Title Page

Variants in the *CYP2B6* 3'UTR alter *in vitro* and *in vivo* CYP2B6 activity: potential role of microRNAs

Kimberly S. Burgess[1,2], Joseph Ipe[2], Marelize Swart[2], Ingrid F. Metzger[2], Jessica Lu[2], Brandon T. Gufford[2], Nancy Thong[2], Zeruesenay Desta[2], Roger Gaedigk[3], Robin Pearce[3], Andrea Gaedigk[3], Yunlong Liu[4,5], Todd C. Skaar[2]

[1]Department of Pharmacology & Toxicology, [2]Department of Medicine, Division of Clinical Pharmacology, [4]Department of Medical & Molecular Genetics, [5]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN; [3]Division of Clinical Pharmacology, Toxicology and Therapeutic Innovation, Children's Mercy Kansas City, Kansas City, MO

Todd C. Skaar

950 W Walnut Street

R2 Building, Room 419

Indianapolis, IN 46202

+1 317 274 2821

+1 317 274 2704

tskaar@iu.edu

References: 34

Figures: 5

Tables: 2

Key words: CYP, pharmacogenetics, efavirenz, bupropion, polymorphism, SNP, metabolism

This is the author's manuscript of the article published in final edited form as:
Burgess, K. S., Ipe, J., Swart, M., Metzger, I. F., Lu, J., Gufford, B. T., Thong, N., Desta, Z., Gaedigk, R., Pearce, R., Gaedigk, A., Liu, Y. and Skaar, T. C. (), Variants in the CYP2B6 3'UTR alter in vitro and in vivo CYP2B6 activity: Potential role of microRNAs. Clin. Pharmacol. Ther.. Accepted Author Manuscript. http://dx.doi.org/10.1002/cpt.892

ABSTRACT:

*CYP2B6*6* and *CYP2B6*18* are the most clinically important variants causing reduced CYP2B6 protein expression and activity. However, these variants do not account for all variability in CYP2B6 activity. Emerging evidence has shown that genetic variants in the 3'UTR may explain variable drug response by altering microRNA regulation. Five 3'UTR variants were associated with significantly altered efavirenz $AUC_{0\text{-}48}$ (8-OH-EFV/EFV) ratios in healthy human volunteers. The rs70950385 (AG>CA) variant, predicted to create a microRNA binding site for miR-1275, was associated with 33% decreased CYP2B6 activity among normal metabolizers (AG/AG vs. CA/CA [p<0.05]). *In vitro* luciferase assays were used to confirm that the CA on the variant allele created a microRNA binding site causing a 11.3% decrease in activity compared to the AG allele when treated with miR-1275 (p=0.0035). Our results show that a 3'UTR variant contributes to variability in CYP2B6 activity.

INTRODUCTION:

CYP2B6 is a highly polymorphic cytochrome P450. Of these variants, the *CYP2B6*6* (516G>T, Q172H and 785A>G, K262R) and *CYP2B6*18* (983T>C, I328T) haplotypes, have been shown to be the most clinically important by causing a reduction in protein expression and activity [1-3]. The *CYP2B6*6* haplotype has been associated with altered elimination and response of many clinically-relevant drugs, such as efavirenz, bupropion, methadone, and cyclophosphamide [1, 4, 5]; however, these and other frequently studied genetic variants account for only a portion of the variability observed in the biotransformation of drugs by CYP2B6.

Efavirenz, a non-nucleoside reverse transcriptase inhibitor, is prescribed as part of the World Health Organization recommended first-line combination antiretroviral therapy to treat HIV type 1 infection [6]. However, efavirenz has been associated with central nervous system adverse events such as

hallucinations and insomnia and with treatment failure in some patients [7]. Higher efavirenz plasma concentrations have also been associated with increased likelihood of adverse CNS effects, whereas lower levels more likely result in treatment failure [8]. Efavirenz has also been shown to have wide variability in plasma concentrations among patients, some of which has been accounted for by genetic polymorphisms in *CYP2B6*, the enzyme primarily responsible for efavirenz biotransformation. Variability in CYP2B6-mediated biotransformation has been observed for other CYP2B6 substrates after the more well-studied polymorphisms have been accounted for, such as the variability in methadone metabolism and clearance [9].

Pharmacogenetics studies have focused primarily on genetic variants in the coding region of drug disposition genes. However, emerging evidence strongly suggests that genetic variants in the 3'UTR also contribute to variability in drug response by altering microRNA (miRNA) regulation [10, 11]. MicroRNAs, approximately 17-22 nucleotides in length, have been predicted to negatively regulate the expression of many genes including drug disposition genes, predominately through imperfect complimentary binding within the 3'UTR. Perfect binding of the seed sequence region, approximately seven nucleotides in length within the miRNA, has been the main focus for predicting miRNA regulation [12]. Variants within the seed sequence of miRNA binding sites may abolish miRNA binding sites, as well as create new miRNA binding sites [13]. These genetic variations in miRNA binding sites of CYPs may explain some of the inter-subject variability in pharmacokinetics and drug response.

In this study, we hypothesized that genetic variants in the 3'UTR of *CYP2B6* are associated with variability in efavirenz pharmacokinetics. To test this hypothesis, we identified genetic variants in the 3'UTR of *CYP2B6* of healthy volunteers that were administered efavirenz. Efavirenz has been shown to be a suitable probe drug substrate to measure CYP2B6 activity *in vitro* and *in vivo* [14]. $AUC_{0\text{-}48}$ and C_{max}

ratios (8-hydroxy-efavirenz/efavirenz) (8-OH-EFV/EFV) have been demonstrated to describe CYP2B6 activity over minor CYP pathways [15, 16]. Genetic variants that were associated with altered CYP2B6 activity were further characterized with *in vitro* functional assays to assess their impact on miRNA regulation of *CYP2B6* expression. Our results show that genetic variants in the 3'UTR of *CYP2B6* functionally alter miRNA regulation and are associated with variability in efavirenz pharmacokinetics.

RESULTS:

Association of *CYP2B6* 3'UTR SNPs with efavirenz AUC_{0-48} metabolite:parent ratios in healthy volunteers

Among 200 healthy volunteers sequenced for *CYP2B6* 3'UTR variants, 197 were also genotyped for *CYP2B6*6* and *CYP2B6*18*. *CYP2B6*6 and CYP2B6*18* genotyping data were used to classify those 197 volunteers as genotypic normal (47%), intermediate (42%), or poor (11%) metabolizers (Supplemental Table 1).

Twenty-five variants were observed in the 3'UTR of this population; however, only seven variants were present at a frequency that allowed us to assess pharmacokinetic differences among the genotypes (Table 1). These variants are not novel and have previously been reported in public databases such as The 1000 Genomes Project [17]. A schematic representation of the *CPY2B6* 3'UTR sequence and variant positions are provided in Supplemental Figure 1. Plasma efavirenz AUC_{0-48} ratios (8-OH-EFV/EFV) were available for 187 volunteers and were used to further describe the metabolic phenotype. The C_{max} ratios were in general agreement with the more robust AUC_{0-48} ratios and are only shown in the supplemental materials (Supplemental Table 2 and 3).

Among the normal metabolizers, an increase in efavirenz metabolism was detected among 3'UTR variants rs3181842, rs7246465 and rs707265 (Figure 1; Supplemental Table 2). For rs3181842, an increase in CYP2B6 activity was observed between volunteers with T/T versus T/C (31.9% increase; $p<0.05$) and C/C genotypes (70.6% increase; $p<0.0001$) (Figure 1A; Supplemental Table 2). An increase in CYP2B6 activity was also seen among volunteers carrying the rs7246465 variant allele (C/C vs. C/T [38.0% increase; $p<0.01$] and T/T [67.9% increase; $p<0.0001$]) and the rs707265 variant allele (G/G vs. G/A [34.2% increase; $p<0.05$] and A/A [72.4% increase; $p<0.0001$]) (Figure 1B, E; Supplemental Table 2). These two variants were in partial linkage disequilibrium in our study cohort ($r^2=0.80$, D'=0.91) (Figure 2).

The rs70950385 and rs1042389 variant alleles were associated with a decrease in CYP2B6 activity when comparing AUC between AG/AG and CA/CA genotypes (32.7%; $p<0.05$) (Figure 1F; Supplemental Table 2). The rs70950385 and rs1042389 variants were in complete linkage disequilibrium ($r^2=1.0$, D'=1.0) in our cohort and thus not shown separately (Figure 2). Comparable changes in efavirenz metabolism among the 3'UTR variants were also observed when normal, intermediate and poor metabolizers were combined (Supplemental Figure 2; Supplemental Table 3).

To assess the changes in efavirenz metabolism among combinations of variants, haplotype and diplotype analyses were performed. Significant decreases in CYP2B6 activity, as measured by AUC ratios, were observed between haplotype 4 versus 2 ($p<0.01$), 12 ($p<0.0001$), 13 ($p<05$), and 24 ($p<0.01$), respectively (Table 2). Haplotypes 2, 4, 12, 13, and 24 varied among the following 3'UTR variants: rs3181842, rs7260525, rs7246465, rs28969420, rs1038376, rs707265, rs70950385, and rs1042389. A significant overall change was also observed in AUC ratios between the different diplotypes, but no significant changes were detected after post-hoc analysis (data not shown). The sample size for many

haplotypes and diplotypes were too small to assess, i.e. many haplotypes and diplotypes were observed only within one of the metabolizer groups (data not shown).

3'UTR SNPs association with altered CYP2B6 activity in human liver microsomes

To further evaluate the changes found in our healthy volunteer population, DNA from 90 human liver microsomal preparations isolated from pediatric and adult liver tissue samples were sequenced and characterized for CYP2B6 activity by measuring bupropion hydroxylation (one sample had an undetectable bupropion hydroxylation rate). In this tissue panel, there were 51 genotypic normal, 31 intermediate, 4 poor metabolizers, and 4 liver samples whose *CYP2B6*6* and *CYP2B6*18* genotypes were unknown (data not shown). A similar decrease in CYP2B6 activity was observed among samples with the rs70950385 homozygous (AG/AG) genotype (543.3 pmol/mg protein/min) versus carriers of the CA allele (326.3 pmol/mg protein/min) in the panel (39.9% decrease; $p<0.05$ based on untransformed data) (Figure 3 – $\log_2$-transformed velocities for visualization purposes; Supplemental Table 4). Differences were not significant in genotypic normal metabolizers, likely because there were too few samples in this subset. The rs70950385 and rs1042389 variants were in complete linkage disequilibrium in this sample population ($r^2=1.0$, D=1.0) (Figure 2). There were no significant changes for any of the other variants, haplotypes, and diplotypes, as their frequencies were low (Supplemental Figure 3 and Table 2).

miR-1275 alters luciferase activity in a plasmid containing the rs70950385 (CA) variant allele

The rs70950385 variant (AG>CA), a combination of rs12979270 (A>C) and rs12979898 (G>A), was found to be in complete linkage disequilibrium in our total population. The rs12979270 (A>C) was predicted by PolymiRTS to create miRNA binding site miR-625-5p (Figure 4A) through perfect complimentary alignment within the seed sequence region, nucleotides 2-8 from the 5' end of the miRNA. Manual seed

sequence comparisons revealed the dinucleotide variant created an additional base pair match for the first nucleotide of miR-1275 (Figure 4B). When this mechanism was tested in a luciferase transfection model in HepG2 cells without co-transfection of predicted miRNAs, there were no differences in luciferase activity between the plasmid containing the wild-type and variant miRNA binding sites with *C. elegans* negative control miRNA or controls without miRNA transfected (Figure 5A). However, when transfected with the miRNAs whose binding was predicted to be altered by the variant, miR-1275 reduced luciferase activity by 11.3% (p=0.0035) in the presence of the variant allele compared to the wild-type allele (Figure 5B). MicroRNA 625-5p did not impact luciferase activity.

DISCUSSION:

Variability in the pharmacokinetics of efavirenz and other drugs that undergo CYP2B6-mediated metabolism has been attributed to genetic variants in the coding region; however, variability in clinical responses and adverse events are not completely explained by these genetic differences. Data from this study indicate that genetic variations in the 3'UTR of *CYP2B6* have an effect on CYP2B6 activity, independent of CYP2B6 genotypic metabolizer status (assigned based on *CYP2B6*6* and **18*). There were 25 unique variants detected in the 3'UTR of *CYP2B6* in our study population. Of these variants, seven were observed at a frequency that allowed us to assess differences in efavirenz pharmacokinetics between the homozygous and heterozygous volunteers. The rs3181842, rs7246465, and rs707265 had similar increases in CYP2B6 activity to one another for homozygous volunteers (Figure 1, panels A, B, and E). These three variants were not predicted to remove or create a miRNA binding site through the PolymiRTS database at the time the database was accessed. Although the mechanism behind these observed changes are unknown, the mechanism may be related to only one of the variants as there is partial linkage disequilibrium among the three variants. The rs3181842 vs. rs7246465, rs3181842 vs. rs707265, and rs7246465 vs. rs707265 have r^2= 0.58, 0.58, and 0.80 values respectively (Figure 2).

The rs70950385 and rs1042389 variants were shown to be in complete LD in both our volunteers and human liver microsomes. These two variants were associated with decreased CYP2B6 activity with respect to efavirenz pharmacokinetics in subjects genotyped as normal metabolizers. They were also associated with decreased rates of bupropion hydroxylation in the panel of human liver microsomes. The rs1042389 variant has previously been associated with a decrease in median efavirenz plasma concentrations among black South African HIV/AIDS patients in a study that assessed the effect of 3'UTR variants on efavirenz response [10]. The rs70950385 and rs1042389 SNPs were shown to be in complete LD in this population as well. We propose that some of the observed decrease in CYP2B6 activity is mediated by miRNA regulation owing to the rs70950385 variant creating a miRNA binding site for miR-1275. This hypothesis is further corroborated by our *in vitro* data (Figure 5). In HepG2 cells treated with miR-1275 that expressed the variant (CA) miRNA binding site, reduced luciferase activity was observed compared to cells that expressed the reference (wild-type) (AG) miRNA binding site. Previously, we showed that miR-1275 expression is increased in pediatric and adult human livers compared to fetal livers, unequivocally demonstrating that this miRNA is expressed in the human liver [18]. These data suggest that this miRNA may play a role in regulating *CYP2B6* activity; however, further studies are warranted to validate that this is the underlying mechanism as opposed to other mechanisms, such as variant-induced structural changes in the mRNA that may affect its stability.

No significant changes in efavirenz pharmacokinetics were observed among the 3'UTR variants within the genetic intermediate and poor metabolizers in our human volunteer population for any variant with the exception of rs7246465 (C/T vs. T/T; $p<0.05$) (data not shown). All other variants yielded similar trends. This is likely due to inadequate sample numbers among the intermediate and poor metabolizer groups to detect these changes. Alternatively, this observation may be due to a limited effect size of the

3'UTR variants in the presence of coding region variants known to reduce CYP2B6 activity. The rs70950385 homozygous variant and other genotypes were only detected in normal metabolizers in our population. Interestingly, many 3'UTR haplotypes and diplotypes were unique to the metabolizer status. This may be due to partial LD with the *CYP2B6*6* and *CYP2B6*18* alleles or that our sample size was too small to detect an overlap.

This study shows that 3'UTR variants impact drug metabolism. Although the mechanisms are not known for all altered efavirenz pharmacokinetics associated with 3'UTR variants in *CYP2B6*, miRNA regulation is a likely contributor. Limitations on linking the other variants to miRNAs may be due to the limitations of miRNA prediction algorithms and tools which focus mostly on the seed sequence region as the main mechanism for miRNA binding. As in the case of rs70950385, our data suggest that miRNA regulation may explain, at least in part, the decreased CYP2B6 activity in healthy volunteers carrying the variant allele, as corroborated by our *in vitro* data. Although pharmacogenetics has primarily focused on genetic variants in the coding and promoter regions, our findings, i.e. that genetic variants in the 3'UTR contribute to variability in drug response, will open up the field to consider 3'UTR variants as a source of variability in the activities of many drug disposition genes.

MATERIALS AND METHODS:

CYP2B6 Activity in Healthy Volunteers

Data from 200 healthy non-pregnant female (n=83) and male (n=117) volunteers (18-55 years old) administered a single of efavirenz dose, a probe drug substrate of CYP2B6 activity, was obtained from the single dose phase of three previous studies, published and unpublished [19, 20]. These studies were approved by the Indiana University School of Medicine Institutional Review Board, conducted at the

Indiana University School of Medicine Clinical Research Center and registered at http://www.clinicaltrials.gov (trial identifiers NCT00668395, NCT01104376, and NCT02401256). After written informed consent, volunteers were thoroughly screened using medical history, physical examination, and laboratory tests such as electrocardiography, HIV test, urinalysis and blood tests. During this screen, venous blood was drawn for DNA isolation. Dietary restrictions and inclusion and exclusion criteria have been previously reported [21, 22].

These open label studies were retrospectively used to evaluate the association of single dose efavirenz pharmacokinetics with the *CYP2B6* 3'UTR variants. Briefly, plasma samples were taken at pre-dose and 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 16, 24, 48, 72 and 144 hours after single dose efavirenz administration. Participants from two studies (127 participants total) were administered 600mg of efavirenz and 73 participants from another study were administered 100mg of efavirenz (Sustiva, Bristol-Myers Squibb). Plasma efavirenz and 8-hydroxy-efavirenz concentrations were measured using a validated LC-MS/MS protocol as described previously [19]. Non-compartmental analysis of data was performed using Phoenix® WinNonlin® (version 7.0, Pharsight Corp., Cary, NC) to determine C_{max} (maximal plasma concentration) and $AUC_{0\text{-}48}$ (area under the plasma concentration time curve from zero to 48 hours). C_{max} was assessed directly from the concentration-time profile and $AUC_{0\text{-}48}$ was determined using the trapezoidal rule with linear up/log down interpolation.

CYP2B6 Activity in Human Liver Microsomes (HLMs)

HLMs were prepared by differential centrifugation, as described by Lu and Levin (1972) [23]. *In vitro* microsomal incubations were performed according to Pearce et al 2015 [24]. Briefly, HLMs (30µg of microsomal protein), potassium phosphate buffer (50mM, pH 7.4), MgCl2 (3mM), EDTA (1mM), and bupropion (500µM) were incubated at the final concentrations listed. Reactions were initiated by the

addition of a NADPH-generating system, consisting of NADP (1mM), glucose-6-phosphate (5mM), and glucose-6-phosphate dehydrogenase (1 U/ml), incubated at 37°C in a Thermo Forma Benchtop Orbital Shaker (Marietta, OH) and terminated after 30 minutes by the addition of 50µL of ice-cold acetonitrile. Incubations were performed in triplicate. Bupropion stock solutions were dissolved in methanol; however, the concentration of methanol present in the incubation mixtures did not exceed 0.25%. Hydroxybupropion and bupropion were resolved by isocratic, reversed-phase high-performance liquid chromatography based on a modification of the method by Faucette et al. 2000 [25] as described previously [24].

Genotyping and Sequencing in Healthy Volunteers

Genomic DNA extracted from whole blood was used for genotyping and Sanger sequencing of the 3'UTR. *CYP2B6* genotyping was performed using TaqMan Genotyping Assays for rs3745274 (516G>T, Q172H) and rs28399499 (983T>C, I328T) (Life Technologies, Foster City, CA) according to the protocol provided. PCR was performed on BioRad iCycler and QuantStudio 12K Flex real-time PCR instruments. *CYP2B6* rs2279343 (785A>G, K262R) genotyping was performed by first amplifying exon 5 with primers 5'-CTCTCTCCCTGTGACCTGCTA-3' (forward) and 5'-CTCCCTCTGTCTTTCATTCTGTC-3' (reverse) (Integrated DNA Technologies, Coralville, IA) as described by Lang et al. [26], then using 1uL exon 5 PCR amplification product as a template for a custom TaqMan Genotyping Assay as previously described [21].

To sequence the 3'UTR region, this region was first amplified by adding the following reagents to each reaction: 1µL of DNA (10 ng), 12µL CloneAmp HiFi PCR Premix (CloneTech, Mountain View, CA), 10µL ultrapure distilled water and 1µL of each primer (final concentration: [0.4µM]) (Integrated DNA Technologies, Coralville, IA). The primer sequences used were 5'-GGCAAAATACCCCCAACATA-3'

(forward) and 5'-AGAGTTGGCATTGAGGTGAGAG-3' (reverse) [10]. The PCR amplification conditions were 98°C for 5 minutes, 40 cycles at 98°C for 10 seconds, 69.5°C for 15 seconds, and 72°C for 2 minutes. Final extension was at 72°C for 7 minutes. PCR purification was performed using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) according to the user manual. DNA concentrations were measured with a Qubit dsDNA BR Assay Kit (Invitrogen, Eugene, OR) according to the user manual. Samples were prepared for sequencing using 10µL of DNA (50 ng) and 2µL of one of four different primers [1.7µM] and Sanger sequenced by ACGT (Germantown, MD). The 1.8 kb PCR product was sequenced in its entirety using the primers mentioned above and forward and reverse primers 5'-CTAAGCCTTGCTCTGTCTCC-3' (forward) and 5'-GGAGAATCACTTGAACCCAGG-3' (reverse).

Genotyping and Sequencing in Human Liver Tissue Samples

Genomic DNA extracted from 90 human liver samples was used for genotyping and Sanger sequencing. Genotyping was performed on Applied Biosystems 7900 HT Real-Time PCR System or QuantStudio 12K Flex real-time PCR instruments using Life Technologies TaqMan assays for *CYP2B6* rs34223104 (282T>C); rs3745274 (516G>T, Q172H), rs28399499 (983T>C, I328T), and rs3211371 (1459C>T, R487C) (Life Technologies, Foster City, CA). To detect *CYP2B6* rs2279343 (785A>G, K262R) a high-resolution melting (HRM) assay was performed as previously described [27].

To sequence the 3'UTR, this region was first amplified by adding the following reagents to each reaction: 1µL of DNA (15 ng), 4µL KAPA LR Hotstart 2X, 2.4µL ultrapure distilled water and 0.3µL of each primer (final concentration: [0.375µM]) (Integrated DNA Technologies, Coralville, IA). The primer sequences used were 5'-AATCTGTTGCAGTGGACATTTG-3' (forward) and 5'-AGAGTTGGCATTGAGGTGAGA-3' (reverse). The PCR amplification conditions were 94°C for 3 minutes, 48 cycles at 94°C for 10 seconds, 61.5°C for 15 seconds, and 68°C for 2 minutes. Final extension was 68°C for 7 minutes. PCR purification

was performed using ExoSAP-IT (USB, Cleveland, OH). A 1:20 dilution of the purified PCR product was used for sequencing, performed with BigDye terminator v3.1 and the STeP cycle sequencing protocol [28]. Briefly, three cycle steps were performed: 1) at 95°C for 10 sec, 50°C for 5 sec and 60°C for 1 min 15 sec for a total 15 cycles, 2) at 95°C for 10 sec, 50°C for 5 sec and 60°C for 1 min 30 sec, a total of 5 cycles and 3) at 95°C for 10 sec, 50°C for 5 sec and 60°C for 2 min for 15 cycles. The sequence reactions were cleaned up with AxyPrep Mag DyeClean (Axygen, Corning Life Sciences, Tewksbury, MA) and electrophoresis was performed on a 3730xl DNA Analyzer (Life Technologies, Foster City, CA). The 1962 bp PCR product was sequenced in its entirety using the following forward primers 5'-TGTGTCTGGGCTTAGGGAC-3', 5'-GTGATTCACCCACCTTAGCC-3', 5'-GTCTCAGCTCCCAAGTAGCTG-3', and the reverse primers 5'-CAGGTCTCTCAGAGGCAGG-3', 5'-AGTTCGAGACCAGCCTGG-3', 5'-CCAGCTCCTCCAGAGGC-3' and 5'-GCAGCTGGGAATGAGGA-3'.

As a quality control, genotyping was repeated in random samples and genotyping and sequencing results were verified by a second person to ensure accurate interpretation of the data. Sequences were analyzed using Sequencher (v.5.4.5, Gene Codes Corporation, Ann Arbor, MI) and BioEdit (v.7.2.5)[29] after alignment to the CYP2B6 reference sequence with GenBank Accession Number NC_000019.10.

CYP2B6 **Allele, Genotype, Haplotype, and Diplotype Analyses**

CYP2B6 star allele designations were assigned in accordance with the Human Cytochrome P450 Allele Nomenclature Database (http://cypalleles.ki.se/). Allele and genotype frequencies, Hardy-Weinberg equilibrium analysis, and linkage disequilibrium analysis were performed using SHEsis [30]. Haplotype construction was performed using PHASE version 2.1.1 [31, 32]. Haplotype and diplotype Venn diagrams were generated using BioVenn [33].

***In Silico* MicroRNA Predictions**

The PolymiRTS Database 3.0 accessed June 2014 was used to identify SNPs predicted to create or abolish miRNA binding sites for *CYP2B6* [13]. For the dinucleotide variant, rs70950385, manual seed sequence alignment was performed.

***In Vitro* Luciferase Assay**

The pIS-0 firefly luciferase vector [34] (Addgene, Cambridge, MA) was used to study 3'UTR function in relation to miRNA regulation. The oligo containing either the wild-type (5'-GCCGTGTAATTCTAGGAGCTCGTCACACACTGCTGTAGTCTTCCCC**AG**TCCTCATCGTTCTAGAGTCGGGGC-3') or variant (5'-GCCGTGTAATTCTAGGAGCTCGTCACACACTGCTGTAGTCTTCCCC**CA**TCCTCATCGTTCTAGAGTCGGGGC-3') miRNA binding sites for the rs70950385 SNP (Integrated DNA Technologies, Coralville, IA) were amplified and then cloned into the pIS-0 vector 3'UTR using the NEBuilder HiFi DNA assembly system as instructed by the user manual (New England BioLabs, Ipswich, MA). The plasmids were then transformed into One Shot OmniMax 2 TI Chemically Competent *E. coli* (Invitrogen, Carlsbad, CA) and plated on ampicillin-treated agar plates. Individual colonies were selected and grown in ampicillin-treated liquid culture overnight. Plasmid DNA was isolated using a NucleoSpin Plasmid (NoLid) DNA Purification Kit (Macherey-Nagel, Düren, Germany) and quantified using a DNA dsDNA BR Assay Kit (Invitrogen, Eugene, OR). DNA was subsequently sequenced using the 5'-GTGGTTTGTCCAAACTCATC-3' reverse primer to ensure clones with the correct wild-type and variant sequences were selected.

HepG2 human liver carcinoma cells were seeded at a density of 90,000 cells per well in 24 well plates and cultured for 48 hours. HepG2 cells were then transfected using Lipofectamine 3000 (Life Technologies, Foster City, CA) for 48 hours with each plasmid (500ng/well) with or without the miRNAs

(hsa-miR-625-5p and hsa-miR-1275; final concentration: [30pM]) (GE Dharmacon, Chicago, IL) predicted to target the *CYP2B6* 3'UTR region encompassing the genetic variant. *Renilla* luciferase (10ng/well) was used as the transfection control. *C. elegans* cel-miR-67 was employed as a negative control. Luciferase activity was measured with the Dual-Glo Luciferase Assay as instructed by the user manual (Promega, Madison, WI).

Statistical Analysis

Efavirenz $AUC_{0\text{-}48}$ and C_{max} metabolite:parent ratios, and bupropion hydroxylase activities (untransformed data) were analyzed by genotype, haplotype, or diplotype using Kruskal-Wallis followed by Dunn's post-hoc tests. Multiple comparisons within each SNP group were accounted for by Dunn's post-hoc tests. Rather than increasing the stringency of our statistical test to reduce the false positive findings resulting from testing multiple *CYP2B6* variants, we followed up our findings with a human liver microsome population study and *in vitro* miRNA functional studies. *In vitro* luciferase assay results were analyzed using paired ratio t-tests of Firefly/*Renilla* luciferase activity ratios. All data analyses were performed using Graphpad Prism version 6 (La Jolla, CA); *p*-values <0.05 were considered statistically significant.

STUDY HIGHLIGHTS:

What is the current knowledge on the topic?

Some of the wide pharmacokinetic variability among efavirenz and other drugs metabolized by CYP2B6 have been accounted for by the *CYP2B6*6* and **18* allelic variants; however, unexplained inter-individual variability still remains. Emerging evidence implicates genetic variants in the 3'UTR to impact variability in drug metabolism by altering microRNA regulation.

What question does this study address?

We hypothesized that genetic variants in the 3'UTR contribute to the wide variability observed in CYP2B6 activity through microRNA regulation. We sequenced the 3'UTR of healthy human volunteers previously administered efavirenz and functionally validated the variants *in vitro*. Findings are supported by data obtained from a panel of human liver tissue samples.

What this study adds to our knowledge.

Genetic variants in the 3'UTR contributes to the variability in efavirenz metabolism among genotypic normal CYP2B6 metabolizers. The observed decrease in CYP2B6 activity in rs70950385 carriers of the variant allele appears to be related to microRNA regulation by miR-1275.

How this might change clinical pharmacology or translational science.

This study demonstrates an association between altered CYP2B6 activity and genetic variants in the 3'UTR. These findings help to further explain the variability in CYP2B6 activity among patients after other well-studied variants have been accounted for. Inclusion of this SNP into test panels may improve the prediction of CYP2B6 activity from genotype.

ACKNOWLEDGEMENT:

This study was supported by funding from NIH/NIGMS F31-GM119401 (Burgess), Clinical Pharmacology Training Grant T32-GM008425 (Gufford), R01-GM078501 (Desta), R01-GM088076 (Skaar) and the Vera Bradley Foundation for Breast Cancer (Swart, Ipe). The National Institute of Child Health and Human Development Brain and Tissue Bank for Developmental Disorders at the University of Maryland is funded by the National Institutes of Health (NIH) contract HHSN275200900011C, reference number, N01-HD-9-0011 and the Liver Tissue Cell Distribution System is funded by NIH contract number N01-DK-7-0004/HHSN267200700004C.

CONFLICT OF INTEREST:

None.

AUTHOR CONTRIBUTIONS:

K.S.B., J.I., M.S., B.T.G., Z.D., R.P., A.G., R.G., and T.C.S wrote the manuscript; K.S.B., Z.D., A.G., and T.C.S. designed the research; K.S.B., M.S., B.T.G., N.T., J.L., I.F.M., R.P., and R.G. performed the research; and K.S.B., J.I, M.S., Y.L., and B.T.G. analyzed the data.

REFERENCES:

1. Desta, Z., et al., *Impact of CYP2B6 polymorphism on hepatic efavirenz metabolism in vitro.* Pharmacogenomics, 2007. **8**(6): p. 547-58.
2. Klein, K., 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): p. 861-73.
3. Hofmann, M.H., 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): p. 284-92.
4. Eap, C.B., et al., *Stereoselective block of hERG channel by (S)-methadone and QT interval prolongation in CYP2B6 slow metabolizers.* Clin Pharmacol Ther, 2007. **81**(5): p. 719-28.
5. Nakajima, M., et al., *Genetic polymorphisms of CYP2B6 affect the pharmacokinetics/pharmacodynamics of cyclophosphamide in Japanese cancer patients.* Pharmacogenet Genomics, 2007. **17**(6): p. 431-45.
6. Organization, W.H., *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach.* 2016.
7. Adkins, J.C. and S. Noble, *Efavirenz.* Drugs, 1998. **56**(6): p. 1055-64; discussion 1065-6.
8. Marzolini, C., et al., *Efavirenz plasma levels can predict treatment failure and central nervous system side effects in HIV-1-infected patients.* AIDS, 2001. **15**(1): p. 71-5.
9. Kharasch, E.D., et al., *Methadone Pharmacogenetics: CYP2B6 Polymorphisms Determine Plasma Concentrations, Clearance, and Metabolism.* Anesthesiology, 2015. **123**(5): p. 1142-53.
10. Swart, M. and C. Dandara, *Genetic variation in the 3'-UTR of CYP1A2, CYP2B6, CYP2D6, CYP3A4, NR1I2, and UGT2B7: potential effects on regulation by microRNA and pharmacogenomics relevance.* Front Genet, 2014. **5**: p. 167.
11. Ramamoorthy, A., et al., *In silico and in vitro identification of microRNAs that regulate hepatic nuclear factor 4alpha expression.* Drug Metab Dispos, 2012. **40**(4): p. 726-33.
12. Liu, J., *Control of protein synthesis and mRNA degradation by microRNAs.* Curr Opin Cell Biol, 2008. **20**(2): p. 214-21.
13. Bhattacharya, A., J.D. Ziebarth, and Y. Cui, *PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways.* Nucleic Acids Res, 2014. **42**(Database issue): p. D86-91.

14. Administration, F.a.D., *Guidance for industry: drug interaction studies-study design, data analysis, implications for dosing, and labeling recommendations.* Center for Drug Evaluation and Research (CDER), Rockville, 2012.
15. Xu, C., et al., *CYP2B6 pharmacogenetics-based in vitro-in vivo extrapolation of efavirenz clearance by physiologically based pharmacokinetic modeling.* Drug Metab Dispos, 2013. **41**(12): p. 2004-11.
16. Jiang, F., et al., *Effects of clopidogrel and itraconazole on the disposition of efavirenz and its hydroxyl metabolites: exploration of a novel CYP2B6 phenotyping index.* Br J Clin Pharmacol, 2013. **75**(1): p. 244-53.
17. Genomes Project, C., et al., *A global reference for human genetic variation.* Nature, 2015. **526**(7571): p. 68-74.
18. Burgess, K.S., et al., *Age-related changes in microRNA expression and pharmacogenes in human liver.* Clin Pharmacol Ther, 2015.
19. Robarge, J.D., et al., *Population Pharmacokinetic Modeling To Estimate the Contributions of Genetic and Nongenetic Factors to Efavirenz Disposition.* Antimicrob Agents Chemother, 2017. **61**(1).
20. Desta, Z., et al., *Inhibition of Cytochrome P450 2B6 Activity by Voriconazole Profiled Using Efavirenz Disposition in Healthy Volunteers.* Antimicrob Agents Chemother, 2016. **60**(11): p. 6813-6822.
21. Michaud, V., et al., *Efavirenz-mediated induction of omeprazole metabolism is CYP2C19 genotype dependent.* Pharmacogenomics J, 2014. **14**(2): p. 151-9.
22. Michaud, V., et al., *Induction of CYP2C19 and CYP3A activity following repeated administration of efavirenz in healthy volunteers.* Clin Pharmacol Ther, 2012. **91**(3): p. 475-82.
23. Lu, A.Y. and W. Levin, *Partial purification of cytochromes P-450 and P-448 from rat liver microsomes.* Biochem Biophys Res Commun, 1972. **46**(3): p. 1334-9.
24. Pearce, R.E., et al., *Developmental Expression of CYP2B6: A Comprehensive Analysis of mRNA Expression, Protein Content and Bupropion Hydroxylase Activity and the Impact of Genetic Variation.* Drug Metab Dispos, 2016. **44**(7): p. 948-58.
25. Faucette, S.R., et al., *Validation of bupropion hydroxylation as a selective marker of human cytochrome P450 2B6 catalytic activity.* Drug Metab Dispos, 2000. **28**(10): p. 1222-30.
26. Lang, T., et al., *Extensive genetic polymorphism in the human CYP2B6 gene with impact on expression and function in human liver.* Pharmacogenetics, 2001. **11**(5): p. 399-415.
27. Twist, G.P., et al., *High-resolution melt analysis to detect sequence variations in highly homologous gene regions: application to CYP2B6.* Pharmacogenomics, 2013. **14**(8): p. 913-22.
28. Platt, A.R., R.W. Woodhall, and A.L. George, Jr., *Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol.* Biotechniques, 2007. **43**(1): p. 58, 60, 62.
29. Hall, T.A. *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT*. in *Nucleic acids symposium series*. 1999.
30. Shi, Y.Y. and L. He, *SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci.* Cell Res, 2005. **15**(2): p. 97-8.

31. Stephens, M. and P. Donnelly, *A comparison of bayesian methods for haplotype reconstruction from population genotype data.* The American Journal of Human Genetics, 2003. **73**(5): p. 1162-1169.
32. Stephens, M., N.J. Smith, and P. Donnelly, *A new statistical method for haplotype reconstruction from population data.* The American Journal of Human Genetics, 2001. **68**(4): p. 978-989.
33. Hulsen, T., J. de Vlieg, and W. Alkema, *BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams.* BMC Genomics, 2008. **9**: p. 488.
34. Yekta, S., I.H. Shih, and D.P. Bartel, *MicroRNA-directed cleavage of HOXB8 mRNA.* Science, 2004. **304**(5670): p. 594-6.

FIGURE LEEGENDS:

Figure 1: Association of *CYP2B6* 3'UTR variants with efavirenz metabolism in healthy volunteers with *CYP2B6* genotypes predicting normal metabolism

Efavirenz AUC metabolite:parent ratios (8-OH-EFV/EFV) among normal metabolizers; subjects with *CYP2B6*6* and **18* alleles excluded. *X-axis*: *CYP2B6* 3'UTR genotypes **F.** rs70950385 variant is in complete LD with rs1042389 (data not shown separately). *Y*-axis: 8-hydroxy-efavirenz/efavirenz ratio. (*=p<0.05, **=p<0.01, ****=p<0.0001)

Figure 2: r^2 linkage disequilibrium diagram of *CYP2B6* variants genotyped in samples from all four studies (3 clinical trials and one liver sample set)

The r^2 values for *CYP2B6* 3'UTR variants and *CYP2B6*6* (SNP 1 and 2) and *CYP2B6*18* (SNP 3).

Figure 3: Association of *CYP2B6* rs70950385 with bupropion metabolism in human liver microsomes

Bupropion hydroxylation activity observed in human liver microsomes among **A.** all samples and **B.** normal metabolizers. *X-axis*: Genotype (rs70950385 in complete linkage disequilibrium with rs1042389). *Y-axis*: $\log_2$ transformed rate of hydroxyl-bupropion (pmol/mg protein/min). (*=p<0.05) Note: Data was log-transformed for visual purposes only.

Figure 4: rs70950385 creates miR-625-5p and miR-1275 binding site in *CYP2B6* 3'UTR

Proposed mechanism for miRNA **A.** 625-5p and **B.** 1275 binding to the *CYP2B6* 3'UTR among carriers of the rs70950385 (CA) allele.

Figure 5: miRNA regulation of *CYP2B6* 3'UTR *in vitro*

MicroRNA regulation of CYP2B6 expression through measurement of luciferase activity in pIS-0 plasmids containing either the reference (wild-type) or variant miRNA binding sites. **A.** *X-axis*: Treatment groups (none, *C. elegans* control, miR-1275, miR-625). *Y-axis*: Wild-type (firefly/*Renilla* luciferase) / Variant (firefly/*Renilla* luciferase). A ratio of 1=no change, >1=wild-type had higher luciferase activity, <1=variant had higher luciferase activity. **B.** *X-axis:* Wild-type and variant plasmids plus miR-1275. *Y-axis*: Firefly/*Renilla* luciferase ratio. (*=p<0.05)

Supplemental Figure 1: Schematic representation of the *CYP2B6* 3'UTR sequence. Variant positions are underlined. * 3'UTR variants associated with changes in efavirenz metabolism in healthy volunteers with *CYP2B6* genotypes predicting normal metabolism.

Supplemental Figure 2: Association of *CYP2B6* 3'UTR variants with efavirenz metabolism in healthy volunteer population (n=187). Efavirenz $AUC_{0\text{-}48}$ metabolite:parent ratios among all volunteers. *X-axis*: *CYP2B6* 3'UTR genotypes **F.** rs70950385 variant is in complete LD with rs1042389 (data not shown separately). *Y*-axis: 8-hydroxy-efavirenz/efavirenz ratio. (**=p<0.01, ***=p<0.001, ****=p<0.0001)

Supplementary Figure 3: Association of *CYP2B6* 3'UTR variants with bupropion metabolism in human liver microsomes (n=89). Bupropion hydroxylation rates among normal metabolizers; subjects with *CYP2B6*6* and **18* alleles excluded. *X-axis*: *CYP2B6* 3'UTR genotypes. **F.** rs70950385 variant is in complete LD with rs1042389 (data not shown separately). *Y*-axis: $\log_2$ transformed bupropion hydroxylation rate (pmol/mg protein/min).

Supplemental Table 1: CYP2B6 Star Allele Frequency

Supplemental Table 2: Association of *CYP2B6* 3'UTR variants with efavirenz metabolism in healthy volunteers with *CYP2B6* genotypes predicting normal metabolism (n=86)

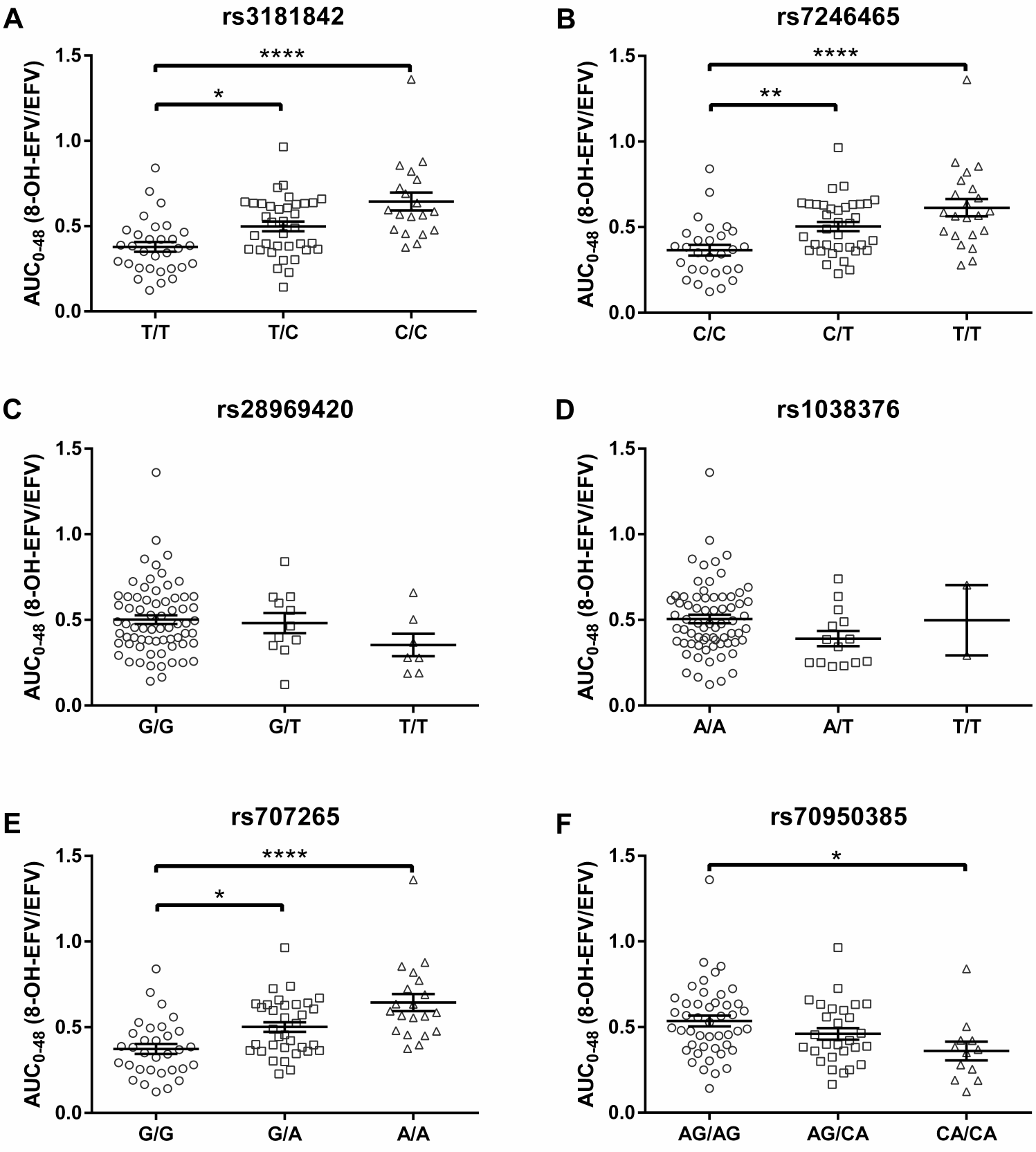
Supplemental Table 3: Association of *CYP2B6* 3'UTR variants with efavirenz metabolism in our total healthy volunteer population (n=187)

Supplemental Table 4: Association of *CYP2B6* 3'UTR variants with bupropion metabolism in human liver microsomes (n=89)

Table 1: *CYP2B6* 3'UTR variant frequencies in healthy volunteers

SNP	Chromosome 19 Position[1]	MAF	HWE	Genotype (n)		
rs3181842	41017111	0.343	Yes	T/T (89)	T/C (85)	C/C (26)
rs7246465	41017398	0.323	No	C/C (99)	C/T (73)	T/T (28)
rs28969420	41017899	0.050	No	G/G (166)	G/T (24)	T/T (10)
rs1038376	41018104	0.333	Yes	A/A (89)	A/T (89)	T/T (22)
rs707265	41018182	0.318	Yes	G/G (101)	G/A (71)	A/A (28)
rs70950385	41018226	0.228	Yes	AG/AG (122)	AG/CA (65)	CA/CA (13)
rs1042389	41018248	0.228	Yes	C/C (122)	C/T (65)	T/T (13)

[1] Chromosome position is according to GenBank Accession Number NC_000019.10 [GRCh37.p17 108 annotation]



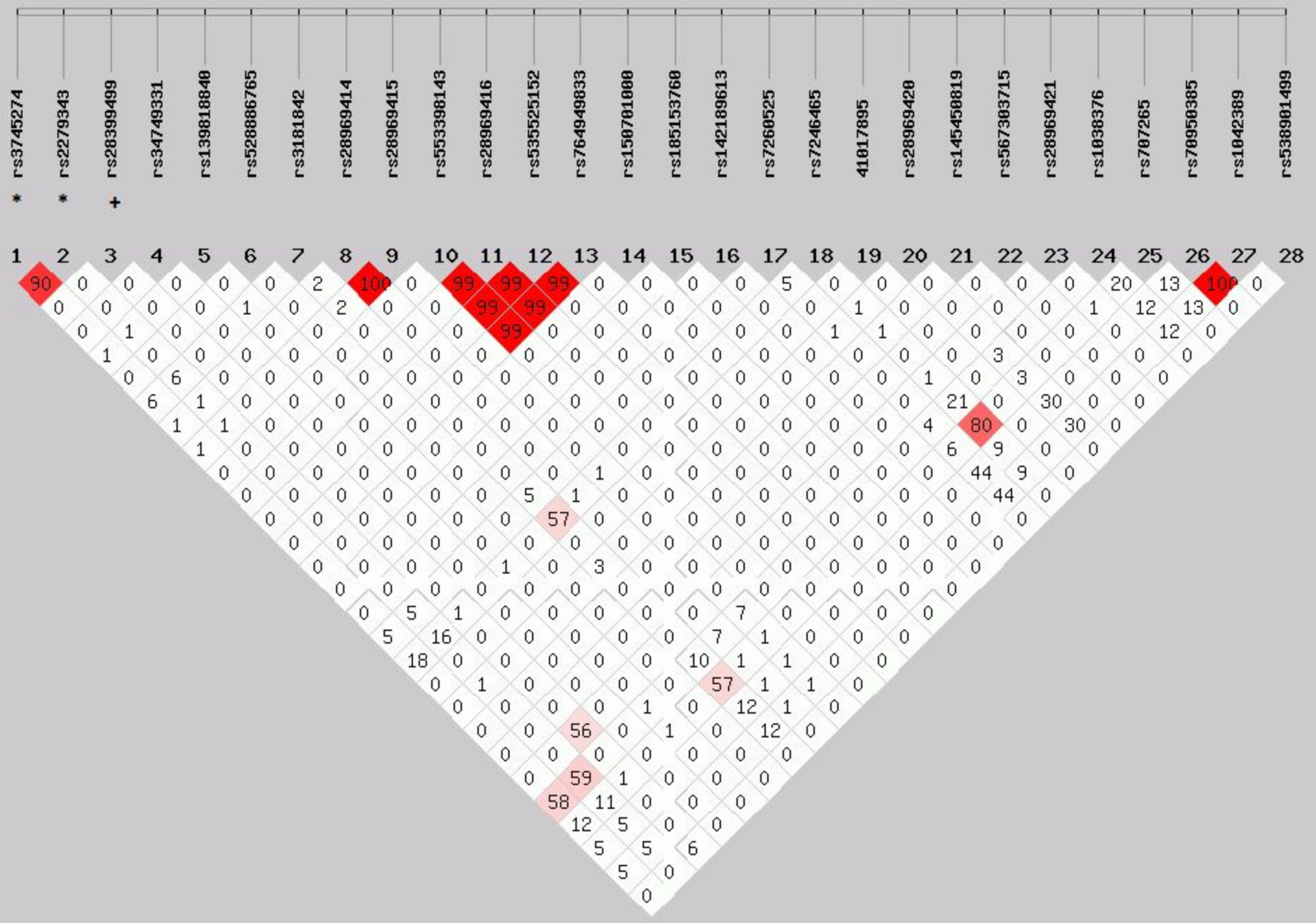


Table 2: *CYP2B6* 3'UTR variant haplotype association with CYP2B6 activity in healthy volunteers and human liver microsomes

Haplotype ID	Haplotype**	Count	Mean AUC_{0-48} (8-OH-EFV/EFV)	Mean OH-BUP Rates (pmol/mg protein/min)
1	G7CCTACACAAGGACCGCCGAGAGTA	7	0.33±0.19	265.7±167.6
2*	G7CCTACACAAGGACCGCCGTGAGTA	25	0.33±0.03*	453.5±375.3
4*	G7CCTACACAAGGATCGCCGAAAGTA	128	0.55±0.03*	572.0±135.8
9	G7CTTACACAAGGACCGCCGAGAGTA	62	0.40±0.03	480.8±104.4
12*	G7CTTACACAAGGACCGCCGTGAGTA	128	0.38±0.02*	512.1±106.1
13*	G7CTTACACAAGGACCTCCGAGCACA	50	0.39±0.03*	153.3±61.7
24*	G7CTTACACAAGGGCCGCCGAGCACA	65	0.40±0.03*	403.9±137.6
34	G7CTAGCACAAGGACCGCCGTGAGTA	22	0.37±0.05	342.4±168.6
37	G8CCTACACAAGGATCGCCGAAAGTA	32	0.51±0.05	444.8±162.7

***** Haplotypes that are statistically significant versus haplotype 4 when comparing AUC ratios.

****** The order of variants composing the haplotypes are corresponding with the individual variants listed in Table 1 as follows: rs34749331, rs139818840, rs528886765, rs3181842, rs28969414, rs28969415, rs553398143, rs28969416, rs535525152, rs764949833, rs150701000, rs185153760, rs142189613, rs7260525, rs7246465, novel variant, rs28969420, rs145450819, rs567303715, rs28969421, rs1038376, rs707265, rs70950385 (AG>CA), rs1042389, rs538901499.