenetic clinical trials require no knowledge of the biological function of the gene containing the variance or variances to be assessed, nor any knowledge of how the therapeutic intervention to be assessed works at a biochemical level. The pharmacogenetics effects of a variance can be addressed at a purely statistical level: either a particular variance or set of variances is consistently associated with a significant difference in a salient drug response parameter (e.g. response rate, effective dose, side effect rate, etc.) or not. On the other hand, if there is information about either the biochemical basis of a therapeutic intervention or the biochemical effects of a variance, then a pharmacogenetic clinical trial can be designed to test a specific hypothesis. In preferred embodiments of the methods of this application the mechanism of action of the compound to be genetically analyzed is at least partially understood.

[0435] Methods for performing clinical trials are well known in the art. (see e.g. Guide to Clinical Trials by Bert Spilker, Raven Press, 1991; The Randomized Clinical Trial and Therapeutic Decisions by Niels Lystrup (Editor), Marcel Dekker; Recent Advances in Clinical Trial Design and Analysis (Cancer Treatment and Research, Cur 75) by Peter F. Thall (Editor) Kluwer Academic Pub, 1995. Clinical Trials: A Methodologic Perspective by Steven Pantelodis, Wiley Series in Probability and Statistics, 1997). However, performing a clinical trial to test the genetic contribution to interpatient variation in drug response entails additional design considerations, including (i) defining the genetic hypothesis or hypotheses, (ii) devising an analytical strategy for testing the hypothesis, including determination of how many patients will need to be enrolled to have adequate statistical power to measure an effect of a specified magnitude (power analysis), (iii) definition of any primary or secondary genetic endpoints, and (iv) definition of methods of statistical genetic analysis, as well as other aspects. In the outline below some of the major types of genetic hypothesis testing, power analysis and statistical testing and their application in different stages of the drug development process are reviewed. One skilled in the art will recognize that certain of the methods will be best suited to specific clinical situations, and that additional methods are known and can be used in particular instances.

A. Performing a Clinical Trial: Overview

[0436] As used herein, a “clinical trial” is the testing of a therapeutic intervention in a volunteer human population for the purpose of determining whether it is safe and/or efficacious in the treatment of a disease, disorder, or condition. The present invention describes methods for achieving superior efficacy and/or safety in a genetically defined subgroup defined by the presence or absence of at least one gene sequence variance, compared to the effect that could be obtained in a conventional trial (without genetic stratification).
A "clinical study" is that part of a clinical trial that involves determination of the effect of a candidate therapeutic intervention on human subjects. It includes clinical evaluation of physiologic responses including pharmacokinetic (bioavailability as affected by drug absorption, distribution, metabolism and excretion) and pharmacodynamic (physiologic response and efficacy) parameters. A pharmacogenetic clinical study (or clinical trial) is a clinical study that involves testing of one or more specific hypotheses regarding the interaction of a genetic variance or variances (or set of variances, i.e. haplotype or haplotypes) on response to a therapeutic intervention. Pharmacogenetic hypotheses are formulated before the study, and may be articulated in the study protocol in the form of primary or secondary endpoints. For example an endpoint may be that in a particular genetic subgroup the rate of objectively defined responses exceeds the response rate in a control group (either the entire control group or the subgroup of controls with the same genetic signature as the treatment subgroup they are being compared to) or exceeds that in the whole treatment group (i.e. without genetic stratification) by some predefined relative or absolute amount.

For a clinical study to commence enrollment and proceed to treat subjects at an institution that receives any federal support (most medical institutions in the US), an application that describes in detail the scientific premise for the therapeutic intervention and the procedures involved in the study, including the endpoints and analytical methods to be used in evaluating the data, should be reviewed and accepted by a review panel, often termed an Institutional Review Board (IRB). Similarly any clinical study that will ultimately be evaluated by the FDA as part of a new drug or product application (or other application as described below), should be reviewed and approved by an MRB. The IRB is responsible for determining that the trial protocol is safe, conforms to established ethical principles and guidelines, has risks proportional to any expected benefits, assures equitable selection of patients, provides sufficient information to patients (via a consent form) to insure that they can make an informed decision about participation, and insures the privacy of participants and the confidentiality of any data collected. (See the report of the National Commission for Protection of Human Subjects of Biomedical and Behavioral Research (1978). The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. Washington, D.C.: DHEW Publication Number (OS) 78-0012. For a recent review see: Coughlin, S. S. (ed.) (1995) Ethics in Epidemiology and Clinical Research. Epidemiology Resources, Newton, Mass.) The European counterpart of the US FDA is the European Medicines Evaluation Agency (EMEA). Similar agencies exist in other countries and are responsible for insuring, via the regulatory process, that clinical trials conform to similar standards as are required in the US. The documents reviewed by an IRB include a clinical protocol containing the information described above, and a consent form.

It is also customary, but not required, to prepare an investigator’s brochure which describes the scientific hypothesis for the proposed therapeutic intervention, the preclinical data, and the clinical protocol. The brochure is made available to any physician participating in the proposed or ongoing trial.

The supporting preclinical data is a report of all the in vitro, in vivo animal or previous human trial or other data that supports the safety and/or efficacy of a given therapeutic intervention. In a pharmacogenetic clinical trial the preclinical data may also include a description of the effect of a specific genetic variance or variances on biochemical or physiologic experimental variables in vitro or in vivo, or on treatment outcomes, as determined by in vivo studies in animals or humans (for example in an earlier trial), or by retrospective genetic analysis of clinical trial or other medical data (see below) used to formulate or strengthen a pharmacogenetic hypothesis. For example, case reports of unusual pharmaco logical responses in individuals with rare alleles (e.g. mutant alleles), or the observation of clustering of pharmaco logical responses in family members may provide the rationale for a pharmacogenetic clinical trial.

The clinical protocol provides the relevant scientific and therapeutic introductory information, describes the inclusion and exclusion criteria for human subject enrollment, including genetic criteria if relevant (e.g. if genotype is to be among the enrollment criteria), describes in detail the exact procedure or procedures for treatment using the candidate therapeutic intervention, describes laboratory analyses to be performed during the study period, and further describes the risks (both known and possible) involving the use of the experimental candidate therapeutic intervention. In a clinical protocol for a pharmacogenetic clinical trial, the clinical protocol will further describe the genetic variance and/or variances hypothesized to account for differential responses in the normal human subjects or patients and supporting preclinical data, if any, a description of the methods for genotyping, genetic data collection and data handling as well as a description of the genetic statistical analysis to be performed to measure the interaction of the variance or variances with treatment response. Further, the clinical protocol for a pharmacogenetic clinical trial will include a description of the genetic study design. For example patients may be stratified by genotype and the response rates in the different groups compared, or patients may be segregated by response and the genotype frequencies in the different responder or nonresponder groups measured. One or more gene sequence variances or a combination of variances and/or haplotypes may be studied.

The informed consent document is a description of the therapeutic intervention and the clinical protocol in simple language (e.g. third grade level) for the patient to read, understand, and, if willing, agree to participate in the study by signing the document. In a pharmacogenetic clinical study the informed consent document will describe, in simple language, the use of a genetic test or a limited set of genetic tests to determine the subject or patient’s genotype at a particular gene variance or variances, and to further ascertain whether, in the study population, particular variances are associated with particular clinical or physiological responses. The consent form should also describe procedures for assuring privacy and confidentiality of genetic information.

The US FDA reviews proposed clinical trials through the process of an Investigational New Drug Application (IND). The SD is composed of the investigator’s brochure, the supporting in vitro and in vivo animal or previous human data, the clinical protocol, and the informed consent forms. In each of the sections of the IND, a specific
description of a single allelic variance or a number of variances to be tested in the clinical study will be included. For example, in the investigator’s brochure a description of the gene or genes hypothesized to account, at least in part, for differential responses will be included as well as a description of a genetic variance or variances in one or more candidate genes. Further, the preclinical data may include a description of in vivo, in vitro or in silico studies of the biochemical or physiologic effects of a variance or variances (e.g., haplotype) in a candidate gene or genes, as well as the predicted effects of the variance or variances on efficacy or toxicology of the candidate therapeutic intervention. The results of retrospective genetic analysis of response data in patients treated with the candidate therapy may be the basis for formulating the genetic hypotheses to be tested in the prospective trial. The US FDA reviews applications with particular attention to safety and toxicological data to ascertain whether candidate compounds should be tested in humans.

[0444] The established phases of clinical development are Phase I, II, III, and IV. The fundamental objectives for each phase become increasingly complex as the stages of clinical development progress. In Phase I, safety in humans is the primary focus. In these studies, dose-ranging designs establish whether the candidate therapeutic intervention is safe in the suspected therapeutic concentration range. However, it is common practice to obtain information about surrogate markers of efficacy even in phase I clinical trials. In a pharmacogenetic clinical trial there may be an analysis of the effect of a variance or variances on Phase I safety or surrogate efficacy parameters. At the same time, evaluation of pharmacokinetic parameters (e.g., adsorption, distribution, metabolism, and excretion) may be a secondary objective; again, in a pharmacogenetic clinical study there may be an analysis of the effect of sequence variation in genes that affect absorption, distribution, metabolism and excretion of the candidate compound on pharmacokinetic parameters such as peak blood levels, half life or tissue distribution of the compound. As clinical development stages progress, trial objectives focus on the appropriate dose and method of administration required to elicit a clinically relevant therapeutic response. In a pharmacogenetic clinical trial, there may be a comparison of the effectiveness of several doses of a compound in patients with different genotypes, in order to identify interactions between genotype and optimal dose. For this purpose the doses selected for late stage clinical testing may be greater, equal or less than those chosen based upon preclinical safety and efficacy determinations. Data on the function of different alleles of genes affecting pharmacokinetic parameters could provide the basis for selecting an optimal dose or range of doses of a compound or biological. Genes involved in drug metabolism may be particularly useful to study in relation to understanding interpatient variation in optimal dose. Genes involved in drug metabolism include the cytochrome P450s, especially 2D6, 3A4, 2C9, 2E1, 2A6 and 1A1; the glucuronyltransferases; the acetyltransferases; the methyltransferases; the sulfotransferases; the glutathione system; the flavine monoxygenases and other enzymes known in the art.

[0445] An additional objective in the latter stages of clinical development is demonstration of the effect of the therapeutic intervention on a broad population. In phase III trials, the number of individuals enrolled is dictated by a power analysis. The number of patients required for a given pharmacogenetic clinical trial will be determined by prior knowledge of variance or haplotype frequency in the study population, likely response rate in the treated population, expected magnitude of pharmacogenetic effect (for example, the ratio of response rates between a genetic subgroup and the unfractonated population, or between two different genetic subgroups); nature of the genetic effect, if known (e.g., dominant effect, codominant effect, recessive effect); and number of genetic hypotheses to be evaluated (including number of genes and/or variances to be studied, number of gene or variance interactions to be studied). Other considerations will likely arise in the design of specific trials.

[0446] Clinical trials should be designed to blind both human subjects and study coordinators from biasing that may otherwise occur during the testing of a candidate therapeutic invention. Often the candidate therapeutic intervention is compared to best medical treatment, or a placebo (a compound, agent, device, or procedure that appears identical to the candidate therapeutic intervention but is therapeutically inert). The combination of a placebo group and blind controls for potentially confounding factors such as prejudice on the part of study participants or investigators, insures that real, rather than perceived or expected, effects of the candidate therapeutic intervention are measured in the trials. Ideally blinding extends not only to trial subjects and investigators but also to data review committees, ancillary personnel, statisticians, and clinical trial monitors.

[0447] In pharmacogenetic clinical studies, a placebo arm or best medical control group may be required in order to ascertain the effect of the allelic variance or variances on the efficacy or toxicology of the candidate therapeutic intervention as well as placebo or best medical therapy. It will be important to assure that the composition of the control and test populations are matched, to the degree possible, with respect to genetic background and allele frequencies. This is particularly true if the variances being investigated may have an effect on disease manifestations (in addition to a hypothesized effect on response to treatment). It is likely that standard clinical trial procedures such as insuring that treatment and control groups are balanced for race, sex and age composition and other non-genetic factors relevant to disease will be sufficient to assure that genetic background is controlled, however a preferred practice is to explicitly test for genetic stratification between test and control groups. Methods for minimizing the possibility of spurious results attributable to genetic stratification between two comparison groups include the use of surrogate markers of geographic, racial and/or ethnic background, such as have been described by Rannala and coworkers. (See, for example: Rannala B, and J L Mountain. 1997 Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA August 19, 94(17):9197-201.) One procedure would be to assure that surrogate markers of genetic background (such as those described by Rannala and Mountain) occur at comparable frequency in two comparison groups.

[0448] Open label trials are unblinded; in single blind trials patients are kept unaware of treatment assignments; in double blind trials both patients and investigators are unaware of the treatment groups; a combination of these procedures may be instituted during the trial period. Pharmacogenetic clinical trial design may include one or a combination of open label, single blind, or double blind clinical trial designs. Reduction of biases attributable to the
knowledge of either the type of treatment or the genotype of the normal subjects or patients is an important aspect of study design. So, for example, even in a study that is single blind with respect to treatment, it should be possible to keep both patients and caregivers blinded to genotype during the study.

[0449] In designing a clinical trial it is important to include termination endpoints such as adverse clinical events, inadequate study participation either in the form of lack of adherence to the clinical protocol or loss to follow up, (e.g. such that adequate power is no longer assured), lack of adherence on the part of trial investigators to the trial protocol, or lack of efficacy or positive response within the test group. In a pharmacogenetic clinical trial these considerations obtain not only in the entire treatment group, but also in the genetically defined subgroups. That is, if a dangerous toxic effect manifests itself predominantly or exclusively in a genetically defined subpopulation of the total treatment population it may be deemed inappropriate to continue treating that genetically defined subgroup. Such decisions are typically made by a data safety monitoring committee, a group of experts not including the investigators, and generally not blinded to the analysis, who review the data from an ongoing trial on a regular basis.

[0450] It is important to note that medicine is a conservative field, and clinicians are unlikely to change their behavior on the basis of a single clinical trial. Thus it is likely that, in most instances, two or more clinical trials will be required to convince physicians that they should change their prescribing habits in view of genetic information. Large scale trials represent one approach to providing increased data supporting the utility of a genetic stratification. In such trials the stringent clinical and laboratory data collection characteristic of traditional trials is often relaxed in exchange for a larger patient population. Important goals in large scale pharmacogenetic trials will include establishing whether a pharmacogenetic effect is detectable in all segments of a population. For example, in the North American population one might seek to demonstrate a pharmacogenetic effect in people of African, Asian, European and Hispanic (i.e. Mexican and Puerto Rican) origin, as well as in native American people. (It generally will not be practical to segment patients by geographical origin in a standard clinical trial, due to loss of power.) Another goal of a large scale clinical trial may be to measure more precisely, and with greater confidence, the magnitude of a pharmacogenetic effect first identified in a smaller trial. Yet another undertaking in a large scale clinical trial may be to examine the interaction of an established pharmacogenetic variable (e.g. a variance in gene A, shown to affect treatment response in a smaller trial) with other genetic variances (either in gene A or in other candidate genes). A large scale trial provides the statistical power necessary to test such interactions.

[0451] In designing all of the above stages of clinical testing investigators should be attentive to the statistical problems raised by testing multiple different hypotheses, including multiple genetic hypotheses, in subsets of patients. Bonferroni's correction or other suitable statistical methods for taking account of multiple hypothesis testing will need to be judiciously applied. However, in the early stages of clinical testing, when the main goal is to reduce the large number of potential hypotheses that could be tested to a few that will be tested, based on limited data, it may be impractical to rigidly apply the multiple testing correction.

B. Phase I Clinical Trials

[0452] 1. Introduction

[0453] Phase I of clinical development is generally focused on safety, although drug companies are increasingly obtaining information on pharmacokinetics and surrogate pharmacodynamic markers in early trials. Phase I studies are typically performed with a small number (<60) of normal, healthy volunteers usually at single institutions. The primary endpoints in these studies usually relate to pharmacokinetic parameters (i.e. adsorption, distribution, metabolism and bioavailability), and dose-related side effects. In a Phase I pharmacogenetic clinical trial, stratification based upon allelic variance or variances of a candidate gene or genes related to pharmacokinetic parameters may allow early assessment of potential genetic interactions with treatment.

[0454] Phase I studies of some diseases (e.g. cancer or other medically intractable diseases for which no effective medical alternative exists) may include patients who satisfy specified inclusion criteria. These safety/limited-efficacy studies can be conducted at multiple institutions to ensure rapid enrollment of patients. In a pharmacogenetic Phase I study that includes patients, or a mixture of patients and normals, the status of a variance or variances suspected to affect the efficacy of the candidate therapeutic intervention may be used as part of the inclusion criteria. Alternatively, analysis of variances or haplotypes in patients with different treatment responses may be among the endpoints. It is not unusual for such a Phase I study design to include a double-blind, balanced, random-order, crossover sequence (separated by washout periods), with multiple doses on separate occasions and both pharmacokinetic and pharmacodynamic endpoints.

[0455] 2. Phase I Trials with Subjects Drawn from Large Populations and/or from Related Volunteer Subjects: the Pharmacogenetic Phase I Unit Concept

[0456] In general it is useful to be able to assess the contribution of genetic variation to treatment response at the earliest possible stage of clinical development. Such an assessment, if accurate, will allow efficient prioritization of candidate compounds for subsequent detailed pharmacogenetic studies; only those treatments where there is early evidence of a significant interaction of genetic variation with treatment response would be advanced to pharmacogenetic studies in later stages of development. In this invention we describe methods for achieving early insight—in Phase I—into the contribution of genetic variation to variation in surrogate treatment response variables. It occurred to the inventors that this can be accomplished by bringing the power of genetic linkage analysis and outlier analysis to Phase I testing via the recruitment of a very large Phase I population including a large number of individuals who have consented in advance to genetic studies (occasionally referred to hereinafter as a Pharmacogenetic Phase I Unit). In one embodiment of a Pharmacogenetic Phase I Unit many of the subjects are related to each other by blood. (Currently Phase I trials are performed in unrelated individuals, and there is no consideration of genetic recruitment criteria, or of genetic analysis of surrogate markers.) There are several novel ways in which a large population, or a population...
comprised at least in part of related individuals, could be useful in early clinical trials. Some of the most attractive applications depend on the availability of surrogate markers for pharmacodynamic drug action which can be used early in clinical development, preferably in normal subjects in Phase I. Such surrogate markers are increasingly used in Phase I, as drug development companies seek to make early yes/no decisions about compounds.

Recruitment of a population optimized for clinical genetic investigation may entail utilization of methods in statistical genetics to select the size and composition of the population. For example powerful methods for detecting and mapping quantitative trait loci in subpairs have been developed. These methods can provide some estimate of the statistical power derived from a given number of groups of closely related individuals. Ideally subjects in the pharmacogenetic Phase I unit are of known ethnic/racial/geographic background and willing to participate in Phase I studies, for pay, over a period of years. The population is preferably selected to achieve a specified degree of statistical power for genetic association studies, or is selected in order to be able to reliably identify a certain number of individuals with rare genotypes, as discussed below. Family participation could be encouraged by appropriate incentive compensation. For example, individual subjects might be paid $200 for participation in a study; two sibs participating in the same study might each be paid $300; if they could encourage another sib (or cousin) to participate the three related individuals might each be paid $350, and so forth. This type of compensation would encourage subjects to recruit their relatives to participate in Phase I studies. (It would also increase the cost of studies, however the type of data that can be obtained cannot be duplicated with conventional approaches.) The optimal location to establish such a Phase I unit is a city with a stable population, many large families, and a positive attitude about gene technology. The Pharmacogenetic Phase I Unit population can then be used to test for the existence of genetic variation in response to any drug as a first step in deciding whether to proceed with extensive pharmacogenetic studies in later stages of clinical development. Specific uses of a large Phase I unit in which some or all subjects are related include:

- It should be possible, for virtually any compound, to assess the magnitude of the genetic contribution to variation in drug response (if any) by comparing variation in drug response traits among related vs. non-related individuals. The rationale is as follows: if a surrogate drug response trait (i.e. a surrogate marker of pharmacodynamic effect that can be measured in normal subjects) is under strong genetic control then related individuals, who share 25% (cousins) or 50% (sibs) of their alleles, should have less divergent responses (less intragroup variance) than unrelated individuals, who share a much smaller fraction of alleles. That is, individuals who share alleles at the genes that affect drug response should be more similar to each other (i.e. have a narrower distribution of responses, whether measured by variance, standard deviation or other means) than individuals who, on average, share very few alleles. By using statistical methods known in the art the degree of variation in a set of data from related individuals (each individual would only be compared with his/her relatives, but such comparisons would be performed within each group of relatives and a summary statistic developed) could be compared to the degree of variation in a set of unrelated individuals (the same subjects could be used, but the second comparison would be across related groups). Account would be taken of the degree of similarity expected between related individuals, based on the fraction of the genome they shared by descent. Thus the extent of variation in the surrogate response marker between identical twins should be less than between sibs, which should be less than between first cousins, which should be less than that between second cousins, and so forth, if there is a genetic component to the variation. It is well known from twin studies (in which, for example, variation between identical twins is compared to variation between fraternal twins) that pharmacokinetic variables (e.g. compound half life, peak concentration) are frequently over 90% heritable; the type of study proposed here (comparison of variation within groups of sibs and cousins to variation between unrelated subjects) would also show this genetic effect, without requiring the recruitment of monozygotic twins. For a summary of pharmacokinetic studies in twins see: Propping, Paul (1978) Pharmacogenetics. Rev. Physiol. Biochem. Pharmacol. 83: 123-173.

- It may be that the pattern of drug responses that distinguishes related individuals from non-related individuals is more complex than, for example, variance or standard deviation. For example, there may be two discrete phenotypes characteristic of intrafamilial variation (a bimodal distribution) that are not a feature of variation between unrelated individuals (where, for example, variation might be more nearly continuous). Such a pattern could be attributable to Mendelian inheritance operating on a restricted set of alleles in a family (or families) with, for example, AA homozygotes giving one phenotype and AB heterozygotes and BB homozygotes giving a second phenotype, all in the context of a relatively homogeneous genetic background. In contrast, variation among non-related subjects would be less discrete due to a greater degree of variation in genetic background and the presence of additional alleles C, D and E at the candidate locus. Statistical measures of the significance of such differences in distribution, including nonparametric methods such as chi square and contingency tables, are known in the art.

- The methods described herein for measuring whether pharmacodynamic traits are under genetic control, using surrogate markers of drug efficacy in phase I studies which include groups of related individuals, will be useful in obtaining an early assessment of the extent of genetically determined variation in drug response for a given therapeutic compound. Such information provides an informed basis for either stopping development at the earliest possible stage or, preferably, continuing with development but with a plan for identifying and controlling for genetic variation so as to allow rapid progression through the regulatory approval process.

- For example, it is well known that Alzheimers trials are long and expensive, and most drugs are only effective in a fraction of patients. Using surrogate measures of response in normals drawn from a population of related individuals would help to assess the contribution of genetic variation to variation in treatment response. For an acetylcholinesterase inhibitor, relevant surrogate pharmacodynamic measures could include testing erythrocyte membrane acetylcholinesterase levels in drug treated normal subjects, or performing psychometric tests that are affected by treatment (and ideally that correlate with clinical efficacy) and
measuring the effect of treatment. As another example, antidepressant drugs can produce a variety of effects on mood in normal subjects—or no effect at all. Careful monitoring and measurement of such responses in related vs. unrelated normal subjects, and statistical comparison of the degree of variation in each group, could provide an early readout on whether there is a genetic component to drug response (and hence clinical efficacy). The observation of similar effects in family members, and comparatively dissimilar effects in unrelated subjects would provide compelling evidence of a pharmacogenetic effect and justify the substantial expenditure necessary for a full pharmacogenetic drug development program. Conversely, the absence of any significant family influence on drug response would provide an early termination point for pharmacogenetic studies. Note that the proposed studies do not require any knowledge of candidate genes, nor is DNA collection or genotyping required—simply a reliable surrogate pharmacodynamic assay and small groups of related normal individuals. Refined statistical methods should permit the magnitude of the pharmacogenetic effect to be measured, which could be a further criteria for deciding whether to proceed with pharmacogenetic analysis. The greater the differential in magnitude or pattern of variance between the related and the unrelated subjects, the greater the extent of genetic control of the trait.

[0462] Not all drug response traits are under the predominant control of one locus. Many such traits are under the control of multiple genes, and may be referred to as quantitative trait loci. It is then desirable to identify the major loci contributing to variation in the drug response trait. This can be done for example, to map quantitative trait loci in a population of drug treated related normals. Either a candidate gene approach or a genome wide scanning approach can be used. (For review of some relevant methods see: Hsu L, Aragaki C, Quijot I. (1999) A genome-wide scan for a simulated data set using two newly developed methods. Genet Epidemiol 17 Suppl 1:S621-6; Zhao L P, Aragaki C, Hsu L, Quijot I. (1998) Mapping of complex traits by single-nucleotide polymorphisms. Am J Hum Genet 63(1):225-40; Stoez M R, Cohen J C, Moos V, et al. (1997) Extension of the Haseman-Elston method to multiple alleles and multiple loci: theory and practice for candidate genes. Ann Hum Genet 61 (Pt 3):263-74.) However, this method would require at least 100 patients (preferably 200, and still more preferably >300) to have adequate statistical power, and each patient would have to be genotyped at a few polymorphic loci (candidate gene approach) or hundreds of polymorphic loci (genome scanning approach).

[0463] b. With a large Phase I population of normal subjects that need not be related (a second type of Pharmacogenetic Phase I Unit) it is possible to efficiently identify and recruit for any Phase I trial a set of individuals comprising virtually any combination of genotypes present in a population (for example, all common genotypes, or a group of genotypes expected to represent outliers for a drug response trait of interest). This method preferably entails obtaining blood or other tissue (e.g. buccal smear) in advance from a large number of the subjects in the Phase I unit. Ideally consent for genotyping would be obtained at the same time. It would be most efficient if blanket consent for genotyping any polymorphic site or sites could be obtained. Second best would be consent for testing any site relevant to any customer project (not specific at the time of initial consent). Third best would be consent to genotype polymorphic sites relevant to specific disease areas. Another, less desirable, solution would be to obtain consent for genotyping on a project by project basis (for example by mailing out reply cards), after the specific polymorphic sites to be genotyped are known.

[0464] One useful way to screen for pharmacogenetic effects in Phase I is to recruit homozygotes for a variance or variances of interest in one or more candidate genes. For example, consider a compound for which there are two genes that are strong candidates for influencing response to treatment. Gene X has alleles A and A', while gene Y has alleles B and B'. If these genes do in fact contribute significantly to response then one would expect that, regardless of the mode of inheritance (recessive, codominant, dominant, polygenic) homozygotes would exhibit the most extreme responses. One would also expect epistatic interactions, if any, to be most extreme in double homozygotes. Thus one would ideally perform a surrogate drug response test in Phase I volunteers doubly homozygous at both X and Y. That is, test AA/BB, A'A'/BB, AA/BB' and A'A'/BB' subjects. If the allele frequencies for A and A' are 0.15 and 0.85, and for B and B' 0.2 and 0.8 then the frequency of AA homozygotes is expected to be 2.25% and BB homozygotes 4%. In the absence of any linkage between the genes, the frequency of AA/BB double homozygotes is expected to be 0.0225x0.04=0.0009, or about 1 subject in 1000. Ideally at least 5 subjects of each genotype are recruited for the Phase I studies, and preferably at least 10 subject. Thus, even for variances of moderately low allele frequency (15%, 20%), the identification of potential outliers (i.e. homozygotes) for the candidate genes of interest will require a large population. Preferably the Phase I unit has enrolled at least 1,000 normal individuals, more preferably 2,000, still more preferably 5,000 and most preferably 10,000 or more. In another application of the large, genotyped Phase I population it may be useful to identify individuals with rare variances in candidates genes (either homozygous or heterozygous), in order to determine whether those variances are predisposing to extreme pharmacological responses to the compound. For example, variances occurring at 5% allele frequency are expected to occur in homozygous form in 0.25% of the population (0.05x0.05), and therefore may rarely, if ever, be encountered in early clinical development. Yet it may be serious adverse effects occurring in just such a small group that create problems in later stages of drug development. In yet another application of the large genotyped Phase I population, subjects may be selected to represent the known common variances in one or more genes that are candidates for influencing the response to treatment. By insuring that all common genotypes are represented in a Phase I trial the likelihood of misleading results due to genetic stratification (resulting in discrepancy with results of later, larger trials can be reduced.

[0465] It would be useful to prospectively genotype the large Phase I population for variances that are commonly the source of interpatient variation in drug response, since demand for genotyped groups of such patients can be anticipated from pharmaceutical companies and contract research organizations (CROs). For example, genotyping might initially focus on common pharmacological targets such as estrogen receptors, adrenergic receptors, or serotonin receptors. The pre-genotyped Phase I population could be part of a package of services (along with genotyping
assay development capability, high throughput genotyping capacity and software and expertise in statistical genetics) designed to accelerate pharmacogenetic Phase I studies. Eventually, as the databank of genotypes built up, individuals with virtually any genotype or combination of genotypes could be called in for precisely designed physiological or toxicological studies designed to test for pharmacogenetic effects.

[0466] One of the most useful aspects of the Pharmacogenetic Phase I Unit is that subjects with rare genotypes can be pharmacologically assessed in a small study. This addresses a serious limitation of conventional clinical trials with respect to the investigation of polygenic traits or the effect of rare alleles. Unfortunately even Phase III studies, as currently performed, are often barely powered to address simple one variance hypotheses about efficacy or toxicity. The problem, of course, is that each time a new genetic variable is introduced the comparison groups are cut in halves or thirds (or even smaller groups if there are multiple haplotypes at each gene). It is therefore a challenging problem to test the interaction of several genes in determining drug response. Yet the character of drug response data in populations—there is often a continuous distribution of responses among different individuals—suggests that drug responses may often be mediated by several genes. (On the other hand, there are an increasing number of well documented single gene, or even single variance, pharmacogenetic effects in the literature, showing that it is possible to detect the effect of a single variance.) One approach to identifying pharmacogenetic effects is to focus on finding the single gene variances that have the largest effects. This approach can be undertaken within the scale of current clinical trials. However, in order to develop a test which predicts a large fraction of the quantitative variation in a drug response trait it may be desirable to test the effect of multiple genes, including the interaction of variances at different genes, which may be non-additive (referred to as epistasis). The Pharmacogenetic Phase I Unit provides a way to efficiently test for gene interactions or multigene effects by, for example, allowing easy identification of individuals who, on account of being homozygous at several loci of interest, should be outliers for the drug response phenotypes of interest if there is a gene x gene interaction. Testing drug response in a small number of such individuals will provide a quick read on gene interaction. Obtaining genetic data on the pharmacodynamic action of a compound in Phase I should also provide a crude measure of allele affects—which variances or haplotypes increase pharmacological responses and which decrease them. This information is of great value in designing subsequent trials, as it constrains the number of hypotheses to be tested, thereby enabling powerful statistical designs. This is because when the effect of variances on drug response measures is unknown one is forced to statistically test all the possible effects of each allele (e.g. two tailed tests). As the number of genetically defined groups increases (e.g. as a result of multiple variances or haplotypes) there is a loss of statistical power due to multiple testing correction. On the other hand, if the relative genotypic effect of each allele at a locus is known (or can be hypothesized) from Phase I data then each individual in a subsequent clinical trial contributes useful information—there is a specific prediction of response based on that individuals combination of genotypes or haplotypes, and testing the fit of the actual data to those predictions provides for powerful statistical designs. (It is also possible to measure allele effects biochemically, of course, to establish which alleles have positive and which negative effects, but at considerable cost.) It is important to note that Phase I trials can provide useful information at almost any stage of clinical development. It is not unusual, for example, for a product in Phase II or even Phase III testing to be remanded to Phase I in order to clarify some aspect of toxicology or physiology. In this context a Pharmacogenetic Phase I Unit would be extremely useful to a drug development company. Phase I studies in defined genetic subgroups drawn from a large genotyped population, or in groups of related individuals, would be the most economical and efficient way to clarify the existence of pharmacogenetic effects, if any, paving the way for future rational development of the product.

C. Phase II Clinical Trials

[0467] Phase II studies generally include a limited number of patients (<100) who satisfy a set of predefined inclusion criteria and do not satisfy any predefined exclusion criteria of the trial protocol. Phase II studies can be conducted at one or more institutions. Inclusion/exclusion criteria may include historical, clinical and laboratory parameters for a disease, disorder, or condition; age; gender; reproductive status (i.e. pre- or postmenopausal); coexisting medical conditions; psychological, emotional or cognitive state, or other objective measures to those skilled in the art. In a pharmacogenetic Phase II trial the inclusion/exclusion criteria may include one or more genotypes or haplotypes. Alternatively, genetic analysis may be performed at the end of the trial. The primary goals in Phase II testing may include (i) identification of the optimal medical indication for the compound, (ii) definition of an optimal dose or range or doses, balancing safety and efficacy considerations (dose-finding studies), (iii) extended safety studies (complementing Phase I safety studies), (iv) evaluation of efficacy in patients with the targeted disease or condition, either in comparison to placebo or to current best therapy. To some extent these goals may be achieved by performing multiple trials with different goals. Likewise, Phase II trials may be designed specifically to evaluate pharmacogenetic aspects of the drug candidate. Primary efficacy endpoints typically focus on clinical benefit, while surrogate endpoints may measure treatment response variables such as clinical or laboratory parameters that track the progress or extent of disease, often at lesser time, cost or difficulty than the definitive endpoints. A good surrogate marker should be convincingly associated with the definitive outcome. Examples of surrogate endpoints include tumor size as a surrogate for survival in cancer trials, and cholesterol levels as a surrogate for heart disease (e.g. myocardial infarction) in trials of lipid lowering cardiovascular drugs. Secondary endpoints supplement the primary endpoint and may be selected to help guide further clinical studies.

[0468] In a pharmacogenetic Phase II clinical trial, retrospective or prospective design will include the stratification of patients based upon a variance or variance in a gene or genes suspected of affecting treatment response. The gene or genes may be involved in mediating pharmacodynamic or pharmacokinetic response to the candidate therapeutic intervention. The parameters evaluated in the genetically stratified trial population may include primary, secondary or surrogate endpoints. Pharmacokinetic parameters—for example, dosage, absorption, toxicity, metabolism, or excre-
tion—may also be evaluated in genetically stratified groups. Other parameters that may be assessed in parallel with genetic stratification include gender, race, ethnic or geographic origin (population history) or other demographic factors.

[0469] While it is optimal to initiate pharmacogenetic studies in phase I, as described above, it may be the case that pharmacogenetic studies are not considered until phase II, when problems relating either to efficacy or toxicity are first encountered. It is highly desirable to initiate pharmacogenetic studies no later than phase I of a clinical development plan because (1) phase III studies tend to be large and expensive—not an optimal setting in which to explore untested pharmacogenetic hypotheses; (2) phase III studies are typically designed to test one fairly narrow hypothesis regarding efficacy of one or a few dose levels in a specific disease or condition. Phase II studies are often numerous, and are intended to provide a broad picture of the pharmacology of the candidate compound. This is a good setting for initial pharmacogenetic studies. Several pharmacogenetic hypotheses may be tested in phase II, with the goal of eliminating all but one or two.

D. Phase III Clinical Trials

[0470] Phase III studies are generally designed to measure efficacy of a new treatment in comparison to placebo or to an established treatment method. Phase II studies are often performed at multiple sites. The design of this type of trial includes power analysis to ensure the sufficient data will be gathered to demonstrate the anticipated effect, making assumptions about response rate based on earlier trials. As a result Phase III trials frequently include large numbers of patients (up to 5,000). Primary endpoints in Phase III studies may include reduction or arrest of disease progression, improvement of symptoms, increased longevity or increased disease-free longevity, or other clinical measures known in the alt. In a pharmacogenetic Phase III clinical study, the endpoints may include determination of efficacy or toxicity in genetically defined-subgroups. Preferably the genetic analysis of outcomes will be confined to an assessment of the impact of a small number of variances or haplotypes at a small number of genes, said variances having already been statistically associated with outcomes in earlier trials. Most preferably variances at only one or two genes will be assessed.

[0471] After successful completion of one or more Phase III studies, the data and information from all trials conducted to test a new treatment method are compiled into a New Drug Application (NDA) and submitted for review by the US FDA, which has authority to grant marketing approval in the US and its territories. The NDA includes the raw (unanalyzed) clinical data, i.e., the patient by patient measurements of primary and secondary endpoints, a statistical analysis of all of the included data, a document describing in detail any observed side effects, tabulation of all patients who dropped-out of trials and detailed reasons for their termination, and any other available data pertaining to ongoing in vitro or in vivo studies since the submission of the investigational new drug (IND) application. If pharmacoeconomic objectives are a part of the clinical trial design then data supporting cost or economic analyses are included in the NDA. In a pharmacogenetic clinical study, the pharmacoeconomic analyses may include genetically stratified assessment of the candidate therapeutic intervention in a cost benefit analysis, cost of illness study, cost minimization study, or cost utility analysis. The analysis may also be simultaneously stratified by standard criteria such as race/ethnicity/geographic origin, sex, age or other criteria. Data from a genetically stratified analysis may be used to support an application for approval for marketing of the candidate therapeutic intervention.

E. Phase IV Clinical Trials

[0472] Phase IV studies occur after a therapeutic intervention has been approved for marketing, and are typically conducted for surveillance of safety, particularly occurrence of rare side effects. The other principal reason for Phase IV studies is to produce information and relationships useful for marketing a drug. In this regard pharmacogenetic analysis may be very useful in Phase IV trials. Consider, for example, a drug that is the fourth or fifth member of a drug class (say statins, or thiazidinediones or fluoropyrimidines) to obtain marketing approval, and which does not differ significantly in clinical effects—efficacy or safety—from other members of the drug class. The first, second and third drugs in the class will likely have a dominant market position (based on their earlier introduction into the marketplace) that is difficult to overcome, particularly in the absence of differentiating clinical effects. However, it is possible that the new drug produces a superior clinical effect—for example, higher response rate, greater magnitude of response or fewer side effects—in a genetically defined subgroup. The generic subgroup with superior response may constitute a larger fraction of the total patient population than the new drug would likely achieve otherwise. In this instance, there is a clear rationale for performing a Phase IV pharmacogenetic trial to identify a variance or variances that mark a patient population with superior clinical response. Subsequently a marketing campaign can be designed to alert patients, physicians, pharmacy managers, managed care organizations and other parties that, with the use of a rapid and inexpensive genetic test to identify eligible patients, the new drug is superior to other members of the class (including the market leader first, second and third drugs introduced). The high responder subgroup defined by a variance or variances may also exhibit a superior response to other drugs in the class (a class pharmacogenetic effect), or the superior efficacy in the genetic subgroup may be specific to the drug tested (a compound-specific pharmacogenetic effect).

[0473] In a Phase IV pharmacogenetic clinical trial, both retrospective and prospective analysis can be performed. In both cases, the key element is genetic stratification based on a variance or variances or haplotype. Phase IV trials will often have adequate sample size to test more than one pharmacogenetic hypothesis in a statistically sound way.

F. Unconventional Clinical Development

[0474] Although the above listed phases of clinical development are well-established, there are cases where strict Phase I, II, III development does not occur, for example, in the clinical development of candidate therapeutic interventions for debilitating or life threatening diseases, or for diseases where there is presently no available treatment. Some of the mechanisms established by the FDA for such studies include Treatment INDs, Fast-Track or Accelerated reviews, and Orphan Drug Status. In a clinical development
program for a candidate therapeutic of this type there is a useful role for pharmacogenetic analysis, in that the candidate therapeutic may not produce a sufficient benefit in all patients to justify FDA approval, however analysis of outcome in genetic subgroups may lead to identification of a variance or variances that predict a response rate sufficient for FDA approval.

[0475] As used herein, “supplemental applications” are those in which a candidate therapeutic intervention is tested in a human clinical trial in order to gain an expanded label indication, expanding recommended use to new medical indications. In these applications, previous clinical studies of the therapeutic intervention, i.e. preclinical safety and Phase I human safety studies can be used to support the testing of the therapeutic intervention in a new indication. Pharmacogenetic analysis is also useful in the context of clinical trials to support supplemental applications. Since these are, by definition, focused on diseases not selected for initial development the overall efficacy may not be as great as for the leading indication(s). The identification of genetic subgroups with high response rates may enable the rapid approval of supplemental applications for expanded label indications. In such instances part of the label indication may be a description of the variance or variances that define the group with superior response.

[0476] As used herein, “outcomes” or “therapeutic outcomes” describe the results and value of healthcare intervention. Outcomes can be multi-dimensional, and may include improvement of symptoms; regression of a disease, disorder, or condition; prevention of a disease or symptom; cost savings or other measures.

[0477] Pharmacoeconomics is the analysis of a therapeutic intervention in a population of patients diagnosed with a disease, disorder, or condition that includes at least one of the following studies: cost of illness study (CIS); cost benefit analysis (CBA), cost minimization analysis (CMA), or cost utility analysis (CUA), or an analysis comparing the relative costs of a therapeutic intervention with one or a group of other therapeutic interventions. In each of these studies, the cost of the treatment of a disease, disorder, or condition is compared among treatment groups. Costs have both direct (therapeutic interventions, hospitalization) and indirect (loss of productivity) components. Pharmacoeconomic factors may provide the motivation for pharmacogenetic analysis, particularly for expensive therapies that benefit only a fraction of patients. For example, interferon alpha is the only treatment that can cure hepatitis C virus infection, however viral infection is completely and permanently eliminated in less than a quarter of patients. Nearly half of patients receive virtually no benefit from alpha interferon, but may suffer significant side effects. Treatment costs are ~$10,000 per course. A pharmacogenetic test that could predict responders would save much of the cost of treating patients not able to benefit from interferon alpha therapy, and could provide the rationale for treating a population in a cost efficient manner, where treatment would otherwise be unaffordable.

[0478] As used herein, “health-related quality of life” is a measure of the impact of a disease, disorder, or condition on a patient’s activities of daily living. An analysis of the health-related quality of life is often included in pharmacoeconomic studies.

[0479] As used herein, the term “stratification” refers to the partitioning of patients into groups on the basis of clinical or laboratory characteristics of the patient. “Genetic stratification” refers to the partitioning of patients or normal subjects into groups based on the presence or absence of a variance or variances in one or more genes. The stratification may be performed at the end of the trial, as part of the data analysis, or may come at the beginning of a trial, resulting in creation of distinct groups for statistical or other purposes.

G. Power Analysis in Pharmacogenetic Clinical Trials

[0480] The basic goal of power calculations in clinical trial design is to insure that trials have adequate patients and controls to fairly assess, with statistical significance, whether the candidate therapeutic intervention produces a clinically significant benefit.

[0481] Power calculations in clinical trials are related to the degree of variability of the drug response phenotypes measured and the treatment difference expected between comparison groups (e.g. between a treatment group and a control group). The smaller the variance within each group being compared, and the greater the difference in response between the two groups, the fewer patients are required to produce convincing evidence of an effect of treatment. These two factors (variance and treatment difference) determine the degree of precision required to answer a specific clinical question.

[0482] The degree of precision may be expressed in terms of the maximal acceptable standard error of a measurement, the magnitude of variation in which the 95% confidence interval must be confined or the minimal magnitude of difference in a clinical or laboratory value that must be detectable (at a statistically significant level, and with a specified power for detection) in a comparison to be performed at the end of the trial (hypothesis test). The minimal magnitude is generally set at the level that represents the minimal difference that would be considered of clinical importance.

[0483] In pharmacogenetic clinical trials there are two countervailing effects with respect to power. First, the comparison groups are reduced in size (compared to a conventional trial) due to genetic partitioning of both the treatment and control groups into two or more subgroups. However, it is reasonable to expect that variability for a trait is smaller within groups that are genetically homogeneous with respect to gene variances affecting the trait. If this is the case then power is increased as a function of the reduction in variability within (genetically defined) groups.

[0484] In general it is pretenable to power a pharmacogenetic clinical trial to see an effect in the largest genetically defined subgroups. For example, for a variance with allele frequencies of 0.7 and 0.3 the common homozygote group will comprise 49% of all patients (0.7^2+0.7×0.3). It is most desirable to power the trial to observe an effect (either positive or a negative) in this group. If it is desirable to measure an effect of therapy in a small genetic group (for example, the 9% of patients homozygous for the rare allele) then genotyping should be considered as an enrollment criterion to insure a sufficient number of patients are enrolled to perform an adequately powered study.

[0485] Statistical methods for powering clinical trials are known in the art. See, for example: Shuster, J. J. (1990)

H. Statistical Analysis of Clinical Trial Data

[0486] There are a variety of statistical methods for measuring the difference between two or more groups in a clinical trial. One skilled in the art will recognize that different methods are suited to different data sets. In general, there is a family of methods customarily used in clinical trials, and another family of methods customarily used in genetic epidemiological studies. Methods in quantitative and population genetics designed to measure the association between genotypes and phenotypes, and to map and measure the effect of quantitative trait loci are also relevant to the task of measuring the impact of a variance on response to a treatment. Methods from any of these disciplines may be suitable for performing statistical analysis of pharmacogenetic clinical trial data, as is known to those skilled in the art.

[0487] Conventional clinical trial statistics include hypothesis testing and descriptive methods, as elaborated below. Guidance in the selection of appropriate statistical tests for a particular data set is provided in texts such as: Biostatistics: A Foundation for Analysis in the Health Sciences, 7th edition (Wiley Series in Probability and Mathematical Statistics, Applied Probability and Statistics) by Wayne W. Daniel, John Wiley & Sons, 1998; Bayesian Methods and Ethics in a Clinical Trial Design (Wiley Series in Probability and Mathematical Statistics, Applied Probability Section) by J. B. Kadane (Editor), John Wiley & Sons, 1996. Examples of specific hypothesis testing and descriptive statistical procedures that may be useful in analyzing clinical trial data are listed below.


[0489] (1) One-sample procedures (binomial confidence interval, Wilcoxon signed rank test, permutation test with general scores, generation of exact permutational distributions)

[0490] (2) Two-sample procedures (t-test, Wilcoxon-Mann-Whitney test, Normal score test, Median test, Van der Waerden test, Savage test, Logrank test for censored survival data, Wilcoxon-Gehan test for censored survival data, Cochran-Armitage trend test, permutation test with general scores, generation of exact permutational distributions)

[0491] (3) RxC contingency tables (Fisher’s exact test, Pearson’s chi-squared test, Likelihood ratio test, Kruskal-Wallis test, Jonckheere-Terpstra test, Linear-by-linear association test, McNemar’s test, marginal homogeneity test for matched pairs)

[0492] (4) Stratified 2x2 contingency tables (test of homogeneity for odds ratio, test of unity for the common odds ratio, confidence interval for the common odds ratio)

[0493] (5) Stratified 2xC contingency tables (all two-sample procedures listed above with stratification, confidence intervals for the odds ratios and trend, generation of exact permutational distributions)


[0495] (7) Analysis of variance and covariance with a nested (hierarchical) structure.

[0496] (8) Designs and randomized plans for nested and crossed experiments (completely randomized design for two treatment, split-plot design, hierarchical design, incomplete block design, latin square design)

[0497] (9) Nonlinear regression models

[0498] (10) Logistic regression for unstratified or stratified data, for binary or ordinal response data, using the logit link function, the normit function or the complementary log-log function.

[0499] (11) Probit, logit, ordinal logistic and gompit regression models.

[0500] (12) Fitting parametric models to failure time data that may be right-, left-, or interval-censored. Tested distributions can include extreme value, normal and logistic distributions, and, by using a log transformation, exponential, Weibull, lognormal, loglogistic and gamma distributions.

[0501] (13) Compute non-parametric estimates of survival distribution with right-censored data and compute rank tests for association of the response variable with other variables.

[0502] B. Descriptive Statistical Methods

[0503] Factor analysis with rotations

[0504] Canonical correlation

[0505] Principal component analysis for quantitative variables.

[0506] Principal component analysis for qualitative data.

[0507] Hierarchical and dynamic clustering methods to create tree structure, dendrogram or phenogram.

[0508] Simple and multiple correspondence analysis using a contingency table as input or raw categorical data.

[0509] Specific instructions and computer programs for performing the above calculations can be obtained from companies such as: SAS/STAT Software, SAS Institute Inc., Cary, N.C., USA; BMDP Statistical Software, BMDP Statistical Software Inc., Los Angeles, Calif., USA; SYSTAT software, SPSS Inc., Chicago, III., USA; StatXact & LogXact, CYTEL Software Corporation, Cambridge, Mass., USA.

C. Statistical Genetic Methods Useful for Analysis of Pharmacogenetic Data

[0510] A wide spectrum of mathematical and statistical tools may be useful in the analysis of data produced in pharmacogenetic clinical trials, including methods employed in molecular, population, and quantitative genetics, as well as genetic epidemiology. Methods developed for plant and animal breeding may be useful as well, particularly methods relating to the genetic analysis of quantitative traits.

The types of statistical analysis performed in different branches of genetics are outlined below as a guide to the relevant literature and publicly available software, some of which is cited.

**Molecular Evolutionary Genetics**

- Patterns of nucleotide variation among individuals, families/populations and across species and genera.
- Alignment of sequences and description of variation/polymorphisms among the aligned sequences, amounts of similarities and dissimilarities.
- Measurement of molecular variation among various regions of a gene, testing of neutrality models.
- Rates of nucleotide changes among coding and the non-coding regions within and among populations.
- Construction of phylogenetic trees using methods such as neighborhood joining and maximum parsimony; estimation of ages of variances using coalescent models.

**Population Genetics**

- Patterns of distribution of genes among genotypes and populations. Hardy-Weinberg equilibrium, departures form the equilibrium.
- Genotype and haplotype frequencies, levels of heterozygosities, polymorphism information contents of genes, estimation of haplotypes from genotypes; the F-M algorithm, and parsimony methods.
- Estimation of linkage disequilibrium and recombination.
- Hierarchical structure of populations, the F-statistics, estimation of inbreeding, selection and drift.
- Genetic admixture/migration and mutation frequencies.
- Spatial distribution of genotypes using spatial autocorrelation methods.
- Kin-structured maintenance of variation and migration.

**Quantitative Genetics**

- Phenotype as the product of the interaction between genotype and environment.
- Additive, dominance and epistatic variance on the phenotype.

**Genetic Epidemiology**

- Determination of number of genes governing a character.
- Determination of quantitative trait loci (QTLs) using family information or population information, and using linkage and/or association studies.
- Determination of quantitative trait nucleotide (QTN) using a combination linkage disequilibrium methods and cladistic approaches.
- Determination of individual causal nucleotide in the diploid or haploid state on the phenotype using the method of measured genotype approaches, and combined effects or synergistic interaction of the causal mutations on the phenotype.
- Determination of relative importance of each of the mutations on a given phenotype using multivariate methods, such as discriminant function, principal component and step-wise regression methods.
- Determination of direct and indirect effect of polymorphisms on a complex phenotype using path analysis (partial regression) methods.
- Determination of the effects of specific environment on a given genotype-genotype-environment interactions using joint regression and additive and multiplicative parameter methods.


**Guidance in the selection of appropriate genetic statistical tests for analysis of data can be obtained from texts such as:** *Fundamentals of Genetic Epidemiology* (Monographs in Epidemiology and Biostatistics, Vol 22) by M. J. Khoury, B. H. Cohen & T. H. Beaty, Oxford Univ Press, 1993; *Methods in Genetic Epidemiology* by Newton.
A. Stratify patients by genotype at one candidate variance in the candidate gene locus.

1. Genetic stratification of patients can be accomplished in several ways, including the following (where 'a' is the more frequent form of the variance being assessed and 'a' is the less frequent form):

(a) AA vs. aa

(b) AA vs. Aa vs. aa

(c) AA vs. (Aa+aa)

(d) (AA+Aa) vs. aa.

2. The effect of genotype on drug response phenotype may be affected by a variety of nongenetic factors. Therefore, it may be beneficial to measure the effect of genetic stratification in a subgroup of the overall clinical trial population. Subgroups can be defined in a number of ways including, for example, biological, clinical, pathological or environmental criteria. For example, the predictive value of genetic stratification can be assessed in a subgroup or subgroups defined by:

a. Biological Criteria:

i. Gender (males vs. females)

ii. Age (for example above 60 years of age). Two, three or more age groups may be useful for defining subgroups for the genetic analysis.

iii. Hormonal status and reproductive history, including pre- vs. post-menopausal status of women, or multiparous vs. nulliparous women

iv. Ethnic, racial or geographic origin, or surrogate markers of ethnic, racial or geographic origin. (For a description of genetic markers that serve as surrogates of racial/ethnic origin see, for example: Rannala, B. and J. L. Mountain. Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA, 94 (17): 9197-9201, 1997. Other surrogate markers could be used, including biochemical markers.)

b. Clinical Criteria:

i. Disease status. There are clinical grading scales for many diseases. For example, the status of Alzheimer’s Disease patients is often measured by cognitive assessment scales such as the mini-mental status exam (MMSE) or the Alzheimer’s Disease Assessment Scale (ADAS), which includes a cognitive component (ADAS-COG). There are also clinical assessment scales for many other diseases, including cancer.

ii. Disease manifestations (clinical presentation).

iii. Radiological staging criteria.

iv. Histopathologic features of disease tissue, or pathological diagnosis. (For example, there are many varieties of lung cancer: squamous cell carcinoma, adenocarcinoma, small cell carcinoma, bronchoalveolar carcinoma, etc., each of which may—which, in combination with genetic variation, may correlate with

ii. Pathological stage. A variety of diseases, particularly cancer, have pathological staging schemes.

iii. Loss of heterozygosity (LOH)

iv. Pathology studies such as measuring levels of a marker protein

v. Laboratory studies such as hormone levels, protein levels, small molecule levels

vi. Measure frequency of responders in each genetic subgroup. Subgroups may be defined in several ways.

2. Stratify by haplotype at one candidate locus where the haplotype is made up of two variances, three variances or greater than three variances.

Data from already completed clinical trials can be retrospectively reanalyzed. Since the questions are new, the data can be treated as if it were a prospective trial, with identified variances or haplotypes as stratification criteria or endpoints in clinically stratified data (e.g. what is the frequency of a particular variance in a response group compared to nonresponders). Care should be taken to in studying a population in which there may be a link between drug-related genes and disease-related genes.

Retroactive pharmacogenetic trials can be conducted at each of the phases of clinical development, if sufficient data is available to correlate the physiologic effect of the candidate therapeutic intervention and the allelic variance or variances within the treatment population. In the case of a retrospective trial, the data collected from the trial can be re-analyzed by imposing the additional stratification on groups of patients by specific allelic variances that may exist in the treatment groups. Retroactive trials can be useful to ascertain whether a hypothesis that a specific variance has a significant effect on the efficacy or toxicity profile for a candidate therapeutic intervention.
A prospective clinical trial has the advantage that the trial can be designed to ensure the trial objectives can be met with statistical certainty. In these cases, power analysis, which includes the parameters of allelic variance frequency, number of treatment groups, and ability to detect positive outcomes can ensure that the trial objectives are met.

In designing a pharmacogenetic trial, retrospective analysis of Phase II or Phase III clinical data can indicate trial variables for which further analysis is beneficial. For example, surrogate endpoints, pharmacokinetic parameters, dosage, efficacy endpoints, ethnic and gender differences, and toxicological parameters may result in data that would require further analysis and re-examination through the design of an additional trial. In these cases, analysis involving statistics, genetics, clinical outcomes, and economic parameters may be considered prior to proceeding to the stage of designing any additional trials. Factors involved in the consideration of statistical significance may include Bonferroni analysis, permutation testing, with multiple testing correction resulting in a difference among the treatment groups that has occurred as a result of a chance of no greater than 20%, i.e. p<0.20. Factors included in determining clinical outcomes to be relevant for additional testing may include, for example, consideration of the target indication, the trial endpoints, progression of the disease, disorder, or condition during the trial study period, biochemical or pathophysiologic relevance of the candidate therapeutic intervention, and other variables that were not included or anticipated in the initial study design or clinical protocol.

V. Variance Identification and Use
A. Initial Identification of Variance in Genes
Selection of Population Size and Composition
Prior to testing to identify the presence of sequence variances in a particular gene or genes, it is useful to understand how many individuals should be screened to provide confidence that most or nearly all pharmacogenetically relevant variances will be found. The answer depends on the frequencies of the phenotypes of interest and what assumptions we make about heterogeneity and magnitude of genetic effects. Prior to testing to identify the presence of sequence variances in a particular gene or genes, it is useful to understand how many individuals should be screened to provide confidence that most or nearly all pharmacogenetically relevant variances will be found. The answer depends on the frequencies of the phenotypes of interest and what assumptions we make about heterogeneity and magnitude of genetic effects. At the beginning we only know phenotype frequencies (e.g. responders vs. non-responders, frequency of various side effects, etc.).

The most conservative assumption (resulting in the lowest estimate of allele frequency, and consequently the largest suggested screening population) is (i) that the phenotype (e.g. toxicity or efficacy) is multifactorial (i.e. can be caused by two or more variances or combinations of variances), (ii) that the variance of interest has a high degree of penetrance (i.e. is consistently associated with the phenotype), and (iii) that the mode of transmission is Mendelian dominant. Consider a pharmacogenetic study designed to identify predictors of efficacy for a compound that produces a 15% response rate in a nonstratified population. If half the response is substantially attributable to a given variance, and the variance is consistently associated with a positive response (in 80% of cases) and the variance need only be present in one copy to produce a positive result then ~10%
of the subjects are likely heterozygotes for the variance that produces the response. The Hardy-Weinberg equation can be used to infer an allele frequency in the range of 3% from these assumptions (given allele frequencies of 5%/95% then: 2×0.05×0.95=0.095, or 9.5% heterozygotes are expected, and 0.05×0.05=0.0025, or 0.25% homozygotes are expected. They sum to 95%/0.25%=9.75% likely responders, 97.5% of whom, or 9.75%, are likely real responders due to presence of the positive response allele. Thus about half of the 15% responders are accounted for). From the Table it can be seen that, in order to have a 99% chance of detecting an allele present at a frequency of 5% nearly 50 subjects should be screened for variances, assuming that the variances occur in the screening population at the same frequency as they occur in the patient population. Similar analyses can be performed for other assumptions regarding likely magnitude of effect, penetrance and mode of genetic transmission.

[0575] At the beginning we only know phenotype frequencies (e.g. responders vs. nonresponders, frequency of various side effects, etc.). As an example, the occurrence of serious 5-FU/FA toxicity—e.g. toxicity requiring hospitalization is often >10%. The occurrence of life threatening toxicity is in the 1-3% range (Buroker et al. 1994). The occurrence of complete remissions is on the order of 2-8%. The lowest frequency phenotypes are thus on the order of ~2%. If we assume that (i) homogeneous genetic effects are responsible for half the phenotypes of interest and (ii) for the most part the extreme phenotypes represent recessive genotypes, then we need to detect alleles that will be present at ~10% frequency (0.1×0.1=0.01, or 1% frequency of homozygotes) if the population is at Hardy-Weinberg equilibrium. To have a ~99% chance of identifying such alleles would require searching a population of 22 individuals (see Table below). If the major phenotypes are associated with heterozygous genotypes then we need to detect alleles present at ~0.5% frequency (2×0.005×0.995=0.00995, or ~1% frequency of heterozygotes). A 99% chance of detecting such alleles would require ~40 individuals (Table below). Given the heterogeneity of the North American population we cannot assume that all genotypes are present in Hardy-Weinberg proportions, therefore a substantial oversampling may be done to increase the chances of detecting relevant variances: For our initial screening, usually, 62 individuals of unknown race/ethnicity are screened for variance. Variance detection studies can be extended to outliers for the phenotypes of interest to cover the possibility that important variances were missed in the normal population screening.

Likelihood of Detecting Polymorphism in a Population as a Function of Allele Frequency & Number of Individuals Genotyped

[0576] Table 4 above shows the probability (expressed as percent) of detecting both alleles (i.e. detecting heterozygotes) at a biallelic locus as a function of (i) the allele frequencies and (ii) the number of individuals genotyped. The chances of detecting heterozygotes increases as the frequencies of the two alleles approach 0.5 (down a column), and as the number of individuals genotyped increases (to the right along a row). The numbers in the table are given by the formula: 1-(p)²-(q)². Allele frequencies are designated p and q and the number of individuals tested is designated n. (Since humans are diploid, the number of alleles tested is twice the number of individuals, or 2n.)

[0577] While it is preferable that numbers of individuals, or independent sequence samples, are screened to identify variances in a gene, it is also very beneficial to identify variances using smaller numbers of individuals or sequence samples. For example, even a comparison between the sequences of two samples or individuals can reveal sequence variances between them. Preferably, 5, 10, or more samples or individuals are screened.

Source of Nucleic Acid Samples

[0578] Nucleic acid samples, for example for use in variance identification, can be obtained from a variety of sources as known to those skilled in the art, or can be obtained from genomic or cDNA sources by known methods. For example, the Coriell Cell Repository (Camden, N.J.) maintains over 6,000 human cell cultures, mostly fibroblast and lymphoblast cell lines comprising the NIGMS Human Genetic Mutant Cell Repository. A catalog (http://locus.umdnj.edu/nigns) provides racial or ethnic identifiers for many of the cell lines. It is preferable to perform polymorphism discovery on a population that mimics the population to be evaluated in a clinical trial, both in terms of racial/ethnic/geographic background and in terms of disease status. Otherwise, it is generally preferable to include a broad population sample including, for example, (for trials in the United States): Caucasians of Northern, Central and Southern European origin, Africans or African-Americans, Hispanics or Mexicans, Chinese, Japanese, American Indian, East Indian, Arabs and Koreans.

Source of Human DNA, RNA and cDNA Samples

[0579] PCR based screening for DNA polymorphism can be carried out using either genomic DNA or cDNA produced

<table>
<thead>
<tr>
<th>Allele frequencies</th>
<th>n = 5</th>
<th>n = 10</th>
<th>n = 15</th>
<th>n = 20</th>
<th>n = 25</th>
<th>n = 30</th>
<th>n = 35</th>
<th>n = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>p = .95</td>
<td>9.56</td>
<td>18.21</td>
<td>26.03</td>
<td>33.10</td>
<td>39.50</td>
<td>45.28</td>
<td>50.52</td>
<td>56.30</td>
</tr>
<tr>
<td>p = .97</td>
<td>26.26</td>
<td>45.62</td>
<td>59.90</td>
<td>70.43</td>
<td>78.19</td>
<td>83.92</td>
<td>88.14</td>
<td>95.24</td>
</tr>
<tr>
<td>p = .95</td>
<td>40.13</td>
<td>64.15</td>
<td>78.53</td>
<td>87.15</td>
<td>92.30</td>
<td>95.39</td>
<td>97.24</td>
<td>99.65</td>
</tr>
<tr>
<td>p = .93</td>
<td>51.60</td>
<td>76.58</td>
<td>88.66</td>
<td>94.51</td>
<td>97.34</td>
<td>98.71</td>
<td>99.38</td>
<td>99.93</td>
</tr>
<tr>
<td>p = .9, q =</td>
<td>65.13</td>
<td>87.84</td>
<td>95.76</td>
<td>98.52</td>
<td>99.48</td>
<td>99.82</td>
<td>99.94</td>
<td>99.9</td>
</tr>
</tbody>
</table>

TABLE 4
from mRNA. For many genes, only cDNA sequences have been published, therefore the analysis of those genes is, at least initially, at the cDNA level since the determination of intron-exon boundaries and the isolation of flanking sequences is a laborious process. However, screening genomic DNA has the advantage that variances can be identified in promoter, intron and flanking regions. Such variances may be biologically relevant. Therefore preferably, when variance analysis of patients with outlier responses is performed, analysis of selected loci at the genomic level is also performed. Such analysis would be contingent on the availability of a genomic sequence or intron-exon boundary sequences, and would also depend on the anticipated biological importance of the gene in connection with the particular response.

**[0580]** When cDNA is to be analyzed it is very beneficial to establish a tissue source in which the genes of interest are expressed at sufficient levels that cDNA can be readily produced by RT-PCR. Preliminary PCR optimization efforts for 19 of the 29 genes in Table 2 reveal that all 19 can be amplified from lymphoblastoid cell mRNA. The 7 untested genes belong on the same pathways and are expected to also be PCR amplifiable.

**PCR Optimization**

**[0581]** Primers for amplifying a particular sequence can be designed by methods known to those skilled in the art, including by the use of computer programs such as the PRIMER software available from Whitehead Institute/MIT Genome Center. In some cases it is preferable to optimize the amplification process according to parameters and methods known to those skilled in the art; optimization of PCR reactions based on a limited array of temperature, buffer and primer concentration conditions is utilized. New primers are obtained if optimization fails with a particular primer set.

**Variance Detection Using T4 Endonuclease VII Mismatch Cleavage Method**

**[0582]** Any of a variety of different methods for detecting variances in a particular gene can be utilized, such as those described in the patents and applications cited in section A above. An exemplary method is a T4 Endo VII method. The enzyme T4 endonuclease VII (T4E7) is derived from the bacteriophage T4. T4E7 specifically cleaves heteroduplex DNA containing single base mismatches, deletions or insertions. The site of cleavage is 1 to 6 nucleotides 3' of the mismatch. This activity has been exploited to develop a general method for detecting DNA sequence variances (Youil et al., 1995; Mashal and Sklar, 1995). A quality controlled T4E7 variance detection procedure based on the T4E7 patent of R. G. H. Cotton and co-workers. (Del Tito et al., in press) is preferably utilized. T4E7 has the advantages of being rapid, inexpensive, sensitive and selective. Further, since the enzyme pinpoints the site of sequence variance, sequencing effort can be confined to a 25-30 nucleotide segment.

**[0583]** The major steps in identifying sequence variations in candidate genes using T4E7 are: (1) PCR amplify 400-600 bp segments from a panel of DNA samples; (2) mix a fluorescently-labeled probe DNA with the sample DNA; (3) heat and cool the samples to allow the formation of heteroduplexes; (4) add T4E7 enzyme to the samples and incubate for 30 minutes at 37°C, during which cleavage occurs at sequence variance mismatches; (5) run the samples on an ABI 377 sequencing apparatus to identify cleavage bands, which indicate the presence and location of variances in the sequence; (6) a subset of PCR fragments showing cleavage are sequenced to identify the exact location and identity of each variance.

**[0584]** The T4E7 Variance Imaging procedure has been used to screen particular genes. The efficiency of the T4E7 enzyme to recognize and cleave at all mismatches has been tested and reported in the literature. One group reported detection of 81 of 81 known mutations (Youil et al. 1995) while another group reported detection of 16 of 17 known mutations (Mashal and Sklar, 1995). Thus, the T4E7 method provides highly efficient variance detection.

**DNA Sequencing**

**[0585]** A subset of the samples containing each unique T4E7 cleavage site is selected for sequencing. DNA sequencing can, for example, be performed on ABI 377 automated DNA sequencers using BigDye chemistry and cycle sequencing. Analysis of the sequencing runs will be limited to the 30-40 bases pinpointed by the T4E7 procedure as containing the variance. This provides the rapid identification of the altered base or bases.

**[0586]** In some cases, the presence of variances can be inferred from published articles which describe Restriction Fragment Length Polymorphisms (RFLP). The sequence variances or polymorphisms creating these RFLPs can be readily determined using convention techniques, for example in the following manner. If the RFLP was initially discovered by the hybridization of a cDNA, then the molecular sequence of the RFLP can be determined by restricting the cDNA probe into fragments and separately hybridizing to a Southern blot consisting of the restriction digestion with the enzyme which reveals the polymorphic site, identifying the subfragment which hybridizes to the polymorphic restriction fragment, obtaining a genomic clone of the gene (e.g., from commercial services such as Genome Systems (Saint Louis, Mo.) or Research Genetics (Alabama) which will provide appropriate genomic clones on receipt of appropriate primer pairs). Using the genomic clone, restrict the genomic clone with the restriction enzyme which revealed the polymorphism and isolate the fragment which contains the polymorphism, e.g., identifying by hybridization to the cDNA which detected the polymorphism. The fragment is then sequenced across the polymorphic site. A copy of the other allele can be obtained by PCT from addition samples.

**Variance Detection Using Sequence Scanning**

**[0587]** In addition to the physical methods, e.g., those described above and others known to those skilled in the art (see, e.g., Housman, U.S. Pat. No. 5,702,890; Housman et al., U.S. patent application Ser. No. 09/045,053), variances can be detected using computational methods, involving computer comparison of sequences from two or more different biological sources, which can be obtained in various ways, for example from public sequence databases. The term “variance scanning” refers to a process of identifying sequence variances using computer-based comparison and analysis of multiple representations of at least a portion of one or more genes. Computational variance detection involves a process to distinguish true variances from
sequencing errors or other artifacts, and thus does not require perfectly accurate sequences. Such scanning can be performed in a variety of ways, preferably, for example, as described in Stanton et al., filed Oct. 14, 1999, Ser. No. 09/419,705, attorney docket number 246/128.

While the utilization of complete cDNA sequences is highly preferred, it is also possible to utilize genomic sequences. Such analysis may be desired where the detection of variances in or near splice sites is sought. Such sequences may represent full or partial genomic DNA sequences for a gene or genes. Also, as previously indicated, partial cDNA sequences can also be utilized although this is less preferred. As described below, the variance scanning analysis can simply utilize sequence overlap regions, even from partial sequences. Also, while the present description is provided by reference to DNA, e.g., cDNA, some sequences may be provided as RNA sequences, e.g., mRNA sequences. Such RNA sequences may be converted to the corresponding DNA sequences, or the analysis may use the RNA sequences directly.

B. Determination of Presence or Absence of Known Variances

The identification of the presence of previously identified variances in cells of an individual, usually a particular patient, can be performed by a number of different techniques as indicated in the Summary above. Such methods include methods utilizing a probe which specifically recognizes the presence of a particular nucleic acid or amino acid sequence in a sample. Common types of probes include nucleic acid hybridization probe fragments and antibodies, for example, monoclonal antibodies, which can differentially bind to nucleic acid sequences differing in one or more variance sites or to polypeptides which differ in one or more amino acid residues as a result of the nucleic acid sequence variance or variances. Generation and use of such probes is well-known in the art and so is not described in detail herein.

Preferably, however, the presence or absence of a variance is determined using nucleotide sequencing of a short sequence spanning a previously identified variance site. This will utilize validated genotyping assays for the polymorphisms previously identified. Since both normal and tumor cell genotypes can be measured, and since tumor material will frequently only be available as paraffin embedded sections (from which RNA cannot be isolated), it will be necessary to utilize genotyping assays that will work on genomic DNA. Thus PCR reactions will be designed, optimized, and validated to accommodate the intron-exon structure of each of the genes. If the gene structure has been published (as it has for some of the listed genes), PCR primers can be designed directly. However, if the gene structure is unknown, the PCR primers may need to be moved around in order to both span the variance and avoid exon-intron boundaries. In some cases one-sided PCR methods such as bubble PCR (Ausubel et al. 1997) may be useful to obtain flanking intronic DNA for sequence analysis.

Using such amplification procedures, the standard method used to genotype normal and tumor tissues will be DNA sequencing. PCR fragments encompassing the variances will be cycle sequenced on ABI 377 automated sequencers using Big Dye chemistry

C. Correlation of the Presence or Absence of Specific Variances with Differential Treatment Response

Prior to establishment of a diagnostic test for use in the selection of a treatment method or elimination of a treatment method, the presence or absence of one or more specific variances in a gene or in multiple genes is correlated with a differential treatment response. (As discussed above, usually the existence of a variable response and the correlation of such a response to a particular gene is performed first.) Such a differential response can be determined using prospective and/or retrospective data. Thus, in some cases, published reports will indicate that the course of treatment will vary depending on the presence or absence of particular variances. That information can be utilized to create a diagnostic test and/or incorporated in a treatment method as an efficacy or safety determination step.

Usually, however, the effect of one or more variances is separately determined. The determination can be performed by analyzing the presence or absence of particular variances in patients who have previously been treated with a particular treatment method, and correlating the variance presence or absence with the observed course, outcome, and/or development of adverse events in those patients. This approach is useful in cases in which observation of treatment effects was clearly recorded and cell samples are available or can be obtained. Alternatively, the analysis can be performed prospectively, where the presence or absence of the variance or variances in an individual is determined and the course, outcome, and/or development of adverse events in those patients is subsequently or concurrently observed and then correlated with the variance determination.

Analysis of Haplotypes Increases Power of Genetic Analysis

In some cases, variation in activity due to a single gene or a single genetic variance in a single gene may not be sufficient to account for a clinically significant fraction of the observed variation in patient response to a treatment, e.g., a drug, there may be other factors that account for some of the variation in patient response. Drug response phenotypes may vary continuously, and such (quantitative) traits may be influenced by a number of genes (Falconer and Mackay, Quantitative Genetics, 1997). Although it is impossible to determine a priori the number of genes influencing a quantitative trait, potentially only one or a few loci have large effects, where a large effect is 5-20% of total variation in the phenotype (Mackay, 1995).

Having identified genetic variation in enzymes that may affect action of a specific drug, it is useful to efficiently address its relation to phenotypic variation. The sequential testing for correlation between phenotypes of interest and single nucleotide polymorphisms may be adequate to detect associations if there are major effects associated with single nucleotide changes; certainly it is useful to this type of analysis. However, there is no way to know in advance whether there are major phenotypic effects associated with single nucleotide changes and, even if there are, there is no way to be sure that the salient variance has been identified by screening cDNAs. A more powerful way to address the question of genotype-phenotype correlation is to assort genotypes into haplotypes. (A haplotype is the cis arrangement of polymorphic nucleotides on a particular chromosome.) Haplotype analysis has several advantages compared
to the serial analysis of individual polymorphisms at a locus with multiple polymorphic sites.

[0596] (1) Of all the possible haplotypes at a locus (2^n haplotypes are theoretically possible at a locus with n binary polymorphic sites) only a small fraction will generally occur at a significant frequency in human populations. Thus, association studies of haplotypes and phenotypes will involve testing fewer hypotheses. As a result there is a smaller probability of Type I errors, that is, false inferences that a particular variant is associated with a given phenotype.

[0597] (2) The biological effect of each variance at a locus may be different both in magnitude and direction. For example, a polymorphism in the 5’ UTR may affect translational efficiency, a coding sequence polymorphism may affect protein activity, a polymorphism in the 3’ UTR may affect mRNA folding and half life, and so on. Further, there may be interactions between variances: two neighboring polymorphic amino acids in the same domain—say cysteine at residue 29 and methyl at residue 166—may, when combined in one sequence, for example, 29cys-166methyl, have a deleterious effect, whereas 29cys-166methyl, 29arg-166methyl and 29arg-166val proteins may be nearly equal in activity. Haplotype analysis is the best method for assessing the interaction of variances at a locus.

[0598] (3) Templeton and colleagues have developed powerful methods for assorting haplotypes and analyzing haplotype/phenotype associations (Templeton et al., 1987). Alleles which share common ancestry are arranged into a tree structure (cladogram) according to their (inferred) time of origin in a population (that is, according to the principle of parsimony). Haplotypes that are evolutionarily ancient will be at the center of the branching structure and new ones (reflecting recent mutations) will be represented at the periphery, with the links representing intermediate steps in evolution. The cladogram defines which haplotype-phenotype association tests should be performed to most efficiently exploit the available degrees of freedom, focusing attention on those comparisons most likely to define functionally different haplotypes (Haviland et al., 1995). This type of analysis has been used to define interactions between heart disease and the apolipoprotein gene cluster (Haviland et al. 1995) and Alzheimer’s Disease and the Apo-E locus (Templeton 1995) among other studies, using populations as small as 50 to 100 individuals. The methods of Templeton have also been applied to measure the genetic determinants of variation in the angiotensin-I converting enzyme gene. (Kearney, B., McKenzie, C. A., Connell, J. M. C., et al. Measured haplotype analysis of the angiotensin-I converting enzyme gene. Human Molecular Genetics 7: 1745-1751.)

Methods for Determining Haplotypes

[0599] The goal of haplotyping is to identify the common haplotypes at selected loci that have multiple sites of variance. Haplotypes are usually determined at the cDNA level. Several general approaches to identification of haplotypes can be employed. Haplotypes may also be estimated using computational methods or determined definitively using experimental approaches. Computational approaches generally include an expectation maximization (E-M) algorithm (see, for example: Excoffier and Slatkin, Mol. Biol. Evol. 1995) or a combination of Parsimony (see below) and E-M methods.

[0600] Haplotypes can be determined experimentally without requirement of a haplotyping method by genotyping samples from a set of pedigrees and observing the segregation of haplotypes. For example families collected by the Centre d’Etude du Polymorphism Humain (CEPH) can be used. Cell lines from these families are available from the Coriell Repository. This approach will be useful for cataloging common haplotypes and for validating methods on samples with known haplotypes. The set of haplotypes determined by pedigree analysis can be useful in computational methods, including those utilizing the E-M algorithm.

[0601] Haplotypes can also be determined directly from cDNA using the T4E7 procedure. T4E7 cleaves mismatched heteroduplex DNA at the site of the mismatch. If a heteroduplex contains only one mismatch, cleavage will result in the generation of two fragments. However, if a single heteroduplex (allele) contains two mismatches, cleavage will occur at two different sites resulting in the generation of three fragments. The appearance of a fragment whose size corresponds to the distance between the two cleavage sites is diagnostic of the two mismatches being present on the same strand (allele). Thus, T4E7 can be used to determine haplotypes in diploid cells.

[0602] An alternative method, allele specific PCR, may be used for haplotyping. The utility of allele specific PCR for haplotyping has already been established (Michalatos-Bejoin et al., 1996; Chang et al. 1997). Opposing PCR primers are designed to cover two sites of variance (either adjacent sites or sites spanning one or more internal variances). Two versions of each primer are synthesized, identical to each other except for the 3’ terminal nucleotide. The 3’ terminal nucleotide is designed so that it will hybridize to one but not the other variant base. PCR amplification is then attempted with all four possible primer combinations in separate wells. Because Taq polymerase is very inefficient at extending 3’ mismatches, the only samples which will be amplified will be the ones in which the two primers are perfectly matched for sequences on the same strand (allele). The presence or absence of PCR product allows haplotyping of diploid cell lines. At most two of four possible reactions should yield products. This procedure has been successfully applied, for example, to haplotype the DPD amino acid polymorphisms.

[0603] Parsimony methods are also useful for classifying DNA sequences, haplotypes or phenotypic characters. Parsimony principle maintains that the best explanation for the observed differences among sequences, phenotypes (individuals, species) etc., is provided by the smallest number of evolutionary changes. Alternatively, simpler hypotheses are preferable to explain a set of data or patterns, than more complicated ones, and ad hoc hypotheses should be avoided whenever possible (Molecular Systematics, Hillis et al., 1996). Parsimony methods thus operate by minimizing the number of evolutionary steps or mutations (changes from one sequence/character) required to account for a given set of data.

[0604] For example, supposing we want to obtain relationships among a set of sequences and construct a structure (tree/topology), we first count the minimum number of mutations that are required for explaining the observed evolutionary changes among a set of sequences. A structure (topology) is constructed based on this number. When once this number is obtained, another structure is tried. This process is continued for all reasonable number of structures. Finally, the structure that required the smallest number of
mutational steps is chosen as the likely structure/evolutionary tree for the sequences studied.

[0605] For haplotypes identified herein, haplotypes were identified by examining genotypes from each cell line. This list of genotypes was optimized to remove variance sites/individuals with incomplete information, and the genotype from each remaining cell line was examined in turn. The number of heterozygotes in the genotype were counted, and those genotypes containing more than one heterozygote were discarded, and the rest were gathered in a list for storage and display.

D. Selection of Treatment Method Using Variance Information

[0606] 1. General

[0607] Once the presence or absence of a variance or variances in a gene or genes is shown to correlate with the efficacy or safety of a treatment method, that information can be used to select an appropriate treatment method for a particular patient. In the case of a treatment which is more likely to be effective when administered to a patient who has at least one copy of a gene with a particular variance or variances (in some cases the correlation with effective treatment is for patients who are homozygous for a variance or set of variances in a gene) than in patients with a different variance or set of variances, a method of treatment is selected (and/or a method of administration) which correlates positively with the particular variance presence or absence which provides the indication of effectiveness. As indicated in the Summary, such selection can involve a variety of different choices, and the correlation can involve a variety of different types of treatments, or choices of methods of treatment. In some cases, the selection may include choices between treatments or methods of administration where more than one method is likely to be effective, or where there is a range of expected effectiveness or different expected levels of contra-indication or deleterious effects. In such cases the selection is preferably performed to select a treatment which will be as effective or more effective than other methods, while having a comparatively low level of deleterious effects. Similarly, where the selection is between method with differing levels of deleterious effects, preferably a method is selected which has low such effects but which is expected to be effective in the patient.

[0608] Alternatively, in cases where the presence or absence of the particular variance or variances is indicative that a treatment or method of administration is more likely to be ineffective or contra-indicated in a patient with that variance or variances, then such treatment or method of administration is generally eliminated for use in that patient.

2. Diagnostic Methods

[0609] Once a correlation between the presence and absence of at least one variance in a gene or genes and an indication of the effectiveness of a treatment, the determination of the presence or absence of that at least one variance provides diagnostic methods, which can be used as indicated in the Summary above to select methods of treatment, methods of administration of a treatment, methods of selecting a patient or patients for a treatment and others aspects in which the determination of the presence or absence of those variances provides useful information for selecting or designing or preparing methods or materials for medical use in the aspects of this invention. As previously stated, such variance determination or diagnostic methods can be performed in various ways as understood by those skilled in the art.

[0610] In certain variance determination methods, it is necessary or advantageous to amplify one or more nucleotide sequences in one or more of the genes identified herein. Such amplification can be performed by conventional methods, e.g., using polymerase chain reaction (PCR) amplification. Such amplification methods are well-known to those skilled in the art and will not be specifically described herein. For most applications relevant to the present invention, a sequence to be amplified includes at least one variance site, which is preferably a site or sites which provide variance information indicative of the effectiveness of a method of treatment or method of administration of a treatment, or effectiveness of a second method of treatment which reduces a deleterious effect of a first treatment method, or which enhances the effectiveness of a first method of treatment. Thus, for PCR, such amplification generally utilizes primer oligonucleotides which bind to or extend through at least one such variance site under amplification conditions.

[0611] For convenient use of the amplified sequence, e.g., for sequencing, it is beneficial that the amplified sequence be of limited length, but still long enough to allow convenient and specific amplification. Thus, preferably the amplified sequence has a length as described in the Summary.

[0612] Also, in certain variance determination, it is useful to sequence one or more portions of a gene or genes, in particular, portions of the genes identified in this disclosure. As understood by persons familiar with nucleic acid sequencing, there are a variety of effective methods. In particular, sequencing can utilize dye terminination methods and mass spectrometric methods. The sequencing generally involves a nucleic acid sequence which includes a variance site as indicated above in connection with amplification. Such sequencing can directly provide determination of the presence or absence of a particular variance or set of variances, e.g., a haplotype, by inspection of the sequence (visually or by computer). Such sequencing is generally conducted on PCR amplified sequences in order to provide sufficient signal for practial or reliable sequence determination.

[0613] Likewise, in certain variance determinations, it is useful to utilize a probe or probes. As previously described, such probes can be of a variety of different types.

VI. Pharmaceutical Compositions, Including Pharmaceutical Compositions Adapted to be Preferentially Effective in Patients Having Particular Genetic Characteristics

A. General

[0614] The methods of the present invention, in many cases will utilize conventional pharmaceutical compositions, but will allow more advantageous and beneficial use of those compositions due to the ability to identify patients who are likely to benefit from a particular treatment or to identify patients for whom a particular treatment is less likely to be effective or for whom a particular treatment is likely to produce undesirable or intolerable effects. However, in some cases, it is advantageous to utilize compositions which are adapted to be preferentially effective in...
patients who possess particular genetic characteristics, i.e., in whom a particular variance or variances in one or more genes is present or absent (depending on whether the presence or the absence of the variance or variances in a patient is correlated with an increased expectation of beneficial response). Thus, for example, the presence of a particular variance or variances may indicate that a patient can beneficially receive a significantly higher dosage of a drug than a patient having a different.

B. Regulatory Indications and Restrictions

[0615] The sale and use of drugs and the use of other treatment methods usually are subject to certain restrictions by a government regulatory agency charged with ensuring the safety and efficacy of drugs and treatment methods for medical use, and approval is based on particular indications. In the present invention it is found that variability in patient response or patient tolerance of a drug or other treatment often correlates with the presence or absence of particular variances in particular genes. Thus, it is expected that such a regulatory agency may indicate that the approved indications for use of a drug with a variance-related variable response or toleration include use only in patients in whom the drug will be effective, and/or for whom the administration of the drug will not have intolerable deleterious effects, such as excessive toxicity or unacceptable side-effects. Conversely, the drug may be given for an indication that it may be used in the treatment of a particular disease or condition where the patient has at least one copy of a particular variance, variances, or variant form of a gene. Even if the approved indications are not narrowed to such groups, the regulatory agency may suggest use limited to particular groups or excluding particular groups or may state advantages of use or exclusion of such groups or may state a warning on the use of the drug in certain groups. Consistent with such suggestions and indications, such an agency may suggest or recommend the use of a diagnostic test to identify the presence or absence of the relevant variances in the prospective patient. Such diagnostic methods are described in this description. Generally, such regulatory suggestion or indication is provided in a product insert or label, and is generally reproduced in references such as the Physician’s Desk Reference (PDR). Thus, this invention also includes drugs or pharmaceutical compositions which carry such a suggestion or statement of indication or warning or suggestion for a diagnostic test, and which may also be packaged with an insert or label stating the suggestion or indication or warning or suggestion for a diagnostic test.

[0616] In accord with the possible variable treatment responses, an indication or suggestion can specify that a patient be heterozygous, or alternatively, homozygous for a particular variance or variances or variant form of a gene. Alternatively, an indication or suggestion may specify that a patient have no more than one copy, or zero copies, of a particular variance, variances, or variant form of a gene.

[0617] A regulatory indication or suggestion may concern the variances or variant forms of a gene in normal cells of a patient and/or in cells involved in the disease or condition. For example, in the case of a cancer treatment, the response of the cancer cells can depend on the form of a gene remaining in cancer cells following loss of heterozygosity affecting that gene. Thus, even though normal cells of the patient may contain a form of the gene which correlates with effective treatment response, the absence of that form in cancer cells will mean that the treatment would be less likely to be effective in that patient than in another patient who retained in cancer cells the form of the gene which correlated with effective treatment response. Those skilled in the art will understand whether the variances or gene forms in normal or disease cells are most indicative of the expected treatment response, and will generally utilize a diagnostic test with respect to the appropriate cells. Such a cell type indication or suggestion may also be contained in a regulatory statement, e.g., on a label or in a product insert.

C. Preparation and Administration of Drugs and Pharmaceutical Compositions Including Pharmaceutical Compositions Adapted to be Preferentially Effective in Patients Having Particular Genetic Characteristics

[0618] A particular compound useful in this invention can be administered to a patient either by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s). In treating a patient exhibiting a disorder of interest, a therapeutically effective amount of a agent or agents such as these is administered. A therapeutically effective dose refers to that amount of the compound that results in amelioration of one or more symptoms or a prolongation of survival in a patient.

[0619] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized.

[0620] For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by HPLC.

[0621] The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient’s condition. (See e.g. Fingl et al., in The Pharmacological Basis of Therapeutics, 1975, Ch. 1 p. 1). It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administrated dose in the management of disorder of interest will vary with the severity of the condition to be treated and the route of administration. The severity of the condition may, for example, be evaluated, in part, by standard prognostic
evaluation methods. Further, the dose and perhaps dose frequency, will also vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

[0622] Depending on the specific conditions being treated, such agents may be formulated and administered systemically or locally. Techniques for formulation and administration may be found in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, Pa. (1990). Suitable routes may include oral, rectal, transdermal, vaginal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrahepatic, direct intraventricular, intravenous, intraperitoneal, intranasal, or intracocular injections, just to name a few.

[0623] For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hank's solution, Ringer's solution, or physiological saline buffer. For such transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

[0624] Use of pharmaceutically acceptable carriers to formulate the compounds herein disclosed for the practice of the invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions of the present invention, in particular, those formulated as solutions, may be administered parenterally, such as by intravenous injection. The compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated.

[0625] Agents intended to be administered intracellularly may be administered using techniques well known to those of ordinary skill in the art. For example, such agents may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the liposome interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are efficiently delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, small organic molecules may be directly administered intracellularly.

[0626] Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. The preparations formulated for oral administration may be in the form of tablets, dragees, capsules, or solutions. The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levitating, emulsifying, encapsulating, entrapping or lyophilizing processes.

[0627] Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Alternatively, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

[0628] Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethyl-cellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carboxypolypegylglycol, and/or titanium dioxide, lacteal solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

[0629] Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.

[0630] The invention described herein features methods for determining the appropriate identification of a patient diagnosed with an inflammatory disease or dysfunction or immunologic disorder based on an analysis of the patient's allele status for a gene listed in Tables 1 and 3. Specifically, the presence of at least one allele indicates that a patient will respond to a candidate therapeutic intervention aimed at treating an inflammatory or immune system clinical symptoms. In a preferred approach, the patient's allele status is rapidly diagnosed using a sensitive PCR assay and a treat-
The findings described herein indicate the predictive value of the target allele in identifying patients at risk for inflammatory or immune disease or dysfunction. In addition, because the underlying mechanism influenced by the allele status is not disease-specific, the allele status is suitable for making patient predictions for diseases not affected by the pathway as well.

The following examples, which describe exemplary techniques and experimental results, are provided for the purpose of illustrating the invention, and should not be construed as limiting.

EXAMPLES

Asthma

I. Description of Asthma

Asthma can be an acute or a chronic condition associated with inflammation of the lower airways and variable levels of airflow obstruction. Asthma symptoms vary among individuals and may include wheezing, shortness of breath, tightness of the chest, trouble controlling a cough, persistent cough at night, difficulty breathing during or soon after physical exertion or exercise, or waking up at night due to one of these symptoms. Episodes of these symptoms (referred to as asthma attacks, flare-ups, or exacerbations) occur when there is sufficiently severe airway constriction to render a patient almost unable or unable to breathe. Although, there can be warning signs, many attacks are sudden and unanticipated.

Individuals with asthma have inflamed airways that are supersensitive to inducers of asthma which exacerbate asthma and enhance underlying inflammation such as allergens, respiratory infections, or industrial pollutants. Provokers of asthma leading to bronchospasm include exercise or physical activities, irritants, emotions and aspirin. Asthma attacks are associated with swollen and inflamed linings of the airways, excess mucus in the airways, and bronchospasm which are reversible. In chronic asthma, there is persistent activation of resident cells (e.g. basophils, eosinophils, neutrophils) lining the airway leading to chronic inflammation which can result in irreversible changes in the airway passages. These permanent changes are part of a remodeling process.

Recent evidence has suggested that airway inflammation is a major factor in the pathogenesis and in the severity of the disease. One theory holds that asthma is a T helper 2 (Th2) cell-driven chronic eosinophilia mediated via dendritic and other antigen-presenting cells. The inflammatory nature of the disease is multicellular in nature, with mast cells, eosinophils, macrophages, basophils, lymphocytes, neutrophils, and epithelial cells participating and therefore immunoglobulins, cytokines, chemokines, adhesion molecules, proteases, inflammatory mediators, and growth factors are involved in various stages and interact to maintain and amplify the inflammatory response. The net result of these interactions is persistent inflammation and repair, ultimately leading to irreversible airway remodeling.

II. Current Therapies for Asthma

Because asthma results from a complex combination of mediators of inflammation, most useful anti-asthma agents affect pathways for these mediators. In acute or chronic asthma, the therapeutic categories include: immunosuppressive agents including glucocorticoids, antiinflammatory agents including leukotriene receptor agonists and mast cell stabilizers (cromolyn sulfate); bronchodilators including β-adrenergic agonists, sympathomimetic agents, and xanthines; and agents to treat cough and excess mucus including expectorants and mucolytics.

Corticosteroids affect the inflammation within the airways by decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

β-Adrenergic agonists and sympathomimetics affect the pulmonary airway lining in a well-characterized mechanism of β-adrenergic receptor activation of adenyl cyclase as well as cAMP independent mechanisms. Bronchodilation is the immediate clinical effect.

Leukotriene modifiers affect the airway by inhibition of 5-lipooxygenase, the initial enzyme of leukotriene biosynthesis, and exert their effect by decreasing leukotriene production, thereby interfering with eosinophil migration and other processes.

Corticosteroids affect the inflammation within the airways: by interfering with growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

Corticosteroids in combination with long-acting β-adrenergic agonists work well as combination therapy.

Cromones are believed to act on the airway by modifying mediator release, and inhibiting mast cell degranulation.

Xanthines are believed to act on the airway in asthma by inhibiting eosinophil cell migration, and enhancing β-adrenergic pathway mediated bronchodilation via the inhibition of phosphodiesterase.

Difficult to treat or therapy-resistant asthma syndromes present a challenge to clinicians. They include difficult acute and chronic, as well as chronic severe, acute severe, therapy-resistant, difficult to control and corticosteroid-dependent asthma.
III. Limitations of Current Therapies for Asthma

Limitations Involving Efficacy

[0645] The therapies discussed above do not reverse the underlying pathological process in asthma: they merely slow or retard the progression of asthma. As thickening of the airways occurs and becomes irreversible, the therapeutic options become limited. Thus, therapies for asthma are aimed at reduction of inflammatory processes and control of symptoms starting at the earliest date (frequently in the pediatric setting).

[0646] The limitations of the adrenergic agonist compounds used for the treatment of asthma include short duration of action and ligand desensitization. Excessive use of short acting \( \beta \)-adrenergic agonists has been proposed to lead to loss of asthma control and consequent increases in morbidity and mortality. Long acting bronchodilators/bronchoprotective agonists acting at adrenergic receptors have supplanted short duration \( \beta \)-agonists.

[0647] Short-acting \( \beta \)-adrenergic agonists are primarily used for the relief of acute asthma symptoms. Excessive reliance on these agents is generally not advised because 1) \( \beta \)-adrenergic receptors undergo a rapid desensitization and the agonist becomes an ineffective bronchodilator, and 2) repetitive high doses of short acting \( \beta \)-adrenergic agonists may be detrimental to the control of asthma by potentially interfering with corticosteroid action. This desensitization occurs through a process involving G-protein receptor coupled-kinases and or cAMP dependent protein kinase or by enhanced degradation of cAMP by phosphodiesterase activity.

[0648] Glucocorticoid associated side effects include: increased appetite, weight gain, fluid retention, acne, ecchymosis, development of Cushingoid facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, septic necrosis, osteoporosis, ulcers, pancreatitis, psuedotumor cerebri, psychosis, glaucoma, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamus-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

[0649] Theophylline or other phosphodiesterase inhibitors have been shown to have a narrow therapeutic window and can result in life-threatening cardiac arrhythmias.

[0650] Difficult to treat asthma involves a spectrum of disease that responds suboptimally to doses of glucocorticoids. In the face of partial response to inhaled or oral steroids, higher doses are administered risking steroid associated side-effects.

[0651] The reduction of clinical symptoms of asthma following anti-inflammatory therapy may only become evident after several weeks to months of therapy. The slow action of these therapies creates problems for the clinician seeking to expeditiously determine optimal therapy for an individual patient. The development of genetic tests to predict response to different agents will allow selection of optimal therapy with less of the time consuming empirical clinical decision making required presently.

Limitations Involving Toxicity or Undesired Side Effects

[0652] There are toxicities and undesired side effects associated with the above current therapies for asthma that require monitoring. Drugs used to treat asthma may cause death, disability, disease, and place a fetus at risk. The undesired side effects or toxicities are listed for each drug category as described above.

IV. Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions for Asthma

[0653] In a recent report, it was demonstrated that the 5-lipoxygenase (5-LO) gene promoter variation among asthma patients is linked to drug response to 5-LO inhibitors (Draken et al., Nature Genetics 22: 1999). In a clinical trial to test efficacy of a potent, selective 5-LO inhibitor (ABT-761), the trial was abruptly closed due to inordinate event rate of abnormal liver function tests. Although the projected enrollment was not reached, the interim data suggested superior efficacy regarding forced expiratory volume in the high dose relative to low dose or placebo groups. The investigators chose to stratify the high dose and placebo group of the enrolled patients based on genotype of the 5-LO gene promoter. The 5-LO gene promoter has been found to contain 3-6 tandem repeats of the Sp1-binding motif. The wild-type allele was designated as 5 tandem repeats and had a frequency of 0.772 in the study population. The forced expiratory volume data indicated that heterozygous patients on high-dose active treatment had, on average, an improvement of forced expiratory volume within one week (23.3 ± 6.0%) and was similar to the wild-type patients (18.8 ± 3.6%). In contrast, the patients with mutant genotype had no benefit from active 5-LO inhibitor treatment (−1.2 ± 2.9%). In Table 12, the trial outcome data is described for two periods following treatment with high dose or placebo.

<table>
<thead>
<tr>
<th>TABLE 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Group</td>
</tr>
<tr>
<td>Wild type, high dose</td>
</tr>
<tr>
<td>Mutant, high dose</td>
</tr>
<tr>
<td>Placebo</td>
</tr>
</tbody>
</table>

*Data extrapolated from published data

[0654] Approximately 6% of asthma patients do not carry a wild-type allele at the 5-LO core promoter locus, and this data indicates that these patients would not benefit from 5-LO inhibitor drug therapy. Further, these data indicate that there is evidence to reasonably identify patients, i.e. stratification based upon 5-LO genotype, to appropriately treat patients with asthma.

[0655] A recent double blind, placebo controlled crossover designed pharmacogenetic retrospective clinical trial of a \( \beta \)-adrenoreceptor polymorphism was implemented to analyze the significance of \( \beta \)-adrenoreceptor polymorphisms (Tan et al. Lancet 350:995-999). In vitro studies have suggested that polymorphism of the \( \beta \)-adrenergic receptor may influence the desensitization induced by \( \beta \) agonists. Twenty-two moderately severe asthmatics were enrolled into a placebo-controlled, cross-over study of formoterol (a \( \beta \)-adrenergic agonist). The patients were divided into
groups by allelic variances: 1) at codon 16, homozygous arginine (n=4), heterozygous arginine/glycine (n=8), and homozygous glycine (n=10); and 2) at codon 27, homozygous glutamine (n=5), heterozygous glutamine/glutamic acid (n=11), and homozygous glutamic acid (n=6). Genotypic analysis determined that individuals who were homozygous for glycine at codon 16 were also homozygous for glutamic acid at codon 27. The results were as follows:

| Polymorphisms of the β2-adrenergic receptor | Degree of Bronchodilator Desensitization after Formoterol Therapy 
| | 6 hour | Maximal | 6 Hour | Maximal |
| Gly 16 (n = 10) | 80% | 48% | 103% | 73% |
| Arg 16 (n = 4) | 28% | 8% | 23% | 35% |
| Gly/Gly16 (n = 8) | 57% | 48% | 70% | 50% |
| Glu27 (n = 6) | 73% | 35% | 90% | 68% |
| Gly16/Glu27 (n = 5) | 47% | 3% | 38% | 15% |
| Glu/Glu27 (n = 11) | 65% | 52% | 70% | 45% |

1Data extrapolated from published graphs.
2All individuals homozygous for Glu27 were also homozygous for Gly16.

The homozygous glycine at position 16 was associated with individuals who were prone to bronchodilator desensitization than at arginine at position 16: the mean FEV1 desensitisation was 80% for Gly16 homozygotes versus 28% for the Arg16 homozygotes. Similar results were observed for the 6 hour FEV1, and the FEF.

For the polymorphism at codon 27, the mean for the Glu27 homozygous individuals demonstrated greater desensitization than those who were homozygous for Gly27.

The allelic variance, glycine at position 16 appeared to dominate over the putative protective effects of the mutation of glutamic acid at position 27.

The effects of the codon 16 and 27 polymorphism in the β2-adrenergic receptor on β2-agonist desensitization, as observed in the above data, suggest that there may be an identifiable subset of patients for whom β2-adrenergic receptor desensitization occurs in the presence of long-acting or repeated use of β2-agonists.

Thus, one skilled in the art, will be able to utilize the presently described pharmacogenetic techniques to identify the allelic variances with the coding region of the β3-adrenergic receptor or other receptor proteins that are similar to the 1-adrenergic receptor, including but not limited to those variances for those genes listed in Tables 1 and 3 and those 7-membrane spanning receptor G-protein coupled receptors. In this way, a skilled practitioner will be able to utilize the methods, protocols, and techniques that are described in the detailed description and those known in the art to identify the gene targets, allelic variance or variances, and candidate drugs that affect these pathways. Further, one can design and implement a strategy that incorporates a diagnostic test to genotype the individual for a given allele or alleles or haplotypes, grouping these candidates by genotype, and testing a β2-adrenergic agonist or another candidate therapeutic product for the effect of the pharmacogenomic difference between or among the groups.

As described above, there is evidence to suggest that there are safety response differences to drug therapy in asthma which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmaceutical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Table 1, and pathway matrix Table 2 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents for use in Asthma

The advantages of a clinical research and drug development program that includes the use of polymorphic genotyping for the stratification of patients for the appropriate selection of candidate therapeutic intervention includes 1) identification of patients that may respond earlier and show signs and symptoms of clinical improvement or side effects and toxicities, 2) identification of the primary gene and relevant polymorphic variance that directly affects manifestation of asthma, 3) identification of pathophysiological relevant variance or variances and potential therapies affecting those allelic genotypes or haplotypes, and 4) identification of allelic variances or haplotypes in genes that indirectly affects efficacy, safety or both.

By identifying subsets of asthma subjects based upon genotype that experience different responses to the administration of a drug, agent, or candidate therapeutic intervention, optimal selection may reduce level and extent of asthma damage to affected joints. Appropriate genotyping and correlation to dosing regimen, or selection of optimal therapy would be beneficial to the patient, caregivers, medical personnel, and the patient’s loved ones.

Based upon these advantages, designing and performing a clinical trial, either prospective or retrospective, which includes a genotype stratification arm will incorporate analysis of clinical outcomes and genetic variation associated with those outcomes, and hypothesis testing of the statistically relevant correlation of the genotypic stratification and therapeutic benefits. If statistical relevance is detectable, these studies will be incorporated into regulatory filings. Ultimately, these clinical trial data will be considered during the approval for marketing process, as well as, incorporated into accepted medical management of asthma.

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 2, genes involved in cytokine-mediated immune regulation, non-cytokine mediated immune regulation (including, but not excluded to, cyclophilins, corticosteroids), cell mediated inflammation involving apoptosis, adhesion and migration, protease and protease inhibitors, complement,
degrenulation (platelets, mast cells, neutrophils, eosinophils), release of inflammatory modulators (including membrane lipids, prostaglandin, platelet activating factor, leukotrienes, histamine, nitric oxide), vasculization mediators (including endothelin and vascular endothelial cell growth factor), neurotransmitters and peptide hormone inflammation modulators (including adrenergic, purinergic, cholinergic, ion channels, tachykinin, neuropeptide, substance P, bradykinin, parathyroid hormone, melanocortin and adreno-corticotrophic hormones, and modulators of general cell growth pathways the optimization of therapy of by an agent can be achieved by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

V. Description of Mechanism of Action Hypotheses for Future Drug Development

[0666] There are many potential mechanisms that may serve as targets for candidate therapeutic interventions. For example, phosphodiesterase inhibitors to PDE4; T-lymphocyte-eosinophil interactions inhibition: targeting the factors involved in the regulation of the TH2(CD4+) differentiation and/or activation by soluble factors (cytokines IL-4, IL-5; co-stimulatory molecules (B7-2/CD86); and transcription factors (GATA-3, AP-1). These targets may be available to limit the TH2 cell involvement in the initiation of asthmatic inflammation.

[0667] Suppression of eosinophil adhesion with consequent inhibition of influx into the lung is a strategy to suppress asthmatic airway inflammation. Such inhibition may be mediated through inhibitors directed towards very late antigen-4 (VLA-4), monoclonal antibodies directed towards VLA-4, intracellular adhesion molecule 1 (ICAM1), and alpha 1,3-fucosyltransferase VII (an enzyme which regulates selectin function). Furthermore, molecules may be targeted to suppress the expression of adhesion molecules (e-selectin, vascular cell-adhesion molecule 1 (VCAM-1) and ICAM1).

[0668] There are a group of chemokines that contain a cysteine-X-cysteine motif, such as IL-8 that are effectors of acute inflammatory episodes, whereas cysteine-cysteine chemokines, such as macrophage inhibitory peptide 1 (MIP-1), eotaxin, RANTES, or macrophage chemotactic peptide 1 (MCP-1) act as chronic mediators of inflammation. These molecules may be appropriate targets for inhibiting either the acute or chronic inflammatory pathway.

[0669] Cysteinyl leukotrienes have a central role in the development of chronic asthma, and antagonists (i.e., CySLL1) may be able to ablate the actions of this ligand. These novel leukotriene receptor agonists may have potential for anti-inflammatory effects. Endothelin receptors may also be a target, with endothelin antagonists to specific receptor subtypes ET(A) or ET(B). Other receptors known to be involved in the inflammatory process that may be potential targets are the tachykinin NK1 receptors and selective ligands to the NK1/NK2 receptors.

[0670] Induction of cyclooxygenase and the consequent increase in prostaglandin release is associated with the development of inflammation. Antisense oligonucleotides directed against the receptor types NK-kB, major basic protein, 5-lipoxygenase, leukotriene C4/LTC4 synthetase, IL-4, IL-8, IL-8 and adenosine have been developed that are invaluable products that can directly block the expression of these mediators of the inflammatory response.

[0671] Other areas of drug target development include immunobiology of the allergicies i.e., TH1 and TH2 and their involvement in the immune response, synthesis of immunoglobulin, IgE, integrins, inhibition of αIL-5 and αIL-5 monoclonal antibody, soluble ILA receptor, neuropeptide receptor antagonist, chemokine inhibitors.

[0672] The inflammatory response is also being evaluated in terms of the effects of NO2, SO2, and ozone on the subsequent effect on airway response to these potential allergens. As well as adhesion molecule expression, cytokine production and cytokine gene transcription factors are also being evaluated.

[0673] Optimization of nonsteroidal or steroidal anti-inflammatory agents, or agents aimed at a mechanism of therapy of the underlying etiology of asthma that further demonstrates the utility of selection of a potential asthma patient that has a predisposing genotype in which selective antiasthmatic or other agents, may be more effective and or have a more desirable safety profile. In considering an optimization protocol, one could potentially predetermine variance or variances within the nonsteroidal antiinflammatory pathway, steroid antiinflammatory pathway, or antiinflammatory mediated intracellular mechanism of action that is preeminently responsible for antiasthmatic drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for asthma.

[0674] A sample of therapies approved or in development for preventing or treating the progression of asthma currently known in the art is shown in Table 16. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

[0675] Pharmacogenomics studies for these drugs, as well as other agents, drugs, compounds or candidate therapeutic interventions, could be performed by: identifying genes that are involved in the function of a drug including, but not limited to its absorption, distribution metabolism, or elimination, interaction of the drug with its target as well as potential alternative targets, the response of the cell to the binding of a drug to a target, the metabolism (including synthesis, distribution elimination) of natural compounds which may alter the activity of the drug by complementary, competitive or allosteric mechanisms that potenti ate or limit the effect of the drug, and genes involved in the etiology of the disease that alter its response to a particular class of therapeutic agents. It will be recognized to those skilled in the art that this broadly includes proteins involved in pharmacokinetics as well as genes involved in pharmacodynamics. This also includes genes that encode proteins homologous to the proteins believed to carry out the above functions, which are also worth evaluation as they may carry
out similar functions. Together, the foregoing proteins comprise the candidate genes for affecting response of a patient to the therapeutic intervention. Using the methods described above, variances in these genes can be identified, and research and clinical studies can be performed to establish an association between a drug response or toxicity and specific variances.

Example 2

Inflammatory Bowel Disease

**Description of Inflammatory Bowel Disease**

[0676] Inflammatory bowel disease (IBD) is a broad clinical term that includes idiopathic chronic inflammatory bowel diseases including Crohn’s disease (CD) and ulcerative colitis (UC) which can be distinguished from inflammatory bowel disease of known origin including diverticulitis, radiation enteritis, colitis, drug or toxin-induced enterocolitis, or vascilitis of the intestinal tract. UC is a term that encompasses a broad category of diffuse, continuous, and superficial inflammation of the colon, which begins within the rectum and extends proximally. The condition is limited to the colon and large intestine, with limited involvement of the small intestine. In UC, the inflammation primarily affects the mucosal process and is not transmural within these anatomical regions. CD is characterized by focal, asymmetric, transmural inflammation affecting any portion of the gastrointestinal tract, i.e. from the mouth to the anus. The focal localization and possible extent of the inflammation distinguishes UC from CD. There are currently approximately 35-100 and 10-100 CD per 100,000 Americans diagnosed with UC or CD, respectively.

[0677] Clinically, patients with UC experience variable stool consistencies from constipation to diarrhea, low-grade fever, malaise, nausea, vomiting associated with defecation, night sweats, arthralgias, dehydration, tachycardia, and symptoms of abdominal tenderness. There can be rectal bleeding, tenesmus, and passage of mucus.

[0678] Patients with Crohn’s disease experience symptoms of peptic ulcer disease, nausea, vomiting, and epigastric pain. Transmural inflammation leads to fibrosis and transmural narrowing. In some cases, the narrowing leads to signs and symptoms of intestinal obstruction including nausea, vomiting, waves of abdominal pain, and a reduced output of stool. Patients with colonic CD are likely to experience abdominal pain, cramping or localized pain, rectal bleeding, and diarrhea. Weight loss is common among CD patients due to malabsorption of nutrients and reduced food intake due to minimization of postprandial symptoms.

[0679] There are extraintestinal manifestations of inflammatory bowel disease affecting the following processes including: nutritional and metabolic abnormalities, hematologic abnormalities, skin and mucous membranes, musculoskeletal, hepatic and biliary abnormalities, renal complications, and optic complications. These complications are associated when the colon or intestinal tract is inflamed. These complications are clinically manifest as joint swelling or pain, erythema nodosum, pyoderma gangrenosum, sclerosing cholangitis, conjunctivitis, or uveitis.

[0680] There is an increased risk for the development of gastrointestinal cancer in patients with IBD. In both UC and CD, there is an increased risk of adenocarcinoma of the intestine. This is not correlated to the intensity of the first attack, subsequent course, or and specific medical therapeutic approach. Therefore routine screening for dysplasia and neoplasia is warranted.

Current Therapy of Inflammatory Bowel Disease

[0681] Strategies for the therapy of inflammatory bowel disease includes antiinflammatory agents and immunomodulation.

[0682] Antiinflammatory agents include the use of glucocorticoids and/or aminosalicylates. Glucocorticoids act by modulation of the immune response. Corticosteroids affect the inflammation within the gastrointestinal tract by decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

[0683] 5-aminosalicylic acid (5ASA) is a salicylate that is used for the treatment of IBD, is not orally active, is poorly absorbed and is inactivated by intestinal bacteria, and is delivered as a suppository or rectal suspension enema. Oral formulations can be used to deliver active drug to the lower intestine which are cogeners of 5ASA. The aminosalicylates are potent antiinflammatory agents that inhibit cyclooxygenase (COX), arane limiting enzyme in the prostaglandin and leukotriene pathway.

[0684] Immunosuppressive agents are also used to modulate the inflammatory/immune response. There are four broad categories of immunosuppressive agents that have distinct mechanisms of action: inhibition of ribonucleotide synthesis which acts to inhibit the proliferation of T-cell clones (6-mercaptopurine), inhibition of folic acid which acts to inhibit T-cell and B-cell function as well as decrease II-1 and II-6 activity (methotrexate), inhibition of T-cell receptor stimulated transcription of lymphokine genes which act to inhibit the production of II-2 and II-2 receptors as well as inhibit certain cytokines (TNF-α, IFN-γ) (cyclosporin and FK506), and inhibition of guanosine nucleotide synthesis which acts as cytosstatic effects on lymphocytes (mycophenolate). Each of these categories of agents have been employed for the therapy of IBD.

[0685] Recently, a chimeric monoclonal antibody was approved for use in the treatment of moderately to severely active Crohn’s disease for those patients that are unresponsive to conventional therapy. This monoclonal antibody is specific for TNF-α and can remove TNF from the bloodstream before it reaches the site of inflammation.

[0686] Crohn’s disease may progress to a level and extent in which surgical removal of the localized inflammation is warranted. Surgery has been indicated for recurrent intestinal obstruction, complicated fistulas, intractable hemorrhage, disease refractory to medical therapy, growth retardation refractory to therapies, or cancer. The surgical procedures vary from excision of a localized, diseased portion of the gastrointestinal tract to removal of large portions, i.e. the entire colon (colectomy). Surgical excision of the infammed region or to correct complications such as
blockage, perforation, abscess, or bleeding can result in a substantial relief of symptoms. Limitations to Current Therapies for IBD

[0687] Salicylate associated side effects include dyspepsia, gastric or small bowel bleeding, ulceration, renal insufficiency, confusion, rash, headache, and hepatic toxicity. NSAIDs also reversibly inhibit platelet aggregation and prolong bleeding time.

[0688] Glucocorticoid associated side effects include increased appetite, weight gain, fluid retention, acne, ecchymosis, development of cushoid facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aspartic necrosis, osteoporosis, ulcers, pancreatitis, pseudotumor cerebri, psychosis, glucose, cataract formation, vascular necrosis, increased susceptibility to infection, impairment of the hypothalamus-pituitary axis, decreased thyroid hormone serum binding proteins, and impaired wound healing.

[0689] Agents involved in immunomodulation have the following undesirable side effects including antimetabolites: hepatic compromise including hepatic fibrosis, ascites, esophageal varices, cirrhosis, pneumonia, myelosuppression; immunosuppressives: myelosuppression, cyclosporine: renal insufficiency anemia, hypertension.

[0690] Monoclonal antibody to TNF proteins therapies have been shown to generate a human-antimouse antibody response (HAMA). However, patients on immunosuppressive agents such as glucocorticoids and others are less likely to generate antibodies to the treatment antibody. Delayed hypersensitivity is demonstrable 2 to 4 years after initial treatment in 25% of the patients treated with the chimeric antibody. Further, there are patients that develop a serum sickness reaction which includes fever, and joint swelling that requiring hospital admission. A positive antinuclear antibody (ANA) occurred in 24-36% of the patients analyzed. Nine percent of the patients developed anti-DNA antibodies, less than 1% developed a lupus-like reaction requiring steroid therapy.

[0691] In surgical therapy of IBD, recurring inflammation and relapse, after excision procedures occurs in 75% of the patients. Attempts have been made to include salicylate therapy after respective surgery, however, the inflammation recurrence rate in that group was 52%.

Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions for Autoimmune Disease

Thiopurine Methyltransferase (TPMT)

[0692] The thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme whose precise physiological role is unknown. This enzyme catalyzes the S-methylation of widely used immunosuppressive or cytokotic thiopurine drugs such as 6-thioguamine, 6-mercaptopurine and azathioprine 8. The in vivo activity of this cytosolic enzyme is characterized by interindividual and interethnic variability caused by the genetic polymorphism of the TPMT gene, which was discovered, using pharmacogenetic techniques, by the existence of three major phenotypes, high (HM), intermediate (IM) and deficient (DM) methylation. As a consequence, individuals greatly differ in detoxification of thiopurine drugs to 6-methylmercaptopurine as well as the occurrence of side effects or therapeutic efficacy. Using genomic techniques, PCR-SSCP (polymerase chain reaction—single strand conformation polymorphism), Spirovayron de la Mureyre et al. 9 have defined the mutational and allelic spectrum of TPMT in a group of 191 Europeans. In this analysis, PCR-SSCP techniques identified allelic variances in the entire coding sequence, the exon-intron boundaries, the promoter region and the 3'-flanking region of the genes. Six mutations were detected throughout the ten exons and seven TPMT alleles were characterized. Within the promoter region, six alleles corresponding to a variable number of repeats (VNTR) were identified. The TPMT phenotype was correctly predicted by genotyping for 87% of individuals. A clear negative correlation between the total number of repeats from both alleles and the TPMT activity level was observed, indicating that VNTRs contribute to inter-individual variations of TPMT activity. This VNTR polymorphism can be considered responsible for shifts to lower or higher TPMT activities observed among discordant individuals. Seven out of the nine phenotyped IMs but genotyped IMs were carrier of a total of eight VNTR repeats. This low number of repeat can account for the switch to high TPMT activities of these samples.

[0693] One in 300 patients with IBD are homozygous-deficient for TPMT. The clinical relevance for this deficiency is that TPMT is the enzyme responsible for the conversion of 6-MP to 6-MMP, and the AZA compounds to 6-TG. In TPMT deficient patients, higher levels of 6-TG and 6-MMP are then produced and are associated with significant leukopenia. In general, patients produce variable levels of 6-TG and 6-MMP as determined by their intrinsic enzyme systems. Higher 6-TG levels are correlated with good therapeutic response, but produce leukopenia. Higher 6-MMP levels correlate with hepatotoxicity and in recent studies with leukopenia.

[0694] There is evidence to suggest that there are safety response differences to drug therapy in IBD which may be attributable to genotypic differences between individuals, one example being the TPMT gene described above. There is provided in this invention examples of other gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmaceutical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Table 1, and pathway matrix Table 2 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents for use in Inflammatory Bowel Disease

[0695] The advantages of a clinical research and drug development program that includes the use of polymorphic genotyping for the stratification of patients for the appropriate selection of candidate therapeutic intervention
Pharmacogenomics studies for these drugs, or other agent, compound, drug, or candidate therapeutic intervention, could be performed by identifying genes that are involved in the function of a drug including, but not limited to absorption, distribution, metabolism, or elimination, the interaction of the drug with its target as well as potential alternative targets, the response of the cell to the binding of a drug to a target, the metabolism (including synthesis, biodistribution or elimination) of natural compounds which may alter the activity of the drug by complementary, competitive or allosteric mechanisms that potentiate or limit the effect of the drug, and genes involved in the etiology of the disease that alter its response to a particular class of therapeutic agents. It will be recognized to those skilled in the art that this broadly includes proteins involved in pharmacokinetics as well as genes involved in pharmacodynamics. This also includes genes that encode proteins homologous to the proteins believed to carry out the above functions are also worth evaluation as they may carry out similar functions. Together the foregoing proteins constitute the candidate genes for affecting response of a patient to the therapeutic intervention. Using the methods described above, variances in these genes can be identified, and research and clinical studies can be performed to establish an association between a drug response or toxicity and specific variances.

V. Description of Mechanism of Action Hypotheses for Future Drug Development

The majority of the hypotheses for future therapeutic interventions for inflammatory bowel disease are based upon the understanding the immunologic mechanisms that cause and perpetuate the inflammation within the gastrointestinal tract. Although the initiating event is elusive, the resulting immunologic events have been studied. All of the gastrointestinal enterocytes have immunologic function. Under physiologic conditions, these enterocytes selectively activate CD8+ nonspecific suppressor cells, in response to inflammation. In patients with IBD, these enterocytes selectively stimulate the development of CD4+ helper T cells which can respond in two ways 1) the Th1 response which involves the activation of IL-2 and IFN-γ and leads to delayed hypersensitivity and cellular immunity and 2) the Th2 response which involves IL-4, IL-5, IL-6, and IL-10 and leads to antibody response and humoral immunity. Both Th1 and Th2 responses are genetically controlled and are coordinate regulated, i.e. Th1 response stimulation results in down regulation of Th2 response and vice versa. It has been demonstrated that in UC patients the Th2 response is favored and in CD patients the Th1 response is favored.

A humanized (95% human, 5% mouse) version of the chimeric antibody (75% human, 25% mouse) to TNF is currently under development. Some antidiotypic antibodies are generated, but it doesn’t appear to stimulate a delayed hypersensitivity, no stimulation of anti-DNA antibodies, or lupus-like reactions.

Mediators of the immune response including intracellular adhesion molecule (ICAM-1) inhibitors (antisense molecules or others), IL-10, IL-11 have been tested in humans. Further, and anti-CD4 monoclonal antibody which has been shown to interfere with the interaction of the CD4 molecule and the HLA class II molecules leading to an inhibition of antigen presentation has been tested.

Thalidomide (inhibitor of TNF, acceleration of the degradation of the TNF mRNA) is also under consideration.

It has been noted that individuals who smoke tobacco products have a lower incidence of IBD. Therefore, understanding the immune response and correlation with nicotinic cholinergic pathways is under investigation.

A sample of therapies approved or in development for preventing or treating the progression of IBD currently known in the art is shown in Table 17. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

Example 3

Hepatitis C

Selecting Optimal Therapy for HCV Patients

Genetically Determined Variation in Response to Interferon α

Treatment of hepatitis C virus (HCV) infection with interferon α is expensive, benefits a minority of patients, and produces side effects in a significant fraction of patients. Addition of ribavirin increases efficacy, but combination therapy remains expensive and still falls well short of providing a lasting benefit to most patients. It would therefore be desirable to identify prospectively those patients likely to have a sustained response to treatment. Ideally, a diagnostic test would also predict what dose of interferon and ribavirin, administered for what length of time, will afford to each patient the best chance of a sustained response. Pre-treatment identification of patients likely to suffer serious toxic side effects would also be desirable.

The best characterized predictors of response to interferon α therapy are viral load and HCV genotype. Low viral load before therapy is predictive of a positive response. However, demonstration of decreased viral load after initiation of therapy is currently the best predictor of response to therapy. There is no consensus on the optimal time after initiation of therapy for measuring viral levels with periods ranging from 2 weeks to 4 months have been proposed. The viral load test is not very effective at discriminating long-term responders from those patients who suffer rebound of HCV infection within 6 months after treatment. Also, the ideal test would be performed in advance of any treatment, thereby saving the considerable costs associated with even short-term therapy. In search of other predictive indices, over 100 controlled clinical studies have examined a variety of viral and host factors in responders and non-responders. Genetic variation in both HCV and host genes has
been shown to independently influence patient response to interferon α treatment. A consensus has emerged regarding the interaction of viral genotype and treatment response, however the contribution of host factors to treatment response has not been as well investigated. There are a number of promising recent findings suggesting that polymorphisms in regulators of human immune function are correlated with response to interferon α.

Viral Genome Variation

Comparison of sequenced HCV genomes reveals considerable variation in viral sequence, with at least 6 major types and well over a dozen minor types recognized. The geographical distribution of viral types is nonrandom, perhaps accounting for some of the apparent racial heterogeneity in the natural history of HCV infection. HCV is present in each patient as a heterogeneous population of viral quasispecies, with the degree of heterogeneity differing among patients. Despite these complexities, there are strong correlations between predominant viral type and treatment response. In general, patients with genotype 1 (especially Ib) respond poorly to interferon α, with many studies showing response rates under 10 percent. Patients with genotype 2 or 3 do well, with response rates typically greater than 40 percent. Most viral genotyping is based on a short variable segment. However, there are multiple segments of the viral genome that vary, and some studies have found that more detailed viral genotyping, for example of the 5' untranslated region, provides stronger correlations with treatment response.

Human Genome Variation

A recent study suggests that there is significant variation in response to interferon α treatment among racial groups in the United States, even after controlling for effect of different HCV types. This finding suggests that host genetic variation may be an important factor in response. A number of candidate genes have recently been tested for correlation with interferon α response. The best studied genes are regulators of immune function such as IL-6, IL-10 and TNFα. One study, for example, found that patients with high expression of IL-10 (attributable to a specific haplotype) tend to respond poorly to interferon, perhaps due to impaired immune response. IL-1A, IL-12 and TGF-β levels have been correlated with treatment response in some (but not all) studies. However, no genetic analysis has been performed. Similarly, hepatic levels of interferon α-β receptor have been correlated with response to interferon, but no genetic analysis has been performed to determine whether polymorphism affects receptor levels. IL-4 alleles have also been correlated with response to interferon, particularly the A24-B54-DR2 haplotype. A number of other compelling candidate genes have not been investigated. For example, a recent report shows that HCV can enter cells via the low density lipoprotein receptor. If so, the well studied amino acid polymorphisms of the LDL-R should be investigated for effects on disease course and response to treatment. There are also likely to be genetic factors that influence response to ribavirin. For example, the drug must be transported across the plasma membrane and then phosphorylated before becoming a substrate for viral enzymes. The transporters and kinases responsible for these processes may be worth genetic investigation.

An optimal test for selecting treatment for HCV infection would (i) suggest the optimal therapeutic regimen (interferon alone, interferon and ribavirin, or some other combination), (ii) suggest the optimal dose and duration of treatment, (iii) predict sustained responders vs. short term responders, and (iv) predict patients likely to suffer serious adverse effects. At least three areas should be further investigated to better predict the response to interferon α treatment. First, it is not clear that conventional viral genotyping methods, focusing on the 5' untranslated region, capture all of the aspects of viral sequence variation that affect viral biology. Additional genetic determinants of viral pathogenicity should be investigated. Second, the human gene variants that have been associated with response need to be more thoroughly investigated, and interactions between human candidate gene alleles, as well as perhaps between human genes and viral genes, should be tested. Third, recent work suggests a number of new host proteins that may affect response to interferon, and proteins that mediate response to ribavirin have not yet been investigated. The genes encoding these proteins should be thoroughly investigated. With additional information on candidate genes available, it should be possible to construct a plan, ideally via retrospective analysis of clinical trial data, for first assessing the impact of variation in each of the candidate genes, then examining genes-gene interactions, and finally reducing the number of tests to a much smaller number for confirmatory prospective trials.

In Table 18, there is a list of the candidate therapeutic interventions that in development for Hepatitis. One skilled in the art could apply, as described in the text, the methods of this invention to ascertain whether there is a gene in the inflammatory pathway that may be involved in the efficacy, safety, or toxicities of these candidate interventions.

Example 4

Method for Producing cDNA

In order to identify sequence variances in a gene by laboratory methods it is in some instances useful to produce cDNA(s) from multiple human subjects. (In other instances it may be preferable to study genomic subjects.). Methods for producing cDNA are known to those skilled in the art, as are methods for amplifying and sequencing the cDNA or portions thereof. An example of a useful cDNA production protocol is provided below. As recognized by those skilled in the art, other specific protocols can also be used.

cDNA Production

1. Make sure that all tubes and pipette tips are RNase-free. (Bake them overnight at 100° C. in a vacuum oven to make them RNase-free.)

2. Add the following to a RNase-free 0.2 ml micro-amp tube and mix gently:

   1. 24 ul water (DEPC treated)
   2. 12 ul RNA (1 ug/ul)
   3. 12 ul random hexamers (50 ng/ul)

3. Incubate on ice for 1 minute.

4. Add the following:

   16 ul 5x Synthesis Buffer
   8 ul 0.1 M DTT
5. Incubate at 42°C for 50 minutes.

6. Heat to 70°C for ten minutes to kill the enzyme, then place it on ice.

7. Add 160 ul of water to the reaction so that the final volume is 240 ul.

8. Use PCR to check the quality of the cDNA. Use primer pairs that will give a ~800 base pair long piece. See “PCR Optimization” for the PCR protocol.

The following chart shows the reagent amounts for a 20 ul reaction, a 80 ul reaction, and a batch of 39 (which makes enough mix for 36) reactions:

<table>
<thead>
<tr>
<th></th>
<th>20 ul × 1 tube</th>
<th>80 ul × 1 tube</th>
<th>80 ul × 39 tubes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>6 ul</td>
<td>24 ul</td>
<td>936 Water</td>
</tr>
<tr>
<td>RNA</td>
<td>3 ul</td>
<td>12 ul</td>
<td>RNA</td>
</tr>
<tr>
<td>random hexamers</td>
<td>3 ul</td>
<td>12 ul</td>
<td>random hexamers</td>
</tr>
<tr>
<td>synthesis buffer</td>
<td>4 ul</td>
<td>16 ul</td>
<td>Synthesis buffer</td>
</tr>
<tr>
<td>0.1 M DTT</td>
<td>2 ul</td>
<td>8 ul</td>
<td>312 0.1 M DTT</td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>1 ul</td>
<td>4 ul</td>
<td>156 10 mM dNTP</td>
</tr>
<tr>
<td>SSRT</td>
<td>1 ul</td>
<td>4 ul</td>
<td>156 SSRT</td>
</tr>
</tbody>
</table>

Example 5

Method for Detecting Variances by Single Strand Conformation Polymorphism (SSCP) Analysis

This example describes the SSCP technique for identification of sequence variances of genes. SSCP is usually paired with a DNA sequencing method, since the SSCP method does not provide the nucleotide identity of variances. One useful sequencing method, for example, is DNA cycle sequencing of 32P labeled PCR products using the Femtomole DNA cycle sequencing kit from Promega (WI) and the instructions provided with the kit. Fragments are selected for DNA sequencing based on their behavior in the SSCP assay.

Single strand conformation polymorphism screening is a widely used technique for identifying an discriminating DNA fragments which differ from each other by as little as a single nucleotide. As originally developed by Orita et al. (Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA, 86(8):2766-70, 1989), the technique was used on genomic DNA, however the same group showed that the technique works very well on PCR amplified DNA as well. In the last 10 years the technique has been used in hundreds of published papers, and modifications of the technique have been described in dozens of papers. The enduring popularity of the technique is due to (1) a high degree of sensitivity to single base differences (>90%), (2) a high degree of selectivity, measured as a low frequency of false positives, and (3) technical ease. SSCP is almost always used together with DNA sequencing because SSCP does not directly provide the sequence basis of differential fragment mobility. The basic steps of the SSCP procedure are described below.

When the intent of SSCP screening is to identify a large number of gene variances it is useful to screen a relatively large number of individuals of different racial, ethnic and/or geographic origins. For example, 32 or 48 or 96 individuals is a convenient number to screen because gel electrophoresis apparatus are available with 96 wells (Applied Biosystems Division of Perkin Elmer Corporation), allowing 3×32, 2×48 or 96 samples to be loaded per gel.

The 32 (or more) individuals screened should be representative of most of the world’s major populations. For example, an equal distribution of Africans, Europeans and Asians constitutes a reasonable screening set. One useful source of cell lines from different populations is the Coriell Cell Repository (Camden, N.J.), which sells EBV immortalized lymphoblastoid cells obtained from several thousand subjects, and includes the racial/ethnic/geographic background of cell line donors in its catalog. Alternatively, a panel of cDNAs can be isolated from any specific target population.

SSCP can be used to analyze cDNAs or genomic DNAs. For many genes cDNA analysis is preferable because for many genes the full genomic sequence of the target gene is not available, however, this circumstance will change over the next few years. To produce cDNA requires RNA. Therefore each cell lines is grown to mass culture and RNA is isolated using an acid/phenol protocol, sold in kit form as Trizol by Life Technologies (Gaithersberg, Md.). The unfractionated RNA is used to produce cDNA by the action of a modified Mäleyn Marine Leukemia Virus Reverse Transcriptase, purchased in kit form from Life Technologies (Superscript II kit). The reverse transcriptase is primed with random hexamer primers to initiate cDNA synthesis along the whole length of the RNAs. This proved useful later in obtaining good PCR products from the 5’ ends of some genes. Alternatively, oligoDT can be used to prime cDNA synthesis.

Material for SSCP analysis can be prepared by PCR amplification of the cDNA in the presence of one 32P labeled dNTP (usually ε32P dCTP). Usually the concentration of nonradioactive dCTP is dropped from 200 uM (the standard concentration for each of the four dNTPs) to about 100 uM, and 32P dCTP is added to a concentration of about 0.1-0.3 uM. This involves adding a 0.3-1 ul (3-10 uCi) of 32P dCTP to a 10 ul PCR reaction. Radioactive nucleotides can be purchased from DuPont/New England Nuclear.

The customary practice is to amplify about 200 base pair PCR products for SSCP, however, an alternative approach is to amplify about 0.8-1.4 kb fragments and then use several cocktails of restriction endonucleases to digest those into smaller fragments of about 0.1-0.4 kb, aiming to have as many fragments as possible between 0.15 and 0.3 kb. The digestion strategy has the advantage that less PCR is required, reducing both time and costs. Also, several different restriction enzyme digests can be performed on each set of samples (for example 96 cDNAs), and then each of the digests can be run separately on SSCP gels. This redundant method (where each nucleotide is surveyed in three different fragments) reduces both the false negative and false positive rates. For example: a site of variance
might lie within 2 bases of the end of a fragment in one digest, and as a result not affect the conformation of that strand; the same variance, in a second or third digest, would likely lie in a location more prone to affect strand folding, and therefore be detected by SSCP.

[0728] After digestion, the radio-labelled PCR products are diluted 1:5 by adding formamide load buffer (80% formamide, 1xSSCP gel buffer) and then denatured by heating to 90°C for 10 minutes, and then allowed to renature by quickly chilling on ice. This procedure (both the dilution and the quick chilling) promotes intra-(rather than inter-) strand association and secondary structure formation. The secondary structure of the single strands influences their mobility on non-denaturing gels, presumably by influencing the number of collisions between the molecule and the gel matrix (i.e., gel sieving). Even single base differences consistently produce changes in intrastrand folding sufficient to register as mobility differences on SSCP.

[0729] The single strands were then resolved on two gels, one a 5.5% acrylamide, 0.5xTBE gel, the other an 8% acrylamide, 10% glycerol, 1xTTE gel. (Other gel recipes are known to those skilled in the art.) The use of two gels provides a greater opportunity to recognize mobility differences. Both glycerol and acrylamide concentration have been shown to influence SSCP performance. By routinely analyzing three different digests under two gel conditions (effectively 6 conditions), and by looking at both strands under all 6 conditions, one can achieve a 12-fold sampling of each base pair of cDNA. However, if the goal is to rapidly survey many genes or cDNAs then a less redundant procedure would be optimal.

Example 6

Method for Detecting Variances by T4 Endonuclease VII (T4E7) Mismatch Cleavage Method

[0730] The enzyme T4 endonuclease VII is derived from the bacteriophage T4. T4 endonuclease is used by the bacteriophage to cleave branched DNA intermediates which form during replication so the DNA can be processed and packaged. T4 endonuclease can also recognize and cleave heteroduplex DNA containing single base mismatches as well as deletions and insertions. This activity of the T4 endonuclease enzyme can be exploited to detect sequence variances present in the general population.

[0731] The following are the major steps involved in identifying sequence variances in a candidate gene by T4 endonuclease VII mismatch cleavage:

[0732] 1. Amplification by the polymerase chain reaction (PCR) of 400-600 bp regions of the candidate gene from a panel of DNA samples. The DNA samples can either be cDNA or genomic DNA and will represent some cross section of the world population.

[0733] 2. Mixing of a fluorescently labeled probe DNA with the sample DNA. Heating and cooling the mixtures causing heteroduplex formation between the probe DNA and the sample DNA.

[0734] 3. Addition of T4 endonuclease VII to the heteroduplex DNA samples. T4 endonuclease will recognize and cleave at sequence variance mismatches formed in the heteroduplex DNA.

[0735] 4. Electrophoresis of the cleaved fragments on an ABI sequencer to determine the site of cleavage.

[0736] 5. Sequencing of a subset of PCR fragments identified by T4 endonuclease VI to contain variances to establish the specific base variation at that location.

A more detailed description of the procedure is as follows:

[0737] A candidate gene sequence is downloaded from an appropriate database. Primers for PCR amplification are designed which will result in the target sequence being divided into amplification products of between 400 and 600 bp. There will be a minimum of a 50 bp of overlap not including the primer sequences between the 5' and 3' ends of adjacent fragments to ensure the detection of variances which are located close to one of the primers.

[0738] Optimal PCR conditions for each of the primer pairs is determined experimentally. Parameters including but not limited to annealing temperature, pH, MgCl₂ concentration, and KCl concentration will be varied until conditions for optimal PCR amplification are established. The PCR conditions derived for each primer pair is then used to amplify a panel of DNA samples (cDNA or genomic DNA) which is chosen to best represent the various ethnic backgrounds of the world population or some designated subset of that population.

[0739] One of the DNA samples is chosen to be used as a probe. The same PCR conditions used to amplify the panel are used to amplify the probe DNA. However, a fluorescently labeled nucleotide is included in the deoxy-nucleotide mix so that a percentage of the incorporated nucleotides will be fluorescently labeled.

[0740] The labeled probe is mixed with the corresponding PCR products from each of the DNA samples and then heated and cooled rapidly. This allows the formation of heteroduplexes between the probe and the PCR fragments from each of the DNA samples. T4 endonuclease VII is added directly to these reactions and allowed to incubate for 30 min. at 37°C. 10 ul of the Formamide loading buffer is added directly to each of the samples and then denatured by heating and cooling. A portion of each of these samples is electrophoresed on an ABI 377 sequencer. If there is a sequence variance between the probe DNA and the sample DNA a mismatch will be present in the heteroduplex fragment formed. The enzyme T4 endonuclease VII will recognize the mismatch and cleave at the site of the mismatch. This will result in the appearance of two peaks corresponding to the two cleavage products when run on the ABI 377 sequencer.

[0741] Fragments identified as containing sequencing variances are subsequently sequenced using conventional methods to establish the exact location and sequence variance.

Example 7

Method for Detecting Variances by DNA Sequencing.

[0742] Sequencing by the Sanger dye-deoxy method or the Maxam Gilbert chemical cleavage method is widely used to determine the nucleotide sequence of genes. Presently, a worldwide effort is being put forward to sequence the entire human genome. The Human Genome Project as it is called has already resulted in the identification and sequencing of
many new human genes. Sequencing can not only be used to identify new genes, but can also be used to identify variations between individuals in the sequence of those genes.

0743] The following are the major steps involved in identifying sequence variations in a candidate gene by sequencing:

0744] 1. Amplification by the polymerase chain reaction (PCR) of 400-700 bp regions of the candidate gene from a panel of DNA samples The DNA samples can either be cDNA or genomic DNA and will represent some cross section of the world population.

0745] 2. Sequencing of the resulting PCR fragments using the Sanger dideoxy method. Sequencing reactions are performed using fluorescently labeled dideoxy terminators and fragments are separated by electrophoresis on an ABI 377 sequencer or its equivalent.

0746] 3. Analysis of the resulting data from the ABI 377 sequencer using software programs designed to identify sequence variations between the different samples analyzed.

A more detailed description of the procedure is as follows:

0747] A candidate gene sequence is downloaded from an appropriate database. Primers for PCR amplification are designed which will result in the target sequence being divided into amplification products of between 400 and 700 bp. There will be a minimum of a 50 bp of overlap not including the primer sequences between the 5' and 3' ends of adjacent fragments to ensure the detection of variations which are located close to one of the primers.

0748] Optimal PCR conditions for each of the primer pairs is determined experimentally. Parameters including but not limited to annealing temperature, pH, MgCl₂ concentration, and KCl concentration will be varied until conditions for optimal PCR amplification are established. The PCR conditions derived for each primer pair is then used to amplify a panel of DNA samples (cDNA or genomic DNA) which is chosen to best represent the various ethnic backgrounds of the world population or some designated subset of that population.

0749] PCR reactions are purified using the QIAquick 8 PCR purification kit (Qiagen cat# 28142) to remove nucleotides, proteins and buffers. The PCR reactions are mixed with 5 volumes of Buffer EB and applied to the wells of the QIAquick spin columns. The liquid is pulled through the strips by applying a vacuum. The wells are then washed two times with 1 ml of buffer PE and allowed to dry for 5 minutes under vacuum. The PCR products are eluted from the strips using 60 ul of elution buffer.

0750] The purified PCR fragments are sequenced in both directions using the Perkin Elmer ABI Prism™ Big Dye™ terminator Cycle Sequencing Ready Reaction Kit (Cat# 4303150). The following sequencing reaction is set up: 8.0 ul Terminator Ready Reaction Mix, 6.0 ul of purified PCR fragment, 20 picomoles of primer, deionized water to 20 ul. The reactions are run through the following cycles 25 times: 96°C. for 10 second, annealing temperature for that particular PCR product for 5 seconds, 60°C. for 4 minutes.

0751] The above sequencing reactions are ethanol precipitated directly in the PCR plate, washed with 70% ethanol, and brought up in a volume of 6 ul of formamide dye. The reactions are heated to 90°C. for 2 minutes and then quickly cooled to 4°C. 1 ul of each sequencing reaction is then loaded and run on an ABI 377 sequencer.

0752] The output for the ABI sequencer appears as a series of peaks where each of the different nucleotides, A, C, G, and T appear as a different color. The nucleotide at each position in the sequence is determined by the most prominent peak at each location. Comparison of each of the sequencing outputs for each sample can be examined using software programs to determine the presence of a variance in the sequence. One example of heterozygote detection using sequencing with dye labeled terminators is described by Kwok et. al. (Kwok, P.-Y.; Carlson, C.; Yager, T. D., Ankener, W., and D. A. Nickerson, Genomics 23, 138-144, 1994). The software compares each of the normalized peaks between all the samples base by base and looks for a 40% decrease in peak height and the concomitant appearance of a new peak underneath. Possible variances flagged by the software are further analyzed visually to confirm their validity.

Example 8
Hardy-Weinberg Equilibrium

0753] Evolution is the process of change and diversification of organisms through time, and evolutionary change affects morphology, physiology and reproduction of organisms, including humans. These evolutionary changes are the result of changes in the underlying genetic or hereditary material. Evolutionary changes in a group of interbreeding individuals or Mendelian population, or simply populations, are described in terms of changes in the frequency of genotypes and their constituent alleles. Genotype frequencies for any given generation is the result of the mating among members (genotypes) of their previous generation. Thus, the expected proportion of genotypes from a random union of individuals in a given population is needed for describing the total genetic variation for a population of any species. For example, the expected number of genotypes that could form from the random union of two alleles, A and a, of a gene are AA, Aa and aa. The expected frequency of genotypes in a large, random mating population was discovered to remain constant from generation to generation; or achieve Hardy-Weinberg equilibrium, named after its discoverers. The expected genotypic frequencies of alleles A and a (AA, 2Aa, aa) are conventionally described in terms of p²+2pq+q² in which p and q are the allele frequencies of A and a. In this equation (p²+2pq+q²=1), p is defined as the frequency of one allele and q as the frequency of another allele for a trait controlled by a pair of alleles (A and a). In other words, p equals all of the alleles in individuals who are homozygous dominant (AA) and half of the alleles in individuals who are heterozygous (Aa) for this trait. In mathematical terms, this is

\[ p = \frac{AA}{AA + 2Aa} \]

Likewise, q equals the other half of the alleles for the trait in the population, or

\[ q = \frac{2Aa + aa}{AA + 2Aa + aa} \]
Because there are only two alleles in this case, the frequency of one plus the frequency of the other must equal 100%, which is to say

\[ p + q = 1 \]

Alternatively,

\[ p = 1 - q \quad \text{OR} \quad q = 1 - p \]

All possible combinations of two alleles can be expressed as:

\[ (p+q)^2 = 1 \]

or more simply,

\[ p^2 + 2pq + q^2 = 1 \]

In this equation, if \( p \) is assumed to be dominant, then \( p^2 \) is the frequency of homozygous dominant (AA) individuals in a population, \( 2pq \) is the frequency of heterozygous (Aa) individuals, and \( q^2 \) is the frequency of homozygous recessive (aa) individuals.

[0754] From observations of phenotypes, it is usually only possible to know the frequency of homozygous dominant or recessive individuals, because both dominant and recessives will express the distinguishable traits. However, the Hardy-Weinberg equation allows us to determine the expected frequencies of all the genotypes, if only \( p \) or \( q \) is known. Knowing \( p \) and \( q \), it is a simple matter to plug these values into the Hardy-Weinberg equation \( (p+2pq+q^2=1) \). This then provides the frequencies of all three genotypes for the selected trait within the population. This illustration shows Hardy-Weinberg frequency distributions for the genotypes AA, Aa, and aa at all values for frequencies of the alleles, \( p \) and \( q \). It should be noted that the proportion of heterozygotes increases as the values of \( p \) and \( q \) approach 0.5.

Linkage Disequilibrium

[0755] Linkage is the tendency of genes or DNA sequences (e.g. SNPs) to be inherited together as a consequence of their physical proximity on a single chromosome. The closer together the markers are, the lower the probability that they will be separated during DNA crossing over, and hence the greater the probability that they will be inherited together. Suppose a mutational event introduces a "new" allele in the close proximity of a gene or an allele. The new allele will tend to be inherited together with the alleles present on the "ancestral," chromosome or haplotype. However, the resulting association, called linkage disequilibrium, will decline over time due to recombination. Linkage disequilibrium has been used to map disease genes. In general, both allele and haplotype frequencies differ among populations. Linkage disequilibrium is varied among the populations, being absent in some and highly significant in others.

Quantification of the Relative Risk of Observable Outcomes of a Pharmacogenomics Trial

Let PlatR be the placebo response rate (0% (PlatR (100%)) and TmTR be the treatment response rate (0% (TmTR (100%)) of a classical clinical trial. ObsRR is defined as the relative risk between TmTR and PlatR:

\[ \text{ObsRR= PlatR/TmTR.} \]

[0756] Suppose that in the treatment group there is a polymorphism in relation to drug metabolism such as the treatment response rate is different for each genotypic sub-group of patients. Let \( q \) be the allele a frequency of a recessive biallelic locus (e.g. SNP) and \( p=1-q \) the allele A frequency. Following Hardy-Weinberg equilibrium, the relative frequency of homozygous and heterozygous patients are as follow:

<table>
<thead>
<tr>
<th>AA: p2</th>
<th>Aa: 2pq</th>
<th>aa: q2</th>
</tr>
</thead>
</table>

with

\[ (p^2+2pq+q^2)=1 \]

Let's define AAR, AaR, aaR as respectively the response rates of the AA, Aa and aa patients. We have the following relationship:

\[ \text{TmTR= AAR}^*p^2+AaR^*2pq+aaR^*q^2. \]

[0757] Suppose that the aa genotypic group of patients has the lowest response rate, i.e. a response rate equal to the placebo response rate (which means that the polymorphism has no impact on natural disease evolution but only on drug action) and let's define ExpRR as the relative risk between AAR and aaR, as

\[ \text{ExpRR= AAR/aaR.} \]

From the previous equations, we have the following relationships:

\[ \text{ObsRR/ExpRR}=\text{PlatR} \]

\[ \text{TmTR/PlatR}=\text{(AAR}^*p^2+AaR^*2pq+aaR^*q^2)/\text{PlatR} \]

The maximum of the expected relative risk, max(ExpRR), corresponding to the case of heterozygous patients having the same response rate as the placebo rate, is such that:

\[ \text{ObsRR= ExpRR}^*p^2+2pq+q^2<\text{ExpRR= (ObsRR-2pq-}\]

\[ q^2)/p^2 \]

The minimum of the expected relative risk, min(ExpRR), corresponding to the case of heterozygous patients having the same response rate as the homozygous non-affected patients, is such that:

\[ \text{ObsRR= ExpRR}^*(p+2pq)\]

[0758] For example, if \( q=0.4 \), PlatR=40% and ObsRR=1.5 (i.e. TmTR=60%), then 1.6 (ExpRR (2.4). This means that the best treatment response rate we can expect in a genotypic subgroup of patients in these conditions would be 95.6% instead of 60%.

[0759] This can also be expressed in terms of maximum potential gain between the observed difference in response rates (TmTR-PlatR) without any pharmacogenetic hypothesis and the maximum expected difference in response rates (max(ExpRR)*PlatR-TmTR) with a strong pharmacogenetic hypothesis:

\[ \text{max(ExpRR)*PlatR-TmTR= (ObsRR-2pq-q^2)/p^2} \]

\[ \text{ExpRR-TmTR} \]

\[ \text{(max(ExpRR)*PlatR-TmTR=)} \]

\[ \text{(TmTR=PlatR^*2pq+q^2)} \]

\[ \text{exp(ExpRR=)} \]

\[ \text{ObsRR= (95.6%)} \]

\[ \text{q^2)} \]

that is for the previous example,

\[ (95.6%)=\{(1-0.62)/0.62\}^*(60%-40%)=35.6% \]
Suppose that, instead of one SNP, we have \( p \) loci of SNPs for one gene. This means that we have \( 2^p \) possible haplotypes for this gene and \( (2p)(2p-1)/2 \) possible genotypes. And with 2 genes with \( p_1 \) and \( p_2 \) SNP loci, we have \( [2(p_1)(p_1-1)/2]^p_1[2(p_2)(2p_2-1)/2]^p_2 \) possibilities; and so on. Examining haplotypes instead of combinations of SNPs is especially useful when there is linkage disequilibrium enough to reduce the number of combinations to test, but not complete since in this last case one SNP would be sufficient. Yet the problem of frequency above still remains with haplotypes instead of SNPs since the frequency of a haplotype cannot be higher than the highest SNP frequency involved.

Statistical Methods to be Used in Objective Analyses

The statistical significance of the differences between variance frequencies can be assessed by a Pearson chi-squared test of homogeneity of proportions with \( n-1 \) degrees of freedom. Then, in order to determine which variance(s) is(are) responsible for an eventual significance, we can consider each variance individually against the rest, up to \( n \) comparisons, each based on a \( 2 \times 2 \) table. This should result in chi-squared tests that are individually valid, but taking the most significant of these tests is a form of multiple testing. A Bonferroni’s adjustment for multiple testing will thus be made to the \( P \)-values, such as \( P^* = 1 - (1 - P)^p \).

The statistical significance of the difference between genotype frequencies associated to every variance can be assessed by a Pearson chi-squared test of homogeneity of proportions with \( 2 \) degrees of freedom, using the same Bonferroni’s adjustment as above.

Testing for unequal haplotype frequencies between cases and controls can be considered in the same framework as testing for unequal variance frequencies since a single variance can be considered as a haplotype of a single locus. The relevant likelihood ratio test compares a model where two separate sets of haplotype frequencies apply to the cases and controls, to one where the entire sample is characterized by a single common set of haplotype frequencies. This can be performed by repeated use of a computer program (Terwilliger and Ott, 1994, Handbook of Human Linkage Analysis, Baltimore, John Hopkins University Press) to successively obtain the log-likelihood corresponding to the set of haplotype frequency estimates on the cases (\( \ln L_{\text{case}} \)) on the controls (\( \ln L_{\text{control}} \)) and on the overall (\( \ln L_{\text{combined}} \)). The test statistic \( 2(\ln L_{\text{case}}) - (\ln L_{\text{control}}) - (\ln L_{\text{combined}}) \) is then chi-squared with \( r-1 \) degrees of freedom (where \( r \) is the number of haplotypes).

To test for potentially confounding effects or effect-modifiers, such as sex, age, etc., logistic regression can be used with case-control status as the outcome variable, and genotypes and covariates (plus possible interactions) as predictor variables.

Example 9

Exemplary Pharmacogenetic Analysis Steps

In accordance with the discussion of distribution frequencies for variances, alleles, and haplotypes, variance detection, and correlation of variances or haplotypes with treatment response variability, the points below list major items which will typically be performed in an analysis of the pharmacogenetic determination of the effects of variances in the treatment of a disease and the selection/optimization of treatment.

1) List candidate gene/genes for a known genetic disease, and assign them to the respective metabolic pathways.

2) Determine their alleles, observed and expected frequencies, and their relative distributions among various ethnic groups, gender, both in the control and in the study (case) groups.

3) Measure the relevant clinical/phenotypic (biochemical/physiological) variables of the disease.

4) If the causal variance/allele in the candidate gene is unknown, then determine linkage disequilibria among variances of the candidate gene(s).

5) Divide the regions of the candidate genes into regions of high linkage disequilibrium and low disequilibrium.

6) Develop haplotypes among variances that show strong linkage disequilibrium using the computation methods.

7) Determine the presence of rare haplotypes experimentally. Confirm if the computationally determined rare haplotypes agree with the experimentally determined haplotypes.

8) If there is a disagreement between the experimentally determined haplotypes and the computationally derived haplotypes, drop the computationally derived rare haplotypes, construct cladograms from these haplotypes using the Templeton (1987) algorithm.

9) Note regions of high recombination. Divide regions of high recombination further to see patterns of linkage disequilibria.

10) Establish association between cladograms and clinical variables using the nested analysis of variance as presented by Templeton (1995), and assign causal variance to a specific haplotype.

11) For variances in the regions of high recombination, use permutation tests for establishing associations between variances and the phenotypic variables.

12) If two or more genes are found to affect a clinical variable determine the relative contribution of each of the genes or variances in relation to the clinical variable, using step-wise regression or discriminant function or principal component analysis.

13) Determine the relative magnitudes of the effects of any of the two variances on the clinical variable due to their genetic (additive, dominant or epistasis) interaction.

14) Using the frequency of an allele or haplotypes, as well as biochemical/clinical variables determined in the in vitro or in vivo studies, determine the effect of that gene or allele on the expression of the clinical variable, according to the measured genotype approach of Boerwinkle et al (Ann. Hum. Genet. 1986).

15) Stratify ethnic/clinical populations based on the presence or absence of a given allele or a haplotype.
[0781] 16) Optimize drug dosages based on the frequency of alleles and haplotypes as well as their effects using the measured genotype approach as a guide.

Example 10

Exemplary Pharmacogenetic Analysis Steps—Biological Function Analysis

[0782] In many cases when a gene which may affect drug action is found to exhibit variances in the gene, RNA, or protein sequence, it is preferable to perform biological experiments to determine the biological impact of the variances on the structure and function of the gene or its expressed product and on drug action. Such experiments may be performed in vitro or in vivo using methods known in the art.

[0783] The points below list major items which may typically be performed in an analysis of the effects of variances in the treatment of a disease and the selection/optimization of treatment using biological studies to determine the structure and function of variant forms of a gene or its expressed product.

[0784] 1) List candidate gene/genes for a known genetic disease, and assign them to the respective metabolic pathways.

[0785] 2) Identify variances in the gene sequence, the expressed mRNA sequence or expressed protein sequence.

[0786] 3) Match the position of variances to regions of the gene, mRNA, or protein with known biological functions. For example, specific sequences in the promoter of a gene are known to be responsible for determining the level of expression of the gene; specific sequences in the mRNA are known to be involved in the processing of nuclear mRNA into cytoplasmic mRNA including splicing and polyadenylation; and certain sequences in proteins are known to direct the trafficking of proteins to specific locations within a cell and to constitute active sites of biological functions including the binding of proteins to other biological constituents or catalytic functions. Variances in sites such as these, and others known in the art, are candidates for biological effects on drug action.

[0787] 4) Model the effect of the variance on mRNA or protein structure. Computational methods for predicting the structure of mRNA are known and can be used to assess whether a specific variance is likely to cause a substantial change in the structure of mRNA. Computational methods can also be used to predict the structure of peptide sequences enabling predictions to be made concerning the potential impact of the variance on protein function. Most useful are structures of proteins determined by X-ray diffraction, NMR or other methods known in the art which provide the atomic structure of the protein. Computational methods can be used to consider the effect of changing an amino acid within such a structure to determine whether such a change would disrupt the structure and/or function of the protein. Those skilled in the art will recognize that this analysis can be performed on crystal structures of the protein known to have a variance as well as homologous proteins expressed from different loci in the human genome, or homologous proteins from other species, or non-homologous but analogous proteins with similar functions from humans or other species.

[0788] 5) Produce the gene, mRNA or protein in amounts sufficient to experimentally characterize the structure and function of the gene, mRNA or protein. It will be apparent to those skilled in the art that by comparing the activity of two genes or their products which differ by a single variance, the effect of the variance can be determined. Methods for producing genes or gene products which differ by one or more bases for the purpose of experimental analysis are known in the art.

[0789] 6) Experimental methods known in the art can be used to determine whether a specific variance alters the transcription of a gene and translation into a gene product. This involves producing amounts of the gene by molecular cloning sufficient for in vitro or in vivo studies. Methods for producing genes and gene products are known in the art and include cloning of segments of genetic material in prokaryotes or eukaryotic hosts, run off transcription and cell-free translation assays that can be performed in cell free extracts, transfection of DNA into cultured cells, introduction of genes into live animals or embryos by direct injection or using vehicles for gene delivery including transfection mixtures or viral vectors.

[0790] 7) Experimental methods known in the art can be used to determine whether a specific variance alters the ability of a gene to be transcribed into RNA. For example, run off transcription assays can be performed in vitro or expression can be characterized in transfected cells or transgenic animals.

[0791] 8) Experimental methods known in the art can be used to determine whether a specific variance alters the processing, stability, or translation of RNA into protein. For example, reticulocyte lysate assays can be used to study the production of protein in cell free systems, transfection assays can be designed to study the production of protein in cultured cells, and the production of gene products can be measured in transgenic animals.

[0792] 9) Experimental methods known in the art can be used to determine whether a specific variant alters the activity of an expressed protein product. For example, protein can be produced by reticulocyte lysate systems or by introducing the gene into prokaryotic organisms such as bacteria or lower eukaryotic organisms such as yeast or fungus, or by introducing the gene into cultured cells or transgenic animals. Protein produced in such systems can be extracted or purified and subjected to bioassays known to those in the art as measures of the action of that particular protein. Bioassays may involve, but are not limited to, binding, inhibition, or catalytic functions.

[0793] 10) Those skilled in the art will recognize that it is sometimes preferred to perform the above experiments in the presence of a specific drug to determine whether the drug has differential effects on the activity being measured. Alternatively, studies may be performed in the presence of an analogue or metabolite of the drug.

[0794] 11) Using methods described above, specific variances which alter the biological function of a gene or its gene product that could have an impact on drug action can be identified. Such variances are then studied in clinical
trial populations to determine whether the presence or absence of a specific variance correlates with observed clinical outcomes such as efficacy or toxicity.

[0795] 12) It will be further recognized that there may be more than one variance within a gene that is capable of altering the biological function of the gene or gene product. These variances may exhibit similar, synergistic effects, or may have opposite effects on gene function. In such cases, it is necessary to consider the haplotype of the gene, namely the combination of variances that are present within a single allele, to assess the composite function of the gene or gene product.

[0796] 13) Perform clinical trials with stratification of patients based on presence or absence of a given variance, allele or haplotype of a gene. Establish associations between observed drug responses such as toxicity, efficacy, drug response, or dose toleration and the presence or absence of a specific variance, allele, or haplotype.

[0797] 14) Optimize drug dosage or drug usage based on the presence of the variant.

Example 11

Stratification of Patients by Genotype in Prospective Clinical Trials.

[0798] In a prospective clinical trial, patients will be stratified by genotype to determine whether the observed outcomes are different in patients having different genotypes. A critical issue is the design of such trials to assure that a sufficient number of patients are studied to observe genetic effects.

The number of patients required to achieve statistical significance in a conventional clinical trial is calculated from:

\[ N = \frac{2(\sigma^2 + 2\beta^2)}{\delta^2} / \text{two tailed test} \]

From this equation it may be inferred that the size of a genetically defined subgroup \( N_p \) required to achieve statistical significance for an observed outcome associated with variance or haplotype “i” can be calculated as:

\[ N_p = \frac{2(\sigma_p^2 + 2\beta_p^2)}{\delta^2} \]

If \( P_i \) is the prevalence of the genotype “i” in the population, the total number of patients that need to be incorporated in a clinical trial \( N_{p} \) to identify a population with haplotype “i” of size \( N_{p} \) is given by:

\[ N_{p} = \frac{N}{P_i} \]

It should be noted that \( N_p \) describes the total number of patients that need to be genotyped in order to identify a subset of \( N_p \) patients with genotype “i”.

If genotyping is used as means for statistical stratification of patients, \( N_p \) represents the number of patients that would need to be enrolled in a trial to achieve statistical significance for subgroup “i”. If genotyping is used as a means for inclusion, it represents the number of patients that needs to be screened to identify a population of \( N_p \) individuals for an appropriately powered clinical trial. Thus, \( N_p \) is a critical determinant of the scope of the clinical trial as well as \( N_p \).

A clinical trial can also be designed to test associations for multiple genetic subgroups “j” defined by a single allele in which case:

\[ N_p = \max(N_{p,i}) \text{ for } i = 1 \ldots j \]

If more than one subgroup is tested, but there is no overlap in the patients contained within the subgroups, these can be considered to be independent hypotheses and no multiple testing correction should be required. If consideration of more than one subgroup constitutes multiple testing, or if individual patients are included in multiple subgroups, then statistical corrections may be required in the values of \( Z_p \) or \( Z_p' \), which would increase the number of patients required.

It should be emphasized that a clinical trial of this nature may not provide statistically significant data concerning associations with any genotype other than “i”. The total number of patients that would be required in a clinical trial to test more than one genetically defined subgroup would be determined by the maximum value of \( N_p \) for any single subgroup.

The power of pharmacogenomics to improve the efficiency of clinical trials arises from the fact it is possible to have \( N_p < N \). The goal of pharmacogenomic analysis is to identify a genetically define subgroup in which the magnitude of the clinical response is greater and the variability in response is reduced. These observations correspond to an increase in the magnitude of the (mean) observed response \( \delta \) or a decrease in the degree of variability \( \alpha \). Since the value of \( N_p \) calculated in equation 1.2 decreases non-linearly as the square of these changes, the total number of patients \( N_p \) can also decrease non-linearly, resulting in a clinical trial that requires fewer patients to achieve statistical significance. If \( \delta_i \) and \( \alpha_i \) are not different than \( \delta \) and \( \alpha \), then \( N_p \) is greater than \( N \) as given by:

\[ N_p = \frac{N}{P_i} \]

It is apparent from this analysis that \( N_p \) is not uniformly less than \( N \), even with modest improvements in the values for \( \delta_i \) and \( \alpha_i \).

[0800] As with a conventional clinical trial, the incorporation of an appropriate control group in the study design is critical for achieving success. In the case of a prospective clinical trial, the control group commonly is selected on the basis of the same inclusion criteria as the treatment group, but is treated with placebo or a standard therapeutic regimen rather than the investigational drug. In the case of a study with subgroups that are defined by haplotype, the ideal control group for a treatment subgroup with haplotype “i” is a placebo-treated subgroup with haplotype “i”. This is often a critical control, since haplotypes which may be associated with the response to treatment may also affect the natural course of the disease.

[0801] A critical issue in considering control groups is that for the control group placebo treated population with haplotype “i” may not be equivalent to that of the control population. If so, 1.5 may overestimate the benefits of any reduction in as in the treatment response group if there is not also a reduction in as in the control group.

[0802] If \( \sigma \) of the treatment and control groups are not equivalent, \( \sigma \) would be still calculated as the difference in the response of the two groups, but car would be different in the two groups with values of \( \sigma_0 \) or \( \sigma_1 \) respectively. In this case, the number of patients in the genetically defined subgroup \( N_p \) would be defined by:

\[ N_p = \frac{(\sigma_0^2 + 2\beta_0^2) \delta^2}{\epsilon^2} \]
The total number of patients that would need to be enrolled in such a trial would be the maximum of $N$ or $N P\pi$.

It will be apparent that such an analysis remains sensitive to increases in $\delta$, but is less sensitive to changes in $\sigma$ which are not also reflected in the control group.

Certain analysis may be performed by comparing individuals with one haplotype against the entire normal population. Such an analysis may be used to establish the selectivity of the response associated with a specific haplotype. For example, it may be desirable to establish that the response or toxicity observed in a specific subgroup is greater than that associated observed with the entire population. It may also be of interest to compare the response to treatment between two different subgroups. If $\sigma$ differs between the groups, then the estimate of the number of patients that need to be enrolled in the trial should be calculated using equations 2.1 with $N$ being the maximum of $N/P$, for the different subgroups.

Another issue in controls is the relative size of the treatment and control groups. In a prospectively designed clinical trial, which selectively incorporates patients with haplotype “i”, the number of patients in the control and treatment group will be essentially equivalent. If the control group is different, or if haplotypes are used for stratification but not inclusion, statistical corrections may need to be made for having populations of different size.

Example 12

Stratification of Patients by Phenotype.

The identification of genetic associations in Phase II or retrospective studies can be performed by stratifying patients by phenotype and analyzing the distribution of genotypes/haplotypes in the separate populations. A particularly important aspect of this analysis is that any gene may have only a partial effect on the observed outcome, meaning that there will be an association value ($A$) corresponding to the fraction of patients in a phenotypically-defined subgroup who exhibit that phenotype due to a specific genotype/phenotype.

It will be recognized to those skilled in the art that the fraction of individuals who exhibit a phenotype due to any specific allele will be less than 1 (i.e. $A<1$). This is true for several reasons. The observed phenotype may occur by random chance. The observed phenotype may be associated with environmental influences, or the observed phenotype may be due to different genetic effects in different individuals. Furthermore, the construction of haplotypes and analysis of recombination may not group all alleles with phenotypically-significant variances within a single haplotype or haplotype cluster. In this case, causative variances at a single locus may be associated with more than one haplotype or haplotype cluster and the association constant $A$ for the locus would be $A=A_1+A_2+\ldots+A_n<1$. It is likely that many phenotypes will be associated with multiple alleles at a given locus, and it is particularly important that statistical methods be sufficiently robust to identify association with a locus even if $A_1$ is reduced by the presence of several causative alleles.

Statistical methods can be used to identify genetic effects on an observed outcome in patient groups stratified by phenotype, e.g. the presence or absence of the observed response. One such method entails determining the allele frequencies in two populations of patients stratified by an observed clinical outcome, for example efficacy or toxicity and performing a maximum likelihood analysis for the association between a given gene and the observed phenotype based on the allele frequencies and a range of values for $A$ (the association constant between a specific allele and the observed outcome used to stratify patients).

This analysis is performed by comparing the observed gene frequencies in a patient population with an observed outcome to gene frequencies in a table in which the predicted frequencies of different alleles of the gene assuming different values of the association constant $A$ for that allele. This table of predicted gene frequencies can be constructed by those skilled in the art based on the frequency of any specific allele in the normal population, the predicted inheritance of the effect (e.g. dominant or recessive) and the fraction of a subgroup with a specific outcome who would have that allele based on the association constant $A$.

For example, if a specific outcome was only observed in the presence of a specific allele of a gene, the expected frequency would be 1. If a specific outcome was never observed in the presence of a specific allele of a gene, the expected frequency would be 0. If there was no association between the allele and the observed outcome, the frequency of that allele among individuals with an observed outcome would be the same as in the general population. A statistical analysis can be performed to compare the observed allele frequencies with the predicted allele frequencies and determine the best fit or maximum likelihood of the association. For example, a chi square analysis will determine whether the observed outcome is statistically similar to predicted outcomes calculated for different modes of inheritance and different potential values of $A$. $P$ values can then be calculated to determine the likelihood that any specific association is statistically significant. A curve can be calculated based on different values of $A$, and the maximal likelihood of an association determined from the peak of such a curve. Methods for chi square analysis are known to those in the art.

A multidimensional analysis can also be performed to determine whether an observed outcome is associated with more than one allele at a specific genetic locus. An example of this analysis considering the potential effects of two different alleles of a single gene is shown. It will be apparent to those skilled in the art that this analysis can be extended to n dimensions using computer programs.

This analysis can be used to determine the maximum likelihood that one or more alleles at a given locus are associated with a specific clinical outcome.

It will be apparent to those skilled in the art that critical issues in this analysis include the fidelity of the phenotypic association and identification of a control group. In particular, it may be useful to perform an identical analysis in patients receiving a placebo to eliminate other forms of bias which may contribute to statistical errors.

Other Embodiments

The invention described herein provides a method for identifying patients with a risk of developing inflamma-
tory or immunological disease or dysfunction by determining the patients allele status for a gene listed in Tables 1 and 3 and providing a forecast of the patients ability to respond to a given drug treatment. In particular, the invention provides a method for determining, based on the presence or absence of a polymorphism, a patient’s likely response to drug therapies as inflammatory or immunological disease or dysfunction. Given the predictive value of the described polymorphisms across two different classes of drug, having different mechanisms of action, the candidate polymorphism is likely to have a similar predictive value for other drugs acting through other pharmacological mechanisms. Thus, the methods of the invention may be used to determine a patient’s response to other drugs including, without limitation, antihypertensives, anti-obesity, anti-hyperlipidemic, or anti-proliferative, antioxidants, or enhancers of terminal differentiation.

[0816] In addition, while determining the presence or absence of the candidate allele is a clear predictor determining the efficacy of a drug on a given patient, other allelic variants of reduced catalytic activity are envisioned as predicting drug efficacy using the methods described herein. In particular, the methods of the invention may be used to treat patients with any of the possible variances, e.g., as described in Table 3 of Stanton & Adams, U.S. application Ser. No. 09/300,747, supra.

[0817] In addition, while the methods described herein are preferably used for the treatment of human patient, non-human animals (e.g., pets and livestock) may also be treated using the methods of the invention.

[0818] All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains. All references cited in this disclosure are incorporated by reference to the same extent as if each reference had been incorporated by reference in its entirety individually.

[0819] One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The methods, variances, and compositions described herein as presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art, which are encompassed within the spirit of the invention, are defined by the scope of the claims.

[0820] It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. For example, using other compounds, and/or methods of administration are all within the scope of the present invention. Thus, such additional embodiments are within the scope of the present invention and the following claims.

[0821] The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms “comprising”, “consisting essentially of” and “consisting of” may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of the invention as defined by the appended claims.

[0822] In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.

Lengthy table referenced here
Please refer to the end of the specification for access instructions.
The patent application contains a lengthy table section. A copy of the table is available in electronic form from the USPTO web site (http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US20070037164A1). An electronic copy of the table will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 1
<210> SEQ ID NO 1
<211> LENGTH: 1643
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 1
17. An isolated nucleic acid molecule comprising 15 to 500 contiguous nucleotides of SEQ ID NO: 1 (tumor necrosis factor alpha) including a nucleotide selected from the group consisting of:

(a) nucleotide 87 of SEQ ID NO: 1 wherein G is replaced by T;

(b) nucleotide 779 of SEQ ID NO: 1 wherein A is replaced by C;

(c) nucleotide 1155 of SEQ ID NO: 1 wherein G is replaced by A; and

(d) nucleotide 1420 of SEQ ID NO: 1 wherein G is replaced by A.

18. The isolated nucleic acid molecule of claim 17, wherein the isolated nucleic acid molecule comprises fewer than 200 nucleotides of SEQ ID NO: 1.
23. The isolated nucleic acid molecule of claim 17, wherein the isolated nucleic acid comprises peptide nucleic acid (PNA).

24. The isolated nucleic acid molecule of claim 17, further comprising a detectable label.

25. The isolated nucleic acid of claim 24, wherein said detectable label is a fluorescent label.

26. The isolated nucleic acid molecule of claim 17 wherein the nucleic acid molecule includes at least two nucleotides selected from the group consisting of:

(a) nucleotide 87 of SEQ ID NO: 1 wherein G is replaced by T;
(b) nucleotide 779 of SEQ ID NO: 1 wherein A is replaced by C;
(c) nucleotide 1155 of SEQ ID NO: 1 wherein G is replaced by A; and
(d) nucleotide 1420 of SEQ ID NO: 1 wherein G is replaced by A.

* * * * *