Published in final edited form as:

Pharmacogenomics. 2010 January ; 11(1): 43–53. doi:10.2217/pgs.09.133.

CYP2D7-2D6 hybrid tandems: identification of novel CYP2D6 duplication arrangements and implications for phenotype prediction

Andrea Gaedigk^{1,†}, Uwe Fuhr², Charlene Johnson³, L Anick Bérard⁴, DiAnne Bradford⁵, and J Steven Leeder¹

¹Section of Developmental Pharmacology & Experimental Therapeutics, The Children's Mercy Hospital and Clinics, 2401 Gillham Road, Kansas City, 62108 MO, USA

²University of Cologne, Cologne, Germany

³Dominion Diagnostics LLC, RI, USA

⁴University of Montreal, Montreal, Quebec, Canada

⁵Morehouse School of Medicine, GA, USA

Abstract

Aims—Allelic variants of cytochrome P450 CYP2D6 (*CYP2D6*), such as gene deletion, duplication, multiplication and conversion, contribute to the wide range of CYP2D6 activity. Novel gene arrangements were discovered and characterized.

Materials & methods—DNA from 32 Caucasian and 59 African–American duplication-positive subjects were analyzed by long-range PCR and genotyping to detect *CYP2D7–2D6* hybrid tandem alleles. Novel allelic variants were sequenced and a strategy for the detection and analysis of hybrid genes was refined.

Results—*CYP2D7–2D6* hybrid tandem alleles were identified in one African–American and four Caucasian subjects. Three novel hybrid genes were found on *CYP2D6*1* and *CYP2D6*2* duplication backgrounds and designated *CYP2D6*76*, **77* and **78*. *CYP2D7* to *2D6* conversion occurred in introns 1 and 4, and exon 9. All carried a T-insertion in exon 1 abolishing activity. In Caucasians, four out of 33 (12%) of the duplication-positive alleles were hybrid tandems, three *CYP2D6*77 + *2* and one *CYP2D6*78 + *2*. By contrast, in African–Americans only one of 60 duplication-positive alleles was identified as a hybrid tandem. This allele was designated *CYP2D6*76 + *1*.

[†]Author for correspondence: Tel.: +1 816 234 3059, Fax: +1 816 855 1958, agaedigk@cmh.edu.

Financial & competing interests disclosure

Sequences have been submitted to GenBank and accession numbers assigned. Sequences will be released upon acceptance of the manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

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

Ethical conduct of research

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

Conclusion—Hybrid tandem alleles occur infrequently (<0.25%) in Caucasians, but may explain why not every subject with a *CYP2D6* duplication presents with an ultrarapid metabolizer phenotype.

Keywords

Cytochrome P450; *CYP2D6**76; *CYP2D6**77; *CYP2D6**78; *CYP2D6*; gene duplication; hybrid genes; polymorphism; ultrarapid metabolism

The cytochrome P450 (*CYP*)2*D* gene locus is highly polymorphic and comprises three genes: *CYP2D6*, *CYP2D7* and *CYP2D8*. Currently, 75 *CYP2D6* alleles and additional sub-alleles have been defined by the Human Cytochrome P450 Allele Nomenclature Committee [101], giving rise to a wide range of activity [1–3]. The majority of alleles are characterized by SNPs and small deletions and insertions. However, the *CYP2D6* gene locus has undergone major structural rearrangements, resulting in the deletion of the entire gene and numerous duplication/multiplication variants. Since duplication and multiplication events have not been distinguished in most studies, we will collectively refer to such events as ‘duplications’.

The distribution of gene duplication alleles and their frequencies differ dramatically among populations [4]. The most common duplication alleles carry two (or more) copies of a *CYP2D6**1 *CYP2D6**2 or *CYP2D6**4 gene, and a *CYP2D6*-derived repetitive element (REP)-6 downstream of each gene [4]. These are referred to as *CYP2D6**1x*N*, *CYP2D6**2x*N* or *CYP2D6**4x*N*. *CYP2D6**6x*N*, *10x*N*, *17x*N* (occurs with a REP6 or CYP2D7-like REP7 region), *35x*N*, *36x*N*, *41x*N*, *43x*N* and *45x*N* duplications have also been observed [5]. To further complicate matters, *CYP2D6**4Nx*N* (duplicated *CYP2D6**4*N* subvariant [5]) and *CYP2D6**36x*N*, both containing a *CYP2D7* exon 9 conversion and a CYP2D7-derived REP-7, have to be added to the duplication repertoire. As their ‘single’ gene counterparts, alleles carrying duplications of nonfunctional genes are associated with poor metabolism, while those with duplications of fully functional genes have been shown to confer ultrarapid metabolism [1]. The only gene duplication known to carry two different gene variants to date is the *CYP2D6**36 + *10 tandem. Since *CYP2D6**36 is nonfunctional [6], activity originates only from the *CYP2D6**10 component of this tandem. Consequently, not every duplication allele is associated with ultrarapid activity. Clearly, the accuracy of phenotype prediction of an individual increases if the nature of the genes located on a duplication allele are determined [5].

Hybrid genes are also the result of rearrangement events, but are composed of a *CYP2D7*-derived 5’-end and a *CYP2D6*-derived 3’-end. These are believed to be products of large deletions that removed parts of each gene within the locus along with intergenic sequences by means of unequal crossover. Prime examples are *CYP2D6**13 [7] (EU098008), *CYP2D6**16 [8] (no GenBank entry available) and *CYP2D6**66 (EU093102). Such hybrid genes are nonfunctional due to the presence of the *CYP2D7*-characteristic T-insertion in exon 1, which causes a frameshift and premature termination of translation. Absence of appreciable CYP2D6 activity has been demonstrated by *in vivo* phenotyping using the probe drug dextromethorphan in a subject initially genotyped as *CYP2D6**5/*16 [9] and revised to

*CYP2D6**5/*66 after the hybrid gene was resequenced [10]. The frequency of hybrid alleles remains somewhat elusive owing to the fact that they appear to be rare and require long-range PCR for their detection [8,10]. Hence, there are only limited data available for hybrid alleles. In our Caucasian population (n = 347), the frequency was approximately 0.1%, which was considerably lower than the 2% we have reported for South African Coloreds (n = 99) [10].

In order to improve accuracy of phenotype prediction from genotype data, we now routinely screen for the presence of gene duplications and hybrid genes. Four Caucasian samples presented with PCR products indicative of the presence of a 'normal' *CYP2D6* gene, a gene duplication and a *CYP2D7*—*2D6* hybrid. These fragments can be reconciled with a heterozygous genotype composed of a duplication allele and a hybrid allele based on the allele frequencies of less than 0.5 and 2% we found in Caucasians and South African Coloreds [10]. However, this allele combination has an expected occurrence of less than one out of 5000, and we would not anticipate to encounter four cases among less than 1000 subjects. We therefore hypothesized that the gene duplication and hybrid are located on the same allele in these cases. Such a tandem gene arrangement would support the formation of all three PCR products from a single allele (Figure 1). This hypothesis was further supported by an African-American subject previously described as carrying a novel duplication arrangement tentatively comprising a hybrid gene [5].

The objectives of this study were to characterize the novel hybrid tandem arrangements within the *CYP2D* gene locus and to provide a refined strategy for their detection.

Materials & methods

Subjects

The subjects included in this analysis gave informed written consent to participate in studies involving genetic testing of the *CYP2D6* locus or testing for genes that contribute to variability in CYP2D6 activity. Studies were approved following guidelines established at each of the participating institutions to protect human subjects in accordance with the Declaration of Helsinki. Subjects were from multiple ongoing or concluded studies performed at Children's Mercy Hospital, MO, USA (n > 500), the University of Cologne, Cologne, Germany (n = 56), the University of Montreal, Montreal, Canada (n = 100), Morehouse School of Medicine, GA, USA (n = 210) and Dominion Diagnostics, RI, USA (n = 53). Genomic DNA was isolated from blood or saliva samples using column-based procedures and stored at 4°C.

Detection of gene duplications and hybrid genes by long-range PCR

Detection of gene duplications was achieved by long-range (XL)-PCR as previously described [5]. This XL-PCR duplex reaction produced a 6.6 kb long amplicon (fragment A) from the *CYP2D6* gene located in the most downstream position within the locus, and a 3.5-kb long fragment (fragment B) from gene duplications containing a REP-6 like repetitive element referred to as REP-DUP [5]. To identify hybrid genes at the same time, an additional primer was added to the reaction mix. Hybrids support formation of a 5-kb long

fragment (fragment H), given that they are comprised of a *CYP2D7* sequence upstream of exon 1 and a *CYP2D6* sequence downstream of *CYP2D6* exon 9. Duplication-positive DNAs were further characterized by amplifying additional fragments across the gene locus. If fragment D was formed, it subsequently served as a template to genotype the duplicated gene. The length of fragment E served as an indicator of whether REP-DUP (*CYP2D6*-derived) or REP-7 (*CYP2D7*-derived) were present. An overview of the specific regions amplified by the various set(s) of primers is provided in Figure 1. XL-PCR was performed with JumpStart™ REDAccuTaq® LA DNA polymerase (Sigma, MO, USA) as prescribed by the manufacturer's protocol. All reactions contained 5% dimethyl sulfoxide, the 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 Table 1.

Genotype analysis of the novel hybrid–duplication tandems

Genotype analysis was performed on products amplified together in a 'triplex' XL-PCR (fragments A, B and H), and/or fragments A, D and H were generated in separate reactions. Sequence variations of interest were detected on respective XL-PCR templates utilizing PCR RFLP-based procedures as described in detail elsewhere [11], and/or commercially available TaqMan® assays (Applied Biosystems, CA, USA). Specifically, for the detection of 137insT, a 203-bp fragment was amplified with the forward primer 5' CCC ACC AGG CCC CCT GCC ACT GCC CGG GCT GGG CAA gCT (lower case, mismatch to generate partial *Hind*III restriction site shown in bold) and the reverse primer 5' CAA ACC TGC TTC CCC TTC TCA GCC and digested with *Hind*III. *CYP2D7*-derived PCR product containing 137insT were cut into 170-bp and 33-bp fragments, while *CYP2D6*-derived product remained uncut. To identify *CYP2D6* sequences in exon 9, a 775-bp long amplicon was produced with primers 5' CTT CCG TGG AGT CTT GCA GG and 5' CCT GGG AGG TAG GTA GCC CTG and digested with *Nco*I. Fragments of 591 bp and 184 bp were generated for *CYP2D6*, while those of *CYP2D7* remained uncut. Allele designation followed that prescribed by the P450 nomenclature committee [101].

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

Long-range PCR reactions covering different regions of the *CYP2D* locus were performed with primers specific for *CYP2D6*, *CYP2D7* and *CYP2D8*. XL-PCR was carried out in the presence of 5% dimethyl sulfoxide as described above with the primer pairs shown in Table 1. Figure 1 provides a graphical display of the approximate region covered in the various reactions. Specificity of the assays was ensured by including DNAs genotyped as *CYP2D6**1/*1, *5/*5, *5/*66 and *4/*35xN as positive and negative controls.

DNA sequence analysis of hybrid genes

Amplicons (fragment H) derived from the *CYP2D7*–*2D6* hybrid genes were generated from genomic DNA as described above and after treatment with ExoSAP-IT® (USB, OH, USA), entirely sequenced using BigDye® terminator v3.1 cycle sequencing chemistry and a 3730 capillary DNA analyzer (Applied Biosystems). Sequences deposited in GenBank (M33387, M33388 and AY545216) served as reference sequences for *CYP2D7* and *CYP2D6*, respectively.

Results

Long-range PCR to detect gene duplications & hybrids

A triplex XL-PCR reaction was successfully established and applied to simultaneously produce a PCR product suitable for genotype analysis and screen for the presence of gene duplications and *CYP2D7—2D6* gene hybrids. As shown in Figure 1, the 6.6-kb long fragment A is generated from the *CYP2D6* gene located in the most downstream position within the locus, the 3.5-kb long fragment B is only produced in the presence of a duplication/multiplication event, and the 5-kb long fragment H is only produced in the presence of a *CYP2D7—2D6* hybrid. Figure 2A demonstrates that the amplification of all three fragments in a one-tube triplex reaction yields results comparable to those obtained in single tube reactions. Designed to improve efficiency, this triplex approach has identified four subjects positive for fragments A, B and H (cases 1—4). Two subjects were from the USA (Kansas City, male, aged 1 year, and North Kingstown, age and gender not recorded), and one each from Montreal, Canada (female, aged 30 years) and Cologne, Germany (male, aged 34 years). Another subject identified earlier and suspected of carrying a novel duplication structure (case 5, a female aged 47 years from the Atlanta cohort [5]) also exhibited all three amplification products. Subsequent screening of additional subjects previously identified as duplication-positive did not reveal any additional individuals with such an amplification pattern. The presence of a gene duplication was further substantiated in the five cases by amplification of the 9.3-kb long fragment C, encompassing the intergenic region between the duplicated genes. Amplification of fragment E yielded 4.9-kb long products in cases 1—4, suggesting that intron 6 and REP-DUP of the duplicated gene originated from *CYP2D6*. Case 5, however, did not produce the 4.9-kb long fragment, suggesting that it differs from cases 1—4 (Figures 1 & 2B); the faint 6.5 kb product seen in Figure 2B is derived from the subject's second allele, *CYP2D6*17xN*, which is a rare variant because it contains the *CYP2D7* spacer as described earlier [5].

Characterization of the *CYP2D7—2D6* hybrid tandems

Fragment H was entirely sequenced for cases 1, 2, 4 and 5; case 3 was partially sequenced. The sequences in cases 1—3 were identical. This hybrid was designated *CYP2D6*77* by the nomenclature committee and deposited into GenBank under accession number [GQ162807](#). Sequence analysis confirmed the T-insertion in exon 1, which renders this allele nonfunctional. The *CYP2D7—2D6* switch region was located at the 3'-end of intron 1, but because nucleotides 746G and 843G matched *CYP2D7* as well as *CYP2D6*2* (Figure 3), the switch region could only be approximated. The functional gene unit within this tandem was also genotyped as *CYP2D6*2* (Figure 1). Hence we refer to this allele as *CYP2D6*77+*2*. Cases 1—3 were Caucasian.

The sequence obtained from case 4, a Caucasian, also revealed the T-insertion in exon 1. The *CYP2D7—2D6* switch in this hybrid occurred in intron 4. Comparison to reference sequences pinpointed the switch to an 8-bp area between M33388 position 2283 and 2292 (Figure 3). This hybrid was designated *CYP2D6*78* and is catalogued under accession number [GQ162808](#). As in *CYP2D6*77*, the *CYP2D6* portion of the hybrid and the

functional gene in the downstream position were *CYP2D6**2. This allele is referred to as *CYP2D6**78 + *2.

Case 5, an African-American subject, has been partially characterized [5]. In this report we describe the full sequence of the hybrid gene we speculated was present in this subject. Again, the T-insertion in exon 1 was confirmed. There were a number of switches from *CYP2D7* to *2D6* and back to *2D7*, with an ultimate switch to *CYP2D6* in the 5'-half of exon 9 (Figure 3). This is just upstream of a region that is converted to *CYP2D7* in some *CYP2D6* alleles, and known as 'exon 9 conversion'. This hybrid was designated *CYP2D6**76 and can be found under accession number GQ162806. Since the functional gene in the downstream position is *CYP2D6**1, this allele is referred to as *CYP2D6**76 + *1.

In Caucasians, four of 33 or 12% of the observed duplication alleles were *CYP2D6**77 + *2 and *78 + *2. Based on the observation that duplications occur at a frequency of 1–2% in Caucasians [4,5], one would expect a combined frequency for these two alleles of approximately 0.25%. Among the African-Americans there was only one subject with a hybrid tandem, *CYP2D6**76 + *1, making this a rare event at an estimated frequency of less than 0.1% (Table 2).

Arrangement of hybrid tandems within the *CYP2D* gene locus

Generation of fragment A indicated that functional *CYP2D6* gene units were located in the most downstream position of the locus in all five cases. However, it remained unclear whether the hybrid tandems possessed an entire *CYP2D7* gene, or whether *CYP2D7* became part of the hybrid. To demonstrate the latter, XL-PCR was performed with a *CYP2D8*-specific primer and a series of *CYP2D6*-specific primers. As shown in Figure 4, the *CYP2D6**77 + *2 allele supported amplification between *CYP2D8* and introns 2, intron 6 and exon 9 of *CYP2D6**77. *CYP2D6**78 + *2 produced fragments for the latter two and *CYP2D6**76 + *1 was positive for a fragment between *CYP2D8* and *CYP2D6* exon 9 only. These amplification patterns were in accordance with the allele structures shown in Figure 1c. Also, in the presence of an additional *CYP2D7* gene, PCR products encompassing the 3'-end of *CYP2D8*, *CYP2D7* and portions of the hybrid would be too large to amplify. Taken together, these results strongly suggested that the hybrid was located immediately downstream of *CYP2D8*.

Genotype analysis of *CYP2D7–2D6* hybrid genes

Figure 5 summarizes a simple approach to genotype hybrids for the presence of the *CYP2D7*-derived T-insertion in exon 1 or any other sequence variation of interest. Corresponding results were obtained using fragment H or triplex XL-PCR products as templates. Of note, because primers are binding to fragments A and H in the triplex mix, the resulting PCR-RFLP patterns appear heterozygous (Figure 5B).

Discussion

In this study, we identified three novel *CYP2D7–2D6* hybrid tandem alleles. While nonfunctional hybrid genes composed of a 5'-*CYP2D7* end and a 3'-*CYP2D6* end have been

known for a long time [7,8], they remained somewhat elusive due to their perceived low frequencies and the reliance of XL-PCR for their detection. However, as we began screening for hybrids such as *CYP2D6**13, *16 and *66 [9,10] using a modified triplex XL-PCR approach, hybrid genes were detected at frequencies ranging between 0.2 and 1% [Unpublished data]. While these single hybrids are located at the downstream end of the *CYP2D* gene cluster and cause poor metabolism [7,9,10,12], the nonfunctional *CYP2D6**76, *77 and *78 hybrids are arranged in tandem with a functional *CYP2D6* gene. Preliminary data from a quantitative gene copy-number assay did not reveal any additional gene copies in these alleles, further substantiating that the hybrid tandems encode activity that compares to *CYP2D6**1 and *CYP2D6**2. Owing to their ability to generate duplication-specific amplicons such as fragment B and C (Figures 1 & 2), these tandem hybrids may inaccurately be designated as duplications, and consequently overestimating *CYP2D6* activity in a subject. After hybrid and duplication characterization, *CYP2D6**2/*77 + *2 (two functional copies; extensive metabolizer [EM]), *CYP2D6**4/*77 + *2 (one functional copy; intermediate metabolizer [IM]) and *CYP2D6**2/*78 + *2 (two functional copies; EM) genotypes were assigned to our cases predicting EM and IM phenotypes. For two of the five cases, phenotype data for the probe substrate dextromethorphan were available. Case 1, a male, aged 34 years with a *CYP2D6**2/*77 + *2 genotype had a dextromethorphan (DM):dextromethorphan (DX) urinary metabolic ratio of 0.0654, which would place this subject into the IM category. Case 4, a 1-year-old male infant with a *CYP2D6**2/*78 + *2 genotype had a DM:DX ratio of 0.009, consistent with an EM phenotype. Urinary DM:DX ratios can vary over the range of two orders of magnitudes for subjects with the same genotype. As previously discussed in detail, subjects with genotypes containing two fully functional genes most likely present as extensive metabolizers, but can also produce DM:DX ratios falling within the intermediate metabolizer range [11]. Since genotype does not solely explain variability in *CYP2D6* activity, other nongenetic factors may be responsible for the sevenfold difference in the DM:DX ratio observed for these two subjects.

The concept of molecular drive first described by Dover in 1986 [13] is believed to be the force behind the diversification of multigene families including the P450s. Unequal crossover, gene conversion, transposition, slippage replication and RNA-mediated exchanges are mechanisms contributing to the formation of gene duplication and deletions [14]. A high degree of homology between the *CYP2D6*, 7 and 8 genes, along with Alu sequences situated downstream of the genes, likely favors unequal crossover events for the generation of duplication and deletion alleles as detailed by Lundqvist *et al.*, Levlie *et al.* and Steen *et al.* [15–17]. Applying this mechanism, the *CYP2D6**5 allele containing an intact *CYP2D7* gene and a complete deletion of *CYP2D6* would have provided an approximately 12-kb long crossover unit to generate a duplication allele [16,17]. For the hybrid tandems described in this report, we hypothesize that the large deletion between *CYP2D7* and a duplicated *CYP2D6**1 or *CYP2D6**2 was also generated by unequal crossover events. It is conceivable that a cross-over unit compared with that lost in *CYP2D6**5 was removed from an allele with multiple copies to form a tandem hybrid, and that this unit was transferred to the other chromosome.

Interestingly, another allele containing a tandem of genes is relatively frequent and appears to be present exclusively or predominantly in Asian subjects and their descen- dents. This

allele, first described by Johansson *et al.* as *CYP2D6Ch2*, and now often called *CYP2D6*36 + *10* in accordance with nomenclature, carries a *CYP2D6*36* gene upstream of a *CYP2D6*10* [18]. Interestingly, the *CYP2D6*36* is not a *CYP2D7—2D6* hybrid, but consists of a 5′-*CYP2D6* portion (exons 1—8) and a 3′-*CYP2D7* portion (exon 9 and downstream REP-7). Even though this allele is technically a duplication (with *CYP2D6*36* being the duplicated gene), due to the presence of REP-7 it does not amplify fragment B, and is therefore not detected as a duplication allele in routine testing (a graphical representation is shown in Figure 1d). Given that the activity encoded by *CYP2D6*36* is negligible *in vitro* [19] and *in vivo* [6], it does not contribute to CYP2D6 activity, and hence does not need to be distinguished from a regular *CYP2D6*10*.

We would also like to stress that formation of the 5-kb long fragment H indicates the presence of a gene with a *CYP2D7*-derived upstream region and a *CYP2D6*-derived exon 9 region, but its presence *per se* does not guarantee that the encoded gene fails to produce a functional protein. While all hybrids known to date, including those reported here, possess the *CYP2D7*-derived T-insertion in exon 1, it is conceivable that other hybrids feature a *CYP2D7—2D6* switch further upstream and provide an ORF. In order to prevent any miscalls, we encourage genotyping all hybrids, regardless of whether they occur in a tandem or by themselves, for the presence of this debilitating insertion.

The importance of accurate determination of *CYP2D6* genotype involving gene duplications is demonstrated by the recent work describing ultrarapid metabolism as a risk factor of neonatal opioid toxicity following maternal use of codeine during breastfeeding [20–22]. The authors suggest that codeine use may not be safe in breastfed infants of mothers with a ultrarapid metabolizer phenotype, and that genotype information may be helpful for treatment decisions or to explain the adverse event. In light of current knowledge regarding the heterogeneity of allele rearrangements involving duplication structures, it cannot be emphasized enough that default assignments are not sufficiently accurate for such applications.

The overall frequency of these hybrid duplications may be low in any given population. They are, however, present in geographically diverse regions, and make up 12% of the duplication alleles detected in the Caucasians studied. The relevance of rare alleles and the necessity for their testing on a routine basis can be debated. However, knowledge that they exist and what their structures are can be very valuable when interpreting unusual genotype results, and may allow for more accurate genotype calls in individual cases.

Conclusion

In summary, we identified and characterized three novel *CYP2D6* variants within duplication structures, namely *CYP2D6*76 + *1*, **77 + *2* and **78 + *2*. While these alleles may only be observed on occasion, they are present in North American and European subjects. Since the *CYP2D6*76*, **77* and **78* gene components are nonfunctional hybrids, subjects may be inaccurately classified as ultrarapid metabolizers.

Acknowledgements

We thank Mrs Liliane Ndjountché (Children's Mercy Hospital) for providing technical assistance with genotype and sequence analyses. We would also like to thank Dr Robin E Pearce (Children's Mercy Hospital) for performing phenotype analysis and Mrs Fatiha Karam (University of Montreal) for her contribution to patient recruitment.

Bibliography

Papers of special note have been highlighted as:

■ of interest

■ ■ of considerable interest

1. Zanger UM, Raimundo S, Eichelbaum M: Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. *Naunyn-Schmiedeberg's Arch. Pharmacol* 369, 23–37 (2004). [PubMed: 14618296] ■ Summarizes important aspects and essential knowledge of CYP2D6.
2. Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C: Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoeypigenetic and clinical aspects. *Pharmacol. Ther* 116, 496–526 (2007). [PubMed: 18001838]
3. Zanger UM, Turpeinen M, Klein K, Schwab M: Functional pharmacogenetics/ genomics of human cytochromes P450 involved in drug biotransformation. *Anal. Bioanal. Chem* 392, 1093–1108 (2008). [PubMed: 18695978] ■ Provides a comprehensive review of P450 pharmacogenetics including CYP2D6.
4. Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, Fuselli S: CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenet. Genomics* 17, 93–101 (2007). [PubMed: 17301689] ■ Provides genotyping data and allele frequencies for numerous populations and ethnic groups across continents.
5. Gaedigk A, Ndjountche L, Divakaran K et al.: Cytochrome P4502D6 (CYP2D6) gene locus heterogeneity: characterization of gene duplication events. *Clin. Pharmacol. Ther* 81, 242–251 (2007). [PubMed: 17259947] ■ Provides additional detailed information for CYP2D6 gene duplications.
6. Gaedigk A, Bradford LD, Alander SW, Leeder JS: *CYP2D6*36* gene arrangements within the *CYP2D6* locus: association of *CYP2D6*36* with poor metabolizer status. *Drug Metab. Dispos* 34, 563–569 (2006). [PubMed: 16415111]
7. Panserat S, Mura C, Gerard N et al.: An unequal cross-over event within the *CYP2D* gene cluster generates a chimeric *CYP2D7/CYP2D6* gene which is associated with the poor metabolizer phenotype. *Br. J. Clin. Pharmacol* 40, 361–367 (1995). [PubMed: 8554938]
8. Daly AK, Fairbrother KS, Andreassen OA, London SJ, Idle JR, Steen VM: Characterization and PCR-based detection of two different hybrid *CYP2D7P/CYP2D6* alleles associated with the poor metabolizer phenotype. *Pharmacogenetics* 6, 319–328 (1996). [PubMed: 8873218]
9. Gaedigk A, Gotschall RR, Forbes NS, Simon SD, Kearns GL, Leeder JS: Optimization of cytochrome P4502D6 (CYP2D6) phenotype assignment using a genotyping algorithm based on allele frequency data. *Pharmacogenetics* 9, 669–682 (1999). [PubMed: 10634130]
10. Gaedigk A, Coetsee C: The CYP2D6 gene locus in South African Coloureds: unique allele distributions, novel alleles and gene arrangements. *Eur. J. Clin. Pharmacol* 64, 465–475 (2008). [PubMed: 18202841] ■ Describes the detection and characterization of 'single' CYP2D6 hybrids adding complementary information to the current report.
11. Gaedigk A, Simon SD, Pearce RE, Bradford LD, Kennedy MJ, Leeder JS: The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin. Pharmacol. Ther* 83, 234–242 (2008). [PubMed: 17971818]
12. Leathart JBS, London SJ, Steward A, Adams JD, Idle JR, Daly AK: CYP2D6 phenotype -genotype relationships in African-Americans and Caucasians in Los Angeles. *Pharmacogenetics* 8, 529–541 (1998). [PubMed: 9918137]

13. Dover GA: Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet.* 2, 159–165 (1986). ■ Summarizes Southern-blot and long-range PCR data for gene duplications and offers mechanistic models that may explain duplication and deletion formation.
14. Hastings PJ, Lupski JR, Rosenberg SM, Ira G: Mechanisms of change in gene copy number. *Nat. Rev. Genet* 10, 551–564 (2009). [PubMed: 19597530] ■■ Summarizes mechanisms leading to copy-number variation.
15. Lundqvist E, Johansson I, Ingelman-Sundberg M: Genetic mechanism for duplication and multiduplication of the *CYP2D6* gene. *Gene* 226, 327–338 (1999). [PubMed: 9931507]
16. Løvlie R, Daly AK, Molven A, Idle JR, Steen VM: Ultrarapid metabolizers of debrisoquine: Characterization and PCR-based detection of alleles with duplication of the *CYP2D6* gene. *FEBS Letters* 392, 30–34 (1996). [PubMed: 8769309]
17. Steen VM, Molven A, Aarskog NK, Gulbrandsen AK: 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 (1995). [PubMed: 8634695]
18. Johansson I, Oscarson M, Q-Yue Y, Bertilsson L, Sjöqvist F, Ingelman-Sundberg M: Genetic analysis of the Chinese cytochrome P4502D locus: characterization of variant *CYP2D6* genes in subjects with diminished capacity for debrisoquine hydroxylation. *Mol. Pharmacol* 46, 452–459 (1994). [PubMed: 7935325]
19. Hanioka N, Okumura Y, Saito Y et al.: Catalytic roles of CYP2D6.10 and CYP2D6.36 enzymes in mexiletine metabolism: *In vitro* functional analysis of recombinant proteins expressed in *Saccharomyces cerevisiae*. *Biochem. Pharmacol* 71, 1386–1395 (2006). [PubMed: 16527257]
20. Koren G, Cairns J, Chitayat D, Gaedigk A, Leeder SJ: Pharmacogenetics of morphine poisoning in a breastfed neonate of a codeine-prescribed mother. *Lancet* 368, 704 (2006). [PubMed: 16920476]
21. Madadi P, Ross CJR, Hayden MR et al.: Pharmacogenetics of neonatal opioid toxicity following maternal use of codeine during breastfeeding: a case-control study. *Clin. Pharmacol. Ther* 85, 31–35 (2009). [PubMed: 18719619]
22. Madadi P, Koren G: Pharmacogenetic insights into codeine analgesia: implications to pediatric codeine use. *Pharmacogenomics* 9, 1267–1284 (2008). [PubMed: 18781855]

Executive summary

CYP2D6 gene duplications

- *CYP2D6* duplication alleles carry two or more copies of a *CYP2D6* gene. The duplicated gene can be functional (e.g., *CYP2D6*1xN*, **2xN*) or nonfunctional (e.g., *CYP2D6*4xN*). Gene duplications may be absent in a given population or be as high as 29%.
- Subjects genotyped with three or more functional genes are predicted to exhibit an ultrarapid metabolizer (UM) phenotype towards CYP2D6 probe substrates.
- Gene duplications are often detected by amplification of a 3.5-kb long PCR product. Generation of this product only indicates that a duplication is present. It does not, however, provide any information whether the duplicated gene is functional or not. Also, it does not determine how many copies there are.
- *CYP2D6*2xN* carrying two or more functional gene copies is the most common. Therefore, duplication assignments are often defaulted to this allele, a practice that may cause inaccurate phenotype assignments.

CYP2D6 hybrid genes

- The 5'-end of hybrid genes are derived from the *CYP2D7* pseudogene, and the 3'-end from *CYP2D6*. A T-insertion in exon 1, a hallmark feature of *CYP2D7*, causes a shift in the reading frame and premature termination of translation in all known hybrid genes. Consequently, the known hybrid genes are nonfunctional.
- A number of hybrid genes, including *CYP2D6*13*, **16* and **66*, have been described. Due to their perceived low frequency and the necessity of long-range PCR for their detection, information as to regarding their frequency is limited.

CYP2D6 hybrid tandem alleles

- We have characterized allelic variants that carry a hybrid gene upstream of a functional gene. These are referred to as hybrid tandems. The hybrid genes were designated *CYP2D6*76*, **77* and **78* by the Human Cytochrome P450 (CYP) Allele Nomenclature Committee.
- The hybrid tandems test positive for the 3.5-kb PCR fragment, which is indicative of the presence of a gene duplication or multiplication event. Because the hybrid genes are nonfunctional, these alleles carry only a single functional gene copy and are therefore not conveying ultrarapid metabolism.
- In our Caucasian population sample, 4/33 (12%) of gene duplications further characterized were tandem hybrids. Their presence may explain, at least in

part, the wide range of activity observed among subjects classified as UMs after limited testing.

Conclusion

- Not every *CYP2D6* allele identified as duplication-positive by the widely used method of amplifying a 3.5-kb product is associated with the UM phenotype. A subset of such alleles carries nonfunctional genes such as *CYP2D6*4xN* or only a single functional gene such as the hybrid tandems described in this report.
- In order to accurately assign UM phenotype status from genotype data, determination of the nature of the duplicated gene is warranted.

Figure 1. Overview of *CYP2D6* gene duplication structures.

CYP2D6, *2D7* and *2D8* gene sequences are shown as pink, green and gray boxes. Repeat elements located downstream of *CYP2D6* (REP-6), *CYP2D7* (REP-7) and *CYP2D6*-derived gene duplications (REP- DUP) are shown in open boxes in respective colors. PCR fragments generated are represented as lines and their sizes are given in kb. Pink, green and gray end points indicate primer specificity to *CYP2D6*, *2D7* and *2D8* sequences. (A) Normal *CYP2D6* gene locus. (B) Most common duplication structure that entails two (or more) genes of the same kind back-to-back, such as *CYP2D6**1xN, *2xN, *4xN and so on. Primers utilized for duplication detection take advantage of sequences at the REP-DUP junction and within respective genes. (C) Summary of the novel duplication structures carrying a functional *CYP2D6**1 or *2 at the 3'-end of the gene locus and a hybrid gene (*CYP2D6**76, *77 or *78) upstream in the duplication position. The intergenic region between the two genes supports amplification of widely used long-range PCR products such as fragment B and C. Depending on where the *CYP2D7*–*2D6* switch occurred, primer pairs

will generate allele-specific PCR fragments. The formation of PCR products with *CYP2D8* and 2D6 primer combinations demonstrates that the hybrid is situated downstream of *CYP2D8*. **(D)** For completion, the structure of the *CYP2D6*36 + *10* tandem is provided. Due to the presence of *CYP2D7* sequences in *CYP2D6*36* - that is, the exon 9 conversion - the spacer sequence and REP-7, fragments B, C and D fail to amplify. Fragment A, however, is generated and when genotyped indicates a *CYP2D6*10* allele assignment. *REP*: *Repetitive element*.

Figure 2. Long-range PCR products generated from the hybrid tandems and control samples. Fragment designation corresponds to that shown in Figure 1 and used throughout this manuscript. The genotypes of positive and negative control DNAs are as indicated. PCR product lengths are given in kb pairs to the right, marker sizes on the left. All PCR reactions were carried out on genomic DNA. **(A)** Fragments A, B and H were generated in separate reactions and in a triplex PCR. The five cases consistently produced all three amplicons indicating the presence of a ‘normal’ *CYP2D6* gene, a gene duplication and a hybrid gene. **(B)** A duplication event was further demonstrated by amplifying the entire intergenic region (fragment C). While cases 1–4 produced fragment E in addition to an internal control amplicon, this fragment failed to amplify from the *CYP2D6**76 case because the forward primer used cannot bind to the *CYP2D7* sequence present in intron 6.

		M33388	EU098008	GQ162807	GQ162808	EU093102	GQ162806	M33387			M33388	EU098008	GQ162807	GQ162808	EU093102	GQ162806	M33387
Position M33388		2D6	*13	*77	*78	*66	*76	2D7	Intron 4								
1344	-276	C	T	T	T	T	T	T	3637	2018	A	A	A	T	T	T	T
1345	-275	C	T	T	T	T	T	T	3638	2019	:	:	:	C	C	C	C
1348	-272	C	T	T	T	T	T	T	3639	2020	A	A	A	G	G	G	G
1352	-268	G	A	A	A	A	A	A	3642	2023	C	C	C	G	G	G	G
1353	-267	G	C	C	C	C	C	C	3644	2025	G	G	G	A	A	A	A
1388	-232	G	C	C	C	C	C	C	3648	2029	A	A	A	:	:	:	:
1395	-225	A	G	G	G	G	G	G	3653	2034	C	C	C	T	T	T	T
1403	-217	G	G	G	G	G	G	C	3659	2050	C	C	C	T	T	T	T
1417	-203	C	T	T	T	C	T	C	3699	2080	G	G	G	C	C	C	C
1418	-202	G	G	G	G	A	G	G	3704	2085	A	A	A	G	G	G	G
1431	-189	T	T	T	T	T	T	C	3716	2097	A	A	A	G	G	A	G
1434	-186	:	:	:	:	:	:	C	3855	2236	T	T	T	:	:	:	:
1434	-186	:	:	:	:	:	:	A	3858	2239	T	T	T	C	C	C	C
1434	-186	:	:	:	:	:	:	C	3859	2240	A	A	A	G	G	G	G
1465	-155	C	C	C	C	C	C	G	3867	2248	C	C	C	T	T	T	T
1466	-154	A	A	A	A	A	A	G	3868	2249	A	A	A	G	G	G	G
1482	-138	A	G	G	G	G	G	G	3887	2268	T	T	T	C	T	T	C
1534	-86	T	C	T	T	T	T	T	3901	2282	:	:	:	G	G	G	G
1602	-18	T	G	G	G	G	G	G	3902	2283	:	:	:	A	A	A	A
1618	-2	G	C	C	C	C	C	C	3911	2292	G	T	G	G	G	G	G
1619	-1	T	C	C	C	C	C	C	3977	2358	A	A	A	A	T	T	T
Intron 1									4004	2385	T	T	T	T	G	G	G
1833	214	G>C	C	C	C	C	C	C	4011	2392	G	G	G	G	A	G	A
1840	221	C>A	A	A	A	A	A	A	4030	2411	G	G	G	G	A	G	G
1842	223	C>G	G	G	G	G	G	G	4043	2424	C	C	C	C	A	C	C
1846	227	T>C	C	C	C	C	C	C	Exon 7								
1851	232	G>C	C	C	C	C	C	C	4776	3157	G	G	G	G	T	T	T
1852	233	A>C	C	C	C	C	C	C	4873	3254	T	T	T	T	T	C	C
1864	245	A>G	G	G	G	G	G	G	4884	3265	G	G	G	G	A	G	G
1873	254	G	T	T	T	T	T	T	4898	3279	G	G	G	G	G	G	A
1889	270	C	T	T	T	T	T	T	4907	3288	G	G	G	G	G	G	A
1929	310	G	G	G	G	G	G	:	4929	3310	G	G	G	G	G	G	A
1936	317	A	G	G	G	G	G	G	Exon 9								
1942	323	G	A	A	A	A	A	A	5743	4124	G	G	G	G	G	C	C
2365	746	C>G	C	G	G	G	C	G	5747	4128	C	C	C	C	C	G	G
2462	843	T>G	T	G	G	G	G	G	5750	4131	A	A	A	A	A	G	G
Exon 2									5752	4133	T	T	T	T	T	C	C
2583	974	C	C	C	A	A	C	A	5774	4155	C	C	C	C	C	C	T
2603	984	A	A	A	G	G	G	G	5775	4156	A	A	A	A	A	A	C
2616	997	C	C	C	G	G	G	G	5777	4158	G	G	G	G	G	G	C
2633	1014	T	T	T	C	C	C	C	5783	4164	T	T	T	T	T	T	G
2641	1022	A	A	A	T	T	T	T	5785	4166	T	T	T	T	T	T	C
2642	1023	C	C	C	A	A	A	A	5786	4167	G	G	G	G	G	G	A
2647	1028	A	A	A	G	G	G	G	5787	4168	C	C	C	C	C	C	G
2655	1036	T	T	T	C	C	C	C	5788	4169	T	T	T	T	T	T	C
2666	1047	G	G	G	G	G	C	G	5791	4172	C	C	C	C	C	C	T
									5799	4180	G>C	G	C	C	C	G	C
									5812	4193	T	T	T	T	T	T	C

Intron 1 conversion

Exon 9 conversion

Figure 3. Alignment of selected regions of hybrid sequences to *CYP2D6* and *2D7* references. Position numbers are given according to M33388, with -1 referring to the first nucleotide 5' of the ATG start codon. *CYP2D6* sequences are shown in yellow, *CYP2D7* in green. Blue highlights deviating nucleotides or SNPs within the novel alleles. Known *CYP2D6* SNPs that correspond to nucleotides found in *CYP2D7* are as indicated (e.g., G>C at position 1833). Note that other known SNPs within *CYP2D6* are not shown. Bold boxes indicate the switch regions. In *CYP2D6**77, two nucleotides in the switch region (orange) could not unequivocally be assigned to *CYP2D6* or *2D7*. Regions referred to as *CYP2D7* intron 1 and exon 9 conversions are as indicated. GenBank accession numbers are as shown.

Figure 4. Further characterization of the tandem hybrid gene locus using long-range PCR.

Fragment designations correspond to those in Figure 1. The genotypes of positive and negative control DNAs are as indicated. PCR product lengths are given in kb to the right, marker sizes on the left. All PCR reactions were carried out on genomic DNA. **(A)** Series of PCR reactions carried out with *CYP2D8* forward and *CYP2D6* reverse primers binding to intron 1, intron 6 and exon 9 sequences, demonstrate that each hybrid is located immediately downstream of *CYP2D8*. According to their respective hybrid composition, *CYP2D6**77 produced all three amplicons, while *CYP2D6**78 and *CYP2D6**76 produced two and one, respectively. **(B)** In this series of reactions, the forward primer is *CYP2D7* specific (also used for fragment H). As in **(A)**, the hybrid tandems amplified fragments according to their hybrid structure. These shorter fragments readily amplify in less than 3 h and provide a convenient way to further analyze hybrids. Note that amplicon sizes vary slightly (middle and bottom panels) due to sequence differences between the hybrids.

Figure 5. Strategy to genotype hybrid-derived templates.

The graph underneath each gel print provides an overview of the hybrid tandem, the long-range PCR fragment(s) present in each genotyping reaction and the sequence variation interrogated. For RFLP genotyping, a PCR fragment was amplified, cut with a restriction enzyme and resulting fragments resolved by agarose gel electrophoresis. Fragment lengths are given in bp to the right, marker sizes on the left. **(A)** Genotyping assay detecting the T-insertion in exon 1 on fragment H. DNA samples with hybrid tandems, *CYP2D7* only (control plasmid) and a single hybrid (*CYP2D6*66*) were completely cut, indicating the presence of the T-insertion. The *CYP2D6* control plasmid remained uncut due to the absence of the insertion. **(B)** Genotyping assay detecting the exon 9 conversion on fragment H. DNA

samples with hybrid tandems, a single hybrid (*CYP2D6*66*) and *CYP2D6* only (plasmid control) were cut, indicating that exon 9 is derived from *CYP2D6*. The *CYP2D7* plasmid control remained uncut due to the presence of the exon 9 conversion. **(C)** Assay as in **(A)**, but performed on triplex long-range PCR products. Since the primers will bind to fragments A (from both alleles) and fragment H (from the tandem hybrid) the resulting band pattern appears heterozygous. A similar (heterozygous) result is obtained when genotyping is performed for the exon 9 conversion (not shown).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Primers for long-range PCR, product designation and length.

PCR fragment *	Primer sequence 5' to 3'	Product length (kb) [‡]
A	5' ATG GCA GCT GCC ATA CAA TCC ACC TG (F) 5' CGA CTG AGC CCT GGG AGG TAG GTA G (R)	6.6
B	5' CCA TGG AAG CCC AGG ACT GAG C (F) 5' CGG CAG TGG TCA GCT AAT GAC (R)	3.5
C	5' GCC ACC ATG GTG TCT TTG CTT TCC TGG (F) 5' CCG GAT TCC AGC TGG GAA ATG CG (R)	9.3
D	5' CCA GAA GGC TTT GCA GGC TTC AG (F) 5' CGG CAG TGG TCA GCT AAT GAC (R)	8.6
E	5' AGG AGG CAA GAA GGA GTG TCA GG (F) 5' CCT GTA GTG TCA GTG ACT CAG AAG GCT G (R)	4.9
H	5' TCC GAC CAG GCC TTT CTA CCA C 5' CGACT GAG CCC TGG GAG GTA GGT AG (R)	5.0
2D8 to hybrid (intron 2)	5' CTC CTG CCC AGG GGA TGA TG (F) 5' CCG GAT TCC AGC TGG GAA ATG CG (R)	6.5
2D8 to hybrid (intron 6)	5' CTC CTG CCC AGG GGA TGA TG (F) 5' CTC GGC CCC TGC ACT GTT TC	8.2
2D8 to hybrid (exon 9)	5' CTC CTG CCC AGG GGA TGA TG (F) 5' GTC ACC AGG AAA GCA AAG ACA CCA TG (R)	9.4
Hybrid (upstream to intron 2)	5' TCC GAC CAG GCC TTT CTA CCA C (F) 5' CCGGATTCCAGCTGGAAATGCG (R)	1.6
Hybrid (upstream to intron 6)	5' TCC GAC CAG GCC TTT CTA CCA C (F) 5' CTC GGC CCC TGC ACT GTT TC (R)	3.3
Hybrid (upstream to exon 9)	5' TCC GAC CAG GCC TTT CTA CCA C (F) 5' GTC ACC AGG AAA GCA AAG ACA CCA TG (R)	4.5

All PCR reactions were carried out in the presence of 5% dimethyl sulfoxide.

* PCR fragments as shown in Figures 1–4.

[‡] fragment sizes from hybrids vary among hybrids, sizes shown correspond to CYP2D6*77.

F: Forward; R: Reverse.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Summary of the duplication variants identified in Caucasian (n = 32) and African–American (n = 59) subjects. In each cohort one subject had a duplication event on both alleles.

Ethnicity	Caucasian	African–American
Alleles with duplications (n)	33 (100%)	60 (100%)
<i>CYP2D6*1xN</i>	9 (27%)	9 (15%)
<i>CYP2D6*2xN</i>	11 (33%)	21 (35%)
<i>CYP2D6*4xN</i>	7 (21%)	25 (42%)
<i>CYP2D6*10xN</i>	0	1 (1.7%)
<i>CYP2D6*17xN</i> [spacer]	0	1 (1.7%)
<i>CYP2D6*35xN</i>	1 (3%)	0
<i>CYP2D6*41xN</i>	0	1 (1.7%)
<i>CYP2D6*45xN</i>	0	1 (1.7%)
<i>CYP2D6*76</i>	0	1 (1.7%)
<i>CYP2D6*77</i>	3 (9%)	0
<i>CYP2D6*78</i>	1 (3%)	0