Novel *CYP2C9* Promoter Variants and Assessment of Their Impact on Gene Expression*

Melissa A. Kramer[1], Allan E. Rettie, Mark J. Rieder, Erwin T. Cabacungan, and Ronald N. Hines

Departments of Pediatrics (M.A.K., E.T.C., R.N.H.) and Pharmacology and Toxicology (M.A.K., R.N.H.), Medical College of Wisconsin and Children's Research Institute (M.A.K., E.T.C., R.N.H.), Children's Hospital and Health System, Milwaukee, Wisconsin, and Departments of Medicinal Chemistry (A.E.R.)and Genome Sciences (M.J.R.), University of Washington, Seattle, Washington

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Running Title: Functional *CYP2C9* Promoter Variants

Corresponding Author: Ronald N. Hines, Ph.D.

Medical College of Wisconsin

TBRC/CRI/CPPT

8701 Watertown Plank Rd.

Milwaukee WI 53226

414 955-4322

Fax: 414 955-6651

rhines@mcw.edu

Number of Text Pages: 44

Number of Tables: 4

Number of Figures: 2

Number of References: 40

Number of Words in Abstract: 247

Number of Words in Introduction: 666

Number of Words in Discussion: 1418

Non-Standard Abbreviations: 24PDR, 24 sample set from the Coriell Polymorphism Discovery Resource; AP1, activator protein 1; CAR, constitutive androstrance receptor; C/EBPα, CCAAT enhancer binding protein α; DCoH, Dimerization Co-Factor HNF1α; HNF1α, hepatic nuclear factor 4α; HNF4α, hepatic nuclear factor 4α, PXR, pregnane X receptor; PCR, polymerase chair reaction; SBE, single-base extension; SNP, single nucleotide polymorphism

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

A considerable number of reports identifying and characterizing genetic variants within the *CYP2C9* coding region have appeared. Much less is known about polymorphic promoter sequences that also might contribute to interindividual differences in *CYP2C9* expression. To address this problem, approximately 10,000 bp of *CYP2C9* upstream information was re-sequenced using 24 DNA samples from the Coriell Polymorphism Discovery Resource. Thirty-one single nucleotide polymorphisms (SNPs) were identified; nine were novel while 22 were previously reported. Using both sequencing and multiplex single-base extension, individual SNP frequencies were determined in 193 DNA samples obtained from unrelated, self-reported Hispanic Americans of Mexican descent and compared to similar data obtained from a non-Latino White cohort. Significant inter-ethnic differences were observed in several SNP frequencies, some of which appeared unique to the Hispanic population. Analysis using PHASE 2.1 inferred nine common (>1%) variant haplotypes, two of which included the g.3608C>T (R144C) *CYP2C9*2* and two the g.42614A>C (I359L) *CYP2C9*3* SNPs. Haplotype variants were introduced into a *CYP2C9*/luciferase reporter plasmid using site-directed mutagenesis and the impact of the variants on promoter activity assessed by transient expression in HepG2 cells. Both constitutive and PXR-mediated inducible activities were measured. Haplotypes 1B, 3A, and 3B each exhibited a 65% decrease in constitutive promoter activity relative to the reference haplotype. Haplotypes 1D and 3B exhibited a 50% decrease and a 40% increase in induced promoter activity, respectively. These data suggest that genetic variation within *CYP2C9* regulatory sequences likely contributes to differences in *CYP2C9* phenotype both within and among different populations.

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Many factors contribute to an individual's response to medications including age, dietary intake, concomitant medication, and various innate pharmacokinetic and pharmacodynamic parameters (Gage *et al.*, 2004; Kamali *et al.*, 2004). Among these, genetic factors that influence drug metabolism play a major role and contribute substantially to observed interindividual variability in response. While numerous enzymes are involved drug metabolism, the cytochrome P450-dependent monooxygenase superfamily is widely recognized as having a prominent role. Consistent with these two facts, a recognized hallmark of cytochrome P450-dependent metabolism is large intersubject variation in the human population. Although environmental factors resulting in induction or suppression contribute to intersubject variation, our current understanding suggests that genetic variability is equally if not more important (Phillips *et al.*, 2001).

CYP2C9 is a major human cytochrome P450 enzyme that accounts for approximately 20% of the total cytochrome P450 protein content in adult human liver (Shimada *et al.*, 1994) and is responsible for the metabolism of approximately 16% of clinically used drugs cleared by oxidative pathways (Williams *et al.*, 2004). Examples of small molecular weight therapeutics for which CYP2C9 is important for disposition include the anticoagulant, warfarin, the antidiabetic agents, tolbutamide and glipizide, the anticonvulsant, phenytoin, the antihypertensive, losartan, the antidepressant, fluoxetine and a number of nonsteroidal anti-inflammatory drugs such as ibuprofen, diclofenac and celecoxib (Miners and Birkett, 1998; Klose *et al.*, 1998; Davies *et al.*, 2000). Over the past several years, multiple *CYP2C9* single nucleotide polymorphisms (SNPs) resulting in amino acid changes have been identified. Currently more than thirty allelic variants are listed on the cytochrome P450 allele website

(http://www.cypalleles.ki.se), however, not all have been fully characterized. The most common *CYP2C9* alleles include *CYP2C9*1A* (reference allele), *CYP2C9*2* (three haplotype variants, all containing the g.3608C>T, R144C SNP), and *CYP2C9*3* (two haplotype variants, all containing the g.42614A>C, I359L SNP). Less common are *CYP2C9*4* (g.42615T>C, I359T) and *CYP2C9*5* (g.42619C>G, D360E), and *CYP2C9*6* (g.10601delA). The *CYP2C9*2* and *CYP2C9*3* variant alleles encode enzymes exhibiting significantly lower intrinsic clearance both *in vivo* and *in vitro* (reviewed in Lee *et al.*, 2002).

As evidenced above, extensive research has been conducted to identify and characterize SNPs within the *CYP2C9* coding region. However, much less is known about other variant sites, such as polymorphic promoter sequences, that also might contribute to observed interindividual differences in *CYP2C9* expression. In a population study of Japanese epileptic patients, seven polymorphic sites were identified within the first 2000 bp upstream of the *CYP2C9* transcription start site (Shintani *et al.*, 2001). Several of the identified SNPs were in linkage disequilibrium, resulting in six unique haplotypes. One of the haplotypes (g.-1912T>C, g.-1885C>, g.-1538G>A, g.-1189C>T, and g.-982G>A) exhibited a 60% reduction in *in vitro* promoter activity. Further, an association between estimated phenytoin intrinsic clearance and promoter activity was observed with respect to the identified *CYP2C9* upstream haplotypes (Shintani *et al.*, 2001). However, because the promoter SNPs defining this pattern also were in linkage disequilibrium with the SNP defining *CYP2C9*3*, it is unclear what impact these variants might have on *CYP2C9* phenotype *in vivo*. Similar findings were reported in a population study of Japanese and non-Latino White subjects using warfarin (Takahashi *et al.*, 2004). However, when differences in S-warfarin CYP2C9 clearance were compared between

Japanese and non-Latino White patients having reference alleles in both promoter (up to position -2100) and coding regions, the Japanese patients exhibited significantly greater intrinsic clearance than the non-Latino White patients. These data not only suggest the likelihood of significant interpopulation genetic differences, but also the possible presence of additional important *CYP2C9* promoter elements and variants upstream of position -2100. Such a conclusion is not only supported by the study of Takahashi *et al.* (2004), but also has precedence based on our knowledge of other cytochrome P450 genes (Martinez-Jimenez *et al.*, 2005).

Given the data supporting the likely presence of important *CYP2C9* regulatory polymorphisms, the evidence for interpopulation differences in both the presence and frequency of genetic variants, and the paucity of pharmacogenetic studies in the Hispanic population, the present study was designed to begin addressing this knowledge gap.

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Materials and Methods

Materials: Custom oliogonucleotides were synthesized by MWG Biotech (High Point, NC). Herculase High Fidelity polymerase and Quick Change site directed mutagenesis kits were obtained from Stratagene (La Jolla, CA). The ExoSap-IT mix and shrimp alkaline phosphatase were purchased from United States Biochemical Corp. (Cleveland, OH). CEQ SNP-Primer Extension and Dye Terminator Cycle Sequencing kits were obtained from Beckman Coulter, Inc. (Fullerton, CA). Restriction endonucleases were purchased from New England Biolabs (Beverly, MA). For SNP discovery, the 24 sample subset of the Polymorphism Discovery Resource (24PDR) was obtained from the Coriell Institute (Camden, NJ). The luciferase reporter plasmid, pGL3Basic, and luciferase reporter assay kit were purchased from Promega (Madison, WI) whereas the luminescent β-galactosidase assay kit was obtained from BD Biosciences (Palo Alto, CA). High purity plasmid purification kits and the QIAmp DNA blood Midi kit were supplied by QIAGEN (Valencia, CA). Cell culture medium, fetal bovine serum, DMSO and rifampicin were purchased from Sigma-Aldrich (St. Louis, MO). Lipofectamine 2000, Opti-MEM reduced serum medium, Taq polymerase and the TA cloning kit were purchased from Invitrogen (Carlsbad, CA). The HepG2 human hepatoblastoma cell line was a gift from Dr. Barbara Knowles (Jackson Laboratories, Bar Harbor, ME). The HNF-4α expression plasmid, pCMVHNF4α (Stoffel and Duncan, 1997), was provided by Dr. Stephen A. Duncan (Medical College of Wisconsin, Milwaukee, WI). The HNF-1α expression plasmid, pBJ5HNF1α, (Kuo *et al.*, 1990), and dimerization cofactor of HNF-1α (DCoH), pBJ5DCoH (Mendel *et al.*, 1991), were generous gifts from Dr. Gerald R. Crabtree (Stanford University School of Medicine, Stanford, CA). The expression vectors for human PXR, pSG5hPXR

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

(Lehmann *et al.*, 1998), and human CAR, pCDM8hCAR (Baes *et al.*, 1994), were kindly provided by Drs. Stephen A. Kliewer (Glaxo Wellcome Research and Development, Research Triangle Park, NC) and David D. Moore (Baylor College of Medicine, Houston, TX), respectively.

Subjects: After consent, blood samples were collected from 193 women of self-reported Hispanic background who traced their Mexican ancestry back a minimum of two generations. Volunteers were recruited after admission to the labor and delivery units of Provena Saint Therese Medical Center or Victory Memorial Hospital, Waukegan, IL. DNA was extracted using the QIAmp DNA blood MIDI kit and stored at 4˚C. This research protocol was approved by all involved Institutional Review Boards.

DNA Amplification for SNP Discovery: All PCR DNA amplification primers were designed using OLIGO V.6.45 (Molecular Biology Insights, Cascade, CO). Each primer pair was designed to have similar melting temperatures in order to facilitate high throughput processing using a 96-well format. Primer sequences are provided in supplemental data (see supplement Table S1).

Common *CYP2C9* genetic variants were identified by sequencing the approximate first 10 kbp of *CYP2C9* 5' flanking sequence in each sample from the Coriell 24PDR. Templates were prepared by PCR DNA amplification using 20 to 25 ng genomic DNA in a 20 μL reaction volume containing 0.2 mM each deoxyribonucleotide triphosphate, 0.5 μM each primer, and 1.25 units of Herculase High Fidelity polymerase (Stratagene). Cycles were as follows: denaturation at 92°C for 40 sec and annealing and extension at 56°C to 62°C for 1 min for a total of 30 cycles. All amplification reactions included an initial 10 sec hold at 94°C and a final

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

5 min hold at 75°C. After amplification, 4 μL of ExoSap-IT (United States Biochemical Corp.) was added and the reaction was incubated at 37°C for 30 min to remove unincorporated deoxy- and dideoxyribonucleotide triphosphates and primers.

***CYP2C9* Sequence Analysis:** Sequence analysis was performed using 60 to 100 fmols of amplicon with 8 μL Quickstart Sequencing reagent (Beckman Coulter) in a MJ Research PTC-225 Peltier thermal cycler per manufacturer's recommendations. Analysis was performed on both DNA strands. SNPs identified on a single allele were verified by a repeat analysis of an independently generated amplicon. In all instances, the coordinates of the identified SNPs follow the recommendations of the HUGO nomenclature working group, wherein the "A" of the ATG start codon is assigned +1 and uses contig NT_030059.12, build 36.1 as a reference. For purposes of clarity, the coordinates of all *CYP2C9* DNA fragments also follow the same convention. Immediately following the sequencing reaction, unincorporated dideoxyribonucleotide triphosphates and primer were removed using CleanSeq magnetic beads and reagents (Agencourt) prior to analysis of samples by capillary electrophoresis in a CEQ8000 Genetic Analysis System (Beckman Coulter).

Genotyping: After reviewing all of the identified *CYP2C9* SNPs, there were five unique clusters and eight remaining SNPs that fell outside of the clusters. Five PCR primer pairs were designed to amplify the DNA sequences containing the SNP clusters, resulting in amplicons ranging in size from 269 bp to 853 bp (see supplemental Table S1). Two *CYP2C9* fragments containing the remaining eight sequence variants were amplified, resulting in amplicons of 1.6 kbp and 2.3 kbp. DNA amplifications were performed as described above using 20 to 25 ng genomic DNA as template. DNA amplification of sequences containing exons 3, 5, and 7 were

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

generated for use as templates for genotyping of the previously identified SNP tags for the *CYP2C9*2, CYP2C9*3, CYP2C9*4, CYP2C9*5,* and *CYP2C9*6* alleles (reviewed in Lee *et al.*, 2002) (see supplemental Table S2).

The frequencies of previously identified *CYP2C9* promoter variants (Veenstra *et al.*, 2005), as well as any new variants identified within the discovery phase of this project, were determined within the Hispanic population using a combination of DNA sequencing (all SNPs that fell within the five unique cluster sets) and multiplexed single base extension (SBE) reactions (Lindblad-Toh *et al.*, 2000) (eight remaining SNPs, as well as the SNP tags for the *CYP2C9*2, CYP2C9*3, CYP2C9*4, CYP2C9*5*, and *CYP2C9*6* alleles) (reviewed in Lee *et al.*, 2002). DNA sequencing was performed as described above. For genotyping by SBE, two muliplex primer sets were designed, consisting of four primers in each set, using OLIGO V. 6.45. The *CYP2C9* -1537G>A variant was not identified in the initial SNP discovery. However, because of the previously reported linkage with the *CYP2C9*3* haplotype (Veenstra *et al.*, 2005), the frequency of this SNP subsequently was determined in the Hispanic study population using SBE. SBE reactions were performed in a 20 μL reaction volume consisting of a mixture of all four dye-labeled dideoxynucleotide terminators and the supplied proprietary polymerase, (Beckman Coulter), 20 fmols template, and approximately 1 to 10 pmols of each primer. Individual primer concentrations were optimized to ensure adequate signal intensities for each primer within the multiplex. SBE reactions were as follows: denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec and extension at 72°C for 30 sec for a total of 25 cycles. Unincorporated dideoxynucleotides were eliminated by adding one unit of shrimp alkaline phosphatase (United States Biochemical Corp.) and incubating for 30 min at 37ºC. A 0.5 μL

aliquot of the SBE reaction and 0.5 μL of SBE size standard 80 (Beckman Coulter) were mixed and analyzed by capillary electrophoresis on a CEQ8000 Genetic Analysis System (Beckman Coulter). As a quality control measure, 10% of the DNA samples were random selected and independently analyzed to confirm the original genotyping calls.

Plasmids and Cloning of *CYP2C9* Promoter Region: A *CYP2C9* genomic fragment containing positions -1,515 to +184 was amplified from a single Hispanic DNA sample that, based on sequence analysis, was homozygous for the *CYP2C9*1* allele. The amplification product was digested with the restriction enzymes *HindIII* (*CYP2C9* position -1,454) and *SacI* (*CYP2C9* position 66) and subsequently cloned into these same sites within the multiple cloning site of the pBluescript II KS+ vector (Stratagene). Site directed mutagenesis was performed to convert the adenosine residue at *CYP2C9* position -1 to cytosine, thereby creating an *Nco*I site at the translation start codon. Following mutagenesis, the complete *CYP2C9* insert was sequenced using M13 forward, M13 reverse and internal primers to verify the identity of the clone and validate its fidelity relative to the reference sequence. The mutagenized amplicon was excised from the pBluescript II KS+ vector using the restriction enzymes *NcoI* and *HindIII*, and cloned into these same sites within the pGL3Basic vector, generating pRNH905.

Due to persistent amplification difficulties, an alternate approach was used to isolate a *CYP2C9* genomic fragment containing positions -5,910 to -1,325. A human BAC clone containing the region of interest (CTD2343I24, chromosome 10, position 96586284 to position 96697248) was purchased from Open Biosystem's Clone Resource. The BAC clone DNA was isolated using a midi prep kit designed for large constructs (Qiagen), and then digested using *Eci*I. The 16,589 bp fragment, representing *CYP2C9* position -7,870 to +8,719, was gel purified

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

and digested with *Hind*III and *Xmn*I and subsequently cloned into the *Hind*III and *Sma*I sites of pRNH905, generating pRNH923.

A similar PCR approach as that described above for the proximal promoter fragment was used to amplify a *CYP2C9* genomic fragment from position –10108 to -5517. The resulting amplicon was treated with Taq polymerase (Invitrogen) in order to add a single deoxyadenosine (A) to the 3' end of the PCR product and subsequently cloned into pCR2.1 (Invitrogen). The relative orientation of the insert was checked by digestion with *Bst1107I* and *SacI*. A clone yielding product sizes of 4509 bp and 3990 bp was selected and the 4509 bp fragment was subsequently cloned into these same sites in pRNH923, generating pRNH924. Thus, this final construct contained *CYP2C9* sequences from position -10,008 to -1 (chromosomal coordinates 96678320 to 96688429) directing the expression of the luciferase gene.

Three *CYP2C9* fragments were excised from pRNH924 and cloned into the pBluescript II KS+ vector to perform site directed mutagenesis and introduce the various haplotype sequences into the *CYP2C9*/luciferease reporter vector. pRNH971 was constructed by digesting pRNH924 with *Hin*dIII and *Sal*I and the resulting 3,376 bp fragment containing the luciferease cassette and CYP2C9 position -1 to -1453 was cloned into the *Hin*dIII and *Sal*I sites of the pBluescript II KS+ vector. pRNH972 was constructed by digesting pRNH924 with *Hin*dIII and *Xba*I and the resulting 4,379 bp fragment cloned into these same sites of pBluescript II KS+. The third mutagenesis plasmid, containing *CYP2C9* sequences from position -5,833 to -10,008, was constructed by digesting pRNH924 with *Xba*I and *Sac*I and cloning the resulting 4,321 bp fragment into these same sites in pBluescript II KS+, generating pRNH973. After DNA sequence analysis of each plasmid, multiple rounds of site-directed mutagenesis were performed

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

to introduce all combinations of SNPs present in each of the nine inferred *CYP2C9* variant haplotypes, as well as the reference. The individual fragments were then re-assembled within the pGL3Basic backbone, thus generating pRNH954 (haplotype 1), pRNH955 (haplotype 1A), pRNH974 (haplotype 1B), pRNH956 (haplotype 1C), pRNH957 (haplotype 1D), pRNH958 (haplotype 1E), pRNH959 (haplotype 2A), pRNH960 (haplotype 2B), pRNH962 (haplotype 3A), and pRNH961 (haplotype 3B).

Cell Culture and Transfection of HepG2 Cells: HepG2 human hepatoma cells were cultured in Eagle's minimal essential media supplemented with 10% fetal bovine serum, penicillin (50 U/mL) and streptomycin (50 μg/mL), and maintained in a humidified incubator at 37°C in an atmosphere of 5% CO_2. Cultures were never allowed to grow beyond approximately 80% confluence. All experiments were performed with cells between passage 4 and 15. For transfection studies, 1.5 x 10^5 cells were seeded into 24-well culture dishes and 24 hours later, were transfected with 1.4 μg of Lipofectamine 2000, 0.5 μg of test luciferase reporter plasmid, and 50 ng each of pCMVβgal, pCMVHNF-4α, pBJ5HNF-1α and pBJ5DCoH in Opti-MEM reduced serum medium. After incubation for 24 hrs at 37ºC, transfection medium was replaced with normal growth medium, and cells were incubated for an additional 24 hrs. For induction experiments, transfection medium was replaced with Opti-MEM reduced serum medium supplemented with 10 μM rifampicin or vehicle (0.1% DMSO). The rifampicin concentration and the time of harvest post-treatment were based on results from initial optimization experiments in which rifampicin concentrations of 5 μM, 10 μM, and 25 μM were added for 10, 16, and 24 hrs (data not shown). Cell lysates were prepared and luciferase assays performed according to the manufacture's instructions. Data were normalized with respect to

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

β-galactosidase activity to correct for transfection efficiency and expressed as relative luciferase activity. Data are reported as the mean ± standard deviation (SD) of at least three determinations performed with two independently prepared luciferase reporter plasmids.

Sequence and Data Analysis: DNA sequence data was analyzed using DNAStar software (LaserGene, Madison, WI); minimal acceptable quality scores values were set at ≥ 12. Sequences were scanned for potential transcription factor recognition sequences using the Match Program and the TRANSFAC Professional V11.3 database (BIOBASE Corporation, Wolfenbuettel, Germany). Both the Liver-Specific and Vertebrate Non-Redundant Matrix Profiles were used with search criteria that minimized the identification of false positives. All SNPs identified were tested for deviations from Hardy-Weinberg equilibrium with the use of a chi-square test. The frequencies of individual SNPs were compared using Fisher's exact test (GraphPad InStat V 3.05, San Diego, CA). Based upon the observed frequencies of each SNP, haplotype analysis was inferred using PHASE V 2.1, with all parameters set at default values except that 5000 iterations were performed with a thinning value of one and a burn in of 1000 (Stephens and Donnelly, 2003). Functional differences among the different haplotypes were assessed using transient expression assays and compared by one-way ANOVA with a Holm-Sidak post hoc test for multiple comparisons (SigmaStat V3.11, Systat Software, Chicago, IL). Inferred haplotype frequencies between non-Latino Whites and Hispanics of Mexican descent were compared using a Student's t-test (SigmaStat V3.11, Systat Software, Chicago, IL). In all instances, an α value of 0.05 was accepted as significant.

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Results

***CYP2C9* SNP Discovery:** SNP discovery was accomplished by sequencing overlapping amplicons spanning approximate 10,000 bp upstream of the *CYP2C9* transcription start site using 24 DNA samples obtained from the Coriell Polymorphism Discovery Resource (Camden, NJ). Thirty-one SNPs were identified, nine novel and 22 previously reported in the literature and/or on the human *CYP2C9* allele nomenclature website (http://www.cypalleles.ki.se/) (Table 1). To determine if any of the novel or previously identified SNPs were located at or near putative transcription factor binding sites, a search within the 10,000 bp of *CYP2C9* upstream sequence was performed using the Match Program along with the TRANSFAC Professional V11.3 database (BioBase Biological Databases) using both the Liver-Specific and Vertebrate Non-Redundant Matrix Profiles. Comparisons also were made against previous reports on functional *CYP2C9* regulatory elements (Ibeanu and Goldstein, 1995; Chen *et al.*, 2005; Kawashima *et al.*, 2006) (Table 1). Only sites whose core or matrix match score was impacted, or which were eliminated or created by variant sequences were considered. None of the identified SNPs were located within any previously identified regulatory elements and only three SNPs were located in putative transcription factor binding sites. The g.-8416T>G SNP falls within a key residue of a putative glucocorticoid receptor IR3 element and would be predicted to eliminate or reduce binding at this site while the g.-3360T>C SNP falls immediately 5' to a putative core NF1/CTF site, reducing the matrix match score. The g.-3597A>G SNP creates a putative new CREB binding site.

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

***CYP2C9* SNP Validation in a Hispanic Population of Mexican Descent:** The PDR24 is an anonymous and blinded DNA panel from individuals representative of the United States population, but provides no insight into the presence or frequency of specific genetic variants within defined ethnic or racial groups. Thus, further characterization of the identified SNPs was needed in the Hispanic population. To define *CYP2C9* promoter region haplotypes, allelic frequencies of the 31 SNPs identified during the discovery phase, the g.-1537G>A SNP previously shown to be linked to the *CYP2C9*3* allele in non-Latino Whites (Veenstra *et al.*, 2005), as well as the SNPs defining the *CYP2C9*2*, CYP2C9**3*, *CYP2C9*4*, *CYP2C9*5* and *CYP2C9*6* alleles were determined in 193 Hispanic Americans of Mexican descent by either sequence analysis (for SNP clusters) or multiplex SBE. All SNPs identified were in Hardy-Weinberg equilibrium except g.-1188T>C (χ^2 = 5.37) and g.-4302C>T (χ^2 = 6.19), both genotyped using multiplex SBE. To eliminate possible assay error for these two positions, all DNA samples with variant alleles at g.-1188T>C or *g*-4302C>T were resequenced and in all instances, the original genotype call was confirmed. Five of the nine novel SNPs identified in the 24PDR were observed in the Hispanic population (g.-8422A>G, g.-5146G>C, g.-5143A>C, g.5140A>T and g.-4302C>T) (Table 2). The g.-4302C>T SNP was observed at a frequency of 10.9% whereas all other novel SNPs were observed at frequencies <0.5%. The most common promoter variant in the Hispanic population was g.-1188T>C at an observed frequency of 20.7%, whereas the g.-8422A>G, g.-8416T>G and g.-7336G>A SNPs were the least commonly observed variants at frequencies <1%. Two previously reported *CYP2C9* structural variants, g.3608C>T (*CYP2C9*2*) and g.42614A>C (*CYP2C9*3*) (Rettie *et al.*, 1994; Haining *et al.*, 1996) also were observed in the Hispanic study population at 7.0% and 4.4%, respectively

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

(Table 3). Other previously reported *CYP2C9* exon variants, g.42615T>C (*CYP2C9*4*), g.42619C>G (*CYP2C9*5*) and g.10601delA (*CYP2C9*6*) (Imai *et al.*, 2000; Dickmann *et al.*, 2001; Kidd *et al.*, 2001) were not observed.

The *CYP2C9* promoter SNP frequencies observed in the Hispanic population were compared to the previously reported frequencies in a non-Latino White cohort (Veenstra *et al.*, 2005) (Table 2). Five upstream genetic variants (g.-8422A>G, g.-5146G>C, *g*-5143A>C, g.-5140A>T and g.-4302C>T) observed in the Hispanic population were absent in the non-Latino White population, one of which, g.-4302C>T, was not found due to lack of ascertainment of that region in the later population (see GenBank accession AY702706, - "region not scanned" due to repetitive sequences). A total of 21 SNPs were observed in both ethnic groups, five of which occurred at different frequencies. For instance, the g.-620G>T, g.-1096A>G, g.-1188T>C, g.-2663delTG and g.-3089G>A SNPs were observed in the Hispanic population at frequencies of 4.9%, 7.0%, 20.7%, 7.3% and 10.9%, respectively, compared to 11.0%, 12.0%, 35.0%, 17.0% and 17.0%, respectively, in the non-Latino White population ($p<0.05$).

Using the determined minor allelic frequencies of upstream and structural variants, a total of 34 *CYP2C9* haplotypes were inferred using PHASE V 2.1, ten of which occurred at a frequency greater than 1%. This latter group was compared to the previously reported haplotypes in a non-Latino White study group (Table 4) (Veenstra *et al.*, 2005). Importantly, for the purposes of comparison, the SNPs previously reported in the non-Latino White cohort were re-analyzed, but only including the promoter and **2* and **3* SNPs. Haplotype 1 was designated as the reference sequence and was inferred to occur at a frequency not significantly different from that reported in the non-Latino White population (Table 4). Four *CYP2C9* haplotypes were

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

deduced in the Hispanic population, but not in the non-Latino White population. Haplotype 1A (g.-8553C>A) was present at 3.2%, 1B (g.-4302C>T) at 10.0%, 1E (g.-1188T>C) at 1.3% and 2A (g.-3597A>G, g-3360T>C, g.-1188T>C, g.-1096A>G, g.-485T>A, g.-484C>A, g.3608C>T) at 2.1%. Five variant *CYP2C9* haplotypes inferred in both the Hispanic and non-Latino White populations were present at different frequencies. For example, haplotype 1C (g.-3089G>A, g.-1188T>C) was deduced in the Hispanic population at a higher frequency than that observed in the non-Latino White population, *i.e.*, 3.1% compared to 0.27%, respectively. In contrast, haplotype 1D was inferred to occur more frequently in the non-Latino White population, 17.2%, compared to the Hispanic population, 6.6%.

In addition to the five haplotypes (1A, 1B, 1C, 1D, and 1E) comprised solely of SNPs within the *CYP2C9* regulatory region, four inferred haplotypes (2A, 2B, 3A, and 3B) consisted of upstream SNPs in linkage disequilibrium with the previously described g.3608C>T (*CYP2C9*2*) and g.42614A>C (*CYP2C9*3*) structural allelic variants (Table 4). Of these four haplotype variants, haplotype 2A (g.-3579A>G, g.-3360T>C, g.-1188T>C, g.-1096A>G, g.-485T>A, g.-484C>A, and g.3608C>T) appeared to be unique to the Hispanic population. Haplotypes 2B, 3A, and 3B were present at different frequencies in the two population groups (Table 4).

Effect of Variant Promoter Haplotypes on Constitutive and Rifampicin-Induced *CYP2C9* Promoter Activity: To explore the possible functional significance of the identified *CYP2C9* variant haplotypes on constitutive promoter activity, we used site-directed mutagenesis to introduce nine common *CYP2C9* haplotype variants into a reporter construct containing 10,008 bp of *CYP2C9* upstream information driving the expression of the luciferase reporter

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

gene. Negligible luciferase activity was observed when the *CYP2C9*/luciferase construct was transiently expressed alone (data not shown). Given the important roles HNF1α, and HNF4α (Chen *et al.*, 2005; Kawashima *et al.*, 2006) have in regulating *CYP2C9* promoter activity, the initial experiments were repeated co-expressing both of these factors along with the *CYP2C9*/luciferase reporter construct. A substantial increase in promoter activity was observed in the presence of both HNF1α and HNF4α, but either factor alone had a minimal effect. As such, all subsequent transient expression experiments included expression vectors for both of these factors in the protocol. Of the nine variant haplotypes, 1B, 3A, and 3B exhibited 2.5-, 3.2-, and 2.6-fold decreased *CYP2C9* promoter activity compared to the reference construct (Figure 1). Comparing the SNPs constituting each of these hypomorphic promoter variants against the variants located within putative transcription factor binding sites (Table 1), only the g.-3360T>C SNP is located immediately 5' to a putative NF1/CTF core element (position -1). The T>C transition reduces the matrix match score from 0.747 to 0.722 and of the eight binding sites selected for defining this matrix, five had a T at position -1 and none had a C.

In addition to the contribution of genetic polymorphisms to interindividual variability in *CYP2C9* activity (Schwarz, 2003; Lee *et al.*, 2002), induction by exogenous agents also can contribute to observed differences (Williamson *et al.*, 1998; Niemi *et al.*, 2001). To determine whether any of the identified *CYP2C9* variant haplotypes alter *CYP2C9* induction by rifampicin, HepG2 cells were co-transfected with pSG5hPXR, a human PXR expression plasmid, and the various *CYP2C9* reporter constructs and treated with 0.1% DMSO (vehicle control) or 10 μM rifampicin for 24 hrs. No effect was observed when HepG2 cells were co-transfected with pSG5hPXR and pRNH954 (reference haplotype) and treated with 0.1% DMSO (Figure 2A).

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

HepG2 cells co-transfected with pSG5hPXR and pRNH954 (reference haplotype) and then treated with 10 μM rifampicin for 24 hrs, exhibited a 3.2-fold induction of luciferase activity (Figure 2A). However, transfection with pRNH957 (haplotype 1D) resulted in only a 1.6-fold induction (49% decrease relative to the reference haplotype), whereas pRNH961 (haplotype 3B) resulted in a 4.6-fold induction (44% increase relative to the reference haplotype) (Figure 2B). No differences in induction were observed with any of the other plasmids representing the other haplotypes.

To determine the effect of promoter haplotypes on phenobarbital-dependent *CYP2C9* induction, HepG2 cells were co-transfected with pCMVhCAR, a human CAR expression plasmid, and the variant *CYP2C9* reporter constructs. No effect was observed with the CAR expression plasmid alone and no induction was observed with the reference or any of the variant haplotypes following treatment with 0.5 mM phenobarbital (data not shown). In contrast, a 2-fold induction was observed in HepG2 cells co-transfected with pGS5hPXR following treatment with 0.5 mM Phenobarbital. Similar differences in induction as those observed with human PXR and rifampicin were seen with the pRNH957 (haplotype 1D) and pRNH961 (haplotype 3B) constructs (data not shown).

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Discussion

Relative to other large population groups, the contribution of genetic polymorphisms to interindividual differences in *CYP2C9* expression in Hispanics is poorly understood. In the current study, five of the nine novel SNPs identified in the 24PDR (g.-8422A>G, g.-5146G>C, g.-5143A>C, g.-5140A>T, and g.-4302C>T) were present in the Hispanics but were absent in the non-Latino Whites. A total of 22 SNPs were present in both populations, five of which (g.-620G>T, g.-1096A>G, g.-1188T>C, g.-2663delTG and g.-3089G>A) occurred at significantly different frequencies (Table 2). Thus, similar to what has been observed for *CYP2C9* structural variants, ethnic differences exist in the presence and frequency of *CYP2C9* regulatory polymorphisms, which may contribute to interpopulation differences in CYP2C9-dependent metabolism.

The observed frequency of two previously reported *CYP2C9* structural variants, g.3608C>T (*CYP2C9*2*) and g.42614A>C (*CYP2C9*3*) were not significantly different between the Hispanics and non-Latino Whites, nor were the determined frequencies different than those reported by Llerena *et al.* (2004) for Hispanic Americans of Mexican descent. In contrast, there was a significant difference when compared to the frequencies reported by Xie *et al.* (2002). However, the ancestral background of the Hispanic population was not defined in the latter study. Thus, the discrepancy in these data is likely explained by a population of mixed ancestry.

Based upon the observed minor allele frequencies of 32 upstream and five structural variants, 10 common (>1%) *CYP2C9* haplotypes were inferred. The observed haplotype 1 (reference) frequency in Hispanics was 60.53%, compared to 63.64% in non-Latino Whites,

consistent with a relatively simple haplotype structure. Four *CYP2C9* haplotypes (1A, 1B, 1E, and 2A) were inferred in Hispanics, but not in non-Latino Whites (Table 2).

The possible contribution of upstream *CYP2C9* genetic variability to interindividual pharmacokinetic differences is more controversial. Variant effects on promoter activity *in vitro* have been observed, but failed to correlate with differences in *in vivo* activity (Shintani *et al.*, 2001). Other studies have failed to show an independent effect of *CYP2C9* upstream haplotype variants on mean warfarin clearance or dose (King *et al.*, 2004; Takahashi *et al.*, 2004; Veenstra *et al.*, 2005). However, the design of these studies precluded the ability to test the impact of the *CYP2C9* upstream polymorphisms independently of the variants impacting CYP2C9 catalytic activity. In the current report, haplotype 1B, consisting of only the g.-4302C>T variant, exhibited a significant decrease in promoter activity *in vitro*. Combined with its relatively high frequency (*i.e.*, 10.0%), these data would be consistent with reduced constitutive *CYP2C9* expression in this population. It is interesting that the frequency of the g.-4302C>T SNP was not in Hardy-Weinberg equilibrium, suggesting a possible selective pressure for the presence of the variant allele, although this deviation may also be due to recent population admixture. Significant decreases in *in vitro* constitutive promoter activity also were observed with haplotypes 3A and 3B. Thus, the results of our study suggest that, in addition to the defective enzymatic function of CYP2C9.3, a decrease in basal *CYP2C9* transcription also may contribute to the overall observed *CYP2C9*3* phenotype.

In vitro studies using primary hepatocytes have shown that CYP2C9 mRNA, protein, and catalytic activity are all increased by drugs such as rifampicin, hyperforin, and phenobarbital through a PXR-dependent mechanism (Chen *et al.*, 2004). Rifampicin treatment

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

also has been reported to enhance the clearance of CYP2C9 substrates, indicative of CYP2C9 induction *in vivo* (Williamson *et al.*, 1998; Niemi *et al.*, 2001). Co-transfection studies in HepG2 cells with a human PXR expression plasmid and *CYP2C*9 reporter constructs were used to investigate the possible functional effects of *CYP2C9* haplotype variants on rifampicin-dependent enhancement of *CYP2C9* promoter activity. With haplotype 1 (reference construct), addition of human PXR alone had no effect on basal *CYP2C9* promoter activity, consistent with results reported by Ferguson *et al.* (2002). These results, however, conflict with those of Chen *et al.* (2004), who reported a 3.8-fold increase in *CYP2C9* promoter activity upon addition of human PXR alone. When the co-transfected cells were treated with 10 μM rifampicin, a 3.2-fold increase in promoter activity was observed, similar to the approximate 3-fold increase reported by Chen *et al.* (2004). Yet Ferguson *et al.* (2002) failed to see an effect with rifampicin. The discrepancies in these data may result from differences in the amount of *CYP2C9* upstream sequences present in the reporter construct or differences in transfection conditions. When co-transfection studies were performed with the *CYP2C9* promoter variants, two of the inferred *CYP2C9* haplotypes exhibited an altered induction profile; haplotype 1D resulted in a 1.6-fold reduction in inducibility, whereas haplotype 3B resulted in a 1.4-fold increase in inducibility. Given the frequency of these two haplotypes (6.6% and 2.4%, respectively) and the magnitude of this observed effect, it is possible these variants contribute to observed interindividual differences in CYP2C9 phenotype.

Previous studies have identified two constitutive androstane receptor (CAR)-responsive elements within the *CYP2C9* promoter at positions -2,898 and -1,839 that bind human CAR *in vitro* and transactivate reporter constructs (Ferguson *et al.*, 2002). Further, Gerbal-Chaloin *et al.*

(2001) described a 4-fold increase in *CYP2C9* mRNA following treatment with phenobarbital, a known human CAR ligand. CAR-mediated phenobarbital induction of *CYP2C9* promoter activity was not observed in the current study, although a two-fold, PXR-mediated phenobarbital enhancement of *CYP2C9* promoter activity was observed. This latter observation supports the recent findings of Chen *et al.* (2004) who determined that human PXR is responsible for the induction of human *CYP2C9* by both rifampicin and phenobarbital.

To gain some insight into which SNP, or combination of SNPs, might be responsible for the observed altered activities, a careful comparison between all observed haplotypes was made. Reduced constitutive activity was observed with haplotypes 1B, 3A and 3B. As indicated earlier, the g.-4302C>T SNP is unique to haplotype 1B and as such, is assumed to be causative, yet does not alter any known or putative transcription factor binding sites. Several SNPs are found in common between haplotypes 3A and 3B, but not in other haplotypes (g.-8897C>A, g.-7419A>G, g.-5813A>G, g.-5661C>A, g.-4877G>A, -1911T>C, -1885C>G, g.-1537G>A, and g.-981G>A). However, similar to the haplotype 1B g.-4302C>T SNP, none of these variants alter known or putative regulatory elements. Altered PXR-mediated inducibility was observed with both haplotypes 1D and 3B. In the case of haplotype 1D (g.-3089G>A, g.-2663delTG, g.-1188T>C), two other inferred haplotypes (1C and 1E) also contain g.-3089G>A and/or g.-1188T>C but exhibit promoter activities no different from the reference construct. The remaining SNP in haplotype 1D, g.-2663delTG, is not present in any other inferred haplotype and as such, may be causative for the reduced inducibility of this promoter variant. Haplotype 3B resulted in an increase in rifampicin-induced *CYP2C9* promoter activity, whereas haplotype 3A exhibited no difference compared to the reference control. Yet, the only

difference between these two haplotypes is the presence of the g.-1188T>C SNP in haplotype 3B. This observation suggests that the g.-1188T>C transition might be responsible for the observed difference in induction profiles. However, two other haplotypes (1C and 1E) also contain g.-1188T>C, yet exhibit no difference relative to the control construct. Further, a recent study by Sandberg *et al.* (2004) found that the g.-1188T>C variant did not affect gene expression *in vitro*. Finally, haplotype 1D also contains the g.-1188T>C SNP, and, in contrast to haplotype 3B, was associated with decreased induction of *CYP2C9* promoter activity. These findings suggest that g.-1188T>C cannot solely be responsible for the observed increase in activity, but rather contributes to a combinatorial effect. The simplest explanation as to how individual or clusters of SNPs might function to alter gene regulation is by modifying or eliminating transcription factor binding. However, this does not appear to be the case for the variants identified herein and is clearly not the case for many regulatory polymorphisms. A survey of 247 known promoters and 647 haplotype variants by Buckland *et al.* (2005) revealed that only 35% of the functional regulatory variants identified were localized within predicted transcription factor binding sites. Thus, many regulatory polymorphisms may impact gene expression through other yet to be identified sequence-specific mechanisms.

In summary, this is the first extensive study of *CYP2C9* haplotype and *in vitro* functional analysis in a Hispanic population of Mexican descent. Interindividual and interethnic differences were observed in the incidence and frequency of *CYP2C9* regulatory polymorphisms. Further, based on *in vitro* assays, several of the inferred haplotypes are predicted to significantly decrease and/or alter PXR-mediated rifampicin-dependent *CYP2C9* induction. These observations suggest that genetic variation within *CYP2C9* regulatory

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

sequences likely contributes to differences in *CYP2C9* phenotype both within and among different populations. However, *in vivo* studies will be required to determine the relevance of these haplotype variants, if any, to clinical outcomes.

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Acknowledgements

The authors would like to acknowledge the expert assistance of Min Le, Ph.D. and Sevasti B. Koukouritaki, Ph.D. in completing these studies.

References

Baes M, Gulick T, Choi HS, Martinoli MG, Simha D and Moore DD (1994) A New Orphan Member of the Nuclear Hormone Receptor Superfamily That Interacts With a Subset of Retinoic Acid Response Elements. *Mol Cell Biol* 14:1544-1552.

Blaisdell J, Jorge-Nebert LF, Coulter S, Ferguson SS, Lee SJ, Chanas B, Xi T, Mohrenweiser H, Ghanayem B and Goldstein JA (2004) Discovery of New Potentially Defective Alleles of Human CYP2C9. *Pharmacogenetics* 14:527-537.

Buckland PR, Hoogendoorn B, Coleman SL, Guy CA, Smith SK and O'Donovan MC (2005) Strong Bias in the Location of Functional Promoter Polymorphisms. *Hum Mutat* 26:214-223.

Chen Y, Ferguson SS, Negishi M and Goldstein JA (2004) Induction of Human CYP2C9 by Rifampicin, Hyperforin, and Phenobarbital Is Mediated by the Pregnane X Receptor. *J Pharmacol Exp Ther* 308:495-501.

Chen Y, Kissling G, Negishi M and Goldstein JA (2005) The Nuclear Receptors Constitutive Androstane Receptor and Pregnane X Receptor Cross-Talk With Hepatic Nuclear Factor 4alpha to Synergistically Activate the Human CYP2C9 Promoter. *J Pharmacol Exp Ther* 314:1125-1133.

Davies NM, McLachlan AJ, Day RO and Williams KM (2000) Clinical Pharmacokinetics and Pharmacodynamics of Celecoxib: a Selective Cyclo-Oxygenase-2 Inhibitor. *Clin Pharmacokinet* 38:225-242.

Dickmann LJ, Rettie AE, Kneller MB, Kim RB, Wood AJJ, Stein CM, Wilkinson GR and Schwarz UI (2001) Identification and Functional Characterization of a New CYP2C9 Variant (CYP2C9*5) Expressed Among African Americans. *Mol Pharmacol* 60:382-387.

Ferguson SS, LeCluyse EL, Negishi M and Goldstein JA (2002) Regulation of Human *CYP2C9* by the Constitutive Androstane Receptor: Discovery of a New Distal Binding Site. *Mol Pharmacol* 62:737-746.

Gage BF, Eby C, Milligan PE, Banet GA, Duncan JR and McLeod HL (2004) Use of Pharmacogenetics and Clinical Factors to Predict the Maintenance Dose of Warfarin. *Thromb Haemost* 91:87-94.

Gerbal-Chaloin S, Pascussi J-M, Pichard-Garcia L, Daujat M, Waechter F, Fabre J-M, Carrère N and Maurel P (2001) Induction of CYP2C Genes in Human Hepatocytes in Primary Culture. *Drug Metab Dispos* 29:242-251.

Haining RL, Hunter AP, Veronese ME, Trager WF and Rettie AE (1996) Allelic Variants of Human Cytochrome P450 2C9: Baculovirus- Mediated Expression, Purification, Structural Characterization, Substrate Stereoselectivity, and Prochiral Selectivity of the Wild-Type and I359L Mutant Forms. *Arch Biochem Biophys* 333:447-458.

Ibeanu GC and Goldstein JA (1995) Transcriptional Regulation of Human CYP2C Genes: Functional Comparison of CYP2C9 and CYP2C18 Promoter Regions. *Biochemistry* 34:8028-8036.

Imai J, Ieiri I, Mamiya K, Miyahara S, Furuumi H, Nanba E, Yamane M, Fukumaki Y, Ninomiya H, Tashiro N, Otsubo K and Higuchi S (2000) Polymorphism of the Cytrochrome

P450 (CYP) 2C9 Gene in Japanese Epileptic Patients: Genetic Analysis of the *CYP2C9* Locus. *Pharmacogenetics* 10:85-89.

Kamali F, Khan TI, King BP, Frearson R, Kesteven P, Wood P, Daly AK and Wynne H (2004) Contribution of Age, Body Size, and CYP2C9 Genotype to Anticoagulant Response to Warfarin. *Clin Pharmacol Ther* 75:204-212.

Kawashima S, Kobayashi K, Takama K, Higuchi T, Furihata T, Hosokawa M and Chiba K (2006) Involvement of Hepatocyte Nuclear Factor 4alpha in the Different Expression Level Between CYP2C9 and CYP2C19 in the Human Liver. *Drug Metab Dispos* 34:1012-1018.

Kidd RS, Curry TB, Gallagher S, Edeki T, Blaisdell J and Goldstein JA (2001) Identification of a Null Allele of CYP2C9 in an African-American Exhibiting Toxicity to Phenytoin. *Pharmacogenetics* 11:803-808.

King BP, Khan TI, Aithal GP, Kamali F and Daly AK (2004) Upstream and Coding Region CYP2C9 Polymorphisms: Correlation With Warfarin Dose and Metabolism. *Pharmacogenetics* 14:813-822.

Klose TS, Ibeanu GC, Ghanayem BI, Pedersen LG, Li L, Hall SD and Goldstein JA (1998) Identification of Residues 286 and 289 As Critical for Conferring Substrate Specificity of Human CYP2C9 for Diclofenac and Ibuprofen. *Arch Biochem Biophys* 357:240-248.

Kuo CJ, Conley PB, Hsieh C-L, Francke U and Crabtree GR (1990) Molecular Cloning, Functional Expression, and Chromosomal Localization of Mouse Hepatocyte Nuclear Factor 1. *Proc Natl Acad Sci USA* 87:9839-9842.

Lee CR, Goldstein JA and Pieper JA (2002) Cytochrome P450 2C9 Polymorphisms: A Comprehensive Review of the in Vitro and Human Data. *Pharmacogenetics* 12:251-263.

Lehmann JM, McKee DD, Watson MA, Wilson TM, Moore JT and Kliewer SA (1998) The Human Orphan Nuclear Receptor PXR Is Activated by Compounds That Regulate CYP3A4 Gene Expression and Cause Drug Interactions. *J Clin Invest* 102:1016-1023.

Lindblad-Toh K, Winchester E, Daly MJ, Wang DG, Hirschhorn JN, Laviolette J-P, Ardlie K, Reich DE, Robinson E, Sklar P, Shah N, Thomas D, Fan J-B, Gingeras T, Warrington J, Patil N, Hudson TJ and Lander ES (2000) Large-Scale Discovery and Genotyping of Single-Nucleotide Polymorphisms in the Mouse. *Nat Genet* 24:381-386.

Llerena A, Dorado P, O'Kirwan F, Jepson R, Licinio J and Wong ML (2004) Lower Frequency of CYP2C9*2 in Mexican-Americans Compared to Spaniards. *Pharmacogenomics J* 4:403-406.

Martinez-Jimenez CP, Gomez-Lechon MJ, Castell JV and Jover R (2005) Transcriptional Regulation of the Human Hepatic CYP3A4: Identification of a New Distal Enhancer Region Responsive to CCAAT/Enhancer-Binding Protein Beta Isoforms (Liver Activating Protein and Liver Inhibitory Protein). *Mol Pharmacol* 67:2088-2101.

Mendel DB, Khavari PA, Conley PB, Graves MK, Hansen LP, Admon A and Crabtree GR (1991) Characterization of a Cofactor That Regulates Dimerization of a Mammalian Homeodomain Protein. *Science* 254:1762-1767.

Miners JO and Birkett DJ (1998) Cytochrome P4502C9: An Enzyme of Major Importance in Human Drug Metabolism. *Br J Clin Pharmacol* 45:525-538.

Niemi M, Backman JT, Neuvonen M, Neuvonen PJ and Kivisto KT (2001) Effects of Rifampin on the Pharmacokinetics and Pharmacodynamics of Glyburide and Glipizide. *Clin Pharmacol Ther* 69:400-406.

Phillips KA, Veenstra DL, Oren E, Lee JK and Sadee W (2001) Potential Role of Pharmacogenomics in Reducing Adverse Drug Reactions: a Systematic Review. *J Amer Med Assn* 286:2270-2279.

Rettie AE, Wienkers LC, Gonzalez FJ, Trager WF and Korzekwa KR (1994) Impaired (S)-Warfarin Metabolism Catalyzed by the R144C Allelic Variant of CYP2C9. *Pharmacogenetics* 4:39-42.

Sandberg M, Johansson I, Christensen M, Rane A and Eliasson E (2004) The Impact of CYP2C9 Genetics and Oral Contraceptives on Cytochrome P450 2C9 Phenotype. *Drug Metab Dispos* 32:484-489.

Schwarz UI (2003) Clinical Relevance of Genetic Polymorphisms in the Human CYP2C9 Gene. *Eur J Clin Invest* 33 Suppl 2:23-30.

Shimada T, Yamazaki H, Mimura M, Inui Y and Guengerich FP (1994) Interindividual Variations in Human Liver Cytochrome P-450 Enzymes Involved in the Oxidation of Drugs, Carcinogens and Toxic Chemicals: Studies With Liver Microsomes of 30 Japanese and 30 Caucasians. *J Pharmacol Exp Ther* 270:414-423.

Shintani M, Ieiri I, Inoue K, Mamiya K, Ninomiya H, Tashiro N, Higuchi S and Otsubo K (2001) Genetic Polymorphisms and Functional Characterization of the 5'-Flanking Region of the Human CYP2C9 Gene: In Vitro and in Vivo Studies. *Clin Pharmacol Ther* 70:175-182.

Stephens M and Donnelly P (2003) A Comparison of Bayesian Methods for Haplotype Reconstruction From Population Genotype Data. *Am J Hum Genet* 73:1162-1169.

Stoffel M and Duncan SA (1997) The Maturity-Onset Diabetes of the Young (MODY1) Transcription Factor HNF4a Regulates Expression of Genes Required for Glucose Transport and Metabolism. *Proc Natl Acad Sci USA* 94:13029-13214.

Takahashi H, Ieiri I, Wilkinson GR, Mayo G, Kashima T, Kimura S, Otsubo K and Echizen H (2004) 5'-Flanking Region Polymorphisms of CYP2C9 and Their Relationship to S-Warfarin Metabolism in White and Japanese Patients. *Blood* 103:3055-3057.

Veenstra DL, Blough DK, Higashi MK, Farin FM, Srinouanprachan S, Rieder MJ and Rettie AE (2005) CYP2C9 Haplotype Structure in European American Warfarin Patients and Association With Clinical Outcomes. *Clin Pharmacol Ther* 77:353-364.

Williams JA, Hyland R, Jones BC, Smith DA, Hurst S, Goosen TC, Peterkin V, Koup JR and Ball SE (2004) Drug-Drug Interactions for UDP-Glucuronosyltransferase Substrates: a Pharmacokinetic Explanation for Typically Observed Low Exposure (AUCi/AUC) Ratios. *Drug Metab Dispos* 32:1201-1208.

Williamson KM, Patterson JH, McQueen RH, Adams KF, Jr. and Pieper JA (1998) Effects of Erythromycin or Rifampin on Losartan Pharmacokinetics in Healthy Volunteers. *Clin Pharmacol Ther* 63:316-323.

Xie HG, Prasad HC, Kim RB and Stein CM (2002) CYP2C9 Allelic Variants: Ethnic Distribution and Functional Significance. *Adv Drug Deliv Rev* 54:1257-1270.

Footnotes

* This work was supported in part by National Institutes of Health Grant GM068797 and funds from the Children Research Institute, Children's Hospital and Health Systems.

[1] Current address: University of Minnesota, College of Pharmacy, Department of Experimental and Clinical Pharmacology, 3-120 Weaver-Densford Hall, 308 Harvard St. SE, Minneapolis, MN 55455

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Figure Legends

Figure 1: Effect of Variant Haplotypes on Constitutive *CYP2C9* Promoter Activity. HepG2 cells were transfected with reference or variant haplotype constructs and analyzed for transient luciferase expression. Individual plasmid numbers and the haplotypes which they represent are shown on the ordinate. Luciferase activities were normalized for transfection efficiency using β-galactosidase activity and were compared to the reference haplotype 1 (pRNH954). Each bar indicates the relative luciferase activity observed following 24 hrs. Data are depicted as the mean ± SD of at least three determinations. (*, $P<0.05$; ANOVA, Holm-Sidak post test).

Figure 2: Effect of Variant Haplotypes on the PXR-Mediated Induction of *CYP2C9* Promoter Activity. HepG2 cells were transfected with reference or variant haplotype constructs and analyzed for transient luciferase expression following treatment with10 μM rifampicin for 24 hrs. Individual plasmid numbers and the haplotypes which they represent are shown on the ordinate. Luciferase activities were normalized for transfection efficiency using β-galactosidase activity and were compared to vehicle treated controls. Data represent the mean ± SD of at least three determinations. (*, $P<0.05$; ANOVA, Holm-Sidak post test). (A) Induced expression as compared to the vehicle-treated control; (B) Fold induction relative to vehicle-treated control.

Table 1: Discovery of *CYP2C9* Promoter SNPs

Discovery performed using the Coriell 24PDR (n=48 chromosomes)

SNP Position[a]	Accession No.[b]	Minor Alleles Observed	SNP and Sequence Context[c]	Putative Transcription Factor Binding Site[d]	Transcription Factor
-8897[e]		2	TTAAC[T>C]TAAAA		
-8553[e]		1	CACTG[C>A]AACCT		
-8435		4	AATTT[C>A]ACCAT		
-8430		4	CACCA[T>G]GTTGG		
-8422		4	TGGCC[A>G]GGCTG		
-8416[e]		4	GGCTG[T>G]TCTCG	**TGGCCA**N_3**TGTTCT**	GR
-8378[e]		4	CCTTG[G>T]CCTCC		
-7982		1	TTGGT[C>A]TAACA		
-7432		1	GGAAA[C>A]TACAA		
-7419[e]		1	AAGTA[A>G]GAAAA		
-7336[e]		2	AAAGA[G>A]TTGAG		

-5813[e]		8	AGAGG[A>G]AATTC		
-5661[e]		3	CCAAT[C>A]GTGTA		
-5146		3	AAAAA[G>C]AAAAC		
-5143		3	AAGAA[A>C]ACAAC		
-5140		2	AAAAC[A>T]ACAAT		
-4877[e]		2	TCATG[G>A]ATATG		
-4302		1	TAACA[C>T]GGTGA		
-3597[e]		1	TGCTC[A>G]TCATT	**C<u>G</u>TCAT**	CREB
-3579[e]		3	ACTAC[G>A]GACCT		
-3360[e]		3	TGCTC[T>C]TTGGT	CRSCTGTBBNN<u>T</u>**TTGGC**ACB	NF1/CTF
-3089[f]		4	CAACC[G>A]TATTA		
-2663[i]		2	GACTG[+/-]GAGGG		
-1911[g]	rs9332902	4	AGTTA[T>C]TGCTT		
-1885[g]	rs9332093	2	AAAGG[C>G]TTCTC		

-1188[g]	rs4918758	3	ATCTT[T>C]TATTG
-1096[h]	rs4917636	3	ACAAT[A>G]GAAAG
-981[g]	rs9332098	4	ATGGA[G>A]AAGGG
-620[h]	rs9332100	1	TTAAT[G>T]GTAAA
-485[h]	rs9332101	3	GGATT[T>A]CATTA
-484[h]	rs9332102	3	GATTT[C>A]ATTAT

[a] Coordinates are numbered relative to the +1 of the ATG start codon and use contig NT_030059.12, build 36.1 as a reference.

[b] Accession number is that reported in dbSNP build 121.

[c] Diallelic insertion/deletion polymorphisms are represented as a plus sign for allele insertion and a minus sign for allele deletion. The + (inserted) allele for site -2663 is TG.

[d] Putative transcription factor binding matrices as reported in the TRANSFAC Professional V7.4 database are shown with the core binding matrix in bold font, the position of the SNP underscored

[e] First reported by Veenstra *et al.* (2005)

[f] First reported by Blaisdell *et al.* (2004)

[g] First reported by Shintani *et al.* (2001)

[h] First reported by Takahashi *et al.* (2004)

[i] First reported by King *et al.* (2004)

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

Table 2: Comparison of *CYP2C9* Promoter SNPs in Hispanic and Non-Latino White Populations

SNP Position[a]	Nucleotide Change[b]	Variant Allelic Frequencies (95% CI) Hispanic (Mexican descent) (n = 386)	Non-Latino White (n = 384)	Linkage Disequilibrium[c]
-8897	T>C	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)	*3
-8553	C>A	0.078 (0.047, 0.108)	0.060 (0.033, 0.087)	*3
-8422	A>G	0.003 (0.000, 0.008)	NO[d]	
-8416	T>G	0.003 (0.000, 0.008)	0.001 (0.000, 0.005)	
-7419	A>G	0.044 (0.021, 0.067)	0.070 (0.041, 0.099)	*3
-7336	G>A	0.003 (0.000, 0.008)	0.001 (0.000, 0.005)	
-5813	A>G	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)	*3
-5661	C>A	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)	*3
-5146	G>C	0.005 (0.000, 0.013)	NO	
-5143	A>C	0.005 (0.000, 0.013)	NO	
-5140	A>T	0.005 (0.000, 0.013)	NO	
-4877	G>A	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)	*3
-4302	C>T	0.109 (0.073, 0.144)[e]	NO	

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

-3597	A>G	0.135 (0.096, 0.173)	0.180 (0.136, 0.224)	
-3579	G>A	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)	*3
-3360	T>C	0.132 (0.094, 0.171)	0.180 (0.136, 0.224)	*2
-3089	G>A	0.109 (0.073, 0.144)[f]	0.170 (0.127, 0.213)	
-2663	+/-	0.073 (0.043, 0.102)[e]	0.170 (0.127, 0.213)	
-1911	T>C	0.044 (0.021, 0.067)	0.070 (0.041, 0.099)	*3
-1885	C>G	0.041 (0.022, 0.061)	0.060 (0.036, 0.084)	*3
-1537	G>A	0.044 (0.021, 0.067)	0.070 (0.041, 0.099)	*3
-1188	T>C	0.207 (0.161, 0.253)[e]	0.350 (0.296, 0.404)	
-1096	A>G	0.070 (0.041, 0.099)[f]	0.120 (0.083, 0.157)	*2
-981	G>A	0.044 (0.021, 0.067)	0.070 (0.041, 0.099)	*3
-620	G>T	0.049 (0.025, 0.074)[f]	0.110 (0.074, 0.146)	*2
-485	T>A	0.070 (0.041, 0.099)	0.110 (0.074, 0.146)	*2
-484	C>A	0.070 (0.041, 0.099)	0.110 (0.074, 0.146)	*2

Molecular Pharmacology Fast Forward. Published on February 29, 2008 as DOI: 10.1124/mol.107.044149
This article has not been copyedited and formatted. The final version may differ from this version.

[a] Coordinates are numbered relative to the +1 of the ATG start codon and use contig NT_030059.12, build 36.1 as a reference.

[b] Diallelic insertion/deletion polymorphisms are represented as a plus sign for allele insertion and a minus sign for allele deletion. The + (inserted) allele for site -2663 is TG.

[c] Linkage Disequilibrium in non-Latino White population

[d] NO, not observed.

[e] Different from non-Latino White study population, $P < 0.001$ (Fisher's exact test).

[f] Different from non-Latino White study population, $P < 0.05$ (Fisher's exact test).

Table 3: Comparison of *CYP2C9* Coding Region SNPs in Hispanic and Non-Latino White Populations

SNP Position[a]	Site	Nucleotide Change	Effect	Allele Designation	Variant Allelic Frequencies (95% CI) Hispanic (Mexican descent) (n = 386)	Non-Latino White (n = 384)
3608	Exon 3	430C>T	Arg144Cys	*CYP2C9*2*	0.070 (0.041, 0.099)	0.110 (0.074, 0.146)
42614	Exon 7	1075A>C	Ile359Leu	*CYP2C9*3*	0.044 (0.021, 0.067)	0.060 (0.033, 0.087)
42615	Exon 7	1076T>C	Ile359Thr	*CYP2C9*4*	NO[b]	NA[c]
42619	Exon 7	1080C>G	Asp360Glu	*CYP2C9*5*	NO	NA
10601	Exon 5	delA	Frame Shift	*CYP2C9*6*	NO	NA

[a] Coordinates are numbered relative to +1 of the ATG start codon and use contig NT_030059.12, build 36.1 as a reference.

[b] NO, not observed.

[c] NA, not applicable.

Table 4: Common (>1%) CYP2C9 Haplotypes in Hispanic and Non-Latino White Populations

Haplotype Number	Plasmid Number	Nucleotide Changes[a,b]	Mean Frequency (mean ± SEM) Hispanic (Mexican) (n = 386)	Non-Latino White (n = 384)
1	954	Reference Sequence	60.53 ± 0.27	63.64 ± 0.12
1A	955	-8553C>A	3.21 ± 0.16[d]	NO[c]
1B	974	-4302C>T	10.01 ± 0.20[d]	NO
1C	956	-3089G>A, -1188T>C	3.10 ± 0.10[e]	0.27 ± 0.10
1D	957	-3089G>A, -2663+>–, -1188T>C	6.61 ± 0.15[d]	17.23 ± 0.13
1E	958	-1188T>C	1.33 ± 0.25	NO
2A	959	-3597A>G, -3360T>C, -1188T>C, -1096A>G, -485T>A, -484C>A, 3608C>T	2.08 ± 0.06[d]	NO
2B	960	-3597A>G, -3360T>C, -1188T>C, -1096A>G, -620G>T, -485T>A, -484C>A, 3608C>T	4.23 ± 0.14[d]	10.93 ± 0.03

3A	962	-8897C>A, -8553C>A, -7419A>G, -5813A>G, -5661C>A, -4877G>A, -3597A>G, -3579G>A, -3360T>C, -1911T>C, -1885C>G, -1537G>A, -981G>A, 42614A>C	1.56 ± 0.05[f]	0.07 ± 0.13
3B	961	-8897C>A, -8553C>A, -7419A>G, -5813A>G, -5661C>A, -4877G>A, -3597A>G, -3579G>A, -3360T>C, -1911T>C, -1885C>G, -1537G>A, -1188T>C, -981G>A, 42614A>C	2.58 ± 0.05[f]	6.35 ± 0.20

[a] Coordinates are numbered relative to the +1 of the ATG start codon and use contig NT_030059.12, build 36.1 as a reference.

[b] Diallelic insertion/deletion polymorphisms are represented as a plus sign for allele insertion and a minus sign for allele deletion. The + (inserted) allele for site -2663 isTG.

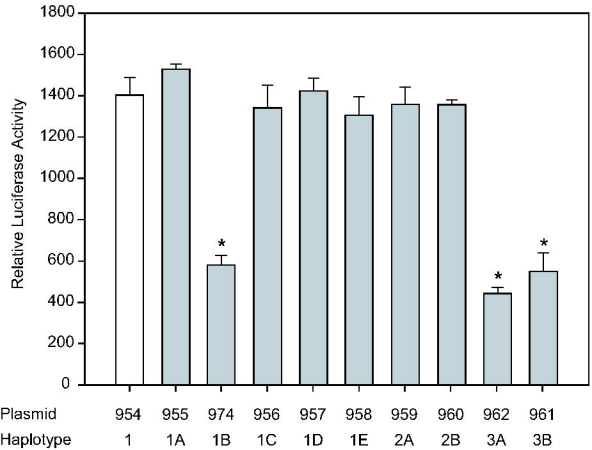
[c] NO, not observed.

[d] Different from Non-Latino White study population, P< 0.001 (Student's t test)

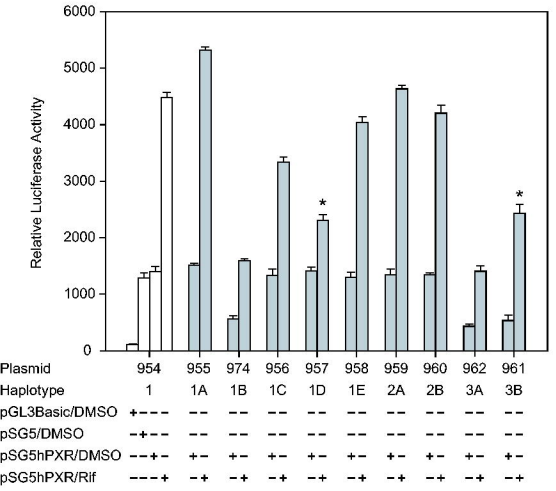
[e] Different from Non-Latino White study population, P< 0.01 (Student's t test)

[f] Different from Non-Latino White study population, P< 0.05 (Student's t test)

Figure 1


Relative Luciferase Activity
1800
1600
1400
1200
1000
800
600
400
200
0
*
*
*
Plasmid 954 955 974 956 957 958 959 960 962 961
Haplotype 1 1A 1B 1C 1D 1E 2A 2B 3A 3B

A


Relative Luciferase Activity
6000
5000
4000
3000
2000
1000
0
*
*
Plasmid 954 955 974 956 957 958 959 960 962 961
Haplotype 1 1A 1B 1C 1D 1E 2A 2B 3A 3B
pGL3Basic/DMSO + − − − − − − − − − − − − − − − − − − −
pSG5/DMSO − − + − − − − − − − − − − − − − − − − −
pSG5hPXR/DMSO − − − + − + − + − + − + − + − + − + − +
pSG5hPXR/Rif − − − + − + − + − + − + − + − + − + − +

B

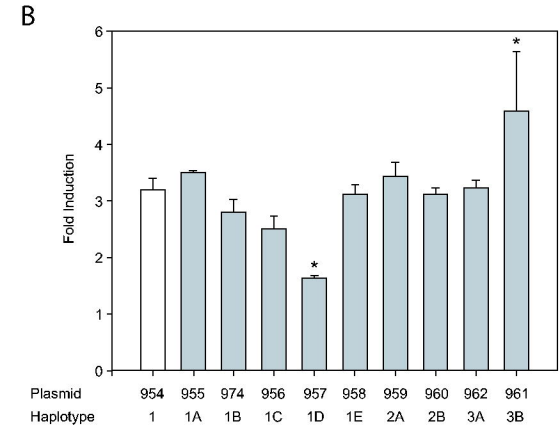

Fold Induction
6
5
4
3
2
1
0
*
*
Plasmid 954 955 974 956 957 958 959 960 962 961
Haplotype 1 1A 1B 1C 1D 1E 2A 2B 3A 3B

Table S1. Polymerase Chain Reaction DNA Amplification and Sequencing Primers

Forward Primer Sequence	Reverse Primer Sequence	*CYP2C9* Position
GAATGCAGAGTACCAGGGGTC	CTGCACCTGTGGATATGATGAT	-10109 to -5517
CATCTTGTCCCACTGTAAGTT	GTACAAAAATTAGCTGGC	-9021 to -8456
GAAGCTGTAATCGCCTATGTAAC	ACAGTCATGCACCATATACC	-8978 to -8249
CATCTCTGAGTTTCAAGCTAT	GTGTTGTCTGTTAGACCAATC	-8546 to -7969
AGGAAGGGAGTTAGTGAAG	AATTAACCCAAGTCTACATTA	-8046 to -7462
GTTATTCACTGATCTACATGAGG	GTCTCTGTCTTTCCAACGG	-8071 to -7219
AGTGCACTACCTATGAAGTAG	AGTGCACTACCTATGAAGTAG	-7516 to -6946
GGGAATAGCTATATTAGTTTC	TTTAATCCCTCTTGAGATTTA	-7079 to -6490
ATAATAAATCTCAAGAGGGAT	GGCCTATTCAAATTATCTA	-6514 to -5963
TCCATAAAGAAACAATCTTC	AGTTCCTCCTCTATTCTTAGT	-6022 to -5444
CCCGTTCTCTACAATCTC	TTGACAACCTGGGACTAC	-5852 to -4223
AAGTCTCATCATCATATCCAC	CATATTTTTGGTGCTAAT	-5545 to -4990
ACTCCTATTCAACATCGTACTA	CTTTCCTGGAGTTTGATTCTTA	-5329 to -4805
AGTATCATTTATATTAGCACC	ATTCATTTTTCCGTATGT	-5019 to -4456
ATGGTCTTTTCAACGAAG	TGCATTTATATTTGGATCTA	-4514 to -3957
TTGAAGTGAGCCGATTGA	GTTTTTAGTTTGCAGGTCCGT	-4180 to 3561
AAATGGGCAACAGATCTGAAC	CAAGCCCTAGCAACAAATAATC	-3673 to -3025
TATGAAAAATGCTCATCATTA	CGGTTGGAATCATACAGT	-3611 to -3089
GGTAGTTGTGCTCTTTGG	AGAAACTAGGGCTTCTCG	-3373 to -1045
AATCTTAAAAGGCTACATACT	GGCAGGATGATCTTGAT	-3124 to -2453
TGAAAAACTGGGATTCTAAGA	AATAAGCTCGGTTGCTGT	-2573 to -1977
TGAGGAATGAAATGATTATTA	GCACAAAGCTTTACTGGT	-2020 to -1444
GAGATTCAAAAGGGACATGAG	TCATGAGAAGAAATAAGGGAT	-1526 to -953
GGGACATGAGGTGTAACAAT	GAAGGAGCATACTTACATTGG	-1515 to +184
CCTTTGCTTTCTGTGCATTATTAC	TGGCCATATTCTATGCCTTG	-1152 to -405
ATTTTTCCCTCAGTTACACT	TTTAAAAAGTATGTCAATGCT	-1027 to -429
CTGCTGTATTTTAGTAGGC	AAGGAGAAGCAAACATGAGAG	-528 to +51
GAGGATGGAAAACAGAGACT	GGTCAGTGATATGGAGTAGGG	+3383 to +3688
GCTTGGTATATGGTATGTA	CAGAACTAGTCAACAAATC	+10359 to +10668
TGTGCATCTGTAACCATCC	TTTGGGGACTTCGAAAAC	+42335 to +42738

Table S2. Single-Base Extension Primers

SNP	Primer Sequence
	Group 1
-5813A>G	C_{15} AAAATGAGTTAAGAATAGAAGAATT
-5661C>A	C_3 AGATTAGCAAATTGCATCCAAT
-4302C>T	C_{22} GAGATCGAGACCATCCTGGCTAACA
-3089G>A	C_{10} GCCTTTTCCAAAATGTAATA
	Group 2
-2663delTG	C_3 GAGTATCAACATTAAGCCCTCC
-1911T>C	C_{15} GTTTCATGAGTCAGGGACCAAGTTA
-1885C>G	C_5 TGCTTTTCTTTGCCCTGTATAAAGG
-1188T>C	C_{22} TAGTGATTTCCCTACCTCCCATCTT
	Group 3
+3608C>T	C_5 GGAAGAGGAGCATTGAGGAC
+10601delA	C_{18} TGCTTCCTGATGAAAATGGAGA
+42614A>C	GCACGAGGTCCAGAGATAC
+42615T>C	C_{18} TGGTGGGGAGAAGSTCA
+42619C>G	C_{11} CAGGCTGGTGGGGAGAAG
	Other
-1537G>A	GGAAATGGGTCAATTTTATTGTAAGCA