Tri-Allelic Haplotypes Determine and Differentiate Functionally Normal Allele CYP2D6*2 and Impaired Allele CYP2D6*41

Ulrich M. Zanger¹,²,*,†, Kyoko Momoi¹,², Ute Hofmann¹,², Matthias Schwab¹,²,³ and Kathrin Klein¹,²,†

CYP2D6 metabolizes 20–25% of all clinically used drugs and its complex genetic polymorphism is a major determinant of drug safety and efficacy. We investigated the basis for the functional difference between the two common alleles *2 (g.2851C>T + g.4181G>C, normal function) and *41 (additional intronic g.2989G>A, reduced function). A recently reported far-distant enhancer polymorphism rs5758550A/G linked to *2 has been suggested to play a decisive role. Genotyping of two white cohorts confirmed strong linkage of rs5758550G to *2, whereas no influence was found on metabolic ratio of sparteine or hepatic expression. Genomic plasmid constructs carrying individual variants or combinations thereof were expressed in COS1 and Huh7 cells. Both g.2851C>T(R296C) and g.2989G>A reduced enzyme activity and protein levels similarly by ~ 50–65% compared to reference (*1), whereas the double variant had only ~ 20% activity. Although the unexpected loss of function caused by g.2851C>T was compensated by g.4181G>C (mimicking the EM-phenotype of *2), the additional loss of function due to intronic g.2989G>A in the triple variant was not compensated (mimicking the IM-phenotype of *41). We also confirmed increased erroneous splicing in carriers of g.2989G>A but not of g.2851C>T as a likely explanation for the impaired function of *41. In conclusion, our data demonstrate g.2989G>A as causal variant of impaired allele CYP2D6*41 whereas triple-haplotypes have to be considered to explain the functional difference between *2 and *41. These data are important for genotyping strategies and clinical implementation of CYP2D6 pharmacogenetics.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
✓ CYP2D6*2 is a functionally normal allele, CYP2D6*41 is an impaired allele with highly similar sequence; an intronic variant and a far downstream enhancer polymorphism have alternatively been proposed as causal mutations.

WHAT QUESTION DID THIS STUDY ADDRESS?
✓ We addressed the questions whether the far downstream enhancer polymorphism influences in vivo phenotype and hepatic expression of CYP2D6 and which genetic variants are functionally relevant in an in vitro genomic expression system.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
✓ The study provides new experimental data on functional effects of several common CYP2D6 variants and how interactions between them change these effects.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?
✓ Our study resolves uncertainties about causal variants for allele *41. This makes CYP2D6 genotyping more reliable and supports implementation of CYP2D6 pharmacogenetics into clinical practice.

Cytochrome P450 2D6 (CYP2D6) is the primary metabolic enzyme for ~ 20% of clinically used drugs, including antidepressants, β-blockers, opioids, tamoxifen, and many others, and because of its well-investigated and complex genetic polymorphism constitutes a paradigm for pharmacogenomics.¹⁻³ Four distinct phenotypes commonly referred to as poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM), and ultrarapid metabolizer are primarily the result of germline inheritance of any two of > 100 variant alleles that vary across different populations and that cause absent, reduced, unchanged, or increased enzyme activity.⁴⁻⁸ These genetic differences have important consequences for drug safety and effectiveness, as reflected by an increasing number of (currently 71) drugs with pharmacogenomic drug label information⁹ and efforts to translate these into therapeutic guidelines¹⁰⁻¹¹ as a basis for clinical implementation of CYP2D6 pharmacogenetics.¹⁵⁻¹⁸

¹Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany; ²University of Tuebingen, Tuebingen, Germany; ³Departments of Clinical Pharmacology, and of Pharmacy and Biochemistry, University Tuebingen, Tuebingen, Germany. *Correspondence: Ulrich M. Zanger (uli.zanger@ikp-stuttgart.de)

†These authors contributed equally to this work.

Received June 26, 2020; accepted October 5, 2020. doi:10.1002/cpt.2078
In white patients, the most common impaired allele is CYP2D6*4, a null allele with a frequency of ~20–25% that leads to absence of CYP2D6 protein expression due to erroneous splicing and that is responsible for the majority of PMs. The second most frequent impaired allele in white patients with a frequency of ~10–15% is CYP2D6*41 that retains low residual activity and is clinically particularly relevant for carriers of a second impaired allele such as *4. For example, the dextromethorphan clearance attributable to one copy of CYP2D6*41 was estimated at 1.7% compared with CYP2D6*1 reference. A large venlafaxine drug monitoring study found the majority of CYP2D6*41/null allele carriers to be phenotypically in the range of PMs. Further recent studies confirmed a strong impact of *41 and/or the IM phenotype also for vortioxetine, risperidone, and aripiprazole, as well as endoxifen. Because of its high frequency and strong functional impact, reliable identification of CYP2D6*41 is important to avoid misclassification in clinical pharmacogenetics, and this requires knowledge of causative mutations. However, low activity associated with the CYP2D6*41 allele proved difficult to explain. Indeed, its sequence was found to be very similar to that of the common normal activity allele CYP2D6*2, with which it shares two amino acid variants R296C (g.2851C>T) and S486T (g.4181G>C). Initially a promoter variant (g.-1584C>G) had been suspected, but subsequent studies identified an intron 6 single-nucleotide polymorphism (SNP; g.2989G>A, rs28371725) as likely causal variant. In human liver samples, g.2989G>A was associated with increased amounts of a variant nonfunctional transcript lacking exon 6 (Δexon6) and lower amounts of normal transcript, suggesting partial erroneous splicing as a possible mechanism.

Since 2014, Wang and colleagues, in a series of papers characterized a novel polymorphic region, located ~115 kb downstream of CYP2D6 and comprising two SNPs rs133333A>G and rs5758550A>G in tight linkage disequilibrium (LD). The rs5758550A>G SNP apparently turns the region into an enhancer that interacts with the CYP2D6 promoter and increases transcription. Because this polymorphism is tightly linked to *2 but not to *41 (Figure 1), Wang and colleagues hypothesized that g.2851C>T could be a damaging variant that may be compensated for by increased transcription of haplotypes with active enhancer G-allele. By contrast, *41, being linked to the inactive enhancer A-allele, would confer impaired function due to lack of compensation of g.2851C>T, whereas g.2989G>A would not be relevant in this scenario. If proven correct, this would require some changes in the guidelines for CYP2D6 genotyping.

In this study, we functionally reassessed the highly related CYP2D6*2 and CYP2D6*41 alleles in vivo and in vitro. We genotyped available DNA samples for the enhancer polymorphisms and determined their impact on in vivo phenotype and liver expression. We then assessed, for the first time, effects of g.2851C>T, g.2989G>A, and g.4181G>C single variants and combinatorial haplotypes in the same genomic expression system. Our results provide insight into an intricate interplay between these polymorphisms at different levels of gene expression and they help to clarify current issues with CYP2D6*2 and CYP2D6*41 genotyping. They further emphasize the need to consider complex haplotypes and they warrant the use of genomic rather than cDNA expression systems for functional evaluation not only of noncoding but also of exonic genetic variants.

METHODS
Study cohorts
DNA samples were available from a cohort of healthy volunteers with known metabolic ratio for sparteine (MRS; n = 207) and from 150 liver donors with previously determined CYP2D6 genotype and expression phenotype. The study protocols were approved by the ethics committees of the medical faculties of the Charité, Humboldt University, and the University of Tübingen. The studies were conducted in accordance with the Declaration of Helsinki, and written informed consent was obtained from each participant. Clinical trial data were used anonymized.

![Figure 1](https://www.pharmvar.org/gene/CYP2D6). Schematic overview of relevant CYP2D6 allelic haplotypes and linkage with downstream enhancer variants. Grey boxes represent CYP2D6 exons and white boxes the downstream enhancer region with major A-allele and minor G-allele indicated. Black symbols denote the approximate location of variant positions as specified above. Some additional intronic variants shared by most *2, *35, and *41 suballeles are not shown (see https://www.pharmvar.org/gene/CYP2D6).
Nomenclature and genotyping

CYP2D6 allele nomenclature and base numbering is according to PharmacVar except that null-alleles were sometimes collectively termed ‘0’. CYP2D6 genotypes were available from previous studies.19,27 Genotyping for rs5758550A>G and rs1333333A>G was performed using the predeveloped assays C___29692254_10 and C___1088576_30, respectively (Thermo Fisher Scientific, Dreieich, Germany) on a TaqMan 7900HT (Thermo Fisher Scientific). About 10% of the genotyped samples were randomly selected and confirmed in repeated runs. Genotypes of both variants were in Hardy-Weinberg equilibrium (P > 0.4).

Genomic expression vectors for CYP2D6 variants

The previously described pCMV derivative of CYP2D6*1 containing the human genomic DNA region (CYP2D6*1.001) from + 1 (ATG) to 4,217 (stop) was used to introduce single mutations by consecutive site directed mutagenesis (Figure S1)27 (Quick change lightening kit; Agilent, Waldbronn, Germany). A 634 bp fragment covering nucleotide positions g.3548-g.4224 containing g.4181C was synthesized by GenScript (Piscataway, NJ) and subcloned into double and single variant plasmids as needed using restriction enzymes SfI/XbaI. Correct sequence including allele variants *2.001 and *41.001 was confirmed by Sanger sequencing (Microsynth Seqlab, Göttingen, Germany).

Cell culture and transfection

Human hepatoma cells Huh7 and COS1 cells were cultured as described.22 Cells were transfected with DNA mix containing 3 µg of plasmid DNA and 600 ng of the β-galactosidase expression plasmid pCMV-LacZ (Clontech Takara Biotech, Mountain View, CA) in 1.4 mL of serum-free medium (Gibco Opti-mem; Thermo Fisher Scientific) using Turbofect reagent (Thermo Fisher Scientific) according to the manufacturer’s instructions. Cells were washed with phosphate-buffered saline (PBS) 48-hour post-transfection and harvested by trypsin treatment (0.25 mM; Gibco, Thermo Fisher Scientific). Total RNA was isolated using Rneasy-mini-kit (Qiagen, Hilden, Germany) with on-column DnaseI treatment and quantification by Nanodrop (Thermo Fisher Scientific). Cells harvested for protein analysis were disrupted using passive lysis buffer (Promega GmbH, Mannheim, Germany). Protein concentration was determined using Pierce BCA protein assay (Thermo Fisher Scientific) on an Empire reader (Perkin Elmer, Waltham, MA). Beta-galactosidase activity was measured using the Galacto-Light system (Thermo Fisher Scientific) on an AutoLumat Plus luminometer (Berthold Technologies, Bad Wildbad, Germany). In total, four independent transfections were performed for each cell line and each transfection was performed in duplicate.

Quantification of protein and enzyme activity

CYP2D6 protein was quantified in 50 µg whole cell lysates by Western blotting using monoclonal antibody mab114.26 Enzyme activity was quantified 48-hour post-transfection by adding 3 mM propafenone to fresh culture medium and incubation for 3 hours at 37°C. Liquid chromatography tandem mass spectrometry analysis of metabolites was performed on an Agilent 1290 LC system coupled to an Agilent 6460 triple quadrupole mass spectrometer.35

RNA transcript analysis

Multiscribe reverse transcriptase (Thermo Fisher Scientific) and random hexamers were used to transcribe 2 µg of total RNA to a final cDNA concentration of 100 ng/µL. Quantification of CYP2D6 normal transcript, Δexo6n-splice variant, and total mRNA was performed by Taqman real-time polymerase chain reaction with specific primer/probe assays (Table S1) on a Taqman 7900HT (Thermo Fisher Scientific) and specificity was confirmed using plasmids containing appropriate inserts.27 Serial dilutions of these plasmids were used for calibration. Expression of glyceraldehyde 3-phosphate dehydrogenase(Hs02758991_g1) or RPLP0 (4326314E) were measured as housekeeping genes (Applied Biosystems). Cellular expression was corrected for transfection efficiency using β-galactosidase, normalized to glyceraldehyde 3-phosphate dehydrogenase, and presented relative to the mean of *1 of four separate duplicate transfections. As reported earlier, the Δexo6n splice variant was highly abundant in all transfected cells (including *1) preventing reliable quantification.27

Statistics

For statistical analysis of differences in expression of transcript variants, protein or activity from CYP2D6 minigenes, repeated measures analysis of variance (ANOVA) was applied on log-transformed data with Bonferroni multiple comparison correction. For comparisons of two groups, t-test statistics was applied. GraphPad Prism version 5 (GraphPad, La Jolla, CA) was used for all calculations with significance level set to P < 0.05.

RESULTS

LD analysis of enhancer polymorphisms and standard CYP2D6 alleles

Genotyping of rs5758550A/G and rs1333333A/G in DNA samples from sparteine-phenotyped volunteers (n = 207)29 and in a human liver cohort (n = 150)32 revealed frequencies of the variant A-allele of 77% and 79%, respectively, corresponding well to European population data (1000G,34 75.4%; gnomAD,35 73.8%). With only three discordant samples, the two SNPs were almost completely linked to each other (> 99%). The minor G-allele was strongly linked to the g.2851T-allele of *2 and related alleles *35 and *59 (Table S1). Genotypes of *35 and *59 carriers were in all cases compatible with linkage of these two *2-related alleles to the enhancer G-allele, whereas no definitive association to *41 was found. Conversely, only one of 58 *1/*1 homozygotes and two of 28 PMs (both with genotype *4/*4) were heterozygous carriers of rs5758550G (Table 1).

Impact of rs5758550 on CYP2D6 drug oxidation and gene expression phenotypes

We first analyzed whether carriers of rs5758550G-allele show a systematic phenotypic deviation from A-allele carriers. As shown in Figure 2, MR*G_s distributions did not differ with respect to rs5758550 in any activity-score subgroup. As evident from Figure 3, phenotypic MR*G_s distributions were very similar for corresponding genotypes encompassing *1, *2, and *35 alleles, with none of the differences being statistically significant. Importantly, this includes three *2/*0 subjects with genotype rs5758550AA (i.e., carriers of *2 alleles with low activity enhancer variant). For comparison, the median MR*G_s of 11 *41/*0 subjects (all rs5758550AA homozygotes) was significantly higher compared with *1/*0 (P = 0.0006) and *2/*0 (P = 0.0007, Figure 3). Further comparisons among other subgroups also did not support a consistent influence of the enhancer genotype on MR*G. The enhancer genotype also failed to show an impact on microsomal CYP2D6 protein and enzyme activity phenotypes (Figure S2 and data not shown).

In vitro functional assessment of isolated variants and combinatorial haplotypes in genomic context

We first constructed genomic minigene expression plasmids on reference (*1) background carrying either g.2851C>T or
g.2989G>A (Figure S1). Surprisingly, transfection of COS1 and HuH7 cells revealed ~ 60–80% reduced enzyme activity, immuno-detectable protein, and normal transcript levels for both single variants (Figure 4). All analyte levels were significantly lower (P < 0.0001) than those obtained with reference or *2 plasmids (e.g., compare haplotypes T-G-G and C-A-G with *1(C-G-G), but they did not differ from each other or from *41 allele construct (Figure S3).

Because the *2 and *41 constructs share several additional variants, we suspected SNP interactions, especially with the exon 9 variant g.4181G>C to influence genotype-phenotype relationships. Indeed, as single variant g.4181G>C enhanced the function of the reference allele at all phenotype levels. In biallelic combinations, it restored protein and activity of both g.2851C>T and g.2989G>A single variants to normal levels and partially also their normal transcript levels (compare T-G-C vs. T-G-G and C-A-C vs. C-A-G; Figure 4). Thus, biallelic haplotypes also failed to explain the phenotypic difference between *2 and *41 alleles.

We next introduced both key variants g.2851C>T and g.2989G>A simultaneously (T-A-G) and analyzed the influence of g.4181G>C in tri-allelic haplotypes. The double variant (T-A-G) resulted in phenotype levels even below those of *41 but strikingly g.4181G>C failed to fully restore phenotype in the triple-variant T-A-C. Thus, only the triple-variant closely mimicked *41 on all phenotype levels (Figure 4).

Taken together, these results demonstrated strong impairment of CYP2D6 expression and function by either single variant, g.2851C>T and g.2989G>A, when tested in a genomic context. The g.4181G>C variant, which occurs on many different alleles including *2 and *41, can rescue both variants individually, but not when both are present as in *41.

**Differential transcript expression in human liver samples**

We reanalyzed expression of the normal CYP2D6 transcript and the Δexon6 splice variant in liver samples using newly developed

---

| Allele  | Allele counta | Allele frequency |
|--------|---------------|-----------------|
| CYP2D6*1 | 269 | 0.390 |
| *1[A]   | 266 | 0.386 |
| *1[G]   | 3  | 0.004 |
| CYP2D6*2 | 105 | 0.152 |
| *2[A]   | 12  | 0.017 |
| *2[G]   | 93  | 0.136 |
| CYP2D6*35 | 42  | 0.061 |
| *35[A]  | 0   | 0.000 |
| *35[G]  | 42  | 0.061 |
| CYP2D6*41 | 58  | 0.084 |
| *41[A]  | 58  | 0.084 |
| *41[G]  | 0   | 0.000 |
| CYP2D6*4 | 144 | 0.209 |
| *4[A]   | 142 | 0.206 |
| *4[G]   | 2   | 0.003 |
| CYP2D6*5 | 20  | 0.029 |
| *5[A]   | 17  | 0.025 |
| *5[G]   | 3   | 0.004 |

*aTotal number of alleles n = 690 (samples with gene duplications are excluded).**
transcript-specific TaqMan assays. Similar to our earlier results, we found ~3-fold higher levels of Δexon6 splice variant in carriers of *41 compared with *1/*1, whereas the normal transcript was significantly reduced (Figure 5). In contrast, transcript distribution of carriers of *2/*2 and *2/*35 was similar as for *1/*1 carriers.

**DISCUSSION**

The goal of this study was to clarify the causal involvement of candidate SNPs and haplotypes in the functionally impaired IM-phenotype associated with the CYP2D6*41 allele. We provide convincing evidence that (i) presence of the intron 6 SNP g.2989G>A is required for the impaired function of *41 on the level of mRNA, protein and activity; (ii) the *2 key variant g.2851C>T (R296C) is similarly damaging as g.2989G>A, if assessed in genomic context; (iii) only tri-allelic haplotypes, including these two and the g.4181 SNP, can explain genotype-phenotype relationships associated with CYP2D6*2 and *41 alleles; (iv) the downstream enhancer polymorphism rs5758550 is neither required to explain the phenotypic difference between *2 and *41 alleles nor does it affect phenotype significantly; and (v) although partial erroneous splicing is likely involved, the exact mechanism(s) by which g.2851C>T and g.2989G>A lead to impaired expression of CYP2D6 remains to be investigated.

According to the hypothesis raised by Wang et al.,28 CYP2D6*2 linked to the low function enhancer A-allele should be impaired and confer the IM-phenotype, at least in combination with a null allele. Our in vivo data do not support this proposition. We identified three null-allele carriers with a *2-allele definitively linked to the enhancer A-allele, and all three had MRₜ values in the EM-range, comparable to other heterozygous *1 and *2 genotypes (Figure 3, Table S2). Considering these and other analyses in our cohorts (Figure 2, Figure S2), a consistent contribution of the functional enhancer allele to phenotype was not apparent. This is in agreement with a report on patients with breast cancer treated with tamoxifen, where rs5758550 had no effect on endoxifen levels in several identified CYP2D6*2 carriers with and without functional enhancer and no improvement of phenotype prediction was obtained after reclassification.36 Another study in human livers also failed to detect an influence of rs5758550 on CYP2D6 protein and activity levels.37

Nevertheless, the causal SNP for functional impairment of *41 remained unclear due to the rare occurrence of informative haplotypes. To prove causal involvement, phenotyped carriers of either SNP, linked to the functional and nonfunctional enhancer variants, would be required in suitable diplotypes. Given the tight LD between these variants such genotypes are too rare to be observable even in large studies. We therefore turned again to in vitro studies, realizing that the relevant variants had never been tested alone (g.2989G>A27) or only as cDNA variants but not in genomic context. Data obtained in different cDNA expression systems and with different substrates indicate that expressed CYP2D6.34 (R296C) and CYP2D6.2 (R296C + S486T) are functionally similar proteins with slightly decreased, similar, or slightly increased function compared with CYP2D6.1, to some degree depending on the substrate.26,38–41 For comparison, the expressed IM-variants CYP2D6.10 (P34S + S486T) and CYP2D6.17 (T107I + R296C + S486T) had only ~10–30% of wild type activity.38,40 Taken together, current in vitro data suggest the isolated
R296C variation to be considered as functionally benign, in agreement with most prediction algorithms. However, as shown by our expression studies using full-length genomic CYP2D6 constructs, cDNA-based expression systems may miss important functional aspects. Unexpectedly, both isolated variants g.2851C>T and g.2989G>A proved to be highly damaging to CYP2D6 expression on mRNA, protein, and activity level (Figure 4). Based on our concept, we would have expected only g.2989G>A to be damaging, whereas according to Wang and colleagues, only g.2851C>T should be damaging (Figure 6). As the enhancer SNP could be ruled out, we suspected SNP-interactions to be involved. Indeed, our double-allelic and triple-allelic haplotype constructs revealed a strong compensatory function of g.4181G>C that was able to fully recover decreased protein and activity levels of g.2851C>T and g.2989G>A in the double-variants, but not in the triple-variant. Thus, only the triple-variant haplotype closely resembled the native *41-construct at all levels of gene expression.

We also readdressed the mechanism by which g.2989G>A and g.2851C>T lower expression. Using newly developed specific TaqMan assays, we found significantly lower levels of CYP2D6 normal transcript and simultaneously increased levels of the Δexon6 splice variant for genotypes that include one *41 allele (Figure 5). This confirms our previous hypothesis of erroneous splicing as cause for functional impairment. However, in contrast to Wang and colleagues, we found no difference in these splice patterns in livers from homozygous *2 carriers, despite a significant reduction of normal transcript levels of the g.2851C>T variant (Figure 4a). It is possible that the variant facilitates the formation of other splice variants than Δexon6 (e.g., a variant lacking exon 3). The mechanism of the damaging properties of g.2851C>T in genomic context thus remains to be clarified. Another interesting question concerns the mechanism by which g.4181G>C compensates for the damaging effects of g.2989G>A and g.2851C>T. In our genomic expression system, the variant itself led to higher protein expression, and in combinatorial variants its compensatory effect was also more efficient on protein than on mRNA levels (Figure 4).

Finally, our new findings resolve current uncertainties regarding definition of the *41 allele. In particular, they confirm g.2989G>A as causal damaging variant and thus also the current core allele definition for *41. This should help to make CYP2D6 genotyping simpler and more reliable, and support its implementation into clinical practice. The g.2989G>A variant has also been found in other constellations in alleles termed *32, *69, *91, *119, and *138. Except for statistical analysis see Figure S3.
for *69, a null allele, no phenotype information is yet available. Given the dominant negative effect of g.2989G>A it seems very likely that all these alleles would be associated with IM-phenotype or PM-phenotype. Regarding g.2851C>T, our study identified this variant as highly damaging, in contrast to current belief and prediction. Interestingly, one individual confirmed by sequencing to carry *34 (isolated g.2851C>T, previously termed *1D) had been described as a misclassified PM. Unfortunately, further data are rare or absent but because it cannot be excluded that *34 occurs at higher frequencies in particular ethnic populations, genotyping for g.4181G>C should always be performed. On the other hand, the far downstream enhancer polymorphism rs5758550A>G. Experimental results of genomic constructs showed that both g.2851C>T (c) and g.2989G>A (d) variants and that both can be fully compensated by g.4181G>C. (e) Only the triple-haplotype construct correctly mimics the impaired*41 allele, because the combined deleterious effects of both damaging variants cannot be fully compensated by g.4181G>C.
In conclusion, we provide novel evidence for a strongly damaging effect of both isolated key variants g.2851C>T (*2) and g.2989G>A (*41) when assessed in a genomic expression system. Although the common exon 9 SNP g.4181G>C (S486T) can functionally compensate the damaging effects of g.2851C>T, thus leading to the EM-phenotype associated with *2, the combined effect of both variants cannot be compensated, leading to the IM-phenotype associated with *41. Because of these intricate functional relationships, and because these variants may not be strongly linked in all ethnic populations, we recommend that genotyping platforms should test for all three variants and not rely on single variants or surrogate markers. Furthermore, in vitro results obtained for coding variants in cDNA expression systems should be interpreted with caution, as genomic effects may escape detection. Finally, the importance of haplotypes for functional classification and prediction by algorithms appears underappreciated.

SUPPORTING INFORMATION
Supplementary information accompanies this paper on the Clinical Pharmacology & Therapeutics website (www.cpt-journal.com).

ACKNOWLEDGMENTS
The authors thank Igor Liebermann, Britta Klumpp, and Markus König, Stuttgart, for excellent technical assistance. Open access funding enabled and organized by ProjektDEAL.

FUNDING
The study was supported by the Robert Bosch Stiftung, Stuttgart, Germany and the European Union Research and Innovation Program Horizon 2020 grant U-PGx 668353.

CONFLICT OF INTEREST
U.M.Z. is co-inventor of several patent applications on the detection of specific CYP2D6 polymorphisms for diagnostic purposes. All other authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS
U.M.Z., K.K., and M.S. wrote the manuscript. U.M.Z. and K.K. designed the research. K.M., U.H., and K.K. performed the research. K.M., U.H., and K.K. analyzed data. M.S. and U.M.Z. contributed analytical tools.

© 2020 The Authors. Clinical Pharmacology & Therapeutics published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

1. Meyer, U.A. Pharmacogenetics – five decades of therapeutic lessons from genetic diversity. Nat. Rev. Genet. 5, 669–676 (2004).
2. Eichelbaum, M., Ingelman-Sundberg, M. & Evans, W.E. Pharmacogenomics and individualized drug therapy. Annu. Rev. Med. 57, 119–137 (2006).
3. Zanger, U.M. & Schwab, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol. Ther. 138, 103–141 (2013).
4. Daly, A.K. et al. Nomenclature for human CYP2D6 alleles. Pharmacogenetics 6, 193–201 (1996).
5. Zanger, U.M., Raimundo, S. & Eichelbaum, M. Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. Naunyn Schmiedebergs Arch. Pharmacol. 369, 23–37 (2004).
6. Ingelman-Sundberg, M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J. 5, 6–13 (2005).
7. Nofziger, C. et al. PharmVar GeneFocus: CYP2D6. Clin. Pharmacol. Ther. 107, 154–170 (2020).
8. Gaedigk, A., Sangkuhl, K., Whirl-Carrillo, M., Klein, T. & Leeder, J.S. Prediction of CYP2D6 phenotype from genotype across world populations. Genet. Med. 19, 69–76 (2017).
9. US Food and Drug Administration (FDA). Table of Pharmacogenomic Biomarkers in Drug Labeling: Content current as of 02/05/2020 <https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling>. Accessed June 25, 2020.
10. Crews, K.R. et al. Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450 2D6 genotype and codeine therapy: 2014 update. Clin. Pharmacol. Ther. 95, 376–382 (2014).
11. Hicks, J.K. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. Clin. Pharmacol. Ther. 98, 127–134 (2015).
12. Bell, G.C. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2D6 genotype and use of ondansetron and trolpsetron. Clin. Pharmacol. Ther. 102, 213–218 (2017).
13. Goetz, M.P. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and tamoxifen therapy. Clin. Pharmacol. Ther. 103, 770–777 (2018).
14. Caudle, K.E. et al. Standardizing CYP2D6 genotype to phenotype translation: consensus recommendations from the clinical pharmacogenomics implementation consortium and Dutch Pharmacogenetics Working Group. Clin. Transl. Sci. 13, 116–124 (2020).
15. Dunnenberger, H.M. et al. Preemptive clinical pharmacogenetics implementation: current programs in five US medical centers. Annu. Rev. Pharmacol. Toxicol. 55, 89–106 (2015).
16. Kalman, L.V. et al. Pharmacogenetic allele nomenclature: international workshop recommendations for test result reporting. Clin. Pharmacol. Ther. 99, 172–185 (2016).
17. van der Wouden, C.H. et al. Implementing pharmacogenomics in Europe: design and implementation strategy of the ubiquitous pharmacogenomics consortium. Clin. Pharmacol. Ther. 101, 341–358 (2017).
18. Cavallari, L.H. et al. Multi-site investigation of strategies for the clinical implementation of CYP2D6 genotyping to guide drug prescribing. Genet. Med. 21, 2255–2263 (2019).
19. Raimundo, S. et al. A novel intrinsic mutation, 2988GA, with high predictivity for impaired function of cytochrome P450 2D6 in white subjects. Clin. Pharmacol. Ther. 76, 128–138 (2004).
20. Abduljalil, K. et al. Assessment of activity levels for CYP2D6*1, CYP2D6*2, and CYP2D6*41 genes by population pharmacokinetics of dextromethorphan. Clin. Pharmacol. Ther. 88, 643–651 (2010).
21. Haslamo, T., Eliasson, E., Jukić, M.M., Ingelman-Sundberg, M. & Molder, E. Significantly lower CYP2D6 metabolism measured as the O/N-desmethylvenlafaxine metabolic ratio in carriers of CYP2D6*41 versus CYP2D6*9 or CYP2D6*10: a study on therapeutic drug monitoring data from 1003 genotyped Scandinavian patients. Br. J. Clin. Pharmacol. 85, 194–201 (2019).
22. Frederiksen, T., Areberg, J., Schmidt, E., Bjergaard, S.T. & Bräsen, K. Quantification of in vivo metabolic activity of CYP2D6 genotypes and alleles through population pharmacokinetic analysis of voroxetine. Clin. Pharmacol. Ther. 109, 150–159 (2021).
23. Jukic, M.M., Smith, R.L., Haslamo, T., Molder, E. & Ingelman-Sundberg, M. Effect of CYP2D6 genotype on exposure and efficacy of risperidone and aripiprazole: a retrospective, cohort study. Lancet Psychiatry 6, 418–426 (2019).
24. Thorén, L. et al. Impairment of endoxifen formation in tamoxifen-treated premenopausal breast cancer patients carrying
reduced-function CYP2D6 alleles. Br. J. Clin. Pharmacol. https://doi.org/10.1111/bcp.14500. [e-pub ahead of print].

25. Raimundo, S., Fischer, J., Eichelbaum, M., Griese, E.U., Schwab, M. & Zanger, U.M. Elucidation of the genetic basis of the common 'intermediate metabolizer' phenotype for drug oxidation by CYP2D6. Pharmacogenet. Genomics 10, 577–581 (2000).

26. Zanger, U.M. et al. Comprehensive analysis of the genetic factors determining expression and function of hepatic CYP2D6. Pharmacogenet. Genomics 11, 573–585 (2001).

27. Toscano, C. et al. Impaired expression of CYP2D6 in intermediate metabolizers carrying the *41 allele caused by the intronic SNP 2988GA: evidence for modulation of splicing events. Pharmacogenet. Genomics 16, 755–766 (2006).

28. Wang, D., Poi, M.J., Sun, X., Gaedigk, A., Leeder, J.S. & Sadee, W. Common CYP2D6 polymorphisms affecting alternative splicing and transcription: long-range haplotypes with two regulatory variants modulate CYP2D6 activity. Hum. Mol. Genet. 23, 268–278 (2014).

29. Wang, D., Papp, A.C. & Sun, X. Functional characterization of CYP2D6 enhancer polymorphisms. Hum. Mol. Genet. 24, 1556–1562 (2015).

30. Ray, B., Ozcagli, E., Sadee, W. & Wang, D. CYP2D6 haplotypes with enhancer single-nucleotide polymorphism rs5758550 and rs16947 (*2 allele): Implications for CYP2D6 genotyping panels. Pharmacogenet. Genomics 29, 39–47 (2019).

31. Boone, E.C. et al. Long-distance phasing of a tentative "enhancer" single-nucleotide polymorphism with CYP2D6 star allele definitions. Front. Pharmacol. 11, 486 (2020).

32. Klein, K. et al. A new panel-based next-generation sequencing method for ADME genes reveals novel associations of common and rare variants with expression in a human liver cohort. Front. Genet. 10, 7 (2019).

33. Feidt, D.M. et al. Profiling induction of cytochrome p450 enzyme activity by statins using a new liquid chromatography-tandem mass spectrometry cocktail assay in human hepatocytes. Drug Metab. Dispos. 38, 1589–1597 (2010).

34. Auton, A. et al. A global reference for human genetic variation. Nature 526, 68–74 (2015).

35. Karczewski, K.J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443 (2020).

36. Sanchez-Spitman, A.B., Moe, D.-J.A., Gelderblom, H., Dezentjé, V.D., Swen, J.J. & Guchelaar, H.J. The effect of rs5758550 on CYP2D6*2 phenotype and formation of endoxifen in breast cancer patients using tamoxifen. Pharmacogenomics 18, 1125–1132 (2017).

37. Ning, M., Duarte, J.D., Rubin, L.H. & Jeong, H. CYP2D6 protein level is the major contributor to interindividual variability in CYP2D6-mediated drug metabolism in healthy human liver tissue. Clin. Pharmacol. Ther. 104, 974–982 (2018).

38. Marcucci, K.A., Pearce, R.E., Crespi, C., Steinmel, D.T., Leeder, J.S. & Gaedigk, A. Characterization of cytochrome P450 2D6.1 (CYP2D6.1), CYP2D6.2, and CYP2D6.17 activities toward model CYP2D6 substrates dextromethorphan, bufuralol, and debrisoquine. Drug Metab. Dispos. 30, 595–601 (2002).

39. Sakuyama, K. et al. Functional characterization of 17 CYP2D6 allelic variants (CYP2D6.2, 10, 14A-B, 18, 27, 36, 39, 47–51, 53–55, and 57). Drug Metab. Dispos. 36, 2460–2467 (2008).

40. Niwa, T., Murayama, N. & Yamazaki, H. Comparison of cytochrome P450 2D6 and variants in terms of drug oxidation rates and substrate inhibition. Curr. Drug Metab. 12, 412–435 (2011).

41. Dong, A.N., Ahemad, N., Pan, Y., Palanisamy, U.D., Yiap, B.C. & Ong, C.E. Functional and structural characterisation of common cytochrome P450 2D6 allelic variants—roles of Pro34 and Thr107 in catalysis and inhibition. Naunyn Schmiedebergs Arch. Pharmacol. 392, 1015–1029 (2019).

42. Dong, C. et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum. Mol. Genet. 24, 2125–2137 (2015).

43. Zhou, Y., Mkrchian, S., Kumondai, M., Hiratsuka, M. & Lauschke, V.M. An optimized prediction framework to assess the functional impact of pharmacogenetic variants. Pharmacogenomics J. 19, 115–126 (2019).

44. Marez, D. et al. Polymorphism of the cytochrome P450 2D6 gene in a European population: characterization of 48 mutations and 53 alleles, their frequencies and evolution. Pharmacogenomics 7, 193–202 (1997).

45. Elias, A.B.R., Araújo, G.S.D., Souza, S.J. & Suarez-Kurtz, G. Distribution and linkage disequilibrium of the enhancer SNP rs5758550 among Latin American populations: influence of continental ancestry. Pharmacogenet. Genomics 30, 67–72 (2020).