JPET Fast Forward. Published on October 18, 2006 as DOI: 10.1124/jpet.106.112268

JPET #112268

Identification and Functional Analysis of Common Human Flavin-containing Monooxygenase 3 (*FMO3)* Genetic Variants

Sevasti B. Koukouritaki, Mark T. Poch, Marilyn C. Henderson, Lisbeth K. Siddens, Sharon K. Krueger, Jonathan E. VanDyke, David E. Williams, Nicholas M. Pajewski, Tao Wang, and Ronald N. Hines

Departments of Pediatrics (S.B.K., M.T.P., R.N.H.), Pharmacology and Toxicology (R.N.H.), and Biostatistics (N.M.P., T.W.), Medical College of Wisconsin, and Children's Research Institute, Children's Hospital and Health System (S.B.K., M.T.P., R.N.H.), Milwaukee, Wisconsin; and Department of Environmental and Molecular Toxicology, The Linus Pauling Institute, Oregon State University, Corvallis, Oregon (M.C.H., L.K.S., S.K.K., J.E.V., D.E.W.)

JPET Fast Forward. Published on October 18, 2006 as DOI: 10.1124/jpet.106.112268
This article has not been copyedited and formatted. The final version may differ from this version.

Running Head: Functional Analysis of Human *FMO3* Haplotypes

Corresponding Author: Ronald N. Hines

Department of Pediatrics

Medical College of Wisconsin

8701 Watertown Plank Rd.

Milwaukee, WI 53226, U.S.A.

Tel: (414) 456-4322

Fax: (414) 456-6651

E-mail: rhines@mail.mcw.edu

Number of Text Pages: 36

Number of Tables: 3

Number of Figures: 1

Number of References: 37

Number of Words in Abstract: 249

Number of Words in Introduction: 730

Number of Words in Discussion: 1510

Non-Standard Abbreviations: AMOVA, analysis of molecular variance; FMO, flavin-containing monooxygenase; SBE, single base extension; SNP, single nucleotide polymorphism

Recommended Section Assignment: Metabolism, Transport and Pharmacogenomics

ABSTRACT

Flavin-containing monooxygenases (FMOs) are important for the disposition of many therapeutics, environmental toxicants, and nutrients. FMO3, the major adult hepatic FMO enzyme, exhibits significant interindividual variation. Eighteen *FMO3* single nucleotide polymorphism (SNP) frequencies were determined in 202 Hispanics (Mexican descent), 201 African Americans, and 200 non-Latino whites. Using expressed recombinant enzyme with methimazole, trimethylamine, sulindac, and ethylenethiourea, the novel structural variants, FMO3 E24D and K416N, were shown to cause a modest changes in catalytic efficiency, while a third novel variant, FMO3 N61K, was essentially devoid of activity. The latter variant was present at an allelic frequency of 5.2% in non-Latino whites and 3.5% in African Americans, but was absent in Hispanics. Inferring haplotypes using PHASE V2.1, the greatest haplotype diversity was observed in African Americans followed by non-Latino whites and Hispanics. Haplotype 2A and 2B, consisting of a hypermorphic promoter SNP cluster (-2650C>G, -2543T>A, and -2177G>C) in linkage with synonymous structural variants was inferred at a frequency of 27% in the Hispanic population, but only 5% in non-Latino whites and African Americans. This same promoter SNP cluster in linkage with one or more hypomorphic structural variant also was inferred in multiple haplotypes at a total frequency of 5.6% in the African American study group, but less than 1% in the other two groups. The sum frequencies of the hypomorphic haplotypes H3[15,167G>A (E158K)], H5B [-2650C>G, 15,167G>A (E158K), 21,375C>T (N285N), 21,443A>G (E308G)], and H6 [15,167G>A (E158K), 21,375C>T (N285N)] were 28% in Hispanics, 23% in non-Latino whites and 24% in African Americans.

INTRODUCTION

The flavin-containing monooxygenases (FMOs; EC 1.14.13.8) are a family of microsomal enzymes that catalyze the NADPH-dependent N- and S-oxidation of a variety of therapeutics, environmental toxicants, carcinogens and nutrients (Krueger and Williams, 2005). The *FMO* multi-gene family consists of a five gene cluster at 1q24.3 (*FMO1-4* and *FMO6p*), a second cluster of five genes at 1q24.2 (*FMO7p-11p*), and a single gene, *FMO5*, at 1q21.1, encoding a total of five active proteins in the human (Hernandez, et al., 2004). The most recent common precursor for all placental mammals is predicted to have already had a cluster containing *FMO1-6P* and a separate *FMO5* locus that arose from duplication of an ancestral gene approximately 210 to 275 million years ago (Hernandez, et al., 2004). Despite the antiquity of the *FMO* gene family, a sequence comparison among modern FMO enzymes reveals 76 to 86% sequence identity between orthologous proteins, suggesting these genes have been highly conserved.

The individual FMO enzymes exhibit broad, but distinct substrate specificities (Krueger and Williams, 2005), as well as species-, sex-, tissue-, and age-dependent differences in expression patterns (reviewed in Hines, 2006). In the human, FMO3 is the predominant adult hepatic enzyme with a specific content of 60 ± 43 pmol/mg microsomal protein (Overby, et al., 1997), comparable to the most abundant adult liver cytochrome P450 enzymes, *i.e.*, CYP3A4 and CYP2C9 (Shimada, et al., 1994). Also similar to the cytochromes P450, 10- to 20-fold interindividual differences in *FMO* expression have been described that may contribute to differences in toxicant susceptibility and/or therapeutic efficacy (Overby, et al., 1997; Yeung, et al., 2000; Koukouritaki, et al., 2002; Hisamuddin, et al., 2005). However, unlike the cytochromes P450, the contribution of *FMO* induction to interindividual differences in expression is at best, controversial. Hukkanen *et al.*

(2005) described a significant increase in FMO3-dependent metabolism during pregnancy while Zhang *et al.* (1996) reported an exacerbation of trimethylaminuria symptoms with menstruation, suggestive of *FMO3* regulation by sex hormones. Yet, no gender differences were observed during postnatal FMO3 developmental expression (Koukouritaki, et al., 2002). Most recently, Tijet *et al.* (2006) reported an 80-fold induction of FMO3 mRNA levels in male mice that was dependent on both the presence of the aryl hydrocarbon receptor and treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin. Given that a strong sexual dimorphism exists in the mouse wherein *FMO3* expression is suppressed in the adult male, but not female (Cherrington, et al., 1998), one must wonder whether or not the observed FMO3 induction was related to an effect on testosterone-dependent *FMO3* suppression in this species.

The existence of polymorphisms affecting human FMO3-dependent metabolism has been well documented (reviewed in Koukouritaki and Hines, 2005). Twenty-seven *FMO3* allelic variants, including missense, nonsense or deletion mutants, that result in a complete or near complete loss of functional activity are causative for trimethylaminuria or "fish-odor syndrome" (http://bsmsol2.biochem.ucl.ac.uk/Human_FMO3/). Because these alleles are rare in most populations, they cannot account for the substantial interindividual differences in FMO3-dependent metabolism. Several studies have documented reduced activity with the relatively common FMO3 E158K and E308G variants *in vitro* (Störmer, et al., 2000) and *in vivo* (Park, et al., 2002; Hisamuddin, et al., 2005), although some substrate dependency is apparent. In some populations, these variants exhibit a high degree of linkage disequilibrium. When present on the same allele, the E158K and E308G exhibit an even more pronounced effect on FMO3 function (Park, et al., 2002), even leading to a mild or transient forms of trimethylaminuria (Zschocke, et al., 1999). Several

other functional FMO3 structural variants have been identified, but their frequency and as such, their relevance to population health is unknown (reviewed in Koukouritaki and Hines, 2005).

In a recent single nucleotide polymorphism (SNP) discovery study, we characterized 40 common *FMO3* SNPs using DNA samples from the Coriell Polymorphism Discovery Resource (Koukouritaki, et al., 2005). Eleven of these polymorphisms were located within exon sequences while seven novel promoter variants also were identified. The latter were used to infer seven promoter haplotypes that differ in frequency among the three different population groups studied; Hispanics of Mexican descent, African Americans and non-Latino whites. Three of the inferred haplotypes significantly altered *FMO3* promoter activity *in vitro*. To extend these findings, the objectives of the current study were to determine a more complete *FMO3* haplotype structure in these same three population groups and conduct functional analysis of structural variants predicted to impact FMO3 catalytic activity.

MATERIALS AND METHODS

Materials: *Platinum Taq* DNA Polymerase High Fidelity, the BaculoDirect C-Term Expression kit, the pENTR/SD-TOPO Cloning kit, and Cellfectin were purchased from Invitrogen (Carlsbad, CA). The ExoSAP-IT mix, containing both exonuclease I and shrimp alkaline phosphatase, as well as shrimp alkaline phosphatase alone, were purchased from USB Corp. (Cleveland, OH). CEQ SNP-Primer Extension and Dye Terminator Cycle Sequencing kits were obtained from Beckman Coulter, Inc. (Fullerton, CA). CleanSEQ magnetic beads were purchased from Agencourt (Beverly, MA). Custom oligonucleotides were synthesized by IDT (Coralville, IA). Methimazole and ethylenethiourea were purchased from Lancaster Synthesis (Pelham, NH) while NADPH, FAD, sulindac and sulindac sulfide were obtained from Sigma-Aldrich (St. Louis, MO). All other reagents were obtained from commercial sources at the purest grade available.

Patient Recruitment and DNA Isolation: Individuals representing various ethnic and/or racial groups were recruited to provide DNA samples as described previously (Koukouritaki, et al., 2005). In all instances, ethnicity and/or race was self-reported. Combining all DNA sources resulted in 202 samples from unrelated individuals of Hispanic (Mexican) descent, 201 samples from unrelated individuals of African-American descent and 200 samples from unrelated individuals of Non-Latino white (Northern European) descent. Research protocols were approved by all Institutional Review Boards involved.

DNA Amplification: Template amplifications were performed in a final volume of 25 μL and contained 50 ng of genomic DNA, 2 mM $MgSO_4$, 0.2 mM of each deoxyribonucleotide triphosphate, 0.2 μM of each primer (see supplementary Table S1), and 1 unit of Platinum *Taq* DNA Polymerase High Fidelity. All polymerase chain reactions had an initial denaturation step at 94°C

for 2 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec, and elongation at 68°C for 1 min. The final elongation step was performed at 68°C for 5 min.

Multiplexed Single Base Extension (SBE) Assay: Genotype frequencies of common *FMO3* variants in different populations were determined using a multiplexed SBE assay as previously described (Koukouritaki, et al., 2005). Sequences of the SBE primers used may be found in supplement Table S2. For a positive control, 20 fmol of linearized pRNH926 (reference sequence), pRNH927 (containing all queried SNPs), or an equal mixture of both were used as templates in the SBE reaction. The *FMO3* sequence reported in NCBI accession number AL021026 was used as the reference for both the *FMO3* gene and, when translated *in silico*, the FMO3 enzyme.

Plasmids: An FMO3 variant 1 transcript, NM_001002294, position 46 to position 1736 was amplified from an adult human liver total RNA sample (Stratagene, La Jolla, CA) by reverse-transcriptase coupled polymerase chain reaction and cloned into the pCR2.1 vector to generate pRNH696. The fidelity of the cloned cDNA was verified by DNA sequence analysis on both strands. An FMO3 cDNA fragment, position +92 to +1698, was amplified using linearized pRNH696 as a template and 5'-GAT GAT TAG GTC AAC ACA AG-3' and 5'-CAC CAT GGG GAA GAA AGT G-3' as forward and reverse primers, respectively. The amplified product was cloned into pENTR/SD/D-TOPO, resulting in pRNH829. Site directed mutagenesis was performed with the QuickChange® Site-Directed Mutagenesis kit (Stratagene) to introduce single nucleotide variants in pRNH829, resulting in: pRNH865 encoding FMO3 E24D, pRNH866 encoding FMO3 N61K, and pRNH904 encoding FMO3 K416N.

To produce a control plasmid for SBE genotyping that contains all eleven structural variants, a fragment representing FMO3 cDNA position 1 to 89 was amplified and cloned into pRNH829 immediately upstream of the existing cDNA fragment, generating pRNH926. Site-directed

mutagenesis was performed using a QuikChange® Multi Site-Directed Mutagenesis kit (Stratagene) to introduce the desired single nucleotide changes into pRNH926, creating pRNH927. Nucleotide changes and the fidelity of non-targeted sequences were verified by DNA sequence analysis.

DNA Sequence Analysis: DNA sequencing was performed using a Dye Terminator Cycle Sequencing kit (Beckman Coulter) according to the manufacturer's instructions. Briefly, plasmid DNA templates were pre-heated at 96°C for 1 min followed by cooling to room temperature on the bench-top. Sequencing reactions were performed in a final volume of 20 μL and contained 50 fmol of plasmid DNA, 1.6 μM forward or reverse primer and 8.0 μL of Dye Terminator Cycle Sequencing Quick Start Master Mix (DTCS). All sequencing reactions included 50 cycles of denaturation at 96°C for 20 sec, annealing at 50°C for 20 sec, and elongation at 60°C for 4 min. Unincorporated primer and nucleotides were removed from the reactions using CleanSEQ magnetic beads (Agencourt) according to the manufacturer's instructions.

FMO3 Variant Expression: The FMO3 cDNA inserts from the pENTR clones were integrated with BaculoDirect linear DNA by LR-mediated Clonase II recombination. Sf9 cell transfection with the recombination reaction mixture was performed with Cellfectin in unsupplemented Grace's Insect Cell medium (Invitrogen). Negative selection was immediately begun with 0.1 mM gancyclovir. After amplification, viral stocks were used to infect Sf9 cells in SF-900 II SFM (Invitrogen) enriched with FAD (10 μg/mL). Cells were harvested 96 h post-infection. Microsomal fractions were prepared by differential centrifugation as described (Krueger, et al., 2001) and resuspended in storage buffer (10 mM potassium phosphate pH 7.6, 20% glycerol, 1 mM EDTA, and 0.4 mM phenylmethylsulphonyl fluoride). Protein concentrations were determined by the method of Bradford (1976). Flavin content, as a measure of FMO3 specific content, was determined as previously described (Henderson, et al., 2004). In all instances, flavin

measurements were corrected for a background of 0.083 nmol/mg microsomal protein based on the measurement of flavin-content in microsomal preparations from Sf9 cells infected with control vector.

Enzyme Assays: FMO3 catalyzed oxidation reactions were performed in 0.1 M Tris-HCl (pH 8.5), 1 mM EDTA, 0.1 mM NADPH, and 40 to 500 μg/mL Sf9 microsomal protein. Trimethylamine (5 to 1200 μM) and ethylenethiourea (Lancaster) (10 to 100 μM) oxidation were assessed by following the oxidation of NADPH spectrophotometrically at 340 nm ($\varepsilon^M = 6.22 \bullet 10^{-3}$ nmol/mL). The limit of detection for NADPH oxidation was 0.161 nmol/min/nmol FAD. Methimazole (5 to 1000 μM) oxidation was monitored at 412 nm following the method of Dixit and Roche (1984). The limit of detection for the methimazole assay was 0.035 nmol/min/nmol FAD. Sulindac (5 to 400 μM) oxidation was measured using high performance liquid chromatography as described by Hamman *et al*. (2000) without differentiating between the S- and R-sulindac S-oxide. All reactions were incubated for 5 min and the concentration of sulindac S-oxide determined by linear regression using racemic sulindac S-oxide standards ranging between 0.1 and 10 nmol (r^2 = 0.999). Calculated rates were corrected for minimal rates of sulindac auto-oxidation observed in the presence of an equal concentration of non-FMO baculovirus infected Sf9 microsomes.

Data Analysis: Identified SNPs were evaluated for deviation from Hardy-Weinberg equilibrium using a chi-square test. Allelic frequencies for individual sequence variants were compared using Fisher's exact test. A Bonferroni adjustment for the comparisons among the three groups was used, reducing the accepted α value from 0.05 to 0.016. Individual haplotypes and their estimated population frequencies were inferred using PHASE V2.1 (Stephens, et al., 2001; Stephens and Donnelly, 2003). Posterior estimates for unknown phase haplotypes were based on 150,000 iterations of the Gibbs sampler implemented in PHASE. A total of 5,000 samples were discarded

as a burn-in. Population recombination parameter estimates were based on 3000 samples using a thinning interval of 50 on the total iterations used to estimate subject haplotypes. The large thinning interval was to account for high auto-correlation amongst the sampled recombination rates. Geweke's Convergence Diagnostic (Geweke, 1992) was used to monitor for chain convergence. The Bayesian Output Analysis package for the R language was used to monitor the PHASE sample output both for convergence and auto-correlation amongst the posterior samples. Inferred mean haplotype frequencies were compared using a one-way ANOVA with a Holm-Sidak post-hoc test (SigmaStat v3.1, Systat Software).

Analysis of Molecular Variance (AMOVA) was used to analyze the haplotype heterogeneity both within and among the three population samples. AMOVA was implemented using Arlequin V 3.0 (Excoffier, et al., 2005). The variance among populations is analogous to Wright's fixation index (F_{ST}). Bootstrap confidence intervals were also computed based on 20,000 bootstrap replications. Pairwise values were also determined among the three populations to compare them individually, computing empirical *p*-values in the same manner as the overall AMOVA. In addition, haplotype diversity (*H*) was evaluated using the estimated haplotype frequencies computed by PHASE as:

$$\hat{H} = \frac{n}{n-1}\left[1-\sum_{i=1}^{k} p_i^2\right]$$

where n = the number of alleles in the population group and *P* = the frequency of the *i*th haplotype of *k* total observed haplotypes for a given population. Asymptotic Confidence Intervals (95%) are presented for *H* using the sampling variance derived by Nei (1978).

Posterior estimates of population baseline recombination between adjacent SNPs ($\bar{\rho}$) and multiplicative factor parameters (λj) with 5%, 10% and 25% posterior quantiles were calculated as described by Li and Stephens (2003).

Enzyme activity data at different substrate concentrations were fit to a single-site, non-linear Michaelis Menten model [goodness of fit (R^2) > 0.99 in all instances] and kinetic parameters determined using SigmaPlot v9.01 with the Enzyme Kinetics Module v1.2 (Systat Software, Point Richmond, CA). Log transformed kinetic parameters for the different FMO3 variants were compared to the reference by one-way ANOVA with a Holm-Sidak post-hoc test (SigmaStat v3.1, Systat Software). Except as noted above, an α of 0.05 was accepted as a significant difference between data sets. For the Holm-Sidak multiple comparisons tests, significance was indicated by a *P* value less than a critical value that is adjusted downward from 0.05 based on the rank of the value and the number of comparisons made.

RESULTS

Individual *FMO3* SNP frequencies. Extending the previously reported analysis of common *FMO3* upstream variant haplotypes (Koukouritaki, et al., 2005), complete genotyping results for 17 *FMO3* SNPs in three study populations are shown in Table 1 (the g.-2099A>G SNP was not observed in any of the studied populations and as such, is not included in Table 1). None of the SNPs deviated from the proportions predicted by Hardy-Weinberg equilibrium. As was observed for the upstream SNPs, significant differences in the frequency of individual SNPs within the *FMO3* structural gene were observed among the study groups. Of the two novel SNPs predicted to impact protein structure (Koukouritaki, et al., 2005), the g.72G>T (E24D) variant was not observed in the African American or Hispanic (Mexican descent) population groups and only occurred in the non-Latino white group at a frequency of 0.5 %. In contrast, the N61K variant occurred at a frequency of 5.2% in the non-Latino white and 3.5% in the African American population groups, but was not observed in Hispanics (Mexican descent). The g.23613G>T (K416N) variant was only observed in the non-Latino white study group at an allelic frequency of 0.2%. The observed frequencies of the g.15167G>A (E158K), g.18281G>A (V257M), and g.21443A>G (E308G) variants were consistent with a previous report in which a haplotype analysis was restricted to these three SNPs (Cashman, et al., 2001).

The synonymous variant, K167K, was not observed in any of the study populations while the T108T synonymous variant was only observed in the non-Latino white group and only at 0.5%. Yet, the K167K and T108T SNPS were observed in 4/48 and 2/48 alleles tested in the Coriell Polymorphism Discovery Resource samples. This observation suggests these synonymous variants may be present at a significantly higher frequency in one of the other population groups represented

in this resource (*i.e.*, Asian Americans or Native Americans). Three other synonymous SNPs, *i.e.*, g.15146C>T (S147S) and g.21375C>T (N285N) occurred at relatively high frequencies that differed among the population groups examined.

Based on an analysis of posterior estimates of population baseline recombination rates and calculated multiplicative factor parameters with posterior quantiles, no significant recombination was detected within the *FMO3* locus beyond the background population rate at an α level of 0.05 (Fig. 1). However, both the Hispanic and African American samples indicated the possibility of a recombination "hotspot" between *FMO3* position 15,437 and 18,281 and between both 15,437 and 18,281 and 18,281 and 21,375, respectively. In each of these instances, the 25% posterior quantile for λ_j between the pairwise SNPs was larger than 1, implying that the posterior probability of increased recombination is greater than 75% (Fig. 1).

Kinetic Analysis of *FMO3* Structural Variants. The impact of *FMO3* genetic variation on promoter activity was previously reported (Koukouritaki, et al., 2005). To investigate the functional significance of the three novel structural variants, E24D, N61K and K416N, baculovirus-expressed FMO3 was employed using methimazole, ethylenethiourea, trimethylamine and sulindac as substrates (Table 2). Compared to the reference enzyme, the FMO3 E24D variant exhibited a significant increase in K_{cat} with each substrate studied except sulindac. However, K_M values were not significantly different. Thus, the E24D amino acid substitution results in a small increase in catalytic efficiency with most substrates, consistent with a prediction that this variant would have a modest positive effect on FMO3 activity *in vivo*. The differences observed with the FMO3 K416N variant are much more equivocal and appeared to be more substrate-dependent. A significant increase in K_M was observed with trimethylamine, but not the other substrates. However, this change was paired with a comparable increase in K_{cat}, resulting in a small decrease in catalytic

efficiency. A significant decrease in the FMO3 K416N K_{cat} was observed with methimazole, resulting in a nearly 3-fold decrease in catalytic efficiency. The most dramatic changes were observed with the FMO3 N61K variant. A complete loss of activity was observed with both ethylenethiourea and trimethylamine (Table 2). Minimal activity was measured with methimazole and sulindac as substrates, both yielding substantially increased K_M and decreased K_{cat} constants. No methimazole oxidation was observed with non-FMO3 baculovirus infected Sf9 microsomes, eliminating substrate auto-oxidation as a possible explanation for the minimal observed oxidation rates (data not shown). Further, the methimazole oxidation activity observed with the FMO3 N61K variant was 2- to 10-fold higher than the limit of detection at all substrate concentrations tested. Rates of sulindac oxidation were corrected for auto-oxidation that was no more than 10% of the FMO3-dependent oxidation rates. Thus, the N61K variant exhibited a 30- to 40-fold decrease in catalytic efficiency with both substrates.

***FMO3* Haplotype Analysis**. Based on the determined allelic frequencies of the upstream and structural variants among the three populations included in this study, *FMO3* haplotypes were inferred using Phase V2.1 (Table 3). Five haplotypes were deduced in the Hispanic study population that occurred at a frequency greater than 5% and accounted for 90% of the variability within this population. Six haplotypes were inferred in the non-Latino white population that exhibited a frequency greater than 5% and accounted for 69% of the variability within this group. Finally, in the African American study population, five haplotypes were deduced that occurred at a frequency equal to or greater than 5% and accounted for 63% of the population variation. AMOVA, which yields a statistic analogous to Wright's fixation index (F_{ST}), indicated the percent variance at the *FMO3* locus among the populations studied was 0.0495% (bootstrap 95% confidence interval, 0.02518,0.07790), suggesting minimal difference in diversity. A calculation of *FMO3* haplotype

diversity was consistent with these results, but did reveal the highest level of diversity in African Americans, followed by non-Latino whites and Hispanics of Mexican descent (Table 3). The 95% Asymptotic Confidence Intervals calculated to provide an indication of the variance in haplotype diversity within each population should be interpreted with some caution, however, as they neglect the inherent variance in estimating unknown phase haplotypes.

Significant differences in haplotype frequencies were observed among the three study populations. Of particular interest were the differences observed in haplotypes involving SNPs known to have a functional impact on either gene expression or enzyme activity. The H2A and H2B haplotypes, consisting of the high activity promoter variant (g.-2650C>G, g.-2543T>A, and g.-2177G>C) (Koukouritaki, et al., 2005) and either two synonymous variants, S147S and N285N (H2A), or a single synonymous variant, S147S (H2B), occurred at a substantially higher frequency in the Hispanic (Mexican descent) population (26.6%) versus either the non-Latino white or African American groups (4.6% and 4.5%, respectively).

The E158K and E308G variants are reported to have altered kinetic constants for the oxidation of several substrates consistent with a 40 to 70% loss of activity. Moreover, the compound variant has been reported to have an even more substantial, 2- to 4-fold loss of enzyme activity (reviewed in Koukouritaki and Hines, 2005). The inferred H3 haplotype frequency, consisting of the single SNP, g.+15,167G>A encoding the E158K variant, was highest in the Hispanic population (22.5%), but was common in all three population groups (Table 3). In contrast, the *FMO3* H5B haplotype that includes both the E158K and E308G variants exhibited inferred frequencies of 5.4% and 7.9% in the Hispanic and non-Latino white population groups, respectively, but less than 5% in the African American group. Yet, the *FMO3* H6 haplotype, consisting of the E158K and the N285N variants, exhibited an inferred frequency of 14.6% in the African American

study population, but less than 5% in Hispanics and was not inferred to be present in non-Latino whites. Interestingly, the sums of the deduced population frequencies for H3, H5B, and H6 that were ≥5% are similar; 27.9% in Hispanics of Mexican descent, 22.7% in non-Latino whites, and 23.7% in African Americans.

The previously described *FMO3* promoter variants (Koukouritaki, et al., 2005) may mask or exacerbate the effect of the structural variants (Table 4). The high activity promoter variant (previously reported as haplotype 2; -2650C>G, -2543T>A, -2177G>C) was inferred in 8 haplotypes in the Hispanic (Mexican descent) population and 14 haplotypes in the non-Latino white population with a total frequency of 29.0% and 6.7%, respectively. Two of these haplotypes in the Hispanic group and four in the non-Latino white group also contain at least one of the reduced activity structural variants (E158K and/or E308G), but occur at an inferred frequency of 1% or less in both groups. In contrast, the high activity promoter variant was inferred in 21 African American haplotypes with a total frequency of 11.7%. Importantly, 13 of these haplotypes also contain at least one of the reduced activity structural variants (E158K and/or E308G) and had a total inferred frequency of 5.6%.

Two low activity promoter haplotype variants were previously identified, referred to as haplotype 8 (-2589C>T and -2106G>A) and 15 (-2106G>A and -1961T>C), both of which occurred at inferred frequencies <5% (Koukouritaki, et al., 2005). Not surprisingly, neither one of these haplotypes were inferred to occur at frequencies >5% in the current study. However, given their predicted impact on *FMO3* expression, a more important measure of their significance is the total frequency of these SNP combinations in all inferred haplotypes. In the Hispanic (Mexican descent) and non-Latino white population groups, only the -2589C>T, -2106G>A SNP combination was inferred in a total of four and six haplotypes and at frequencies of <1% and 4%, respectively. Both

low activity promoter variants were inferred in the African American population group; the -2589C>T, -2106G>A in 3 haplotypes (total frequency <1%) and the -2106G>A, -1961T>C in 7 haplotypes (total frequency of 4%). Three of the latter haplotypes also contained one of the reduced activity structural variants, *i.e.*, E158K, E308G or N61K.

Neither the FMO3 E24D or N61K variants were inferred in any of the common *FMO3* haplotypes. However, given the near complete loss of catalytic activity observed with the N61K variant, the individual SNP frequency for this variant is the most important measure of its functional impact. Although absent in the Hispanic population study group, the N61K variant was inferred in 25 non-Latino white haplotypes, total frequency of 5.2%, 12 of which (total frequency of 3.0%) also involve one or more of the FMO3 reduced activity variants, E158K and/or E308G. Similarly, the N61K variant was inferred in 17 haplotypes in the African American population group at a total frequency of 3.5%. Nine of these 17 haplotypes also contain one or more of the E158K and/or E308G structural variants at an inferred frequency of 2.2%.

DISCUSSION

Genetic variation in drug metabolizing enzymes are well known determinants of interindividual differences in drug and toxicant disposition and response. In many cases, the differences are substantial with important clinical implications (Evans and McLeod, 2003). Underlying knowledge of such genetic variation and their associated functional changes is increasingly considered critical for drug discovery and development (Roses, 2004) and for the design of effective therapeutic regimens (*e.g.*, Hisamuddin, et al., 2005). Few previous studies have attempted to examine *FMO3* haplotype structure and its impact on overall gene function. The results of an *FMO3* SNP discovery effort that included all exons, splice sites and approximately 1 kbp of upstream information was recently reported (Koukouritaki, et al., 2005). This effort also examined the potential functional significance of upstream haplotypes identified in three study populations. The current study extended these findings, testing the functional impact of two novel structural variants predicted to effect secondary structure, but perhaps more importantly, integrating the previous and current data into an overall *FMO3* haplotype structure that allows a better prediction of genotype/phenotype relationships.

Employing expressed recombinant FMO3, the kinetic parameters of the FMO3 E24D, N61K, and K416N variants were determined using four different substrates (methimazole, ethylenethiourea, trimethylamine and sulindac). The E24D variant had a modest effect on FMO3 enzyme activity with an increase in K_{cat} and resultant increase in catalytic efficiency with most substrates. The results observed with the FMO3 K416N variant were much more equivocal and were substrate-dependent. In contrast, the N61K variant exhibited a complete loss of catalytic activity towards both trimethylamine and ethylenethiorurea and a 30- to 40-fold decrease in catalytic efficiency with both methimazole and sulindac. These observations are consistent with the previous report by Dolphin

et al. (2000) who reported that substitution of a serine for asparagine at the same position (N61S) abolished FMO3's capacity to catalyze the N-oxidation of trimethylamine. Thus, similar to the FMO3 N61S variant, the N61K variant is predicted to contribute to the incidence of trimethylaminuria, an autosomal recessive disorder that is characterized by a deficiency in FMO3 enzyme activity and an inability of individuals to metabolize malodorous trimethylamine (TMA) to the odorless trimethylamine N-oxide (Ayesh, et al., 1993). Predictably, the N61K variant was not inferred in any of the common *FMO3* haplotypes. However, the individual SNP frequency for this variant is of greater interest, given its functional impact. Not observed in the Hispanic population, the N61K variant occurred at an allelic frequency of 5.2% and 3.5% and was inferred in 25 and 17 haplotypes in the non-Latino white and African American populations, respectively. Interestingly, 12 of the non-Latino white and 9 of the African American haplotypes also involved at least one of the previously characterized, hypomorphic structural variants, g.15167G>A (E158K) and g.21443A>G (E308G), which would, if anything, exacerbate the loss of activity. In such individuals, genotype/phenotype correlations based on the presence or absence of the E158K and/or E308G variants alone also would over-estimated the contribution of these latter SNPs to the phenotype.

Eswaramoorthy *et al.* (2006) recently reported the crystal structure of *Schizosaccharomyces pombe* FMO as enzyme-FAD, enzyme-FAD-NADPH, and enzyme-FAD-methimazole complexes. Alignment of human FMO3 with the *S. pombe* FMO amino acid sequence revealed an overall identify of only 21%, but conservation of the asparagine residue wherein the human FMO3 N61 position is analogous to yeast N91. Indeed, an alignment of 56 reported eukaryotic FMO sequences from 19 species, including *Schizosaccharomyces pombe* (Eswaramoorthy, et al., 2006), *Sachromyces cerivisiae* (AY357358), *Caenorhabditis elegans* (AJ582070, AJ582071, AJ582072,

AJ581300, AJ582073, NM_069567, NM_069571, NM_066955, NM_073969, NM_070951), *Drosophila melanogaster* (NM_138015, NM_136373), *Crassostrea gigas* (AJ585074), *Danio rerio* (NM_198910), *Xenopus tropicalis* (NM_001030424), *Galus galus* (AJ431490), *Canis familiaris* (AF384053, XM_537197, AF384054), *Mus musculus* (NM_010231, NM_018881, NM_008030, NM_144878, NM_010232), *Rattus norvegicus* (NM_012792, NM_144737, NM_053433, NM_144561, NM_144739), *Oryctolagus cuniculus* (M32030, M32029, L10391, L10392, L08449), *Cavia porcellus*, (L10037, L37081), *Sus Scrofa* (NM_214064), *Bos taurus* (XM_866868, XM_580516, NM_174057, XM_586631, XM_598393), *Macaca mulatta* (U59453, AY063498), *Pongo pygmaeus* (CR860815, CR857503), *Pan troglodytes* (XM_513039, NM_001009008, NM_001009092), and *Homo sapiens* (NM_002021, NM_001460, NM_006894, NM_002022, NM_001461) revealed 100% conservation of this asparagine, consistent with the data from the *S. pombe* FMO crystal structures that this asparagine moiety is the only amino acid residue directly involved in catalysis (Eswaramoorthy, et al., 2006). The glutamate at human FMO3 position 24 is not as highly conserved as asparagine position 61, being observed in the FMO from all of the above vertebrate species, but not in most of the non-vertebrate FMO. Further, the impact of the E24D variant on FMO3 activity was less than predicted in an earlier report (Koukouritaki, et al., 2005). Human FMO3 glutamate position 24 residue aligns with *S. pombe* glutamate position 28, which is involved in a turn between helix 1 and sheet 2, but structurally, is positioned well away from both the co-factor and substrate binding sites (Eswaramoorthy, et al., 2006). Thus, its modest but consistent effect of increasing the FMO3 k_{cat} is likely due to minor confirmation changes in the overall protein structure that facilitate catalysis. Somewhat similarly, the lysine at position 416 in human FMO3 also is not that highly conserved among the 56 FMO proteins examined, although it is 100% conserved among mammalian FMO3 enzymes. Aligning the human FMO3 and S. pombe

FMO primary sequence suggests that FMO3 lysine position 416 would fall within the carboxy terminal end of alpha helix 6 (Eswaramoorthy, et al., 2006) and, given the lack of predicted impact on secondary structure (Koukouritaki, et al., 2005), it is not surprising that this amino acid substitution has minimal effects on FMO3 catalytic properties. Although the comparisons between human FMO3 and *S. pombe* FMO are interesting and consistent with the reported functional data, because the overall sequence identity is only 21%, conclusions drawn from these comparisons should be taken cautiously.

An important objective of the current study was to determine whether or not linkage between previously identified hypo- or hypermorphic promoter variants (Koukouritaki, et al., 2005) and any of the *FMO3* structural variants might contribute to discrepancies between *in vitro* and *in vivo* observations that were based on the more common analysis of structural variants alone. In the African American population, the high activity promoter haplotype [previously reported as haplotype 2 (Koukouritaki, et al., 2005)] was inferred to occur in 21 haplotypes with a total frequency of 11.7%. Importantly, thirteen of these haplotypes also contain at least one of the reduced activity structural variants (E158K and/or E308G) and occurred at a total frequency of 5.6%. In such individuals, genotype/phenotype association based solely on the analysis of structural variants would be misleading as one would predict the high activity promoter variant would mask the impact of the structural variants. Similar linkage in the Hispanic and non-Latino white population groups was much less frequent, occurring at less than 1%.

Structural variants previously characterized as causative for reduced activity phenotypes based on *in vitro* studies would be exacerbated by low activity promoter variants. Indeed, in the non-Latino white population, the promoter SNP cluster exhibiting low activity (previously reported as haplotype 8) occurred in six haplotypes with a total frequency of 4%, five of which contained at

least one of the reduced activity structural variants (E158K and/or E308G). In the African American population, the promoter SNP cluster exhibiting low activity previously reported as haplotype 15 occurred in 7 haplotypes with a total frequency of 4%. Two of these haplotypes also contained at least one of the reduced activity structural variants (E158K and/or E308G), whereas one haplotype contained the N61K variant. The absence of low activity promoter variants in the Hispanic (Mexican) study population combined with the absence of the N61K allele, as well as the relative abundance of the high activity haplotype allele, would be consistent with higher mean *FMO3* expression in the Hispanic (Mexican) population.

FMO3 haplotype diversity was highest in African-Americans, followed by non-Latino whites and Hispanics of Mexican descent. The greater genetic diversity observed among the African-Americans is consistent with the findings of other studies (Salisbury, et al., 2003). A possible recombination "hotspot" also was identified with a probability of a recombination rate above the African American and Hispanic population baselines of at least 75%. While these "hotspots" are not statistically significant at the 95% level, they merit consideration given the fact that the current study was not powered to detect small increases in recombination (Li and Stephens, 2003).

In summary, the reported haplotype analysis provides strong evidence that genetic variation in both the *FMO3* promoter and structural gene contributes to the observed interindividual differences in *FMO3* expression. Further, the haplotype analysis strongly suggests that variation in promoter sequences can modulate the effects of previously characterized structural variants and should be taken into account when assessing genotype/phenotype association studies, or even when prescribing therapeutics for which FMO3 is important for disposition, *e.g.*, sulindac (Hamman, et al., 2000), itopride (Mushiroda, et al., 2000) or pyrazoloacridine (Reid, et al., 2004). Although there are minimal differences in overall *FMO3* diversity among the population groups studied, significant

differences in individual haplotype frequencies were observed which would contribute to inter-population differences in the FMO3-dependent metabolism of therapeutics, environmental chemicals and dietary constituents.

References

Ayesh R, Mitchell SC, Zhang A and Smith RL (1993) The fish-odour syndrome: Biochemical, familial and clincal aspects. *Br Med J* **307**:655-657.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248-254.

Cashman JR, Zhang J, Leushner J and Braun A (2001) Population distribution of human flavin-containing monooxygenase form 3: Gene polymorphisms. *Drug Metab Dispos* **29**:1629-1637.

Cherrington NJ, Cao Y, Cherrington JW, Rose RL and Hodgson E (1998) Physiological factors affecting protein expression of flavin-containing monooxygenases 1,3, and 5. *Xenobiotica* **28**:673-682.

Dixit A and Roche TE (1984) Spectrophotometric assay of the flavin-containing monooxygenase and changes in its activity in female mouse liver with nutritional and diurnal conditions. *Arch Biochem Biophys* **233**:50-63.

Dolphin CT, Janmohamed A, Smith RL, Shephard EA and Phillips IR (2000) Compound heterozygosity for missense mutations in the flavin-containing monooxygenase 3 (FMO3) gene in patients with fish-odour syndrome. *Pharmacogenetics* **10**:799-807.

Eswaramoorthy S, Bonanno JB, Burley SK and Swaminathan S (2006) Mechanism of action of a flavin-containing monooxygenase. *Proc Natl Acad Sci U S A* **103**:9832-9837.

Evans WE and McLeod HL (2003) Drug Therapy: Pharmacogenomics - Drug disposition, drug targets, and side effects. *N Engl J Med* **348**:538-549.

Excoffier L, Laval G and Schneider S (2005) Arlequin version 3.0: An integrated software package for population genetics data analysis. *Evol Bioinformatics Online* **1**:47-50.

Geweke J (1992) Evaluating the accuracy of sampling-based approaches to calculating posterior moments, in *Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting* (Bernardo JM, Berger JO, Dawid AP and Smith AFM eds) Oxford University Press, Oxford.

Hamman MA, Haehner-Daniels BD, Wrighton SA, Rettie AE and Hall SD (2000) Stereoselective sulfoxidation of sulindac sulfide by flavin-containing monooxygenases. *Biochem Pharmacol* **60**:7-17.

Henderson MC, Krueger SK, Siddens LK, Stevens JF and Williams DE (2004) S-Oxygenation of the thioether organophosphate insecticides phorate and disulfoton by human lung flavin-containing monooxygenase 2. *Biochem Pharmacol* **68**:959-967.

Hernandez D, Janmohamed A, Chandan P, Phillips IR and Shephard EA (2004) Organization and evolution of the flavin-containing monooxygenase genes of human and mouse: identification of novel gene and pseudogene clusters. *Pharmacogenetics* **14**:117-130.

Hines RN (2006) Developmental and tissue-specific expression of human flavin-containing monooxygenase 1 and 3. *Expert Opinion in Drug Metabolism andToxicology* **2**:41-49.

Hisamuddin IM, Wehbi MA, Schmotzer B, Easley KA, Hylind LM, Giardiello FM and Yang VW (2005) Genetic polymorphisms of flavin monooxygenase 3 in sulindac-induced regression of colorectal adenomas in familial adenomatous polyposis. *Cancer Epidemiol Biomarkers Prev* **14**:2366-2369.

Hukkanen J, Dempsey D, Jacob P, III and Benowitz NL (2005) Effect of pregnancy on a measure of FMO3 activity. *Br J Clin Pharmacol* **60**:224-226.

Koukouritaki SB and Hines RN (2005) Flavin-containing monooxygenase genetic polymorphism: impact on chemical metabolism and drug development. *Pharmacogenomics* **6**:807-822.

Koukouritaki SB, Poch MT, Cabacungan ET, McCarver DG and Hines RN (2005) Discovery of Novel Flavin-Containing Monooxygenase 3 (*FMO3*) Single Nucleotide Polymorphisms and Functional Analysis of Upstream Haplotype Variants. *Mol Pharmacol* **68**:383-392.

Koukouritaki SB, Simpson P, Yeung CK, Rettie AE and Hines RN (2002) Human hepatic flavin-containing monooxygenase 1 (*FMO1*) and 3 (*FMO3*) developmental expression. *Pediatric Res* **51**:236-243.

Krueger SK and Williams DE (2005) Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacol Ther* **106**:357-387.

Krueger SK, Yueh M-F, Martin SR, Pereira CB and Williams DE (2001) Characterization of expressed full-length and truncated FMO2 from Rhesus monkey. *Drug Metab Dispos* **29**:693-700.

Li N and Stephens M (2003) Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**:2213-2233.

Mushiroda T, Douya R, Takahara E and Nagata O (2000) The involvement of flavin-containing monooxygenase but not CYP3A4 in metabolism of itopride hydrochloride, a gastroprokinetic agent: Comparison with cisapride and mosapride citrate. *Drug Metab Dispos* **28**:1231-1237.

Nei M (1978) Estimation of average heterozygoxity and genetic distance from a small number of individuals. *Genetics* **89**:583-590.

Overby LH, Carver GC and Philpot RM (1997) Quantitation and kinetic properties of hepatic microsomal and recombinant flavin-containing monooxygenases 3 and 5 from humans. *Chem Biol Interact* **106**:29-45.

Park C-S, Kang J-H, Chung W-G, Yi H-G, Pie J-E, Park D-K, Hines RN, McCarver DG and Cha Y-N (2002) Ethnic differences in allelic frequency for two flavin-containing monooxygenase 3 (*FMO3*) polymorphisms: Linkage and effects on in vivo and in vitro FMO activities. *Pharmacogenetics* **12**:77-80.

Reid JM, Walker DL, Miller JK, Benson LM, Tomlinson AJ, Naylor S, Blajeski AL, LoRusso PM and Ames MM (2004) The metabolism of pyrazoloacridine (NSC 366140) by cytochromes P450 and flavin monooxygenase in human liver microsomes. *Clin Cancer Res* **10**:1471-1480.

Roses AD (2004) Pharmacogenetics and drug development: the path to safer and more effective drugs. *Nat Rev Genet* **5**:645-656.

Salisbury BA, Pungliya M, Choi JY, Jiang R, Sun XJ and Stephens JC (2003) SNP and haplotype variation in the human genome. *Mutat Res* **526**:53-61.

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.

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.

Stephens M, Smith NJ and Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**:978-989.

Störmer E, Roots I and Brockmöller J (2000) Benzydamine *N*-oxidation as an index reaction reflecting FMO activity in human liver microsomes and impact of *FMO3* polymorphisms on enzyme activity. *Br J Clin Pharmacol* **50**:553-561.

Tijet N, Boutros PC, Moffat ID, Okey AB, Tuomisto J and Pohjanvirta R (2006) Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. *Mol Pharmacol* **69**:140-153.

Yeung CK, Lang DH, Thummel KE and Rettie AE (2000) Immunoquantitation of FMO1 in human liver, kidney, and intestine. *Drug Metab Dispos* **28**:1107-1111.

Zhang AQ, Mitchell SC and Smith RL (1996) Exacerbation of symptoms of fish-odour syndrome during menstruation. *Lancet* **348**:1740-1741.

Zschocke J, Kohlmüeller D, Quak E, Meissner T, Hoffmann GF and Mayatepek E (1999) Mild trimethylaminuria caused by common variants in *FMO3* gene. *Lancet* **354**:834-835.

Footnotes

This study was supported in part by PHS grants CA53106 (R.N.H.) and HL38650 (D.E.W.) and with funds from the Children's Research Institute, Children Hospital and Health Systems.

Figure Legends

Figure 1. *Identification of possible FMO3 recombination 'hotspots'.* Posterior estimates of population-specific baseline recombination rates ($\bar{\rho}$) and multiplicative factor parameters (λj) between adjacent SNPs were determined with 5% (●), 10% (○) and 25% (▼) posterior quantiles (PQ) for the (A) non-Latino white, (B) Hispanic of Mexican descent, and (C) African-American populations (Li and Stephens, 2003). Numbers along the ordinate are indicative of the space between adjacent SNP pairs, beginning with g.-2650C>G and g.-2589C>T and proceeding in a 5' to 3' direction along the *FMO3* gene (see Table 1).

TABLE 1. Individual *FMO3* SNP Frequency in Three Study Populations[a]

Position	Location	Amino Acid Change	Variant Allelic Frequency (95% CI) Hispanic (Mexican) (n = 404)	Non-Latino White (n=400)	African American (n=402)
-2650C>G	Upstream		0.450 (0.402,0.499)[bc]	0.301 (0.256,0.346)[d]	0.259 (0.216,0.302)[d]
-2589C>T	Upstream		0.007 (0.000,0.016)[c]	0.040 (0.021,0.059)[bd]	0.007 (0.000,0.016)[c]
-2543T>A	Upstream		0.290 (0.245,0.334)[bc]	0.204 (0.165,0.243)[d]	0.192 (0.153,0.230)[d]
-2177G>C	Upstream		0.290 (0.245,0.334)[bc]	0.067 (0.043,0.092)[bd]	0.124 (0.092,0.157)[cd]
-2106G>A	Upstream		0.007 (0.000,0.016)[c]	0.040 (0.021,0.059)[d]	0.027 (0.011,0.043)
-1961T>C	Upstream		0.005 (0.000,0.012)	0.005 (0.000,0.012)	0.020 (0.006, 0.034)
-1732G>T	Exon 1		0.000	0.007 (0.000,0.016)	0.000
72G>T	Exon 2	E24D	0.000	0.005 (0.000,0.012)	0.000
11177C>A	Exon 3	N61K	0.000	0.052 (0.030,0.074)[d]	0.035 (0.017,0.053)[d]
15019A>C	Exon 4	T108T	0.000	0.005 (0.000,0.012)	0.000
15136C>T	Exon 4	S147S	0.285 (0.241,0.329)[bc]	0.070 (0.045,0.095)[d]	0.102 (0.072, 0.132)[d]
15167G>A	Exon 4	E158K	0.304 (0.259,0.349)[bc]	0.400 (0.353,0.448)[d]	0.418 (0.370,0.466)[d]
15437A>G	Exon 5	K167K	0.000	0.000	0.000
18281G>A	Exon 6	V257M	0.097 (0.068,0.125)[b]	0.067 (0.043,0.092)	0.065 (0.041, 0.089)[d]
21375C>T	Exon 7	N285N	0.381 (0.334,0.429)[bc]	0.259 (0.216,0.302)[bd]	0.475 (0.426,0.524)[cd]
21443A>G	Exon 7	E308G	0.067 (0.042,0.091)[c]	0.157 (0.121,0.192)[d]	0.052 (0.030,0.074)
23613G>T	Exon 8	K416N	0.000	0.002 (0.000,0.007)	0.000

[a] The SNPs at position -2650, -2589, -2543, -2177, -2106, and -1961 were previously reported (Koukouritaki, et al., 2005).

[b] Different from African American study population, $P < 0.016$ (Fisher's exact test, Bonferroni adjustment)

[c] Different from Non-Latino White study population, $P < 0.016$ (Fisher's exact test, Bonferroni adjustment)

[d] Different from Hispanic study population, $P < 0.016$ (Fisher's exact test, Bonferroni adjustment)

JPET Fast Forward. Published on October 18, 2006 as DOI: 10.1124/jpet.106.112268
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 2. Kinetic Analysis of FMO3 Structural Variants

Substrate	FMO3 Enzyme	K_M (μM)	K_{cat} (nmol/min/nmol FAD)	K_{cat}/K_M
Methimazole	Reference	71.8 ± 8.1	123.8 ± 9.1	1.72
	E24D	85.4 ± 9.9	219.8 ± 13.4[a]	2.57
	N61K	123.3 ± 29.1[a]	4.3 ± 0.4[a]	0.04
	K416N	77.3 ± 3.6	46.8 ± 1.7[a]	0.61
Ethylenethiourea	Reference	55.7 ± 12.1	36.3 ± 4.2	0.65
	E24D	76.1 ± 12.3	50.1 ± 4.3[a]	0.66
	N61K	NA	NA	---
	K416N	50.2 ± 6.0	42.2 ± 2.0	0.84
Trimethylamine	Reference	20.9 ± 4.8	33.7 ± 2.2	1.61
	E24D	22.8 ± 4.0	64.9 ± 2.0[a]	2.85
	N61K	NA	NA	---
	K416N	37.3 ± 7.0[a]	51.0 ± 3.4[a]	1.37
Sulindac	Reference	150.1 ± 39.8	46.8 ± 10.9	0.31
	E24D	120.8 ± 0.1	43.6 ± 1.4	0.36
	N61K	207.1 ± 93.5[a]	2.9 ± 2.8[a]	0.01
	K416N	66.5 ± 25.9	68.4 ± 10.9	1.03

[a] Significantly different than reference enzyme (calculated *P* value less than critical level, one-way ANOVA with Holm-Sidak post-hoc)

[b] NA ≡ No activity

TABLE 3. Common (>5%) *FMO3* Haplotypes in Three Study Populations

Haplotype	Nucleotide Changes	Mean % Frequency ± SEM[a]		
		Hispanic (Mexican) (n=404)	Non-Latino White (n=400)	African American (n=402)
H1	Reference (AL021026)	26.9 ± 0.4	32.0 ± 0.9	20.3 ± 1.2
H2A	-2650C>G, -2543T>A, -2177G>C, 15136C>T (S147S), 21375C>T (N285N)	26.6 ± 0.2	4.6 ± 0.4	1.0 ± 0.4
H2B	-2650C>G, -2543T>A, -2177G>C, 15136C>T (S147S)	1.1 ± 0.1	0.5 ± 0.1	4.5 ± 0.6
H3	15167G>A (E158K)	22.5 ± 0.3	14.8 ± 0.6	9.1 ± 1.2
H4B	-2650C>G, 18281G>A (V257M)	8.8 ± 0.5	5.0 ± 0.3	2.1 ± 0.2
H5B	-2650C>G, 15167G>A (E158K), 21375C>T (N285N), 21443A>G (E308G)	5.4 ± 0.0	7.9 ± 0.8	1.1 ± 0.2
H1E	-2543T>A	NI	5.0 ± 0.8	1.0 ± 0.4
H6	15167G>A (E158K), 21375C>T (N285N)	0.7 ± 0.2	NI	14.6 ± 1.1
H7	21375C>T (N285N)	2.6 ± 0.5	0.9 ± 0.2	14.6 ± 1.2
	SUM	94.6	70.7	71.3

Haplotype Diversity (*H*) (95% CI)	0.796 (0.795,0.797)	0.856 (0.855,0.858)	0.892 (0.891,0.893)

[a] Inferred haplotypes frequencies for each population group were significantly different from other population groups, $P < 0.001$ (ANOVA with Holm-Sidak post hoc test). NI indicates that this haplotype was not inferred to be present in this population group.

JPET Fast Forward. Published on October 18, 2006 as DOI: 10.1124/jpet.106.112268
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 1

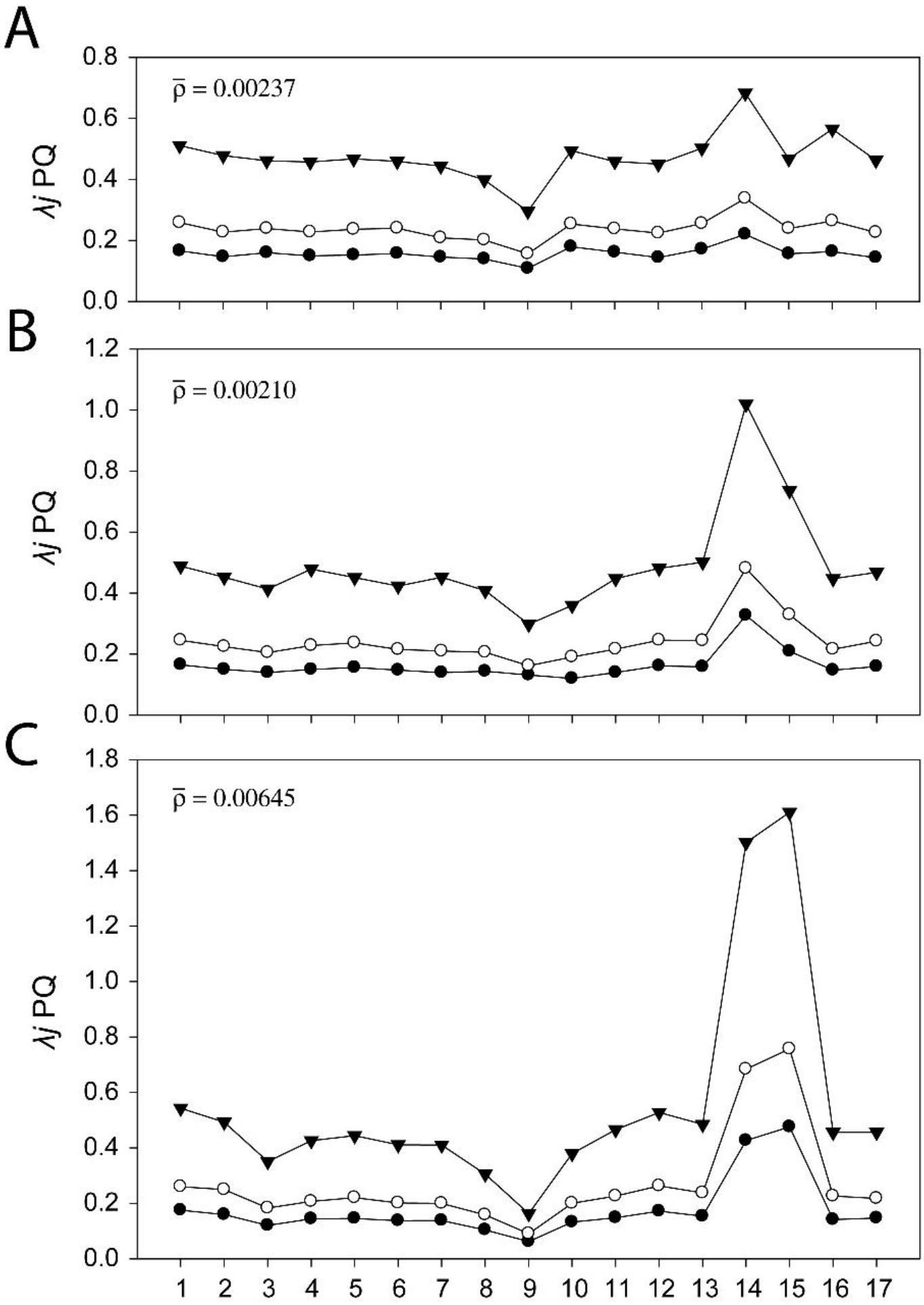

Table S1. *FMO3* Amplification Primers for SBE

Primer Set	Primer Sequence	FMO3 Coordinates[a]	Target
1	5' TTGGCGTGTGCTCACCCATTC 3' 5' GTTGGTCAGGCTGGTCTCAAG 3'	-1869 to -1849 -1618 to -1638	Exon 1
2	5' GGCGTGAGCTACCATACTCAG 3' 5' GGAGAATTTCAAAGCCCATTGTC 3'	-151 to -131 -319 to -341	Exon 2
3	5' CCAGCCCTGACCATGATCAG 3' 5' GAGTTAGAGAGGCCACACTG 3'	+11060 to +11079 +11449 to +11430	Exon 3
4	5' CACACCAACAGGCAGCCATTTAG 3' 5' CCTCCCCCACATTTCATATCACAC 3'	+14451 to +14473 +15651 to +15628	Exons 4-5
5	5' TGGGATGGTCCTTTGGCATTG 3' 5' CCAGCAGGCATATCACGTTTAG 3'	+17847 to +17867 +18372 to +18351	Exon 6
6	5' ACACCGAGGTACCCTCACTTCAG 3' 5' GGCTGGATGTTGGCTCGTAAAC 3'	+21023 to +21045 +22105 to +22084	Exon 7
7	5' GTGATGCCAATTCACCTCTG 3' 5' AGAGAATTTATGGGCTAGAG 3'	+23238 to +23257 +23989 to +23970	Exon 8

[a] Coordinates are based on assigning the "A" of the ATG translation initiation codon +1 (den Dunnen JT and Antonarakis SE (2001) Nomenclature for the description of human sequence variations. *Hum.Genet.* **109**:121-124.

Table S2. *FMO3* Single Base Extension Primers

Primer	SNP Queried	Quantity in Assay (pmol)
Group 1		
5' $(C)_{28}$TTGGCCTCCATCAGGAGCTGTCTGGAAGA 3'	72G>T	1.0
5' $(C)_{23}$TAGCATTTACAAATCAGTCTTTTCCAA 3'	11,177C>A	1.5
5' $(C)_{18}$GATGTTTATTTACACTGGATACAAA 3'	15,019A>C	0.5
5' $(C)_{11}$TCTTTGATGCTGTAATGGTTTGTTC 3'	15,136C>T	7.0
5' $(C)_{6}$ATGTGTATCCCAACCTACCAAAA 3'	15,167G>A	7.0
5' $(C)_{3}$TGCTGTGGAAGCATTTGCC 3'	15,437A>G	10.0
Group 2		
5' AACTGCCCAGACGGTTGGACA 3'	-1732G>T	3.0
5' $(C)_{5}$GAATCTTGCATTCATCTGCTTCA 3'	18,281G>A	5.0
5' $(C)_{13}$TGAGGAAAGAGCCTGTATTTAA 3'	21,375C>T	3.0
5' $(C)_{16}$TCCCATCCTCAAAAATGGCCGAGGTC 3'	21,443A>G	0.8
5' $(C)_{21}$ATGATATTAATGAGAAAATGGAGAAAAA 3'	23,613G>T	10.0