Published in final edited form as:

Pharmacogenomics. 2012 April ; 13(5): 555–570. doi:10.2217/pgs.11.160.

Worldwide variation in human drug-metabolism enzyme genes *CYP2B6* and *UGT2B7*: implications for HIV/AIDS treatment

Jing Li¹, Vincent Menard², Rebekah L Benish³, Richard J Jurevic⁴, Chantal Guillemette², Mark Stoneking¹, Peter A Zimmerman³, and Rajeev K Mehlotra^{*,3}

¹Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

²Pharmacogenomics Laboratory, CHUQ Research Center & Faculty of Pharmacy, Laval University, QC, Canada

³Center for Global Health & Diseases, Case Western Reserve University School of Medicine, Cleveland, OH, USA

⁴Department of Biological Sciences, Case Western Reserve University School of Dental Medicine, Cleveland, OH, USA

Abstract

Aim—Hepatic enzymes, *CYP2B6* and *UGT2B7* play a major role in the metabolism of the widely used antiretroviral drugs efavirenz, nevirapine and zidovudine. In the present study, we provide a view of *UGT2B7* haplotype structure, and quantify the genetic diversity and differentiation at both *CYP2B6* and *UGT2B7* genes on a worldwide scale.

Materials & methods—We genotyped one intronic and three promoter SNPs, and together with three nonsynonymous SNPs, inferred *UGT2B7* alleles in north American (n = 326), west African (n = 133) and Papua New Guinean (n = 142) populations. We also included genotype data for five *CYP2B6* and six *UGT2B7* SNPs from an additional 12 worldwide populations (n = 629) analyzed in the 1000 Genomes Project.

Results—We observed significant differences in certain SNP and allele frequencies of *CYP2B6* and *UGT2B7* among worldwide populations. Diversity values were higher for *UGT2B7* than for *CYP2B6*, although there was more diversity between populations for *CYP2B6*. For both genes, most of the genetic variation was observed among individuals within populations, with the Papua New Guinean population showing the highest pairwise differentiation values for *CYP2B6*, and the Asian and European populations showing higher pairwise differentiation values for *UGT2B7*.

Conclusion—These new genetic distinctions provide additional insights for investigating differences in antiretroviral pharmacokinetics and therapy outcomes among ethnically and geographically diverse populations.

© 2012 Future Medicine Ltd

* Author for correspondence: Tel.: +1 216 368 6172 Fax: +1 216 368 4825 rkm@case.edu.

Financial & competing interests disclosure The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Keywords

1000 Genomes Project; CYP2B6; efavirenz; HIV/AIDS; nevirapine; Papua New Guinea; UGT2B7; zidovudine

The introduction of highly active antiretroviral therapy (HAART) as a standard of care has markedly decreased AIDS incidence and improved HIV/AIDS prognosis [1]. However, a number of clinical factors, including regimen, adherence and stage of infection/disease at the initiation of therapy, are known to influence outcomes of HAART [2–4]. In addition, host-related genetic factors, including polymorphisms in antiretroviral drug-metabolism enzyme and transporter genes, may influence outcomes of HAART [5,6].

Among a variety of antiretroviral drugs that are included in HAART, the non-nucleoside reverse transcriptase inhibitors efavirenz (EFV) and nevirapine (NVP) and, the nucleoside reverse transcriptase inhibitor zidovudine (AZT) are widely used, particularly in resource-limited regions because of cost considerations. For example, NVP is widely used in Africa, especially to prevent mother-to-child transmission [101]. These drugs have narrow therapeutic windows [7–12], but show large interindividual variability in their pharmacokinetics [13,14], contributing to the variable outcomes seen among HIV-infected patients. Clinical studies have suggested that subtherapeutic plasma concentrations are associated with treatment failure, whereas supratherapeutic concentrations are associated with toxicities of these drugs [7–12]. Repeated exposure to subtherapeutic concentrations may also facilitate the emergence of antiretroviral drug resistance [9].

The metabolism of EFV, NVP and AZT is relatively well defined. The hepatic enzyme CYP2B6 plays a major role in the hydroxylation of EFV and NVP [15,16]. The *CYP2B6* gene (chromosome 19) is highly polymorphic, containing 25 nonsynonymous, several synonymous and numerous promoter SNPs [17]. To date, a total of 29 alleles of *CYP2B6* (*1A [wild-type] to *29) have been described [102], many of which are associated with increased, decreased or abolished enzyme activity [17–19]. In particular, *CYP2B6**6 (516G>T [rs3745274, exon-4, Gln172His] and 785A>G [rs2279343, exon-5, Lys262Arg]) is associated with ‘slow metabolism’ of EFV [19,20]. Hofmann *et al.* demonstrated that 516G>T in *CYP2B6**6 was alone responsible for aberrant splicing, resulting in a high-splice variant and low CYP2B6 expression phenotype [21].

A number of clinical studies conducted in HIV-infected patients of European and African ethnicities, have analyzed the association between *CYP2B6* SNP genotypes, particularly 516G>T, a marker for *6, and/or 983T>C (rs28399499, exon-7, Ile328Thr), and the pharmacokinetics of EFV and NVP [22–29], as well as between the genotypes or pharmacokinetics and outcomes of treatment with these drugs [25,26,30,31]. Recently, similar studies have been conducted in other ethnically diverse patients [32–39]. Taken collectively, the results pertaining to the association between the genotypes and the pharmacokinetics of these drugs are mostly consistent across populations: 516TT and/or 983CC genotypes are associated with up to fivefold higher drug concentrations (~30-fold in one case [40]) compared with other genotypes. However, the results pertaining to the association between the genotypes or pharmacokinetics and outcomes of treatment with these drugs are less clear [25,26,31,37,39], possibly because the *CYP2B6* genotype associations may be cohort-specific [31].

Another hepatic enzyme, UGT2B7, catalyzes direct glucuronidation of AZT, EFV and EFV metabolites [41,42]. The *UGT2B7* gene (chromosome 4) is also polymorphic, containing three nonsynonymous, and many synonymous, intronic and promoter SNPs [43–45]. To

date, a total of four alleles of *UGT2B7* (*1a–*4) have been described [103], which may affect *UGT2B7* expression and glucuronidation activity [44–46]. Among these alleles, *UGT2B7**2 (802C>T [rs7439366, exon-2, His268Tyr]) affects the *in vitro* glucuronidation of AZT, but not of EFV [41,42]. To our knowledge, influence of *UGT2B7* SNP genotypes on antiretroviral pharmacokinetics and clinical response was not reported until recently: 802C>T was found to be significantly associated with lower plasma EFV concentrations in HIV-infected African patients in univariate, but not in multivariate, analysis [47]. A synonymous SNP 735A>G (rs28365062, exon-2, Thr245Thr), linked to several intronic SNPs, was found to be associated with faster AZT clearance in African patients coinfecting with HIV–tuberculosis, and with higher AZT glucuronidation *in vitro* [48]. However, another study, conducted in Belgian patients of European and African ethnicities, did not observe a significant effect of 802C>T or 735A>G on the pharmacokinetics of EFV [49]. Recently, in European–American (EA) and African–American (AFA) patients, we did not observe a significant association between 802C>T or the intronic SNP IVS1+985A>G (rs62298861, intron-1) genotypes and the HAART-related time to achieve virologic success [31]. These studies indicate that the effects of *UGT2B7* SNP genotypes could also be cohort-specific [31,47–49].

Despite the fact that EFV, NVP and AZT are widely used antiretroviral drugs, the information regarding worldwide variation in *CYP2B6* is limited [17] and, in *UGT2B7*, is not available. In recent years, the information regarding the prevalence of *CYP2B6* alleles in ethnically diverse populations has become increasingly available [50–54], and two general conclusions can be drawn from these studies: *CYP2B6**6 is the most prevalent mutant allele and there are significant interethnic differences in *CYP2B6* allele frequencies. However, the information regarding worldwide patterns of genetic diversity and differentiation for this gene is still not available. Therefore, it is not known whether the observed differences in the allele frequencies between populations are due to demographic changes or, if they are indicative of the action of local (i.e., geographically restricted) natural selection. Furthermore, the data regarding haplotype structure of *UGT2B7* in ethnically diverse populations are still scarce [44,55–57]. The present study takes the first step towards closing some of these gaps.

Previously, based on six *CYP2B6* nonsynonymous SNPs, we reported the prevalence of alleles *1A–*7, *9 and *18 in six populations (four major north American ethnic groups, west African and Papua New Guinean [PNG]) [53,58]. The same populations were also analyzed for three *UGT2B7* nonsynonymous SNPs [59]. In the present study, our aims were to provide a view of *UGT2B7* haplotype structure and to quantify the genetic diversity and differentiation in both *CYP2B6* and *UGT2B7* genes on a worldwide scale. For these aims, we genotyped one intronic and three promoter SNPs, and together with the nonsynonymous SNPs, inferred *UGT2B7* haplotypes (alleles) in north American, west African and PNG populations; we also included genotype data for five *CYP2B6* and six *UGT2B7* SNPs from an additional 12 worldwide populations analyzed in the 1000 Genomes Project [60]. Our results provide new insights regarding comparative variation in both *CYP2B6* and *UGT2B7* genes, which may help guide investigations into antiretroviral drug metabolism and therapy outcomes.

Materials & methods

North American, west African & PNG populations (study populations)

The study populations and sample collection procedures have been previously described [53,59,61,62]. Blood samples were collected from north American random blood donors, belonging to four major ethnic groups (EA, AFA, Asian–American [ASA] and Hispanic–American [HA]) [62], and from PNGs from the Wosera area, East Sepik Province

[53,59]. DNA samples were collected from west Africans from five countries (Senegal, Guinea, Sierra Leone, Ivory Coast and Ghana) [61]. All samples were collected under protocols including the procedures for informed consent approved by the corresponding institutional review boards. The numbers of samples analyzed in this study ($n = 607$ [*CYP2B6*] and $n = 601$ [*UGT2B7*]) are provided in supplementary table s1 (see online at: www.futuremedicine.com/doi/suppl/10.2217/pgs.11.160). The ethnicity of the north American individuals was self-identified. All west African samples were collected from the West Atlantic subgroup of the Niger-Congo language family, and were analyzed as one population. The PNG samples were collected from individuals speaking Abelam, which belongs to the Sepik language family. Genomic DNA was extracted from the north American and PNG blood samples using the QIAamp® 96 DNA blood kit (Qiagen) [53,59].

PCR & SNP genotyping

The aforementioned study populations have previously been genotyped for six *CYP2B6* (64C>T [rs8192709, exon-1, Arg22Cys], 516G>T, 777C>A [rs45482602, exon-5, Ser259Arg], 785A>G, 983T>C, and 1459C>T [rs3211371, exon-9, Arg487Cys]) [53,58] and three *UGT2B7* (211G>T [rs12233719, exon-1, Ala71Ser], 802C>T and 1192G>A [no rs number, exon-5, Asp398Asn]) [59] SNPs using the oligonucleotide ligation detection reaction-fluorescent microsphere assay on the Bio-Plex™ suspension array system (Bio-Rad Laboratories). In the present study, these populations were genotyped for the *UGT2B7* intronic SNP IVS1+985A>G, which is a haplotype-tagging SNP [44], and promoter SNPs -327G>A (rs7662029), -138G>A (rs73823859) and -125T>C (rs7668282) using the same ligation detection reaction-fluorescent microsphere assay. The method for genotyping IVS1+985A>G has been recently published [31]. The primers and conditions to selectively amplify the *UGT2B7* promoter region (574 bp) are described in supplementary table s2. The ligase detection reaction primers and conditions to genotype the promoter SNPs, and the mean and 95% CI of log-transformed fluorescent values corresponding to each of the genotypes are also presented in supplementary table s2.

1000 Genomes Project populations (reference populations)

In addition, we included *CYP2B6* and *UGT2B7* data, pertaining to the SNPs that were genotyped in the study populations, from the 1000 Genomes Project [60]. The genotype calls were based on 629 individuals, belonging to 12 populations, from the 2010–08–04 sequence and alignment release of the project (northern and western European from UT, USA; Toscani from Italy; British from England and Scotland; Finnish from Finland [FIN]; African ancestry from southwest, USA; Yoruba from Ibadan, Nigeria; Luhya from Webuye, Kenya; Han Chinese from Beijing, China; Han Chinese from south China; Japanese from Tokyo, Japan; Mexican ancestry from Los Angeles, CA, USA; and Puerto Rican from Puerto Rico, USA [PUR]). Sample sizes of these globally distributed reference populations are provided in supplementary table s1.

SNP data

The 1000 Genomes Project provides low-coverage (2–4×) sequence data for each sample; at this depth, sequencing cannot provide the complete genotype of each sample, but may allow the detection of most variants with frequencies as low as 1%. Using BEAGLE genetic analysis software package [104], we included high-quality phased genotype data for the five *CYP2B6* (64C>T, 516G>T, 785A>G, 983T>C and 1459C>T) and six *UGT2B7* (-327A>G, -138G>A, -125T>C, 211G>T, IVS1+985A>G and 1192G>A) SNPs from the aforementioned 12 reference populations. The *CYP2B6* 777C>A and *UGT2B7* 802C>T SNP data were not available, possibly because these sites did not pass quality filters.

Statistical analysis

GenePop v4.0.10 [105] was used to calculate *CYP2B6* and *UGT2B7* SNP frequencies and to perform the Hardy–Weinberg exact tests (estimation of p-values were calculated using the Markov chain method). SHEsis online version [106] was used to compute Lewontin's D' and correlation coefficient (r^2) parameters in order to measure linkage disequilibrium (LD) between SNPs. Since the six SNPs included in this study define well-known *CYP2B6* alleles [102], *CYP2B6* alleles in the study populations were identified manually. However, since the newly characterized SNP IVS1+985A>G is not yet part of the *UGT2B7* alleles [103], *UGT2B7* alleles in these populations were inferred using PHASE v2.1.1 [107]. An online 2×2 contingency table for Fisher's exact test [108] was used to calculate differences in the SNP/allele frequencies between populations, and a two-tail value of <0.05 was considered to be significant. DnaSP v5.10 [109] was used to compute haplotype diversity (H), and Arlequin v3.11 [110] was used to compute analysis of molecular variance (AMOVA) and pairwise population differentiation ($PA-F_{st}$). Both H and F_{st} values range from 0 (no diversity, no differentiation) to 1 (every individual has a different haplotype, fixed difference between populations).

Results

SNP frequencies, LD patterns & allele profiles

***CYP2B6* SNP frequencies**—The *CYP2B6* SNP frequencies are presented in table 1. 64C>T had a low prevalence, 1459C>T had a low-to-moderate prevalence, whereas 516G>T and 785A>G were highly prevalent in all populations. Among all of the 18 populations examined, the lowest frequency of 516G>T was observed in the FIN population. The 516G>T and 785A>G SNP frequencies in the PNG population were significantly higher than in any other population. 983T>C was low to moderately prevalent in the populations of African ancestry, and was also present in the PUR (0.05) and HA (0.01) populations. The PUR population arose as a result of admixing among Europeans, west Africans and Native Americans [63]. Overall, our *CYP2B6* SNP frequency results are in agreement with those previously reported for comparable populations [52,54,64,65].

***UGT2B7* SNP frequencies**—The *UGT2B7* SNP frequencies are presented in table 2. –138G>A and –125T>C had a low prevalence, whereas –327A>G and IVS1+985A>G were highly prevalent in all populations. Notably, the frequency of –327A>G in the European group (~0.5), and the frequencies of IVS1+985A>G in the Asian group (0.04–0.11) were lower than in any other group. 802C>T was also highly prevalent in the study populations, with the frequency of 802T (0.53) being significantly higher in the EA population than in any other population. The 802C>T SNP data for the reference populations were not available. However, the frequencies of 802T in northern and western Europeans from UT, USA; Yoruba from Ibadan, Nigeria; Han Chinese from Beijing, China plus Japanese from Tokyo, Japan populations (all are part of the low-coverage panel) have been reported as 0.49, 0.26 and 0.32, respectively [111]. These frequencies are comparable to those in EA, AFA and ASA study populations, respectively (table 2), suggesting worldwide interethnic differences. 211G>T was highly prevalent in the populations of Asian ancestry, and was also present in the FIN (0.04), Luhya from Webuye, Kenya (0.01) and HA (0.02) populations. The origin of the population of Finland has been debatable, and both European and Asian populations seem to have provided a contribution to the Finnish gene pool [66–68]. Luhya from Webuye, Kenya, is an admixed population in which individuals have varying continental ancestry proportions [69]. Overall, our *UGT2B7* SNP frequency results are in agreement with those previously reported for comparable populations [44,56,57,70,71].

Populations in Hardy–Weinberg equilibrium

We calculated expected genotype numbers for each of the *CYP2B6* and *UGT2B7* SNPs in the study and reference populations. For all SNPs analyzed, the expected genotype numbers did not differ significantly from the observed genotype numbers in any of the populations (data not shown). However, a significant deficit of heterozygosity at *CYP2B6* 516G>T ($p = 0.016$) and 1459C>T ($p = 0.003$) was observed in the PNG population, which indicates either nonrandom mating or population substructure.

LD patterns

We quantified the extent of LD among the *CYP2B6* and *UGT2B7* SNP pairs in all populations. Strong LD, defined by high values for both D' (> 0.8) and r^2 (> 0.5) parameters [72], was only observed between SNP pairs *CYP2B6* 516–785 (all populations) and *UGT2B7* –327–802 (study populations; 802C>T SNP data not available for the reference populations) (table 3). The *UGT2B7* SNP pair –327–IVS1+985 had high D' (> 0.8) and moderate r^2 (0.12–0.351) values in all populations except the Asian group. The *UGT2B7* SNP pair –327–138, had high D' (1) and moderate r^2 (0.106–0.259) values in the African group. All other SNP pairs had highly variable D' (0–0.8) and low r^2 (< 0.1) values in all populations (data not shown). *CYP2B6* & *UGT2B7* allele profiles Results regarding *CYP2B6* six-SNP alleles in the study populations have been previously reported [53,58]. Briefly, a total of ten alleles were identified in the study populations, of which the wild-type **1A* and the mutant **6* were the most prevalent alleles in all populations. The frequency of **6* in the PNG population (0.62) was significantly higher than in any north American (0.23–0.34) or west African (0.42) population [53]. *CYP2B6* **18* (983C) was prevalent in the populations of African ancestry (AFA [0.08] and west African [0.05]) [58]. Since we did not have the complete six-SNP frequency data for the reference populations (table 1), we could not perform the allele profile analysis on this set of populations. However, since we had the 516G>T, 785A>G, and 983T>C SNP data for all of the reference populations, we estimated frequencies of haplotypes 516T_785G (presumably **6*), 516G_785A_983C (presumably **18*) and 516G_785G_983C (presumably **16*) (table 4). Overall, the results of this analysis were similar to those for the study populations [53,58]. Taken collectively, these results indicate that among all the populations examined, the frequency of **6* is lowest in the FIN population (0.07) and highest in the PNG population (0.62). In the African group and the PUR population, 983T>C was observed predominantly as **18*. Previously, 983T>C has been found either as **18* in AFA, Ghanaian and other west African populations [52,58], or as **16* in Tanzanian and Turkish populations [73].

We inferred *UGT2B7* six-SNP allele profiles in the study populations using PHASE v2.1.1 (table 5). Assuming random associations among the six polymorphic *UGT2B7* SNPs, one can predict 64 different alleles (2^6). However, only seven alleles ($6 + 1$) would be possible if mutations are the only evolutionary forces acting to create new alleles, and other forces, such as recombination, recurrent mutation and gene conversion, do not occur [74]. We inferred a total of seven alleles in the study populations, which is the theoretical minimum. Of these, three alleles together accounted for the majority of total chromosomes examined: **1a*, **1m*, and **2a* occurred at frequencies > 0.1 in most of the populations. Notably, these were the only alleles present in the PNG population, whereas all other populations had five to six alleles. The new haplotype, tentatively designated as *UGT2B7* **1m*, contains IVS1+985G, together with –327G and 802C, and was highly prevalent (table 5). The frequency of *UGT2B7* **1m* was significantly higher in HA (0.37) than in any other north American population or in the PNG population. *UGT2B7* **2a* (–327A_802T) was also highly prevalent; its frequency was significantly higher in EA (0.49) than in HA, AFA, west African or PNG population. Finally, *UGT2B7* **3* (211T) was prevalent (~ 0.1) in the ASA population, as previously reported for other populations of Asian ancestry [56,57].

Genetic diversity & differentiation

CYP2B6 and *UGT2B7* H values for the various populations are presented in Figure 1. Diversity values were higher for *UGT2B7* (mean H = 0.708) than for *CYP2B6* (mean H = 0.608); with particularly low H-values observed for *CYP2B6* in the FIN (0.408, Han Chinese from Beijing, China (0.344); Han Chinese from south China (0.254); and Japanese from Tokyo, Japan (0.399) populations. Interpopulation diversity varied much more for *CYP2B6* (H = 0.254–0.690) than for *UGT2B7* (H = 0.603–0.737). Comparison of the genetic variation between *CYP2B6* and *UGT2B7* genes in all populations by AMOVA revealed that 90% of the variation was partitioned within populations, with some small but significant variation among groups (table 6), indicating significant between-population differences. These results are largely in agreement with previous findings, including those pertaining to absorption–distribution–metabolism–excretion genes, that the overwhelming majority of human genetic variation is found among individuals within populations [75,76].

The $PA-F_{st}$ values for *CYP2B6* and *UGT2B7* in all populations are presented in supplementary table s3a & s3b, respectively. The mean values for these genes were 0.062 and 0.061, respectively. For *CYP2B6*, the highest values were observed in comparisons involving a PNG population, in keeping with the very different allele-frequency distribution for this population [53]. For *UGT2B7*, the highest values tend to involve the populations with Asian ancestry (mean pairwise differentiation [$PA-F_{st}$] = 0.050–0.089), followed by the values for those with European ancestry (mean $PA-F_{st}$ = 0.042–0.076). We also analyzed these results according to ethnic groups (European, African, Asian, Hispanic and PNG). In this analysis, the PNG population showed the highest differentiation at *CYP2B6*, and the Asian ($PA-F_{st}$ = 0.043–0.068) and European ($PA-F_{st}$ = 0.019–0.080) groups showed higher differentiation at *UGT2B7* (table 7).

Discussion

This study attempted to provide comprehensive comparative information regarding variation in both *CYP2B6* and *UGT2B7* genes on a worldwide scale, by including north American, west African and PNG, as well as 1000 Genomes Project populations. Taken collectively, we observed that:

- There are significant differences in certain SNP and allele frequencies of *CYP2B6* and *UGT2B7* among worldwide populations, which in turn produce differences in H and $PA-F_{st}$;
- Diversity values were higher for *UGT2B7* than for *CYP2B6*, although there was more diversity between populations for *CYP2B6*;
- For both genes, most of the genetic variation was observed among individuals within populations, with the PNG population showing the highest $PA-F_{st}$ values for *CYP2B6*, and the Asian and European populations showing higher $PA-F_{st}$ values for *UGT2B7*.

These observations are meant to help guide antiretroviral drug metabolism and therapy outcome studies that are designed to optimize drug effectiveness, while minimizing toxicity.

Differences in SNP/allele frequencies *CYP2B6*

Our *CYP2B6* SNP/allele frequency data suggest that among the four population groups, European, African, Asian and Hispanic, the slow metabolism phenotype of EFV, associated with *CYP2B6**6, is likely to be more prevalent in the African (frequencies of *6; 0.21–0.41) and Hispanic (0.23–0.32) groups than in the European (0.07–0.31) and Asian (0.1–0.23) groups. Thus, it is conceivable that slower hepatic metabolism and the resulting increased systemic circulation level of EFV are more frequent in African and Hispanic groups than in

European and Asian groups. Based on such allele frequency differences, as well as some recent proof-of-concept and pharmacokinetic model-based studies, it is being advocated that there is a need for prospective clinical studies to evaluate the utility of genotype-driven EFV dose adjustments in populations of diverse ancestry [77]. Another suggested approach is therapeutic drug monitoring to identify improperly dosed individuals within a population [13,78]. Genotype-based prediction and/or therapeutic drug monitoring could possibly lead to EFV dose reduction (an economic advantage) if many individuals in a given population are shown to be exposed to supratherapeutic drug concentrations. Results of a number of recent clinical studies conducted in the populations of Africa, where the frequency of *6 ranges from 0.3 to 0.5, support the use of such an approach [17,77,78]. Our worldwide *6 allele frequency data substantiates the need to consider such an approach in a PNG population.

Recent published estimates suggest that the PNG population has among the highest adult HIV prevalences in the Asia-Pacific region, estimated at 1.28% among people aged 15–49 years in 2007, although more recent estimates suggest national prevalence may be closer to 1.0% [112,113]. Regardless, HIV prevalence is projected to increase up to almost 3% by 2015 [113]. Current antiretroviral therapy coverage is ~75% [113]. EFV is used in first-line antiretroviral therapy as a substitute for NVP, and is preferred for cases of HIV–tuberculosis coinfection [114]. Despite the unique high frequency (0.62) of *6, data on EFV pharmacokinetics, clinical response and toxicity are not available in the PNG population. Therefore, in this population it would be most sensible to pursue genotype-driven EFV clinical studies. In all likelihood, this would permit a decreased overall use of EFV per patient, allowing the treatment of more patients. It would also contribute to a clearer understanding of the effects of high-frequency genetic mutations in the therapeutic setting.

On the other hand, we noted that among all the populations, the frequency of *6 was lowest in the FIN population (0.07). It is anticipated that the area under the 24-h plasma EFV concentration–time curve ($AUC_{0-24\text{ h}}$) may be lower among FIN than among other populations, although a formal pharmacokinetic analysis would be needed to confirm this prediction. Thus, the PNG and FIN populations, compared with the other populations in the present study, would present two different dosing approaches for EFV to be both effective and safe.

The populations of African ancestry, and with African admixture (PUR and HA in this study), may show further interindividual differences in EFV and NVP pharmacokinetics and/or treatment outcomes due to the presence of 983T>C, compared with the populations of non-African ancestry where this SNP is absent. 983T>C, together with 516G>T, was associated with fivefold higher mean plasma EFV concentrations in HIV-infected African patients [73]. Subsequently, a number of studies reported significant associations between 983T>C genotype and lower clearance/higher plasma concentrations of EFV [23] and NVP [28–30], and between 983T>C genotype and NVP-related adverse events [30,79] in HIV-infected African patients.

UGT2B7

Compared with *CYP2B6* polymorphisms, fewer and inconclusive data are available in the literature regarding the effects of *UGT2B7* polymorphisms on AZT/EFV metabolism [41,42], pharmacokinetics [47–49] and clinical response [31]. Here, it is important to consider that the pharmacokinetics and clinical response studies were conducted in patients of European and African ethnicities only, and it is possible that the outcome is cohort-specific. Therefore, further investigations that include more diverse study populations are needed. The *UGT2B7* SNPs –327A>G, IVS1+985A>G, 802C>T and 211G>T, showing noticeable frequency differences in the present study, are functionally important. The

–327A>G SNP may not be of direct functional importance, but is a part of the haplotype II that includes other promoter SNPs, including –161T>C (previously known as –102T>C). In an *in vitro* assay, the –161C variant resulted in a twofold increase in the transcriptional activity [80]. However, in patients, the effect of –161T>C genotypes on the glucuronidation of morphine is not clear and may depend on medical condition, daily dosage of morphine and administration of other drugs [80]. In human liver microsomes, IVS1+985A>G, a part of the haplotype 4, was significantly associated with increased mRNA expression and the glucuronidation of morphine [44]. A negative impact of 802C>T on the protein content [41] and specific activity of the enzyme has been indicated [70]. However, the SNP did not affect the *in vitro* glucuronidation of AZT [81] and EFV [42]. Finally, compared with the 211G allele, heterologous expression of the 211T variant in HeLa cells resulted in lower protein expression and a significant decrease in the glucuronidation of carvedilol, an α - and β -adrenergic blocking drug [82]. However, the effect of 211G>T on the pharmacokinetics of carvedilol may be cohort-specific [83,84]. Thus, it is important to determine whether the IVS1+985G and 211T enzyme variants result in altered catalytic activity towards AZT and EFV, and whether certain promoter haplotypes, in linkage with coding SNPs (–327A>G is in linkage with 802C>T; table 3), influence antiretroviral pharmacokinetics and clinical response.

A small number of studies that have analyzed the influence of a few *UGT2B7* SNP genotypes on antiretroviral pharmacokinetics and clinical response have yielded varying results [31,47–49]. Using the common *UGT2B7* haplotypes may be more meaningful to interpret HIV/AIDS treatment outcome measures than by analyzing those SNPs singly. In this direction, our *UGT2B7* haplotype frequency data show substantial differences among ethnically diverse populations (table 5), thus providing highly valuable information for future pharmacogenetic studies of antiretroviral therapy.

Differences in *CYP2B6* & *UGT2B7* diversity & differentiation

In addition to *CYP2B6* and *UGT2B7* SNP/allele frequency data, our genetic diversity and differentiation data may also be helpful in interpreting differences in outcomes of antiretroviral therapy among worldwide populations. Worldwide, interpopulation diversity varied much more for *CYP2B6* ($H = 0.254$ – 0.690) than for *UGT2B7* ($H = 0.603$ – 0.737), which reflects more pronounced frequency differences in the *CYP2B6* SNPs than in the *UGT2B7* SNPs. This suggests that interpopulation differences in outcomes of antiretroviral therapy are likely to be more correlated with the differences in *CYP2B6* diversity than with the differences in *UGT2B7* diversity. This is also supported by larger between-population differences for *CYP2B6* than for *UGT2B7* in the AMOVA (10.9 and 6.65%, respectively) and pairwise population differentiation (mean $PA-F_{st} = 0.017$ – 0.226 and 0.030 – 0.089 , respectively) results. Based on the pairwise population differentiation results (supplementary tables s3a & s3b), the PNG (mean $PA-F_{st}$ [*CYP2B6*] = 0.226), FIN (mean $PA-F_{st}$ [*CYP2B6*] = 0.101), Asian (mean $PA-F_{st}$ [*UGT2B7*] = 0.050–0.089) and European (mean $PA-F_{st}$ [*UGT2B7*] = 0.042–0.076) populations are likely to show the most noticeable differences in outcomes of antiretroviral therapy.

Genetic differentiation among populations reflects the action of both random processes (e.g., genetic drift) and natural selection. An interesting question is: does natural selection appear to have played a role in the patterns of genetic differentiation observed at *CYP2B6* and *UGT2B7*, particularly at *CYP2B6* in the PNG population? The present data are limited, and further analyses, including in-depth investigation of the patterns of genetic variation in and around these loci, would be required to investigate this question. However, regardless of whether demography or natural selection is responsible for the SNP frequency differences and corresponding patterns of population differentiation, our results can still be considered

as a starting point in identifying the contribution of genetic variation in *CYP2B6* and *UGT2B7* to (variable) HIV/AIDS treatment outcomes among diverse populations.

In summary, despite the complexity of HIV/AIDS treatment-related genetic investigations, many associations have been described between *CYP2B6* variants, antiretroviral pharmacokinetics and/or therapy outcomes; however, some of those associations must still be validated. On the other hand, only a few studies have taken *UGT2B7* variation into consideration; these studies have yielded inconsistent results. It is worth noting that for each reported association, or where there was no apparent association, the implicated *CYP2B6*/*UGT2B7* variant differs considerably in frequency among ethnic groups. Although *CYP2B6* and *UGT2B7* variation has important clinical implications for HIV/AIDS treatment, recent studies indicate that the contributions of other drug-metabolism enzyme genes [47,85], drug-transporter genes [86], transcription factor genes [64,87] and protein–protein interactions [88] should not be discounted. Furthermore, in addition to SNPs, other key genomic mechanisms are likely to underly phenotypic variability; these include recently discovered alternative splicing of *UGT2B7*, which gives rise to multiple mRNA splice variants with novel functions [89]. Finally, the presence of key functional SNPs, such as *CYP2B6* 983T>C and *UGT2B7* 211G>T in admixed populations emphasizes the need to consider population history while assessing the genotype associations. In the future, antiretroviral treatment-related clinical studies need to consider more comprehensive genetic information, which would enable a systematic understanding of the interplay between ethnicity and HIV/AIDS treatment outcomes. Using this knowledge for optimizing antiretroviral effectiveness, while minimizing toxicity, could have substantial public health benefits, particularly in resource-limited countries.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are thankful to D McNamara, C Myers, C Henry-Halldin, B John, W Blank, A Ramesh and T Phipps for thorough reading and constructive criticism of the manuscript. RK Mehlotra thanks ST Small for reading and discussing this manuscript multiple times, and L Gray for her help with haplotype analysis during the initial stages of this study. The authors are also thankful to the north American, west African, and Papua New Guinean individuals for contributing samples analyzed in this study. Results of this study were presented, in part, at the 12th International Glucuronidation and UGT Workshop, 24–27 July 2008, QC, Canada, and the Cold Spring Harbor/Wellcome Trust Conference on Pharmacogenomics, 19–22 November 2008, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

This study was supported by a Development Award (to RK Mehlotra and C Guillemette) from the Center for AIDS Research, University Hospitals Case Medical Center, OH, USA (NIH grant #AI36219); an Infectious Diseases Research Support (to RK Mehlotra) from STERIS Corporation, Mentor, OH; a Large Pilot Grant (to RK Mehlotra) from the Case Western Reserve University/Cleveland Clinic CTSA grant #UL1RR024989 (National Center for Research Resources); and in part by a grant from National Institute of Dental and Craniofacial Research (NIH grant #IP01DE019759–01 to RJ Jurevic). J Li and M Stoneking are supported by funds from the Max Planck Society. V Menard is a recipient of a CIHR Frederick Banting and Charles Best studentship award. C Guillemette holds a Canada Research Chair in Pharmacogenomics (CIHR grant MOP-42392).

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest

1. Palella FJ Jr, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N. Engl. J. Med. 1998; 338(13):853–860. [PubMed: 9516219] ■ Key paper describing the impact of highly active antiretroviral therapy on HIV/AIDS prognosis.
2. Conway B. The role of adherence to antiretroviral therapy in the management of HIV infection. J. Acquir. Immune Defic. Syndr. 2007; 45(Suppl. 1):S14–S18. [PubMed: 17525686]
3. Easterbrook PJ, Phillips AN, Hill T, et al. Patterns and predictors of the use of different antiretroviral drug regimens at treatment initiation in the UK. HIV Med. 2008; 9(1):47–56. [PubMed: 18199172]
4. Kitahata MM, Gange SJ, Abraham AG, et al. Effect of early versus deferred antiretroviral therapy for HIV on survival. N. Engl. J. Med. 2009; 360(18):1815–1826. [PubMed: 19339714]
5. Telenti A, Zanger UM. Pharmacogenetics of anti-HIV drugs. Annu. Rev. Pharmacol. Toxicol. 2008; 48:227–256. [PubMed: 17883329]
6. Tozzi V. Pharmacogenetics of antiretrovirals. Antiviral Res. 2010; 85(1):190–200. [PubMed: 19744523]
7. Duong M, Buisson M, Peytavin G, et al. Low trough plasma concentrations of nevirapine associated with virologic rebounds in HIV-infected patients who switched from protease inhibitors. Ann. Pharmacother. 2005; 39(4):603–609. [PubMed: 15713788]
8. de Requena, D Gonzalez; Nunez, M.; Jimenez-Nacher, I.; Soriano, V. Liver toxicity caused by nevirapine. AIDS. 2002; 16(2):290–291. [PubMed: 11807315]
9. Marzolini C, Telenti A, Decosterd LA, Greub G, Biollaz J, Buclin T. Efavirenz plasma levels can predict treatment failure and central nervous system side effects in HIV-1-infected patients. AIDS. 2001; 15(1):71–75. [PubMed: 11192870]
10. Mentre F, Escolano S, Diquet B, Golmard JL, Mallet A. Clinical pharmacokinetics of zidovudine: inter and intraindividual variability and relationship to long term efficacy and toxicity. Eur. J. Clin. Pharmacol. 1993; 45(5):397–407. [PubMed: 8112367]
11. Sasomsin P, Mentre F, Diquet B, Simon F, Brun-Vezinet F. Relationship between exposure to zidovudine and decrease of P24 antigenemia in HIV-infected patients in monotherapy. Fundam. Clin. Pharmacol. 2002; 16(5):347–352. [PubMed: 12602459]
12. Stahle L, Moberg L, Svensson JO, Sonnerborg A. Efavirenz plasma concentrations in HIV-infected patients: inter- and intraindividual variability and clinical effects. Ther. Drug Monit. 2004; 26(3):267–270. [PubMed: 15167626]
13. Dahri K, Ensom MH. Efavirenz and nevirapine in HIV-1 infection: is there a role for clinical pharmacokinetic monitoring. Clin. Pharmacokinet. 2007; 46(2):109–132. [PubMed: 17253884] ■ A comprehensive article discussing the utility of therapeutic drug monitoring in HIV/AIDS patients.
14. Escolano S, Mentre F, Golmard JL, Diquet B, Mallet A. Plasma concentrations of zidovudine. N. Engl. J. Med. 1992; 326(23):1571. [PubMed: 1579149]
15. Erickson DA, Mather G, Trager WF, Levy RH, Keirns JJ. Characterization of the *in vitro* biotransformation of the HIV-1 reverse transcriptase inhibitor nevirapine by human hepatic cytochromes P-450. Drug Metab. Dispos. 1999; 27(12):1488–1495. [PubMed: 10570031]
16. Ward BA, Gorski JC, Jones DR, Hall SD, Flockhart DA, Desta Z. The cytochrome P450 2B6 (CYP2B6) is the main catalyst of efavirenz primary and secondary metabolism: implication for HIV/AIDS therapy and utility of efavirenz as a substrate marker of CYP2B6 catalytic activity. J. Pharmacol. Exp. Ther. 2003; 306(1):287–300. [PubMed: 12676886]
17. Zanger UM, Klein K, Saussele T, Blievernicht J, Hofmann MH, Schwab M. Polymorphic *CYP2B6*: molecular mechanisms and emerging clinical significance. Pharmacogenomics. 2007; 8(7):743–759. [PubMed: 17638512] ■ Comprehensive review describing the molecular/genetic and functional aspects of *CYP2B6*.
18. Watanabe T, Sakuyama K, Sasaki T, et al. Functional characterization of 26 *CYP2B6* allelic variants (*CYP2B6.2–CYP2B6.28*, except *CYP2B6.22*). Pharmacogenet. Genomics. 2010; 20(7):459–462. [PubMed: 20517174]
19. Zhang H, Sridar C, Kanaan C, Amunugama H, Ballou DP, Hollenberg PF. Polymorphic variants of cytochrome P450 2B6 (*CYP2B6.4–CYP2B6.9*) exhibit altered rates of metabolism for bupropion

and efavirenz: a charge-reversal mutation in the *K139E* variant (*CYP2B6*.8) impairs formation of a functional P450-reductase complex. *J. Pharmacol. Exp. Ther.* 2011; 338(3):803–809. [PubMed: 21659470] ■ Key paper describing efavirenz metabolism by heterologously expressed, reconstituted *CYP2B6* variants.

20. Desta Z, Saussele T, Ward B, et al. Impact of *CYP2B6* polymorphism on hepatic efavirenz metabolism *in vitro*. *Pharmacogenomics.* 2007; 8(6):547–558. [PubMed: 17559344] ■ First study describing efavirenz metabolism by human liver microsomes.
21. Hofmann MH, Blievernicht JK, Klein K, et al. Aberrant splicing caused by single nucleotide polymorphism c.516G>T [Q172H], a marker of *CYP2B6**6, is responsible for decreased expression and activity of CYP2B6 in liver. *J. Pharmacol. Exp. Ther.* 2008; 325(1):284–292. [PubMed: 18171905] ■ Key paper describing the molecular/genetic mechanism of ‘slow metabolism’ phenotype associated with the most common *CYP2B6**6 allele.
22. Cabrera SE, Santos D, Valverde MP, et al. Influence of the cytochrome P450 2B6 genotype on population pharmacokinetics of efavirenz in human immunodeficiency virus patients. *Antimicrob. Agents Chemother.* 2009; 53(7):2791–2798. [PubMed: 19433561]
23. Haas DW, Gebretsadik T, Mayo G, et al. Associations between *CYP2B6* polymorphisms and pharmacokinetics after a single dose of nevirapine or efavirenz in African Americans. *J. Infect. Dis.* 2009; 199(6):872–880. [PubMed: 19239339]
24. Mahungu T, Smith C, Turner F, et al. Cytochrome P450 2B6 516G>T is associated with plasma concentrations of nevirapine at both 200 mg twice daily and 400 mg once daily in an ethnically diverse population. *HIV Med.* 2009; 10(5):310–317. [PubMed: 19228205]
25. Motsinger AA, Ritchie MD, Shafer RW, et al. Multilocus genetic interactions and response to efavirenz-containing regimens: an adult AIDS clinical trials group study. *Pharmacogenet. Genomics.* 2006; 16(11):837–845. [PubMed: 17047492]
26. Ribaud HJ, Liu H, Schwab M, et al. Effect of *CYP2B6*, *ABCB1*, and *CYP3A5* polymorphisms on efavirenz pharmacokinetics and treatment response: an AIDS Clinical Trials Group study. *J. Infect. Dis.* 2010; 202(5):717–722. [PubMed: 20662624]
27. Rotger M, Tegude H, Colombo S, et al. Predictive value of known and novel alleles of *CYP2B6* for efavirenz plasma concentrations in HIV-infected individuals. *Clin. Pharmacol. Ther.* 2007; 81(4):557–566. [PubMed: 17235330]
28. Schipani A, Wyen C, Mahungu T, et al. Integration of population pharmacokinetics and pharmacogenetics: an aid to optimal nevirapine dose selection in HIV-infected individuals. *J. Antimicrob. Chemother.* 2011; 66(6):1332–1339. [PubMed: 21441248] ■ Key paper describing a population pharmacokinetic–pharmacogenetic model of nevirapine for dosage optimization.
29. Wyen C, Hendra H, Vogel M, et al. Impact of *CYP2B6* 983T>C polymorphism on non-nucleoside reverse transcriptase inhibitor plasma concentrations in HIV-infected patients. *J. Antimicrob. Chemother.* 2008; 61(4):914–918. [PubMed: 18281305]
30. Ciccacci C, Borgiani P, Ceffa S, et al. Nevirapine-induced hepatotoxicity and pharmacogenetics: a retrospective study in a population from Mozambique. *Pharmacogenomics.* 2010; 11(1):23–31. [PubMed: 20017669]
31. Mehlotra RK, Cheruvu VK, Zikursh MJB, et al. Chemokine (C-C motif) receptor 5-2459 genotype in patients receiving highly active antiretroviral therapy: race-specific influence on virologic success. *J. Infect. Dis.* 2011; 204(2):291–298. [PubMed: 21673041]
32. Carr DF, la Porte CJ, Pirmohamed M, Owen A, Cortes CP. Haplotype structure of *CYP2B6* and association with plasma efavirenz concentrations in a Chilean HIV cohort. *J. Antimicrob. Chemother.* 2010; 65(9):1889–1893. [PubMed: 20639527]
33. Chen J, Sun J, Ma Q, et al. CYP2B6 polymorphism and nonnucleoside reverse transcriptase inhibitor plasma concentrations in Chinese HIV-infected patients. *Ther. Drug Monit.* 2010; 32(5): 573–578. [PubMed: 20625352]
34. Gatanaga H, Hayashida T, Tsuchiya K, et al. Successful efavirenz dose reduction in HIV type 1-infected individuals with cytochrome P450 2B6 *6 and *26. *Clin. Infect. Dis.* 2007; 45(9):1230–1237. [PubMed: 17918089] ■ An important study describing that efavirenz dose reduction can be achieved in HIV patients carrying *CYP2B6* slow metabolism genotypes.

35. Leger P, Dillingham R, Beauharnais CA, et al. CYP2B6 variants and plasma efavirenz concentrations during antiretroviral therapy in Port-au-Prince, Haiti. *J. Infect. Dis.* 2009; 200(6): 955–964. [PubMed: 19659438]
36. Lindfelt T, O'Brien J, Song JC, Patel R, Winslow DL. Efavirenz plasma concentrations and cytochrome 2B6 polymorphisms. *Ann. Pharmacother.* 2010; 44(10):1572–1578. [PubMed: 20841522]
37. Puthanakit T, Tanpaiboon P, Aupibul L, Cressey TR, Sirisanthana V. Plasma efavirenz concentrations and the association with *CYP2B6*-516G>T polymorphism in HIV-infected Thai children. *Antivir. Ther.* 2009; 14(3):315–320. [PubMed: 19474465]
38. Ramachandran G, Ramesh K, Kumar AK Hemanth, et al. Association of high T allele frequency of *CYP2B6* G516T polymorphism among ethnic south Indian HIV-infected patients with elevated plasma efavirenz and nevirapine. *J. Antimicrob. Chemother.* 2009; 63(4):841–843. [PubMed: 19218571]
39. To KW, Liu ST, Cheung SW, Chan DP, Chan RC, Lee SS. Pharmacokinetics of plasma efavirenz and *CYP2B6* polymorphism in southern Chinese. *Ther. Drug Monit.* 2009; 31(4):527–530. [PubMed: 19531981]
40. Hasse B, Gunthard HF, Bleiber G, Krause M. Efavirenz intoxication due to slow hepatic metabolism. *Clin. Infect. Dis.* 2005; 40(3):e22–e23. [PubMed: 15668854]
41. Barbier O, Turgeon D, Girard C, et al. 3'-azido-3'-deoxythymidine (AZT) is glucuronidated by human UDP-glucuronosyltransferase 2B7 (UGT2B7). *Drug Metab. Dispos.* 2000; 28(5):497–502. [PubMed: 10772627]
42. Belanger AS, Caron P, Harvey M, Zimmerman PA, Mehlotra RK, Guillemette C. Glucuronidation of the antiretroviral drug efavirenz by UGT2B7 and an *in vitro* investigation of drug–drug interaction with zidovudine. *Drug Metab. Dispos.* 2009; 37(9):1793–1796. [PubMed: 19487252]
43. Guillemette C. Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *Pharmacogenomics J.* 2003; 3(3):136–158. [PubMed: 12815363]
44. Innocenti F, Liu W, Fackenthal D, et al. Single nucleotide polymorphism discovery and functional assessment of variation in the UDP-glucuronosyltransferase 2B7 gene. *Pharmacogenet. Genomics.* 2008; 18(8):683–697. [PubMed: 18622261]
45. Mackenzie PI, Gregory PA, Lewinsky RH, et al. Polymorphic variations in the expression of the chemical detoxifying UDP glucuronosyltransferases. *Toxicol. Appl. Pharmacol.* 2005; 207(2 Suppl.):77–83. [PubMed: 15979674]
46. Shipkova M, Wieland E. Glucuronidation in therapeutic drug monitoring. *Clin. Chim. Acta.* 2005; 358(1–2):2–23. [PubMed: 15893300]
47. Kwara A, Lartey M, Sagoe KW, Kenu E, Court MH. *CYP2B6*, *CYP2A6* and *UGT2B7* genetic polymorphisms are predictors of efavirenz mid-dose concentration in HIV-infected patients. *AIDS.* 2009; 23(16):2101–2106. [PubMed: 19779319]
48. Kwara A, Lartey M, Boamah I, et al. Interindividual variability in pharmacokinetics of generic nucleoside reverse transcriptase inhibitors in TB/HIV-coinfected Ghanaian patients: *UGT2B7*1c* is associated with faster zidovudine clearance and glucuronidation. *J. Clin. Pharmacol.* 2009; 49(9):1079–1090. [PubMed: 19628728]
49. Elens L, Vandercam B, Yombi JC, Lison D, Wallemacq P, Haufroid V. Influence of host genetic factors on efavirenz plasma and intracellular pharmacokinetics in HIV-1-infected patients. *Pharmacogenomics.* 2010; 11(9):1223–1234. [PubMed: 20860463]
50. Arenaz I, Vicente J, Fanlo A, et al. Haplotype structure and allele frequencies of *CYP2B6* in Spaniards and Central Americans. *Fundam. Clin. Pharmacol.* 2010; 24(2):247–253. [PubMed: 19682083]
51. Davaalkham J, Hayashida T, Tsuchiya K, Gatanaga H, Nyamkhuu D, Oka S. Allele and genotype frequencies of cytochrome P450 2B6 gene in a Mongolian population. *Drug Metab. Dispos.* 2009; 37(10):1991–1993. [PubMed: 19581387]
52. Klein K, Lang T, Saussele T, et al. Genetic variability of *CYP2B6* in populations of African and Asian origin: allele frequencies, novel functional variants, and possible implications for anti-HIV therapy with efavirenz. *Pharmacogenet. Genomics.* 2005; 15(12):861–873. [PubMed: 16272958]

53. Mehlotra RK, Ziats MN, Bockarie MJ, Zimmerman PA. Prevalence of *CYP2B6* alleles in malaria-endemic populations of west Africa and Papua New Guinea. *Eur. J. Clin. Pharmacol.* 2006; 62(4): 267–275. [PubMed: 16506047]
54. Xu BY, Guo LP, Lee SS, et al. Genetic variability of *CYP2B6* polymorphisms in four southern Chinese populations. *World J. Gastroenterol.* 2007; 13(14):2100–2103. [PubMed: 17465455]
55. Holthe M, Rakvag TN, Klepstad P, et al. Sequence variations in the UDP-glucuronosyltransferase 2B7 (*UGT2B7*) gene: identification of 10 novel single nucleotide polymorphisms (SNPs) and analysis of their relevance to morphine glucuronidation in cancer patients. *Pharmacogenomics J.* 2003; 3(1):17–26. [PubMed: 12629580]
56. Saeki M, Saito Y, Jinno H, et al. Single nucleotide polymorphisms and haplotype frequencies of *UGT2B4* and *UGT2B7* in a Japanese population. *Drug Metab. Dispos.* 2004; 32(9):1048–1054. [PubMed: 15319348]
57. Saito K, Moriya H, Sawaguchi T, et al. Haplotype analysis of UDP-glucuronocyltransferase 2B7 gene (*UGT2B7*) polymorphisms in healthy Japanese subjects. *Clin. Biochem.* 2006; 39(3):303–308. [PubMed: 16466707]
58. Mehlotra RK, Bockarie MJ, Zimmerman PA. *CYP2B6* 983T>C polymorphism is prevalent in West Africa but absent in Papua New Guinea: implications for HIV/AIDS treatment. *Br. J. Clin. Pharmacol.* 2007; 64(3):391–395. [PubMed: 17391322]
59. Mehlotra RK, Bockarie MJ, Zimmerman PA. Prevalence of *UGT1A9* and *UGT2B7* nonsynonymous single nucleotide polymorphisms in west African, Papua New Guinean, and north American populations. *Eur. J. Clin. Pharmacol.* 2007; 63(1):1–8. [PubMed: 17115150]
60. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467(7319):1061–1073. [PubMed: 20981092]
61. Zimmerman PA, Dadzie KY, De Sole G, Remme J, Alley ES, Unnasch TR. *Onchocerca volvulus* DNA probe classification correlates with epidemiologic patterns of blindness. *J. Infect. Dis.* 1992; 165(5):964–968. [PubMed: 1569351]
62. Zimmerman PA, Woolley I, Masinde GL, et al. Emergence of *FY*A(null)* in a *Plasmodium vivax*-endemic region of Papua New Guinea. *Proc. Natl Acad. Sci. USA.* 1999; 96(24):13973–13977. [PubMed: 10570183]
63. Burchard EG, Borrell LN, Choudhry S, et al. Latino populations: a unique opportunity for the study of race, genetics, and social environment in epidemiological research. *Am. J. Public Health.* 2005; 95:2161–2168. [PubMed: 16257940]
64. Lamba V, Lamba J, Yasuda K, et al. Hepatic *CYP2B6* expression: gender and ethnic differences and relationship to *CYP2B6* genotype and CAR (constitutive androstane receptor) expression. *J. Pharmacol. Exp. Ther.* 2003; 307(3):906–922. [PubMed: 14551287]
65. Solus JF, Arietta BJ, Harris JR, et al. Genetic variation in eleven Phase I drug metabolism genes in an ethnically diverse population. *Pharmacogenomics.* 2004; 5(7):895–931. [PubMed: 15469410] ■ First comprehensive report of drug-metabolism enzyme genetic variation in ethnically diverse populations.
66. Ingman M, Gyllenstein U. A recent genetic link between Sami and the Volga-Ural region of Russia. *Eur. J. Hum. Genet.* 2007; 15(1):115–120. [PubMed: 16985502]
67. Kittles RA, Perola M, Peltonen L, et al. Dual origins of Finns revealed by Y chromosome haplotype variation. *Am. J. Hum. Genet.* 1998; 62(5):1171–1179. [PubMed: 9545401]
68. Salmela E, Lappalainen T, Franssón I, et al. Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in northern Europe. *PLoS ONE.* 2008; 3(10):e3519. [PubMed: 18949038]
69. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010; 467(7311):52–58. [PubMed: 20811451]
70. Bhasker CR, McKinnon W, Stone A, et al. Genetic polymorphism of UDP-glucuronosyltransferase 2B7 (*UGT2B7*) at amino acid 268: ethnic diversity of alleles and potential clinical significance. *Pharmacogenetics.* 2000; 10(8):679–685. [PubMed: 11186130]
71. Hwang MS, Lee SJ, Jeong HE, Lee S, Yoo MA, Shin JG. Genetic variations in UDP-glucuronosyltransferase 2B7 gene (*UGT2B7*) in a Korean population. *Drug Metab. Pharmacokinet.* 2010; 25(4):398–402. [PubMed: 20814162]

72. Ferlin A, Ganz F, Pengo M, Selice R, Frigo AC, Foresta C. Association of testicular germ cell tumor with polymorphisms in estrogen receptor and steroid metabolism genes. *Endocr. Relat. Cancer*. 2010; 17(1):17–25. [PubMed: 19776291]
73. Wang J, Sonnerborg A, Rane A, et al. Identification of a novel specific *CYP2B6* allele in Africans causing impaired metabolism of the HIV drug efavirenz. *Pharmacogenet. Genomics*. 2006; 16(3): 191–198. [PubMed: 16495778]
74. Bonnen PE, Wang PJ, Kimmel M, Chakraborty R, Nelson DL. Haplotype and linkage disequilibrium architecture for human cancer-associated genes. *Genome Res*. 2002; 12(12):1846–1853. [PubMed: 12466288]
75. Li J, Zhang L, Zhou H, Stoneking M, Tang K. Global patterns of genetic diversity and signals of natural selection for human ADME genes. *Hum. Mol. Genet*. 2011; 20(3):528–540. [PubMed: 21081654] ■ A comprehensive study describing genetic/evolutionary characteristics of many drug-metabolism enzymes and transporter genes in worldwide populations.
76. Myles S, Tang K, Somel M, Green RE, Kelso J, Stoneking M. Identification and analysis of genomic regions with large between-population differentiation in humans. *Ann. Hum. Genet*. 2008; 72(Pt 1):99–110. [PubMed: 18184145]
77. Rakhmanina NY, van den Anker JN. Efavirenz in the therapy of HIV infection. *Expert Opin. Drug Metab. Toxicol*. 2010; 6(1):95–103. [PubMed: 20001610] ■ A comprehensive article describing the need and approaches to efavirenz dose reduction in diverse populations.
78. Rotger M, Telenti A. Optimizing efavirenz treatment: *CYP2B6* genotyping or therapeutic drug monitoring. *Eur. J. Clin. Pharmacol*. 2008; 64(4):335–336. [PubMed: 18188549] ■ An expert opinion article discussing the utility of therapeutic drug monitoring in HIV/AIDS patients.
79. Yuan J, Guo S, Hall D, et al. Toxicogenomics of nevirapine-associated cutaneous and hepatic adverse events among populations of African, Asian, and European descent. *AIDS*. 2011; 25(10): 1271–1280. [PubMed: 21505298]
80. Duguay Y, Baar C, Skorpen F, Guillemette C. A novel functional polymorphism in the uridine diphosphate-glucuronosyltransferase 2B7 promoter with significant impact on promoter activity. *Clin. Pharmacol. Ther*. 2004; 75(3):223–233. [PubMed: 15001974]
81. Court MH, Krishnaswamy S, Hao Q, et al. Evaluation of 3'-azido-3'-deoxythymidine, morphine, and codeine as probe substrates for UDP-glucuronosyltransferase 2B7 (UGT2B7) in human liver microsomes: specificity and influence of the *UGT2B7*2* polymorphism. *Drug Metab. Dispos*. 2003; 31(9):1125–1133. [PubMed: 12920168]
82. Takekuma Y, Takenaka T, Yamazaki K, Ueno K, Sugawara M. Stereoselective metabolism of racemic carvedilol by UGT1A1 and UGT2B7, and effects of mutation of these enzymes on glucuronidation activity. *Biol. Pharm.Bull*. 2007; 30(11):2146–2153. [PubMed: 17978490]
83. Honda M, Toyoda W, Shimizu T, et al. UGT2B7*3 did not affect the pharmacokinetics of *R*- and *S*-carvedilol in healthy Japanese. *Drug Metab. Pharmacokinet*. 2007; 22(5):382–386. [PubMed: 17965522]
84. Takekuma Y, Takenaka T, Kiyokawa M, et al. Evaluation of effects of polymorphism for metabolic enzymes on pharmacokinetics of carvedilol by population pharmacokinetic analysis. *Biol. Pharm.Bull*. 2007; 30(3):537–542. [PubMed: 17329852]
85. di Iulio J, Fayet A, Arab-Alameddine M, et al. In vivo analysis of efavirenz metabolism in individuals with impaired CYP2A6 function. *Pharmacogenet. Genomics*. 2009; 19(4):300–309. [PubMed: 19238117]
86. Grady BJ, Torstenson ES, McLaren PJ, et al. Use of biological knowledge to inform the analysis of gene–gene interactions involved in modulating virologic failure with efavirenz-containing treatment regimens in ART-naïve ACTG clinical trials participants. *Pac. Symp. Biocomput*. 2011; 2011:253–264. [PubMed: 21121053]
87. Wyen C, Hendra H, Siccardi M, et al. Cytochrome P450 2B6 (*CYP2B6*) and constitutive androstane receptor (*CAR*) polymorphisms are associated with early discontinuation of efavirenz-containing regimens. *J. Antimicrob. Chemother*. 2011
88. Ishii Y, Takeda S, Yamada H. Modulation of UDP-glucuronosyltransferase activity by protein–protein association. *Drug Metab. Rev*. 2010; 42(1):145–158. [PubMed: 19817679]

89. Menard V, Eap O, Roberge J, Harvey M, Levesque E, Guillemette C. Transcriptional diversity at the *UGT2B7* locus is dictated by extensive pre-mRNA splicing mechanisms that give rise to multiple mRNA splice variants. *Pharmacogenet. Genomics*. 2011; 21(10):631–641. [PubMed: 21881541]

Websites

101. WHO. HIV/AIDS Antiretroviral therapy. www.who.int/hiv/topics/treatment/en/index.html
102. *CYP2B6* allele nomenclature. www.cypalleles.ki.se/cyp2b6.htm
103. *UGT2B7* allele nomenclature. www.pharmacogenomics.pha.ulaval.ca/webdav/site/pharmacogenomics/shared/Nomenclature/UGT2B/UGT2B7.htm
104. BEAGLE Genetic Analysis Software Package. faculty.washington.edu/browning/beagle/beagle.html
105. GENEPOP. www.genepop.curtin.edu.au/
106. SHEsis. www.analysis.bio-x.cn/myAnalysis.php
107. Center for Commercialization University of Washington UW. PHASE. www.depts.washington.edu/uwc4c/expresslicenses/assets/phase/
108. Fisher's Exact Test. www.langsrud.com/fisher.htm
109. DnaSP v5.10. www.ub.edu/dnasp
110. Arlequin ver 3.11. www.cmpg.unibe.ch/software/arlequin3/
111. dbSNP short geneic variations. www.ncbi.nlm.nih.gov/projects/SNP/
112. WHO programmes and projects: HIV. www.who.int/hiv
113. HIV and AIDS data hub. www.aidsdatahub.org/en/country-profiles/papua-new-guinea
114. National AIDS Council. Papua New Guinea. www.staging.nacs.org.pg/resources/documents/Introduction_of_ART_Mar05.pdf

Executive summary

Background

- The key to a long life with HIV is highly active antiretroviral therapy, the HIV medications that changed the lives of all those affected and infected with HIV. Highly active antiretroviral therapy is defined as treatment with at least three antiretroviral medications, typically two nucleoside or nucleotide reverse transcriptase inhibitors plus a non-nucleoside reverse transcriptase inhibitor or a protease inhibitor.
- The non-nucleoside reverse transcriptase inhibitors efavirenz or nevirapine and nucleoside reverse transcriptase inhibitor zidovudine are included in many highly active antiretroviral therapy regimens used worldwide. Pharmacokinetics of, and outcomes of treatment with these drugs vary interindividually as well as interethnically. Genetic changes in the enzyme metabolism of these medications may contribute to clinical variability.
- The hepatic enzymes CYP2B6 and UGT2B7 play a major role in the metabolism of efavirenz, nevirapine and zidovudine. The information regarding comparative variation in both *CYP2B6* and *UGT2B7* genes on a worldwide scale is limited.

Materials & methods

- In the present study, we provide a view of *UGT2B7* haplotype structure and quantify the genetic diversity and differentiation at both *CYP2B6* and *UGT2B7* genes on a worldwide scale, by analyzing key SNPs in north American, west African, Papua New Guinean and 1000 Genomes Project populations.

Results

- We found significant differences in certain SNP and allele frequencies of *CYP2B6* and *UGT2B7* among worldwide populations, which in turn produce differences in diversity and pairwise differentiation. Diversity values were higher for *UGT2B7* than for *CYP2B6*, although there was more diversity between populations for *CYP2B6*. For both genes, most of the genetic variation was observed among individuals within populations, with the Papua New Guinean population showing the highest pairwise differentiation values for *CYP2B6*, and the Asian and European populations showing higher pairwise differentiation values for *UGT2B7*.

Conclusion

- These findings are meant to help guide investigations into antiretroviral drug metabolism and therapy outcomes, aimed at optimizing effectiveness and minimizing toxicity of these drugs.

Figure 1. *CYP2B6* and *UGT2B7* haplotype diversity in various populations

Details regarding the study (EA, AFA, ASA, HA, west African and PNG) and the reference (CEU, TSI, GBR, FIN, ASW, YRI, LWK, CHB, CHS, JPT, MXL and PUR) populations are provided in supplementary table s1 and in the 'Materials and methods' section.

AFA: African-American; ASA: Asian-American; ASW: African ancestry from southwest USA; CEU: Northern and western European from UT, USA; CHB: Han Chinese from Beijing, China; CHS: Han Chinese from south China; EA: European-American; FIN: Finnish; HA: Hispanic-American; JPT: Japanese from Tokyo, Japan; LWK: Luhya from Webuye, Kenya; MXL: Mexican ancestry from Los Angeles, CA, USA; PNG: Papua New Guinean; PUR: Puerto Rican from Puerto Rico, USA; TSI: Toscani from Italy; YRI: Yoruba from Ibadan, Nigeria.

Table 1

Frequencies of *CYP2B6* SNPs in various populations.

Population [‡]	n	64C>T	516G>T	777C>A	785A>G	983T>C	1459C>T
		<i>f(T)</i> [‡]	<i>f(T)</i> [‡]	<i>f(A)</i> [‡]	<i>f(G)</i> [‡]	<i>f(C)</i> [‡]	<i>f(T)</i> [‡]
European-American	59	0.03	0.34	0	0.38	0	0.07
Northern and western European from UT (USA)	90	0.04	0.25	NA	0.20	0	0.07
Toscani from Italy	92	0.07	0.29	NA	0.21	0	0.09
British from England and Scotland	43	0.03	0.26	NA	0.14	0	0.10
Finnish	36	0.04	0.08	NA	0.10	0	0.08
African-American	85	0.02	0.36	0	0.37	0.08	0.04
African ancestry from southwest USA	24	0.06	0.38	NA	0.35	0.04	0.02
Yoruba from Ibadan (Nigeria)	78	0.04	0.35	NA	0.22	0.12	0.01
West African	153	0.04	0.5	0	0.42	0.05	0.02
Luhya from Webuye (Kenya)	67	0.08	0.37	NA	0.33	0.06	0
Asian-American	61	0.07	0.23	0	0.27	0	0.02
Han Chinese from Beijing	68	0.03	0.13	NA	0.15	0	0
Han Chinese from south China	25	0.04	0.10	NA	0.10	0	0
Japanese from Tokyo (Japan)	84	0.01	0.21	NA	0.19	0	0.01
Hispanic-American	77	0.03	0.37	0	0.35	0.01	0.07
Mexican ancestry from Los Angeles (CA, USA)	17	0.06	0.29	NA	0.24	0	0.12
Puerto Rican from Puerto Rico (USA)	5	0	0.30	NA	0.30	0.05	0.05
Papua New Guinean	172	0	0.65 [*]	0	0.64 [*]	0	0.03

[‡] As described in the 'Materials and methods' section.

[‡] Mutant allele.

^{*} p< 0.05 compared with any other population.

f: Frequency; NA: Not available.

Table 2

Frequencies of *UGT2B7* SNPs in various populations.

Population [‡]	n	-327A>G	-138G>A	-125T>C	211G>T	IVS1+985A>G	802C>T	1192G>A
		$f(G)$ [‡]	$f(A)$ [‡]	$f(C)$ [‡]	$f(T)$ [‡]	$f(G)$ [‡]	$f(T)$ [‡]	$f(A)$ [‡]
European-American	79	0.48	0.03	0.03	0	0.12	0.53*	0
Northern and western European from UT (USA)	90	0.51	0.02	0	0	0.16	NA	0
Toscani from Italy	92	0.48	0.08	0	0	0.13	NA	0
British from England and Scotland	43	0.47	0.01	0	0	0.13	NA	0
Finnish	36	0.51	0.04	0	0.04	0.24	NA	0
African-American	82	0.66	0.02	0.02	0	0.23	0.34	0
African ancestry from southwest USA	24	0.79	0.04	0.08	0	0.38	NA	0
Yoruba from Ibadan (Nigeria)	78	0.80	0.03	0.04	0	0.33	NA	0
West African	133	0.79	0.02	0.05	0	0.34	0.21	0
Luhya from Webuye (Kenya)	67	0.72	0.09	0.04	0.01	0.29	NA	0
Asian-American	81	0.61	0.01	0.09	0.09	0.09	0.4	0
Han Chinese from Beijing	68	0.68	0	0.07	0.12	0.04	NA	0
Han Chinese from south China	25	0.76	0	0.1	0.16	0.04	NA	0
Japanese from Tokyo (Japan)	84	0.67	0	0.07	0.13	0.11	NA	0
Hispanic-American	84	0.72	0.01	0.01	0.02	0.37	0.28	0
Mexican ancestry from Los Angeles (CA, USA)	17	0.66	0.09	0	0	0.31	NA	0
Puerto Rican from Puerto Rico (USA)	5	0.7	0.1	0	0	0.45	NA	0
Papua New Guinean	142	0.72	0	0	0	0.27	0.28	0

[‡] As described in the 'Materials and methods' section.

^{*} Mutant allele.

* p < 0.05 compared with African-American, west African, Asian-American, Hispanic-American or Papua New Guinean population.
f: Frequency; NA: Not available.

Table 3

Pairwise linkage disequilibrium in *CYP2B6* and *UGT2B7* SNPs in various populations.

Population [†]	SNP pair	
	<i>CYP2B6</i> 516–785 <i>D'</i> (<i>r</i> ²)	<i>UGT2B7</i> –327–802 <i>D'</i> (<i>r</i> ²)
European–American	0.873 (0.657)	1 (0.977)
Northern and western European from UT (USA)	0.778 (0.454)	–
Toscani from Italy	0.82 (0.447)	–
British from England and Scotland	1 (0.472)	–
Finnish	0.815 (0.561)	–
African–American	0.973 (0.901)	1 (1)
African ancestry from southwest USA	1 (0.914)	–
Yoruba from Ibadan (Nigeria)	0.955 (0.48)	–
West African	1 (0.719)	0.976 (0.911)
Luhya from Webuye (Kenya)	0.964 (0.763)	–
Asian–American	1 (0.803)	1 (1)
Han Chinese from Beijing	0.869 (0.63)	–
Han Chinese from south China	1 (1)	–
Japanese from Tokyo (Japan)	0.882 (0.695)	–
Hispanic–American	0.877 (0.727)	1 (1)
Mexican ancestry from Los Angeles (CA, USA)	1 (0.738)	–
Puerto Rican from Puerto Rico (USA)	0.762 (0.58)	–
Papua New Guinean	0.987 (0.926)	1 (1)

[†] As described in the 'Materials and methods' section.

–: 802C>T SNP data not available.

Table 4

Frequencies of certain *CYP2B6* haplotypes in reference populations.

Population [†]	Haplotype				
	516T_785G [‡]	516G_785A_983C [§]	516G_785G_983C [¶]		
Northern and western European from UT (USA)	0.17	0	0	0	
Toscani from Italy	0.18	0	0	0	
British from England and Scotland	0.14	0	0	0	
Finnish	0.07	0	0	0	
African ancestry from southwest USA	0.35	0.04		0	
Yoruba from Ibadan (Nigeria)	0.21	0.11		0.01	
Luhya from Webuye (Kenya)	0.32	0.05		0	
Han Chinese from Beijing	0.12	0		0	
Han Chinese from south China	0.10	0		0	
Japanese from Tokyo (Japan)	0.17	0		0	
Mexican ancestry from Los Angeles (CA, USA)	0.23	0		0	
Puerto Rican from Puerto Rico (USA)	0.25	0.05		0	

[†] As described in the 'Materials and methods' section.

[‡] Presumably *6.

[§] Presumably *18 [52,58].

[¶] Presumably *16 [73].

Table 5

Frequencies of *UGT2B7* alleles in study populations.

Allele [†]	SNP							Population [‡]					
	- 327	- 138	- 125	211	IVS1+985	802	1192	European-American (n = 79)	Hispanic-American (n = 82)	Asian-American (n = 81)	African-American (n = 84)	West African (n = 133)	Papua New Guinean (n = 142)
*1a	G	G	T	G	A	C	G	0.34	0.32	0.33	0.40	0.40	0.45
*1b	G	G	C	G	A	C	G	0.03	0.01	0.09	0.02	0.05	0
1m	G	G	T	G	G	C	G	0.11	0.37	0.09	0.23	0.34	0.27
*2a	A	G	T	G	A	T	G	0.49**	0.27	0.38	0.32	0.19	0.28
*2g	A	A	T	G	A	T	G	0.03	0.01	0.01	0.02	0.02	0
*2i	A	G	T	G	G	T	G	0.01	0	0	0	0	0
*3	G	G	T	T	A	C	G	0	0.02	0.09	0	0	0
*4	G	G	T	G	A	C	A	0	0	0	0	0	0

[†]UGT2B7*1m and *2i represent new haplotypes (alleles), require validation.

[‡]Based on [103].

^{*}As described in the 'Materials and methods' section.

^{*}p < 0.05 compared with any other north American or Papua New Guinean population;

^{**}p < 0.05 compared with any other north American (except Asian-American; p = 0.07), west African or Papua New Guinean population.

Table 6

Distribution of *CYP2B6* and *UGT2B7* genetic variation among and within ethnic groups.

Source of variation	<i>CYP2B6</i>		<i>UGT2B7</i>	
	Variation (%)	<i>p</i> -value	Variation (%)	<i>p</i> -value
Among groups	9.3	<0.001	6.25	<0.001
Among populations within groups	1.6	<0.001	0.4	0.039
Among individuals within populations	89.1	<0.001	93.35	<0.001

Ethnic groups: European (European–American; northern and western European from UT, USA; Toscani from Italy, British from England and Scotland; and Finnish), African (African–American; African ancestry from the southwest USA, Yoruba from Ibadan, Nigeria; west African; and Luhya from Webuye, Kenya), Asian (Asian–American; Han Chinese from Beijing; Han Chinese from south China; and Japanese from Tokyo, Japan), Hispanic (Hispanic–American; Mexican ancestry from Los Angeles, CA, USA; and Puerto Rican from Puerto Rico, USA), and Papua New Guinean. All these populations are described in the ‘Materials and methods’ section.

Table 7

Pairwise genetic differentiation among various ethnic groups.

	European	African	Asian	Hispanic	Papua New Guinean
European	–	0.080 [‡]	0.044 [‡]	0.019 [‡]	0.059 [‡]
African	0.043 [‡]	–	0.068 [‡]	0.021 [‡]	0.003 [‡]
Asian	0.012 [‡]	0.076 [‡]	–	0.043 [‡]	0.059 [‡]
Hispanic	0.012 [‡]	0.011 [‡]	0.048 [‡]	–	0.008 [‡]
Papua New Guinean	0.254 [‡]	0.112 [‡]	0.338 [‡]	0.163 [‡]	–

[‡] Values represent pairwise differentiation at UGT2B7.

[‡] Values represent pairwise differentiation at CYP2B6.

[‡] Ethnic groups: European (European–American; northern and western European from UT, USA; Toscani from Italy; British from England and Scotland; and Finnish), African (African–American; African Ancestry from southwest USA; Yoruba from Ibadan, Nigeria; west African; and Luhya from Webuye, Kenya), Asian (Asian–American; Han Chinese from Beijing, China; Han Chinese from south China; and Japanese from Tokyo, Japan), Hispanic (Hispanic–American; Mexican ancestry from Los Angeles, CA, USA; and Puerto Rican from Puerto Rico, USA) and Papua New Guinean. All these populations are described in the ‘Materials and methods’ section.