Int J Clin Exp Med 2015;8(10):18917-18926
www.ijcem.com /ISSN:1940-5901/IJCEM0013929

Original Article

Analysis of genetic variations in *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* genes using oligonucleotide microarray

Yuanyuan Dong[1], Huasheng Xiao[2], Qi Wang[2], Chunxiu Zhang[2], Xiuming Liu[1], Na Yao[1], Haihui Sheng[2], Haiyan Li[1]

[1]*Ministry of Education Engineering Research Center of Bioreactor and Pharmaceutical Development, Jilin Agricultural University, Changchun, China;* [2]*National Engineering Center for Biochip at Shanghai, Shanghai, China*

Received August 4, 2015; Accepted October 4, 2015; Epub October 15, 2015; Published October 30, 2015

Abstract: The cytochrome P450 enzymes play a critical role in the metabolism of many commonly prescribed drugs. Among them, the most important enzymes are highly polymorphic *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5*, which are responsible for about 40% of the metabolism of clinical used drugs. Here we developed a novel CYP450 oligonucleotide microarray that allow for detection of 32 known variations of CYP genes from a single multiplex reaction, including 19 polymorphisms of *CYP2D6* gene, 8 polymorphisms of *CYP2C9* gene, 4 polymorphisms of *CYP2C19* gene and 1 polymorphism of *CYP3A5* gene. 229 genomic DNA samples from unrelated Han subjects were analyzed. The microarray results showed to have high call rate and accuracy according to concordance with genotypes identified by independent bidirectional sequencing. Furthermore, we found that the major *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* alleles in Chinese Han population were *CYP2C9**3 (allelic frequency of 10.7%), *CYP2C9**2 (20.31%), *CYP2C19**2 (5.68%), *CYP2D6**10 (58.52%), *CYP2D6**2 (13.76) and *CYP3A5**3 (70.69%). With flexible DNA preparation, the microarray can significantly facilitates the process of detecting genetics variations in *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* gene and provide safe and effective therapy for individual patients.

Keywords: Polymorphism, microarray, CYP2D6, CYP2C9, CYP2C19, CYP3A5

Introduction

Genetic factors, sequence variants in genes encoding drug-metabolizing enzymes, drug transporters and drug targets, influence the response of an individual to a given drug and account for 25-95% of variability in drug disposition and effects [1]. Genetic determinants of drug response, unlike others, remain stable throughout the lifespan. It is critical for correct prediction of drug response to genotype genetic variants contributing to interindividuals differences in drug response.

Human cytochrome P450 (CYP) enzymes, a superfamily of heme-thiolate proteins, are the most important enzymes that catalyze phase I drug metabolism, involved in the metabolism of most commonly used drugs [2]. At least 57 genes encoding cytochrome P450 proteins and 58 pseudogenes have been identified so far [3]. Among those, *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* genes are highly polymorphic and accounts together for about 40% of the metabolism of clinical used drugs [4-10]. Many variations in CYP genes cause changes in the function or level of expression of their correspondingly encoding enzymes, resulting in severe adverse effects or resistance to therapy. According to catalytic activity of CYP enzymes coded by CYP genes of carrier, phenotype can be classified into categories: poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM) or ultrarapid metabolizer (UM) phenotype, which only exists in carriers of duplicated or multiduplicated active *CYP2D6* genes. Therefore, medication dosages should be tailored to these genotypic-phenotypic classes to increase drug efficacy and reduce adverse effects, especially for those drugs with a narrow therapeutic index [11, 12].

Table 1. Primers used in the PCR reaction and detected CYP450 alleles

Primer name	Sequence (5'-3')	Detected polymorphism	Product size (bp)
2C19E4F	GAAAACATCAGGATTGTAAGCACCC	636G>A, 681G>A	1679
2C19E5R	TTAAGTTCCTCCTGTGCTGATCTCA		
2C19P1F	TCAAGCCCTTAGCACCAAATTCTCT	-806C>T, 1A>G	1291
2C19E1R	GCAAGCCACTGAAGGAGCATACT		
3A5E4F	GTTTGGTGAGAGCAGTGGATGAGGT	6986A>G	510
3A5E4R	CAAGAGTCTCACACAGGAGCCACC		
2C9E3F	TAGAAGCCTGTGTGGCTGAATAAAAG	269T>C, 374G>T, 389C>T, 449G>T	636
2C9E3R	TAAGGTCAGTGATATGGAGTAGGG	430C>T	
2C9E7F	CCATCCAGGTCAGTAACAGGTCAGT	1003C>T, 1075A>C, 1080C>G	914
2C9E7R	CTGCCAGAAATTCCAGCCCAAGGTT		
2D6E1F	GCCCATTTGGTAGTGAGGCAGGT	100C>T, 124G>A, 138 insT	1090
2D6E1R	CACCGCTGCTTGCCTTGGGAA		
2D6E3F	GGGTTGGAGTGGGTGGTGGAT	1661G>C, 1707delT, 1758G>A	508
2D6E3R	CCCTGAAGAGACTCCTCGGTCT		
2D6E4F	GCTGTCCCCGTCCTCCTGCAT	2539 del ACCT, 2573 ins C, 2850C>T	581
2D6E4R	CGGGTGTCCCAGCAAAGTTCAT	2935A>C, 2950G>C, 2988G>A	
2D6E6F	GGGTCCCAGCATCCTAGAGTC	4180G>C	673
2D6E6R	CTGCTCAGCCTCAACGTACCCCT		
2D6P1F	GGCAGCTGCCATACAATCCACCT	-1584C>G	616
2D6P1R	GGCACCACCCAGCCTAATTCTTT		
2D6E2N1F	GACCCGGTTCAAACCTTTTGC	883G>C, 974C>A, 984A>G, 1023C>A	607
2D6E2N1R	GGCCTGTTTCATGTCCACGAC	1039C>T	

The *CYP2D6* gene is more polymorphic than other CYP genes and presently more than 90 alleles have been identified (http://www.cypalleles.ki.se/cyp2d6.htm). The *CYP2D6* allele frequency varies in different population in the world. In Chinese, *CYP2D6*10* causes decreased (but not deficient) enzyme activity and is the most common allele with an allele frequency of more than 50% [13]. *CYP2D6*17* causes greatly decreased (but not deficient) enzyme activity and is frequent in Black African [14]. Among Caucasian, major variant *CYP2D6* alleles are *CYP2D6*4*, *CYP2D6*10* and *CYP2D6*5*, accounting for 17-27% of alleles [15, 16]. *CYP2C9* and *CYP2C19* gene, both located on chromosome 10, belong to the human *CYP2C* subfamily. Two common alleles of *CYP2C9*, *CYP2C9*2* and *CYP2C9*3*, are associated with markedly reductions in (S)-warfarin 79-hydroxylation and a greater risk of bleeding [17]. *CYP2C9*2* and *CYP2C9*3* occur in 0.1-2% of Caucasians, but *CYP2C9*2* is absent in East Asian [5]. Several polymorphisms of the *CYP2C19* gene have been identified and produce an inactive enzyme, and two inactive genetic variants (*CYP2C19*2* and *CYP2C19*3*) account for more than 95% of cases of poor metabolism of the relevant medications. *CYP3A5* is polymorphically expressed in human liver and small intestinal mucosa. The variable expression of the *CYP3A5* gene is strong correlative with a single nucleotide polymorphism (SNP) in intron 3 of the *CYP3A5* gene (designated *CYP3A5*3*) [18]. The frequency of *CYP3A5*3* is 88% in Caucasians, 75% in Asian and 35% in Africans [19].

Detection of genetic variations in CYP genes can be used to predict the phenotype and thereby suffice to optimize drug treatment. Various approaches to investigate SNPs have been applied to CYP genes genotyping, including allele-specific PCR (ASPCR) [13, 20], ASPCR-MP [21], TaqMan [22], Invader assay [23], SnaPshot [24] and oligonucleotide microarray [25-27]. Among them, only microarray technology offers the ability to analyze simultaneously a large number of polymorphisms. In the present study, we have developed a novel microarray for detecting 35 known variations in *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* genes from a one-tube reaction. The microarray can

discriminate 30 variant alleles of these four CYP genes, covering most common alleles of four CYP genes in Orientals, Caucasians and African. Results showed that the microarray has high call rate and accuracy. With simple DNA preparation method and high information contention, the microarray significantly facilitates the process of detecting variations in CYP genes and can be used in clinical practice and research investigations.

Materials and methods

Subjects

A total of 229 unrelated healthy individuals participated in the study. They were informed about the experimental procedure and the purpose of the study, and written consent was obtained from each participant. All subjects in this study are ethnically Eastern Chinese Han. Venous blood (2 ml) was obtained from each subject, and genomic DNA was extracted from whole blood using the Flexi Gene DNA Kit (Qiagen, Germany) according to the manufacture's protocol.

Variations and primers

A panel of 32 variants in CYP2D6, CYP2C9, CYP2C19 and CYP3A5 genes were identified from the literature and the CYP allele web site (http://www.cypalleles.ki.se/). All PCR primers were designed using primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www_slow.cgi/primer3.cgi) and were selected with melting temperature of >60°C. All primers sequencing were checked to avoid complimentary sequences (more than three bases) between 3'-end of any primers. BLAST (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html) and SNP BLAST (http://www.ncbi.nlm.nih.gov/SNP/snp_blastByOrg.cgi) were carried out to ensure specificity of primer sequence and avoid polymorphic site in the primer sequences, respectively. The sequences of the primers were listed in **Table 1**.

Multiplex PCR

DNA fragments spanning 32 variations were coamplified in a single multiplex reaction. Amplification reactions were carried out in a total volume of 30 μl containing 0.3 μM of each deoxynucleoside triphosphate, 40 mM Tricine-KOH (pH 8.7), 16 mM KCl, 3.5 mM $MgCl_2$, 3.75 μg/ml BSA, 2 μM of each primer, 100 ng of DNA and 1.2× Titanium polymerase (Clontech Laboratories Inc., USA). Cycling conditions were as follows: 95°C for 3 min, followed by 40 cycles of 95°C for 30 s, 68°C for 4.5 min, followed by 68°C for 10 min.

Fragmentation and labeling

PCR products were purified with QIAquick PCR Purification Kit (Qiagen, Germany). Purified PCR products were fragmented for 18 min at 37°C using DNase I (at 0.001 U/μg of DNA) in 20 μl volume containing 1 mM Tris-HCl, 0.25 mM $MgCl_2$, 0.05 mM $CaCl_2$ and 2.5 U of alkaline phosphatase (Fermentas, Lithuania), followed by heat inactivated at 95°C for 6 min. The 3' ends of the fragmented amplicons were fluorescenced by adding 0.3 μl 1 mM Cy3-dCTP (GE Healthcare, USA) using terminal deoxynucleotidyl transferase (Fermentas, Lithuania) in a 25 μL volume containing 200 mM potassium cacodylate, 25 mM Tris-HCl (pH 7.2), 1 mM $CoCl_2$, 0.01% Triton X-100, for 1 h at 37°C, followed by heat inactivation at 95°C for 5 min.

CYP oligonucleotide microarray

For each variant site three probes were designed, two of which are allele-specific probes, perfectly complementary to the target allele (PM), and one with a single base mismatch at central position (MM), which acts as a control for nonspecific hybridization. Two control probes were also designed to verify the performance of the CYP microarray. All 64 of the amino-modified oligonucleotides with 17 T residues spacers were dissolved at 25 μM concentration in 6× SSC buffer (pH 7.0) with 0.05% SDS and 25% DMSO, and printed in triplicate on the aldehyde-coated glass slides (Cell Associates Inc., TX, USA) as 3 subarrays using a OmniGrid™ 100 microarrayer (GeneMachine, San Carlos, CA, USA). A total of 64 oligonucleotides were printed on each glass slide in a 4 mm×8 mm. Spot spacing was 320 μm, and spot size was 120 μm. After printing, the CYP microarray was dried at room temperature for 2 hour, and then stored desiccated at room temperature.

For hybridization, 15 μl fragmented and fluorescenced PCR products were dissolved in 15 μl 2× hybridization buffer (10× SSPE, 10×

Handart's solution, 5% DMSO, 0.2% Triton X-100 and 2 nM fluorescenced control oligonucleotide). The hybridization solutions were heated at 95°C for 5 min, immediately placed on ice and 20 μl were transferred to the subarrays. Hybridization was carried out at 48°C for 2 hour. After hybridization, microarrays were washed with 2× SSC and 0.1% SDS at 42°C for 5 min, and subsequently washed with 1× SSC and 0.1% SDS at 42°C for 5 min and with 0.5× SSC at room temperature for 5 min.

Data analysis

Slides were scanned on the GenePix 4000B (Axon Instruments, Foster City, CA, USA) scanner. The fluorescence signal intensity values were determined using the GenePix Pro (Axon Instruments, Foster City, CA, USA). The signal intensities of the oligonucleotides that were spotted in triplicate were averaged. The mean of background signal plus 2 SD was used as the cutoff value to block weak or unreliable signals before classification. Typically, if the signal intensity of a variation divided cutoff value from a particular sample is ≤0, the hybridization signal is considered weak. Average values of allele probe were regard as the real signal values after subtraction of mean background signal and MM probe signal. Allelic fractions (AF) for each variant was calculated from the normalized mean by using follow formula: Allele B/(Allele A + Allele B) (http://www.bioinfo.helsinki.fi/SNPSnapper/). Theoretically, homozygous allele 1, heterozygous and homozygous allele 2 genotypes will give AF=0.0, 0.5 and 1.0, respectively. Experimentally, the AF is affected by various factors such as nonspecific hybridization and variation in labeling. Therefore, we set AF limit values as AF<0.25 for allele A homozygous genotypes, 0.4<AF<0.6 for heterozygous genotypes, and AF>0.75 for allele B homozygous genotypes. With regard to AF values between 0.25 and 0.4 or 0.6 and 0.75, variation isn't genotyped due to lower specificity of the hybridization signal. The AF limit values were confirmed by comparing the microarray results with those obtained by independent bidirectional sequencing.

Data analysis for determination of *CYP2D6* gene copy number was similar to previous report and allelic ratio (AR) of the signal intensity for *CYP2D6* gene to the signal intensity for *CYP2C9* gene was calculated (Allelic Ratio = Signal intensity for *CYP2D6* gene/Signal intensity for *CYP2C9* gene) [23]. Three samples known to contain one, two and three (or more) copies of *CYP2D6* gene, respectively, were used to ascertain AR limit values. AR<0.35, 0.65<AR<1.35, 1.65<AR<2.35 and AR>2.65 were scored as a homozygous *CYP2D6* gene deletion, heterozygous single copy, two gene copies and three or more gene copies, respectively. ARs between 0.35 and 0.65, 1.35 and 1.65 or 2.35 and 2.65 were designated as equivocal and no genotype was assigned.

Results

Primer design and multiplex amplification

It is necessary for improving multiplexing level to minimize and avoid primer-primer interactions or mitigating such interactions by making additional experimental efforts. Primers, with four or more consecutive complementary bases between the 3'-end of any primers, were excluded and redesigned to avoid artefacts. Because *CYP2C*, *CYP2D* and *CYP3A* subfamily have high sequence homology [28-30], primer sequences were submitted to the NCBI Web sites for BLAST (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html) searches to ensure specificity of primer sequence, only one hit in the human genome. To eliminate poor amplification efficiency due to polymorphic site in primer sequences, SNP BLAST (http://www.ncbi.nlm.nih.gov/SNP/snp_blastByOrg.cgi) searches were also performed. To boost amplification sensitivity and specificity, primers were designed with an annealing temperature of 64-70°C to allow fast cycling between 95°C and 68°C, except for primer pair CYP2C19P. Singleplex PCR was run to validate amplification efficiency of each primer pair. If no band or very weak band was observed, the primers were redesigned. Moreover, the singleplex PCR products were purified and then sequenced both directions to confirm correct PCR fragments amplified. To simplify and stabilize amplification of all targets in the same PCR reaction, multiplex PCR was optimized to obtain an acceptable yield and a similar amount of each fragment (**Figure 1A**, **1B**).

CYP450 oligonucleotide microarray

The microarray is composed of 106 oligonucleotide probes to detect a total of 35 polymorphisms of CYP genes, including 22 polymorphisms and mutations of the *CYP2D6* gene, 6 polymorphisms of the *CYP2C9* gene, 6 polymor-

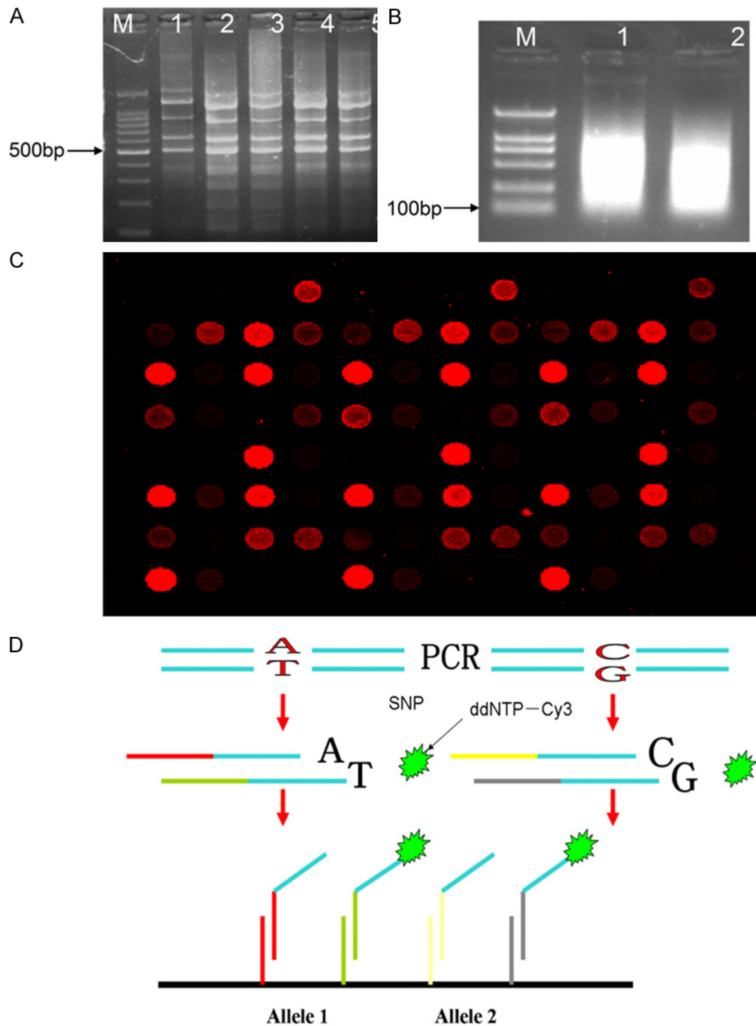

Figure 1. Overview of the *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* genes. A. Multiplex PCR amplified with *CYP450* primers. B. Fragment samples of purified PCR products. C. *CYP450* microarray results overview. D. Schematic overview of *CYP450* microarray.

phisms of the *CYP2C19* gene and 1 polymorphism of the *CYP3A5* gene. The microarray can discriminate 25 alleles, including 10 *CYP2D6* alleles (*CYP2D6*2*, **3*, **4*, **5*, **6*, **7*, **8*, **9*, **10*), 6 *CYP2C9* alleles (*CYP2C9*2*, **3*, **5*, **6* and **8*), 7 *CYP2C19* alleles (*CYP2C19*2*, **3*, **5*, **7*, **10* and **12*) and 2 *CYP3A5* alleles (*CYP3A5*3*). Each variant probe set was arranged in pairs, two of which are PM probes to discriminate between the two alleles, and one of which is MM probe to act as internal controls for PM probes to mitigate nonspecific hybridization. Every oligonucleotide was printed one time in three blocks, respectively (**Figure 1C**).

CYP450 microarray performance

To block cross-hybridization and weak signal, we designed MM probe for every variation and set cutoff value. After normalization of signal intensities of PM probes, the genotype for a given polymorphic locus (AA, AB or BB) is determined by the ratio between the signal intensity contributed from the allele B and signals from both alleles using AF limit values as cutoffs, which divided variations into three groups, one heterozygous (0.4<AF<0.6) and two homozygous (AF<0.25 or AF>0.75). 29 samples were carried out bidirectional sequencing and thereby were used as control for validation of AF limit values. The AF limit values based on concordance between microarray results and genotypes obtained by sequencing.

Genomic DNA samples extracted from 229 subjects were further analyzed for CYP genes genotype with the CYP microarray to demonstrate microarray quality. All DNA samples were genotyped in triplicate, yielding a total of 687 assays. We successfully called 21984 genotypes based on AF and visual inspection. Only four variations due to signal intensities under the cutoff value failed to call, which were caused by relatively low amount of PCR product. These four genotypes were successfully called through PCR re-amplification of the DNA samples. To further validate our microarray results, in the case of a heterozygote at the polymorphic site, the genotype was also determined by bidirectional sequencing. All results were found to be consistent with the microarray results. These results showed that the microarray have a high call rate, accuracy and reproducibility.

Table 2. Frequencies of the CYP2C9, CYP2C19, CYP2D6 and CYP3A5 alleles identified

	Alleles	Number of alleles	Frequency
CYP2C9	*CYP2C9*1*	316	69.00%
	*CYP2C9*2*	93	20.31%
	*CYP2C9*3*	49	10.70%
	*CYP2C9*5*	0	0
	*CYP2C9*6*	0	0
	*CYP2C9*8*	0	0
CYP2C19	*CYP2C19*1*	345	75.33%
	*CYP2C19*2*	26	5.68%
	*CYP2C19*3*	50	11%
	*CYP2C19*5*	37	8%
	*CYP2C19*7*	0	0
	*CYP2C19*10*	0	0
	*CYP2C19*12*	0	0
CYP2D6	*CYP2D6*1*	90	19.65%
	*CYP2D6*2*	63	13.76%
	*CYP2D6*3*	0	0
	*CYP2D6*4*	0	0
	*CYP2D6*5*	23	5.02%
	*CYP2D6*6*	0	0
	*CYP2D6*7*	0	0
	*CYP2D6*8*	0	0
	*CYP2D6*9*	14	3.06%
	*CYP2D6*10*	268	58.52%
CYP3A5	*CYP3A5*1*	134	29.26%
	*CYP3A5*3*	324	70.74%
	Total	458	100

CYP2D6, CYP2C9 and CYP2C19 and CYP3A5 allele frequencies in Easter Han Chinese population

CYP2D6, *CYP2C9*, *CYP2C19* and *CYP3A5* alleles and genotype frequencies in this study were summarized in **Table 2**. The frequency of *CYP2C9*3* was 10.70%, accounted for 1.7% (1/58) of PM in this population. No other deficient alleles had identified in the study population. The frequency of *CYP2C19*2* and **3* was 5.68% and 10.92%, respectively, accounting together for 28% of PM in this population. Among the *CYP2D6* variant alleles analyzed, *CYP2D6*10* was the most common allele (58.52%), followed by *CYP2D6*1* (19.65%), *CYP2D6*2* (13.76%) and *CYP2D6*5* (5.02%). Other deficient CYP2D6 alleles had not observed in the study population. The frequency of *CYP3A5*3* was 70.74%, accounting for 60% of PM in the population.

Discussion

CYP2C9, *CYP2C19*, *CYP2D6* and *CYP3A5* are highly polymorphic and contribute to a significant number of adverse drug reactions (ADRs), which cost US society more than 19 billion [8]. In addition, incidence of ADRs in patients who possess polymorphic forms of P450s is higher than those who possess non-polymorphic alleles [8]. Prediction of phenotype by detecting polymorphisms of CYP genes that are important for drug metabolism is instrumental in drug therapy and drug development. Considering many a variation that exists in CYP genes, routine genotyping methods, such as restriction fragment length polymorphism analysis, ASPCR and TaqMan, are not suitable due to labouriousness, time-consumingness and/or expensiveness. In contrast, high throughput, information-rich microarray can detect many variations in multiple genes at a time. To obtain same information, microarray platforms require less work and labor than routine traditional genotyping approaches. Application of microarray in clinical practice is necessary to advance significantly personalized medicine.

In this report, we described a rapid and high-throughput microarray-based genotyping method for detecting variations of CYP genes. The microarray offers the unique ability to discriminate simultaneously a total of 32 of the most relevant variations of CYP genes, including 19 polymorphisms and mutations of the *CYP2D6* gene, 8 polymorphisms of the *CYP2C9* gene, 4 polymorphisms of the *CYP2C19* gene and 1 polymorphism of the *CYP3A5* gene, which can analyze 19 *CYP2D6* alleles, 8 *CYP2C9* alleles, 4 *CYP2C19* alleles and 1 *CYP3A5* alleles. One variation, *CYP2C9* 1465 C>T, was excluded on account of cross-hybridization, which can not be reduced or eliminated by enhancement of printing buffer and hybridization condition. Basically, common alleles of these genes in Asians, Caucasians and African can be detected using the microarray. One of advantages of the method is that all DNA targets were coamplified in a single multiplex reaction, no singleplex PCR amplification required before genotype determination. To assess CYP microarray quality, we measure concordance of genotypes determined by independent direct sequencing.

CYP microarray results showed a high call rate and accuracy as generated by sequencing. In addition, a total of 229 subjects was analyzed using the CYP microarray and the allele frequencies of CYP genes in this Chinese Han population were similar to those found in previous studies [9, 13, 25, 31, 32].

There are two commercial chips for genotyping CYP in the market, CodeLink P450 Bioarray (http://www4.amershambiosciences.com/aptrix/upp00919.nsf/(FileDownload)?OpenAgent&docid =567810EC64A91CBCC1256EB400417FE8&file =74004181AC.pdf) and AmpliChip CYP Test [26]. CodeLink P450 Bioarray is developed by GE Healthcare and its multiplex level is the lowest among the microarrays for detecting variations of CYP genes. Although it can be used to detect a total of 110 polymorphisms of 9 CYP genes, 12 multiplex reactions are needed to be carried out to amplify targets. This DNA preparation method is laborious and cumbersome, a vital drawback for its routine use in practice. AmpliChip CYP Test, developed by Roche, is the first microarray-based pharmacogenomic test for diagnostic use. AmpliChip is able to distinguish common variations of *CYP2C19* and *CYP2D6* gene using a single multiplex reaction to amplify targets. However, low information content of AmpliChip CYP Test may limit its wide use in clinical practices and research investigations. Therefore, our CYP microarray fills the gap between AmpliChip CYP Test and CodeLink P450 Bioarray. It is important that *CYP3A5* and *CYP3A4* genes have overlapping substrate specificity and essentially no functional variant forms of *CYP3A4* gene have been observed in Orientals and Caucasians [8]. Genetic variations contributing to polymorphic expression of *CYP3A5* gene result in interindividual differences in the clearance of many CYP3A substrates [7, 8]. In general, the major genetic causes of polymorphism of P450s drug metabolism are polymorphisms of *CYP2D6*, *CYP2C9*, *CYP2C19* and *CYP3A5* genes. Our microarray may satisfy the needs of clinical practice and research investigations.

Genetic factors contributed to variability in drug response are polymorphisms in genes encoding drug-metabolizing enzymes, drug transporters, drug receptors, and drug targets. Detection of all these genetic polymorphisms (at least several thousand) related to drug response conduces to identify the right drug and dose for individual and may ultimately realize personal medicine. Although SNP microarray can genotype simultaneously tens of thousands of SNPs, the throughput level of SNP microarray is limited by DNA preparation and not by microarray technology [33]. With the development of SNP genotyping technology, there are many over 1000-plex DNA preparation methods invented, such as GoldenGate™, one-primer assay and molecular inversion probes [34-38]. However, these DNA preparation approaches aren't suitable for analysis of polymorphisms of CYP genes because of pseudogenes and high homology between members of CYP subfamily, even different CYP subfamilies [3, 28-30]. It is the reason why both AmpliChip CYP Test and CodeLink P450 Bioarray use multiplex PCR to amplify targets. Multiplex level of most multiplex PCR remains usually restricted to about five to ten targets due to primer dimerization [39, 40]. Consequently, it is difficult for coamplifying loci of highly homologous CYP genes to improve multiplexing levels of multiplex PCR to tens-of-plex, even according to rigorous standard developed by Wang et al [40]. Although fragments of four genes were coamplified in a one-tube reaction, we failed to coamplify these fragments with long fragments for the presence of *CYP2D6* duplication and deletion, respectively, in a single multiplex reaction. Because various kinds of microarray have been used successfully for comparative genomic hybridization experiments [33] and microarray technology is adapt for quantitative analysis, we developed a method for determination of *CYP2D6* gene copy number by comparison of signal intensities of probes for the presence of *CYP2D6* and *CYP2D9* gene, respectively. In addition, samples with different copies of the *CYP2D6* gene show different hybridization intensity patterns and thus duplicated alleles could be possible determinated by appropriate algorithm. The method needs further evaluation in large population due to rare incidence of *CYP2D6* gene duplication. Although multiplex level of single reaction described here is higher than others, new DNA preparation method remains to be developed to amplify most loci of CYP genes or drug-metabolism genes in a one-tube reaction, even including all drug-metabolism genes, drug transporter genes and drug target genes. Simple and rapid DNA preparation approaches and high information content

is the trend in development of mutation and polymorphism detection or screening microarray for clinical practices and research investigations. Thus, systematic and comprehensive genetic analysis, being critical for the understanding of drug response, could become feasible and cost-effective. This in turn will accelerate the discovery and development of personalized medicines.

In conclusion, we have developed and validated a CYP oligonucleotide microarray for genotyping variations of *CYP2C9*, *CYP2C19*, *CYP2D6* and *CYP3A5* genes in Eastern Han Chinese population. In addition to conventional SNP genotyping for detection of polymorphisms of CYP genes, we anticipate a future in which the CYP microarray used in routine clinical practice and may provide safe and effective therapy for individual patients.

Acknowledgements

This work was supported by the National High Technology Research and Development Program of China (863 Program) (Grant No. 2011AA100606), the National Natural Science Foundation of China (Grant Nos. 31101091, 31101172, 31271746, 31201144, 31501366).

Disclosure of conflict of interest

None.

Address correspondence to: Haihui Sheng, National Engineering Center for Biochip at Shanghai, No. 151, Libing Road, Shanghai 210203, China. E-mail: haihui_sheng@shbiochip.com; Haiyan Li, Ministry of Education Engineering Research Center of Bioreactor and Pharmaceutical Development, Jilin Agricultural University, N0 2888, Xincheng Street, Changchun 130118, China. E-mail: hyli99@163.com

References

[1] Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics 1998; 8: 283-289.

[2] Michalets EL. Update: clinically significant cytochrome P-450 drug interactions. Pharmacotherapy 1998; 18: 84-112.

[3] 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.

[4] Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2004; 5: 6-13.

[5] Xie HG, Prasad HC, Kim RB, Stein CM. CYP2C9 allelic variants: ethnic distribution and functional significance. Adv Drug Deliv Rev 2002; 54: 1257-1270.

[6] Lee CR, Goldstein JA, Pieper JA. Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. Pharmacogenetics 2002; 12: 251-263.

[7] Lamba JK, Lin YS, Schuetz EG, Thummel KE. Genetic contribution to variable human CYP3A-mediated metabolism. Adv Drug Deliv Rev 2002; 54: 1271-1294.

[8] Ingelman-Sundberg M. Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. Trends Pharmacol Sci 2004; 25: 193-200.

[9] Hu YF, He J, Chen GL, Wang D, Liu ZQ, Zhang C, Duan LF, Zhou HH. CYP3A5*3 and CYP3A4*18 single nucleotide polymorphisms in a Chinese population. Clin Chim Acta 2005; 353: 187-192.

[10] Wu L, Zhong W, Liu J, Han W, Zhong S, Wei Q, Liu S, Tang L. Human microsomal cyttrochrome P450-mediated reduction of oxysophocarpine, an active and highly toxic constituent derived from Sophora flavescens species, and its intestinal absorption and metabolism in rat. Fitoterapia 2015; 105: 26-36.

[11] Brockmoller J, Kirchheiner J, Meisel C, Roots I. Pharmacogenetic diagnostics of cytochrome P450 polymorphisms in clinical drug development and in drug treatment. Pharmacogenomics 2000; 1: 125-151.

[12] Tan ML, Lim LE. The effects of Andrographis paniculata (Burm.f.) Nees extract and diterpenoids on the CYP450 isoforms' activities, a review of possible herb-drug interaction risks. Drug Chem Toxicol 2015; 38: 241-53.

[13] 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: 983-988.

[14] Masimirembwa C, Persson I, Bertilsson L, Hasler J, Ingelman-Sundberg M. A novel mutant variant of the CYP2D6 gene (CYP2D6*17) common in a black African population: association with diminished debrisoquine hydroxylase activity. Br J Clin Pharmacol 1996; 42: 713-719.

[15] Ledesma MC, Agundez JA. Identification of subtypes of CYP2D gene rearrangements among carriers of CYP2D6 gene deletion and duplication. Clin Chem 2005; 51: 939-943.

[16] Bank PC, Swen JJ, Guchelaar HJ, van der Straaten T. GenoChip CYP2D6 macroarray as a method to genotype for CYP2D6 variants: results of a validation study in a Caucasian population. Pharmacogenomics 2015; 16: 681-687.

[17] Sanderson S, Emery J, Higgins J. CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: a HuGEnet systematic review and meta-analysis. Genet Med 2005; 7: 97-104.

[18] Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, Watkins PB, Daly A, Wrighton SA, Hall SD, Maurel P, Relling M, Brimer C, Yasuda K, Venkataramanan R, Strom S, Thummel K, Boguski MS, Schuetz E. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 2001; 27: 383-391.

[19] Lee SJ, Usmani KA, Chanas B, Ghanayem B, Xi T, Hodgson E, Mohrenweiser HW, Goldstein JA. Genetic findings and functional studies of human CYP3A5 single nucleotide polymorphisms in different ethnic groups. Pharmacogenetics 2003;13: 461-472.

[20] Hiratsuka M, Agatsuma Y, Omori F, Narahara K, Inoue T, Kishikawa Y, Mizugaki M. High throughput detection of drug-metabolizing enzyme polymorphisms by allele-specific fluorogenic 5' nuclease chain reaction assay. Biol Pharm Bull 2000; 23: 1131-1135.

[21] Ishiguro A, Kubota T, Soya Y, Sasaki H, Yagyu O, Takarada Y, Iga T. High-throughput detection of multiple genetic polymorphisms influencing drug metabolism with mismatch primers in allele-specific polymerase chain reaction. Anal Biochem 2005; 337: 256-261.

[22] 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: 476-485.

[23] Nevilie M, Selzer R, Aizenstein B, Maguire M, Hogan K, Walton R, Welsh K, Neri B, de Arruda M. Characterization of cytochrome P450 2D6 alleles using the Invader system. Biotechniques Suppl 2002; 34-38: 40-33.

[24] Sistonen J, Fuselli S, Levo A, Sajantila A. CYP2D6 genotyping by a multiplex primer extension reaction. Clin Chem 2005; 51: 1291-1295.

[25] Liu CH, Peck K, Huang JD, Lin MS, Wang CH, Hsu WP, Wang HW, Lee HL, Lai ML. Screening CYP3A single nucleotide polymorphisms in a Han Chinese population with a genotyping chip. Pharmacogenomics 2005; 6: 731-747.

[26] Chou WH, Yan FX, Robbins-Weilert DK, Ryder TB, Liu WW, Perbost C, Fairchild M, de Leon J, Koch WH, Wedlund PJ. Comparison of two CYP2D6 genotyping methods and assessment of genotype-phenotype relationships. Clin Chem 2003; 49: 542-551.

[27] Qi DC, Wu B, Tao SL, Zhou J, Qian HX, Wang D. Analysis of differentially expressed genes in malignant biliary strictures. Genet Mol Res 2014; 13: 2674-2682.

[28] Flockhart DA, Rae JM. Cytochrome P450 3A pharmacogenetics: the road that needs traveled. Pharmacogenomics J 2003; 3: 3-5.

[29] Lewis LD. Personalized drug therapy; the genome, the chip and the physician. Br J Clin Pharmacol 2005; 60: 1-4.

[30] Kimura S, Umeno M, Skoda RC, Meyer UA, Gonzalez FJ. 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 1989; 45: 889-904.

[31] Chou FC, Tzeng SJ, Huang JD. Genetic polymorphism of cytochrome P450 3A5 in Chinese. Drug Metab Dispos 2001; 29: 1205-1209.

[32] Niu CY, Luo JY, Hao ZM. Genetic polymorphism analysis of cytochrome P4502C19 in Chinese Uigur and Han populations. Chin J Dig Dis 2004; 5: 76-80.

[33] Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet 2005; 37: 549-554.

[34] Hardenbol P, Baner J, Jain M, Nilsson M, Namsaraev EA, Karlin-Neumann GA, Fakhrai-Rad H, Ronaghi M, Willis TD, Landegren U, Davis RW. Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol 2003; 21: 673-678.

[35] Hardenbol P, Yu F, Belmont J, Mackenzie J, Bruckner C, Brundage T, Boudreau A, Chow S, Eberle J, Erbilgin A, Falkowski M, Fitzgerald R, Ghose S, Iartchouk O, Jain M, Karlin-Neumann G, Lu X, Miao X, Moore B, Moorhead M, Namsaraev E, Pasternak S, Prakash E, Tran K, Wang Z, Jones HB, Davis RW, Willis TD, Gibbs RA. Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay. Genome Res 2005; 15: 269-275.

[36] Matsuzaki H, Dong S, Loi H, Di X, Liu G, Hubbell E, Law J, Berntsen T, Chadha M, Hui H, Yang G, Kennedy GC, Webster TA, Cawley S, Walsh PS, Jones KW, Fodor SP, Mei R. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 2004; 1: 109-111.

[37] Matsuzaki H, Loi H, Dong S, Tsai YY, Fang J, Law J, Di X, Liu WM, Yang G, Liu G, Huang J, Kennedy GC, Ryder TB, Marcus GA, Walsh PS,

Shriver MD, Puck JM, Jones KW, Mei R. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. Genome Res 2004; 14: 414-425.

[38] Shen R, Fan JB, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, Steemers F, Garcia F, Kermani BG, Gunderson K, Oliphant A. High-throughput SNP genotyping on universal bead arrays. Mutat Res 2005; 573: 70-82.

[39] Broude NE, Zhang L, Woodward K, Englert D, Cantor CR. Multiplex allele-specific target amplification based on PCR suppression. Proc Natl Acad Sci U S A 2001; 98: 206-211.

[40] Wang HY, Luo M, Tereshchenko IV, Frikker DM, Cui X, Li JY, Hu G, Chu Y, Azaro MA, Lin Y, Shen L, Yang Q, Kambouris ME, Gao R, Shih W, Li H. A genotyping system capable of simultaneously analyzing >1000 single nucleotide polymorphisms in a haploid genome. Genome Res 2005; 15: 276-283.