Comparison of Two CYP2D6 Genotyping Methods and Assessment of Genotype-Phenotype Relationships

WEN-HWEI CHOU,1* FENG-XIANG YAN,1† DORIS K. ROBBINS-WEILERT,2 THOMAS B. RYDER,3 WEI WEI LIU,3 CLOTILDE PERBOST,3 MAUREEN FAIRCHILD,4 JOSÉ DE LEÓN,5 WALTER H. KOCH,4 and PETER J. WEDLUND1‡

Background: There have been no published reports comparing the CYP450 GeneChip® microarray assay with more standard methods of genetic testing.

Methods: We collected 20-mL blood samples from 236 volunteers for DNA isolation and testing before each individual ingested 60 mg of dextromethorphan, and collected their urine. CYP2D6 alleles *3 to *7, *9, *17, and *41, and multiple CYP2D6 gene copies were tested by allele-specific PCR (AS-PCR), whereas alleles *2 to *4 and *6 to *11 were tested by the Affymetrix CYP450 GeneChip assay. Five of the CYP2D6 alleles (*3, *4, *6, *7, and *9) were tested by both AS-PCR and the CYP450 GeneChip assay in an independent and blinded fashion in 232 of the 236 healthy volunteers. The combined CYP2D6 genotype from both methods was used to divide the population into four subgroups, poor metabolizers (PMs), intermediate metabolizers (IMs), extensive metabolizers (EMs), and ultrarapid metabolizers (UMs), based on their relative function and ability to express the CYP2D6 gene. The urinary elimination of dextromethorphan was assessed in each of these CYP2D6 subgroups.

Results: The CYP2D6*3, *4, *6, *7, and *9 alleles showed a high degree of concordance between the CYP450 GeneChip and AS-PCR methods (>99% concordance). The mean (SD) of the log[dextromethorphan metabolic ratio (MR)] in the four CYP2D6 subgroups was PM = 0.49 (0.38); IM = 1.24 (0.53); EM = 2.35 (0.61); and UM = 2.43 (0.38).

Conclusions: Oligonucleotide microarray technology is an efficient and reliable way to test for CYP2D6 gene variation based on five alleles compared by separate methods. The methodology is influenced by the quality and amount of DNA present. The log(dextromethorphan MR) is a highly variable index that appears to reflect the crude nature of the dextromethorphan MR as an indicator of CYP2D6 in vivo enzyme activity.

© 2003 American Association for Clinical Chemistry

A total of 67 point mutations and 9 insertions or deletions of ≥3 bp account for 43 cytochrome P450-2D6 (CYP2D6) alleles reported to date (1). Many of the CYP2D6 gene variations affect the expression or activity of the CYP2D6 enzyme to various extents. Interindividual variability in CYP2D6 enzyme activity observed within the general population ranges from complete absence of this enzyme to its overexpression. Gene expression in individuals with inactive CYP2D6 alleles [poor metabolizers (PMs); 5–8% of Caucasians] and individuals with multiple functional CYP2D6 gene copies [ultrarapid metabolizers (UMs); 2–5% of Caucasians] are purported to define the two extremes in CYP2D6 enzyme activity within the population. However, among the vast majority of individuals...
who express the CYP2D6 enzyme, its activity appears to show substantial variation as reflected by a typical metabolic ratio (phenotype) measurement. Some of this variation may be accounted for by a more complete CYP2D6 genotype, but testing for numerous alleles in this gene remains a major undertaking today (2). Thus, it is necessary to ask two questions: (a) can newer technology that determines genetic variation at multiple sites simultaneously reliably replace methods that assess genetic variation at one polymorphism at a time; and (b) if an accurate and extensive assessment of large numbers of CYP2D6 alleles can be done, what are its real merits or limits over a far less sophisticated and more direct assessment of CYP2D6 enzyme activity provided by a simple phenotype? To address these two issues, we compared the reliability of the microarray GeneChip® technology with allele-specific PCR (AS-PCR) methods at five major allelic sites on the CYP2D6 gene. The relative potential of a genetic test for assessing in vivo CYP2D6 enzyme activity was also compared with the dextromethorphan (DXT) urinary metabolic ratio (MR) that has found broad use as an indicator of in vivo CYP2D6 enzyme activity.

Recent results from several laboratories have demonstrated that a limited CYP2D6 allele set may predict the vast majority of Caucasian individuals lacking CYP2D6 enzyme activity with close to 100% accuracy (2–8). As more CYP2D6 null and reduced activity alleles have been discovered in various populations, the ability to account for more subtle variations in CYP2D6 enzyme activity has improved. Before focusing on how much CYP2D6 genotype detail is required in clinical settings, a better understanding of the general reliability provided by new genetic testing methods and a relative appreciation of the limits inherent in the phenotype and genotype are needed.

Materials and Methods

Participants

As a part of the CYP2D6 polymorphism inclusion criteria for a clinical study, a total of 236 healthy individuals were recruited for this research. The study protocol was reviewed by an independent ethics committee/Institutional Review Board for Bio-Kinetic Clinical Applications in Kansas City, Kansas. Volunteers were recruited from a single site, and each volunteer signed an informed consent before participating. A medical history was obtained from all volunteers to determine that there was no alcohol or drug abuse in the last year, no donation of blood or plasma in the last 30 days, no prescription or nonprescription drug use in the last 2 weeks, and no participation in any investigational drug study in the last 30 days. In addition, a urine drug screen was done at the time of recruitment to verify the absence of any drugs of abuse. Volunteers were asked to refrain from alcohol, coffee, tea, chocolate, and cola beverages for 24 h before the start of the study and through its completion. We first collected a 20-mL blood sample for CYP2D6 genetic testing from each volunteer; volunteers then collected a blank urine sample before taking 60 mg of DXT and collecting all of their urine over the next 8 h in an unsupervised setting.

DNA isolation methods

DNA was extracted from whole blood according to a previously described method at the University of Kentucky, College of Pharmacy (4). Briefly, this method involves lysis of all blood cells with a hypotonic salt solution and Igepal CA630 detergent, followed by the centrifugation and collection of cell nuclei. The cell nuclei are lysed with a hypertonic salt solution, and the proteins are degraded in a protease K–sodium dodecyl sulfate solution. Residual proteins are precipitated by 6 mol/L NaCl, and the supernatant containing the DNA is transferred to a fresh tube. The DNA is precipitated with an equal volume of absolute ethanol and washed with 700 mL ethanol/L of water. The precipitated DNA is transferred to a clean, sterile Eppendorf tube to which 500 μL of Tris-EDTA buffer (pH 8.0) is added to reconstitute the precipitated DNA. In some instances, residual blood was extracted with the Qiagen Blood Kit for the GeneChip CYP450 assay (Qiagen Inc., Valencia, CA).

Genetic testing

Consenting participants were tested for CYP2D6 alleles by AS-PCR or PCR followed by restriction enzyme digestion at the University of Kentucky for the CYP2D6*3, *4, *5, *6, *7, *9, *17, *41, *1xn, *2xn/*35xn, and *4xn alleles based on published methods with some modifications (4, 9–16). All of these genetic tests use CYP2D6 intron-specific primers for the initial amplification, followed by either restriction digestion or a second allele-specific amplification for the allele test (Table 1). The genetic tests do not distinguish a *2xn allele from a *35xn allele, but both the *2xn and *35xn alleles are associated with overexpression of the CYP2D6 enzyme. The assignment of *41 may include several as yet undefined CYP2D6*2-related alleles. The *17 allele designation also includes the *40 allele. The *1 allele designation is a composite of *1 plus *13, *15, *16, *18, *22 to *27, *33, *34, and *36 to *39.

Multiple-copy CYP2D6 alleles were first detected by amplifying across adjacent CYP2D6 genes between exon 9 through exon 1 to generate a 9.5-kb amplicon. The subsequent tests assume that CYP2D6 sequential genes are identical and that these are the only duplicated alleles of CYP2D6. The relative hybridization of amplicon fragments to the GeneChip supports the assumption that sequential CYP2D6 genes are indeed duplicated copies of the same allele, not different CYP2D6 alleles.

The execution of this work over a period of several years meant that it was not possible to always go back and retest every sample for every allele. For example, seven DNA samples containing the 1661C, 2850T, and 4180C point mutations and initially classified as *2 by the GeneChip were not available for later −1584G/C promoter analysis after this polymorphism was reported in
### Table 1. Primers, primer pairs, amplicon size, allele test, and PCR conditions for all amplifications.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Primer pair</th>
<th>Amplicon size and/or allele test</th>
<th>PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6F</td>
<td>5'-GTGTGTCAGAGAGAAGCAT-3'</td>
<td>2D6R + 2D6F</td>
<td>4.4-kb amplicon</td>
<td>93 °C for 1 min; 93 °C for 1 min, 65 °C for 30 s, 68 °C for 5 min × 35 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D6R</td>
<td>5'-GTGCTGCTCCATACGAACGCT-3'</td>
<td>UPF14 + 2D6RPI</td>
<td>1.9-kb amplicon</td>
<td>95 °C for 2 min; 95 °C for 15 s, 61 °C for 30 s, 72 °C for 2 min × 33 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>UPF14</td>
<td>5'-GCTGCTGAGGTCATGTCCTCCTC-3'</td>
<td>AFFY-D12R + UPF14</td>
<td>2.7-kb amplicon</td>
<td>95 °C for 2 min; 95 °C for 15 s, 61 °C for 30 s, 72 °C for 2 min × 3 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D62AR</td>
<td>5'-GCATTCCCAGCTTGGAAATCC-3'</td>
<td>2D62AR + 2D61BF</td>
<td>1.8-kb amplicon</td>
<td>93 °C for 1 min; 60 °C for 30 s, 72 °C for 5 min × 28 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D618F</td>
<td>5'-CCAGCCCTTACCTCCTCTTCT-3'</td>
<td>DPF + DRP</td>
<td>3.5 kb (*5 allele test)</td>
<td>93 °C for 1 min; 65 °C for 30 s, 68 °C for 5 min × 35 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D617WT</td>
<td>5'-GCAGCTGCTGAGCCAGCTGCTG-3'</td>
<td>Lx2F + Lx2R</td>
<td>9.5-kb amplicon test for CYP2D6 multiple gene copies</td>
<td>94 °C for 1 min; 94 °C for 30 s, 68 °C for 12 min × 16 cycles; 94 °C for 30 s, 68 °C for 12 min × 15 s/cycle × 16 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D61496F</td>
<td>5'-GTGGTGGGCCCTGACACTCCTTCT-3'</td>
<td>P92-1R + Lx2F</td>
<td>264 bp (*2xn, *2xn, and *4xn alleles)</td>
<td>94 °C for 1 min; 94 °C for 1 min, 58 °C for 1 min, 72 °C for 1.5 min × 25 cycles; 72 °C for 10 min (*1xn cut by restriction enzyme HindIII; +2xn and *4xn uncut)</td>
</tr>
<tr>
<td>2D61496R3</td>
<td>5'-GCTGCTGAGAATGGAATGGAATGG-3'</td>
<td>P11-4F + P124R with 9.5-kb amplicon in place of genomic DNA</td>
<td>433 bp (distinguish *4xn from *2xn and *1xn alleles)</td>
<td>94 °C for 2 min; 94 °C for 1 min, 53 °C for 1 min, 72 °C for 1.5 min × 20 cycles; 72 °C for 10 min (*4xn cut by restriction enzyme HpaII; +2xn and *1xn uncut)</td>
</tr>
<tr>
<td>2D665F</td>
<td>5'-CGGCTGCTCCATACGAACGCTG-3' (wild type)</td>
<td>2D665F + 2D62AR</td>
<td>577 bp [test for *3 allele (2549Adel)]</td>
<td>94 °C for 1 min; 48 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D666F</td>
<td>5'-CGGCTGCTCCATACGAACGCTG-3' (mutant)</td>
<td>2D666F + 2D62AR with 1.8-kb amplicon</td>
<td>577 bp [test *4 allele (1846G→A)]</td>
<td>94 °C for 1 min; 48 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D67BR</td>
<td>5'-GCGGAGGGGGGCCCTCC-3' (wild type)</td>
<td>2D67BR + 2D61BF</td>
<td>577 bp [test *4 allele (1846G→A)]</td>
<td>94 °C for 1 min; 48 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D68BR</td>
<td>5'-GCGGAGGGGGGCCCTCC-3' (mutant)</td>
<td>2D68BR + 2D61BF with 1.8-kb amplicon</td>
<td>470-bp fragment [test *7 allele (2935A→C)]</td>
<td>94 °C for 1 min; 65 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>EF20</td>
<td>5'-GGCGAAACCTCAGCAGCTCCCTCC-3'</td>
<td>EF20 + ER1 and EF20 + ER2 with 1.8-kb amplicon.</td>
<td>470-bp fragment [test *6 allele (1707Tdel)]</td>
<td>94 °C for 1 min; 65 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>ER1</td>
<td>5'-CTCAGCTGCTCCATACGAACGCTG-3' (wild type)</td>
<td>T2R + 2D6MT and TR2 + 2D6MT with 1.8-kb amplicon</td>
<td>470-bp fragment [test *6 allele (1707Tdel)]</td>
<td>94 °C for 1 min; 65 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>ER2</td>
<td>5'-CTCAGCTGCTCCATACGAACGCTG-3' (mutant)</td>
<td>2D661CF</td>
<td>517-bp fragment [test *9 allele (2613–2615del)]</td>
<td>94 °C for 1 min; 65 °C for 1 min, 72 °C for 1.5 min × 12 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D69CF</td>
<td>5'-CTCTTCTGCTGAGGAGGAG-3' (wild type)</td>
<td>2D69CF + 2D62AR and 2D610CF + 2D62AR with 1.8 kb amplicon</td>
<td>237-bp fragment [test *17 allele (1023 C→T)]</td>
<td>94 °C for 1 min; 60 °C for 1 min, 72 °C for 1.5 min × 15 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D610CF</td>
<td>5'-CTCTTCTGCTGAGGAGGAG-3' (mutant)</td>
<td>2D617WT and 2D617WT and 2D617MT and 2D617MT with 4.4-kb amplicon</td>
<td>237-bp fragment [test *17 allele (1023 C→T)]</td>
<td>94 °C for 1 min; 60 °C for 1 min, 72 °C for 1.5 min × 15 cycles; 72 °C for 10 min</td>
</tr>
<tr>
<td>2D617WT</td>
<td>5'-CCAGGCTGCTGAGGAGGAG-3' (wild type)</td>
<td>2D617F + 2D617WT</td>
<td>203-bp fragment from 1.9, 2.7, and 9.5-kb amplicons [test for (−1584 C→G) for *2 and *4 alleles]</td>
<td>94 °C for 5 min; 94 °C for 30 s, 64 °C for 30 s, 72 °C for 30 s × 25 cycles; 72 °C for 10 min. Cut with restriction enzyme Bsal (*2 cut; *41 uncut). Use of 9.5-kb fragment to generate 203-bp fragment to identify type of multiple copy (*2xn or *41xn allele)</td>
</tr>
<tr>
<td>2D617MT</td>
<td>5'-CCAGGCTGCTGAGGAGGAG-3' (mutant)</td>
<td>2D617F and 2D617F with 1.9, 9.5, and 2.7-kb amplicons</td>
<td>203-bp fragment from 1.9, 2.7, and 9.5-kb amplicons [test for (−1584 C→G) for *2 and *4 alleles]</td>
<td>94 °C for 5 min; 94 °C for 30 s, 64 °C for 30 s, 72 °C for 30 s × 25 cycles; 72 °C for 10 min. Cut with restriction enzyme Bsal (*2 cut; *41 uncut). Use of 9.5-kb fragment to generate 203-bp fragment to identify type of multiple copy (*2xn or *41xn allele)</td>
</tr>
<tr>
<td>2D61496F</td>
<td>5'-CGTCTGCTGAGAATGGAATGGAATGG-3'</td>
<td>2D61496F + 2D61496F with 1.9, 9.5, and 2.7-kb amplicons</td>
<td>203-bp fragment from 1.9, 2.7, and 9.5-kb amplicons [test for (−1584 C→G) for *2 and *4 alleles]</td>
<td>94 °C for 5 min; 94 °C for 30 s, 64 °C for 30 s, 72 °C for 30 s × 25 cycles; 72 °C for 10 min. Cut with restriction enzyme Bsal (*2 cut; *41 uncut). Use of 9.5-kb fragment to generate 203-bp fragment to identify type of multiple copy (*2xn or *41xn allele)</td>
</tr>
</tbody>
</table>
the literature (15, 16, 18). The remaining DNA samples that were available were used in an attempt to generate a 1.9-kb ampiclon with the primers described in Table 1 (16–18). All 1.9-kb amplicons generated were tested to assess the −1584G/C polymorphism in the promoter region of the CYP2D6 gene. The 1.9-kb ampiclon relies on the presence of a CYP2D7P gene conversion element within intron 1 of the CYP2D6 gene. A few *2 alleles (~4%) do not contain this CYP2D7P gene conversion element in intron 1, and thus no 1.9-kb ampiclon can be generated. This 4% of *2 alleles represents *2-related alleles containing the 1661C, 2850T, 4180C constellation of point mutations. All identified *2 alleles for which the 1.9-kb fragment was not generated and all samples that contained two *2 alleles were restested using a second reverse primer located in intron 2 of the CYP2D6 gene (Table 1). This generates a 2.7-kb ampiclon from all CYP2D6 genes that was tested for the presence of the −1584G/C variation regardless of whether the CYP2D7P gene conversion element is present in the CYP2D6 gene (Table 1). The regulatory regions of duplicated *2xn samples were similarly evaluated. An ampiclon of the CYP2D6 regulatory region (203 bp) was generated directly from the 9.5-kb amplicon obtained from the duplicated CYP2D6 allele. That ampiclon was then tested for the −1584G variation to confirm the *2xn allele assessment.

CYP2D6 alleles (*2, *3, *4, *6, *7, *8, *9, *10, and *11) were tested by Roche Molecular Systems (Alameda, CA) and Affymetrix (Santa Clara, CA) with the GeneChip CYP450 assay, an oligonucleotide microarray hybridization method. Briefly, CYP2D6 exons 1–9 and adjacent flanking regions were coamplified in a multiplex reaction together with amplicons encompassing exons 4 and 5 of CYP2C19 with primers and amplification conditions provided by the manufacturer. Ten microliters of the PCR amplicon targets were treated with a fragmentation reagent for 20 min at 25 °C (DNaseI at ~0.1 U/μg of DNA) in the presence of 0.5 U of alkaline phosphatase (Roche Molecular Biochemicals) to remove excess deoxynucleotide triphosphates, followed by heat inactivation (95 °C for 10 min). The fragmented DNA amplicons were 3’-end labeled using 25 μmol/L fluorescein-conjugated dideoxy-ATP (NEB) and 10 U of Terminal Transferase (Roche Molecular Biochemicals) at 37 °C for 35 min and heat inactivated for 5 min at 95 °C. The labeled fragmentated target was hybridized to the GeneChip CYP450 probe array in 500 μL of 5X sodium chloride–sodium phosphate–EDTA solution [SSPE; 1× contains, per liter, 8.76 g of NaCl, 10 mL of 1 mol/L NaH2PO4, 2 mL of 0.5 mol/L EDTA (pH 7.4)] containing 1× Denhardt’s solution, 0.5 mL/L Triton, 1 mmol/L hexadeceyltrimethyl ammonium bromide (CTAB), and 1 nmol/L F1 fluoresceinated control oligonucleotide, and subsequently washed with 6X SSPE in the GeneChip Fluidics Station 400 according to automated fluidics protocols provided by the manufacturer. Fluorescence hybridization intensity images were collected using the GeneChip Scanner 50 and analyzed using proprietary analysis algorithms that compare perfect match and mismatch probe sets to determine a call at each polymorphic site.

Fifteen polymorphic sites within CYP2D6 were assessed, and the combination of polymorphisms observed was reported as defined alleles according to standard nomenclature (1). For example, the *10 allele was assigned when 100T, 1661C, 1039T, and 4180C were present, whereas the *4 allele was assigned when these polymorphisms were observed together with 1846A. Many CYP2D6 allelic variants contain the *2A constellation (G1661C, C2850T, and G4180C), and this allele was assumed in the absence of 1758T (*8) or 883C (*11). However, the *2 allele could be incorrectly assigned because the additional polymorphisms associated with the *2 allele designation include *12, *14, *19, *20, *21, *28, *29, *30, *31, *32, *35, and *40, which were not queried by any alternative methods. It is noteworthy that the nested AS-PCR test for the *17 allele was based on an AS-PCR test for the C1023→T point mutation and that all tests for the *17 allele were identified as *2 alleles by the GeneChip based on the constellation of 1661C, 2850T, and 4180C point mutations. The CYP2D6*1 allele was assumed when all CYP2D6 alleles tested by both laboratories were absent.

**GENOTYPIC CLASSIFICATIONS**

Participants were placed in one of four groups based on the number and activity of the combined CYP2D6 alleles from both classic and GeneChip methods. Participants were classified as UM if they carried at least three functional copies of the CYP2D6 gene (i.e., *1/*1xn, *1/*2xn, *2/*1xn, or *2/*2xn). Individuals were classified as PMs if they carried any two of the following CYP2D6 alleles: *3, *4, *4xn, *5, *6; *7, *8, or *11. Individuals were classified as intermediate metabolizers (IMs) if they carried one CYP2D6 allele associated with diminished CYP2D6 enzyme activity (*9, *10, *17, or *41) and one nonfunctional CYP2D6 allele (*3, *4, *4xn, *5, *6, *7, *8, or *11) (19). Genotypes that did not fall into one of these three categories (e.g., individuals identified as carrying one or two functional alleles, *1 and/or *2) were by default placed in the efficient metabolizer (EM) group.

Five individuals who carried two partially functional alleles [*41/*41 (n = 3), *41/*17 (n = 1), and *41/*9 (n = 1)] were classified prospectively based on their mean log(MR) measurement [log(MR) = −1.16] relative to the UM [log(MR) = −2.43], EM [log(MR) = −2.35], IM [log(MR) = −1.24], and PM [log(MR) = 0.49] groups. Inadequate information was available in the literature to decide whether the CYP2D6 activity associated with two partially functional CYP2D6 gene copies should relegate these individuals to the EM or IM group. To avoid introduction of bias into the statistical analysis, we excluded individuals with two partially active CYP2D6 alleles from the initial analysis to determine how varia-
tions in the CYP2D6 genotype were related to the CYP2D6 phenotype.

PHENOTYPES
Each volunteer was asked to collect a blank urine sample and then ingest 60 mg of DXT and collect urine over the next 8 h in an unsupervised setting. The urine volume measurements were available for all participants, and 45-mL aliquots were saved in polypropylene tubes and frozen (~20 °C) until analyzed for DXT and dextorphan (DRP).

Urine was assayed for DXT and DRP by adding 60 μg of buprenorphine (internal standard) to 1000 μL of urine followed by 1.0 mL of acetate buffer (0.10 mol/L, pH 4.5) and 100 μL of freshly prepared β-glucuronidase solution (2000 U/sample). Samples were incubated at 37 °C for 12–18 h. After the incubation, the sample pH was adjusted to above pH 11 with 1 mL of sodium phosphate (pH 12) and 200 μL of 3 mol/L NaOH. The analytes were extracted into 10 mL of n-butanol–hexane (10:90 by volume), and then back-extracted into water by the addition of 300 μL of 0.01 mol/L HCl. A 50-μL aliquot of the aqueous phase was injected into the HPLC. The HPLC system consisted of an Alltech Platinum phenyl column [100 × 4.6 mm (i.d.); 3 μm] connected to a Kratos fluorescence detector (excitation at 228 nm with a 300 nm emission cutoff filter). The mobile phase consisted of acetonitrile–water (80:20 by volume) containing 10 mmol/L KH$_2$PO$_4$ (pH 4.0) and was pumped at a rate of 1.0 mL/min. Positive quality controls (previously tested urine samples) were included in each phenotyping run. Blank urine samples from each patient were used as negative quality controls to detect potential unknown interfering urinary peaks. The intra- and interday CVs were <10% for the DXT and DRP concentrations over the concentration range tested. The assay had lower limits of detection of 33 μg/L for DXT and 300 μg/L for DRP. The limits of quantification were 60 μg/L for DXT and 600 μg/L for DRP. The upper limits were 5.4 mg/L for DXT and 49 mg/L for DRP. In situations in which DXT concentrations fell below these limits, a larger sample (up to 5 mL) was used for analysis, pushing the limit of quantitation to 15 ng/5 mL. Samples were diluted when urinary DRP concentrations exceeded the limits of the assay.

The DXT MR was used as an index of in vivo CYP2D6 enzyme activity. The urinary DXT MR was determined for each individual based on the following equation:

$$\text{MR} = \frac{\mu\text{moles of DXT in 0–8 h urine sample}}{\mu\text{moles of DRP in 0–8 h urine sample}}$$

DATA ANALYSIS
We compared log-transformed MR values between groups of different genotypes and DXT and DRP excretion amounts by one-way ANOVA (SAS program). Significance was defined as $P < 0.05$. Post hoc analysis was corrected for multiple comparisons by use of the Bonferroni method.

ASSOCIATION OF THE CYP2D6 GENOTYPE WITH THE DXT PHENOTYPE
Of the 236 participants, 206 (87.3%) were Caucasian, 13 (5.5%) were African American, 6 (2.5%) were multiracial, 4 (1.7%) were Hispanic, and 3 (1.2%) were Asian based on self-descriptions. The race of four other participants (1.7%) was undefined. We placed 229 of the volunteers in four genotype subgroups based on the number and purported activities of the CYP2D6 alleles. The remaining seven volunteers had genotypes that were considered ambiguous with regard to genotype group: five samples had CYP2D6 genotypes of *41/*41, *41/*9, or *41/*17, and two samples could not be retested because the *2 allele in the promoter region prevented definitive subgroup assignment (*4/*2 and *2/*2; e.g., EM or IM).

A comparison of the MRs for various CYP2D6 genotypes is summarized in Fig. 1. A trend toward a smaller mean log(MR) (higher enzyme activity) is evident as the number of functional CYP2D6 gene copies increases, but substantial scatter exists within the specific genotypes and overlap between genotype groups is apparent. When we placed individuals in only four major groups (UM, EM, IM, and PM) based on their genotype and anticipated relative expression of the CYP2D6 enzyme, the trend toward higher MRs with lower CYP2D6 expression became evident (Fig. 2A). Although there was still substantial variation within the groups, the overall difference in the group mean log[MR] values was significant (Fig. 2A; $P < 0.001$, one-way ANOVA). The mean log[MR] for each group was significantly different from that of all other groups except for the differences between the UM and EM groups, which failed to reach significance.

Among 13 individuals exhibiting a MR $>0.3$, 10 carried two nonfunctional CYP2D6 alleles (combinations of the *3, *4, and *5 alleles). However, three individuals classified with only the CYP2D6*1, *2, and *41 alleles were identified with MRs near the antimode separating the EM from the PM phenotype [0.339 [2(41)]/2(41)], DNA not available for retesting for $−1584G/C$ polymorphism and thus not plotted], 0.365 [41/*1], and 0.529 [41/*1]). The racial composition of these three individuals was African American, multiracial, and Caucasian, respectively. One Caucasian with a CYP2D6 genotype of *4/*5 had a MR of 0.404, overlapping with the MR for the three individuals who appeared to carry only *1, *2, and/or *41 alleles. The Caucasian and multiracial DNA samples have since been retested for a broader array of CYP2D6 alleles (virtually every allele between *1 and *41), but the genotypes in both samples remained unchanged (*41/*1; data not shown). We have no DNA remaining from the African-American sample to permit its retesting.

The 0–8 h urinary recovery of DXT differed 40-fold
between the CYP2D6 genotype subgroups, with 3.2% of dose recoverable in the urine of the PM group and <0.008% in one individual with three or more functional CYP2D6 gene copies (Fig. 2B; \( P < 0.001 \), one-way ANOVA). The mean DRP excretion increased and then decreased as the number of functional CYP2D6 gene copies exceeded two, with mean urinary recoveries of DRP for the PM, IM, EM, and UM groups of 0.93%, 18.6%, 26.6%, and 16.7%, respectively (Fig. 2C). The overall difference in DRP recovery among the groups was not significant when corrected for multiple comparisons.

### CYP2D6 Allele Frequencies in US Population

A summary of frequencies for the CYP2D6 alleles tested in this study is provided in Table 2. The four inactive alleles observed in PMs (CYP2D6*3, *4, *5, and *6) accounted for 24.2% (114 of 472) of the total alleles in this population. The remaining alleles associated with inactive enzyme expression (CYP2D6*7, *8, *11, and *4xn) were not observed in this population. The frequency of alleles reported to be associated with diminished CYP2D6 enzyme activity (*9, *10, *17, and *41) was 15% (70 of 472) of the total alleles in the population. The frequency of the gene duplication alleles was 2.3% (11 of 472) of the total alleles in the population (Table 2).

### Comparison of CYP2D6 Genotyping by AS-PCR with Oligonucleotide Microarray Hybridization (Affymetrix CYP450 GeneChip)

We tested 232 individuals by AS-PCR and the Affymetrix CYP450 GeneChip assay in blinded fashion without previous knowledge of the genotype or phenotype of the individuals. In two individuals, the blind was broken before their analysis by the GeneChip. In one case, it was not possible to repeat the GeneChip analysis with the available DNA, and in the last case, the GeneChip could not call the genotype because the relative signal intensity...
of the second allele was too low with respect to the *1xn allele.

The GeneChip CYP450 multiplex PCR reaction was sensitive to the quality of DNA used and initially failed to yield sufficient amounts of the longer PCR products in the multiplex reaction /H11011/20% of the time (the seven multiplex products varied in size from 159 to 1125 bp—CYP2D6 exons 1–2). In these cases, the use of smaller DNA volumes or DNA prepared with the Qiagen Blood Amp Kit (a minimum of 300 ng/reaction) was consistently found to support the multiplex reaction (data not shown). Table 3 summarizes the concordance between AS-PCR and the GeneChip assay. Genotyping results for the CYP2D6*3, *4, *6, *7, and *9 alleles showed a high degree of concordance between the Affymetrix GeneChip and the AS-PCR methods. However, two discordant alleles were identified on comparison of the AS-PCR results with those of the GeneChip assay; both were associated with genotype errors by the AS-PCR method. During blind testing, the AS-PCR assay failed to detect a CYP2D6*4 A splice site polymorphism in a heterozygous sample (*4/*1). Reassessment of the gel picture indicated that the PCR amplification was weak, and retesting of the DNA sample two more times generated a consistent *4/*1 genotype. Another discrepancy resulted from an initial AS-PCR failure to identify a *9 allele in one sample (*9/*4), an error that led to an initial misassignment to the EM instead of the IM group. This high concordance (>99%) has been typical with several subpopulations compared to date by both methods (W-H. Chou, F-X. Yan, T.B. Ryder, W-W. Liu, C. Perbost,
and false negatives in the sample set.


### Discussion

**Comparison of CYP2D6 Genotyping by AS-PCR with Oligonucleotide Microarray Hybridization (Affymetrix GeneChip CYP450)**

Genotypes of 232 individuals obtained by AS-PCR and Affymetrix GeneChip CYP450 were compared with >99% concordance between these methods for the five allelic variants tested (Table 3). Four individuals were excluded from the comparison because of unsuccessful DNA amplification (*1/*3; DNA quality/quantity for the GeneChip assay) or an inability to obtain an unambiguous genotype by the GeneChip assay (*1xn/*4). The remaining two genotypes (*1/*5 and *1/*4) were eventually run and corrected identified, but not before the unblinding of the DNA samples; thus, they were omitted from the comparison. Only CYP2D6*3, *4, *6, *7, and *9 alleles were included in the comparison because only these five alleles were detected by both assays at the time this study was initially conceived. Because it is important to consider false negatives and false positives in the error estimate, the total error was calculated based on the entire number of samples tested by both methods, not simply the sample number that contained just these alleles.

The Affymetrix GeneChip CYP450 assay and AS-PCR methods performed similarly in the current study. Historically, the error rate with oligonucleotide arrays has been similar to that of AS-PCR. We have identified discordant results between these two methods with other sample sets, which underscored the need to validate the oligonucleotide microarray system as thoroughly as any other new assay before assuming that it will perform flawlessly under all conditions. Some of the weaknesses of the oligonucleotide array are relatively easy to address; in other cases, the weaknesses could require changes in the oligonucleotides that are used in the microarray. For example, in an early modified research version of the assay, errors were made in calling some *4 alleles and a *3 allele that were traced to the PCR procedures. These problems were corrected in the final product version of the PCR multiplexing amplification. However, the seven-plex PCR reaction remains sensitive to the quality and quantity of DNA used, with the longer PCR products generated being the ones most adversely affected by these variables. A *4 allele (e.g., *4E) call was complicated by difficulty in assessing the T3975C polymorphism. That problem was eventually eliminated by ignoring the *4E subtyping polymorphism in the CYP2D6 gene because all *4 variants are null alleles. The annealing of primers to certain oligonucleotide probes on the microarray does not always follow the expected or desired pattern for the homozygous and heterozygous sites. There is currently no method to reliably predict oligonucleotide hybridization behavior or avoid every type of problem that might arise with large numbers of oligonucleotide probes until they are evaluated with real DNA samples. On the positive side, once oligonucleotide probes have been tested and validated, their incorporation into new or different microarrays is straightforward and does not require further optimization if assay conditions are kept constant.

The GeneChip hybridization intensity patterns were also affected in samples with multiple copies of the CYP2D6 gene, an issue that was identified only after several samples with multiple gene copies were tested by both methods. The GeneChip software failed to make a genotype call of some multiple-copy alleles, necessitating one or more retests of the sample. One DNA sample could not be called even after three separate tries (*1xn/*4). The problem was traced to the base-calling algorithm cutoffs that were trained with single CYP2D6 gene copy allele samples. Because a skewed hybridization intensity signal is generated by the nonduplicated gene at certain mutation sites in the CYP2D6 gene relative to those arising from the multiple-copy allele, the signals sometimes fell outside of the cutoff range established for typical single gene copy biallelic distributions. This was not a frequent problem, but it did sometimes cause the assay to fail data quality tests performed by the analysis software. As such it was recognized as a technical issue that could be

---

**Table 3. Comparison of genotyping by AS-PCR with oligonucleotide microarray hybridization (Affymetrix GeneChip CYP450) in 232 individuals.**

<table>
<thead>
<tr>
<th>Methods</th>
<th>*3</th>
<th>*4</th>
<th>*6</th>
<th>*7</th>
<th>*9</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-PCR</td>
<td>4</td>
<td>93</td>
<td>4</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>GeneChip</td>
<td>4</td>
<td>94</td>
<td>4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Errors, n (%)</td>
<td>0/464 (0%)</td>
<td>1/464 (0.22%)</td>
<td>0/464 (0%)</td>
<td>0/464 (0%)</td>
<td>1/464 (0.22%)</td>
</tr>
</tbody>
</table>

*a* The GeneChip assay failed to call two samples; two more were not tested before breaking the study blind. One sample was not called because of limited amounts of DNA and the inability to reamplify after the initial seven-plex PCR failed. The second sample failed because of a multiple-copy CYP2D6 allele coupled with a *4* allele. The results from these four samples were excluded from comparison. Genotypes of the four samples were as follows: *1xn/*4 (n = 1); *1/*4 (n = 1); *1/*4 (n = 1); and *1/*5 (n = 1) based on only AS-PCR methods. A total of 464 alleles were tested by both methods. The error calculation is based on the total false positives and false negatives in the sample set.

*b* Four *4* alleles were coupled with the *5 deletion and appear homozygous by the GeneChip assay and initial AS-PCR test. A *4* and a *9* allele were missed by the AS-PCR method.
resolved either by retraining the analysis algorithm or by assay design changes. A newer version of the CYP450 microarray assay that directly identifies the presence of gene duplications has been developed at Roche Molecular Systems and appears to have eliminated this problem. The newer CYP450 array and assay was not used in this research so that consistency could be maintained throughout the evaluation process.

Considering the magnitude of the variation that exists in the CYP2D6 gene, the high-throughput, information-rich oligonucleotide microarray is a very practical solution to the more traditional genotyping approaches that typically assess one polymorphism at a time (e.g., restriction fragment length polymorphism analysis, AS-PCR, TaqMan®) (2, 4). To acquire the same information obtainable from a single GeneChip assay requires far more work and labor and remains too slow and/or expensive to justify in a routine clinical environment. It currently takes ~4 days to do the separate long amplifications, nested PCR, and retests of amplicons to generate a limited CYP2D6 genotype by the AS-PCR methods used by our laboratory. As the number of CYP2D6 alleles tested has increased, so has the time required for their detection. In contrast, the microarray can detect all the alleles it tests simultaneously in a single day. Looking to the future, where it may be necessary to assess multiple genetic variations in multiple genes, informative multilocus screening technology (like that of the oligonucleotide microarray) may represent the only practical way to rapidly examine the genetic variation that exists in a population at important gene loci.

There are two caveats to this notion. One is that the technology for GeneChip analysis remains fairly expensive, and the other is that the CYP450 GeneChip itself is not inexpensive. There thus must be a fairly strong incentive for clinical laboratories to make this investment. Its advantage is the level of genetic detail it can provide within a relatively modest length of time and with a significantly reduced investment of human effort. Moving forward, researchers should be aware that oligonucleotide microarrays are not foolproof and that it will require more than just a cursory effort to establish their reliability and validate performance. Some of the limitations of microarray technology are not apparent until after many different genomic sample types have been tested from random groups of individuals.

CYP2D6 allele frequencies in the US population
The CYP2D6 allele frequencies in this US population were similar to those found in previous European and US studies (2, 7, 8, 20, 21). Four inactive CYP2D6 alleles (CYP2D6*3, *4, *5, and *6) accounted for a total of 24.2% of the total CYP2D6 alleles vs 25.98% in European populations and 24.5% in other US groups. We did not detect other alleles, including CYP2D6*7, *8, *11, and *2xn (associated with inactive enzyme expression), in this study group. Genetic testing for the four inactive alleles of the CYP2D6 gene (*3, *4, *5, *6) accounted for nearly all individuals classified by the DXT MR as exhibiting the PM phenotype. Testing for this reduced CYP2D6 allele set is probably adequate for most research done in Caucasians, but eventually a higher confidence level will be required if CYP2D6 genetic testing finds its way into the therapeutic arena or is used in other racial groups.

Purists may point out that use of a MR to separate an EM from a PM for all races is open to question. However, it should also be noted there has been considerable intermingling of races in the US and that attempts to relate phenotypes and genotypes separately for Caucasians, African Americans, Hispanics, and Asians is both artificial and inappropriate. If the CYP2D6 genotype can accurately predict CYP2D6 enzyme activity, it should be expected to be predictive of that enzyme activity regardless of the ethnic group. Although this implies that more alleles must be characterized within multiracial populations, that characterization will probably be necessary anyway if genetic testing is to find use as a therapeutic tool in today’s multiracial society. Although the frequency of the CYP2D6 alleles in just Caucasians is reported in Table 2, this was done because the mixture of racial groups in this subpopulation was not representative of the entire US population.

CYP2D6 genotype vs CYP2D6 phenotype
The DXT MR is a rough indicator of in vivo CYP2D6 enzyme activity, as suggested by the marked variation in this phenotypic characteristic in nearly every CYP2D6 genotype (Fig. 1). The DXT MR is also adversely affected by virtually the same urinary recovery of the major metabolite DRP in the IM and UM groups. The need to administer a probe drug and to collect urine over the ensuing 8 h before defining CYP2D6 enzyme activity has always limited the practical utility of a patient phenotype. This time-consuming, patient- and therapy-dependent procedure has never found very widespread clinical acceptance, and it provides only descriptive information about a patient under a specific set of clinical conditions. Fortunately, the CYP2D6 genotype is a fairly reliable predictor of CYP2D6 control enzyme expression and activity, at least for most individuals with the greatest extremes in CYP2D6 enzyme activity [Figs. 1 and 2; Ref. (18)]. The fact that a CYP2D6 genotype can be assessed directly from blood or other biological samples makes the CYP2D6 genotype a far more attractive way to characterize potential CYP2D6 enzyme expression in clinical settings. One advantage of the CYP2D6 genotype lies with the realization that it is unaffected by the clinical environment. Whereas CYP2D6 enzyme activity may vary for one clinical reason or another, the CYP2D6 genotype provides a point of reference from which to understand differences in CYP2D6 enzyme expression among patients that are not possible with the more descriptive phenotype assessment.

There has been substantial interest in characterizing CYP2D6 in vivo activity through a more detailed CYP2D6
genotype (2, 5, 6, 8, 15, 18, 19). This work sheds additional light on those efforts and supports the use of a CYP2D6 genotype to characterize individual CYP2D6 phenotypes as PM, IM, EM, or UM. However, it is the clinical environment where the actual potential of a CYP2D6 genotype must be demonstrated (22) and eventually assessed for its ability to predict therapeutic outcomes and influence patient care costs. It is the impact of the CYP2D6 genotype on outcome and costs that will ultimately define its true therapeutic utility (23). Before focusing too much effort on finer and finer subdivisions of the CYP2D6 genotype, it might be better to focus on whether the more extreme variations in this enzyme can be shown to actually affect clinical outcomes in patients, not just its ability to affect the dose, drug concentrations, or a specific side effect. These more limited efforts have dominated much of the clinical focus on this and other genetic polymorphisms to the detriment of efforts to determine whether genetic testing could actually serve as a viable therapeutic tool.

We would like to thank Bonnie Fijal for help with the plotting routine for the genotype/MR plots and Run-Mei Pan for contributions to the genetic testing work. This research was supported in part by Hoechst Marion Rossel, Inc. (now Aventis Pharmaceuticals, Inc.), which directed and performed all the clinical research and provided nearly all resources for the cost of the phenotype and genotype tests performed at the University of Kentucky laboratories. Roche Molecular Systems, Inc. and Affymetrix provided GeneChip testing for CYP2D6 at no cost to the study and some PCR supplies. Dr. P.J. Wedlund is a paid consultant of Roche Molecular Systems, Inc.

References