Identification of novel CYP2D7-2D6 hybrids: non-functional and functional variants

Andrea Gaedigk1*, Lazara Karelia Montane Jaime2, Joseph S. Bertino Jr.3, Anick Bérard4, Victoria M. Pratt5, L. DiAnne Bradford6 and J. Steven Leeder1

1 Division of Developmental Pharmacology and Experimental Therapeutics, The Children’s Mercy Hospital and Clinics, Kansas City, MO, USA
2 Pharmacology Unit, Department of Paraclinical Sciences, Faculty of Medical Sciences, The University of The West Indies, St. Augustine, Trinidad and Tobago
3 Bertino Consulting, Schenectady, NY, USA
4 Faculty of Pharmacy, University of Montreal, and Research Center, CHU Ste-Justine, Montreal, QC, Canada
5 Quest Diagnostics Nichols Institute, Chantilly, VA, USA
6 Department of Psychiatry and Medicine, Morehouse School of Medicine, Atlanta, GA, USA

*Correspondence: Andrea Gaedigk, Division of Developmental Pharmacology and Experimental Therapeutics, Children’s Mercy Hospital and Clinics, 2401 Gillham Road, Kansas City, MO 64108, USA. e-mail: agaedigk@cmh.edu

INTRODUCTION

CYP2D is a highly polymorphic drug metabolizing enzyme. To date 80 allelic variants including those described in this report and numerous subvariants have been defined by the cytochrome P450 nomenclature committee1. The collection of known alleles comprise variants encoding no, reduced, full, or ultrarapid CYP2D6 activity and give rise to corresponding poor (PM), intermediate (IM), extensive (EM), and ultrarapid (UM) metabolizer phenotypes. While the most frequent sequence variations are single nucleotide polymorphisms (SNPs) or small insertions/deletions, this gene locus also undergoes major recombination events creating alleles harboring large deletions, duplications, multiplications, conversions, and combinations thereof (Zanger et al., 2004; Gaedigk et al., 2007; Kramer et al., 2009; Gaedigk et al., 2010). The CYP2D gene locus contains direct repeat regions and highly homologous gene units which facilitate unequal crossing-over that can lead to allelic variants with large deletions and duplications (Hastings et al., 2009). Most prominent examples are the CYP2D6*5 gene deletion, CYP2D6*1xN, *2xN, and *4xN gene duplications/multiplications and gene tandems such as CYP2D6*36+*10 (the CYP2D6*36 gene carries a CYP2D7 conversion in exon 9). CYP2D7/2D6 hybrid genes are recombination products that likely resulted from large deletion events that fused respective ends of both genes together and carry a T-insertion in exon 1 that causes premature termination. These hybrids are non-functional. We have recently described three alleles carrying non-functional hybrids upstream of intact CYP2D6*1 and *2 genes (CYP2D6*76 + *1, *77 + *2, and *78+ *2). Notably, these alleles presented as gene duplications when tested on certain platforms/assays and hence may be misclassified as CYP2D6*1xN and *2xN, if not fully characterized (Gaedigk et al., 2010). Kramer et al. (2009) have also observed CYP2D7/2D6 hybrid genes in single and tandem arrangements providing further evidence that hybrid genes may not be as rare as generally believed. A graphical overview is provided in Figure 1.

Testing of hybrid genes is rarely performed because their detection typically requires long-range PCR which is not amenable to automation (Figure 1). Hence, despite major CYP2D6 genotyping efforts and the development of many commercially available platforms and test kits, hybrid genes remain elusive. In order to detect CYP2D7/2D6 hybrid genes on a routine basis, we have modified our long-range (XL) PCR to not only generate the 6.6-kb long CYP2D6-specific amplicon which is subsequently used for genotyping, but carries a CYP2D7 conversion in exon 9).

Polymorphic expression of CYP2D6 contributes to the wide range of activity observed for this clinically important drug metabolizing enzyme. In this report we describe novel CYP2D7/2D6 hybrid genes encoding non-functional and functional CYP2D6 protein and a CYP2D7 variant that mimics a CYP2D7/2D6 hybrid gene. Five-kilobyte-long PCR products encompassing the novel genes were entirely sequenced. A quantitative assay probing in different gene regions was employed to determine CYP2D6 and 2D7 copy number variations and the relative position of the hybrid genes within the locus was assessed by long-range PCR. In addition to the previously known CYP2D6*13 and *66 hybrids, we describe three novel non-functional CYP2D7/2D6 hybrids with gene switching in exon 2 (CYP2D6*79), intron 2 (CYP2D6*80), and intron 5 (CYP2D6*67). A CYP2D7-specific T-ins in exon 1 causes a detrimental frame shift. One subject revealed a CYP2D7 conversion in the 5′-flanking region of a CYP2D6*35 allele, was otherwise unaffected (designated CYP2D6*35B). Finally, three DNAs revealed a CYP2D7 gene with a CYP2D6-like region downstream of exon 9 (designated CYP2D7(REP6)). Quantitative copy number determination, sequence analyses, and long-range PCR mapping were in agreement and excluded the presence of additional gene units. Undetected hybrid genes may cause overestimation of CYP2D6 activity (CYP2D6*1/*1 vs *1/*hybrid, etc), but may also cause results that may interfere with the genotype determination. Detection of hybrid events, “single” and tandem, will contribute to more accurate phenotype prediction from genotype data.

Keywords: CYP2D6, hybrid genes, CYP2D6*35B, CYP2D6*67, CYP2D6*78, CYP2D6*80, CYP2D6 poor metabolizer
The subjects included in this analysis gave written informed consent to participate in studies involving genetic testing of the CYP2D6 gene. Studies were approved by the Ethic Committees or Institutional Review Boards of the Institutions where the participants were enrolled (cohort 1: Children's Mercy Hospital, Kansas City, n = 490; cohort 2: University of Montreal, CHU Ste-Justine, n = 268; cohort 3: the University of the West Indies, n = 279; Morehouse School of Medicine, Atlanta, n = 272; or independent institutional review boards for samples obtained at Bassett Healthcare, Cooperstown, NY, n = 1 and Ordway Research Institute Drug Development Center, Albany, NY, n = 1). One sample was also generate duplication-specific and hybrid-specific amplicons (Gaedigk et al., 2010) in a convenient triplex reaction. Prospective and retrospective application of this strategy resulted in the discovery of 12 subjects, which were hybrid-positive, in addition to the CYP2D6*13 (EU098008) and CYP2D6*66 (EU093102) index cases previously characterized. We also found an additional subject positive for the CYP2D6*76+*2 tandem hybrid corroborating our previous finding (Gaedigk et al., 2010).

The goal of the present study was to characterize the hybrid genes within the CYP2D locus in these subjects in detail, assess their functionality using in vivo phenotype data with dextromethorphan as probe drug in a subset of individuals and describe strategies for their detection and analysis.

**MATERIALS AND METHODS**

**SUBJECTS**

The subjects included in this analysis gave written informed consent to participate in studies involving genetic testing of the CYP2D6 gene. Studies were approved by the Ethic Committees or Institutional Review Boards of the Institutions where the participants were enrolled (cohort 1: Children's Mercy Hospital, Kansas City, n = 490; cohort 2: University of Montreal, CHU Ste-Justine, n = 268; cohort 3: the University of the West Indies, n = 279; Morehouse School of Medicine, Atlanta, n = 272; or independent institutional review boards for samples obtained at Bassett Healthcare, Cooperstown, NY, n = 1 and Ordway Research Institute Drug Development Center, Albany, NY, n = 1). One sample was

---

**FIGURE 1 | Overview of the CYP2D7/2D6 hybrid genes.** CYP2D6, CYP2D7, and CYP2D8 genes are shown in light gray, red, and dark gray boxes, respectively. The 600-bp repeat element immediately downstream of exon 9 is shown in blue. CYP2D6 and CYP2D7-derived repetitive elements (REP) are in red and light gray; REP-DEL indicates a fused repeat element generated by a large deletion involving parts of those elements from both genes. PCR fragments generated are represented as lines and their sizes are given in kilobyte. Red and black endpoints denote primer specificity to CYP2D6 and CYP2D7. (A) Graph represents the CYP2D reference locus. Areas affected by large deletions and implicated in CYP2D7/2D6 hybrid formation and the CYP2D6*5 gene deletion are as indicated. (B) Graphic display of the CYP2D6*5 gene deletion allele. XL-PCR amplicons utilized for detection are shown. (C) Graphic display of CYP2D7/2D6 hybrid genes and their detection by amplification of fragment H. Other depicted fragments are only amplified, if respective rearrangements are present in a sample. (D) Representation of an allele with a CYP2D7 gene lacking the 1.6-kb spacer. This CYP2D7 variant also supports formation of fragment H although the CYP2D7/2D6 switch occurs in the downstream region.
obtained through physician-requested diagnostic testing (Quest Diagnostics Nichols Institute). Genomic DNA was isolated from whole blood or saliva specimen collected in Oragen DNA sample collection disks (DNA Genotek, Kanata, ON, Canada) using column-based procedures and stored at 4°C.

**CYP2D6 Activity determined by in vivo phenotyping**

Urinary concentrations of dextromethorphan (DM) and its metabolite dextropropoxyphene (DP) were determined by reversed-phase, high-performance liquid chromatography as described (Blake et al., 2007). Phenotype classification was made based on the urinary DM/DP ratio in 4-, 5-, and 12-h urine collections sampled in Kansas City and Atlanta, Trinidad and Cooperstown, respectively. UM, EM, IM, and PM categories are as defined by Gaedigk et al. (2008).

**Detection of hybrid genes by long-range PCR**

The simultaneous detection of the CYP2D6 gene, hybrids and duplications were achieved by long-range (XL) PCR essentially as described in detail elsewhere (Gaedigk et al., 2010). Briefly, this XL-PCR triplex reaction produced a 6.6-kb-long amplicon from the CYP2D6 gene, a 5-kb-long fragment (fragment H) (Figure 1) and a 3.5-kb-long fragment from gene duplications containing a CYP2D6[REP6] downstream sequence. Hybrid genes support formation of fragment H given that they are comprised of a CYP2D7 sequence upstream of exon 1 and a CYP2D6-specific sequence downstream of CYP2D6 exon 9. XL-PCR was performed at a final concentration of 5% DMSO with KAPA Long Range HotStart DNA Polymerase (KAPA Biosystems, Woburn, MA, USA) as described by the manufacturer’s protocol. Annealing temperature was 68°C and extension times 1 min/1 kb. Reaction volumes were 8 μl containing 15–25 ng genomic DNA. Typically 1.5–3 μl of PCR product were analyzed by agarose gel electrophoresis. Primers and PCR amplicon lengths are given in (Gaedigk et al., 2010). Genotype analysis was performed on diluted triplex XL-PCR amplicons using PCR-RFLP-based assays as previously described (Gaedigk et al., 2010) to determine whether the hybrid genes carried CYP2D6 or 2D7-specific sequence variations in exons 1 and 9, respectively. For subsequent sequence analysis, fragment H was generated as sole amplicon under identical conditions. Genotyping analysis for all other hybrid gene, hybrids and duplications were subsequently resolved by agarose gel electrophoresis.

**DNA sequence analysis**

For sequence analysis, fragment H was purified with the GenElute PCR Clean-Up Kit (Sigma, St Louis, MO, USA) and entirely sequenced using BigDye terminator v3.1 cycle sequencing chemistry and a 3730 capillary DNA analyzer (Applied Biosystems). Sequences deposited in GenBank (M33387, M33388, AY545216, EU098008, and EU093102 served as reference sequences for CYP2D7, CYP2D6, CYP2D6*13 and *16, respectively. To determine the CYP2D7 upstream sequence in more detail, a 5-kb-long amplicon was generated from five subjects (CYP2D6*1/*5; CYP2D6*4/*4; CYP2D6*5/*5; CYP2D6*2/*41, and CYP2D6*1/*1 [Coriell ID NA17111]) with the following CYP2D7-specific primers: 5’- TCAGACGGCCTTTCCACCAC and 5’-CAGGAGAAAGCTGACGAGAGA. Exon 1 and the 276 bp of the upstream region were sequenced as described above.

**Characterization of the CYP2D locus containing hybrid-duplication tandems**

Long-range PCRs covering different regions of the CYP2D locus were performed with primers specific for CYP2D6, CYP2D7, and CYP2D8. XL-PCR was carried out as described previously (Daly et al., 1996, Gaedigk et al., 2010). Figure 1 provides a graphical display of the approximate region covered in the various reactions. Note that the 8-kb ampiclon depicted in Figure 1 corresponds to the XL-PCR first devised by Daly et al. (1996) for the detection of CYP2D6*13 and *16.

**Genotyping assay detecting the CYP2D7 exon 9 conversion and CYP2D7[REP6]**

A 597-bp-long ampiclon was amplified with a CYP2D7-specific forward primer directed at exon 9 (5’-AGCCACTCTCCTGTGCTGTCGCT and a CYP2D6-specific reverse primer binding downstream of the gene (same binding site as the primer used to generate the 6.6-kb-long genotyping template; 5’-CGACTGAGCCGCTGGAGGTAGTAG). A second pair of CYP3A7 primers was added to the reaction that generated an 860-bp-long product (internal control) independently of the CYP2D6 genotype. PCR was carried out in 8-μl reactions with RedJumpStart DNA polymerase (Sigma) and ~15-ng genomic DNA. Cycling conditions comprised an initial denaturing step for 3 min at 94°C that was followed by 42 cycles at 94°C for 10 s, 68°C for 10 s, and 72°C for 30 s and a final hold at 10°C. PCR fragments were subsequently resolved by agarose gel electrophoresis.

An RFLP-based assay discriminating CYP2D6 from CYP2D7 exon 9 sequences was carried out as following. A PCR product was amplified with primers encompassing exon 9 (forward primer 5’-GGCAAGGATTAAGATCTC and reverse primer 5’-CTCGGAGGTAGTACGCT), with the XL-PCR first devised by Daly et al. (1996) and its location immediately downstream of CYP2D8 demonstrated by the generation of XL-PCR amplicons that span CYP2D8 and different regions within the hybrid (the product covering CYP2D8 through CYP2D6 intron 6 is shown in Figure 2C). Also, as depicted in Figure 2C, none of the subjects was positive for a duplication event, i.e. produced the duplication-specific 3.5-kb-long XL-PCR product. Presence of other rearranged genes was excluded by a quantitative copy number assay (Ramamoorthy et al., 2010).

**RESULTS**

Twelve subjects were identified by the triplex XL-PCR to carry CYP2D7*1*2 hybrid gene structures (Table 1). The hybrid genes presented themselves by the presence of a 5-kb ampiclon (fragment H) as shown in Figure 2A. The hybrid was confirmed by “solo” XL-PCR (Figure 2B) and its location immediately downstream of CYP2D8 demonstrated by the generation of XL-PCR amplicons that span CYP2D8 and different regions within the hybrid (the product covering CYP2D8 through CYP2D6 intron 6 is shown in Figure 2C). Also, as depicted in Figure 2C, none of the subjects was positive for a duplication event, i.e. produced the duplication-specific 3.5-kb-long XL-PCR product. Presence of other rearranged genes was excluded by a quantitative copy number assay (Ramamoorthy et al., 2010).
Table 1 | Summary of cases.

| Subject # | Origin* | Ethnicitya | Genotype | Activity score | Predicted Phenotypeb | DM/DXc |
|-----------|---------|------------|----------|----------------|----------------------|--------|
| Index Case | Kansas City | C | *5/*66 | 0 | PM | 1.511 (PM) |
| Case 1 | Quebec | C | *9/*66 | 0.5 | PM | n/a |
| Case 2 | Quebec | C | *5/*66 | 0 | PM | n/a |
| Case 3 | Cooperstown | C | *1/*66 | 1 | slow-EM | n/a |
| Case 4 | Quebec | C | *5/*66 | 0 | PM | 2.17502 (PM) |
| Case 5 | Quebec | C | *6/*66 variant | 0 | PM | n/a |
| Case 6 | Albany | C | *6/*67 | 0 | PM | n/a |
| Case 7 | Coriell | AA | *29/*80 | 0.5 | IM | n/a |
| Case 8 | Quebec | C | *2/*79 | 1 | slow-EM | n/a |
| Case 9 | Cooperstown | C | *4/*35B | 1 | slow-EM | 0.03783 (IM) |
| Case 10 | Kansas City | AA | 2D7[REP6]/*1/*17daughter | 1.5 | fast-EM | n/a |
| Case 11 | Trinidad | Afro-Trinidadian | 2D7[REP6]/*1/*2daughter | 2 | fast-EM | 0.02102 (slow-EM) |
| Case 12 | Atlanta | AA | 2D7[REP6]/*1/*2daughter | 2 | fast-EM | 0.003 (fast-EM) |
| Case 13 | Trinidad Indo-Trinidadian | *4/*76 + *1 | 1 | fast-EM | 0.0094 (fast-EM) |

*Origin denotes the city, province or state the participant was living in when enrolled; Coriell denotes that the DNA sample was obtained from the Coriell Institute.
'C, Caucasian; AA, African American. *Predicted phenotypes are according to Gaedigk et al (Gaedigk et al., 2008). *Observed phenotype using urinary ratio (DM/DX) as measure of CYP2D6 activity; corresponding phenotype group is shown in brackets. n/a, phenotype not available. *Coriell ID NA17165. *This CYP2D6*66A was partially sequenced (exon 7 switch region). *The CYP2D6*1 allele carries a CYP2D7[REP6] (Figure 3, top panel).

![Figure 2](https://example.com/figure2.png)

**FIGURE 2** | Long-range (XL) PCR products generated from hybrid genes and control samples. Fragment designation corresponds to that provided in Figure 1 and used throughout this manuscript. The genotypes of the cases are shown in Table 1. PCR product lengths are given in kilobase to the right, marker sizes on the left. All PCRs were carried out on genomic DNA. (A) Triplex XL-PCR showing the formation of fragment A (entire CYP2D6 gene), fragment B (presence of gene duplication), and fragment H (indicating the presence of a CYP2D7/2D6 hybrid gene). (B) Fragment H was generated by itself. Note that this fragment is weak in the triplex reaction in panel A for case 9, but unequivocally present when amplified "solo." (C) As expected from sequence data, cases 5, 6, and 7 generated a PCR product spanning CYP2D8 to CYP2D6 intron 6. CYP2D6*66 (index case) and case 9 serve as negative controls (see Figure 3 for hybrid details).
Taken together, these results suggested that large deletion events removed parts of the 3′-region of \textit{CYP2D7}, the intragenic region and parts of the 5′-region of \textit{CYP2D6} to form \textit{CYP2D7}/2D6 hybrid genes as displayed in Figure 1.

Sequence analysis revealed that three of the 12 \textit{CYP2D7}/2D6 hybrid-positive cases (Table 1) were identical and corresponded to the \textit{CYP2D6}*66 sequence previously described for our index case and deposited under accession number EU093102. This sequencing effort identified a G-insertion (TCAGGGAGGGAT) in exon 4 that introduced an additional frame shift and further corroborated the non-functional nature of the hybrid (the original sequence EU093102 was corrected accordingly). One subject (case 5) had a hybrid with an identical sequence in the \textit{CYP2D7}/2D6 switch region in exon 7 (Figure 3), but differed from \textit{CYP2D6}*66 in six positions, three of them being upstream of exon 1 (Figure 4). This allele is referred to here as \textit{CYP2D6}*66[variant].

Three subjects (cases 6, 7, and 8) had hybrid genes with a switch from \textit{CYP2D7} to \textit{CYP2D6} in exon 2, exon 3, and exon 5, respectively (Figure 3). These alleles were designated by the nomenclature committee as \textit{CYP2D6}*79, *80, and *67; their corresponding GenBank accession numbers are HM641839, HM641840, and EU098009. Notably, all three variants carry a T-insertion in exon 1, a hallmark feature of \textit{CYP2D7}, which does not support formation of functional protein due to premature translation termination.

In contrast, subject (case 9) was negative for the T-insertion and encoded an open reading frame that corresponded to \textit{CYP2D6}*35 and therefore the formation of active protein. However, the upstream region contained a 110-bp-long cassette (or inversion) that corresponded to \textit{CYP2D7} and contained the primer binding site for the primer used to generate fragment H (Table 1). When this subject was initially genotyped using a shorter XL-PCR template (Gaedigk et al., 1999), we determined a \textit{CYP2D6}*4/*4 genotype, which was discordant with the subject’s phenotype. In contrast, analysis with our current procedure using a 6.6-kb-long template, two alleles were detected and the revised \textit{CYP2D6}*4/*35 genotype was consistent with the DM/DX ratio of 0.05783. While we could not explain this discrepancy at the time this subject was first genotyped over ten years ago, these findings are consistent with the discovery of the \textit{CYP2D7} conversion cassette that prevents binding of the \textit{CYP2D6}-specific primer we originally utilized (primer binding site is shown in Figure 4). To confirm the \textit{CYP2D6} conversion, and determine additional sequence in the upstream region, the 6.6-kb-long XL-PCR product was sequenced in addition to fragment H. This allele was designated \textit{CYP2D6}*35B and has been catalogued under accession number HM641838.

Sequence analysis of the hybrid genes amplified from cases 10, 11, and 12 revealed a \textit{CYP2D7} gene lacking the 1.6 kb spacer sequence that is typically found downstream of \textit{CYP2D7} between a 600 bp repeat region and REP7 (Figures 1 and 3 and Gaedigk et al., 2010).
Due to the absence of the spacer sequence, the CYP2D6-specific reverse primer bound to this gene unit and supported amplification of a CYP2D7/2D6 hybrid PCR product. In line with primer design (CYP2D7-specific forward and CYP2D6-specific reverse primer), the CYP2D7[REP6] gene was also detected by an assay we originally developed to screen genomic DNA for the presence of the exon 9 conversion in CYP2D6 variants such as CYP2D6*4N and *36. As shown in Figure 5A, cases 10–12 produced a PCR product suggesting that the CYP2D7/2D6 switch occurred downstream of exon 9. In contrast, these subjects were negative for the exon 9 conversion when the 6.6-kb-long CYP2D6-specific XL-PCR product was used as genotyping template excluding the possibility that the amplicon shown in Figure 5A is derived from an allele similar to CYP2D6*4N or *36 (Figure 5B). The quantitative copy number assay (to be published elsewhere), revealed two copies for CYP2D6 and CYP2D7, respectively indicating that no additional gene units are present. Taking all findings together we concluded that this “hybrid” is most likely derived from the CYP2D7 gene possessing a CYP2D6-like downstream region, i.e., REP6 lacking the 1.6-kb-long spacer. Furthermore, because all three cases have a CYP2D6*1 allele in common, CYP2D7[REP6] is most likely associated with CYP2D6*1 as depicted in Figure 3.

Following are details for three cohorts for which a full set of experimental and demographic data were available to estimate hybrid allele frequencies: cohort 1 consisted of 490 subjects collected in the US who defined themselves as Caucasian (n = 329), African American (n = 83), Asian (n = 50), Hispanic (n = 10), Indian (n = 8), or did not fit into those categories (unknown, other and admixed, n = 10). Among those we discovered our CYP2D6*66 index case (frequency of CYP2D6*66 in the Caucasian subset was 0.1%) and one CYP2D7[REP6] carrier (f = 0.6% in the African American subset). Cohort 2 was a predominantly Caucasian cohort collected in Canada and the US and consisted of 268 subjects (n = 232 Caucasian, n = 13 African American, n = 9 Hispanic, n = 4 Asian, n = 10 admixed or of other ethnicity). Two CYP2D6*66, one *66[variant], one novel *79 hybrid and one *77 + *2 tandem hybrid allele were found in the Caucasians of this cohort (frequency of all single hybrids combined was 4/232 or 0.86%). A third cohort was composed of 174 Indo Trinidadians and 105 Afro Trinidadians among which only a single CYP2D7[REP6] was discovered in a subject with African ancestry (f = 0.5%). Notably an Indo Trinidadian subject revealed a CYP2D6*76 + *1 tandem hybrid, an allele so far observed in only one other subject who was African American (Gaedigk et al., 2010). To estimate the frequency for CYP2D6*35B, we screened an additional 28 carriers of CYP2D6*35 from cohort 1. However, none were detected. Given a frequency of about 5% of CYP2D6*35 in Caucasians (de Leon et al., 2009) the CYP2D6*35B subvariant appears to be rare (<0.1%).

An alignment of the hybrid genes with CYP2D6 and CYP2D7 reference sequences revealed that the hybrid genes matched with the CYP2D7 sequence given in ENSG000001205702 (Ensemble database), but not GenBank entry M33387 in the immediate upstream region of exon 1. Due to the differences between the two genes in this region, it is a preferred binding site for CYP2D6-specific gene amplification.
CYP2D6* hybrid genes

Gaedigk et al.

and frequently targeted by investigators. We resequenced CYP2D7 of 10 alleles of selected subjects with CYP2D6*1/*1, *1/*5, *4/*4, *5/*5, and *2/*41 genotypes to explore sequence variations within the upstream region of CYP2D7. As summarized in Figure 4, seven alleles corresponded to ENSG00000205702 and three exhibited CA at positions −203/−202, which was also observed in some of the hybrids. However, none matched M33387.

DISCUSSION

The CYP2D6 gene locus has been extensively studied for allelic variation and ethnic differences. It is fair to state that this gene is amongst the most polymorphic drug metabolizing genes with 80 defined allelic variants, numerous recognized subvariants, many variants with trivial names and some that remain unnamed. In addition to simple sequence variations such as SNPs, the locus is also known for major recombination events that created gene duplications, multiplications, conversions, and deletions (Steen et al., 1995; Kramer et al., 2009). Mechanisms that lead to changes in gene copy number and new combinations of exons between different genes via translocation, insertion or deletion have recently been reviewed by Hastings et al. (2009). The CYP2D locus contains three highly homologous genes (CYP2D6, 2D7, and 2D8) along repeat motifs constituting a perfect target for homologous recombination events. Such events are believed to be the mechanism by which CYP2D7/2D6 hybrids were formed: the deletion of a 3′-portion of the CYP2D7, the intergenic region and a 5′-portion of CYP2D6 results in a fused gene unit composed of CYP2D7 and CYP2D6. The first hybrid genes described were CYP2D6*13 and *16 in which the switch occurred in intron 1 and somewhere between intron 7 and exon 8, respectively (Panserat et al., 1995; Daly et al., 1996) (Figure 3); these have recently also been detected by Kramer et al. (2009). We now report additional hybrids switching to CYP2D6 in exon 2, intron 2, and intron 5, a variant of CYP2D6*66 and a CYP2D7 variant with a CYP2D6-like REP6 region mimicking a CYP2D7/2D6 hybrid. The majority of subjects carrying hybrid genes or a hybrid in tandem arrangements (Gaedigk et al., 2010) are of Caucasian origin (Table 1). To date we have identified only one subject of African ancestry with a “single” hybrid gene (case...
and only two non-Caucasian subjects with a tandem hybrid (case 13, an Indo-Trinidadian and an African American described previously (Gaedigk et al., 2010), both having a \textit{CYP2D6*76 + *1}). In contrast, all three subjects with a \textit{CYP2D7*REP6} are of African ancestry (cases 10, 11, and 12).

We estimate the combined frequency of “single” hybrid genes at 0.1–1% in Caucasians and <0.1% in African Americans. This estimate is based on three cohorts which have systematically been interrogated for single as well as tandem hybrids. We have also screened an African American cohort for both single and tandem hybrids and only found a single individual with a hybrid, i.e., \textit{CYP2D6*76 + *1} (Gaedigk et al., 2010) further supporting a low frequency of such alleles in subjects of African ancestry. Before we started to routinely screen for such alleles we selected only samples presenting with rare homozygous genotypes such as \textit{CYP2D6*5/*5} (index case) or *6/*6 (case 6) for hybrid testing. The latter case, a Caucasian who was revised to \textit{CYP2D6*6/*6}, was part of an ethnically mixed cohort of 285 subjects collected in Albany, NY. However, it remains unknown whether any other homozygous individuals of that sample cohort carry hybrids. Other cases were selected for a more complete characterization for a number of reasons. Case 3 for instance was tested with the AmpliChip P450 test in a diagnostic setting and revealed inconsistencies with the genotype of confirmed family members. Identification of a \textit{CYP2D6*66} resolved these inconsistencies. Case 9 on the other hand initially revealed a genotype that was discordant with phenotype. This example illustrates how an unknown sequence variation can interfere with the genotyping procedure, i.e., prohibit primer binding, and thereby cause an incorrect genotype assignment. Unfortunately, a systematic re-evaluation of entire corresponding cohorts from which these cases were extracted was not feasible due to DNA availability, quality and/or expired patient consent.

We did not observe any single or tandem hybrids other than \textit{CYP2D6*36 + *10} in the 50 Asian subjects among our collection. The \textit{CYP2D6*36} gene harboring the \textit{CYP2D7} exon 9 conversion on an otherwise \textit{CYP2D6} background in combination with a \textit{CYP2D6*10} (also referred to as \textit{CYP2D6*36 + *10}) is frequently detected when discriminated from the “single” \textit{CYP2D6*10} allele (Ramamoorthy et al., 2010). While we are not aware of any other evidence that would suggest the presence of \textit{CYP2D7/2D6} hybrids in Asians, larger population samples would need to be investigated.

In single and tandem configuration, hybrid genes with a \textit{CYP2D7/2D6} switch beyond exon 1 are non-functional due to a T-insertion in this exon that leads to premature termination. A relatively simple strategy for hybrid detection and genotyping is to amplify a 5-kb-long PCR product and test for the presence of the T-insertion. We now routinely co-amplify this 5-kb-long ampli- fon (i.e., fragment H) along a duplication-specific fragment and a 6.6-kb-long product encompassing \textit{CYP2D6} as described previously (Gaedigk et al., 2010). This triplex PCR not only provides information regarding the presence or absence of a hybrid and/or duplication, but can also be used as template to genotype for the T-insertion. An alternative strategy to detect hybrid genes in single and tandem configuration is to determine gene copy number in a quantitative manner. However, as discussed by Ramamoorthy et al. (2010), the selection of assay, i.e. which part of the gene is interrogated, is crucial. We have established quantitative assays that probe in four different regions of the \textit{CYP2D6} gene (exon 1, intron 5, intron 6, and exon 9) and two regions each of \textit{CYP2D7} and \textit{CYP2D8} (to be publishes elsewhere) and have applied these assays to the cases reported here. Results were consistent with all genotyping results obtained from genomic DNA and XL-PCR templates as well as additional XL-PCR amplifications spanning different regions of the \textit{CYP2D6} gene locus (Figures 1C, D and 2).

Of the four additional \textit{CYP2D6*66} alleles we have resequenced, one differed in six positions suggesting that not all \textit{CYP2D6*66} may have arisen from the same recombination event or that one variant accumulated additional SNPs. The \textit{CYP2D6*66[variant]} matched with the apparently more common \textit{CYP2D7} sequence corresponding to ENSG00000205702, while \textit{CYP2D6*66} aligned with the minor \textit{CYP2D7} allele (Figure 4).

Interestingly, one initially discordant subject (case 9) revealed a \textit{CYP2D7} conversion in the upstream region of a \textit{CYP2D6*35} (designated \textit{CYP2D6*35B}) that first prevented its detection due to the lack of the \textit{CYP2D6} primer binding site. Furthermore, this allele supported the formation of a PCR product described by Daly et al. (1996) to detect \textit{CYP2D6*16}; hence the initial \textit{CYP2D6*4/*16} genotype assignment. When the genotyping template, however, was generated with a \textit{CYP2D6} primer located further upstream, the \textit{CYP2D6*35B} allele was detected. This case nicely illustrates how unknown sequence variability in a region that is frequently used as primer binding site can cause an incorrect genotype assignment that is discordant with phenotype.

Hybrid genes remain largely undetected using conventional genotyping strategies or platforms. However, depending on primer binding sites utilized and which part of a hybrid gene is derived from \textit{CYP2D6} and \textit{CYP2D7}, amplification may or may not proceed and thereby interfere with the genotype call. For example, if the \textit{CYP2D6} portion is amplified, a wild-type call for a particular SNP would “mimic” the presence of a wild-type or \textit{CYP2D6*1} allele. If no amplification takes place, the result reflects only that of the other allele leading to a homozygous genotype call. Thereby, “hidden” (undetected) hybrids may cause over-estimation of a subject’s phenotype (e.g. \textit{CYP2D6*1/*1 vs 1/hybrid}) or may interfere with a genotyping results by producing unexplained/inconsistent genotyping results preventing unequivocal assignments. Indeed, a subset of the subjects described by Kramer et al. (2009) were triaged for further characterization because of inconsistencies including data that could not be interpreted or homozygosity for rare alleles suggesting allele drop out or the presence of a hybrid gene. This strategy was also pursued by us to identify our first hybrid genes including our index case that was first published as \textit{CYP2D6*5/*16} and later revised to \textit{CYP2D6*5/*66} (Gaedigk et al., 1999; Gaedigk and Coetsee, 2008).

Phenotype information, i.e. urinary metabolic ratios of DM/DX was available for four subjects in addition to case 9 described in detail in the section “Results” and the index case. For eight subjects phenotyping with DM was either not part of the study protocol or there was not sufficient urine available for analysis (case 10). Consistent with genotype, the index case and case 4 were phenotypic PMs with a DM/DX ratio > 0.3. Without information for hybrid alleles, these two cases along case 2, 5, and 6 would have been assigned homozygous \textit{null/null} genotypes (Table 1), which would be consistent with the observed phenotype. The DM/DX ratios observed for cases 11–13 were also in general agreement. In contrast, \textit{CYP2D6} activity towards DM would be overestimated for
cases 1, 3, 7, and 8 because these subjects carry only one and not two copies of a functional or reduced activity allele. For example, case 1 would be misclassified as CYP2D6*9/*9 with an activity score (AS) of 1.5 predicting an EM phenotype while the accurate CYP2D6*9/*66 genotype with an AS of 0.5 predicted this individual to present as an intermediate metabolizer (Gaedigk et al., 2008).

CYP2D7 genes lacking the 1.6-kb spacer sequence (referred to as CYP2D7 [REP6]) and positioned upstream of a functional CYP2D7 complicated the characterization and interpretation of true CYP2D7/2D6 hybrid genes by also forming the 3-kb-long fragment H. It was distinguished from the other hybrid genes by being CYP2D7 throughout exon 9 and being negative for any amplification with primers specific for CYP2D7 and a sequence on the far 3′-end of the CYP2D locus as outlined in Figures 1B, C. In contrast, single hybrid genes support such amplicons. Furthermore, quantitative copy number determination including CYP2D6 and CYP2D7 supported the CYP2D7 [REP6] structure shown in Figure 3.

In summary, our recent discoveries of recombination events creating hybrid genes in single and tandem rearrangements suggest that such events contribute significantly to the high degree of heterogeneity of the CYP2D gene locus. Furthermore, consistent discovery of additional cases with known and novel hybrids in different configurations and subjects of different ethnic and geographic origin demonstrates that such alleles are more common than initially believed. Their detection may improve the accuracy of genotype assignments and phenotype prediction. Assays amenable to automated and high(er) throughput testing, such as quantitative copy number detection, is a promising route to capture a variety of rearrangement event including the CYP2D7/2D6 hybrids described in this report.

ACKNOWLEDGMENTS
We are grateful for the technical assistance of Mrs Liliane Nd joumché. In addition we would also like to thank the following individuals for their contributions to this manuscript: Dr Roger Gaedigk (CMH; art work); Mrs Fatihha Karam (University of Montreal/CHU Ste-Justine; patient recruitment and sample collection) and Mr. Anthony Lalla (The University of the West Indies; patient recruitment, DNA isolation and phenotyping) and Mrs Renea Springe for her excellent work on allele characterization during her practicum semester at CMH.

REFERENCES
Blake, M. J., Gaedigk, A., Pearce, R. E., Bomgaars, L. R., Christensen, M. L., Stowe, G., James, L. P., Wilson, J. T., Kearns, G. L., and Leeder, J. S. (2007). Ontogeny of dextromethorphan O- and N-demethylation in the first year of life. Clin. Pharmacol. Ther. 81, 510–516.

Daly, A. K., Fairbrother, K. S., Andreassen, O. A., London, S. J., Idle, J. R., and Steen, V. M. (1996). Characterization and PCR-based detection of two different hybrid CYP2D7P/CYP2D6 alleles associated with the poor metabolizer phenotype. Pharmacogenomics 6, 319–328.

de Leon, J., Susce, M. T., Johnson, M., Hardin, M., Maw, L., Shao, A., Allen, A. C., Chiafari, F. A., Hillman, G., and Nikoloff, D. M. (2009). DNA microarray technology in the clinical environment: the AmpliChip CYP450 test for CYP2D6 and CYP2C19 genotyping. CNS Spectrums 14, 19–34.

Gaedigk, A., and Coetsee, C. (2008). The CYP2D6 gene locus in South African Coloureds: unique allele distributions, novel alleles and gene arrangements. Eur. J. Clin. Pharmacol. 64, 465–475.

Gaedigk, A., Fuhr, U., Johnson, C., Berard, L. A., Bradford, D., and Leeder, J. S. (2010). CYP2D7-2D6 hybrid tandems: identification of novel CYP2D6 duplication arrangements and implications for phenotype prediction. Pharmacogenomics 11, 43–53.

Gaedigk, A., Gotschall, R. R., Forbes, N. S., Simon, S. D., Kearns, G. L., and Leeder, J. S. (1999). Optimization of cytochrome P4502D6 (CYP2D6) phenotype assignment using a genotyping algorithm based on allele frequency data. Pharmacogenomics 9, 669–682.

Gaedigk, A., Ndjountché, L., Divakaran, K., Bradford, L. D., Zineh, I., Oberlander, T. F., Brousseau, D. C., McCarver, G. D., Johnson, J. A., Alander, S. W., Riggs, K. W., and Leeder, J. S. (2007). Cytochrome P4502D6 (CYP2D6) gene locus heterogeneity: characterization of gene duplication events. Clin. Pharmacol. Ther. 81, 242–251.

Gaedigk, A., Simon, S. D., Pearce, R. E., Bradford, L. D., Kennedy, M. J., and Leeder, J. S. (2008). The CYP2D6 activity score: Translating genotype information into a qualitative measure of phenotype. Clin. Pharmacol. Ther. 83, 234–242.

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., and Ira, G. (2009). Mechanisms of change in gene copy number. Nat. Rev. 10, 551–564.

Kramer, W. E., Walker, D. L., O’Kane, D. J., Mrazek, D. A., Fisher, P. K., Dukek, B. A., Bruflat, J. K., and Black, J. L. (2009). CYP2D6: novel genomic structures and alleles. Pharmacogenet. Genom. 19, 813–822.

Panserat, S., Mura, C., Gerard, N., Vincent-Viry, M., Galateau, M. M., Jasc-Aigrain, E., and Krishnamoorthy, R. (1995). An unequal cross-over event within the CYP2D6 gene cluster generates a chimeric CYP2D7/CYP2D6 gene which is associated with the poor metabolizer phenotype. Br. J. Clin. Pharmacol. 40, 361–367.

Ramamoorthy, A., Fleckhart, D.A., Hosono, N., Kubo, M., Nakamura, Y., and Skaar, T. C. (2010). Differential quantification of CYP2D6 gene copy number by four different quantitative real-time PCR assays. Pharmacogenet. Genom. 20, 451–454.

Steen, V. M., Molven, A., Aarskog, N. K., and Gulbrandsen, A.-K. (1995). Homologous unequal cross-over involving a 2.8 kb direct repeat as a mechanism for the generation of allelic variants of the cytochrome P450 CYP2D6 gene. Hum. Mol. Genet. 4, 2251–2257.

Zanger, U. M., Ratnunodo, S., and Eichelbaum, M. (2004). Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry.