Acta Pharmacol Sin 2007 Feb; 28 (2): 279–286

Full-length article

Allelic distributions of *CYP2D6* gene copy number variation in the Eastern Han Chinese population[1]

Hai-hui SHENG[2,6], Ai-ping ZENG[3,6], Wen-xiang ZHU[2], Ren-fang ZHU[3], Hong-mei LI[2], Zhi-dong ZHU[2], Ying QIN[2], Wei JIN[4], Yan LIU[4], Yun-lan DU[5], Jian SUN[5], Hua-sheng XIAO[2,7]

[2]*National Engineering Center for Biochip at Shanghai, Shanghai 201203, China;* [3]*Department of Clinical Laboratory, Wenling No 1 People's Hospital, Wenzhou Medical College, Wenzhou 317500, China;* [4]*Department of Cardiology, Ruijin Hospital, Medical College, Shanghai Jiaotong University, Shanghai 200025;* [5]*Department of Neurology, Shanghai Gongli Hospital, Shanghai 200135, China*

Key words

CYP2D6; long PCR; deletion; duplication; copy number variation; rearrangement

[1] Project supported by grants from the Shanghai Science and Technology Committee Fund (No 05dz52022 and No 05dz19318).

[6] These authors contributed equally to this work.
[7] Correspondence to Dr Hua-sheng XIAO.
Phn 86-21-5132-0288.
Fax 86-21-5132-0266.
E-mail huasheng_xiao@shbiochip.com

Received 2006-06-21
Accepted 2006-08-29

doi: 10.1111/j.1745-7254.2007.00479.x

Abstract

Aim: The human cytochrome P450 2D6 (CYP2D6) gene copy number variation, involving *CYP2D6* gene deletion (*CYP2D6*5*) and duplication or multiduplication (*CYP2D6*×N*), can result in reduced or increased metabolism of many clinically used drugs. The identification of *CYP2D6*5* and *CYP2D6*×N* and the investigation of their allelic distributions in ethnic populations can be important in determining the right drug and dosage for each patient. **Methods:** The *CYP2D6*5* and *CYP2D6* genes, and *CYP2D6* gene duplication were identified by 2 modified long PCR, respectively. To determine duplicated alleles, a novel long PCR was developed to amplify the entire duplicated *CYP2D6* gene which was used as template for subsequent PCR amplification. A total of 363 unrelated Eastern Han Chinese individuals were analyzed for *CYP2D6* gene copy number variation. **Results:** The frequency of *CYP2D6*5* and *CYP2D6*×N* were 4.82% (*n*=35) and 0.69% (*n*=5) in the Eastern Han Chinese population, respectively. Of the 5 duplicated alleles, 3 were *CYP2D6*1×N* and 2 were *CYP2D6*10×N*. One individual was a carrier of both *CYP2D6*5* and *CYP2D6*1×N*. Taken together, the *CYP2D6* gene rearrangements were present in 10.74% of subjects. **Conclusion:** Allelic distributions of the *CYP2D6* gene copy number variation differ among Chinese from different regions, indicating ethnic variety in Chinese. Long PCR are convenient, cost effective, specific and semiquantitative for the detection of the *CYP2D6* gene copy number variation, and amplification of the entire duplicated *CYP2D6* gene is necessary for the accurate identification of duplicated alleles.

Introduction

Human cytochrome P450 (CYP) enzymes play important roles in the metabolism of a wide variety of exogenous and endogenous compounds. Approximately 57 *CYP* genes encoding cytochrome P450 proteins and 58 pseudogenes are present in the human genome and are classified into distinct families and subfamilies according to their sequence similarity[1]. The *CYP2D* subfamily comprises the *CYP2D6* gene and 2 pseudogenes (*CYP2D7* and *CYP2D8*), located in tandem on chromosome 22q13.1. The *CYP2D6* gene is the only functional gene present in the human *CYP2*D gene locus. Although *CYP2D6* is expressed at low levels in the liver, it plays crucial roles in the metabolism of over 65 commonly used drugs, including β-adrenergic blocking agents, antiarrhythmics, antipsychotics, antidepressants, and narcotic analgesics[2].

The *CYP2D6* gene is highly polymorphic and has the most variations among the *CYP450* gene superfamily, with more than 80 variations identified so far[3]. The variations can result in absent, decreased, normal, increased or qualitatively altered catalytic activity of CYP2D6, and consequently cause clinically relevant interindividual differences in therapeutic efficacy or adverse drug reactions. The *CYP2D6* genotype importantly determines the metabolism of approximately 12% of all clinically used drugs[4].

The functionally deficient alleles are caused by detrimental mutations that range from single base pair changes to partial or whole gene deletion. The incidence of poor metabolizers (PMs), the homozygous or compound homozygous carriers of 2 functionally deficient alleles, is approximately 3%–10% in Caucasians, but only 1%–2% in Orientals[5]. In Caucasians, common deficient alleles include *CYP2D6*3*, *4, *5, and *6, accounting for about 98% of PMs[6]. However, the most common allele in Han Chinese is *CYP2D6*10*, which is associated with reduced activity[7–10]. The occurrence of deficient alleles of *CYP2D6* is less frequent in Chinese, except for *CYP2D6*5*. *CYP2D6*5*, representing the deletion of the entire *CYP2D6* gene, occurs in 2%–7% of Caucasians[9] and 3%–7% of Chinese[7,8,10,11]. The frequency of *CYP2D6*5* is similar across populations, and therefore, *CYP2D6*5* is not a major cause of the difference of PM prevalences between populations[5,8]. In contrast to PMs, ultrarapid metabolizers (UMs) usually carry a duplicated, or even multiduplicated (up to 13 copies of *CYP2D6*), active *CYP2D6* allele (*CYP2D6*×N*). *CYP2D6*2×N* and *CYP2D6*41×N* are the most common alleles with CYP2D6 gene duplications[9]. UM may have therapeutic failure with drugs on account of increased enzymatic activity. The frequencies of *CYP2D6*×N* vary greatly between races[9], which is significantly different from the allelic distribution of *CYP2D6*5*. *CYP2D6*×N* is relatively rare in South-East Asians and Northern Europeans, but occurs at a frequency of 29% in an Ethiopian population owing to dietary selective pressure in the past that favored preservation of duplicated genes. The North-South gradient of the presence of the *CYP2D6*×N* in the European population is due to migrations of subjects from North-East Africa to the Mediterranean areas[9,12]. Both *CYP2D6*5* and *CYP2D6*×N* result from *CYP2D6* gene rearrangement[13] and comprise *CYP2D6* gene copy number variation.

The methods for determining the *CYP2D6* gene copy number variation can be divided into quantitative and qualitative analyses. Quantitative methods, including pyrosequencing, Invader™ and real-time PCR, are based on the assessment of the relative *CYP2D6* gene copy number by comparison between the amount of PCR product reflecting the number of *CYP2D6* genes and the co-amplified region from an unrelated constitutive single-copy gene[12,14–17]. However, quantitative methods sometimes result in errors when determining the *CYP2D6* gene copy number[12,14,17] and are unable to discriminate alleles with duplicated *CYP2D6* genes. Qualitative methods often use long PCR spanning the repeated regions (CYP-REP) flanking the *CYP2D6* gene to identify *CYP2D6*5* and *CYP2D6*×N*[18–21]. Long PCR is one of the most commonly used methods for the detection of *CYP2D6*5* and *CYP2D6*×N*, with its simplicity, convenience and cost effectiveness. Due to high polymorphism, it is important for the identification of the duplicated alleles to obtain the entire duplicated *CYP2D6* gene. Two methods, the restriction digestion of genomic DNA[22] and PCR amplification, can obtain fragments containing the entire duplicated gene. The former is very time-consuming and not suitable for large scale and high-throughput detection and clinical practice. Although PCR is highly efficient and convenient, a new method which can specifically amplify the entire duplicated *CYP2D6* gene is still needed.

The *CYP2D6* gene copy number variation obviously affects the metabolic rates of drugs which are substrates of CYP2D6[23], and its alleles vary in frequency among populations. Although *CYP2D6* gene rearrangements have been studied in the Central Han Chinese population, there are no data regarding the Eastern Han Chinese population. Therefore, the aims of this study were to develop assays for detecting *CYP2D6* gene copy number variation and to assess the prevalence of the *CYP2D6* gene copy number variation in the Eastern Han Chinese population.

Materials and methods

Subjects This study included 363 unrelated healthy individuals (51% women; mean±SD, 55±10.39 years) recruited through several hospitals in Shanghai and the Zhejiang province. All subjects were ethnically Eastern Han Chinese and were informed about the experimental procedure and the purpose of the study. Written consent was obtained from each participant. Genomic DNA was extracted from blood samples of the subjects using the Flexi Gene DNA Kit (Qiagen, Hilden, Germany) according to the manufacture's protocol.

Detection for *CYP2D6*5* To identify the *CYP2D6*5* allele, the assay was carried out by a duplex long PCR method. A forward primer DuplF, binding in the 5'UTR of the *CYP2D6* gene, was designed by using Primer Premier 5 (PREMIER Biosoft International, Palo Alto, CA, USA) and BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) (Table 1) and a reverse primer DPKlow is specific for the downstream of the *CYP2D6* gene[24]. Primers cyp-13 and cyp-24 are specific for the downstream of the *CYP2D7* and *CYP2D6* genes, respectively. Duplex long PCR was carried out in a total volume of 50 μL containing 0.32 μmol/L of primers cyp-13 and cyp-24, 0.4 μmol/L of primers DuplF and DPKlow, 0.4

Table 1. PCR primer sequences.

Primer	Sequence	Position[b]	Expected amplicon size (bp)
cyp-13	5'-ACCGGGCACCTGTACTCCTCA-3'	21925178-21925198	3501
cyp-24	5'-GCATGAGCTAAGGCACCCAGAC-3'	21909541-21909562	
DuplF	5'-CCCATTTGGTAGTGAGGCAGGT-3'	4179-4200	4776
DPKlow	5'-GCCGACTGAGCCCTGGGAGGTAGGTA-3'	21926453-21912578	
cyp-17	5'-TCCCCCACTGACCCAACTCT-3'	21926555-21926574/ 21912840-21912859	3616/5174
cyp-32	5'-CACGTGCAGGGCACCTAGAT-3'	21921401-21921420	
DuplR[a]	5'-CCACGTGCAGGGCACCTAGATT-3'	21921400-21921421	8181
2D6P1R[a]	5'-GGTGTCAGCAGAAGGGACTTT-3'	4651-4671	494
2D6P2F	5'-ACCCGGTTCAAACCTTTTGC-3'	4756-4775	605
2D6P2R	5'-GCCTGTTTCATGTCCACGAC-3'	5341-5360	
2D6P3F	5'-GGTTGGAGTGGGTGGTGGAT-3'	5736-5755	506
2D6P3R	5'-CCTGCAGAGACTCCTCGGTCT-3'	6221-6241	
2D6P4F	5'-ATTGAGACCCCGTTCTGTCT-3'	6615-6634	676
2D6P4R	5'-CTCCTATGTTGGAGGAGGTC-3'	7271-7290	
2D6P5F	5'-GGGTCCCAGCATCCTAGAGTC-3'	7945-7965	670
2D6P5R	5'-CTCAGCCTCAACGTACCCCT-3'	8595-8614	
2D6P6F	5'-CTGTAAGCCTGACCTCCTCCAA-3'	7261-7282	752
2D6P6R	5'-TGAGTGTCGTTCCCTGGGCAGGA-3'	7990-8012	

[a] Entire duplicated *CYP2D6* gene was amplified with primers DuplF and DuplR. Exon 1 of the *CYP2D6* gene was amplified with primer 2D6DuplF and 2D6P1R.
[b] GeneBank accession number AY545216 for the CYP2D6 gene and NT_011520.11 for downstream sequences of *CYP2D*6 and *CYP2D7* genes.

mmol/L of each deoxynucleoside triphosphate, 1× PCR reaction buffer, 2.85 mmol/L $MgCl_2$, 2.5 U LA *Taq* (TaKaRa, Otsu, Shiga, Japan) and 200 ng of genomic DNA. Cycling conditions were as follows: 94 °C for 3 min, followed by 35 cycles of 94 °C for 35 s, 66 °C for 1 min and 72 °C for 5 min, followed by 72 °C for 6 min. Five µL of amplification products were then separated on 0.8% agarose gel electrophoresis and identified.

Detection of the duplicated *CYP2D6* gene To identify *CYP2D6* gene duplications, the assay was carried out by a modified long PCR method, as described by Lovlie *et al*[20]. The PCR was performed in a total volume of 25 µL containing 0.36 µmol/L of each primer, 0.4 mmol/L of each deoxynucleoside triphosphate, 1× PCR reaction buffer, 2.65 mmol/L $MgCl_2$, 1.25 U LA *Taq* (TaKaRa, Otsu, Shiga, Japan) and 150 ng of genomic DNA. Cycling conditions were as follows: 94 °C for 3 min, followed by 35 cycles of 94 °C for 35 s, 64 °C for 1 min and 72 °C for 5.2 min, followed by 72 °C for 6 min. Five µL of PCR products were then separated on 0.8% agarose gel electrophoresis and identified.

Amplification of the entire duplicated *CYP2D6* gene and sequencing To amplify the entire duplicated *CYP2D6* gene, a long PCR was performed using specific primers DuplF and DuplR (Table 1). Long PCR reactions were carried out in a total volume of 50 µL containing 0.4 µmol/L of each primer, 0.4 mmol/L of each deoxynucleoside triphosphate, 1× PCR reaction buffer, 2.85 mmol/L $MgCl_2$, 3 U LA *Taq* (TaKaRa, Otsu, Shiga, Japan) and 240 ng of genomic DNA. Cycling parameters were 2.5 min at 94 °C, followed by 35 cycles at 94 °C for 35 s, 66 °C for 1 min and 72 °C for 8.5 min, and then a final extended step of 72 °C for 10 min. The PCR product was purified by QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) to avoid contamination of genomic DNA, and was then used as template for subsequent PCR amplification.

To amplify the coding region and part of the introns of the duplicated *CYP2D6* genes, primers were designed using primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www_slow.cgi); the amplicons were 494–752 bp. The sequences of the primers are listed in Table 1. Each fragment was amplified by touchdown[25]. Amplification reactions were carried out in a total volume of 30 μL containing 0.3 mmol/L of each deoxynucleoside triphosphate, 10 mmol/L Tris-HCl, 50 mmol/L KCl, 2 mmol/L $MgCl_2$, 20% Q solution (Qiagen, Hilden, Germany), 0.16 μmol/L of each primer, 10 ng purified long PCR product, and 1.2 U *Taq* (TaKaRa, Otsu, Shiga, Japan). Cycling conditions were as follows: 94 °C for 3 min, followed by 10 cycles of 94 °C for 30 s, 66 °C for 30 s with a 0.5 °C decrement of the annealing temperature per cycle and 72 °C for 30 –45 s, followed by 30 cycles of 94 °C for 30 s, 61 °C for 30 s and 72 °C for 30–45 s, followed by 72 °C for 10 min. Amplified products were purified by the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and sequenced with forward and reverse primer by ABI 3700 sequencer according to the Big-Dye chemistry reaction protocol (Applied Biosystems, Foster City, CA, USA). For definition of CYP2D6 alleles, see the CYP allele website[3].

Determination of genotype of carriers of *CYP2D6*×N* To determine the genotype of carriers of *CYP2D6*×N*, all exons and most parts of introns were amplified using genomic DNA as templates and sequenced. The primers used for amplification and sequencing and the reaction conditions were as described earlier.

Statistical analyses The frequencies were compared using the χ^2 test and Fisher's exact test. A value of $P<0.05$ was considered statistically significant. These analyses were carried out with SAS (SAS Institute Inc, Cary, NC, USA).

Results

With primer combination DuplF and DPKlow, a 4.8 kb fragment was amplified from the *CYP2D*6 gene locus, indicating the presence of the *CYP2D6* gene (Figure 1). Because individuals homozygous for *CYP2D6*5* are very rare, with a frequency of less than 0.5%[10,26], the 4.8 kb fragment also functions as a positive control for the amplification. In order to increase the efficiency of amplification, we used primer DuplF instead of DPKup, since the former does not form a hairpin structure and yields a shorter PCR product

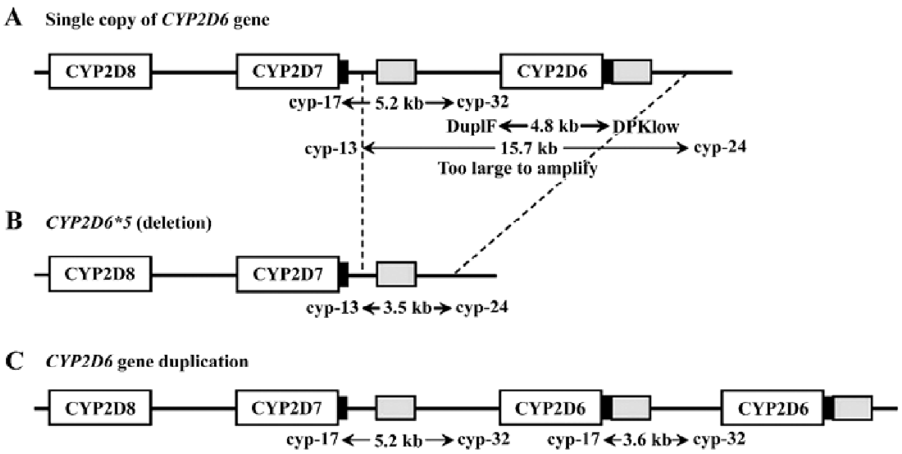

Figure 1. Schematic overview of the *CYP2D* genes and methods for the detection of *CYP2D6*5* and *CYP2D6* gene duplication. Large boxes represent the *CYP2D* genes. Small black and gray boxes represent 0.6 kb repeats and 2.8 kb CYP-REP units, respectively. The 2.8 kb CYP-REP units are almost completely identical the downstream of the *CYP2D7* and *CYP2D6* gene and duplicated *CYP2D6* gene. The 0.6 kb repeats and 2.8 kb CYP-REP units the downstream of the *CYP2D7* gene are separated by a 1.6 kb DNA fragment. The horizontal arrowheads show specific primer sets. (A) For individuals carrying single copy of *CYP2D6* gene, 5.2 kb fragment and 4.8 kb fragment are amplified by primer sets, cyp-17 and cyp-32, and DuplF and DPKlow, respectively. The 15.7 kb fragment is too large to amplify with primers cyp-13 and cyp-24. (B) When the *CYP2D6* gene is deleted (*CYP2D6*5*), a 3.5 kb fragment is amplified with primers cyp-13 and cyp-24. (C) When duplicated or multiduplicated *CYP2D6* genes are present, a 3.6 kb and a 5.2 kb fragment are co-amplified with primers cyp-17 and cyp-32. The 4.8 kb fragment for the presence of the *CYP2D6* gene may be potential mixture of PCR products of duplicated and single copy of the *CYP2D6* gene. An 8.2 kb fragment yielded by primers DuplF and DuplR only contains the upstream and central *CYP2D6* genes.

when combined with primer DPKlow[24]. A 3.5 kb fragment produced with primers cyp-13 and cyp-24 indicated the deletion of the *CYP2D6* gene[19]. The 4.8 kb fragments were observed in all 363 samples, indicating a reliable and effective long PCR amplification in every sample. No subject with homozygous *CYP2D6*5* was found. A long PCR analysis showed the existence of a 3.5 kb fragment in 35 of 363 subjects (Figure 2). *CYP2D6*5* was found with a frequency of 4.82% in the study population (Table 2).

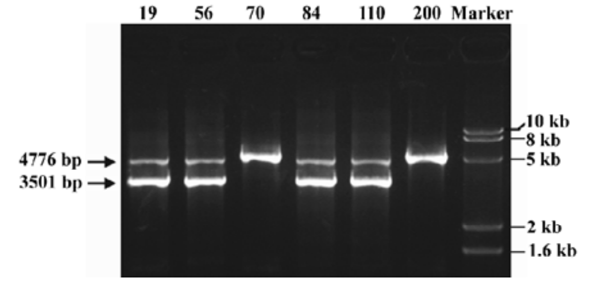

Figure 2. Long-PCR for the detection of *CYP2D6*5*. The 4.8 kb fragment for the presence of the *CYP2D6* gene and the 3.2 kb for *CYP2D6*5* were co-amplified in a single multiples reaction. Samples 19, 56, 84, and 110 are heterozygous for *CYP2D6**5, whereas samples 70 and 200 have no *CYP2D6* gene deletion.

Table 2. Comparison of *CYP2D6* gene rearrangement across populations.

	Eastern Han Chinese (*n*=363)	Central Han Chinese (*n*=223)	Malaysian Chinese (*n*=236)
Deletion	35 (4.82%)	32 (7.17%)	12 (2.54%)
Duplication	5 (0.69%)	6 (1.35%)	2 (4.24%)
[a]1×2	3 (0.41%)	5 (1.12%)	
[a]2×2	0 (0)		
[a]10×2	2 (0.28%)	1 (0.22%)	
Rearrangement	40 (5.51%)	38 (8.52%)	14 (2.97)

[a] Entire duplicated *CYP2D6* gene was amplified with primers DuplF and DuplR. Exon 1 of the *CYP2D6* gene was amplified with primer 2D6DuplF and 2D6P1R.
[#] GeneBank accession number AY545216 for the CYP2D6 gene and NT_011520.11 for downstream sequences of *CYP2D*6 and *CYP2D7* genes.

CYP2D6×N* was detected by long PCR with primers cyp-17 and cyp-32 (Figure 1). The 5.2 kb fragment, an internal control for the reaction, was observed in all 363 samples. The 3.6 kb fragment representing the duplicated *CYP2D6* gene was seen in 5 individuals (0.69%), whereas no amplification product was observed in the others. One of the 5 carriers of *CYP2D6*×N* is also a carrier of *CYP2D6*5*. Taken together, 39 individuals carried *CYP2D6* gene rearrangements with an incidence of 10.74%. The *CYP2D6* gene copy number variation showed no statistically significant difference between Eastern and Central Han Chinese (χ^2=4.215, *P*=0.112), or between Eastern Han Chinese and Malaysian Chinese (χ^2=4.323, *P*=0.115), but a statistically significant difference between Central Han Chinese and Malaysian Chinese (χ^2=13.252, *P*=0.0008; Table 2).

To characterize the duplicated *CYP2D6* gene, 2 specific primers DuplF and DuplR were used to amplify the entire duplicated *CYP2D6* gene (Figure 1). Forward primer DuplF is specific for the 5'UTR of the *CYP2D*6 gene and reverse primer DuplR can bind in *CYP2D6–CYP2D6* intergenic regions and the downstream of the *CYP2D7* gene. The primer DuplR was modified from cyp-32 in order to increase amplification specificity and yield. This primer combination can amplify an 8.2 kb fragment spanning the 5'UTR of the *CYP2D6* gene and *CYP2D6–CYP2D6* intergenic regions, which only allows for the amplification of the upstream and central *CYP2D6* genes, not the downstream or single-copy *CYP2D6* gene. Therefore, the 8.2 kb fragment contains the entire duplicated *CYP2D6* genes, not the single-copy *CYP2D*6 gene. The 8.2 kb fragment was observed in all 5 carriers of the duplicated *CYP2D6* gene, whereas no amplification product was seen in any other sample, as expected (Figure 3). Direct sequencing showed that the 8.2 kb fragment contained all 9 exons and indicated that the entire duplicated *CYP2D6* genes were specifically amplified by the long PCR. Among the 5 carriers of the duplicated *CYP2D6* gene, the most frequent duplicated allele was *CYP2D6*1* (60%), followed by *CYP2D6*10* (40%; Table 2). Since we were not able to exclusively sequence the downstream *CYP2D6* gene, the downstream alleles were indirectly determined by sequencing both genomic DNA and the entire duplicated *CYP2D6* gene. The downstream *CYP2D6* alleles in the 5 carriers were identical to the corresponding upstream and/or central *CYP2D6* alleles. The genotypes of 2 of the carriers of

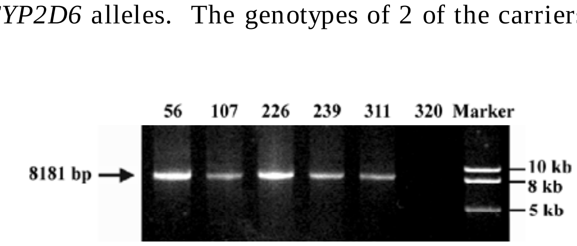

Figure 3. Results of the PCR amplified fragments containing the entire duplicated *CYP2D6* gene. The 8.2 kb fragment was obtained from samples 56, 107, 226, 239, and 311, whereas this fragment was not seen in sample 320 with a single copy of the *CYP2D6* gene.

CYP2D6×N* were *CYP2D6*1×N/*1*, while the others were *CYP2D6*1×N/*5*, *CYP2D6*10×N/*1* and *CYP2D6*10×N/*2*, respectively.

Discussion

Recently, several studies revealed that numerous copy number variations are in the human genome[27–31]. Many copy number variations contain entire genes and their numbers lead to differential levels of gene expression. Copy number variations account for many normal phenotypic variations and are involved in diseases[32], thus becoming a major focus of research in human genetics. In the present study, *CYP2D6*5* and *CYP2D6*×N* were detected to characterize the *CYP2D6* gene copy number variation in a large Eastern Han Chinese population. The frequency of *CYP2D6*5* is lower (4.82%) than in previous findings of 7.17% in the Central Han Chinese population[10] and 5.7% in Chinese living in Sweden[11], and very close to the 4.6% found in a Hong Kong Chinese population[7]. The frequency of *CYP2D6*×N* was also lower (0.66%) in the study population than that in the Central Han Chinese (1.35%)[10]. In Malaysian Chinese, the frequency of *CYP2D6*5* and *CYP2D6*×N* are both lower than those found on the Chinese mainland[8]. Our study showed that the *CYP2D6* gene copy number variation was the frequent variant in the Eastern Han Chinese population. Furthermore, the most common duplicated allele was *CYP2D6*1*, followed by *CYP2D6*10*, which was inconsistent with the high prevalence of *CYP2D6*10* in Chinese[7,8,10,11] and different from Central Han Chinese and other populations[9,10]. In the Chinese population, the frequency of *CYP2D6*10* ranges from 51% to 70%[7,8,10,11]. The structure, *CYP2D6*36+*10* tandem, was not observed in 5 carriers of *CYP2D6*×N*. However, our results need further verification in a larger population since only a few cases of *CYP2D6* gene duplication were included in our study. Differences of allelic distributions of the *CYP2D6* gene were observed between not just Eastern and Central Han Chinese, but also Taiwanese[33] and Hong Kong Chinese[7], implying that there is genetic diversity in Chinese from different regions.

The sequences of the *CYP2D7* and *CYP2D6* genes, as well as their downstream sequences, are almost completely identical, thus it is difficult to design PCR-based methods for detecting the *CYP2D6* gene copy number variation. Long PCR is the original method for the determination of *CYP2D6* gene rearrangement and has been widely used. In this study, we combined 2 long PCR for the detection of the *CYP2D6*5* and *CYP2D6* gene, and *CYP2D6*×N*, respectively, to investigate the *CYP2D6* gene copy number variation. This approach offers a unique ability to distinguish among 0, 1, 2 or more *CYP2D6* genes. It is a semiquantitative method for the detection of the *CYP2D6* gene copy number. Unusual *CYP2D6* gene rearrangement may confuse the determination of the *CYP2D6* gene copy number, but it occurs very rarely[26,34,35]. Long PCR are specific, costless and convenient for detecting *CYP2D6* gene rearrangements. Since TaqMan PCR, Invader™ and pyrosequencing for the detection of the *CYP2D6* gene copy number variation require pre-amplifying specific regions of the *CYP2D6* gene, it seems unlikely that these methods are less time-consuming than long PCR. Furthermore, polymorphic sites and gene conversions in target sequences may affect the accuracy of these methods[12,14,17]. The AmpliChip CYP450 test can detect simultaneously *CYP2D6* alleles and determine 7 duplicated alleles[36], but it remains relatively expensive.

The *CYP2D* genes are rich in specific DNA elements for recombination, and thus the *CYP2D* locus is a hot spot region for unequal crossover events[13]. The breakpoint is located downstream of both the *CYP2D7* and *CYP2D6* genes. According to the recombination pattern of the *CYP2*D locus, several methods for identification of the *CYP2D6* gene deletion and duplication, respectively, were developed by amplification of fragments spanning the potential crossover sites[19,20]. In order to identify alleles with duplicated *CYP2D6* genes, additional assays should be carried out. Routine methods are to amplify fragments spanning exon 9 in the upstream extra *CYP2D6* gene to intron 2 in the downstream *CYP2D6* gene used as template for restriction fragment length polymorphism (RFLP) assay[20,24] and allele-specific PCR[10]. However, gene duplications such as *CYP2D6*36×N* and the *CYP2D6*36+*10* tandem can not amplify due to the gene conversion in exon 9. The new method described by Gaedigk *et al* can amplify all duplication arrangements with forward and reverse primers binding to intron 6 and intron 2, respectively[37]. Because tag single nucleotide polymorphisms (SNPs) and/or mutations are distributed widely across the entire *CYP2D6* gene, duplicated alleles could be misclassified when analyzed on the basis of only part of the sequence of the duplicated *CYP2D6* gene. Amplification of the fragment containing the entire duplicated CYP2D6 gene is necessary for accurate identification of the duplicated alleles. Although the method described by Johansson *et al* can amplify a 5.1 kb fragment containing the entire *CYP2D6* gene by the DPKup/DPKlow primer pair[24], both single-copy and duplicated *CYP2D6* genes were amplified. The method described by Gaedigk *et al*[37] suffers from the same drawback. Therefore, duplicated alleles can not be exactly identified as a result of the mixture of PCR products. To accurately determine variant alleles with duplication of the *CYP2D6* gene, a novel long PCR was developed to specifically amplify the entire duplicated *CYP2D6* gene and confirmed by sequencing. The long PCR can always amplify upstream and central *CYP2D6* genes whenever sequential *CYP2D6* genes are duplicated comprising same allele or *CYP2D6*36* in tandem with *CYP2D6*10*. The 8.2 kb fragment, containing the entire duplicated CYP2D6 gene, can be used as template for the identification of alleles with duplication of the *CYP2D6* gene. The long PCR is less time-consuming than both the restriction digestion of genomic DNA, which takes at least 2 days[22], and the previously described PCR-based methods[20,24,37], since our method yields a shorter PCR product. In addition, Lovlie *et al* showed that the amplification of larger genomic DNA fragments, in contrast to shorter fragments, is more prone to failure[20]. Therefore, the long PCR is more maneuverable than the aforementioned methods[20,24,37]. The amplified duplicated *CYP2D6* gene is useful for further genotyping of duplicated genes to avoid misclassification of PM as UM due to the duplication of an inactive allele, or extensive metabolizer (EM) as UM due to the duplication of an allele associated with reduced activity. In addition, the long PCR can be readily adapted for other applications, such as the detection of CYP2D6 gene duplication.

In conclusion, we screened *CYP2D6* gene rearrangements by long PCR in the Eastern Han Chinese population and developed a long PCR to amplify the entire duplicated CYP2D6 gene. The allelic distributions of the CYP2D6 gene copy number variation vary among Chinese from different regions, indicating ethnic variety in Chinese.

Acknowledgement

We thank Dr Yi QU for reviewing this manuscript.

References

1 Nelson DR, Zeldin DC, Hoffman SM, Maltais LJ, Wain HM, Nebert DW. Comparison of cytochrome P450 (*CYP*) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. Pharmacogenetics 2004; 14: 1–18.
2 Wilkinson GR. Drug metabolism and variability among patients in drug response. N Engl J Med 2005; 352: 2211–21.
3 The Human Cytochrome P450 (CYP) Allele Nomenclature committee tHCPCAN. [Updated 2006 May 17, cited 2006 Jun 15] Available from: http://www.cypalleles.ki.se/cyp2d6.htm.
4 Ingelman-Sundberg M. The human genome project and novel aspects of cytochrome P450 research. Toxicol Appl Pharmacol 2005; 207: 52–6.
5 Bertilsson L, Dahl ML, Dalen P, Al-Shurbaji A. Molecular genetics of *CYP2D6*: clinical relevance with focus on psychotropic drugs. Br J Clin Pharmacol 2002; 53: 111–22.
6 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 1999; 9: 669–82.
7 Garcia-Barcelo M, Chow LY, Chiu HF, Wing YK, Lee DT, Lam KL, *et al*. Genetic analysis of the *CYP2D*6 locus in a Hong Kong Chinese population. Clin Chem 2000; 46: 18–23.
8 Ismail R, Teh LK, Amir J, Alwi Z, Lopez CG. Genetic polymorphism of *CYP2D6* in Chinese subjects in Malaysia. J Clin Pharm Ther 2003; 28: 279–84.
9 Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2004; 5: 6–13.
10 Ji L, Pan S, Marti-Jaun J, Hanseler E, Rentsch K, Hersberger M. Single-step assays to analyze *CYP2D6* gene polymorphisms in Asians: allele frequencies and a novel **14B* allele in mainland Chinese. Clin Chem 2002; 48: 983–8.
11 Johansson I, Oscarson M, Yue QY, Bertilsson L, Sjoqvist F, Ingelman-Sundberg M. Genetic analysis of the Chinese cytochrome P4502D locus: characterization of variant *CYP2D6* genes present in subjects with diminished capacity for debrisoquine hydroxylation. Mol Pharmacol 1994; 46: 452–9.
12 Soderback E, Zackrisson AL, Lindblom B, Alderborn A. Determination of *CYP2D6* gene copy number by pyrosequencing. Clin Chem 2005; 51: 522–31.
13 Lundqvist E, Johansson I, Ingelman-Sundberg M. Genetic mechanisms for duplication and multiduplication of the human *CYP2D6* gene and methods for detection of duplicated *CYP2D*6 genes. Gene 1999; 226: 327–38.
14 Nevilie M, Selzer R, Aizenstein B, Maguire M, Hogan K, Walton R, *et al*. Characterization of cytochrome P450 2D6 alleles using the Invader system. Biotechniques 2002; Suppl: 34–38, 40–43.
15 Muller B, Zopf K, Bachofer J, Steimer W. Optimized strategy for rapid cytochrome P450 2D6 genotyping by real-time long PCR. Clin Chem 2003; 49: 1624–31.
16 Bodin L, Beaune PH, Loriot MA. Determination of cytochrome P450 2D6 (CYP2D6) gene copy number by real-time quantitative PCR. J Biomed Biotechnol 2005; 2005: 248–53.
17 Schaeffeler E, Schwab M, Eichelbaum M, Zanger UM. *CYP2D6* genotyping strategy based on gene copy number determination by TaqMan real-time PCR. Hum Mutat 2003; 22: 476–85.
18 Hersberger M, Marti-Jaun J, Rentsch K, Hanseler E. Rapid detection of the *CYP2D6*3*, *CYP2D6*4*, and *CYP2D6*6* alleles by tetra-primer PCR and of the *CYP2D6*5* allele by multiplex long PCR. Clin Chem 2000; 46: 1072–7.
19 Steen VM, Andreassen OA, Daly AK, Tefre T, Borresen AL, Idle JR, *et al*. Detection of the poor metabolizer-associated *CYP2D6* (D) gene deletion allele by long-PCR technology. Pharmacogenetics 1995; 5: 215–23.
20 Lovlie 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 Lett 1996; 392: 30–4.
21 Dorado P, Caceres MC, Pozo-Guisado E, Wong ML, Licinio J, Llerena A. Development of a PCR-based strategy for *CYP2D6* genotyping including gene multiplication of worldwide potential

use. Biotechniques 2005; 39: 571–4.

22 Johansson I, Lundqvist E, Bertilsson L, Dahl ML, Sjoqvist F, Ingelman-Sundberg M. 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 1993; 90: 11 825–9.

23 Ma MK, Woo MH, McLeod HL. Genetic basis of drug metabolism. Am J Health Syst Pharm 2002; 59: 2061–9.

24 Johansson I, Lundqvist E, Dahl ML, Ingelman-Sundberg M. PCR-based genotyping for duplicated and deleted *CYP2D6* genes. Pharmacogenetics 1996; 6: 351–5.

25 Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 1991; 19: 4008.

26 Ledesma MC, Agundez JA. Identification of subtypes of *CYP2D* gene rearrangements among carriers of *CYP2D6* gene deletion and duplication. Clin Chem 2005; 51: 939–43.

27 McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, *et al*. Common deletion polymorphisms in the human genome. Nat Genet 2006; 38: 86–92.

28 Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA. Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet 2006; 38: 82–5.

29 Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, *et al*. Detection of large-scale variation in the human genome. Nat Genet 2004; 36: 949–51.

30 Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, *et al*. Large-scale copy number polymorphism in the human genome. Science 2004; 305: 525–8.

31 Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, *et al*. Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 2005; 77: 78–88.

32 Feuk L, Marshall CR, Wintle RF, Scherer SW. Structural variants: changing the landscape of chromosomes and design of disease studies. Hum Mol Genet 2006; 15: R57–66.

33 Wang SL, Lai MD, Huang JD. G169R mutation diminishes the metabolic activity of CYP2D6 in Chinese. Drug Metab Dispos 1999; 27: 385–8.

34 Ishiguro A, Kubota T, Sasaki H, Iga T. A long PCR assay to distinguish *CYP2D6*5* and a novel *CYP2D6* mutant allele associated with an 11-kb *Eco*RI haplotype. Clin Chim Acta 2004; 347: 217–21.

35 Fukuda T, Maune H, Ikenaga Y, Naohara M, Fukuda K, Azuma J. Novel structure of the *CYP2D6* gene that confuses genotyping for the *CYP2D6*5* allele. Drug Metab Pharmacokinet 2005; 20: 345–50.

36 Juran BD, Egan LJ, Lazaridis KN. The AmpliChip CYP450 Test: principles, challenges, and future clinical utility in digestive disease. Clin Gastroenterol Hepatol 2006; 4: 822–30.

37 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 2006; 34: 563–9.