

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/084133 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2006/003846

(22) International Filing Date: 1 February 2006 (01.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/648,927 1 February 2005 (01.02.2005) US
60/658,429 4 March 2005 (04.03.2005) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/648,927 (CIP)
Filed on	1 February 2005 (01.02.2005)
US	60/658,429 (CIP)
Filed on	4 March 2005 (04.03.2005)

(71) Applicant (for all designated States except US): **UNIVERSITY OF CHICAGO** [US/US]; 640 Ellis Avenue, 661N, Chicago, IL (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **RATAIN, Mark J.** **GRABER, Andrea.** **INNOCENTI, Federico.** **RAMIREZ, Jacqueline.**

(74) Agent: **SHISHIMA, Gina N.**; FULBRIGHT & JAWORSKI L.L.P., 600 Congress Avenue, Suite 2400, Austin, TX 78701 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR DOSING L-THYROXINE

(57) Abstract: The present invention is directed to methods and compositions for determining 5 the level of L-thyroxine glucuronidation in a patient. In some examples of the invention the number of TA repeats in the UGT1A1 gene is determined or evaluated to predict the level of activity. Such methods and compositions are used to optimize dosing for L-thyroxine, a common medication for the treatment of thyroid diseases, including hyperthyroidism.



WO 2006/084133 A2

METHODS AND COMPOSITIONS FOR DOSING L-THYROXINE

BACKGROUND OF THE INVENTION

This application claims priority to U.S. Provisional Patent Application Serial
5 Nos. 60/648,927, filed on February 1, 2005, and 60/658,429, filed on March 4, 2005,
which are hereby incorporated by reference in their entirety. The government may
own rights in the present invention pursuant to grant number GM61393 from the
National Institutes of Health.

1. Field of the Invention

10 The present invention relates generally to the fields of molecular genetics,
pharmacogenetics, and hormone therapy. In particular, the present invention is
directed to methods and compositions for evaluating how much L-thyroxine, a
particular drug which undergoes glucuronidation, is to be prescribed to a patient in
15 need of the drug. More specifically, the present invention is directed to methods and
compositions for achieving a euthyroid state in a patient who suffers from
hyperthyroidism.

2. Description of Related Art

20 Thyroxine (T₄) is prescribed in patients with hypothyroidism, and is known to
undergo glucuronidation. Previous studies propose 1A1 and 1A9 as the main UGT
isoforms responsible for T₄G formation, but a complete 1A screening has not been
performed.

Hyperthyroidism occurs in approximately 5% of the US population (Hollowell
25 *et al.*, 2002). L-thyroxine is the most widely used treatment for hypothyroidism and
the third prescribed drug in USA in 2003. Titration of L-thyroxine dosage in
hypothyroid patients is required to prevent continued hypothyroidism or iatrogenic
hyperthyroidism which can lead to cardiac dysfunction (Klein *et al.*, 2001). There is a
high variability in the L-thyroxine dosage required to achieve a euthyroid state.

30 L-thyroxine is inactivated by metabolism via glucuronidation in the liver.
However, it was not known whether specific UGT polymorphisms affected the
glucuronidation of L-thyroxine. Therefore, improved methods and compositions for
optimizing dosing of L-thyroxine are still needed.

SUMMARY OF THE INVENTION

The present invention is based on identification and characterization of correlations between the genotype of the *UGT1A1* gene and phenotype relating to the inactivation of L-thyroxine. The inventors have identified the UGT enzymes involved in the hepatic and extrahepatic metabolism of L-thyroxine. Using liver specimens from donors, the inventors have demonstrated a significant correlation between a common polymorphism in the *UGT1A1* gene (a TA indel in the gene promoter, also referred to as (TA)_n) and L-thyroxine glucuronidation rates. Significant differences between *UGT1A1* -53(TA)_{6>7} genotypes 6/6, 6/7 and 7/7 were observed (ANOVA, $P=0.003$, $n=54$). In a post hoc Tukey multiple comparison test, 7/7 showed significantly lower T₄G formation than 6/6 ($P<0.01$). There was not a significant difference between the other two pairs. There was a significant trend in T₄G formation between *UGT1A1* -53(TA)_{6>7} genotypes 6/6, 6/7, and 7/7 with 6/6 > 6/7 > 7/7 (trend analysis $P=0.002$, $n=54$). Hence, genetic differences in the way L-thyroxine is metabolized can affect the efficacy of L-thyroxine in hypothyroid patients. Thus, the present invention provides methods and compositions that exploit correlations between genotype and phenotype concerning L-thyroxine. It is contemplated that such methods and compositions can be used to optimize dosing in patients. Moreover, such methods and compositions can prevent or reduce the risk of iatrogenic hyperthyroidism and/or its side effects.

The present invention involves methods for determining or evaluating the level of L-thyroxine glucuronidation activity in a patient. This is accomplished by directly or indirectly determining the number of TA repeats in the promoter of the *UGT1A1* promoter. Additionally, the claims provide a summary of the invention. It will be understood that the term “determine” is used according to its ordinary and plain meaning to indicate “to ascertain definitely by observation, examination, calculation, etc.,” according to the Oxford English Dictionary (2nd ed.). It will also be understood that the phrase “determining the sequence at position X” means that the nucleotide(s) at that position is directly or indirectly identified. In some embodiments, the sequence at a particular position is determined, while in other embodiments, what is determined at a particular position is that a particular nucleotide is *not* at that position. It is specifically contemplated that embodiments of the invention include methods in which a medical practitioner determines information regarding a patient’s

genotype by performing a test or assay or ordering that the patient have the assay or test performed, such as by a laboratory. Therefore, it is contemplated that a person may “determine” information by receiving the results of an assay or test revealing the results of a determination.

5 Based on the level of activity determined, it is contemplated that a patient is given a different dosage than he or she would have otherwise received had the genotyping not been performed. Thus, in some embodiments of the invention, a typical dosage is adjusted for a particular person (individualized therapy).

10 Moreover, it is contemplated that some methods of the invention are practiced by a laboratory and include, but are not limited to, the following steps: receiving or obtaining a biological sample from a patient; assaying the sample for one or more polymorphisms (such as are described herein); providing the results of the assay to a medical practitioner of the patient's. It is contemplated that the results may be provided in the form of a report or lab result.

15 Certain methods of the invention concern prescribing an amount of L-thyroxine for a patient comprising: a) determining the number of TA repeats in one or both alleles of the *uridine diphosphate glucuronosyltransferase 1A1* (*UGT1A1*) promoter, wherein the presence of seven TA repeats correlates with decreased inactivation of L-thyroxine; and, b) prescribing an average daily dose of L-thyroxine
20 for the patient. The term “average daily dose” means an amount of L-thyroxine that is the average amount of L-thyroxine on a daily basis. Thus, it is contemplated that a patient may be prescribed X amount to be taken over a three day period, which amounts to X/3 as the average daily dose over that period even though the patient may have taken 0.5X on day 1 and 0.5X on day 3. Furthermore, in some embodiments an
25 average daily dose can be expressed as a total amount to be taken in a day or as one or more amounts that are to be taken a certain number of times a day. For example, it is contemplated that a patient is prescribed L-thyroxine having a particular dose, such as 50 mcg, to be taken twice a day (b.i.d.). Typically, however, L-thyroxine is taken once a day, so any dose prescribed is the total amount for the day.

30 In other embodiments of the invention, there are methods for reducing the risk of hypothyroidism or hyperthyroidism in a patient comprising: a) predicting the level of L-thyroxine glucuronidation in the patient; and, b) prescribing an average daily dose of L-thyroxine to the patient based on the level of L-thyroxine glucuronidation predicted in the patient. The risk of either hyperthyroidism or hypothyroidism is

reduced by prescribing a therapeutically appropriate amount of L-thyroxine to address the patient's thyroid problem based on the predicted level of L-thyroxine glucuronidation in that patient.

The present invention also concerns methods for evaluating a patient's L-thyroxine glucuronidation capacity comprising: a) obtaining DNA from the patient; b) determining from the DNA the number of TA repeats in one or both alleles of the *UGT1A1* promoter; and, c) informing the patient or the patient's physician about the patient's L-thyroxine glucuronidation capacity. It is contemplated that either instead of c) or as a way of achieving c), a report or other document providing the results from steps a) and b) is generated. The report or document can be provided to the patient and/or the patient's caregiver (physician or nurse).

In further embodiments of the invention there are methods for optimizing the dose of L-thyroxine for a patient comprising: a) determining from the patient's DNA the number of TA repeats in one or both alleles of the *UGT1A1* promoter; and b) prescribing a dose of L-thyroxine to the patient based on the number of TA repeats.

The level of L-thyroxine glucuronidation is predicted by determining the number of TA repeats in one or both alleles of the *UGT1A1* promoter in some embodiments of the invention. The presence of seven TA repeats in a single allele indicates the same or a lower level of L-thyroxine glucuronidation than the presence of six TA repeats in a single allele. However, the presence of six TA repeats in one allele and seven TA repeats in the other allele indicates a lower level than the presence of six TA repeats in both alleles, but a higher level than the presence of seven TA repeats in both alleles. In some embodiments, only one allele is or can be evaluated, while in other embodiments, at least one allele or both alleles is or can be evaluated.

The number of TA repeats correlates with the level of inactivation of L-thyroxine inactivation as a result of glucuronidation by the gene product of *UGT1A1*. Consequently, patients having at least one *UGT1A1* allele with seven TA repeats in the promoter region generally have a lower level of L-thyroxine inactivation than patients who do not have a (TA)₇. Consequently, the highest doses will generally be prescribed for those with six TA repeats in both alleles, the lowest doses will generally be prescribed for those with seven TA repeats in both alleles, and doses between the highest and lowest will generally be prescribed to those with six TA repeats in one allele and seven TA repeats in the other allele. Based on these averages,

the dose prescribed for the patient will be lower if the patient is found to have seven TA repeats in at least one allele than the dose prescribed for a patient who has six TA repeats in both alleles. Conversely, in further embodiments, the dose prescribed for the patient will be higher if the patient has six TA repeats in one allele than the dose prescribed for a patient who has a seven TA repeats in both alleles. Moreover, the dose prescribed for the patient will be lower if the patient has seven TA repeats in both alleles than the dose prescribed for a patient who has six TA repeats in one allele and seven TA repeats in the other allele. Also, the dose prescribed for the patient will be higher if the patient has six TA repeats in both alleles than the dose prescribed for a patient who has six TA repeats in one allele and seven TA repeats in the other allele. A person of ordinary skill in the art would be able to adjust the amount typically prescribed in the absence of evaluating the number of TA repeats. It is contemplated that the dose may be adjusted by about, at least about or at most about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300% or more, or any range derivable therein. While the word "dose" refers to a single administration of L-thyroxine, it is contemplated that adjustments may be made with respect to the frequency of dosing

In certain cases, methods include determining the presence of seven TA repeats in at least one allele. In certain embodiments, methods include determining the presence of six TA repeats in at least one allele. In other embodiments, methods including determining the presence of six or seven TA repeats in both alleles or six repeats in one and seven repeats in the other. In some cases, methods include prescribing a dose of L-thyroxine for a patient having six TA repeats in both alleles, wherein the dose is higher than the dose for a patient having seven TA repeats in one or both alleles. Other methods involve prescribing a dose of L-thyroxine for a patient having seven TA repeats in both alleles, wherein the dose is lower than the dose for a patient having six TA repeats in one or both alleles. Additional methods include prescribing a dose of L-thyroxine for a patient having six TA repeats in one allele and seven TA repeats in the other allele, wherein the dose is lower than the dose for a patient having seven TA repeats in both alleles but higher than the dose for a patient having having six TA repeats in both alleles.

It is contemplated that a dose may be adjusted with respect to the amount of L-thyroxine that a patient takes at any one time and/or the frequency of the intake. In certain instances the dose will be adjusted with respect to the total amount a patient takes a day (24-hour period) (referred to as average daily dose), while in others, the adjustment is with respect to the total amount a patient takes in a week (seven 24-hours periods) (average weekly dose). It is contemplated that any embodiment discussed with respect to "dose" can be implemented with respect to "average daily dose" and "average weekly dose." In other words, if embodiments concern a first patient who is prescribed a lower "dose" than a second patient with a different genotype, then the embodiment will be understood to describe also a first patient who is prescribed a lower "average daily dose" or an "average weekly dose" than the second patient with a different genotype.

In even further instances, only the frequency of intake will be adjusted, while in others, both frequency and total intake is adjusted. It is contemplated that any embodiment discussed with respect to "dose" can be implemented with respect to frequency of intake. For example, if embodiments concern a first patient who is prescribed a lower "dose" than a second patient with a different genotype, then the embodiment will be understood to describe also a first patient who is prescribed to intake L-thyroxine less frequently than the second patient with a different genotype. Frequency can be discussed in terms of any measurement of time, but in specific embodiments, the measurements are expressed in terms of number of times a day (*e.g.*, one, two, three, four, five, or more times) or number of times a week (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more times). It is contemplated that adjustments can be such that a patient takes L-thyroxine less frequently or more frequently than someone else with a different genotype.

The invention is not limited by the way in which the number of TA repeats in one or both alleles is determined. In some embodiments, the number of TA repeats is determined by evaluating directly the TA repeats in one or both alleles of the *UGT1A1* promoter. This can be accomplished, for example, by sequencing one or both alleles at that location in the gene or by amplifying the number of TA repeats. The number of TA repeats can be amplified by polymerase chain reaction. Alternatively, the number of TA repeats can be determined using an invasive cleavage assay. In certain embodiments, the number of TA repeats is determined by using gel

electrophoresis as part of the assay. In certain embodiments, it is contemplated that only one allele is evaluated, while in other embodiments, both alleles are evaluated.

In certain methods of the invention, embodiments further include obtaining DNA from the patient. This may be accomplished by obtaining any biological sample
5 from the patient so long as the sample contains DNA from the patient.

Other aspects of invention include methods in which other UGT enzyme polymorphisms that affect glucuronidation of L-thyroxine are evaluated. The polymorphism may be in the gene of any other UGT enzyme. In some embodiments, the polymorphism is in the *UGT1A* gene. In other embodiments, the polymorphism is
10 in the *UGT1A3* gene. The *UGT1A3* polymorphism may be one or more of the following polymorphisms: 17(A>G), 31(T>C), 81(G>A), 133(C>T), 140(T>C), or 477(A>G), or a combination thereof. In still further embodiments, the polymorphism is in the *UGT1A8* gene. The *UGT1A8* polymorphism may be one or more of the following polymorphisms: 518(C>G), 765(A>G), 830(G>A), or a combination
15 thereof. It is contemplated that the polymorphism is selected from the group consisting of: *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T. In other embodiments, the polymorphism is in the *UGT1A10* gene. The *UGT1A10* polymorphism may be one or more of the following polymorphisms: 126(G>A), 177(G>A), 415(G>A), 597(T>C), 605(C>T), 622 (T>C), 693(C>T), 719(C>T),
20 730(C>A), or a combination thereof. It is contemplated that a combination of polymorphisms or a specific haplotype may be evaluated in certain embodiments of the invention.

Methods of the present invention involve determining the number of TA repeats in the promoter region of the *UGT1A1* gene by evaluating a polymorphism in
25 linkage disequilibrium (LD) with the number of repeats. In certain embodiments, this involves evaluating at least one of the following polymorphisms: *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A. Other polymorphisms in *UGT1A1* may also be included in embodiments of the invention. The term "linkage disequilibrium" is used according to its ordinary meaning in the art.
30 In certain embodiments, the polymorphism is in complete LD with the number of TA repeats.

Other embodiments also involve identifying a patient in need of L-thyroxine. Methods and compositions of the present invention are of particular interest to such patients. In certain embodiments, the patient does not have or is not suspected of

having Gilbert's Syndrome. Gilbert's Syndrome is correlated with the number of repeats, however, it is not clear that the presence or absence of clinical Gilbert's Syndrome indicates a particular phenotype with respect to a particular UGT1A1 drug substrate.

5 Methods also include identifying a starting daily dose or starting average daily dose for a patient in need of L-thyroxine. A "starting daily dose" refers to the amount of L-thyroxine a patient is initially prescribed on a daily basis after the prescriber takes into consideration the number of TA repeats in one or both alleles of the patient's *UGT1A1* genes. A "starting average daily dose" is the same as a daily dose
10 except that it is expressed as the average amount that is take as an initial dose. It is contemplated that the dose prescribed for a patient may subsequently be altered depending on the patient's health, the prognosis of the patient's thyroid conditions, any side effects, etc.

 Prior to the invention, dosages were altered as a result of titrating the amount
15 of L-thyroxine in the patient. In some embodiments, it is contemplated that further refinements may involve titrating the amount of L-thyroxine in the patient. Such refinements may be implemented to address issues concerning specific patients.

 Moreover, it is contemplated that such methods may be implemented with respect to multiple patients. In certain embodiments, DNA from multiple patients is
20 obtained and the number of TA repeats determined for each patient. Assays may be carried out concurrently for multiple samples (*i.e.*, patients). In addition, assays may be fully or partially automated.

 Methods of the invention also include monitoring for toxicity or adverse events once L-thyroxine is administered, and possibly, adjusting or modifying dosage
25 based on those results.

 Reference to nucleotides (or residues) may be according to their well known abbreviations. A "C" refers to a cytosine; "T" refers to "thymine"; "A" refers to adenine; and, "G" refers to guanine. If mRNA is used to determine a nucleotide sequence, "U" refers to uracil.

30 Positions are indicated by conventional numbering where a negative sign (-) refers to nucleotides upstream (5') from the transcriptional start site (+1) (these sequences are in the promoter), unless otherwise designated. Sequences in the 5' untranslated region (5' UTR) may also be referred to using a negative sign, and in these cases, the positioning is with respect to the translated portion, where the first

nucleotide of a codon is understood as +1. Positions downstream of the translational start site may or may not have a plus sign (+). Furthermore, unless otherwise indicated or understood, identification of a position downstream of the transcriptional start site refers to a position with respect to only the coding region of the gene, that is, its exons and not the introns. In some instances, positions within introns are referred to and the numbering for these positions is typically with respect to that intron alone, and not the gene as a whole.

It is contemplated that in methods of the invention, one or more sequences in one or both alleles of the *UGT1A1* gene is determined. In some embodiments, both alleles of the patient are evaluated, while in others, only one allele is evaluated.

In further embodiments of the invention, methods also include obtaining a sample from a patient and using the sample to determine the sequence at a particular position or the number of repeats at a particular position. The sample may contain whole blood, serum, plasma, or a tissue biopsy.

Sequences may be determined by performing or conducting a hybridization assay, an amplification assay, particularly one that is allele-specific, a sequencing or microsequencing assay.

The number of TA repeats in the *UGT1A1* gene may be determined directly or indirectly. A direct determination involves performing an assay with respect to that position. An indirect determination means that the sequence or number of repeats at that position is determined based on data regarding a different position, particularly by evaluating the sequence of a position in linkage disequilibrium (LD) with that polymorphism. In some embodiments, the sequence in LD with the number of TA repeats is in complete linkage disequilibrium with another polymorphism. In additional embodiments, the position in linkage disequilibrium with the number of TA repeats is selected from the group consisting of positions -3156G>A, -3279T>G, -3156G>A, 211G>A, and/or 686C>A of the *UGT1A1* gene. In some cases, more than one position in linkage disequilibrium with the TA repeat is evaluated to determine the sequence at position -53. Therefore, in some embodiments of the invention, a haplotype that includes the number of TA repeats is evaluated. In these embodiments, a determination of one or more sequences in one or both alleles of a gene in the haplotype is included in methods of the invention.

In methods of the invention, in some embodiments, an additional step of administering L-thyroxine to the patient is included. In some cases, the amount,

formulation, or timing of the administration is based on the analysis of the number of TA repeats or the level of L-thyroxine glucuronidation capacity. In some embodiments of the invention, a patient is also provided additional thyroid therapy.

5 The invention also concerns kits. Kits can include reagents for detecting the number of TA repeats and one or more of the polymorphisms discussed above and elsewhere in the application. Such reagents may be nucleic acids, for example, sequencing primers, amplification primers, or probes. Alternatively, the reagents may be compatible with specific cleavage assays.

10 It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein. Similarly, any embodiment discussed with respect to one aspect of the invention may be used in the context of any other aspect of the invention.

15 Throughout this application, the term “about” is used to indicate that a value includes the standard deviation of error for the device or method being employed to determine the value.

The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims and/or the specification may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.”

20 The use of the term “or” in the claims is used to mean “and/or” unless explicitly indicated to refer to alternatives only or the alternative are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and “and/or.”

25 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

30

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be

better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

FIG. 1: *UGT1A1* (TA)_n Genotype versus L-thyroxine glucuronidation.

5 **FIG. 2:** Screen of UGTs for T₄G formation. All of the *UGT1A* activities were normalized by relative protein expression as determined by western blot.

FIG. 3A-B. Correlation of T₄ with SN-38 (A) and Flavopiridol glucuronidation (B). Data represent means of duplicate determinations from a single experiment.

10 **FIG. 4.** Comparison of T₄G formation between *UGT1A1* (TA)₆₋₇ genotypes. Data represent means of duplicate determinations from a single experiment. Kruskal-Wallis test $p < 0.01$.

FIG. 5. Comparison of T₄G formation between *UGT1A9* -118T>10 genotypes. Data represent means of duplicate determinations from a single
15 experiment. Kruskal-Wallis test $p > 0.05$.

FIG. 6. Michaelis-Menten kinetics of T₄ glucuronidation in HLM. Data represent the mean $\pm$ SE of a single experiment performed in triplicate.

FIG. 7A-B. Screen of recombinant microsomes expressing functional human UGTs for T₄G formation (A) non-normalized (B) normalized by relative protein
20 expression. Data represent means of duplicate determinations from a single experiment. UGT1A activities were normalized for relative protein expression by western blot. The UGT1A standard was designated as the basal UGT expression level.

FIG. 8A-B. Correlation of T₄ glucuronidation with (A) SN-38 and (B)
25 flavopiridol glucuronidation. Data represent means of duplicate determinations from a single experiment.

FIG. 9. Comparison of T₄G formation between *UGT1A1* -53(TA)_{6>7} genotypes. Data represent means of duplicate determinations from a single
30 experiment.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention provides methods and compositions for evaluating the level of L-thyroxine glucuronidation in a subject who will be administered L-thyroxine (or has already been administered L-thyroxine). The development of these methods and compositions allows for the use of such an evaluation to optimize treatment of a patient and to lower the risk of toxicity or adverse events.

I. L-THYROXINE

L-thyroxine (T_4) is the generic name for a commonly prescribed drug to treat hypothyroidism. Hypothyroidism is a condition caused by the effects of too little thyroid hormone on tissues of the body. L-thyroxine may also be used to treat Pituitary TSH Suppression. L-thyroxine is the most widely used treatment for hypothyroidism, a condition which occurs in approximately 5% of the U.S. population (Hollowell et al., 2002). Due to a low therapeutic index, T_4 dosage requires titration in small increments (Roberts and Ladenson, 2004). Interindividual differences in T_4 metabolism could result in variation in the patient exposure to T_4 . This variation could potentially affect the steady state dose required to reach a euthyroid state.

T_4 is metabolized by deiodination, sulfation, and glucuronidation. Deiodination acts to regulate the bioactivity of the molecule. Outer ring deiodination converts T_4 to the biologically active form, 3,3',5-triiodothyronine (T_3). Deiodination of the inner ring produces the biologically inactive 3,3',5'-triiodothyronine (rT_3). Conjugation of T_4 with a glucuronide or sulfate group increases urinary and biliary clearance by rendering the molecule more water-soluble. T_4 glucuronide (T_4G) is stable and may undergo enterohepatic recirculation through conversion to T_4 by bacterial β -glucuronidase in the intestine (Visser, 1996). T_4 sulfate is very unstable because sulfation accelerates inner ring deiodination by approximately 200-fold. Thus, little T_4 sulfate is found in bile, urine, or serum.

Deiodination is widely considered to be the major route of T_4 metabolism. However, mass balance studies including conjugates are not currently available. On the other hand, animal studies suggest that T_4 glucuronidation may significantly affect T_4 clearance. Gunn rats, which are UGT1A deficient, have circulating T_4 concentrations that are 1.8 fold higher than those in normal rats, suggesting that decreased UGT activity leads to higher T_4 levels (Kato et al., 2005). In addition,

several other studies have demonstrated that the administration of hepatic inducers of UGTs leads to a dose-dependent decrease in serum T₄ concentrations in rats (Beetstra et al., 1991; Liu et al., 1995; Hood et al., 1999). One study has also demonstrated a reduction in serum T₄ concentrations in Gunn rats treated with phenobarbital, suggesting that other inducible proteins may also mediate T₄ clearance (Kato et al., 2005). Treatment with rifampin (a potent UGT1A1 inducer) also results in increased T₄ clearance in humans (Ohnhaus et al., 1981).

Concerning T₃ metabolism, previous studies have demonstrated that T₃ undergoes glucuronidation by UGT2B2 in rats (Visser et al., 1993b). However, no significant T₃ glucuronidation has been observed in human liver or kidney microsomes, suggesting that glucuronidation is a minor pathway of T₃ metabolism in humans (Findlay et al., 2000).

Many studies have demonstrated considerable interindividual variation in UGT activities for specific drugs. In many cases, this variation has been correlated with polymorphisms in UGT genes (Burchell, 2003). The availability of T₄ for conversion to T₃ might depend upon the rate of T₄ metabolism after oral administration. Identifying the UGTs responsible for T₄ glucuronidation and investigating the effects of polymorphisms within the genes coding for these UGTs is an important step in determining the role of the glucuronidation pathway in T₄ metabolism.

Previous studies in rats suggested that both UGT1A1 and UGT1A9 are involved in T₄ glucuronidation (Visser et al., 1993b). Comparisons of T₄ glucuronidation activity in normal human liver microsomes (HLM), Crigler-Najjar HLM (where UGT1A1 is absent), and human kidney microsomes (which highly express UGT1A9) indicated that UGT1A1 plays a greater role in human T₄ glucuronidation than UGT1A9 (Findlay et al., 2000). Screens of cell lines expressing human UGT1A1 and UGT1A9 have also demonstrated higher T₄G formation with UGT1A1 than UGT1A9 (Visser et al., 1993a; Findlay et al., 2000). However, a complete screen of all of the functional human UGTs had not been previously reported. Moreover, the effect of common UGT1A polymorphisms on the rate of T₄ glucuronidation had never been investigated.

L-thyroxine is sold in the United States under a variety of names including, but not limited to, Synthroid (levothyroxine sodium tablets), Levotheroid, Levoxyl, Unithroid, and Levo-T.

It is typically taken once a day. It is contemplated that the daily average dose of L-thyroxine may be adjusted after the level of a patient is evaluated using methods and compositions of the invention.

Hepatic glucuronidation results from the activities of a multigene family of UGT enzymes, the members of which exhibit specificity for a variety of endogenous substrates and xenobiotics. The UGT enzymes are broadly classified into two distinct gene families. The UGT1 locus codes for multiple isoforms of UGT, all of which share a C-terminus encoded by a unique set of exons 2-5, but which have a variable N-terminus encoded by different first exons, each with its own independent promoter (Bosma *et al.*, 1992; Ritter *et al.*, 1992). The variable first exons confer the substrate specificity on the enzyme. Isoforms of the UGT2 family are unique gene products of which at least eight isozymes have been identified (Clarke *et al.*, 1994). The UGT1A1 isoform is the major bilirubin glucuronidation enzyme. Genetic defects in the UGT1A1 gene can result in decreased glucuronidation activity which leads to abnormally high levels of unconjugated serum bilirubin that may enter the brain and cause encephalopathy and kernicterus (Owens & Ritter, 1995). As described above, this condition is commonly known as Gilbert's syndrome (which is frequently diagnosed based on elevated total bilirubin levels—a biochemical diagnosis). The molecular defect in Gilbert's Syndrome is a change in the TATA box within the UGT1A1 promoter (Bosma *et al.*, 1995; Monaghan *et al.*, 1996). This promoter usually contains a (TA)₆ TAA element, but another allele, termed UGT1A1*28 or allele 7, is also present in human populations at high frequencies, and contains the sequence (TA)₇ TAA. This polymorphism in the promoter of the UGT1A1 gene results in reduced expression of the gene and accounts for most cases of Gilbert's Syndrome (Bosma *et al.*, 1995). As discussed below, overall, gene expression levels for the UGT1A1 promoter alleles are inversely related to the length of the TA repeat in the TATA box.

The variation observed in this promoter may also account for the inter-individual and inter-ethnic variation in drug metabolism and response to xenobiotic exposure. UGTs have been shown to contribute to the detoxification and elimination of both exogenous and endogenous compounds. Examples of how UGT polymorphisms affect drugs other than L-thyroxine are provided in, for example, U.S. Patent Nos. 6,472,157 and 6,395,481, which are both incorporated by reference with respect to their teaching about *UGT1A1* sequences.

II. NUCLEIC ACIDS

The TA repeats in the *UGT1A1* gene of humans have been reported as GenBank Accession numbers AY533181 (eight TA repeats) (SEQ ID NO:4);
 5 AY533180 (seven TA repeats) (SEQ ID NO:3); AY533179 (six TA repeats) (SEQ ID NO:2); and AY533178 (five TA repeats) (SEQ ID NO:1), which are all incorporated by reference. The sequence of *UGT1A9* containing the polymorphisms discussed herein can be found at GenBank Accession No. AF297093.1, which is hereby incorporated by reference.

10 Certain embodiments of the present invention concern various nucleic acids, including amplification primers, oligonucleotide probes, and other nucleic acid elements involved in the analysis of genomic DNA. In certain aspects, a nucleic acid comprises a wild-type, a mutant, or a polymorphic nucleic acid.

The term "nucleic acid" is well known in the art. A "nucleic acid" as used
 15 herein will generally refer to a molecule (*i.e.*, a strand) of DNA, RNA or a derivative or analog thereof, comprising a nucleobase. A nucleobase includes, for example, a naturally occurring purine or pyrimidine base found in DNA (*e.g.*, an adenine "A," a guanine "G," a thymine "T" or a cytosine "C") or RNA (*e.g.*, an A, a G, an uracil "U" or a C). The term "nucleic acid" encompasses the terms "oligonucleotide" and
 20 "polynucleotide," each as a subgenus of the term "nucleic acid." The term "oligonucleotide" refers to a molecule of between about 3 and about 100 nucleobases in length. The term "polynucleotide" refers to at least one molecule of greater than about 100 nucleobases in length. A "gene" refers to coding sequence of a gene product, as well as introns and the promoter of the gene product. In addition to the
 25 *UGT1A1* gene, other regulatory regions such as enhancers for *UGT1A1* are contemplated as nucleic acids for use with compositions and methods of the claimed invention.

In some embodiments, nucleic acids of the invention comprise or are complementary to all or 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,
 30 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370,

380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 or more contiguous
5 nucleotides of the human *UGT1A1* gene or other UGT gene discussed herein.

These definitions generally refer to a single-stranded molecule, but in specific embodiments will also encompass an additional strand that is partially, substantially or fully complementary to the single-stranded molecule. Thus, a nucleic acid may encompass a double-stranded molecule or a triple-stranded molecule that comprises
10 one or more complementary strand(s) or "complement(s)" of a particular sequence comprising a molecule. As used herein, a single stranded nucleic acid may be denoted by the prefix "ss", a double stranded nucleic acid by the prefix "ds", and a triple stranded nucleic acid by the prefix "ts."

In particular aspects, a nucleic acid encodes a protein, polypeptide, or peptide.
15 In certain embodiments, the present invention concerns novel compositions comprising at least one proteinaceous molecule. As used herein, a "proteinaceous molecule," "proteinaceous composition," "proteinaceous compound," "proteinaceous chain," or "proteinaceous material" generally refers, but is not limited to, a protein of greater than about 200 amino acids or the full length endogenous sequence translated
20 from a gene; a polypeptide of greater than about 100 amino acids; and/or a peptide of from about 3 to about 100 amino acids. All the "proteinaceous" terms described above may be used interchangeably herein.

A. UGT Nucleic Acids

As discussed above, embodiments of the invention concern evaluating a
25 patient's *UGT1A1* genes and determining how many TA repeats the patient has in each allele. In addition, other polymorphisms in *UGT1A1* may be evaluated and/or polymorphisms in other UGT genes may be evaluated. These include the following polymorphisms discussed herein.

Other *UGT1A1* polymorphisms are discussed in published PCT application
30 number WO 04/108954, which is hereby incorporated by reference in its entirety. In particular embodiments, the polymorphism is the following: -3440C>A, -3401T>C, -3279G>T, -3177C>G, -3175A>G, -3156G>A, or any combination thereof. The notation -3440C>A, for example indicates that cytosine nucleotide (C) at the -3440 position is replaced by an Adenosine (A), where the negative sign indicates what

residue upstream from the mRNA start site (+1) the residue is. In additional embodiments, a UGT1A1 polymorphism is one or more of the following: -3401T>C, -3345C>-, -3279T>G (originally reported as T>G but G is the ancestral allele), -3177C>G, -3175G>A, -3156G>A, -2951A>G, -2828G>C, -2473G>T, -2353C>T, -1516T>G, -1352A>C, -484G>A, -364C>T, -129C>A, -106T>C, -53(TA)⁵⁻⁸, 141C>T, 211G>A, 674T>G, 686C>A, 924C>T, 953C>T, 1206C>T, 1395A>G, 3705G/T, 3726A>G, 3729G/T, 3788T>G, 4254C>T, 4305G>A, 4375C>T, 4528T>C, 4578A/AAAAGGGAGGGA, 4654G>C, 4754C>G>T, 4969A>G, 6339T>G, 6601C>T, 6633C>T, 6895C>T, 7023T>G, 7184A>G, 7262T>C, 7462T/TGT, 7478C>T, 7479T>C, 7524T>C, 7526C>A, 7853G>A, 7872G>A, 8045A>C, 11514CCCT>-, 11691G>A, 11745A>C, 11800G>C, 12125T>G, 12318T>C, 12482C>T, 12610C>G, 12711C>G, 13122C>A, 13128C>T, 13146A>G, 13185G>A, 17208C>T, 17223C>T, 17410T>A, 17497T>C, 17551C>T, 17667C>T, 17764G>T, 17901T>A, 17974A>G, 18184C>T, 18196C>A, 18223T>C, 18275G>A, 18294A>G, 18420A>G, 18484T>A, 18640T>G, 18649T>C, 18994G>A, 19029A>C, 19056G>C, 19102A>C, 19393A>G, 19408T>C, 19719A>G, 19720T>C, or a combination thereof. See on the World Wide Web pharmgkb.org and Genbank Accession number AF297093, which is hereby incorporated by reference.

The present invention in some embodiments also concerns other UGT genes, such as UGT1A3, UGT1A8, UGT1A9, and UGT1A10. The following tables describe polymorphisms, genotypes, and/or haplotypes that may also be evaluated in the context of the invention. It is contemplated that one or more of the polymorphisms, genotypes, or haplotypes identified below may be evaluated as part of methods of the invention or that kits of the invention will include a nucleic acid(s) capable of detecting one or more of such polymorphisms, genotypes, or haplotypes. Moreover, the GenBank or other accession numbers provided in the tables below are specifically incorporated by reference.

TABLE 1

UGT1A3							
Allele naming	Protein	Nucleotide change	Amino acid change	Effect	Enzyme activity		References
					<i>In vivo</i>	<i>In vitro</i>	
UGT1A3*1	UGT1A3.1			Wild-type	Normal	Normal	NM_019093
UGT1A3*2a	UGT1A3.2	31(T>C)/81(G>A)/140(T>C)	W11R/E27E/V47A				AF465194 Ehmer U, 2004
UGT1A3*2b	UGT1A3.2	31(T>C)/81(G>A)/140(T>C)/477(A>G)	W11R/E27E/V47A/A159A			High activity (Iwai et al., 2004)	AB120361 Iwai M, 2004
UGT1A3*3a	UGT1A3.3	31(T>C)/81(G>A)	W11R/E27E				AF465193 Ehmer U, 2004
UGT1A3*3b	UGT1A3.3	31(T>C)/81(G>A)/477(A>G)	W11R/E27E/A159A				AB120359 Iwai M, 2004
UGT1A3*4	UGT1A3.4	133(C>T)	R45W				AB120362 Iwai M, 2004
UGT1A3*5	UGT1A3.5	17(A>G)/31(T>C)/81(G>A)/477(A>G)	Q6R/W11R/E27E/A159A				AB120360 Iwai M, 2004

Base change	Amino acids change	SNP frequency (n=200, Japanese)	Mutations	Reference
31(T>C)	W11R	0.28	Nonsynonymous	Ehmer et al., 2004 Iwai et al., 2004
81(G>A)	E27E	0.28	Silent	
140(T>C)	V47A	0.125	Nonsynonymous	Iwai et al., 2004
17(A>G)	Q6R	0.055	Nonsynonymous	
133(C>T)	R45W	0.11	Nonsynonymous	
447(A>G)	A159A	0.28	Silent	

5 * W11R, E27E and V47A were detected in 65%, 65% and 58% of blood donors (normal controls) (n=81, German Caucasians, Ehmer et al., 2004)

TABLE 2

UGT1A8									
Allele naming	Protein	Nucleotide change	Amino acid change	Effect	Enzyme activity		Genbank	References	Notes
					<i>In vivo</i>	<i>In vitro</i>			
UGT1A8*1a	UGT1A8.1			Wild-type	Normal	Normal	AF462268		
UGT1A8*1b	UGT1A8.1	765(A>G)	T255T	Silent			N/A	Huang et al. 2002	Renamed from *1a
UGT1A8*2	UGT1A8.2	518(C>G)	A173G			Similar to *1 (Huang et al. 2002; Bernard et al. 2004; Thibaudau et al., 2006)	AF462267	Huang et al. 2002; Bernard et al. 2004; Thibaudau et al., 2006	
UGT1A8*3	UGT1A8.3	830(G>A)	C277Y			Severely reduced (Huang et al. 2002; Bernard et al. 2004; Thibaudau et al., 2006)	N/A	Huang et al. 2002; Bernard et al. 2004; Thibaudau et al., 2006	

5 N/A : Data not available in the manuscript.

TABLE 3 – UGT1A10

Allele naming	Protein	Nucleotide change	Amino acid change	Effect	Enzyme activity		Genbank	Reference	Notes
					In vivo	In vitro			
UGT1A10*1a	UGT1A10.1			Wild-type		Normal	U39550, AF297093		
UGT1A10*1b	UGT1A10.1	126(G>A)	Q42Q				N/A	Elahi et al., 2003	Renamed from UGT1A10*1a
UGT1A10*1c	UGT1A10.1	597(T>C)	H199H				N/A	Elahi et al., 2003	Renamed from UGT1A10*1b
UGT1A10*1d	UGT1A10.1	597(T>C)/693(C>T)	H199H/A231A				N/A	Elahi et al., 2003	Renamed from UGT1A10*1c
UGT1A10*2a	UGT1A10.2	415(G>A)	E139K				N/A	Elahi et al., 2003	
UGT1A10*3a	UGT1A10.3	719(C>T)	T240M				N/A	Elahi et al., 2003	
UGT1A10*3b	UGT1A10.3	693(C>T)/719(C>T)	A231A/T240M				N/A	Elahi et al., 2003	Renamed from UGT1A10*3a
UGT1A10*4a	UGT1A10.4	730(C>A)	L244I				N/A	Elahi et al., 2003	
UGT1A10*4b	UGT1A10.4	693(C>T)/730(C>A)	A231A/L244I				N/A	Elahi et al., 2003	Renamed from UGT1A10*4a
UGT1A10*4c	UGT1A10.4	597(T>C)/693(C>T)/730(C>A)	H199H/A231A/L244I				N/A	Elahi et al., 2003	Renamed from UGT1A10*4b
UGT1A10*5	UGT1A10.5	177(G>A)	M59I			No change (Jinno et al., 2003)	N/A	Saeki et al., 2002, Jinno et al., 2003	
UGT1A10*6	UGT1A10.6	605(C>T)	T202I			Reduced by 50% (Jinno et al., 2003)	N/A	Saeki et al., 2002, Jinno et al., 2003	
UGT1A10*7	UGT1A10.7	622 (T>C)	I211T			No activity (Martineau et al., 2004)	N/A	Martineau et al., 2004	

N/A : Data not available in the manuscript.

As for UGT1A9, any of the following polymorphisms, or a combination thereof, may be implemented in the context of the present invention: -2208(C>T), -2152(C>T), -2141(C>T), -1887(T>G), -1818(T>C), -665(C>T), -440 (T>C), -331(C>T), -275(T>A), -109 to -98 T_n(9>10), -87 (G>A), and codon 33 (T>C).

- 5 Any embodiment discussed in the context of UGT1A1 may be applied with respect to any other UGT polymorphism discussed herein using the relevant sequence or sequences.

B. Preparation of Nucleic Acids

- 10 A nucleic acid may be made by any technique known to one of ordinary skill in the art, such as for example, chemical synthesis, enzymatic production or biological production. Non-limiting examples of a synthetic nucleic acid (*e.g.*, a synthetic oligonucleotide), include a nucleic acid made by *in vitro* chemical synthesis using phosphotriester, phosphite or phosphoramidite chemistry and solid phase techniques such as described in European Patent 266,032, incorporated herein by reference, or
15 via deoxynucleoside H-phosphonate intermediates as described by Froehler *et al.*, 1986 and U.S. Patent 5,705,629, each incorporated herein by reference. In the methods of the present invention, one or more oligonucleotide may be used. Various different mechanisms of oligonucleotide synthesis have been disclosed in for example, U.S. Patents 4,659,774, 4,816,571, 5,141,813, 5,264,566, 4,959,463,
20 5,428,148, 5,554,744, 5,574,146, 5,602,244, each of which is incorporated herein by reference.

- A non-limiting example of an enzymatically produced nucleic acid include one produced by enzymes in amplification reactions such as PCR™ (see for example, U.S. Patent 4,683,202 and U.S. Patent 4,682,195, each incorporated herein by
25 reference), or the synthesis of an oligonucleotide described in U.S. Patent 5,645,897, incorporated herein by reference. A non-limiting example of a biologically produced nucleic acid includes a recombinant nucleic acid produced (*i.e.*, replicated) in a living cell, such as a recombinant DNA vector replicated in bacteria (see for example, Sambrook *et al.* 2001, incorporated herein by reference).

C. Purification of Nucleic Acids

- 30 A nucleic acid may be purified on polyacrylamide gels, cesium chloride centrifugation gradients, chromatography columns or by any other means known to one of ordinary skill in the art (see for example, Sambrook *et al.*, 2001, incorporated herein by reference). In some aspects, a nucleic acid is a pharmacologically

acceptable nucleic acid. Pharmacologically acceptable compositions are known to those of skill in the art, and are described herein.

In certain aspects, the present invention concerns a nucleic acid that is an isolated nucleic acid. As used herein, the term "isolated nucleic acid" refers to a nucleic acid molecule (*e.g.*, an RNA or DNA molecule) that has been isolated free of, or is otherwise free of, the bulk of the total genomic and transcribed nucleic acids of one or more cells. In certain embodiments, "isolated nucleic acid" refers to a nucleic acid that has been isolated free of, or is otherwise free of, bulk of cellular components or *in vitro* reaction components such as for example, macromolecules such as lipids or proteins, small biological molecules, and the like.

D. Nucleic Acid Segments

In certain embodiments, the nucleic acid is a nucleic acid segment. As used herein, the term "nucleic acid segment," are fragments of a nucleic acid, such as, for a non-limiting example, those that encode only part of a *UGT1A1* gene locus or a *UGT1A1* gene sequence. Thus, a "nucleic acid segment" may comprise any part of a gene sequence, including from about 2 nucleotides to the full length gene including promoter regions to the polyadenylation signal and any length that includes all the coding region.

Various nucleic acid segments may be designed based on a particular nucleic acid sequence, and may be of any length. By assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all nucleic acid segments can be created:

$$n \text{ to } n + y$$

where n is an integer from 1 to the last number of the sequence and y is the length of the nucleic acid segment minus one, where $n + y$ does not exceed the last number of the sequence. Thus, for a 10-mer, the nucleic acid segments correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the nucleic acid segments correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the nucleic segments correspond to bases 1 to 20, 2 to 21, 3 to 22 ... and so on. In certain embodiments, the nucleic acid segment may be a probe or primer. As used herein, a "probe" generally refers to a nucleic acid used in a detection method or composition. As used herein, a "primer" generally refers to a nucleic acid used in an extension or amplification method or composition.

E. Nucleic Acid Complements

The present invention also encompasses a nucleic acid that is complementary to a nucleic acid. A nucleic acid is "complement(s)" or is "complementary" to another nucleic acid when it is capable of base-pairing with another nucleic acid according to the standard Watson-Crick, Hoogsteen or reverse Hoogsteen binding complementarity rules. As used herein "another nucleic acid" may refer to a separate molecule or a spatial separated sequence of the same molecule. In preferred embodiments, a complement is a hybridization probe or amplification primer for the detection of a nucleic acid polymorphism.

As used herein, the term "complementary" or "complement" also refers to a nucleic acid comprising a sequence of consecutive nucleobases or semiconsecutive nucleobases (*e.g.*, one or more nucleobase moieties are not present in the molecule) capable of hybridizing to another nucleic acid strand or duplex even if less than all the nucleobases do not base pair with a counterpart nucleobase. However, in some diagnostic or detection embodiments, completely complementary nucleic acids are preferred.

III. NUCLEIC ACID DETECTION

Some embodiments of the invention concern identifying polymorphisms in *UGT1A1* such as the number of TA repeats, correlating genotype or haplotype to phenotype, wherein the phenotype is altered *UGT1A1* activity or expression, and then identifying such polymorphisms in patients who have or will be given L-thyroxine or other drugs or compounds that are *UGT1A1* substrates. Thus, the present invention involves assays for identifying polymorphisms and other nucleic acid detection methods. Such assays involve identifying the number of TA repeats in the *UGT1A1* gene, as shown in SEQ ID NOs:1-4. Other embodiments also involve identifying other polymorphisms in *UGT1A1* or in other *UGT* genes, including, but not limited to *UGT1A3*, *UGT1A8*, *UGT1A9*, and *UGTA10*.

It is contemplated that probes and primers can be prepared using the sequences disclosed in SEQ ID NOs:1-4. Nucleic acids, therefore, have utility as probes or primers for embodiments involving nucleic acid hybridization. They may be used in diagnostic or screening methods of the present invention. Detection of nucleic acids encoding *UGT1A1*, as well as nucleic acids involved in the expression or stability of *UGT1A1* polypeptides or transcripts, are encompassed by the invention. General

methods of nucleic acid detection methods are provided below, followed by specific examples employed for the identification of polymorphisms, including single nucleotide polymorphisms (SNPs) or the number of TA repeats.

A. Hybridization

5 The use of a probe or primer of between 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 50, 60, 70, 80, 90, or 100 nucleotides, preferably between 17 and 100 nucleotides in length, or in some aspects of the invention up to 1-2 kilobases or more in length, allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over contiguous
10 stretches greater than 20 bases in length are generally preferred, to increase stability and/or selectivity of the hybrid molecules obtained. One will generally prefer to design nucleic acid molecules for hybridization having one or more complementary sequences of 20 to 30 nucleotides, or even longer where desired. Such fragments may be readily prepared, for example, by directly synthesizing the fragment by chemical
15 means or by introducing selected sequences into recombinant vectors for recombinant production.

 In certain embodiments, amplification is employed to determine the number of TA repeats. See, *e.g.*, U.S. Patent Nos. 6,472,157 and 6,395,481; Te *et al.*, 2000; and, Innocenti *et al.*, 2004, all of which are hereby incorporated by reference for their
20 teachings regarding determining the number of TA repeats in the *UGT1A1* gene.

 Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of DNAs and/or RNAs or to provide primers for amplification of DNA or RNA from samples. Depending on the application envisioned, one would desire to employ varying
25 conditions of hybridization to achieve varying degrees of selectivity of the probe or primers for the target sequence.

 For applications requiring high selectivity, one will typically desire to employ relatively high stringency conditions to form the hybrids. For example, relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about
30 0.10 M NaCl at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch between the probe or primers and the template or target strand and would be particularly suitable for isolating specific genes or for detecting a specific polymorphism. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide.

For example, under highly stringent conditions, hybridization to filter-bound DNA may be carried out in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1 x SSC/0.1% SDS at 68°C (Ausubel *et al.*, 1989).

5 Conditions may be rendered less stringent by increasing salt concentration and/or decreasing temperature. For example, a medium stringency condition could be provided by about 0.1 to 0.25M NaCl at temperatures of about 37°C to about 55°C, while a low stringency condition could be provided by about 0.15M to about 0.9M salt, at temperatures ranging from about 20°C to about 55°C. Under low stringent conditions, such as moderately stringent conditions the washing may be carried out
10 for example in 0.2 x SSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989). Hybridization conditions can be readily manipulated depending on the desired results.

In other embodiments, hybridization may be achieved under conditions of, for example, 50mM Tris-HCl (pH 8.3), 75mM KCl, 3mM MgCl₂, 1.0mM dithiothreitol, at temperatures between approximately 20°C to about 37°C. Other hybridization
15 conditions utilized could include approximately 10mM Tris-HCl (pH 8.3), 50mM KCl, 1.5mM MgCl₂, at temperatures ranging from approximately 40°C to about 72°C.

In certain embodiments, it will be advantageous to employ nucleic acids of defined sequences of the present invention in combination with an appropriate means,
20 such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of being detected. In preferred embodiments, one may desire to employ a fluorescent label or an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other
25 environmentally undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known that can be employed to provide a detection means that is visibly or spectrophotometrically detectable, to identify specific hybridization with complementary nucleic acid containing samples. In other aspects, a particular nuclease cleavage site may be present and detection of a particular nucleotide
30 sequence can be determined by the presence or absence of nucleic acid cleavage.

In general, it is envisioned that the probes or primers described herein will be useful as reagents in solution hybridization, as in PCR, for detection of expression or genotype of corresponding genes, as well as in embodiments employing a solid phase.

In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to hybridization with selected probes under desired conditions. The conditions selected will depend on the particular circumstances (depending, for example, on the G+C content, type of target nucleic acid, source of nucleic acid, size of hybridization probe, *etc.*). Optimization of hybridization conditions for the particular application of interest is well known to those of skill in the art. After washing of the hybridized molecules to remove non-specifically bound probe molecules, hybridization is detected, and/or quantified, by determining the amount of bound label. Representative solid phase hybridization methods are disclosed in U.S. Patents 5,843,663, 5,900,481 and 5,919,626. Other methods of hybridization that may be used in the practice of the present invention are disclosed in U.S. Patents 5,849,481, 5,849,486 and 5,851,772. The relevant portions of these and other references identified in this section of the Specification are incorporated herein by reference.

B. Amplification of Nucleic Acids

Nucleic acids used as a template for amplification may be isolated from cells, tissues or other samples according to standard methodologies (Sambrook *et al.*, 2001). In certain embodiments, analysis is performed on whole cell or tissue homogenates or biological fluid samples with or without substantial purification of the template nucleic acid. The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to first convert the RNA to a complementary DNA.

The term "primer," as used herein, is meant to encompass any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process. Typically, primers are oligonucleotides from ten to twenty and/or thirty base pairs in length, but longer sequences can be employed. Primers may be provided in double-stranded and/or single-stranded form, although the single-stranded form is preferred.

Pairs of primers designed to selectively hybridize to nucleic acids corresponding to the UGT1A1 gene locus, or variants thereof, and fragments thereof are contacted with the template nucleic acid under conditions that permit selective hybridization. Depending upon the desired application, high stringency hybridization conditions may be selected that will only allow hybridization to sequences that are

completely complementary to the primers. In other embodiments, hybridization may occur under reduced stringency to allow for amplification of nucleic acids that contain one or more mismatches with the primer sequences. Once hybridized, the template-primer complex is contacted with one or more enzymes that facilitate template-dependent nucleic acid synthesis. Multiple rounds of amplification, also referred to as "cycles," are conducted until a sufficient amount of amplification product is produced.

The amplification product may be detected, analyzed or quantified. In certain applications, the detection may be performed by visual means. In certain applications, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of incorporated radiolabel or fluorescent label or even via a system using electrical and/or thermal impulse signals (Affymax technology; Bellus, 1994).

A number of template dependent processes are available to amplify the oligonucleotide sequences present in a given template sample. One of the best known amplification methods is the polymerase chain reaction (referred to as PCR™) which is described in detail in U.S. Patents 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*, 1988, each of which is incorporated herein by reference in their entirety.

Another method for amplification is ligase chain reaction ("LCR"), disclosed in European Application No. 320 308, incorporated herein by reference in its entirety. U.S. Patent 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence. A method based on PCR™ and oligonucleotide ligase assay (OLA) (described in further detail below), disclosed in U.S. Patent 5,912,148, may also be used.

Alternative methods for amplification of target nucleic acid sequences that may be used in the practice of the present invention are disclosed in U.S. Patents 5,843,650, 5,846,709, 5,846,783, 5,849,546, 5,849,497, 5,849,547, 5,858,652, 5,866,366, 5,916,776, 5,922,574, 5,928,905, 5,928,906, 5,932,451, 5,935,825, 5,939,291 and 5,942,391, Great Britain Application 2 202 328, and in PCT Application PCT/US89/01025, each of which is incorporated herein by reference in its entirety. Qbeta Replicase, described in PCT Application PCT/US87/00880, may also be used as an amplification method in the present invention.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[alpha-thio]-triphosphates in one strand of a restriction site may also be

useful in the amplification of nucleic acids in the present invention (Walker *et al.*, 1992). Strand Displacement Amplification (SDA), disclosed in U.S. Patent 5,916,779, is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.*, nick
5 translation.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR (Kwoh *et al.*, 1989; PCT Application WO 88/10315, incorporated
10 herein by reference in their entirety). European Application 329 822 disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention.

PCT Application WO 89/06700 (incorporated herein by reference in its entirety) disclose a nucleic acid sequence amplification scheme based on the
15 hybridization of a promoter region/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic, *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" and "one-sided PCR" (Frohman, 1990; Ohara *et al.*, 1989).

20 C. Detection of Nucleic Acids

Following any amplification, it may be desirable to separate the amplification product from the template and/or the excess primer. In one embodiment, amplification products are separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods (Sambrook *et al.*, 2001).
25 Separated amplification products may be cut out and eluted from the gel for further manipulation. Using low melting point agarose gels, the separated band may be removed by heating the gel, followed by extraction of the nucleic acid.

Separation of nucleic acids may also be effected by spin columns and/or chromatographic techniques known in art. There are many kinds of chromatography
30 which may be used in the practice of the present invention, including adsorption, partition, ion-exchange, hydroxylapatite, molecular sieve, reverse-phase, column, paper, thin-layer, and gas chromatography as well as HPLC.

In certain embodiments, the amplification products are visualized, with or without separation. A typical visualization method involves staining of a gel with

ethidium bromide and visualization of bands under UV light. Alternatively, if the amplification products are integrally labeled with radio- or fluorometrically-labeled nucleotides, the separated amplification products can be exposed to x-ray film or visualized under the appropriate excitatory spectra.

5 In one embodiment, following separation of amplification products, a labeled nucleic acid probe is brought into contact with the amplified marker sequence. The probe preferably is conjugated to a chromophore but may be radiolabeled. In another embodiment, the probe is conjugated to a binding partner, such as an antibody or biotin, or another binding partner carrying a detectable moiety.

10 In particular embodiments, detection is by Southern blotting and hybridization with a labeled probe. The techniques involved in Southern blotting are well known to those of skill in the art (see Sambrook *et al.*, 2001). One example of the foregoing is described in U.S. Patent 5,279,721, incorporated by reference herein, which discloses an apparatus and method for the automated electrophoresis and transfer of nucleic
15 acids. The apparatus permits electrophoresis and blotting without external manipulation of the gel and is ideally suited to carrying out methods according to the present invention.

Other methods of nucleic acid detection that may be used in the practice of the instant invention are disclosed in U.S. Patents 5,840,873, 5,843,640, 5,843,651,
20 5,846,708, 5,846,717, 5,846,726, 5,846,729, 5,849,487, 5,853,990, 5,853,992, 5,853,993, 5,856,092, 5,861,244, 5,863,732, 5,863,753, 5,866,331, 5,905,024, 5,910,407, 5,912,124, 5,912,145, 5,919,630, 5,925,517, 5,928,862, 5,928,869, 5,929,227, 5,932,413 and 5,935,791, each of which is incorporated herein by reference.

25 **D. Other Assays**

Other methods for genetic screening may be used within the scope of the present invention, for example, to detect mutations in genomic DNA, cDNA and/or RNA samples. Methods used to detect point mutations include denaturing gradient gel electrophoresis ("DGGE"), restriction fragment length polymorphism analysis
30 ("RFLP"), chemical or enzymatic cleavage methods, direct sequencing of target regions amplified by PCR™ (see above), single-strand conformation polymorphism analysis ("SSCP") and other methods well known in the art.

One method of screening for point mutations is based on RNase cleavage of base pair mismatches in RNA/DNA or RNA/RNA heteroduplexes. As used herein,

the term "mismatch" is defined as a region of one or more unpaired or mispaired nucleotides in a double-stranded RNA/RNA, RNA/DNA or DNA/DNA molecule. This definition thus includes mismatches due to insertion/deletion mutations, as well as single or multiple base point mutations.

5 U.S. Patent 4,946,773 describes an RNase A mismatch cleavage assay that involves annealing single-stranded DNA or RNA test samples to an RNA probe, and subsequent treatment of the nucleic acid duplexes with RNase A. For the detection of mismatches, the single-stranded products of the RNase A treatment, electrophoretically separated according to size, are compared to similarly treated
10 control duplexes. Samples containing smaller fragments (cleavage products) not seen in the control duplex are scored as positive.

Other investigators have described the use of RNase I in mismatch assays. The use of RNase I for mismatch detection is described in literature from Promega Biotech. Promega markets a kit containing RNase I that is reported to cleave three
15 out of four known mismatches. Others have described using the MutS protein or other DNA-repair enzymes for detection of single-base mismatches.

Alternative methods for detection of deletion, insertion or substitution mutations that may be used in the practice of the present invention are disclosed in U.S. Patents 5,849,483, 5,851,770, 5,866,337, 5,925,525 and 5,928,870, each of
20 which is incorporated herein by reference in its entirety.

E. Specific Examples of Polymorphism Screening Methods

Spontaneous mutations that arise during the course of evolution in the genomes of organisms are often not immediately transmitted throughout all of the members of the species, thereby creating polymorphic alleles that co-exist in the
25 species populations. Often polymorphisms are the cause of genetic diseases. Several classes of polymorphisms have been identified. For example, variable nucleotide type polymorphisms (VNTRs), arise from spontaneous tandem duplications of di- or trinucleotide repeated motifs of nucleotides. If such variations alter the lengths of DNA fragments generated by restriction endonuclease cleavage, the variations are
30 referred to as restriction fragment length polymorphisms (RFLPs). RFLPs are been widely used in human and animal genetic analyses.

Another class of polymorphisms are generated by the replacement of a single nucleotide. Such single nucleotide polymorphisms (SNPs) rarely result in changes in a restriction endonuclease site. Thus, SNPs are rarely detectable restriction fragment

length analysis. SNPs are the most common genetic variations and occur once every 100 to 300 bases and several SNP mutations have been found that affect a single nucleotide in a protein-encoding gene in a manner sufficient to actually cause a genetic disease. SNP diseases are exemplified by hemophilia, sickle-cell anemia, hereditary hemochromatosis, late-onset alzheimer disease *etc.*

SNPs can be the result of deletions, point mutations and insertions and in general any single base alteration, whatever the cause, can result in a SNP. The greater frequency of SNPs means that they can be more readily identified than the other classes of polymorphisms. The greater uniformity of their distribution permits the identification of SNPs "nearer" to a particular trait of interest. The combined effect of these two attributes makes SNPs extremely valuable. For example, if a particular trait (*e.g.*, inability to efficiently metabolize irinotecan) reflects a mutation at a particular locus, then any polymorphism that is linked to the particular locus can be used to predict the probability that an individual will exhibit that trait.

Several methods have been developed to screen polymorphisms and some examples are listed below. The reference of Kwok and Chen (2003) and Kwok (2001) provide overviews of some of these methods; both of these references are specifically incorporated by reference.

SNPs or other polymorphisms relating to UGT1A1, or any other UGT1 enzyme can be characterized by the use of any of these methods or suitable modification thereof. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes where the respective alleles of the site create or destroy a restriction site, the use of allele-specific hybridization probes, the use of antibodies that are specific for the proteins encoded by the different alleles of the polymorphism, or any other biochemical interpretation.

i) DNA Sequencing

The most commonly used method of characterizing a polymorphism is direct DNA sequencing of the genetic locus that flanks and includes the polymorphism. Such analysis can be accomplished using either the "dideoxy-mediated chain termination method," also known as the "Sanger Method" (Sanger *et al.*, 1975) or the "chemical degradation method," also known as the "Maxam-Gilbert method" (Maxam *et al.*, 1977). Sequencing in combination with genomic sequence-specific amplification technologies, such as the polymerase chain reaction may be utilized to facilitate the recovery of the desired genes (Mullis *et al.*, 1986; European Patent

Application 50,424; European Patent Application. 84,796, European Patent Application 258,017, European Patent Application. 237,362; European Patent Application. 201,184; U.S. Patents 4,683,202; 4,582,788; and 4,683,194), all of the above incorporated herein by reference.

5 **ii) Exonuclease Resistance**

Other methods that can be employed to determine the identity of a nucleotide present at a polymorphic site utilize a specialized exonuclease-resistant nucleotide derivative (U.S. Patent. 4,656,127). A primer complementary to an allelic sequence immediately 3'-to the polymorphic site is hybridized to the DNA under investigation.
10 If the polymorphic site on the DNA contains a nucleotide that is complementary to the particular exonucleotide-resistant nucleotide derivative present, then that derivative will be incorporated by a polymerase onto the end of the hybridized primer. Such incorporation makes the primer resistant to exonuclease cleavage and thereby permits its detection. As the identity of the exonucleotide-resistant derivative is known one
15 can determine the specific nucleotide present in the polymorphic site of the DNA.

iii) Microsequencing Methods

Several other primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described (Komher *et al.*, 1989; Sokolov, 1990; Syvanen 1990; Kuppuswamy *et al.*, 1991; Prezant *et al.*, 1992; Ugozzoli *et al.*, 1992;
20 Nyren *et al.*, 1993). These methods rely on the incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. As the signal is proportional to the number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide result in a signal that is proportional to the length of the run (Syvanen *et al.*, 1990).

25 **iv) Extension in Solution**

French Patent 2,650,840 and PCT Application WO91/02087 discuss a solution-based method for determining the identity of the nucleotide of a polymorphic site. According to these methods, a primer complementary to allelic sequences immediately 3'-to a polymorphic site is used. The identity of the nucleotide of that
30 site is determined using labeled dideoxynucleotide derivatives which are incorporated at the end of the primer if complementary to the nucleotide of the polymorphic site.

v) Genetic Bit Analysis or Solid-Phase Extension

PCT Application WO92/15712 describes a method that uses mixtures of labeled terminators and a primer that is complementary to the sequence 3' to a

polymorphic site. The labeled terminator that is incorporated is complementary to the nucleotide present in the polymorphic site of the target molecule being evaluated and is thus identified. Here the primer or the target molecule is immobilized to a solid phase.

5 **vi) Oligonucleotide Ligation Assay (OLA)**

This is another solid phase method that uses different methodology (Landegren *et al.*, 1988). Two oligonucleotides, capable of hybridizing to abutting sequences of a single strand of a target DNA are used. One of these oligonucleotides is biotinylated while the other is detectably labeled. If the precise complementary
10 sequence is found in a target molecule, the oligonucleotides will hybridize such that their termini abut, and create a ligation substrate. Ligation permits the recovery of the labeled oligonucleotide by using avidin. Other nucleic acid detection assays, based on this method, combined with PCR have also been described (Nickerson *et al.*, 1990). Here PCR is used to achieve the exponential amplification of target DNA, which is
15 then detected using the OLA.

vii) Ligase/Polymerase-Mediated Genetic Bit Analysis

U.S. Patent 5,952,174 describes a method that also involves two primers capable of hybridizing to abutting sequences of a target molecule. The hybridized product is formed on a solid support to which the target is immobilized. Here the
20 hybridization occurs such that the primers are separated from one another by a space of a single nucleotide. Incubating this hybridized product in the presence of a polymerase, a ligase, and a nucleoside triphosphate mixture containing at least one deoxynucleoside triphosphate allows the ligation of any pair of abutting hybridized oligonucleotides. Addition of a ligase results in two events required to generate a
25 signal, extension and ligation. This provides a higher specificity and lower "noise" than methods using either extension or ligation alone and unlike the polymerase-based assays, this method enhances the specificity of the polymerase step by combining it with a second hybridization and a ligation step for a signal to be attached to the solid phase.

30 **vii) Invasive Cleavage Reactions**

Invasive cleavage reactions can be used to evaluate cellular DNA for a particular polymorphism. A technology called INVADER® employs such reactions (*e.g.*, de Arruda *et al.*, 2002; Stevens *et al.*, 2003, which are incorporated by reference). Generally, there are three nucleic acid molecules: 1) an oligonucleotide

upstream of the target site ("upstream oligo"), 2) a probe oligonucleotide covering the target site ("probe"), and 3) a single-stranded DNA with the target site ("target"). The upstream oligo and probe do not overlap but they contain contiguous sequences. The probe contains a donor fluorophore, such as fluorescein, and an acceptor dye, such as Dabcyl. The nucleotide at the 3' terminal end of the upstream oligo overlaps ("invades") the first base pair of a probe-target duplex. Then the probe is cleaved by a structure-specific 5' nuclease causing separation of the fluorophore/quencher pair, which increases the amount of fluorescence that can be detected. *See Lu et al., 2004.*

In some cases, the assay is conducted on a solid-surface or in an array format.

ix) Other Methods To Detect SNPs

Several other specific methods for SNP detection and identification are presented below and may be used as such or with suitable modifications in conjunction with identifying polymorphisms of the *UGT1* genes in the present invention. Several other methods are also described on the SNP web site of the NCBI at the website on the World Wide Web at ncbi.nlm.nih.gov/SNP, incorporated herein by reference.

In a particular embodiment, extended haplotypes may be determined at any given locus in a population, which allows one to identify exactly which SNPs will be redundant and which will be essential in association studies. The latter is referred to as 'haplotype tag SNPs (htSNPs)', markers that capture the haplotypes of a gene or a region of linkage disequilibrium. *See Johnson et al. (2001) and Ke and Cardon (2003)*, each of which is incorporated herein by reference, for exemplary methods.

The VDA-assay utilizes PCR amplification of genomic segments by long PCR methods using TaKaRa LA Taq reagents and other standard reaction conditions. The long amplification can amplify DNA sizes of about 2,000-12,000 bp. Hybridization of products to variant detector array (VDA) can be performed by a Affymetrix High Throughput Screening Center and analyzed with computerized software.

A method called Chip Assay uses PCR amplification of genomic segments by standard or long PCR protocols. Hybridization products are analyzed by VDA, Halushka *et al.* (1999), incorporated herein by reference. SNPs are generally classified as "Certain" or "Likely" based on computer analysis of hybridization patterns. By comparison to alternative detection methods such as nucleotide sequencing, "Certain" SNPs have been confirmed 100% of the time; and "Likely" SNPs have been confirmed 73% of the time by this method.

Other methods simply involve PCR amplification following digestion with the relevant restriction enzyme. Yet others involve sequencing of purified PCR products from known genomic regions.

5 In yet another method, individual exons or overlapping fragments of large exons are PCR-amplified. Primers are designed from published or database sequences and PCR-amplification of genomic DNA is performed using the following conditions: 200 ng DNA template, 0.5 μ M each primer, 80 μ M each of dCTP, dATP, dTTP and dGTP, 5% formamide, 1.5mM MgCl₂, 0.5U of Taq polymerase and 0.1 volume of the Taq buffer. Thermal cycling is performed and resulting PCR-products
10 are analyzed by PCR-single strand conformation polymorphism (PCR-SSCP) analysis, under a variety of conditions, *e.g.* 5 or 10% polyacrylamide gel with 15% urea, with or without 5% glycerol. Electrophoresis is performed overnight. PCR-products that show mobility shifts are reamplified and sequenced to identify nucleotide variation.

15 In a method called CGAP-GAI (DEMIGLACE), sequence and alignment data (from a PHRAP.ace file), quality scores for the sequence base calls (from PHRED quality files), distance information (from PHYLIP dnadist and neighbour programs) and base-calling data (from PHRED '-d' switch) are loaded into memory. Sequences are aligned and examined for each vertical chunk ('slice') of the resulting assembly for
20 disagreement. Any such slice is considered a candidate SNP (DEMIGLACE). A number of filters are used by DEMIGLACE to eliminate slices that are not likely to represent true polymorphisms. These include filters that: (i) exclude sequences in any given slice from SNP consideration where neighboring sequence quality scores drop 40% or more; (ii) exclude calls in which peak amplitude is below the fifteenth
25 percentile of all base calls for that nucleotide type; (iii) disqualify regions of a sequence having a high number of disagreements with the consensus from participating in SNP calculations; (iv) removed from consideration any base call with an alternative call in which the peak takes up 25% or more of the area of the called peak; (v) exclude variations that occur in only one read direction. PHRED quality
30 scores were converted into probability-of-error values for each nucleotide in the slice. Standard Bayesian methods are used to calculate the posterior probability that there is evidence of nucleotide heterogeneity at a given location.

In a method called CU-RDF (RESEQ), PCR amplification is performed from DNA isolated from blood using specific primers for each SNP, and after typical cleanup protocols to remove unused primers and free nucleotides, direct sequencing using the same or nested primers.

5 In a method called DEBNICK (METHOD-B), a comparative analysis of clustered EST sequences is performed and confirmed by fluorescent-based DNA sequencing. In a related method, called DEBNICK (METHOD-C), comparative analysis of clustered EST sequences with phred quality > 20 at the site of the mismatch, average phred quality $\geq$ 20 over 5 bases 5'-FLANK and 3' to the SNP, no
10 mismatches in 5 bases 5' and 3' to the SNP, at least two occurrences of each allele is performed and confirmed by examining traces.

In a method identified by ERO (RESEQ), new primers sets are designed for electronically published STSs and used to amplify DNA from 10 different mouse strains. The amplification product from each strain is then gel purified and sequenced
15 using a standard dideoxy, cycle sequencing technique with ^{33}P -labeled terminators. All the ddATP terminated reactions are then loaded in adjacent lanes of a sequencing gel followed by all of the ddGTP reactions and so on. SNPs are identified by visually scanning the radiographs.

In another method identified as ERO (RESEQ-HT), new primers sets are
20 designed for electronically published murine DNA sequences and used to amplify DNA from 10 different mouse strains. The amplification product from each strain is prepared for sequencing by treating with Exonuclease I and Shrimp Alkaline Phosphatase. Sequencing is performed using ABI Prism Big Dye Terminator Ready Reaction Kit (Perkin-Elmer) and sequence samples are run on the 3700 DNA
25 Analyzer (96 Capillary Sequencer).

FGU-CBT (SCA2-SNP) identifies a method where the region containing the SNP were PCR amplified using the primers SCA2-FP3 and SCA2-RP3. Approximately 100 ng of genomic DNA is amplified in a 50 ml reaction volume containing a final concentration of 5mM Tris, 25mM KCl, 0.75mM MgCl_2 , 0.05%
30 gelatin, 20pmol of each primer and 0.5U of Taq DNA polymerase. Samples are denatured, annealed and extended and the PCR product is purified from a band cut out of the agarose gel using, for example, the QIAquick gel extraction kit (Qiagen) and is sequenced using dye terminator chemistry on an ABI Prism 377 automated DNA sequencer with the PCR primers.

In a method identified as JBLACK (SEQ/RESTRICT), two independent PCR reactions are performed with genomic DNA. Products from the first reaction are analyzed by sequencing, indicating a unique FspI restriction site. The mutation is confirmed in the product of the second PCR reaction by digesting with Fsp I.

5 In a method described as KWOK(1), SNPs are identified by comparing high quality genomic sequence data from four randomly chosen individuals by direct DNA sequencing of PCR products with dye-terminator chemistry (see Kwok *et al.*, 1996). In a related method identified as KWOK(2) SNPs are identified by comparing high quality genomic sequence data from overlapping large-insert clones such as bacterial
10 artificial chromosomes (BACs) or P1-based artificial chromosomes (PACs). An STS containing this SNP is then developed and the existence of the SNP in various populations is confirmed by pooled DNA sequencing (see Taillon-Miller *et al.*, 1998). In another similar method called KWOK(3), SNPs are identified by comparing high quality genomic sequence data from overlapping large-insert clones BACs or PACs.
15 The SNPs found by this approach represent DNA sequence variations between the two donor chromosomes but the allele frequencies in the general population have not yet been determined. In method KWOK(5), SNPs are identified by comparing high quality genomic sequence data from a homozygous DNA sample and one or more pooled DNA samples by direct DNA sequencing of PCR products with dye-
20 terminator chemistry. The STSs used are developed from sequence data found in publicly available databases. Specifically, these STSs are amplified by PCR against a complete hydatidiform mole (CHM) that has been shown to be homozygous at all loci and a pool of DNA samples from 80 CEPH parents (see Kwok *et al.*, 1994).

In another such method, KWOK (OverlapSnpDetectionWithPolyBayes), SNPs
25 are discovered by automated computer analysis of overlapping regions of large-insert human genomic clone sequences. For data acquisition, clone sequences are obtained directly from large-scale sequencing centers. This is necessary because base quality sequences are not present/available through GenBank. Raw data processing involves analyzed of clone sequences and accompanying base quality information for
30 consistency. Finished ('base perfect', error rate lower than 1 in 10,000 bp) sequences with no associated base quality sequences are assigned a uniform base quality value of 40 (1 in 10,000 bp error rate). Draft sequences without base quality values are rejected. Processed sequences are entered into a local database. A version of each sequence with known human repeats masked is also stored. Repeat masking is

performed with the program "MASKERAID." Overlap detection: Putative overlaps are detected with the program "WUBLAST." Several filtering steps followed in order to eliminate false overlap detection results, *i.e.* similarities between a pair of clone sequences that arise due to sequence duplication as opposed to true overlap. Total
5 length of overlap, overall percent similarity, number of sequence differences between nucleotides with high base quality value "high-quality mismatches." Results are also compared to results of restriction fragment mapping of genomic clones at Washington University Genome Sequencing Center, finisher's reports on overlaps, and results of the sequence contig building effort at the NCBI. SNP detection: Overlapping pairs of
10 clone sequence are analyzed for candidate SNP sites with the 'POLYBAYES' SNP detection software. Sequence differences between the pair of sequences are scored for the probability of representing true sequence variation as opposed to sequencing error. This process requires the presence of base quality values for both sequences. High-scoring candidates are extracted. The search is restricted to substitution-type
15 single base pair variations. Confidence score of candidate SNP is computed by the POLYBAYES software.

In method identified by KWOK (TaqMan assay), the TaqMan assay is used to determine genotypes for 90 random individuals. In method identified by KYUGEN(Q1), DNA samples of indicated populations are pooled and analyzed by
20 PLACE-SSCP. Peak heights of each allele in the pooled analysis are corrected by those in a heterozygote, and are subsequently used for calculation of allele frequencies. Allele frequencies higher than 10% are reliably quantified by this method. Allele frequency = 0 (zero) means that the allele was found among individuals, but the corresponding peak is not seen in the examination of pool. Allele
25 frequency = 0-0.1 indicates that minor alleles are detected in the pool but the peaks are too low to reliably quantify.

In yet another method identified as KYUGEN (Method1), PCR products are post-labeled with fluorescent dyes and analyzed by an automated capillary electrophoresis system under SSCP conditions (PLACE-SSCP). Four or more
30 individual DNAs are analyzed with or without two pooled DNA (Japanese pool and CEPH parents pool) in a series of experiments. Alleles are identified by visual inspection. Individual DNAs with different genotypes are sequenced and SNPs identified. Allele frequencies are estimated from peak heights in the pooled samples after correction of signal bias using peak heights in heterozygotes. For the PCR

primers are tagged to have 5'-ATT or 5'-GTT at their ends for post-labeling of both strands. Samples of DNA (10 ng/ul) are amplified in reaction mixtures containing the buffer (10mM Tris-HCl, pH 8.3 or 9.3, 50mM KCl, 2.0mM MgCl₂), 0.25μM of each primer, 200μM of each dNTP, and 0.025 units/μl of Taq DNA polymerase premixed
5 with anti-Taq antibody. The two strands of PCR products are differentially labeled with nucleotides modified with R110 and R6G by an exchange reaction of Klenow fragment of DNA polymerase I. The reaction is stopped by adding EDTA, and unincorporated nucleotides are dephosphorylated by adding calf intestinal alkaline phosphatase. For the SSCP: an aliquot of fluorescently labeled PCR products and
10 TAMRA-labeled internal markers are added to deionized formamide, and denatured. Electrophoresis is performed in a capillary using an ABI Prism 310 Genetic Analyzer. Genescan softwares (P-E Biosystems) are used for data collection and data processing. DNA of individuals (two to eleven) including those who showed different genotypes on SSCP are subjected for direct sequencing using big-dye
15 terminator chemistry, on ABI Prism 310 sequencers. Multiple sequence trace files obtained from ABI Prism 310 are processed and aligned by Phred/Phrap and viewed using Consed viewer. SNPs are identified by PolyPhred software and visual inspection.

In yet another method identified as KYUGEN (Method2), individuals with
20 different genotypes are searched by denaturing HPLC (DHPLC) or PLACE-SSCP (Inazuka *et al.*, 1997) and their sequences are determined to identify SNPs. PCR is performed with primers tagged with 5'-ATT or 5'-GTT at their ends for post-labeling of both strands. DHPLC analysis is carried out using the WAVE DNA fragment analysis system (Transgenomic). PCR products are injected into DNASep column,
25 and separated under the conditions determined using WAVEMaker program (Transgenomic). The two strands of PCR products that are differentially labeled with nucleotides modified with R110 and R6G by an exchange reaction of Klenow fragment of DNA polymerase I. The reaction is stopped by adding EDTA, and unincorporated nucleotides are dephosphorylated by adding calf intestinal alkaline
30 phosphatase. SSCP followed by electrophoresis is performed in a capillary using an ABI Prism 310 Genetic Analyzer. Genescan softwares (P-E Biosystems). DNA of individuals including those who showed different genotypes on DHPLC or SSCP are subjected for direct sequencing using big-dye terminator chemistry, on ABI Prism 310

sequencer. Multiple sequence trace files obtained from ABI Prism 310 are processed and aligned by Phred/Phrap and viewed using Consed viewer. SNPs are identified by PolyPhred software and visual inspection. Trace chromatogram data of EST sequences in Unigene are processed with PHRED. To identify likely SNPs, single
5 base mismatches are reported from multiple sequence alignments produced by the programs PHRAP, BRO and POA for each Unigene cluster. BRO corrected possible misreported EST orientations, while POA identified and analyzed non-linear alignment structures indicative of gene mixing/chimeras that might produce spurious SNPs. Bayesian inference is used to weigh evidence for true polymorphism versus
10 sequencing error, misalignment or ambiguity, misclustering or chimeric EST sequences, assessing data such as raw chromatogram height, sharpness, overlap and spacing; sequencing error rates; context-sensitivity; cDNA library origin, etc.

In method identified as MARSHFIELD(Method-B), overlapping human DNA sequences which contained putative insertion/deletion polymorphisms are identified
15 through searches of public databases. PCR primers which flanked each polymorphic site are selected from the consensus sequences. Primers are used to amplify individual or pooled human genomic DNA. Resulting PCR products are resolved on a denaturing polyacrylamide gel and a PhosphorImager is used to estimate allele frequencies from DNA pools.

20

6. Linkage Disequilibrium

Polymorphisms in linkage disequilibrium with the number of TA repeats may also be used with the methods of the present invention. "Linkage disequilibrium" ("LD" as used herein, though also referred to as "LED" in the art) refers to a situation
25 where a particular combination of alleles (*i.e.*, a variant form of a given gene) or polymorphisms at two loci appears more frequently than would be expected by chance. "Significant" as used in respect to linkage disequilibrium, as determined by one of skill in the art, is contemplated to be a statistical p or α value that may be 0.25 or 0.1 and may be 0.1, 0.05, 0.001, 0.00001 or less. "Haplotype" is used according to
30 its plain and ordinary meaning to one skilled in the art. It refers to a collective genotype of two or more alleles or polymorphisms along one of the homologous chromosomes.

V. FORMULATIONS AND DOSAGES

L-thyroxine is commercially available as SYNTHROID®. SYNTHROID® is supplied as a tablet containing synthetic crystalline L-3,3',5,5'-tetraiodothyronine sodium salt [levoxythyroxine (T₄) sodium]. Synthetic T₄ is chemically identical to the hormone produced in the thyroid gland of a human.

Typically, a patient takes a single daily dose of SYNTHROID®. Dosages of SYNTHROID® include 100-125 mcg (µg)/day for a 70 kg adult. Occasionally, patients take doses above that, such as between 200 and 300 µg/day, or lower. Average daily dosages that are taken include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427,

428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444,
445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461,
462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478,
479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495,
5 496, 497, 498, 499, 500 mcg/day. It is contemplated that a typical amount for an adult
patient, for exmaple 100-125 mcg/day, is adjusted based on the predicted L-thyroxine
glucuronidation activity in the patient. If the activity is relatively high, the dosage will
be higher than the dosage for a patient whose activity is predicted to be relatively low
or at least lower. It is contemplated that dosages may be adjusted by timing as well so
10 that the patient effectively receives relatively more or less on an average daily basis.

It is contemplated that doses can be adjusted to be less than or more than the
concentrations discussed above or less frequently or more frequently than the timing
discussed above. It is contemplated treatment cycles may be repeated and that there
may be a respite between cycles. One of ordinary skill in the art is familiar with
15 dosages regimens. The overall amount of the drug administered to the patient in a
single regimen or for the treatment overall may be increased or decreased by about,
by at least about, or by at most about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65,
70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200,
1300, 1400, 1500, 1600, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800,
20 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100,
4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000 mcg/day, or any ranges
derivable therein. The term "dose" and "dosage" are used interchangeably as
synonyms herein.

The dosages of L-thyroxine that are administered to patients is well known to
25 those of skill in the art. These dosages may be reduced or increased relative to a
dosage that would have been administered in the absence of genotyping. It is
specifically contemplated that the dosages of any of those drugs may be similarly
altered or modified based on genotypic analysis described herein.

III. KITS

30 Any of the compositions described herein may be comprised in a kit. In a
non-limiting example, reagents for determining the sequence or number one or more
polymorphisms are included in a kit. In some embodiments, the kit includes reagents

for determining the number of TA repeats in the upstream region of the *UGT1A1* gene. It is contemplated that kits may include nucleic acids for amplifying, priming, or hybridizing to any of the polymorphisms discussed herein, such as the following: *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A, 5 *UGT1A3* 17(A>G), *UGT1A3* 31(T>C), *UGT1A3* 81(G>A), *UGT1A3* 133(C>T), *UGT1A3* 140(T>C), or *UGT1A3* 477(A>G), *UGT1A8* 518(C>G), *UGT1A8* 765(A>G), *UGT1A8* 830(G>A), *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T. *UGT1A10* 126(G>A), *UGT1A10* 177(G>A), *UGT1A10* 415(G>A), *UGT1A10* 597(T>C), *UGT1A10* 605(C>T), *UGT1A10* 622 (T>C), 10 *UGT1A10* 693(C>T), *UGT1A10* 719(C>T), *UGT1A10* 730(C>A), or a combination thereof, any of which may or may not be combined with such nucleic acids for detecting the number of TA repeats in *UGT1A1*.

Kits for implementing methods of the invention described herein are specifically contemplated. The kits will thus comprise in some embodiments, in 15 suitable container means, one or more of the following: an enzyme for amplifying or detecting a nucleic acid sequence with the polymorphism or capable of detecting whether a nucleic acid sequence has or does not have the polymorphism, one or more buffers, such as reaction buffer, or a hybridization buffer, compounds for preparing amplified nucleic acid sequences or detecting nucleic acid sequences, and components 20 for isolating sequences. Other kits of the invention may include components for making a nucleic acid array comprising polymorphic sequences or sequences that will allow the detection of polymorphisms and/or the absence of such polymorphisms, and thus, may include, for example, a solid support.

All or part of any nucleic acid discussed above may be implemented as part of 25 a kit.

The components of the kits may be packaged either in aqueous media or in lyophilized form. The container means of the kits will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which a component may be placed, and preferably, suitably aliquoted. Where there is more than one 30 component in the kit, the kit also will generally contain a second, third or other additional container into which the additional components may be separately placed. However, various combinations of components may be comprised in a vial. The kits of the present invention also will typically include a means for containing the nucleic acids, and any other reagent containers in close confinement for commercial sale.

Such containers may include injection or blow-molded plastic containers into which the desired vials are retained.

When the components of the kit are provided in one and/or more liquid solutions, the liquid solution is an aqueous solution, with a sterile aqueous solution being particularly preferred.

However, the components of the kit may be provided as dried powder(s). When reagents and/or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent. It is envisioned that the solvent may also be provided in another container means. In some embodiments, labeling dyes are provided as a dried power. It is contemplated that 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 130, 140, 150, 160, 170, 180, 190, 200, 300, 400, 500, 600, 700, 800, 900, 1000 μ g or at least or at most those amounts of dried dye are provided in kits of the invention. The dye may then be resuspended in any suitable solvent, such as DMSO.

The container means will generally include at least one vial, test tube, flask, bottle, syringe and/or other container means, into which the nucleic acid formulations are placed, preferably, suitably allocated. The kits may also comprise a second container means for containing a sterile, pharmaceutically acceptable buffer and/or other diluent.

The kits of the present invention will also typically include a means for containing the vials in close confinement for commercial sale, such as, *e.g.*, injection and/or blow-molded plastic containers into which the desired vials are retained.

Such kits may also include components that facilitate isolation of nucleic acids. It may also include components that preserve or maintain the nucleic acids or that protect against their degradation. Such components may be nuclease-free or protect against nucleases, such as DNAses or RNAses. Such kits generally will comprise, in suitable means, distinct containers for each individual reagent or solution.

A kit will also include instructions for employing the kit components as well the use of any other reagent not included in the kit. Instructions may include variations that can be implemented.

Kits of the invention may also include one or more of the following: Control nucleic acids; nuclease-free water; PEG or dextran; ethanol; acetic acid; sodium

acetate; ammonium acetate; guanidinium; detergent; nucleic acid size marker; tube tips; and RNase or DNase inhibitors.

It is contemplated that such reagents are embodiments of kits of the invention. Such kits, however, are not limited to the particular items identified above and may include any reagent used for the manipulation or characterization of nucleic acids that determine or detect polymorphisms discussed herein.

EXAMPLES

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

EXAMPLE 1

UGT1A1 AND UGT1A9 VARIANTS AFFECT THYROXINE GLUCURONIDATION IN HUMAN LIVERS

Methods

Thirty human liver microsomes were genotyped for the *1A1* (TA)_n and the *1A9* -118T_{9>10} promoter polymorphisms. cDNA transfected 1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A9 and 1A10 were screened for T₄ glucuronidation activity and normalized by relative protein expression. T₄G formation was measured by HPLC.

Results

There was a significant correlation between both *1A1* (TA)_{6>7} and *1A9* -118T_{9>10} genotypes and T₄ glucuronidation (P<0.01). The highest T₄G formation was observed with 1A3 followed by 1A8, 1A1, 1A10, 1A9, and 1A7 and was undetectable with 1A4 and 1A6. **CONCLUSIONS:** *1A1* (TA)_n and the *1A9* -118(T)_{9>10} promoter polymorphisms affect T₄ glucuronidation rates. Moreover, the data propose 1A3 as another 1A isoform potentially involved in T₄G formation.

Moreover, using liver specimens from donors, the inventors have demonstrated a significant correlation between a common polymorphism in the UGT1A1 gene (a TA indel in the gene promoter, also referred to as (TA)_n) and L-thyroxine glucuronidation rates. Individuals with the 7/7 UGT1A1 genotype had reduced L-thyroxine glucuronidation compared to either 6/7 or 6/6 genotypes (P<0.01). Hence, genetic differences in the way L-thyroxine is metabolized can affect the efficacy of L-thyroxine in hypothyroid patients. See FIG. 1.

EXAMPLE 2

UGT1A1 AND UGT1A9 VARIANTS AFFECT THYROXINE GLUCURONIDATION IN HUMAN LIVERS

Material and Methods

Pooled human liver microsomes (HLMs) were used to optimize incubation conditions and determine the affinity constant (K_m) of L-thyroxine (T4).

Normal human liver tissue (n=55) used for the correlation study between genotype and phenotype was processed through Dr. Mary Relling's laboratory at St. Jude Children's Research Hospital and was provided by the Liver Tissue Procurement and Distribution System and by the Cooperative Human Tissue Network.

Microsomes from cDNA-transfected baculovirus-insect cells expressing human UGTs (supersomes) and control insect cell microsomes from wild-type baculovirus-infected cells were obtained from BD Gentest (Woburn, MA).

HLMs and supersomes were screened for T4 glucuronidation levels by incubation with T4 under the following conditions: 0.2 mg of HLMs, 5 mM UDPGA, 1 mM MgCl₂, 5 mM sacharolactone, 10 µg alamethicin, and 100 mM Tris-HCl pH 7.1 in a final volume of 200 µL. All screening and correlation incubations were performed with 91 µM T4. Kinetics experiments in HLMs were performed with 0-300 µM T4. Incubations were carried out for 90 minutes at 37°C. T₄G was measured by HPLC. HPLC was implemented as follows:

Column: µBondapak C18 column (10 µM, 3.9 X 300 mM, Waters Corp., Milford, Massachusetts)

Detection: UV 234 nm

Flow: 1 mL/min

Mobile Phase:

Time (minutes)	Acetonitrile	Potassium phosphate 20 mM, pH 2
0-25	25%	75%
25.1-30	25-70%	75-30%
30.1-45	70%	30%
45.1-55	25%	75%

5 Correlation studies were performed with HLMs and UGT1A1 and UGT1A9 substrates SN-38 and flavopiridol as previously described (Innocenti *et al.*, 2002, Ramirez *et al.*, 2002).

Livers were genotyped for *UGT1A1* -53(TA)₆₋₇ and *UGT1A9* -118T>10 polymorphisms by PCR and sizing.

10 **Results**

The K_m for T₄ in HLMs was determined to be 91 μM.

UGT screening showed UGT1A8 to have the highest T₄G formation followed by UGT1A1 > 1A3 > 1A10 > 1A7 > 1A9 (FIG. 2). A high level of correlation was observed between T₄ and SN-38 glucuronidation levels (Spearman r = 0.8234, p < 0.0001, n = 55) (FIG. 3A). T₄ and flavopiridol glucuronide formation levels were also significantly correlated (Spearman r = 0.5775, p < 0.001, n = 49) (FIG. 3B). *UGT1A1* -53(TA)₆₋₇ genotypes had significantly different levels of T₄G formation with 6/6 > 6/7 > 7/7 (Kruskal-Wallis test p < 0.01) (FIG. 4). There was no significant difference in T₄ glucuronidation between *UGT1A9* -118T>10 genotypes (FIG. 5).

20

EXAMPLE 3

ADDITIONAL DETAILS AND RESULTS REGARDING AN *IN VITRO* PHARMACOGENETIC STUDY OF THYROXINE GLUCURONIDATION

Material and Methods

25 *Chemicals and Reagents*

Unless otherwise stated, all chemicals were obtained from Sigma-Aldrich (St. Louis, MO). Irinotecan (CPT-11) was kindly provided by Yakult Honsha Co., Ltd.

(Tokyo, Japan). Microsomes from cDNA-transfected baculovirus-insect cells expressing UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9, UGT1A10, UGT2B4, UGT2B7, UGT2B15, and UGT2B17 as well as control insect cell microsomes from wild-type baculovirus-infected cells were obtained from BD Gentest (Woburn, MA). Rabbit anti-human UGT1A primary and secondary antibodies were also purchased from BD Gentest. Pre-cast 10% acrylamide gels were obtained from BioRad (Hercules, CA). Blotting was purchased from Pierce (Rockford, IL). Chemiluminescence reagents were obtained from Amersham Biosciences (Arlington Heights, IL). Methanol and acetonitrile (both HPLC grade) were purchased through VWR (West Chester, PA). All reagents were of the highest grade commercially available.

Human Livers

Fifty-four Caucasian human livers were processed through Dr. Mary Relling's laboratory at St. Jude Children's Research Hospital and were provided by the Liver Tissue Procurement and Distribution System and by the Cooperative Human Tissue Network, funded through NIH contract #N01-DK-9-2310.

In Vitro T₄ Glucuronidation Assay

Experiments were initially performed with pooled human liver microsomes (n=10) to determine optimal incubation conditions. A typical incubation contained 0.2 mg of HLM, 90 μ M T₄, 5 mM UDPGA, 1 mM MgCl₂, 5 mM saccharolactone, 10 μ g alamethicin, and 100 mM Tris-HCl (pH 7.1) in a final volume of 200 μ L. T₄ was dissolved in DMSO. The concentration of DMSO in the incubation was 1% (v/v). Microsomes were preincubated with alamethicin for 5 min on ice. The reaction was initiated with the addition of UDPGA. Samples were incubated at 37°C for 90 min. The reaction was terminated with 0.4 mL of ice-cold methanol. Internal standard (15 μ L of 100 μ M CPT-11) was added to the samples before centrifuging for 20 min at 14,000 rpm, 4°C. The supernatant was evaporated to dryness at 37°C under nitrogen gas. Samples were reconstituted with 200 μ L of initial mobile phase (25/75 acetonitrile/20 mM potassium phosphate buffer, pH 2). Aliquots of 150 μ L were analyzed by high performance liquid chromatography (HPLC). Control incubations without UDPGA, T₄, and microsomes were simultaneously performed.

β -glucuronidase Reactions

Hydrolysis with β -glucuronidase was used to confirm the identity of the T₄ glucuronide peak. T₄ incubations with HLM described above were dried down under nitrogen gas and reconstituted in 0.2 mL of 0.1 M sodium phosphate buffer, pH 6.4
5 containing 2000 units of β -glucuronidase (Type VII, from Escherichia coli, Sigma, St Louis, MO) and incubated for 24 hours at 37°C. The reaction was stopped with 0.4 mL of ice-cold methanol. Samples were processed and analyzed by HPLC as described above. Control incubations containing no β -glucuronidase were treated identically. Incubation with β -glucuronidase resulted in the disappearance of T₄G
10 peaks. T₄G peaks were observed in control incubations with the buffer when β -glucuronidase was absent (data not shown).

Measurement of T₄ Glucuronidation

T₄G formation was measured with a modification of a method previously published (Jemnitz and Vereczkey, 1996). Separation was achieved using a reverse-
15 phase μ Bondapak C₁₈ column (10 μ M, 3.9 X 300 mm, Waters Corp., Milford, MA). The mobile phase consisted of acetonitrile (A) and 20 mM potassium phosphate monobasic, pH 2 (B). The following gradient was applied at a flow rate of 1 mL/min: 0-25 min 25% A and 75% B, 25.1-30 min from 25% A and 75% B to 70% A and 30% B, 30.1-45 min 70% A and 30% B, and 45.1-55 min 25% A and 75% B. The
20 eluate was monitored using UV detection at 234 nm. Retention times for CPT-11, T₄G, and T₄ were 11.4, 16.0 and 33.4 min, respectively.

T₄G is not commercially available, so T₄G formation was calculated using a T₄ standard curve made by combining T₄ stock in DMSO (0.1% final concentration) with incubation buffer, incubated at 37°C for 90 min and processed as described
25 above. The quantitation limit was 32.2 nM. Intra-assay reproducibility was determined by measuring T₄G in ten separate incubations with the same HLM pool. The coefficient of variation was 3.2%. Inter-assay reproducibility was determined by incubating four pooled HLM samples each day for three days. The coefficient of variation was 10.2%.

Screening of Recombinant Microsomes Expressing Human UGT Isoforms for T₄G Formation

Incubations with 0.2 mg human UGT-expressing and control insect microsomes were performed with 90 μ M T₄ as described for human liver microsomes.

5 All incubations were performed in duplicate.

Kinetic Analyses

Kinetic studies were performed with pooled HLM and recombinant human UGT1A1, UGT1A3, UGT1A7, UGT1A8, UGT1A9, and UGT1A10 expressing insect cell microsomes with the incubation conditions described above. T₄ concentrations
10 ranged from 10 to 400 μ M. Incubations with HLM were performed in triplicate, while kinetic studies with recombinant UGTs were done in duplicate. The maximum T₄G formation for each UGT was normalized by relative protein expression as determined by western blot.

Western Blot

15 Recombinant UGTs (UGT1A1, UGT1A3, UGT1A8, UGT1A9 and UGT1A10) were prepared for western blot analysis by adding loading buffer and incubating at 100°C for 5 min. Proteins were separated by gel electrophoresis on a 10% acrylamide gel, transferred to a PVDF membrane, and blocked overnight with Blotto. The membrane was washed with tris buffered saline containing 0.05% tween
20 20 and incubated with anti-human UGT1A primary antibody (1:500 dilution in Blotto) for 1 hour. Subsequently, the membrane was washed with tris buffered saline containing 0.05% tween 20 and incubated with horseradish peroxidase conjugated goat anti-rabbit secondary antibody (1:500 dilution in blotto) for 1 hour. Immunocomplexes were visualized with chemiluminescence using a ChemiDoc
25 (BioRad, Hercules, CA) and quantified by BioImage Visage 110s (Genomic Solution Inc., Ann Arbor, Michigan).

Correlation Studies in HLM

Human liver microsomes were incubated with T₄ as described above to determine the correlation between T₄ glucuronidation and UGT1A1 and UGT1A9
30 activities. UGT1A1 and UGT1A9 activities were determined by incubating HLM with SN-38 and flavopiridol, respectively, as previously described (Iyer *et al.*, 1999;

Ramirez *et al.*, 2002). The six samples were not assayed for flavopiridol glucuronidation because of insufficient HLM tissue.

Genotyping of UGT1A1 -53(TA)₆₋₇ and UGT1A9 -118T_{9>10} Polymorphisms

The liver samples had been previously genotyped for UGT1A1 -53(TA)_{6>7} (UGT1A1*28) and UGT1A9 -118T_{9>10} (UGT1A9*22) polymorphisms (Innocenti *et al.*, 2002; Innocenti *et al.*, 2005).

Data Analysis and Statistics

Activities are presented as mean $\pm$ SE of a single experiment performed in duplicate unless otherwise stated. Apparent Michaelis-Menten parameters were estimated by nonlinear regression using the equation $V = V_{\max} * S / (K_m + S)$ where $V_{\max}$ is the maximum velocity of the enzyme, S is the substrate concentration and K_m is the Michaelis-Menten constant. Frequency distribution analyses and the Kolmogorov-Smirnov test showed that the distribution of T₄G formation rates was not normal (KS distance=0.2237, p=0.0081), while the glucuronidation levels of SN-38 and flavopiridol were normally distributed (KS distance=0.07326; p>0.10 and KS distance=0.1151; p>0.10, respectively). T₄G formation rates were log transformed prior to statistical analyses, after which they passed the normality test (KS=0.09980, p>0.10). Pearson correlations were used to test the level of correlation between glucuronidation levels of T₄ and other UGT substrates. To investigate further the relationship between the glucuronidation of T₄G and both SN-38 and flavopiridol glucuronidation, additive and interactive multivariate models were tested. To test for significance in the relationship between genotypes and T₄G formation, we performed a one way analysis of variance (ANOVA test) followed by the Tukey test and trend analysis. The threshold value for statistical significance was set at 0.05. Statistical calculations were performed using Graph Pad Prism version 4.00 for Windows (GraphPad Software Inc., San Diego, CA), Stata/SE 9.0 for Windows (StataCorp, College Station, TX) and the R package (R Development Core Team, 2005).

Results

T₄ Glucuronidation in HLM

T₄G formation rates are linear with microsomal protein concentration up to 1.5 mg/mL and time up to 2 hours. T₄ glucuronidation follows Michaelis-Menten

kinetics through 300 μM T_4 . T_4G formation slowly decreased at T_4 concentrations higher than 300 μM . Hence, kinetic parameters were determined using the Michaelis-Menten equation for 10-300 μM . The apparent K_m , V_{max} and intrinsic clearance were determined to be $85.7 \pm 14.6 \mu\text{M}$ (mean $\pm$ SE), $6.7 \pm 0.4 \text{ pmol/min/mg}$ (mean $\pm$ SE), and 0.078 ml/min/mg (FIG. 6). Interindividual variation in T_4G formation was analyzed in 54 Caucasian HLMs. The median T_4G formation was 2.6 pmol/min/mg, range 0.8 - 10.5 pmol/min/mg.

T₄ Glucuronidation in Microsomes Expressing Human UGT cDNA

T_4G was formed mainly by UGT1A8 (mean T_4G formation = 119 pmol/min/mg protein) followed by UGT1A10 (81.5 pmol/min/mg protein) > UGT1A3 (30.2 pmol/min/mg protein) > UGT1A1 (15.0 pmol/min/mg protein). Minor T_4G production was observed with UGT1A7 (4.7 pmol/min/mg protein) and UGT1A9 (4.1 pmol/min/mg protein). Negligible T_4 glucuronidation was observed with UGT1A4, UGT1A6, UGT2B4, UGT2B7, UGT2B15 and UGT2B17 (FIG. 7A-B).

To better characterize their relative capacities for T_4 glucuronidation, kinetic parameters were determined in the four isoforms showing the highest activity in the screen and also in UGT1A9 because it has previously been reported to play a role in T_4 glucuronidation (Visser et al., 1993a; Findlay et al., 2000). Each UGT initially followed Michaelis-Menten kinetics, but a pattern of decreasing T_4G formation was observed in the high end of the T_4 concentration range. Kinetic parameters are shown in Table 1 and were calculated using the Michaelis-Menten model with T_4 concentrations ranging to the T_4 concentration where saturation was first observed (400 μM for UGT1A1, 150 μM for UGT1A3 and UGT1A8, 225 μM for UGT1A9 and 90 μM for UGT1A10). The highest intrinsic clearance was observed with UGT1A10 > UGT1A8 > UGT1A3 > UGT1A1 > UGT1A9 (Table 4).

TABLE 4

Apparent kinetic parameters for the formation of T₄G by human UGT enzymes

UGT	Apparent K _m (μM)	Relative UGT protein level	<u>T₄G formation</u>		<u>T₄G formation</u> (normalized by protein levels)	
			V _{max} (pmol/min/mg)	Intrinsic clearance (mL/min/mg)	V _{max} (pmol/min/mg)	Intrinsic clearance (mL/min/mg)
1A1	144.2 ± 29.8	0.803	13.7 ± 1.2	0.095	17.1 ± 1.5	0.118
1A3	36.7 ± 9.2	0.428	7.1 ± 0.7	0.195	16.7 ± 1.5	0.454
1A8	67.3 ± 15.5	0.718	71.9 ± 7.4	1.068	100 ± 10	1.487
1A9	58.5 ± 24.9	0.595	1.1 ± 0.2	0.019	1.9 ± 0.3	0.032
1A10	21.8 ± 4.6	0.238	12.8 ± 0.9	0.586	53.6 ± 3.9	2.466

All incubations were performed in duplicate. To better assess the relative T₄ glucuronidation activities of UGT isoforms, V_{max} was normalized for relative protein expression by western blot. The UGT1A standard was designated as the basal UGT expression level.

Correlation Studies in HLM

Correlation studies were performed with substrates for UGT1A1 and UGT1A9 (FIG. 8A-B). A high level of correlation was observed between T₄ and SN-38 glucuronidation levels (Pearson $r=0.82$, $P<0.0001$, $n=54$). T₄ and flavopiridol glucuronide formation levels were also significantly correlated, but at a lower extent compared with SN-38 glucuronidation (Pearson $r=0.55$, $P<0.0001$, $n=48$). The interactive model $\log(T_4G) \sim \text{SN-38G} * \text{flavopiridol-7-G}$ was found to be significant (multiple $r^2=0.76$, $P<0.001$).

10 *Association of T₄G Formation Phenotype with UGT1A1 -53(TA)_{6>7} and UGT1A9 -118T_{9>10} Polymorphisms*

Significant differences between UGT1A1 -53(TA)_{6>7} genotypes 6/6, 6/7 and 7/7 were observed (ANOVA, $P=0.003$, $n=54$). In a post hoc Tukey multiple comparison test, 7/7 showed significantly lower T₄G formation than 6/6 ($P<0.01$). There was not a significant difference between the other two pairs. There was a significant trend in T₄G formation between UGT1A1 -53(TA)_{6>7} genotypes 6/6, 6/7, and 7/7 with $6/6 > 6/7 > 7/7$ (trend analysis $P=0.002$, $n=54$) (FIG. 9). No significant variation in T₄G formation was observed between UGT1A9 -118T_{9>10} genotypes 9/9, 9/10, and 10/10 ($P>0.05$, $n=48$).

20 This is the first study to conduct a full screen of functional and commercially available human UGTs for T₄ glucuronidation. Past UGT screenings had been limited to UGT1A1, UGT1A6 and UGT1A9. Our data demonstrates that UGT1A1 is the main enzyme involved in T₄ glucuronidation in the human liver.

25 Concerning the extrahepatic UGTs, our experiments suggest that UGT1A8 and UGT1A10 are involved in the glucuronidation of T₄. Both enzymes had relatively high catalytic efficiencies when compared to that of UGT1A1. UGT1A8 and UGT1A10 are exclusively expressed in the gastrointestinal tract (Tukey and Strassburg, 2000).

* * * *

30

All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure.

While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

U.S. Patent 4,582,788
U.S. Patent 4,656,127
U.S. Patent 4,659,774
U.S. Patent 4,682,195
U.S. Patent 4,683,194
U.S. Patent 4,683,195
U.S. Patent 4,683,202
U.S. Patent 4,683,202
U.S. Patent 4,683,202
U.S. Patent 4,800,159
U.S. Patent 4,816,571
U.S. Patent 4,883,750
U.S. Patent 4,946,773
U.S. Patent 4,959,463
U.S. Patent 5,141,813
U.S. Patent 5,264,566
U.S. Patent 5,279,721
U.S. Patent 5,428,148
U.S. Patent 5,554,744
U.S. Patent 5,574,146
U.S. Patent 5,602,244
U.S. Patent 5,645,897
U.S. Patent 5,705,629
U.S. Patent 5,840,873
U.S. Patent 5,843,640
U.S. Patent 5,843,650
U.S. Patent 5,843,651
U.S. Patent 5,843,663

U.S. Patent 5,846,708
U.S. Patent 5,846,709
U.S. Patent 5,846,717
U.S. Patent 5,846,726
U.S. Patent 5,846,729
U.S. Patent 5,846,783
U.S. Patent 5,849,481
U.S. Patent 5,849,483
U.S. Patent 5,849,486
U.S. Patent 5,849,487
U.S. Patent 5,849,497
U.S. Patent 5,849,546
U.S. Patent 5,849,547
U.S. Patent 5,851,770
U.S. Patent 5,851,772.
U.S. Patent 5,853,990
U.S. Patent 5,853,992
U.S. Patent 5,853,993
U.S. Patent 5,856,092
U.S. Patent 5,858,652
U.S. Patent 5,861,244
U.S. Patent 5,863,732
U.S. Patent 5,863,753
U.S. Patent 5,866,331
U.S. Patent 5,866,337
U.S. Patent 5,866,366
U.S. Patent 5,900,481
U.S. Patent 5,905,024
U.S. Patent 5,910,407
U.S. Patent 5,912,124
U.S. Patent 5,912,145
U.S. Patent 5,912,148
U.S. Patent 5,916,776
U.S. Patent 5,916,779

U.S. Patent 5,919,626
U.S. Patent 5,919,630
U.S. Patent 5,922,574
U.S. Patent 5,925,517
U.S. Patent 5,925,525
U.S. Patent 5,928,862
U.S. Patent 5,928,869
U.S. Patent 5,928,870
U.S. Patent 5,928,905
U.S. Patent 5,928,906
U.S. Patent 5,929,227
U.S. Patent 5,932,413
U.S. Patent 5,932,451
U.S. Patent 5,935,791
U.S. Patent 5,935,825
U.S. Patent 5,939,291
U.S. Patent 5,942,391
U.S. Patent 5,952,174
U.S. Patent 6,395,481
U.S. Patent 6,395,481
U.S. Patent 6,472,157
U.S. Patent 6,472,157

Ausubel *et al.*, In: *Current Protocols in Molecular Biology*, Green Pub. Assoc., Inc.,
and John Wiley & Sons, Inc., NY, (I):2.10.3, 1989.

Basu *et al.*, *J Biol Chem* 279:1429-1441, 2004

Beetstra *et al.*, *Endocrinology* 128:741-746, 1998.

Bellus, *J. Macromol. Sci. Pure Appl. Chem.*, A31(1): 1355-1376, 1994.

Beutler *et al.*, *Proc Natl Acad Sci U S A* 95:8170-8174, 1991.

Bosma *et al.*, *Hepatology*, 15:941-947, 1992.

Bosma *et al.*, *N. Eng. J. Med.*, 333:1171-1175, 1995.

Burchell, *Am J Pharmacogenomics* 3:37-52, 2003.

Clarke *et al.*, *Handbook of Experimental Pharmacology*, 112:3-43, 1994.

de Arruda *et al.*, *Expert. Rev. Mol. Diagn.*, 2:487-496, 2002.

- European Appln. 201,184
European Appln. 237,362
European Appln. 258,017
European Appln. 329 822
European Appln. 50,424
European Appln. 84,796
European Patent 266,032
European Patent 320 308
Findlay *et al.*, *J Clin Endocrinol Metab* 85:2879-2883, 2000.
Fisher *et al.*, *Drug Metab Rev* 33:273-297, 2003.
French Appln. 2,650,840
Froehler *et al.*, *Nucleic Acids Res.*, 14(13):5399-5407, 1986.
Frohman, In: *PCR Protocols: A Guide To Methods And Applications*, Academic Press, N.Y., 1990.
Girard *et al.*, *Pharmacogenetics* 14:501-515, 2004.
Great Britain Appln. 2 202 328
Halushka *et al.*, *Nat. Genet.*, 22(3):239-247, 1999.
Hollowell *et al.*, *J. Clin. Endocrinol. Metab.*, 87(2):489-499, 2002.
Hood *et al.*, *Toxicol Appl Pharmacol* 160:163-170, 1999.
Inazuka *et al.*, *Genome Res*, 7(11):1094-1103, 1997.
Innis *et al.*, *Proc. Natl. Acad. Sci. USA*, 85(24):9436-9440, 1988.
Innocenti *et al.*, *Pharmacogenetics* 12:725-733, 2002.
Innocenti *et al.*, *J. Clin. Oncol.*, 22:1382, 2004.
Innocenti *et al.*, *Pharmacogenet Genomics* 15:295-301, 2005.
Iyer *et al.*, *Clin Pharmacol Ther* 65:576-582, 1999.
Jemnitz *et al.*, *J Chromatogr B Biomed Appl* 681:385-389, 1996.
Johnson *et al.*, *Nat. Genet.*, 29(2):233-237, 2001.
Kato *et al.*, *Drug Metab Dispos* in press.
Ke and Cardon, *Bioinformatics*, 19(2):287-288, 2003.
Klein *et al.*, *N. Engl. J. Med.*, 344(7):501-509, 2001.
Komher, *et al.*, *Nucl. Acids. Res.* 17:7779-7784, 1989.
Kuppuswamy, *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:1143-1147, 1991.
Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:1173, 1989.
Kwok and Chen, *Curr Issues Mol. Biol.*, Apr;5(2):43-60, 2003.

- Kwok *et al.*, *Genomics*, 23(1):138-144, 1994.
- Kwok, *Annu. Rev. Genomics Hum. Genet.*, 2:235-258, 2001.
- Landegren, *et al.*, *Science* 241:1077-1080, 1988.
- Liu *et al.*, *J Pharmacol Exp Ther* 273:977-985, 1995.
- Lu *et al.*, *Biopolymers*, 73:606-613, 2004.
- Maxam, *et al.*, *Proc. Natl. Acad. Sci. USA*, 74:560, 1977.
- Monaghan *et al.*, *Lancet*, 347:578-581, 1996.
- Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273, 1986.
- Nickerson *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:8923-8927, 1990.
- Nyren *et al.*, *Anal. Biochem.* 208:171-175, 1993.
- Ohara *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:5673-5677, 1989.
- Owens and Ritter, *Prog. Nucleic Acid Res. Mol. Biol.*, 51:305-338, 1995.
- PCT Appln. PCT/US87/00880
- PCT Appln. PCT/US89/01025,
- PCT Appln. WO 88/10315
- PCT Appln. WO 89/06700
- PCT Appln. WO91/02087
- PCT Appln. WO92/15712
- Prezant *et al.*, *Hum. Mutat.*, 1:159-164, 1992.
- Ramirez *et al.*, *Pharm Res* 19:588-594, 2002.
- Ritter *et al.*, *J. Biol. Chem.*, 267:3257-3261, 1992.
- Roberts *et al.*, *Lancet* 363:793-803, 2004.
- Sambrook *et al.*, *In: Molecular cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
- Sanger *et al.*, *J. Molec. Biol.*, 94:441, 1975.
- Sokolov, *Nucl. Acids Res.* 18:3671, 1990.
- Stevens *et al.*, *Biotechniques*, 34:198-203, 2003.
- Stone *et al.*, *Drug Metab Dispos* 31:1086-1089, 2003.
- Syvanen *et al.*, *Genomics* 8:684-692, 1990.
- Te *et al.*, *Transplantation*, 69:1882-1886, 2000.
- Taillon-Miller *et al.*, *Genome Res*, 8(7):748-754, 1998.
- Tukey *et al.*, *Annu Rev Pharmacol Toxicol* 40:581-616, 2000.
- Ugozzoli *et al.*, *GATA* 9:107-112, 1992.
- Visser, *Acta Med Austriaca* 23:10-16, 1996.

Visser *et al.*, *FEBS Lett* 324:358-360, 1993a.

Visser *et al.*, *Endocrinology* 133:2177-2186, 1993b.

Walker *et al.*, *Nucleic Acids Res.*, 20(7):1691-1696, 1992.

Yamanaka *et al.*, *Pharmacogenetics* 14:329-332, 2004.

CLAIMS

1. A method for prescribing an amount of L-thyroxine for a patient comprising:
 - a) determining the number of TA repeats in one or both alleles of the *uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1)* promoter, wherein the presence of seven TA repeats correlates with decreased inactivation of L-thyroxine; and,
 - b) prescribing an average daily dose of L-thyroxine for the patient.
2. The method of claim 1, wherein the number of TA repeats is determined by evaluating directly the TA repeats in one or both alleles of the *UGT1A1* promoter.
3. The method of claim 1, wherein if the patient has seven TA repeats in at least one allele, the average daily dose prescribed for the patient is lower than if the patient has six TA repeats in both alleles.
4. The method of claim 1, wherein if the patient has six TA repeats in at least one allele, the average daily dose prescribed for the patient is higher than if the patient has seven TA repeats in both alleles.
5. The method of claim 1, wherein both alleles of the *UGT1A1* promoter are evaluated.
6. The method of claim 5, wherein if the patient has seven TA repeats in both alleles, the average daily dose prescribed for the patient is lower than if the patient has six TA repeats in one or both alleles.
7. The method of claim 5, wherein if the patient has six TA repeats in one allele and seven TA repeats in the other allele, the average daily dose prescribed for the patient is lower than if the patient has six TA repeats in both alleles, but is higher than if the patient has seven TA repeats in both alleles.
8. The method of claim 1, further comprising obtaining DNA from the patient.

9. The method of claim 1, further comprising identifying a patient in need of L-thyroxine.
10. The method of claim 1, further comprising identifying a starting daily dose for a patient in need of L-thyroxine.
11. The method of claim 9, wherein the patient does not have or is not suspected of having Gilbert's Syndrome.
12. The method of claim 1, further comprising titrating the amount of L-thyroxine in the patient.
13. The method of claim 2, wherein the number of TA repeats is evaluated by sequencing one or both alleles.
14. The method of claim 2, wherein the number of TA repeats is evaluated by amplifying the number of TA repeats.
15. The method of claim 14, wherein the number of TA repeats is amplified by polymerase chain reaction.
16. The method of claim 15, wherein the number of TA repeats is determined by gel electrophoresis.
17. The method of claim 1, further comprising evaluating one or more other polymorphisms in a UGT enzyme gene, wherein the polymorphism affects glucuronidation.
18. The method of claim 17, wherein the polymorphism is in a *UGT1A* gene.
19. The method of claim 18, wherein the polymorphism is in a *UGT1A3* gene.
20. The method of claim 18, wherein the polymorphism is in a *UGT1A8* gene.

21. The method of claim 18, wherein the polymorphism is in a *UGT1A10* gene.
22. The method of claim 18, wherein the polymorphism is selected from the group consisting of: *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T.
23. The method of claim 1, further comprising determining the presence of seven TA repeats in at least one allele.
24. The method of claim 1, further comprising determining the presence of seven TA repeats in both alleles.
25. The method of claim 1, further comprising determining the presence of six TA repeats in both alleles.
26. The method of claim 1, wherein the number of TA repeats is determined by evaluating a polymorphism at one or more of *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A.
27. A method for reducing the risk of hypothyroidism or hyperthyroidism in a patient comprising:
 - a) predicting the level of L-thyroxine glucuronidation in the patient;
 - b) prescribing an average daily dose of L-thyroxine to the patient based on the level of L-thyroxine glucuronidation predicted in the patient.
28. The method of claim 27, wherein the level of L-thyroxine glucuronidation is predicted by determining the number of TA repeats in one or both alleles of a *UGT1A1* promoter, wherein the presence of six TA repeats in one allele and seven TA repeats in the other allele predicts a level of L-thyroxine glucuronidation that is lower than the level predicted for the presence of six TA repeats in both alleles, but higher than the level predicted for the presence of seven TA repeats in both alleles.
29. The method of claim 27, further comprising identifying a patient in need of L-thyroxine.

30. The method of claim 28, wherein the number of TA repeats is determined by evaluating directly the TA repeats in one or both alleles of the *UGT1A1* promoter.
31. The method of claim 27, wherein if the patient has seven TA repeats in at least one allele, the average daily dose prescribed for the patient is lower than if the patient has six TA repeats in both alleles.
32. The method of claim 27, wherein if the patient has six TA repeats in one allele, the average daily dose prescribed for the patient is higher than if the patient has seven TA repeats in both alleles.
33. The method of claim 27, wherein if the patient has six TA repeats in one allele and seven TA repeats in the other allele, the average daily dose prescribed for the patient is lower than if the patient has six TA repeats in both alleles, but is higher than if the patient has seven TA repeats in both alleles.
34. The method of claim 28, wherein both alleles of the *UGT1A1* promoter are evaluated.
35. The method of claim 34, wherein the average daily dose prescribed for the patient is lower if the patient has seven TA repeats in both alleles than if the patient has six TA repeats in one or both alleles.
36. The method of claim 27, further comprising obtaining DNA from the patient.
37. The method of claim 27, further comprising identifying a patient in need of L-thyroxine.
38. The method of claim 27, wherein the patient is prescribed a starting daily dose of L-thyroxine.
39. The method of claim 27, wherein the patient does not have or is not suspected of having Gilbert's Syndrome.

40. The method of claim 27, further comprising titrating the amount of L-thyroxine in the patient.
41. The method of claim 30, wherein the number of TA repeats is evaluated by sequencing one or both alleles.
42. The method of claim 30, wherein the number of TA repeats is evaluated by amplifying the number of TA repeats.
43. The method of claim 42, wherein the number of TA repeats is amplified by polymerase chain reaction.
44. The method of claim 43, wherein the number of TA repeats is determined by gel electrophoresis.
45. The method of claim 28, further comprising evaluating one or more other polymorphisms in a UGT enzyme gene, wherein the polymorphism affects glucuronidation.
46. The method of claim 45, wherein the polymorphism is in a *UGT1A* gene.
47. The method of claim 46, wherein the polymorphism is in a *UGT1A3* gene.
48. The method of claim 46, wherein the polymorphism is in a *UGT1A8* gene.
49. The method of claim 46, wherein the polymorphism is in a *UGT1A10* gene.
50. The method of claim 45, wherein the polymorphism is selected from the group consisting of: *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T.
51. The method of claim 28, further comprising determining the presence of seven TA repeats in at least one allele.

52. The method of claim 28, further comprising determining the presence of seven TA repeats in both alleles.
53. The method of claim 28, further comprising determining the presence of six TA repeats in both alleles.
54. The method of claim 28, wherein the number of TA repeats is determined by evaluating a polymorphism at one or more of *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A.
55. A method for evaluating L-thyroxine glucuronidation capacity in a patient comprising:
- a) obtaining DNA from the patient;
 - b) determining from the DNA the number of TA repeats in one or both alleles of a *UGT1A1* promoter;
 - c) informing the patient or the patient's physician about the patient's L-thyroxine glucuronidation capacity based on the number of TA repeats.
56. The method of claim 55, wherein the presence of seven TA repeats in one allele predicts a lower level of L-thyroxine glucuronidation than the presence of six TA repeats in both alleles.
57. The method of claim 55, wherein DNA from multiple patients is obtained and the number of TA repeats determined for each patient.
58. The method of claim 55, wherein the number of TA repeats is determined by evaluating directly the TA repeats in one or both alleles of the *UGT1A1* promoter.
59. The method of claim 55, further comprising prescribing L-thyroxine to the patient.
60. The method of claim 59, further comprising prescribing an average daily dose of L-thyroxine for a patient having seven TA repeats in at least one allele, wherein the

average daily dose is lower than the dose for a patient having six TA repeats in both alleles.

61. The method of claim 59, further comprising prescribing an average daily dose of L-thyroxine for a patient having six TA repeats in at least one allele, wherein the average daily dose is higher than the dose for a patient having seven TA repeats in both alleles.

62. The method of claim 55, wherein both alleles of the *UGT1A1* promoter are evaluated.

63. The method of claim 62, further comprising prescribing an average daily dose of L-thyroxine for a patient having six TA repeats in both alleles, wherein the average daily dose is higher than the dose for a patient having seven TA repeats in one or both alleles.

64. The method of claim 62, further comprising prescribing an average daily dose of L-thyroxine for a patient having seven TA repeats in both alleles, wherein the average daily dose is lower than the dose for a patient having six TA repeats in one or both alleles.

65. The method of claim 62, further comprising prescribing an average daily dose of L-thyroxine for a patient having six TA repeats in one allele and seven TA repeats in the other allele, wherein the average daily dose is lower than the dose for a patient having seven TA repeats in both alleles but higher than the dose for a patient having six TA repeats in both alleles.

66. The method of claim 55, further comprising identifying a patient in need of L-thyroxine.

67. The method of claim 55, further comprising identifying a starting daily dose for a patient in need of L-thyroxine.

68. The method of claim 55, wherein the patient does not have or is not suspected of having Gilbert's Syndrome.
69. The method of claim 59, further comprising titrating the amount of L-thyroxine in the patient after the patient has taken L-thyroxine.
70. The method of claim 58, wherein the number of TA repeats is evaluated by sequencing one or both alleles.
71. The method of claim 58, wherein the number of TA repeats is evaluated by amplifying the number of TA repeats.
72. The method of claim 71, wherein the number of TA repeats is amplified by polymerase chain reaction.
73. The method of claim 72, wherein the number of TA repeats is determined by gel electrophoresis.
74. The method of claim 55, further comprising evaluating one or more other polymorphisms in a UGT enzyme gene, wherein the polymorphism affects glucuronidation.
75. The method of claim 74, wherein the polymorphism is a *UGT1A* gene.
76. The method of claim 75, wherein the polymorphism is in a *UGT1A3* gene.
77. The method of claim 75, wherein the polymorphism is in a *UGT1A8* gene.
78. The method of claim 75, wherein the polymorphism is in a *UGT1A10* gene.
79. The method of claim 74, wherein the polymorphism is selected from the group consisting of: *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T.

80. The method of claim 55, further comprising determining the presence of seven TA repeats in at least one allele.
81. The method of claim 55, further comprising determining the presence of seven TA repeats in both alleles.
82. The method of claim 55, wherein the number of TA repeats is determined by evaluating a polymorphism at one or more of *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A.
83. A method for optimizing the dose of L-thyroxine for a patient comprising:
- a) identifying a patient in need of L-thyroxine;
 - b) obtaining DNA from the patient;
 - c) determining from the DNA the number of TA repeats in one or both alleles of a *UGT1A1* promoter; and
 - d) prescribing an average daily dose of L-thyroxine to the patient based on the number of TA repeats.
84. The method of claim 83, wherein the number of TA repeats is determined by evaluating directly the TA repeats in one or both alleles of the *UGT1A1* promoter.
85. The method of claim 83, wherein if the patient has seven TA repeats, the dose prescribed for the patient is lower than the average daily dose for a patient having six TA repeats.
86. The method of claim 83, wherein if the patient has six TA repeats, the dose prescribed for the patient is higher than the average daily dose for a patient having seven TA repeats.
87. The method of claim 83, wherein both alleles of the *UGT1A1* promoter are evaluated.

88. The method of claim 87, wherein if the patient has seven TA repeats in both alleles, the average daily dose prescribed for the patient is lower than the average daily dose for a patient having six TA repeats in one or both alleles.

89. The method of claim 87, wherein if the patient has six TA repeats in both alleles, the average daily dose prescribed for the patient is lower than the average daily dose for a patient having seven TA repeats in one or both alleles.

90. The method of claim 87, wherein if the patient has six TA repeats in one allele and seven TA repeats in the other allele, the average daily dose prescribed for the patient is lower than for a patient having six TA repeats in both alleles but higher than a patient having seven TA repeats in both alleles.

91. The method of claim 83, further comprising identifying a starting average daily dose for a patient in need of L-thyroxine.

92. The method of claim 83, wherein the patient does not have or is not suspected of having Gilbert's Syndrome.

93. The method of claim 91, further comprising titrating the amount of L-thyroxine in the patient after the patient has taken L-thyroxine.

94. The method of claim 84, wherein the number of TA repeats is evaluated by sequencing one or both alleles.

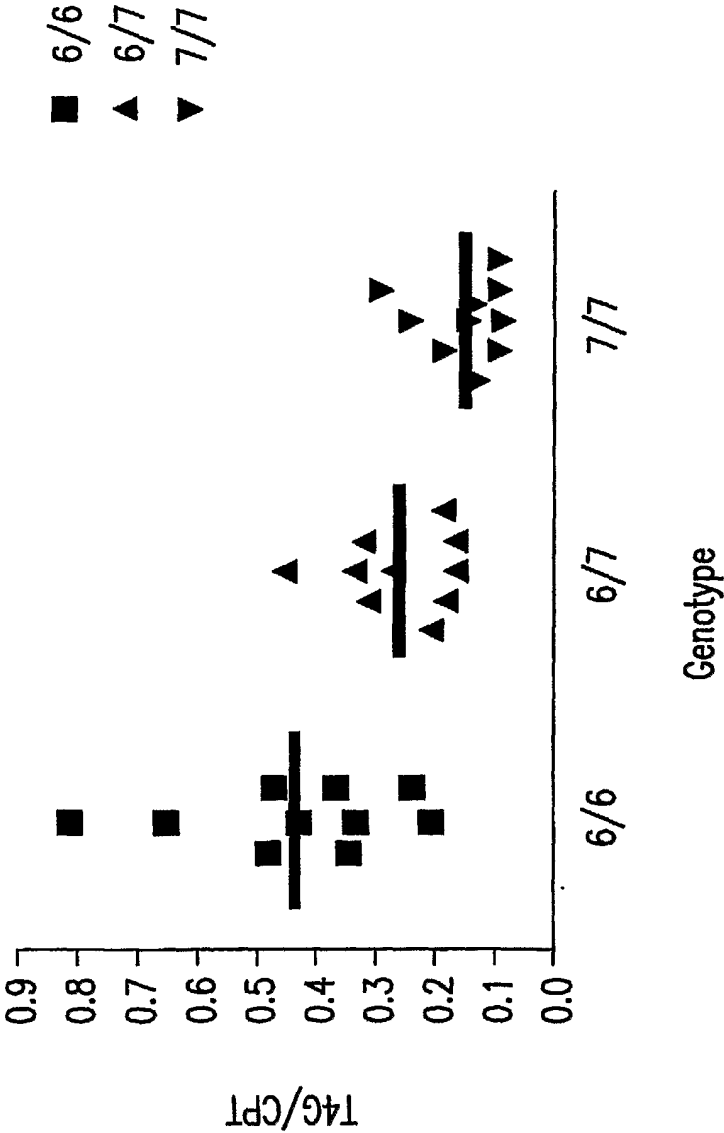
95. The method of claim 84, wherein the number of TA repeats is evaluated by amplifying the number of TA repeats.

96. The method of claim 95, wherein the number of TA repeats is amplified by polymerase chain reaction.

97. The method of claim 96, wherein the number of TA repeats is determined by gel electrophoresis.

98. The method of claim 83, further comprising evaluating one or more other polymorphisms in a UGT enzyme gene, wherein the polymorphism affects glucuronidation.
99. The method of claim 98, wherein the polymorphism is in a *UGT1A* gene.
100. The method of claim 99, wherein the polymorphism is in a *UGT1A3* gene.
101. The method of claim 99, wherein the polymorphism is in a *UGT1A8* gene.
102. The method of claim 99, wherein the polymorphism is in a *UGT1A10* gene.
103. The method of claim 99, wherein the polymorphism is selected from the group consisting of: *UGT1A9* -118(T)_{9>10}, *UGT1A9* -275T>A, and *UGT1A9* -2152C>T.
104. The method of claim 83, further comprising determining the presence of seven TA repeats in at least one allele.
105. The method of claim 83, further comprising determining the presence of seven TA repeats in both alleles.
106. The method of claim 83, wherein the number of TA repeats is determined by evaluating a polymorphism at one or more of *UGT1A1* -3156G>A, *UGT1A1* -3279T>G, *UGT1A1* 211G>A, *UGT1A1* 686C>A.

1A1 (TA)n Genotype vs. Thyroxine Glucuronidation Level



Nonparametric Trend: $7/7 < 6/7 < 6/6$, $P < 0.01$, $z = -3.72$
ANOVA: $P = 0.0002$
Unpaired t test: $7/7$ vs $6/6$ $P = 0.0003$
 $7/7$ vs $6/7$ $P = 0.0113$
 $6/7$ vs $6/6$ $P = 0.0180$

FIG.1

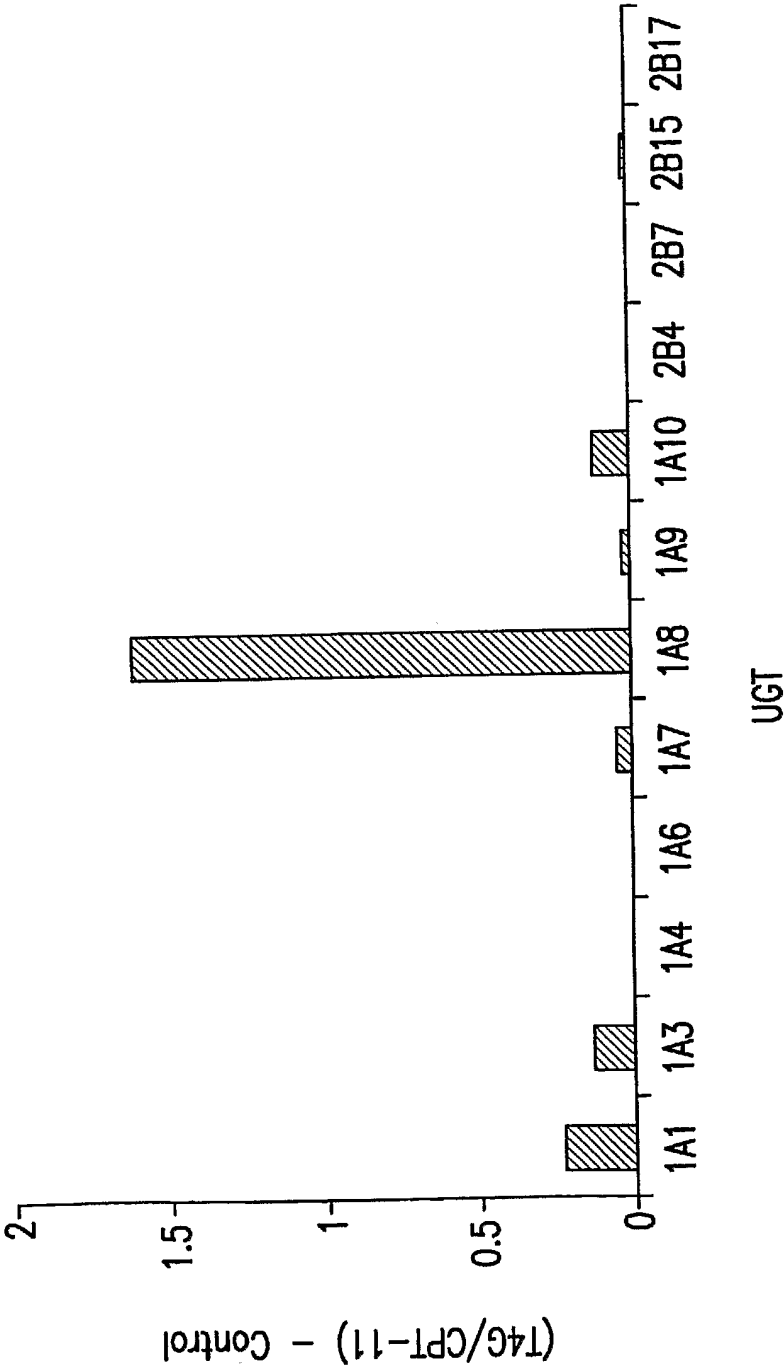


FIG.2

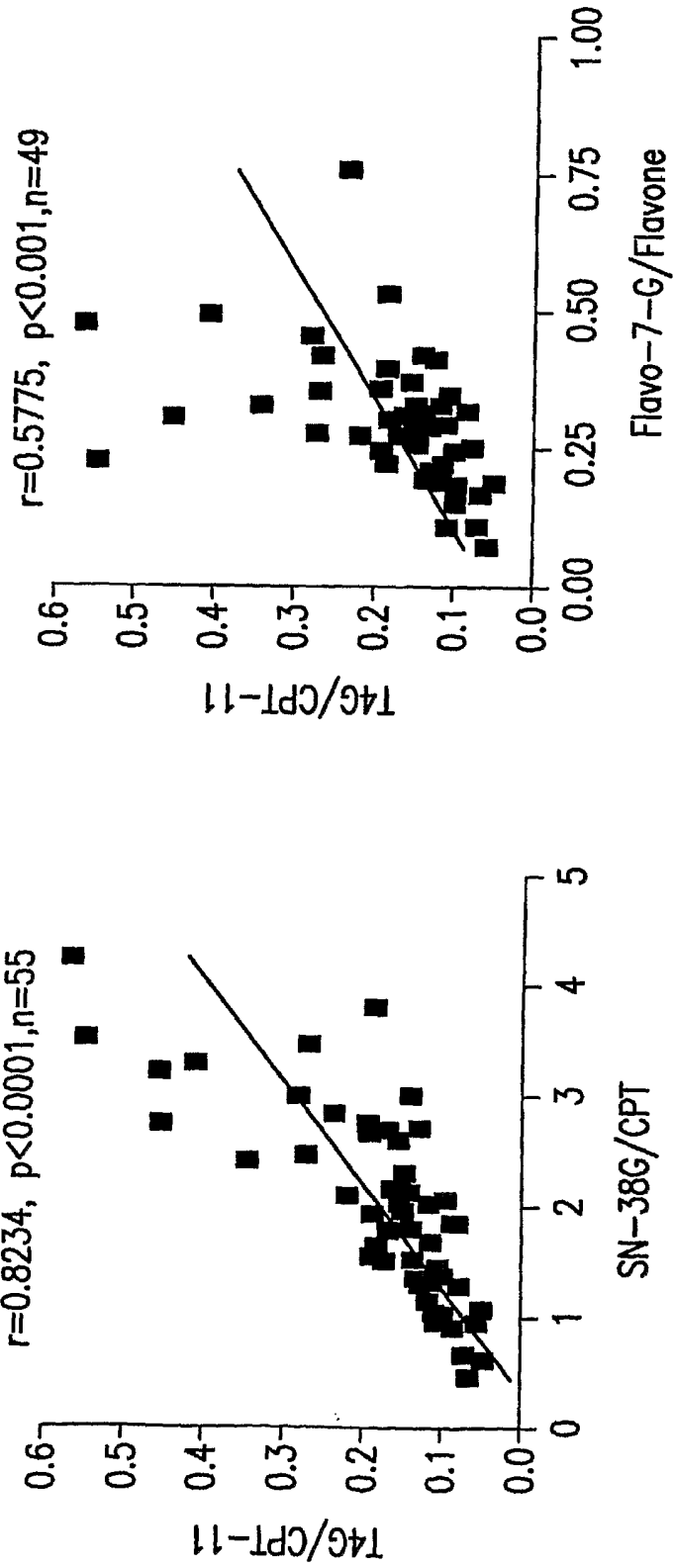


FIG.3B

FIG.3A

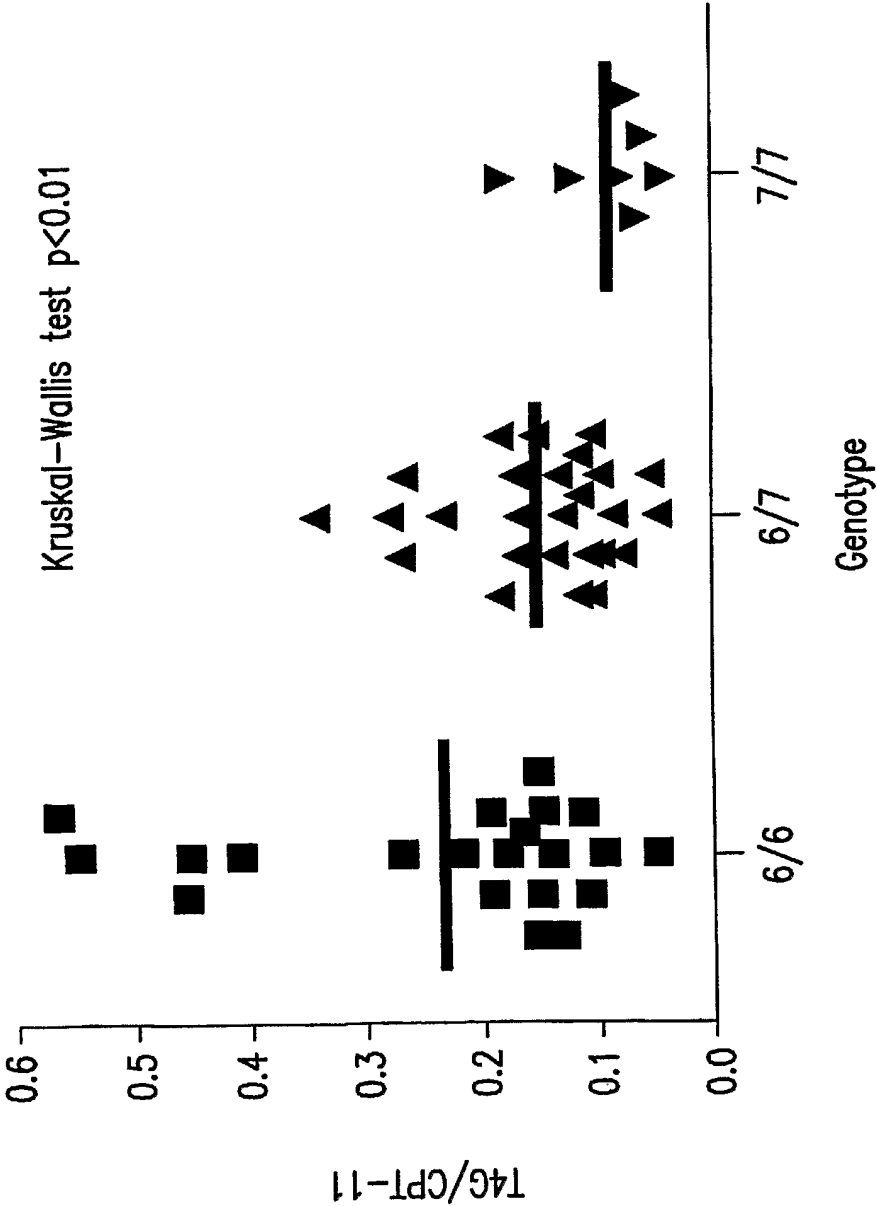


FIG.4

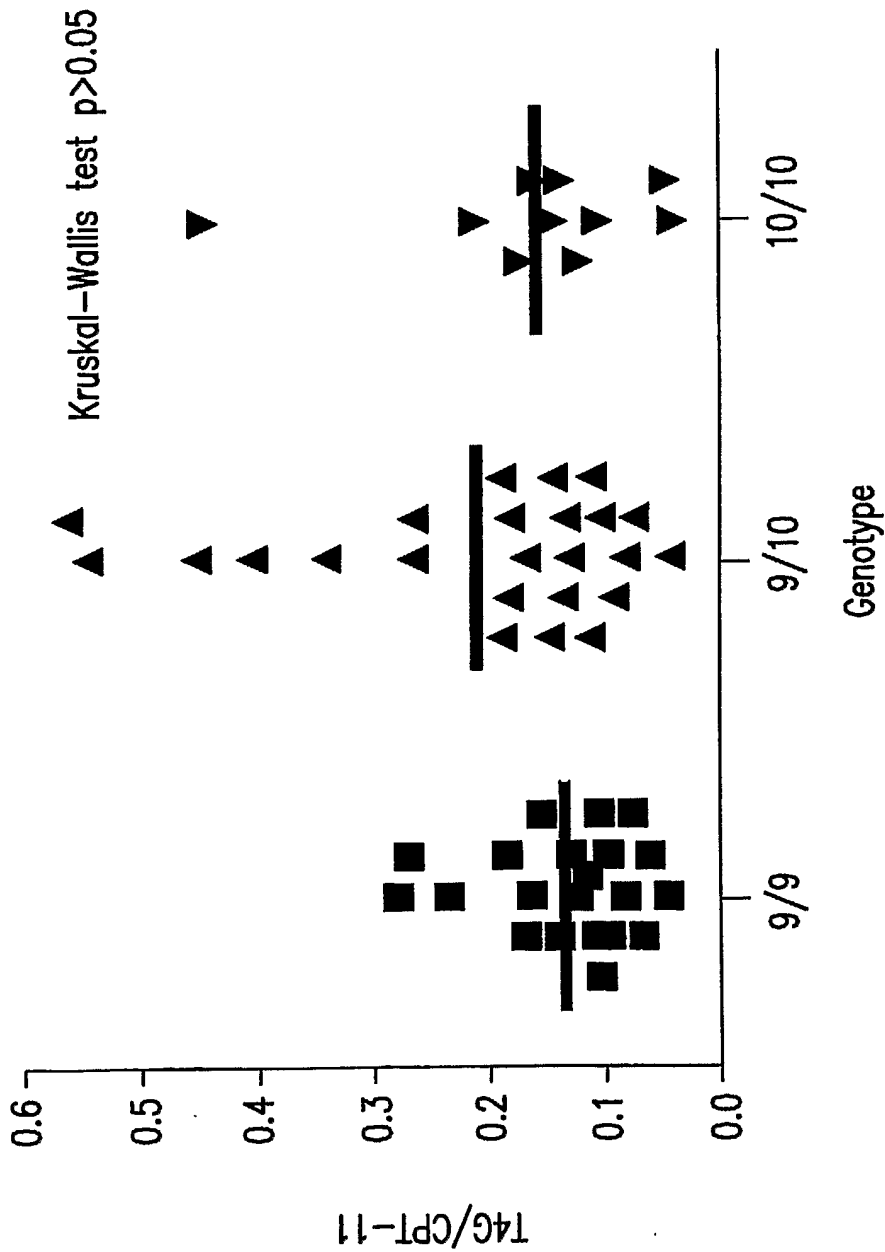


FIG.5

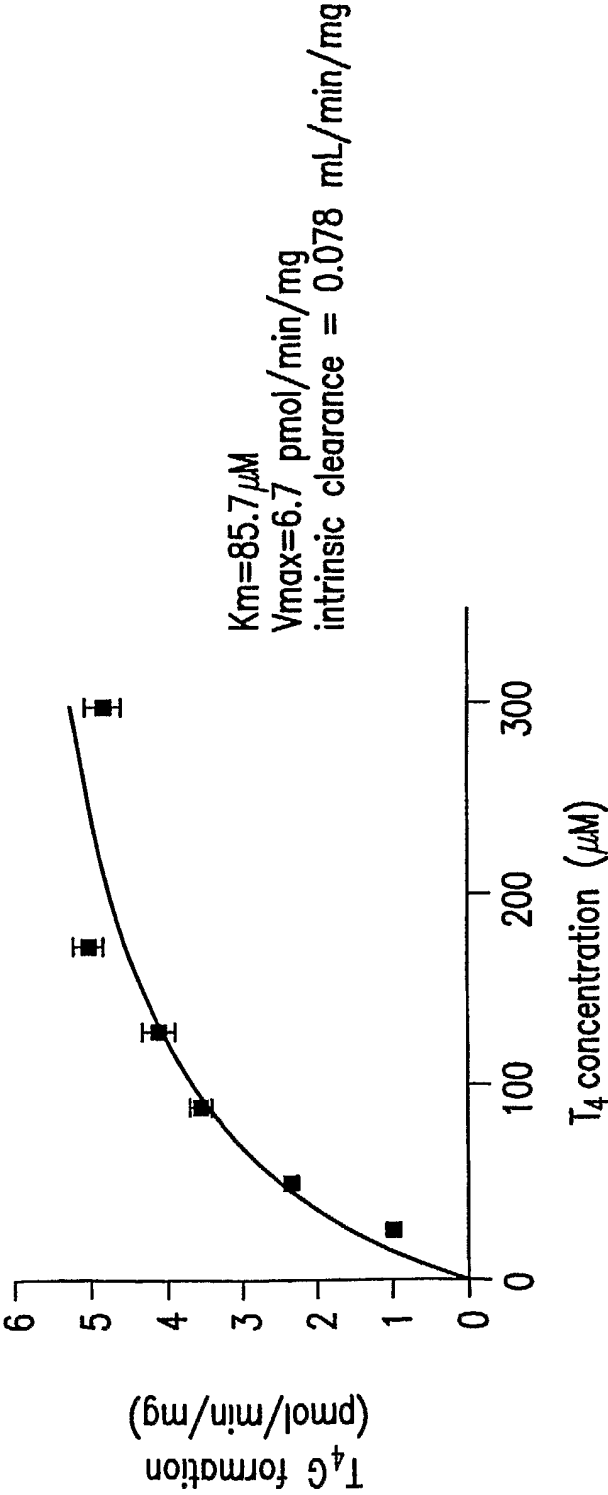
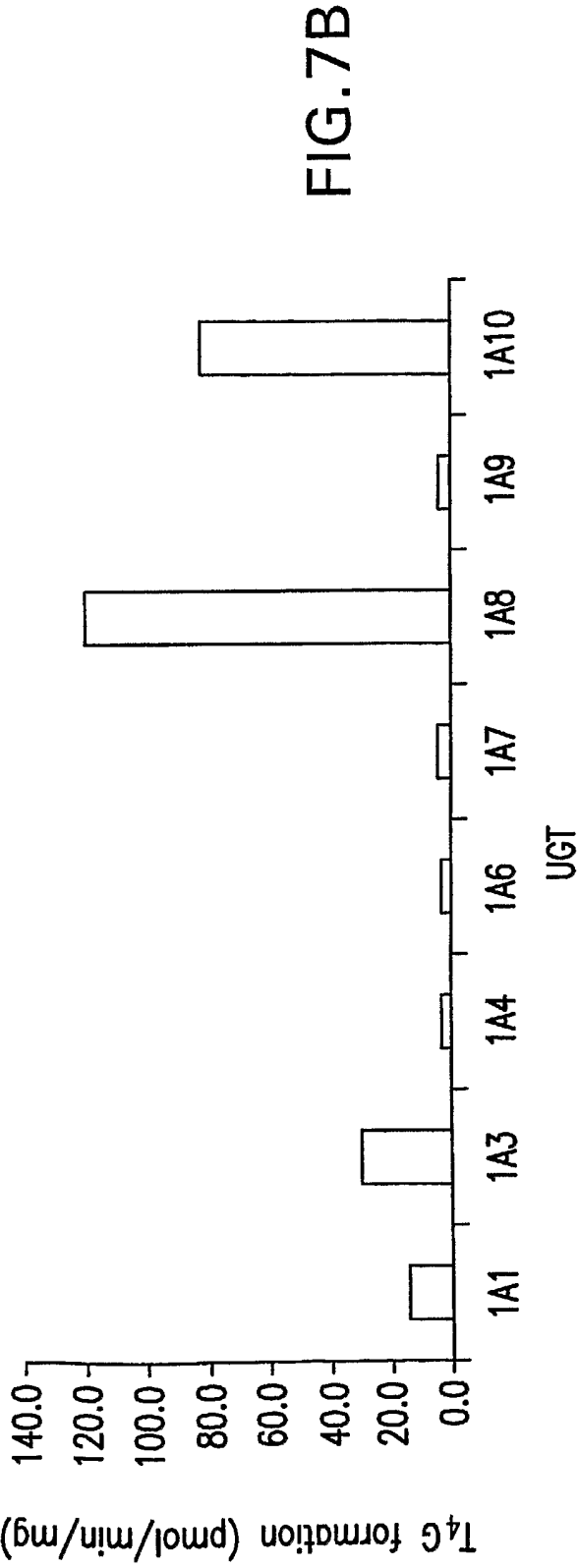
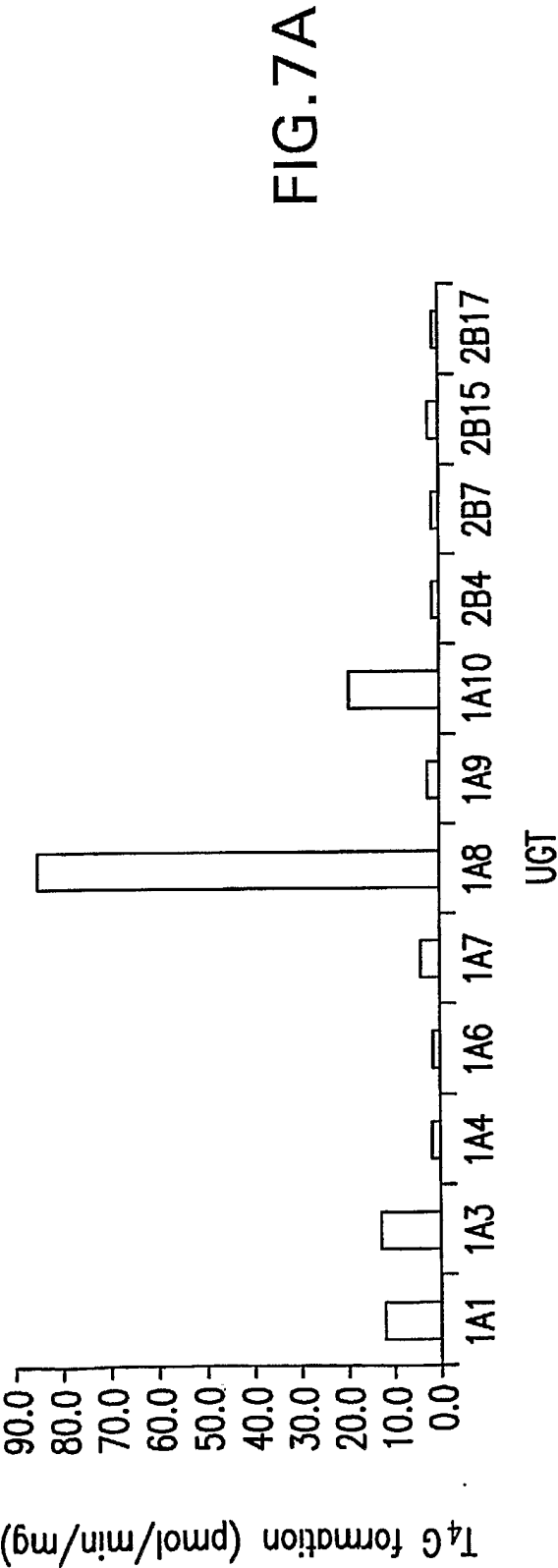


FIG.6



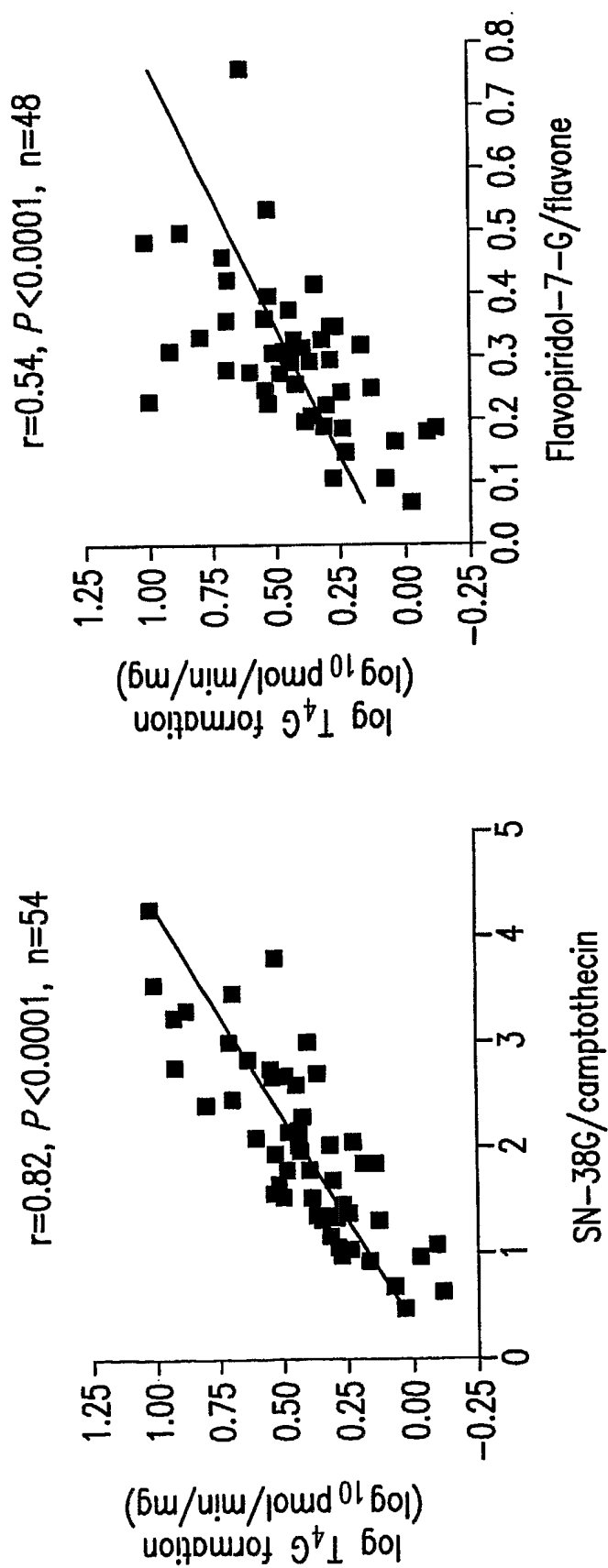


FIG.8B

FIG.8A

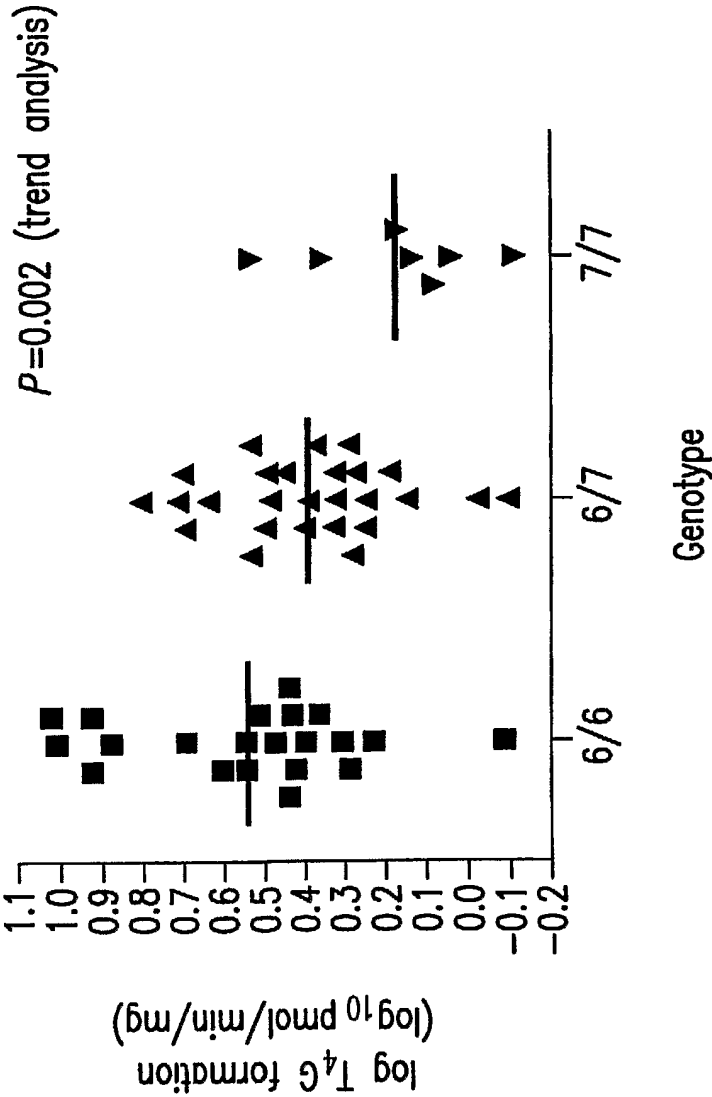


FIG.9

SEQUENCE LISTING

<110> RATAIN, MARK J.
 GRABER, ANDREA YODER
 INNOCENTI, FEDERICO
 RAMIREZ, JACQUELINE

<120> METHODS AND COMPOSITIONS FOR DOSING L-THYROXINE

<130> ARCD:421WO

<140> UNKNOWN

<141> 2006-02-01

<150> 60/658,429

<151> 2005-03-04

<150> 60/648,927

<151> 2005-02-01

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 488

<212> DNA

<213> Homo sapiens

<400> 1

```

caagtga gca ggcagtagc ggggagctgt ggagtgaggca ctcttacagg tttccatggc 60
gaaagcg ggggtacagttgt gttcttttct ttctaaaagg ctttctaaaa agccttctgt 120
ttaattt ctg gaaaagaagc ctaacttggt cactacatag tcgtccttct tcctctctgg 180
taacactt gt tggctctgtg aaataactaat ttaatggatc ctgaggttct ggaagtactt 240
tgctgtgt tc actcaagaat gtgatttgag tatgaaattc cagccagttc aactggtgtt 300
gcctatta ag aaacctaata agctccacc ttctttatct ctgaaagtga actccctgct 360
acctttgt gg actgacagct ttttacagtc acgtgacaca gtcaaacatt aacttggtgt 420
atcgattg gt ttttgccata tatatatata agtaggagag ggcgaacctc tggcaggagc 480
aaaggcgc

```

<210> 2

<211> 490

<212> DNA

<213> Homo sapiens

<400> 2

```

caagtga gca ggcagtagc ggggagctgt ggagtgaggca ctcttacagg tttccatggc 60
gaaagcg ggggtacagttgt gttcttttct ttctaaaagg ctttctaaaa agccttctgt 120
ttaattt ctg gaaaagaagc ctaacttggt cactacatag tcgtccttct tcctctctgg 180
taacactt gt tggctctgtg aaataactaat ttaatggatc ctgaggttct ggaagtactt 240
tgctgtgt tc actcaagaat gtgatttgag tatgaaattc cagccagttc aactggtgtt 300
gcctatta ag aaacctaata agctccacc ttctttatct ctgaaagtga actccctgct 360
acctttgt gg actgacagct ttttatagtc acgtgacaca gtcaaacatt aacttggtgt 420
atcgattg gt ttttgccata tatatatata taagtaggag agggcgaacc tctggcagga 480
gcaaaggcgc

```

<210> 3

<211> 492

<212> DNA

<213> Homo sapiens

<400> 3

```

caagtgagca ggcagtaccg ggggagctgt ggagtgggca ctcttacagg tttccatggc 60
gaaagcgggg gtacagttgt gttcttttct ttctaaaagg ctttctaaaa agccttctgt 120
ttaatttttg gaaaagaagc ctaacttggt cactacatag tcgtccttct tcctctctgg 180
taacacttgt tggctctgtgg aaataactaat ttaatggatc ctgagggttct ggaagtactt 240
tgctgtgttc actcaagaat gtgatttgag tatgaaattc cagccagttc aactgttggt 300
gcctattaag aaacctaata aagctccacc ttctttatct ctgaaagtga actccctgct 360
acctttgtgg actgacagct ttttatagtc acgtgacaca gtcaaacatt aacttggtgt 420
atcgattggg ttttgccata tatatatata tataagtagg agagggcgaa cctctggcag 480
gagcaaaggc gc                                     492

```

<210> 4

<211> 494

<212> DNA

<213> Homo sapiens

<400> 4

```

caagtgagca ggcagtaccg ggggagctgt ggagtgggca ctcttacagg tttccatggc 60
gaaagcgggg gtacagttgt gttcttttct ttctaaaagg ctttctaaaa agccttctgt 120
ttaatttttg gaaaagaagc ctaacttggt cactacatag tcgtccttct tcctctctgg 180
taacacttgt tggctctgtgg aaataactaat ttaatggatc ctgagggttct ggaagtactt 240
tgctgtgttc actcaagaat gtgatttgag tatgaaattc cagccagttc aactgttggt 300
gcctattaag aaacctaata aagctccacc ttctttatct ctgaaagtga actccctgct 360
acctttgtgg actgacagct ttttatagtc acgtgacaca gtcaaacatt aacttggtgt 420
atcgattggg ttttgccata tatatatata tatataagta ggagagggcg aacctctggc 480
aggagcaaag gcgc                                     494

```