Characterization of Cytochrome P450 2D6 Alleles Using the Invader™ System

Matt Neville¹, Rebecca Selzer², Brian Aizenstein¹, Megan Maguire¹, Kirk Hogan², Robert Walton³, Ken Welsh⁴, Bruce Neri¹, and Monika de Arruda¹

¹Third Wave Technologies, Madison, WI, USA; ²University of Wisconsin-Madison, Department of Anesthesiology, Madison, WI, USA; ³Cancer Research UK GPRG, Department of Clinical Pharmacology, Oxford, UK; and ⁴Clinical Genomics, London, UK

BioTechniques 32:S34-S43 (June 2002)

ABSTRACT

The cytochrome p450 (CYP) superfamily comprises enzymes that play an essential role in the transformation of medically relevant compounds. Accurate genotyping of polymorphisms in members of this family is drawing increasing interest because certain allelic variants may result in either loss of efficacy or toxic accumulation of therapeutic agents. Debrisoquine 4-hydroxylase, or CYP2D6, is among the most widely studied of the CYPs. The complexity of the CYP2D6 genomic region, including pseudogenes, gene deletions, and gene duplications, has offered numerous challenges to developing a genotyping strategy. We describe a comprehensive CYP2D6 genotyping strategy that employs both a PCR/Invader™ genotyping assay system and an Invader® genomic copy number assay. The Invader system is a homogeneous, isothermal, highly specific, and robust signal amplification system. Results from 11 CYP2D6 assays in an allele frequency study compare well to published allele frequency values for Caucasians. Further, Invader assays provided unambiguous genotyping determinations for 100% of the 171 samples that yielded a visible PCR product on an agarose gel. A copy number assay yielded only one equivocal result in 205 samples. We identified 17 single-copy individuals and 17 three-copy (or more) individuals.

INTRODUCTION

The field of pharmacogenetics is advancing rapidly as increasing numbers of functional polymorphisms are identified in genes encoding proteins essential for drug action. One clinically important protein is an enzyme in the cytochrome P450 family, debrisoquine 4-hydroxylase or cytochrome p450 2D6 (CYP2D6), encoded by the *CYP2D6* gene on chromosome band 22q13.1. The enzyme metabolizes about 25% of all therapeutic drugs (reviewed in Reference 6) as well as many environmental xenobiotic substances. Hence, variation in enzyme activity has been linked to a wide range of illnesses, including liver cancer (1) and Parkinson's disease (35).

Currently, more than 70 polymorphisms within the exonic and promoter regions of *CYP2D6* have been identified, as have more than 70 haplotypes (see <http://www.imm.ki.se/CYPalleles/cyp2d6.htm>). Many genetic variations (more than 20 polymorphisms, a gene deletion, and numerous gene conversion events) decrease CYP2D6 activity. Depending on ethno-geographic ori-

gins, the overall incidence of poor metabolizer (PM) status ranges between 1% and 8% (32). In addition, duplications in *cis* and *trans* have been associated with ultra-rapid metabolic (UM) states, which have an incidence of 1% to 7% in distinct populations (16,23).

Two *CYP2D6* pseudogenes, *CYP2D7* and *CYP2D8*, lie directly upstream of the approximately 5-kb *CYP2D6* gene (Figure 1A). Both are highly homologous (97% and 92%, respectively) to the exonic sequence of *CYP2D6* (19), so that many polymorphic alleles found in *CYP2D6* also occur in the pseudogenes. Despite the importance of the CYP2D6 enzyme in drug metabolism and adverse drug effects, the complexity of its genomic region has hampered development of genetic tests for predicting clinically relevant phenotypes. Here we describe a simple, scalable, and comprehensive *CYP2D6* genotyping strategy that couples selective amplification of the *CYP2D6* locus with the specificity of the Invader reaction.

PCR/Invader Strategy

To prevent false-positive results or inflated wild-type signals caused by hybridization of Invader oligos to the two adjacent pseudogenes, we designed PCR primers to specifically amplify *CYP2D6* and used these amplicons as templates for interrogation by the Invader reaction.

The Invader DNA assay relies on the structure-specific activity of Cleavase® DNA enzymes and on the sequence-specific hybridization of synthetic oligos to a DNA target. Cleavase enzymes are members of the structure-specific 5' flap endonuclease (FEN) family (13,25) that play critical roles in DNA replication, recombination, and repair (reviewed in Reference 21). The preferred enzyme substrate consists of distinct upstream, downstream, and template strands, which resemble the replication fork generated during DNA synthesis. In nature, these enzymes perform DNA repair by removing the overlapping 5' fragments generated by DNA polymerases during strand-displacement synthesis (20). In this process, they create a nicked strand, which becomes a substrate for DNA ligase (18,24).

The synthetic components of the Invader assay mimic the upstream and downstream strands of the natural substrate. When an upstream Invader oligo and a downstream allele-spe-

cific probe hybridize in tandem to the complementary DNA target (PCR product or genomic DNA), they create an invasive structure, the substrate for the Cleavase enzyme (18,26,27) (Figure 2A). The probe has two regions, a hybridizing target-specific region and a non-hybridizing "5' flap." The Cleavase enzyme recognizes this invasive structure and severs the unpaired 5' flap of the probe, releasing it as a target-specific product.

Rapid cycling of cleaved and uncleaved probes is achieved by incubating the reaction near the melting temperature (T_m) of the probes, which are present in excess (27,31). The Invader oligo is designed to remain hybridized to the DNA target at the reaction temperature. Thus, under the isothermal reaction conditions, multiple probes are cleaved for each target molecule, resulting in a linear amplification of the number of released 5' flaps with respect to target concentration (27). In a concurrent secondary reaction within the same reaction well, the cleaved 5' flaps hybridize to a fluorescence resonance energy transfer (FRETTM) cassette to create another invasive structure that is cleaved by the Cleavase enzyme (Figure 2B). Cleavage of the FRET cassette separates the fluorophore (F) and quencher (Q) dyes, generating a fluorescence signal that can be readily detected with standard fluorescence plate readers.

The Cleavase enzyme produces approximately 10^3 – 10^4 cleavage events per target sequence per hour. As described by Hall and co-workers (12), the coupled, serial version of the Invader assay yields signal amplification on the order of 10^6 – 10^7 .

Because the sequences of the cassette and 5' flap of the probe are independent of the target, these components can remain constant in assays designed to detect different DNA targets. Consequently, the universal nature of the cassettes and 5' flaps simplifies assay design and reduces costs in new assay development. Further, assays for different targets can be configured to have identical reaction conditions, enabling Invader assays for multiple targets to be performed simultaneously on the same microtiter plate. These features make this technology well-suited for high-throughput applications in 96-, 384-, or 1536-well microtiter plate formats.

Figure 2A illustrates how to utilize the Invader assay to discriminate between an A or a G nucleotide in an A/G polymorphism. When the "A" allele-specific probe hybridizes to the "A" allele target (as in Figure 2A), the Invader oligo and the "A" probe form a cleavable invasive substrate. The 5' flap is cleaved, the secondary reaction proceeds, and detectable signal is generated. Conversely, when the "A" allele-specific probe hybridizes to the "G" allele target in Figure 2A, the target, the Invader oli-

Figure 1. The *CYP2D6* genomic region and triplex PCR strategy. (A) The position of *CYP2D6* in relation to its two pseudogenes *CYP2D7* and *CYP2D8*. (B) The positions of polymorphisms found within or bordering the nine *CYP2D6* exons. The length of the vertical arrow indicates the relative frequency of the different polymorphisms. Solid arrows indicate non-synonymous polymorphisms and hatched arrows synonymous polymorphisms. The position and base change of each polymorphism are indicated at the end of each arrow. The asterisks (*) below the arrows indicate the 11 polymorphisms investigated in this study. The position and size of the three PCR products in the triplex PCR are indicated below the exons. (C) An example of the PCR products generated in triplex PCR as viewed on an agarose gel with ethidium bromide staining.

RESEARCH REPORT

go, and the "A" probe do not form the invasive substrate. Thus, the probe is not cleaved, and signal generation is suppressed.

The biplex format enables simultaneous detection of two DNA sequences in a single well, such as two variants of a particular polymorphism or two distinct loci. This format features two different allele-specific probes, each with a unique 5' flap, and two different FRET cassettes, each with a spectrally distinct fluorophore. By design, the released 5' flaps will hybridize only to their respective cassettes to generate a target-specific signal.

The combination of sequence-specific hybridization and structure-specific cleavage has enabled the development of ho-

mogeneous, isothermal, and highly specific assays for accurate resolution of genetic variations, such as single nucleotide polymorphisms (SNPs) or insertions/deletions from both unamplified genomic DNA (12,14,26) and PCR products (29,30,37).

Although this paper focuses on DNA applications using FRET detection, this technology can use multiple detection platforms (4,11,15) and can be applied equally well to qualitative and quantitative RNA detection (3,7).

MATERIALS AND METHODS

CYP2D6-Specific Triplex PCR for Genotyping

The *CYP2D6* region spans 5 kb of genomic sequence. Thus, the success of long-range PCR amplification depends on template quality. To improve the success rate of *CYP2D6* PCR amplification, the entire gene sequence was amplified as three shorter, non-overlapping fragments in a single triplex reaction. All primers were designed within 5', 3', or intronic *CYP2D6* regions and had a T_m of 68°C (see Table 1 for primer sequences). Primer pair 1 amplifies exons 1 and 2, primer pair 2 amplifies exons 3 to 6, and primer pair 3 amplifies exons 7 to 9, generating 2036-bp, 1683-bp, and 1754-bp products, respectively.

DNA samples from a group of 181 anonymous donors were used in the allele frequency study. The samples were numbered, unlinked, and tested anonymously. DNA was isolated using the QIAamp® whole blood kit (Qiagen, Valencia, CA, USA). The *CYP2D6*-specific triplex PCRs were performed using the Hercules® Hotstart PCR system (Stratagene, La Jolla, CA, USA) with 10–200 ng of genomic DNA, 250 µM dNTPs, 0.4 µM of each primer, 2% dimethyl sulfoxide, and 2.5 U of the supplied enzyme in a final volume of 50 µL. The reactions were incubated on a PCR Express thermal cycler (Thermo Hybaid, Franklin, MA, USA) using the following cycling parameters: 95°C for 5 minutes, followed by 35 cycles of 95°C for 30 seconds and 68°C for 4 minutes, then a final 10-minute 68°C extension cycle. For verification purposes, 10 µL of the PCR product was observed on a 1% agarose gel containing ethidium bromide. Figure 1C provides an example of the three PCR products generated in this step.

CYP2D6 Genotyping Assays

Assays were designed for the following *CYP2D6* polymorphisms: *CYP2D6**2-2850C→T; *2-4180G→C; *3-2549A→Del; *4-1846G→A; *6-1707T→Del; *10-100C→T; *11-883G→C; *18-4125GTGCCCCACT→Duplication; *33-2483G→T; *35-31G→A; and *37-1943G→A. The number immediately after the asterisk represents the *CYP2D6* allelic variant and the number after the hyphen represents the position of the polymorphism in relation to the translational start codon (5). The asterisks in Figure 1B indicate the relative positions of the 11 genotyping assays. The *CYP2D6**2, *10, *33, and *35 alleles are among the most common functional alleles in Caucasians aside from *CYP2D6**1. Conversely, the *CYP2D6**3, *4, and *6 alleles are among the most common nonfunctional alleles in Caucasians apart from the deletion allele *CYP2D6**5 (9,28). Each PCR product was detected by at least two Invader assays. Each assay used a synthetic oligo complementary to both

Figure 2. Schematic illustration of the Invader reaction, showing discrimination of single base-pair changes. (A) Enzyme-substrate recognition and sequence-specific discrimination components of the Invader assay for mutation/polymorphism detection. The creation of a cleavable substrate and production of fluorescence signal in the secondary reaction depend both on complementarity between the probe and the target sequence at the base to be discriminated and an overlap of at least one base by the Invader oligo. (B) FRET format of the Invader biplex assay. The cleaved 5' flaps from the primary reaction hybridize to their respective FRET cassettes to create an invasive substrate recognized and cleaved by the Cleavase enzyme. Cleavage separates the two ends on the cassette, producing a fluorescence signal detectable above background.

Table 1. PCR and Invader Locus-Specific Oligos

Oligonucleotide Locus	Reaction Type	Oligonucleotide Type	Sequence ^a
Triplex PCR 1	PCR	PCR sense PCR anti-sense	CTGGGCTGGGAGCAGCCTC CACTCGCTGGCCTGTTTCATGTC
Triplex PCR 2	PCR	PCR sense PCR anti-sense	CTGGAATCCGGTGTCGAAGTGG CTCGGCCCTGCACTGTTTC
Triplex PCR 3	PCR	PCR sense PCR anti-sense	GAGGCAAGAAGGAGTGTCAGGG AGTCCTGTGGTGAGGTGACGAGG
<i>CYP2D6</i> *2 <i>XN</i> (duplication)	PCR	PCR sense PCR anti-sense	GCCACCATGGTGCTTTGCTTTC ACCGGATTCCAGCTGGGAAATG
<i>CYP2D6</i> *5 (deletion)	PCR	PCR sense	ACCGGGCACCTGTACTCCTCA
	PCR	anti-sense	GCATGAGCTAAGGCACCCAGAC
<i>CYP2D6</i> Copy Number assay	INVADER	2D6 Primary Probe Actin Primary Probe 2D6 Invader oligo Actin Invader oligo	<u>ACGGACGCGGAGTT</u> ACAGCACAGGTGC <u>CGCGCCGAGGC</u> AGGTAGTCGGTGAGATC CCCGCGCCACCCACACTGAGCC AAGAGTAGCCACGCTCGGTGAGGATCTTCATT
<i>CYP2D6</i> *10-100 C>T	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	GCAGTGGCAGGGGGCTGGTGT <u>ATGACGTGGCAGAC</u> GGTAGCGTGACGC <u>CGCGCCGAGGAGT</u> AGCGTGACGCC
<i>CYP2D6</i> *6-1707 T>Deletion	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	CAGGCGGCCTCCTCGGTACCT <u>CGCGCCGAGGC</u> ACTGCTCCAGCGA <u>ATGACGTGGCAGAC</u> CTGCTCCAGCGA
<i>CYP2D6</i> *4-1846 G>A	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	CCTTACCCGCATCTCCACCCCAT <u>CGCGCCGAGGAG</u> ACGCCCTTTTCG <u>ATGACGTGGCAGAC</u> GGACGCCCTTTTCG
<i>CYP2D6</i> *3-2549 A>Deletion	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	GCTGGGCTGGGTCCCAGGTACATCT <u>CGCGCCGAGGCT</u> GTGCTCAGTTAGCAG <u>ATGACGTGGCAGAC</u> CGTGCTCAGTTAGCAG
<i>CYP2D6</i> *2-2850 C>T	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	GGCAGAGAACAGGTGAGCCACCATATGCT <u>ATGACGTGGCAGAC</u> GCAGTTCTCATCATTGAA <u>CGCGCCGAGGAC</u> AGGTTCTCATCATTGAAG
<i>CYP2D6</i> *2-4180 G>C	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	GCCACCATGGTGCTTTGCTTCTCGGTGAT <u>CGCGCCGAGGCC</u> CCATCCCCCTATG <u>ATGACGTGGCAGAC</u> GCCCATCCCCCTATG
<i>CYP2D6</i> *18-4125 GTGCCCACT>Dup.	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	CCGGGGCTGTCCAGTGGGCAT <u>CGCGCCGAGGC</u> AGTGGGCACCGA ^b <u>ATGACGTGGCAGAC</u> CCGAGAAGCTGAAGTG ^b
<i>CYP2D6</i> *11-883 G>C	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	CCCGAAGCGGCGCCGCAAT <u>CGCGCCGAGGCT</u> GCAGAGGGAGGG ^b <u>ATGACGTGGCAGAC</u> TGCAGAGGGAGGG ^b
<i>CYP2D6</i> *35-31 G>A	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	GGCTAGAAGCACTGRTGCCCTGGCCT <u>ATGACGTGGCAGAC</u> GTGATAGTGGCCATCTTC <u>CGCGCCGAGGAT</u> GATAGTGGCCATCTTC
<i>CYP2D6</i> *33-2483 G>T	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	AGCCTTTTGAAGCGTAGGACCTTGCCAGT <u>ATGACGTGGCAGAC</u> CCAGCGCTGGGATA <u>CGCGCCGAGGAC</u> AGCGCTGGGATAT
<i>CYP2D6</i> *37-1943 G>A	INVADER	Invader Oligo Primary Probe 1 Primary Probe 2	CTGAGCTAGGTCCAGCAGCCTGAGGAAGA <u>ATGACGTGGCAGAC</u> CGAGGGTCGTCTGTAC ^b <u>CGCGCCGAGGT</u> GAGGGTCGTCTGTAC ^b

^aUnderlined sequence represents the 5' flap of the probe; ^bamino-blocked 3' end.

the Invader oligo and probes as a positive control. The Invader reactions were performed using 384-well Invader assay FRET detection plates with assay components dried down in each well, consisting of: Cleavase enzyme, both F (FAM) dye and R (Redmond Red™) dye (Epoch Biosciences, Redmond, WA, USA)

FRET cassettes, and reaction buffer.

Three microliters of a 1/20 dilution of the *CYP2D6*-specific PCR products, a negative control ($T_{10}e_{0.1}$ buffer (10 mM Tris, pH 8, 0.1 mM EDTA)), or a positive control (allele-specific synthetic DNA target) were added to the appropriate wells, fol-

RESEARCH REPORT

lowed by addition of 3 μL of the appropriate probes/Invader oligo/MgCl₂ mix, and an overlay of 7 μL of molecular biology-grade mineral oil to prevent evaporation (Sigma-Aldrich, Steinheim, Germany). Each 6- μL reaction contained 10 ng Cleavase enzyme, 4% PEG 8000, 2% glycerol, 0.06 % Nonidet® P40 (NP 40), 0.06% Tween® 20, 12 $\mu\text{g}/\text{mL}$ bovine serum albumin (BSA), 0.58 μM each of F dye and R dye FRET cassettes, 8 mM MgCl₂, 0.7 μM of each allele-specific probe, and 0.07 μM Invader oligo. Following reagent dispensing, plates were spun for 10 seconds at 1000 rpm (120 $\times$ g), then sequentially incubated at 95°C for 5 minutes and 63°C for 30 minutes in a PCR Express thermal cycler (Thermo Hybaid). After the incubation period, fluorescence was measured directly on a CytoFluor® 4000 fluorescence plate reader (Applied Biosystems, Foster City, CA, USA). The settings used were: 485/20 nm excitation/bandwidth and 530/25 nm emission/bandwidth for F dye detection, and 560/20 nm excitation/bandwidth and 620/40 nm emission/bandwidth for R dye detection.

CYP2D6 Copy Number Assay

We used the quantitative capability of the Invader system to determine gene copy number by comparing the target gene signal (*CYP2D6*) with that of a reference gene, such as α -actin, known not to be polymorphic for either duplication or deletion. The *CYP2D6* probe set was designed to react specifically with only the *CYP2D6* sequence. The relative ratios of the *CYP2D6* and reference gene signals from each assay enabled us to identify and quantitate the deletion and duplication alleles of *CYP2D6*.

With the approval of the University of Wisconsin–Madison Institutional Review Board, the *CYP2D6* copy number was assayed in 205 individuals. Genomic DNA was isolated from whole blood using the PUREGENE™ DNA Isolation Kit (Gen-

tra Systems, Minneapolis, MN, USA). Invader detection of the *CYP2D6* copy number was performed in duplicate using 96-well dry-down plates. In brief, 7 μL of pre-denatured DNA samples (15–20 ng/ μL) or negative control (10 ng/ μL solution of tRNA in T₁₀e_{0.1} buffer) were added to the appropriate wells followed by addition of 8 μL of the appropriate probes/Invader oligo/MgCl₂ mix and an overlay of 15 μL of molecular biology-grade mineral oil. Each 15- μL reaction contained 100 ng Cleavase enzyme, 4% PEG 8000, 2% glycerol, 0.06 % NP 40, 0.06% Tween 20, 12 $\mu\text{g}/\text{mL}$ BSA, 0.35 μM each of F dye and R dye FRET cassettes, 8 mM MgCl₂, 0.7 μM of each allele-specific probe, and 0.07 μM Invader oligo. Following reagent dispensing, plates were spun for 10 seconds at 1000 rpm (120 $\times$ g), incubated at 63°C for 4 hours in a PTC-100™ thermal cycler (MJ Research, Waltham, MA, USA), and then directly read in a CytoFluor 4000 fluorescence plate reader using the settings given above.

Invader assignments were confirmed by long-range PCR, using the primers described in Table 1. The *CYP2D6* deletion (*CYP2D6**5) produces a 3.5-kb PCR product (36), as illustrated in Figure 3. Duplicated or multiple copy *CYP2D6* alleles (e.g., *CYP2D6**2XN) yield a 10-kb PCR product (17). Samples identified as containing either one or three-plus copies of the *CYP2D6* allele were subjected to PCR. Gene deletion and duplication PCR assays were performed with the GeneAmp® XL PCR kit (Applied Biosystems). The deletion primers (36) were used in a 50- μL PCR with 200 ng DNA, 1 $\times$ XL kit reaction buffer, 1.1 mM Mg(OAc)₂, 200 mM of each dNTP, 0.3 mM of each primer, and 1 U of XL kit DNA polymerase. The cycling parameters used were as follows: 94°C for 1 minute followed by 35 cycles of 94°C for 1 minute, 65°C for 30 seconds, and 68°C for 5 minutes, then a final 12-minute 72°C extension cycle. The resulting 3.5-kb PCR products were observed on a 1% agarose gel containing ethidium bromide.

Figure 3. Conventional confirmation of *CYP2D6**2XN (duplication) and *CYP2D6**5 (deletion) by long-range PCR. The shaded boxes represent regions of almost complete identity. (A) The 15-kb product from samples containing only one copy is too large to amplify. (B) When the *CYP2D6* allele is deleted (*CYP2D6**5), a 3.5-kb region can be amplified. (C) When duplicated or multiple copy *CYP2D6* alleles are present, a 10-kb region can be amplified.

RESEARCH REPORT

The duplication PCR primers amplified a fragment between exon 9 of the proximal *CYP2D6* copy and intron 2 of a distal *CYP2D6* copy in regions specific to *CYP2D6* (17) (Table 1). The 50-μL PCR contained 400 ng DNA, 1× XL reaction buffer, 1.0 mM Mg(OAc)₂, 200 mM of each dNTP, 0.3 mM of each primer, and 3 U of XL kit DNA polymerase. The cycling parameters used were as follows: 94°C for 1 minute followed by 35 cycles of 94°C for 1 minute, 61.4°C for 30 seconds, and 68°C for 10 minutes, then a final 12-minute 72°C extension cycle. The resulting 10-kb PCR products were detected on a 1% agarose gel containing ethidium bromide. We observed no amplification in alleles lacking the duplication. However, conventional PCR could not determine the exact number of *CYP2D6* duplications.

Data Analysis for Genotype and Copy Number Determination

Data were exported into Microsoft® Excel. For each allele of a given polymorphism, the net signal/background, or Net Fold Over Zero (FOZ - 1), values are calculated as follows:

$$\text{Net F dye FOZ} = \frac{\text{F dye raw counts from sample}}{\text{F dye raw counts from negative control}} - 1$$

$$\text{Net R dye FOZ} = \frac{\text{R dye raw counts from sample}}{\text{R dye raw counts from negative control}} - 1$$

Determination of the genotype or copy number was based on the ratio of the Net R dye FOZ to the Net F dye FOZ value as shown below:

$$\text{Allelic Ratio} = \frac{\text{Net R dye FOZ}}{\text{Net F dye FOZ}}$$

When the Net FOZ value was equal to or less than zero, the Net FOZ value was adjusted to 0.01 to avoid the generation of negative values or division by zero. An allelic ratio of less than or equal to 0.25 was scored as homozygous for the F dye allele, a ratio equal to or greater than 4.0 was scored as homozygous for the R dye allele, and a ratio equal to or greater than 0.4 but less than or equal to 2.5 was scored as heterozygous. Values that fell within two intermediate ranges were designated as equivocal and no genotype was assigned. Instances in which both the F dye and R dye Net FOZ values were less than 2 were recorded as low-signal. The allelic ratio was calculated separately for each sample duplicate. If the two results were discordant, the sample result was considered invalid.

For the copy number assay, the Net R dye and F dye FOZ values (for *CYP2D6* and *α-actin*) were calculated as above and the ratio of the *CYP2D6* to *α-actin* Net FOZ was calculated to identify *CYP2D6* copy number (R dye Net FOZ/F dye Net FOZ, as with the allelic ratio formula above). To accommodate differences in reactivity between *CYP2D6* and *α-actin* probe sets, we used the allelic ratio from a sample known to contain two copies of the *CYP2D6* gene as a reference. We first divided the allelic ratio of each sample by the allelic ratio of the reference sample, and then multiplied that value by 2. A normalized ratio of less than or equal to 0.35 was scored as a homozygous *CYP2D6**5 (*CYP2D6* deletion), a normalized ratio equal to or greater than 0.65 and less than or equal to 1.35 was scored as heterozygous single copy, a normalized ratio equal to or greater than 1.65 and less than or

equal to 2.35 was scored as two gene copies, and a normalized ratio equal to or greater than 2.65 was scored as three or more gene copies. Ratios that fell within the three intermediate ranges were designated as equivocal and no genotype was assigned. Samples yielding normalized Net FOZ values of less than 1.3 for both dyes were recorded as low-signal. The normalized allelic ratio was calculated separately for each sample duplicate. If the two results were discordant, the sample result was considered invalid. The values for each range were assigned based on cluster segregation.

RESULTS

Of the 181 genomic DNA samples used for the *CYP2D6* gene-specific amplification assays, 171 were visible by standard agarose gel electrophoresis with ethidium bromide staining. Ten DNA samples generated no visible PCR product; Invader assays could genotype three of these and they were included in the analysis. The remaining seven DNA samples were considered degraded and not used. All Invader reactions were performed in duplicate and each reaction was scored independently for genotype. A final genotyping score was recorded only if the duplicate results matched. From the possible 1914 results for the 11 loci and 174 DNA samples, 1904 unambiguous genotyping scores were recorded; only 10 genotyping scores could not be assigned. The 10 invalid assays were from the three DNA samples that did not generate visible triplex PCR products. Of these 10 assays, four were invalid because of low-signal in both duplicates, and six were invalid because signal ratios fell within the equivocal ranges. The Invader assay genotyping efficiency for all 174 DNA samples (including the three that yielded no visible triplex PCR products) was 99.5%. The reaction yielded unambiguous genotype determinations for 100% of the 171 DNA samples that produced visible triplex PCR products.

Figure 4 shows four graphs as examples of the Net FOZ values (FOZ - 1) from the 11 different assays; Table 2 shows the corresponding allele frequencies. All samples yielded heterozygous or homozygous variant signals with three exceptions: *CYP2D6**11-883, *18-4125, and *37-1943. These alleles have a previously reported frequency of 0.1% or less (28), so their absence in the 174 individuals analyzed here is not unusual. The allele frequencies found in this study are comparable to published allele frequency values for Caucasians (9,10,28,32). Although *CYP2D6**2-2850 assay results do not satisfy the Hardy-Weinberg equilibrium, graph 1 in Figure 4 shows very distinct cluster separation. In addition, all duplicates were in agreement and no unpredicted haplotypes were detected. Further studies on additional samples are needed to verify this finding.

We constructed individual *CYP2D6* haplotypes from information on the Web site (<http://www.imm.ki.se/CYPalleles/cyp2d6.htm>) and assigned compound haplotypes to individuals. We also independently calculated the frequencies of individual polymorphisms and haplotypes using the expectation maximization (EM) algorithm implemented in Arlequin software (<http://lgb.unige.ch/arlequin/>) (33,34). The EM algorithm identified nine distinct haplotypes (Table 3) that co-segregated into 22 different compound haplotypes (Table 4) within the 171 samples that yielded unambiguous genotypes with the Invader reaction. Ten individuals carried two nonfunctional *CYP2D6* alleles and 70

Table 2. Summary of Data Generated from Screening 174 DNA Samples with 11 *CYP2D6* Invader Genotyping Assays

Position	31	100	883	1707	1846	1943	2483	2549	2850	4125–4133	4180
Allele	*35	*10	*11	*6	*4	*37	*33	*3	*2	*18	*2
PCR product	PCR1	PCR1	PCR1	PCR2	PCR2	PCR2	PCR2	PCR2	PCR2	PCR3	PCR3
Polymorphism	G>A	C>T	G>C	T>Del	G>A	G>A	G>T	A>Del	C>T	duplication	G>C
Effect	V11M	P34S	splice defect	frameshift	splice defect	R201H	A237S	frameshift	R296C	468VPT470	S486T
n=	172	171	174	174	173	174	173	172	173	174	173
Common allele frequency	157	97	174	170	102	174	169	164	89	174	41
Heterozygous frequency	14	66	0	4	64	0	4	8	60	0	76
Rare allele frequency	1	8	0	0	7	0	0	0	24	0	56
Hardy Weinberg	$\chi^2_{1.2}$ ($P=0.28$)	$\chi^2_{0.6}$ ($P=0.44$)		$\chi^2_{0.0}$ ($P=0.88$)	$\chi^2_{0.6}$ ($P=0.4$)		$\chi^2_{0.0}$ ($P=0.88$)	$\chi^2_{0.1}$ ($P=0.75$)	$\chi^2_{26.4}$ ($P=0.011$)		$\chi^2_{2.3}$ ($P=0.13$)
Rare allele frequency	4.6	24	0	1.2	22.5	0	1.2	2.3	31.2	0	54.3
EM frequency ^a	4.6	23.8	0	1.2	22.7		1.2	2.3	31.4		54.7
Published frequency ^b (%)	6.7	18 (Asian 16)	0.1	1.8	18 (African 6.3)	0.1 (African 0.01)	0.6	1.7	34 (Asian 18)	Asian 0.7	53 (Asian 65)

^aAllele frequency as generated by the EM algorithm implemented in the Arlequin genetic software; ^ballele frequency described in References 8, 9, 10, 32.

individuals carried a single functional allele.

The copy number assay (205 samples) identified 17 single-copy individuals, 170 two-copy individuals, and 17 three-copy (or more) individuals. Only one assay yielded results that fell into an equivocal range (described above). As graph 5 in Figure 4 illustrates, the groupings of one, two, and three (or more) copies of the *CYP2D6* gene are distinctly separated. Sixteen of the 17 gene deletions and 11 of the 17 gene duplications detected by the Invader assay were confirmed by long-range PCR. Generating lengthy PCR products requires pure and intact genomic DNA, so any fragmentation of the DNA template will lead to PCR failure. Therefore, the absence of a long-PCR product cannot unambiguously assign the absence of the *CYP2D6* duplication or deletion.

DISCUSSION

This study demonstrates the Invader assay's ability to generate both qualitative and quantitative data and to surmount the complex genotyping challenges presented by *CYP2D6*. In addition, there is no need to observe or purify PCR products before performing the Invader reactions (29).

Figure 4. Invader assay fluorescence data. Graphs 1–4 represent four examples of the 11 *CYP2D6* genotyping assays. Graph 5 shows the distribution of normalized ratios in the copy number assay. The areas between the dashed and solid lines represent the intermediate equivocal ranges.

RESEARCH REPORT

Table 3. CYP2D6 Haplotype Predictions from 175 Genomic Samples Using the EM Algorithm Implemented in Arlequin Genetic Software

Haplotype	Frequency	Std. Dev.	Polymorphism site											
CYP2D6*1	0.401	0.027	G31	C100	G883	INS1707	G1846	G1943	G2483	INS2549	C2850	DEL4125	G4180	
CYP2D6*2	0.262	0.024	G31	C100	G883	INS1707	G1846	G1943	G2483	INS2549	T2850	DEL4125	C4180	
CYP2D6*4	0.221	0.024	G31	T100	G883	INS1707	A1846	G1943	G2483	INS2549	C2850	DEL4125	C4180	
CYP2D6*35	0.047	0.013	A31	C100	G883	INS1707	G1846	G1943	G2483	INS2549	T2850	DEL4125	C4180	
CYP2D6*3	0.029	0.012	G31	C100	G883	INS1707	G1846	G1943	G2483	DEL2549	C2850	DEL4125	G4180	
CYP2D6*6	0.012	0.007	G31	C100	G883	DEL1707	G1846	G1943	G2483	INS2549	C2850	DEL4125	G4180	
CYP2D6*33	0.012	0.007	G31	C100	G883	INS1707	G1846	G1943	T2483	INS2549	C2850	DEL4125	G4180	
CYP2D6*10	0.012	0.007	G31	T100	G883	INS1707	G1846	G1943	G2483	INS2549	C2850	DEL4125	C4180	
CYP2D6*4k	0.006	0.004	G31	T100	G883	INS1707	A1846	G1943	G2483	INS2549	T2850	DEL4125	C4180	

Table 4. Compound CYP2D6 Haplotypes for 171 DNA Samples Genotyped by the Invader System and Categorized by Number of Functional Alleles

Number of Functional Alleles	Compound Haplotype	Number of Subjects
2	*1/*1	31
	*1/*2	26
	*2/*2	15
	*2/*35	8
	*1/*35	3
	*1/*33	3
	*1/*10	2
	*2/*10	1
	*2/*33	1
1	*35/*35	1
	*1/*4	35
	*2/*4	22
	*1/*3	4
	*2/*3	2
	*4/*35	3
	*1/*6	2
	*4/*10	1
0	*2/*6	1
	*4/*4	7
	*3/*4	2
	*4/*6	1

In the allele frequency study, the results from the 11 Invader CYP2D6 assays provided an unambiguous genotyping determination for 100% of the 171 samples that yielded a visible PCR product on an agarose gel. All invalid genotype results arose from the same three samples that yielded no visible triplex PCR products. The overall unambiguous genotyping determination was 99.5%, but this lower success rate can be attributed to PCR amplification failure. Poor PCR amplification most likely arose from partial degradation of the genomic samples; more than 40% of the same 181 samples also failed to yield a 5-kb PCR product in

initial CYP2D6 long-range PCR amplification attempts (data not shown). Our triplex PCR approach generated a CYP2D6-specific template from all but the most degraded DNA samples, and may be more robust than protocols involving long-range PCR.

High failure rates inherent in long-range PCR-based methods make PCR-based methods to detect CYP2D6 copy number cumbersome (17). The quantitative screening strategy used to resolve CYP2D6 copy number complements the PCR/Invader genotyping assays and avoids problems associated with previous long-range PCR (17,23) or restriction fragment length polymorphism (RFLP) methods (8,16). Only one of the 205 samples yielded an equivocal result in the Invader assay. CYP2D6 copy numbers greater than three could not be resolved in this study. However, our data suggest that complementing the PCR/Invader genotyping assay with the genomic DNA copy number assay would be invaluable for identifying both ultra-rapid and poor metabolizers. Twenty-seven of the DNA samples expressed two nonfunctional alleles or possessed a deleted allele (Table 4) and 17 samples possessed three or more alleles. With poor metabolizers, caregivers must prescribe a lower dose of medication for the same therapeutic effect. On the other hand, with ultra-rapid metabolizers, caregivers must consider whether potentially toxic metabolites would accumulate or whether a therapeutic level of medication would be reached. For example, individuals homozygous for CYP2D6*35 (2,22) or with multiple gene copies have elevated enzyme activity and may need higher doses of medication to achieve therapeutic levels.

In summary, a refined set of CYP2D6 Invader polymorphism assays combined with the copy number assay offer a simple, rapid, and precise genotyping solution for the complex CYP2D6 locus to enable accurate predictions of CYP2D6 phenotypes. In addition, the Invader CYP2D6 panel described here can be expanded to accommodate either a more comprehensive set of allelic variants or newly discovered polymorphisms.

ACKNOWLEDGMENTS

We thank Peggy Eis, Kafryn W. Lieder, Stephane Perrey, and Eric Rasmussen for critical revisions and for insightful scientific discussions. This work was supported by Grant No. 1 R43 GM64317-01 from the National Institutes of Health to R.W. Kwiatkowski.

REFERENCES

- Agundez, J.A., M.C. Ledesma, J. Benitez, J.M. Ladero, A. Rodriguez-Lescure, E. Diaz-Rubio, and M. Diaz-Rubio. 1995. CYP2D6 genes and risk of liver cancer. *Lancet* 345:830-831.
- Allorge, D., J. Harlow, O. Boulet, G.P. Hayhurst, J. Chowdry, E. Roth, K. Crewe, J.M. Lo-Guidice, M. Lhermitte, F. Broly, G.T. Tucker, and S.W. Ellis. 2001. In-vitro analysis of the contribution of CYP2D6.35 to ultra-rapid metabolism. *Pharmacogenetics* 11:739-741.
- Berggren, W.T., T. Takova, M.C. Olson, P.S. Eis, R.W. Kwiatkowski, and L.M. Smith. 2002. Multiplexed gene expression analysis using the Invader RNA assay with MALDI-TOF mass spectrometry detection. *Anal. Chem.* 74:1745-1750.
- Cooksey, R.C., B.P. Holloway, M.C. Oldenburg, S. Listenbee, and C.W. Miller. 2000. Evaluation of the Invader assay, a linear signal amplification method, for identification of mutations associated with resistance to rifampin and isoniazid in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* 44:1296-1301.
- Daly, A.K., J. Brockmoller, F. Broly, M. Eichelbaum, W.E. Evans, E.J. Gonzalez, J.D. Huang, J.R. Idle, et al. 1996. Nomenclature for human CYP2D6 alleles. *Pharmacogenetics* 6:193-201.
- Daly, A.K., S. Cholerton, W. Gregory, and J.R. Idle. 1993. Metabolic polymorphisms. *Pharmacol. Ther.* 57:129-160.
- Eis, P.S., M.C. Olson, T. Takova, M.L. Curtis, S.M. Olson, T.I. Vener, H.S. Ip, K.L. Vedvik, et al. 2001. An invasive cleavage assay for direct quantitation of specific RNAs. *Nat. Biotechnol.* 19:673-676.
- Gaedigk, A., M. Blum, R. Gaedigk, M. Eichelbaum, and U.A. Meyer. 1991. Deletion of the entire cytochrome P450 CYP2D6 gene as a cause of impaired drug metabolism in poor metabolizers of the debrisoquine/sparteine polymorphism. *Am. J. Hum. Genet.* 48:943-950.
- Gaedigk, A., R.R. Gotschall, N.S. Forbes, S.D. Simon, G.L. Kearns, and J.S. Leeder. 1999. Optimization of cytochrome P4502D6 (CYP2D6) phenotype assignment using a genotyping algorithm based on allele frequency data. *Pharmacogenetics* 9:669-682.
- Griese, E.U., U.M. Zanger, U. Brudermanns, A. Gaedigk, G. Mikus, K. Morike, T. Stuvén, and M. Eichelbaum. 1998. Assessment of the predictive power of genotypes for the in-vivo catalytic function of CYP2D6 in a German population. *Pharmacogenetics* 8:15-26.
- Griffin, T.J., J.G. Hall, J.R. Prudent, and L.M. Smith. 1999. Direct genetic analysis by matrix-assisted laser desorption/ionization mass spectrometry. *Proc. Natl. Acad. Sci. USA* 96:6301-6306.
- Hall, J.G., P.S. Eis, S.M. Law, L.P. Reynaldo, J.R. Prudent, D.J. Marshall, H.T. Allawi, A.L. Mast, et al. 2000. Sensitive detection of DNA polymorphisms by the serial invasive signal amplification reaction. *Proc. Natl. Acad. Sci. USA* 97:8272-8277.
- Harrington, J.J. and M.R. Lieber. 1994. The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J.* 13:1235-1246.
- Hessner, M.J., K.D. Friedman, K.V. Voelkerding, S. Huber, D. Ryan, B. Nucie, and M. Ledford. 2001. Multisite study for genotyping of the factor II (prothrombin) G20210A mutation by the Invader assay. *Clin. Chem.* 47:2048-2050.
- Hsu, T.M., S.M. Law, S. Duan, B.P. Neri, and P.Y. Kwok. 2001. Genotyping single-nucleotide polymorphisms by the Invader assay with dual-color fluorescence polarization detection. *Clin. Chem.* 47:1373-1377.
- Johansson, I., E. Lundqvist, L. Bertilsson, M.L. Dahl, F. Sjoqvist, and M. Ingelman-Sundberg. 1993. Inherited amplification of an active gene in the cytochrome P450 CYP2D locus as a cause of ultrarapid metabolism of debrisoquine. *Proc. Natl. Acad. Sci. USA* 90:11825-11829.
- Johansson, I., E. Lundqvist, M.L. Dahl, and M. Ingelman-Sundberg. 1996. PCR-based genotyping for duplicated and deleted CYP2D6 genes. *Pharmacogenetics* 6:351-355.
- Kaiser, M.W., N. Lyamicheva, W. Ma, C. Miller, B. Neri, L. Fors, and V.I. Lyamichev. 1999. A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. *J. Biol. Chem.* 274:21387-21394.
- Kimura, S., M. Umeno, R.C. Skoda, U.A. Meyer, and F.J. Gonzalez. 1989. The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *Am. J. Hum. Genet.* 45:889-904.
- Kwiatkowski, R.W., V. Lyamichev, M. de Arruda, and B. Neri. 1999. Clinical, genetic, and pharmacogenetic applications of the Invader assay. *Mol. Diagn.* 4:353-364.
- Lieber, M.R. 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* 19:233-240.
- Lovlie, R., A.K. Daly, G.E. Matre, A. Molven, and V.M. Steen. 2001. Polymorphisms in CYP2D6 duplication-negative individuals with the ultrarapid metabolizer phenotype: a role for the CYP2D6*35 allele in ultrarapid metabolism? *Pharmacogenetics* 11:45-55.
- Lovlie, R., A.K. Daly, A. Molven, J.R. Idle, and V.M. Steen. 1996. Ultrarapid metabolizers of debrisoquine: characterization and PCR-based detection of alleles with duplication of the CYP2D6 gene. *FEBS Lett.* 392:30-34.
- Lundquist, R.C. and B.M. Olivera. 1982. Transient generation of a displaced single stranded DNA during nick translation. *Cell* 31:53-60.
- Lyamichev, V., M.A.D. Brow, and J.E. Dahlberg. 1993. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* 260:778-783.
- Lyamichev, V., A.L. Mast, J.G. Hall, J.R. Prudent, M.W. Kaiser, T. Takova, R.W. Kwiatkowski, T.J. Sander, et al. 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* 17:292-296.
- Lyamichev, V.I., M.W. Kaiser, N.E. Lyamicheva, A.V. Vologodskii, J.G. Hall, W.P. Ma, H.T. Allawi, and B.P. Neri. 2000. Experimental and theoretical analysis of the invasive signal amplification reaction. *Biochemistry* 39:9523-9532.
- Marez, D., M. Legrand, N. Sabbagh, J.M. Guidice, C. Spire, J.J. Lafitte, U.A. Meyer, and F. Broly. 1997. Polymorphism of the cytochrome P450 CYP2D6 gene in a European population: characterization of 48 mutations and 53 alleles, their frequencies and evolution. *Pharmacogenetics* 7:193-202.
- Mein, C.A., B.J. Barratt, M.G. Dunn, T. Siegmund, A.N. Smith, L. Esposito, S. Nutland, H.E. Stevens, et al. 2000. Evaluation of single nucleotide polymorphism typing with Invader on PCR amplicons and its automation. *Genome Res.* 10:330-343.
- Ohnishi, Y., T. Tanaka, K. Ozaki, R. Yamada, H. Suzuki, and Y. Nakamura. 2001. A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* 46:471-477.
- Reynaldo, L.P., A.V. Vologodskii, B.P. Neri, and V.I. Lyamichev. 2000. The kinetics of oligonucleotide replacements. *J. Mol. Biol.* 297:511-520.
- Sachse, C., J. Brockmoller, S. Bauer, and I. Roots. 1997. Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am. J. Hum. Genet.* 60:284-295.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: a software for population genetics data analysis. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Switzerland.
- Slatkin, M. and L. Excoffier. 1996. Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* 76(Pt 4):377-383.
- Smith, C.A., A.C. Gough, P.N. Leigh, B.A. Summers, A.E. Harding, D.M. Maraganore, S.G. Sturman, A.H. Schapira, et al. 1992. Debrisoquine hydroxylase gene polymorphism and susceptibility to Parkinson's disease. *Lancet* 339:1375-1377.
- Steen, V.M., O.A. Andreassen, A.K. Daly, T. Tefre, A.L. Borresen, J.R. Idle, and A.K. Gulbrandsen. 1995. Detection of the poor metabolizer-associated CYP2D6(D) gene deletion allele by long-PCR technology. *Pharmacogenetics* 5:215-223.
- Wilkins Stevens, P., J.G. Hall, V. Lyamichev, B.P. Neri, M. Lu, L. Wang, L.M. Smith, and D.M. Kelso. 2001. Analysis of single nucleotide polymorphisms with solid phase invasive cleavage reactions. *Nucleic Acids Res.* 29:E77.

Address correspondence to:

Dr. Monika de Arruda
Third Wave Technologies
502 S. Rosa Rd.
Madison, WI 53719-1256, USA
e-mail: mdearruda@tw.com