

# Genotyping and haplotyping of CYP2C19 functional alleles on thin-film biosensor chips

Kaori Nakamoto<sup>a</sup>, Judith R. Kidd<sup>b</sup>, Robert D. Jenison<sup>c</sup>, Curtis D. Klaassen<sup>a</sup>, Yu-Jui Yvonne Wan<sup>a</sup>, Kenneth K. Kidd<sup>b</sup> and Xiao-bo Zhong<sup>a</sup>

**Objectives** Numerous functional polymorphisms in the CYP2C19 gene have been identified; some alleles (e.g. CYP2C19\*2 and CYP2C19\*3) are associated with poor metabolism of CYP2C19 substrate drugs. Studies have found that the proportion of poor metabolizers, explained by CYP2C19\*2 and CYP2C19\*3, varies from less than 50% to more than 90% of poor metabolizers. Therefore, phenotype-genotype correlation studies should cover more than CYP2C19\*2 and CYP2C19\*3. A broader coverage, however, requires an easy-to-use and high-throughput genotyping platform. This broader coverage should also include the recently identified functional allele, CYP2C19\*10, which involves a nucleotide change adjacent to the altered nucleotide change in CYP2C19\*2. The currently used restriction fragment length polymorphism-based method for genotyping CYP2C19\*2 cannot distinguish between CYP2C19\*2 and CYP2C19\*10. We aim to develop a simple platform that can genotype all CYP2C19 functional alleles.

**Methods** We have developed a thin-film biosensor chip platform to genotype 16 exonic CYP2C19 variants, including two sets of two adjacent single nucleotide polymorphisms and 12 single single nucleotide polymorphisms, using a ligation strategy.

**Results** We demonstrate that this is a rapid, accurate, and inexpensive method for genotyping CYP2C19 variants using individual's genomic DNA samples. We further

demonstrate that this genotyping platform can be used to construct a haplotype structure of the CYP2C19 variants in a population, and to assign a haplotype combination to each individual on the basis of his/her genotype results.

**Conclusion** This assay can be applied in pharmacogenomic studies in both basic research and clinical laboratories.

It is also an ideal technology for pharmacogenomic tests in both developed and developing countries.

*Pharmacogenetics and Genomics* 17:103–114 © 2007 Lippincott Williams & Wilkins.

*Pharmacogenetics and Genomics* 2007, 17:103–114

**Keywords:** CYP2C19, genotyping, haplotyping, thin-film biosensor chip

<sup>a</sup>Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, Kansas City, Kansas, <sup>b</sup>Department of Genetics, Yale University School of Medicine, New Haven, Connecticut and <sup>c</sup>Inverness Medical-Biostar, Louisville, Colorado, USA

Correspondence and requests for reprints to Dr Xiaobo Zhong, PhD, Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, Kansas 66160, USA  
Tel: +1 913 588 0400; fax: +1 913 588 7501; e-mail: xzhong@kumc.edu

Sponsorship: This research is supported by NIH grant RR021940 (XB Zhong), NIH grant AA012081 (Y.Y. Wan), and NIH grant GM057672 (K.K. Kidd).

Received 5 June 2006 Accepted 27 September 2006

## Introduction

CYP2C19, a member of the cytochrome-P450 enzyme superfamily, is a monooxygenase that metabolizes a wide variety of drugs, including anticonvulsants (*S*-mephenytoin and diazepam) [1,2], antidepressants (barbiturates, citalopram, and imipramine) [3–5], antimalarial drugs (proguanil) [6,7], proton pump inhibitors (e.g. omeprazole) [8],  $\beta$ -adrenergic blockers (e.g. propranolol) [9], and HIV-protease inhibitors (e.g. nelfinavir) [10]. On the basis of the metabolic rates for the drugs, individuals are classified as extensive metabolizers (EM), intermediate metabolizers (IM), and poor metabolizers (PM). PMs have a much lower ability than average to oxidize these drugs and they may experience undesirable adverse effects at standard dosages, such as decreased awareness after administration of diazepam [2].

The frequency of CYP2C19 PMs is different in various ethnic populations and ranges from 2 to 5% in individuals of European ancestry [11,12], 4 to 7.5% in people of African descent [12–14], 13 to 20% in East Asians [12,15], and 38 to 79% in Pacific Islanders [16]. Numerous functional alleles have been identified [11,17–21] and named by the Human Cytochrome-P450 (CYP) Allele Nomenclature Committee (<http://www.imm.ki.se/CYPalleles/>). The most common variants or PM alleles are named as CYP2C19\*2 series of \*2A, \*2B, and \*2C and contain a variant that is a splice defective mutation in exon 5, with a G→A nucleotide change at 681 of cDNA sequence [single nucleotide polymorphism (SNP) ID in the National Council of Bioinformatics SNP database is rs4244285] [11]. This change alters the reading frame of the mRNA starting

with amino-acid 215 and produces a premature stop codon 20 amino acids downstream, which results in a truncated, nonfunctional protein. Another common PM allele (CYP2C19\*3) has a G→A mutation at position 636 of exon 4 (rs28399504), which also creates a premature stop codon [17]. Other null variants include a mutation (rs28399504) in the initiation codon (CYP2C19\*4) [18], a mutation in the stop codon (stop491Cys) (CYP2C19\*12) [19], and a splice mutation in intron 5 (CYP2C19\*7) [20]. Numerous synonymous and nonsynonymous polymorphisms have also been identified, including a T→C nucleotide change at 50 in the cDNA sequence, resulting in a Leu→Pro amino-acid change (CYP2C19\*14), an A→C (rs17882687) at 55 resulting in an Ile→Leu (CYP2C19\*15), a G→C at 276 (rs17878459) resulting in a Glu→Asp (CYP2C19\*2B), a T→C at 358 resulting in a Trp→Arg (CYP2C19\*8), a G→A at 395 resulting in an Arg→Gln (CYP2C19\*6), a G→A at 431 (rs17884712) resulting in an Arg→His (CYP2C19\*9), a G→A at 449 resulting in an Arg→His (CYP2C19\*11), a C→T at 680 (rs6413438) resulting in a Pro→Leu (CYP2C19\*10), an A→G at 991 (rs3758581) resulting in an Ile→Val (CYP2C19\*1B), and a C→T at 1228 (rs17879685) resulting in an Arg→Cys (CYP2C19\*13) [18,20,21]. Most nonsynonymous alleles produce allelic proteins showing lower enzyme activity than the common protein by in-vivo or in-vitro assays [11,17–22]. It should be noted that there are two pairs of adjacent polymorphisms existing in the CYP2C19 gene at 680C→T (rs6413438) and 681G→A (rs4244285), and at 990C→T (rs3758580) and 991A→G (rs3758581). Linkage disequilibrium between 681G and 990C, and between 681A and 990T are observed in several populations [19].

Homozygotes of the \*2/\*2 and \*3/\*3 or heterozygotes of the \*2/\*3 are associated with PM for drugs dependent on CYP2C19 metabolism, such as *S*-mephenytoin [23–25], omeprazole [26], and proguanil [15,27–29]. Frequencies of these two alleles have been established in global-wide populations, for example: in Africa (Ethiopia [23], Tanzania [13,30], Egypt [31], South Africa, and Zimbabwe [32]); Middle East (Israel [33,34] and Turkey [35]); Europe (Italy [36] and Russia [25,37]); South Asia (India [38,39] and Thailand [40]); East Asia (China [41,42], Korea [24,26], and Japan [43]); Australia [27] and Pacific Islands (Papua New Guinea [44] and Vanuatu [16]); North America (European Americans [45,46], African-Americans [12,46], Mexican-Americans [46]); and South America (European ancestry in Bolivia [47]). Some studies showed that \*2 and \*3 could explain more than 90% PM phenotypes [26,41,42], but other studies indicate that \*2 and \*3 could explain less than 50% of PMs [30,38,46]. Additional defective alleles that have been found to contribute to PM of *S*-mephenytoin include CYP2C19\*4 [18] and CYP2C19\*5 [20], but their frequencies in different populations have not been

established. To fully understand the genetic mechanism of CYP2C19 in PMs, all known functional variants need to be studied simultaneously in individuals from a set of global populations. Such studies would require a high-throughput genotyping platform with relatively low costs.

In this report, we describe development of thin-film biosensor chips for genotyping 16 CYP2C19 exonic variants (Fig. 1), including the two pairs of adjacent SNPs at 680/681 and 990/991. Ten out of the 16 polymorphisms have been characterized to have no or decreased CYP2C19 activity by in-vivo or in-vitro assays [11,17–21]. Because of their high sensitivity, ease-of-use, low cost, and optical characteristics, thin-film biosensor chips have been used in point-of-care doctor's offices for rapid detection of infectious pathogens, such as respiratory virus of influenza A and B [48], and sexually transmitted pathogens such as *Chlamydia* [49], *Gonorrhea*, and *Streptococcus B* [50]. We have applied this platform for high-throughput genotyping of SNPs [51] and micro-satellite polymorphisms [52]. Details of the experimental procedure for genotyping 16 CYP2C19 exonic variants, including the two pairs of adjacent SNPs, are described in this report. Accuracy of genotyping on thin-film biosensor chips is validated by comparison with the HapMap database and a previously studied population of Mexican-Americans. Application of high-throughput genotyping and haplotyping of CYP2C19 variants is demonstrated in a European-American population. The advantages of this system compared with other commercial CYP2C19 genotyping platforms are also discussed in this report, specially focusing on its potential application in clinical pharmacogenomic tests in hospitals.

## Methods

### Human participants

Twelve HapMap DNA samples were used for validating genotyping accuracy; one African YRI family trio (child NA18500, father NA18501, and mother NA18502); one European-American CEU family trio (maternal grandmother NA06985, mother NA06991, and maternal grandfather NA06993); three independent Chinese Han CHB individuals (NA18524, NA18526, and NA18529), and three independent Japanese JPT individuals (NA18940, NA18942, and NA18943). The genomic DNAs were purchased from Coriell Cell Repositories (Camden, New Jersey). A Mexican-American population ( $n = 70$ ) was also used in this study for validating genotyping accuracy. These DNA samples were collected by another research project on pharmacogenomics of CYP enzymes in Mexican-Americans [46]. Samples of European-Americans ( $n = 85$ ) were used for genotyping and haplotyping of CYP2C19 exonic variants. This population sample has been used in our global population studies (sample SA000020C in ALFRED, <http://alfred.med.yale.edu>). Studies on these human participants were approved by

the Human Subjects Committee of the University of Kansas Medical Center.

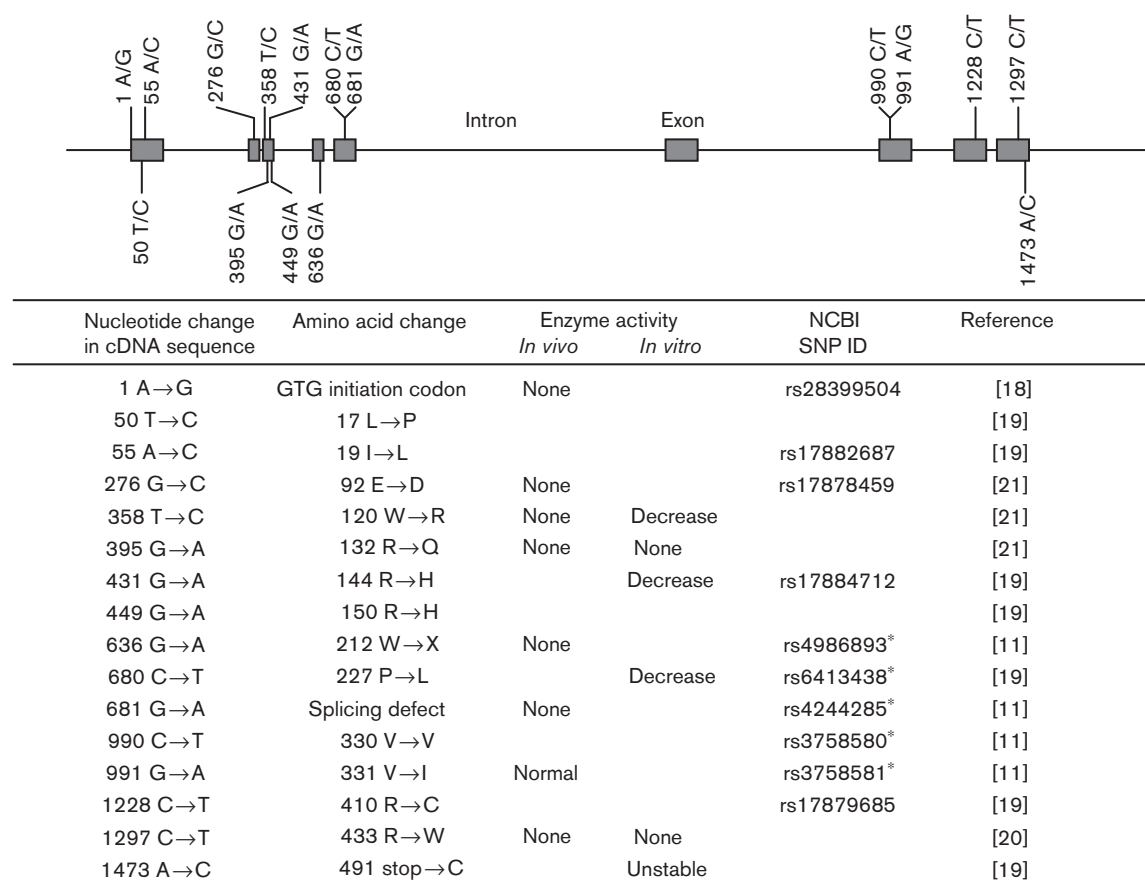
### Genotyping on thin-film biosensor chip

Sixteen exonic CYP2C19 SNPs, including 12 single SNPs and two pairs of adjacent SNPs, were selected from the Human Cytochrome-P450 (CYP) Allele Nomenclature. They are located in all exons except exon 6 (Fig. 1). Target DNA molecules for each exon were amplified by polymerase chain reaction (PCR). PCR primers were designed based on the following criteria to make the PCR reactions uniform: (i) product size was 120–250 bp with about 50–150 bases of flanking sequences in both directions around the SNP site and (ii) annealing temperature was about 60°C for the PCR reaction. The best primer sets were selected by DS Gene Software, version 1.5 (Accelrys, Cambridge, UK). The primer sequences and PCR product size for each exon are listed in Table 1. The selected primer sequences were synthesized by Invitrogen (Carlsbad, California, USA). Single or multiplex PCR reactions were performed at cycling conditions of 95°C for 15 min, 40 cycles of 94°C

for 15 s, 60°C for 15 s, and 72°C for 30 s, followed by 72°C for 5 min, using Go Taq Polymerase provided by Promega (Madison, Wisconsin, USA).

For each single SNP, three oligonucleotide probes were synthesized. A pair of allele-specific P-1 oligos, differing only in their 3'-terminal nucleotide sequence, have 40 nucleotides complementary to the corresponding target sequences on one side of the SNP, and an additional 10-dA residue at their 5'-ends, which constitutes a 'spacer'. The 5'-terminal nucleotide is modified with an aldehyde group, allowing covalent attachment to the chip surface containing a layer of hydrazine groups [51]. A second oligonucleotide probe (biotin-P2) with 20 nucleotides complementary to the other side of the SNP, and immediately adjacent to the SNP nucleotide, carries a biotin at the 3' end for detection, and a phosphate at its 5' end for ligation. For the two adjacent SNPs, four allele-specific P-1 oligos were synthesized with different combinations of the last two nucleotides, one for the common sequence, one for the single variant at the last base, one for the single variant at the second to the last

Fig. 1



\*SNPs are genotyped by the HapMap project.

Location of the 16 selected exonic variants in the CYP2C19 gene.

Table 1 Oligonucleotide sequences of capture probe (P1), detection probe (P2), synthetic targets and PCR primers

| Nucleotide change in cDNA sequence | Oligonucleotide sequence  |
|------------------------------------|---|
| <b>(a) Adjacent SNPs</b>           |   |
| Exon 5 (PCR product: 213 bp)       |   |
| Forward primer                     | AATTACAACCAGAGCTTGGC  |
| Reverse primer                     | TGTTGATGCCATCGATTCTTG   |
| 680-681 P1-CG                      | ALD-AAAAAAAAAAAAATATGCAATAATTTCCCACTATCATTGATTATTTCCCG          |
| 680-681 P1-TG                      | ALD-AAAAAAAAAAAAATATGCAATAATTTCCCACTATCATTGATTATTTCC <u>TG</u>  |
| 680-681 P1-CA                      | ALD-AAAAAAAAAAAAATATGCAATAATTTCCCACTATCATTGATTATTTCC <u>CA</u>  |
| 680-681 P1-TA                      | ALD-AAAAAAAAAAAAATATGCAATAATTTCCCACTATCATTGATTATTTCC <u>TA</u>  |
| 680-681 P2                         | Phosphate-GGAACCCATAACAAATTACT-biotin                           |
| 680-681 target for P1-CG           | GTAATTTGTTATGGGTTCCCGGAAATAATCAATGATAGT                         |
| 680-681 target for P1-TG           | GTAATTTGTTATGGGTTCC <u>AG</u> GAAATAATCAATGATAGT                |
| 680-681 target for P1-CA           | GTAATTTGTTATGGGTTCC <u>TG</u> GAAATAATCAATGATAGT                |
| 680-681 target for P1-TA           | GTAATTTGTTATGGGTTCC <u>TA</u> GAAATAATCAATGATAGT                |
| Exon 7 (PCR product: 163 bp)       |   |
| Forward primer                     | TGTTCCATTCTCCTTTTCC   |
| Reverse primer                     | ATGTATCTCTGGACCTCGTG  |
| 990-991 P1-CA                      | ALD-AAAAAAAAAAGTCTTGTCCAGCTAAAGTCCAGGAAGAGATTGAACGTGTCA         |
| 990-991 P1-TA                      | ALD-AAAAAAAAAAGTCTTGTCCAGCTAAAGTCCAGGAAGAGATTGAACGTGT <u>TA</u> |
| 990-991 P1-CG                      | ALD-AAAAAAAAAAGTCTTGTCCAGCTAAAGTCCAGGAAGAGATTGAACGTGT <u>CG</u> |
| 990-991 P1-TG                      | ALD-AAAAAAAAAAGTCTTGTCCAGCTAAAGTCCAGGAAGAGATTGAACGTGT <u>TG</u> |
| 990-991 P2                         | Phosphate-TTGGCAGAAACCGGAGCCCC-biotin                           |
| 990-991 target for P1-CA           | GGGGCTCCGGTTTCTGCCAATGACACGTTCAATCTCTTCT                        |
| 990-991 target for P1-TA           | GGGGCTCCGGTTTCTGCCAATGACACGTTCAATCTCTTCT                        |
| 990-991 target for P1-CG           | GGGGCTCCGGTTTCTGCCAATGACACGTTCAATCTCTTCT                        |
| 990-991 target for P1-TG           | GGGGCTCCGGTTTCTGCCAATGACACGTTCAATCTCTTCT                        |
| <b>(b) Single SNP</b>              |   |
| Exon 1 (PCR product: 202 bp)       |   |
| Forward primer                     | AGGCACACACACTTAATTAGC   |
| Reverse primer                     | ATCACTGGGAGAGGAGTAGG  |
| 1 P1-A                             | ALD-AAAAAAAAAAGCAAGCTCACGGTTGTCTTAACAAGAGGAGAAGGCTTCAA          |
| 1 P1-G                             | ALD-AAAAAAAAAAGCAAGCTCACGGTTGTCTTAACAAGAGGAGAAGGCTTCA <u>G</u>  |
| 1 P2                               | Phosphate-TGGATCCTTTTGTGGTCTT-biotin                            |
| 1 target for P1-A                  | AAGGACCACAAAAGGATCCACTGAAGCCTTCTCCTTGT                          |
| 1 target for P1-G                  | AAGGACCACAAAAGGATCCACTGAAGCCTTCTCCTTGT                          |
| 50 P1-T                            | ALD-AAAAAAAAAATTGTGGTCTTTGTGCTCTGTCTCATGTTTGTCTTCTCT            |
| 50 P1-C                            | ALD-AAAAAAAAAATTGTGGTCTTTGTGCTCTGTCTCATGTTTGTCTTCTC <u>CC</u>   |
| 50 P2                              | Phosphate-TTCAATCTGGAGACAGAGCT-biotin                           |
| 50 target for P1-T                 | AGCTCTGTCTCCAGATTGAAAGGAGAAGCAACATGAGAGA                        |
| 50 target for P1-C                 | AGCTCTGTCTCCAGATTGAAAGGAGAAGCAACATGAGAGA                        |
| 55 P1-A                            | AAAAAAAAAAGTCCCTTGTGCTCTGTCTCATGTTTGTCTTCTTCAA                  |
| 55 P1-C                            | AAAAAAAAAAGTCCCTTGTGCTCTGTCTCATGTTTGTCTTCTTCA <u>C</u>          |
| 55 P2                              | Phosphate-TCTGAGACAGAGCTCTGGG-biotin                            |
| 55 target for P1-A                 | CCCAGAGCTCTGTCTCCAGATTGAAAGGAGAAGCAACATG                        |
| 55 target for P1-C                 | CCCAGAGCTCTGTCTCCAGATTGAAAGGAGAAGCAACATG                        |
| Exon 2 (PCR product: 146 bp)       |   |
| Forward primer                     | CTCTGTATTTGGCCTGGAAC  |
| Reverse primer                     | TGCACACCTACCAAATCCTC  |
| 276 P1-G                           | ALD-AAAAAAAAAATATGAAGTGGTGAAGGAAGCCCTGATTGATCTTGGAGAG           |
| 276 P1-C                           | ALD-AAAAAAAAAATATGAAGTGGTGAAGGAAGCCCTGATTGATCTTGGAGAG <u>C</u>  |
| 276 P2                             | Phosphate-GAGTTTTCTGGAAGAGGCCA-biotin                           |
| 276 target for P1-G                | TGGCCTCTTCCAGAAACTCCTCTCCAAGATCAATCAGGGC                        |
| 276 target for P1-C                | TGGCCTCTTCCAGAAACTC <u>C</u> TCTCCAAGATCAATCAGGGC               |
| Exon 3 (PCR product: 226 bp)       |   |
| Forward primer                     | CAGTGTGAGCTTCTCTTTTC  |
| Reverse primer                     | TATGTTACCCACCCTTGG  |
| 358 P1-T                           | ALD-AAAAAAAAAACTCTTCTGTTAGGAATCGTTTTTCAGCAATGAAAGAGAT           |
| 358 P1-C                           | ALD-AAAAAAAAAACTCTTCTGTTAGGAATCGTTTTTCAGCAATGAAAGAGAG <u>C</u>  |
| 358 P2                             | Phosphate-GGAAGGAGATCCGGCGTTTT-biotin                           |
| 358 target for P1-T                | GAAACGCCGGATCTCCTTCCATCTCTTTCCATTGCTGAAAA                       |
| 358 target for P1-C                | GAAACGCCGGATCTCCTTCCG <u>T</u> CTCTTTCCATTGCTGAAAA              |
| 395 P1-G                           | ALD-AAAAAAAAAAGATGGAAGGAGATCCGGCGTTTCTCCCTCATGACGCTGC           |
| 395 P1-A                           | ALD-AAAAAAAAAAGATGGAAGGAGATCCGGCGTTTCTCCCTCATGACGCTGC <u>A</u>  |
| 395 P2                             | Phosphate-GAATTTTGGATGGGGAAGA-biotin                            |
| 395 target for P1-G                | TCTTCCCCTCCCAAATTCAGCAGCTCATGAGGGAGAAA                          |
| 395 target for P1-A                | TCTTCCCCTCCCAAATTCAGCAGCTCATGAGGGAGAAA                          |
| 431 P1-G                           | ALD-AAAAAAAAAATCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACCG           |
| 431 P1-A                           | ALD-AAAAAAAAAATCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGAC <u>C</u>    |
| 431 P2                             | Phosphate-TGTTCAAGAGGAAGCCCGCT-biotin                           |
| 431 target for P1-G                | AGCGGGCTTCTCTTGAACACGGTCTCAATGCTCCTCTTC                         |
| 431 target for P1-A                | AGCGGGCTTCTCTTGAACATGGTCTCAATGCTCCTCTTC                         |
| 449 P1-G                           | ALD-AAAAAAAAAAGGAAGGAGGAGCATTGAGGACCGTGTTCAGAGGAAGCCCG          |
| 449 P1-A                           | ALD-AAAAAAAAAAGGAAGGAGGAGCATTGAGGACCGTGTTCAGAGGAAGCC <u>C</u>   |
| 449 P2                             | Phosphate-CTGCCTTGTGAGGAGTTGA-biotin                            |
| 449 target for P1-G                | TCAACTCTCCACAAGGCAGCGGGCTTCTCTTGAACACGG                         |
| 449 target for P1-A                | TCAACTCTCCACAAGGCAGTGGGCTTCTCTTGAACACGG                         |

Table 1 (continued)

| Nucleotide change in cDNA sequence | Oligonucleotide sequence                                 |
|------------------------------------|--|
| Exon 4 (PCR product: 179 bp)       |  |
| Forward primer                     | CTGTGATCCCACTTTCATCC                                     |
| Reverse primer                     | CAGGAAGCAAAAACTTGGC                                      |
| 636 P1-G                           | AAAAAAAAAAAAAATTGAATGAAAACATCAGGATTGTAAGCACCCCCTGG       |
| 636 P1-A                           | AAAAAAAAAAAAAATTGAATGAAAACATCAGGATTGTAAGCACCCCCTGA       |
| 636 P2                             | Phosphate-ATCCAGGTAAGGCCAAGTT-biotin                     |
| 636 target for P1-G                | AAACTTGGCCTTACCTGGATCCAGGGGGTGCTTACAATCCT                |
| 636 target for P1-A                | AAACTTGGCCTTACCTGGATCCAGGGGGTGCTTACAATCCT                |
| Exon 8 (PCR product: 141 bp)       |  |
| Forward primer                     | GGGCACAACCATATTAACCTCC                                   |
| Reverse primer                     | GCTGAGAAAGGCATGAAGTAG                                    |
| 1228 P1-C                          | ALD-AAAAAAAAAAAAACAACAAGAATTTCCCAACCCAGAGATGTTTGACCCTC   |
| 1228 P1-T                          | ALD-AAAAAAAAAAAAACAACAAGAATTTCCCAACCCAGAGATGTTTGACCCTC   |
| 1228 P2                            | Phosphate-GTCACTTCTGGATGAAGGT-biotin                     |
| 1228 target for P1-C               | ACCTTCATCCAGAAAGTGACGAGGGGTCAAACATCTCTGGGT               |
| 1228 target for P1-T               | ACCTTCATCCAGAAAGTGACAGGGGTCAAACATCTCTGGGT                |
| Exon 9 (PCR product: 294 bp)       |  |
| Forward primer                     | CTCACAGTTACACATGAGGAG                                    |
| Reverse primer                     | TTGGACCAGAGGAAAGAGAG                                     |
| 1297 P1-C                          | ALD-AAAAAAAAAAAAAGGAGTAACCTCTCCCTATGTTTGTATTTTCAGGAAAAAC |
| 1297 P1-T                          | ALD-AAAAAAAAAAAAAGGAGTAACCTCTCCCTATGTTTGTATTTTCAGGAAAAAC |
| 1297 P2                            | Phosphate-GGATTTGTGTGGGAGAGGGC-biotin                    |
| 1297 target for P1-C               | GCCCTCTCCCACACAATCCGTTTTCTGAAAATAACAAAC                  |
| 1297 target for P1-T               | GCCCTCTCCCACACAATCCGTTTTCTGAAAATAACAAAC                  |
| 1473 P1-A                          | ALD-AAAAAAAAAATGTCCCGCCCTTATCAGCTGTGCTTCATTCTGTCTGA      |
| 1473 P1-C                          | ALD-AAAAAAAAAATGTCCCGCCCTTATCAGCTGTGCTTCATTCTGTCTGC      |
| 1473 P2                            | Phosphate-AGAAGCACAGATGGTCTGGC-biotin                    |
| 1473 target for P1-A               | GCCAGACCATCTGTGCTTCTCAGACAGGAATGAAGCACAG                 |
| 1473 target for P1-C               | GCCAGACCATCTGTGCTTCTCAGACAGGAATGAAGCACAG                 |

PCR, polymerase chain reaction; SNP, single nucleotide polymorphism. Underlines indicate SNP sites.

base, and one for the double variant sequence. To test genotyping specificity, oligonucleotide targets (40-mer) perfectly matched for each P1 and P2 probe combination were also synthesized. The P1, P2, and target sequences are listed in Table 1. The synthesized P1 oligos were diluted to 100  $\mu\text{mol/l}$  in 0.1 mol/l phosphate buffer, pH 7.8. A P1 working solution of 1  $\mu\text{mol/l}$  in 0.1 mol/l phosphate buffer, pH 7.8, and 10% glycerol was prepared for each P-1 probe before spotting. The thin-film biosensor chips were purchased from Inverness Medical-Biostar (Louisville, Colorado, USA) (<http://www.biostar.com/>). A 10 cm diameter wafer costs \$33. Each wafer was cut into  $\sim 150$   $7 \times 7 \text{ mm}^2$  chips. P1-working solution (20 nl) was spotted on a  $7 \times 7 \text{ mm}^2$  chip surface in an  $8 \times 8$  dot format on a wafer, by a BioDot PC-controlled dispense arrayer AD3200. After the spotted chips were incubated in a humidity-controlled chamber for at least 2 h, the chips were washed with 0.1% sodium dodecyl sulfate in water and then air dried. Each of the spotted chips was transferred into an  $8 \times 8 \text{ mm}^2$  well in a 96-well microtiter plate, and fixed by double-sided tape. A standard operating procedure for genotyping SNPs on the printed biosensor chips as previously described [51] was applied with a slight modification. An arrayed chip was assembled into a square well of a 96-well microtiter plate for hybridization. A ligation reaction was carried out in the microtiter plate well containing an arrayed chip. A reaction solution (100  $\mu\text{l}$ ) contained 100 femtomoles of each relevant PCR amplicon for the 16 CYP2C19 SNPs,

10 nmol/l P-2 probe (one for each SNP), and 5 units of mutant Ampligase in a buffer of 20 mmol/l Tris-HCl, pH 8.3, 25 mmol/l KCl, 10 mmol/l  $\text{MgCl}_2$ , 0.5 mmol/l NAD, 0.01% Triton X-100, 10% formamide, and 5 mg/ml alkaline-treated casein (ATC). The ligation reaction was incubated for 20 min at 62°C. All 96 chips in a 96-well plate were processed simultaneously. After a stringent wash (3 times in 0.01 mol/l NaOH at room temperature and 3 times in  $0.1 \times \text{SSC}$ ), the chips were incubated with an antibiotin horse-radish peroxidase (HRP) conjugate (Jackson ImmunoResearch Lab, West Grove, Pennsylvania, USA; 1  $\mu\text{g/ml}$  in  $5 \times \text{SSC}$  buffer containing 5 mg/ml ATC) for 10 min, and the chips were rinsed with  $0.1 \times \text{SSC}$ . One hundred microliter of a precipitate-generating HRP substrate 3,3',5,5' tetramethyl benzidine (BioFex, Owings Mills, Maryland, USA) was added to each chip and incubated for 5 min, rinsed in  $\text{ddH}_2\text{O}$ , and air dried. Signal images were captured by a black-white digital camera. Genotypes were manually recorded on the basis of the signal's position for each pair of capture probes.

#### Analysis of haplotypes by statistical tools

Genotypes generated by the thin-film biosensor chips were saved in a linkage format file and uploaded into HAPLOT (Yale University, New Haven, Connecticut, USA), a software package for graphical illustration of linkage disequilibrium (LD) blocks, haplotype frequency, and haplotype-tagging SNPs for multiple populations [53]. Four different block partition algorithms are

incorporated into HAPLOT to generate (LD) block structure, on the basis of LD measure  $D'$  (confidence interval) [54], the four-gamete test [55], the solid spine of LD [56], and LD measure  $r^2$  (Kidd  $r^2$ ) [53]. Frequencies of common and less common alleles, genotypes of homozygous common allele, homozygous less common allele, and heterozygous, frequency of observed heterozygous and predicted heterozygous genotypes, and Hardy-Weinberg test were determined and calculated for each variant by the HAPLOT software package. LD blocks were defined by using the Kidd  $r^2$  partition algorithms. Haplotypes were generated through the program HAPLO [57], which implements the EM algorithm [58].

## Results

### Genotyping of two-adjacent SNPs on thin-film biosensor chips

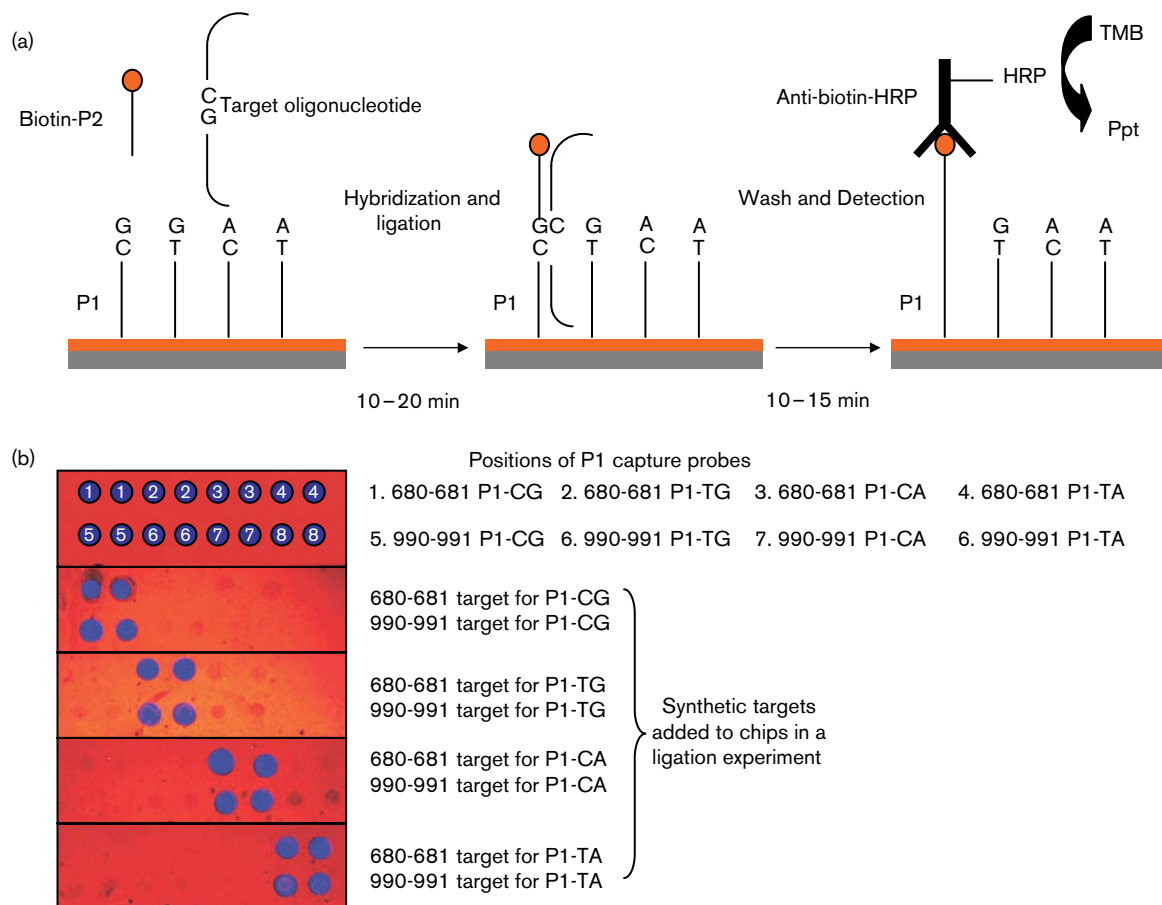
Although two-adjacent SNPs do not frequently occur, they occasionally exist in the human genome, for example, at the loci of 580–581 and 990–991 in the CYP2C19 cDNA sequence. Most platforms for genotyping a single SNP may not be able to distinguish easily the four possible allelic combinations of two adjacent SNPs. For example, the restriction fragment length polymorphism-based SNP genotyping platform relies on recognition of polymorphic sequences by restriction enzymes. At least three or four different enzymes are required to recognize each combination of the two adjacent polymorphic nucleotides. In some situations, it is not possible to find appropriate restriction enzymes. For example at 580–581 of CYP2C19, a *Sma*I site is present in the common allele CYP2C19\*1 (TCCCGGGAAC). Many previous studies identified the CYP2C19\*2 allele as failure to be digested by *Sma*I based on the de Morais' method [11]. The *Sma*I enzyme, however, cannot recognize the variant allele at position 580 (TCCTGGGAAC), the variant allele at position 581 (TCCCAGGAAC), nor the double-variant allele (TCCTAGGAAC). Thus, CYP2C19\*2 alleles defined by *Sma*I digestion may not be all CYP2C19\*2, but a mixture of different sequences with different functional consequences. Allele-specific primers for PCR amplification can identify the presence of two variants at two adjacent SNPs by using different strands as PCR templates in two separate amplification reactions. But this assay cannot distinguish whether the two adjacent variants occur on the same chromosome or on different chromosomes. In the present study, a ligation strategy is used to detect haplotype status of two adjacent SNPs in a single step on thin-film biosensor chips (Fig. 2a). Four capture probes (P1), differing at the last two nucleotides at their 3' terminals, are immobilized on the thin-film biosensor chip surface by their 5' modified nucleotide [51]. When a reaction solution, containing a biotin-labeled detection probe (P2), a synthetic single-stranded oligonucleotide (target) with a specific sequence perfectly matched to one of the capture probes, and a

thermostable DNA ligase (Ampligase), is applied on the chip surface, the biotinylated P2 is only ligated to the P1 probe, which has a sequence perfectly matched to the target. The last two nucleotides are very critical for the ligation by a thermostable ligase. After a stringent wash with NaOH to remove all noncovalently binding molecules, followed by a color-signal-generation procedure, a positive signal was observed at the position where the biotinylated P2 probe is immobilized to the P1 probe that is perfectly matched to the target at both sites. Specificity for genotyping two adjacent SNPs of CYP2C19 580C→T-581G→A and 990C→T-991A→G is robustly achieved by this strategy. The P1, P2, and synthetic targets (see Table 1) were synthesized by Invitrogen. The P1 probes were arrayed on a thin-film biosensor chip surface in duplicate by a BioDot PC-controlled dispense arrayer AD3200 (Fig. 2b top panel). The synthetic targets were applied on a chip surface for allele-specific ligation. In a test experiment, specificity was perfect at a ligation condition with 10% formamide at 62°C (Fig. 2b).

### Genotyping of 16 functional CYP2C19 SNPs on thin-film biosensor chips

Our thin-film biosensor chip simultaneously genotypes the 16 exonic CYP2C19 SNPs (list in Fig. 1), including two pairs of adjacent SNPs at 580–581 and 990–991, and 12 single SNPs. The capture P1 probes were arrayed in duplicate in a format of 8 dots per row × 8 dots per column on a 7 × 7 mm<sup>2</sup> chip (Fig. 3, left panel). Ligation specificity for each SNP was tested by applying a pool of synthetic single-stranded oligonucleotide targets. As expected, only perfectly matched probes showed positive signals (Fig. 3, right top panel). To further verify genotyping accuracy on this platform from a human participant, we performed a genotyping comparison with the HapMap database using 12 selected HapMap samples (three from each population of CEU, CHB, JPT, and YRI). Of the 16 selected exonic CYP2C19 polymorphisms, 10 are included in the NCBI SNP database, of which five have been genotyped by the HapMap project, including rs4986893 (636 G→A), rs6413438 (680 C→T), rs4244285 (681 G→A), rs3758580 (990 C→T), and rs3758581 (991 G→A). Genotypes of the five SNPs (including both of the pairs of adjacent SNPs) were generated by the Illumina Bead-Array platform using allele-specific primers for PCR in the HapMap project. Two sets of primers were used to detect the adjacent SNPs using opposite strands as PCR templates. The genotypes of the 12 HapMap samples were downloaded from the HapMap database and listed in Table 2. In our thin-film biosensor chip genotyping platform, target molecules from each participant were prepared by single or multiplex PCR amplification. The amplified targets were applied to the chip surface to generate genotypes by an allele-specific ligation reaction. An image of the signals on each chip was captured by a

Fig. 2



(a) A strategy for detecting two-adjacent SNPs in the CYP2C19 gene on thin-film biosensor chip using a thermostable DNA ligase. (b) Images show detection specificity for the two adjacent SNPs using synthetic single-stranded oligonucleotides as the targets. SNP, single nucleotide polymorphism; TMB, 3,3',5,5' tetramethyl benzidine; HRP, horse-radish peroxidase; Ppt, precipitate.

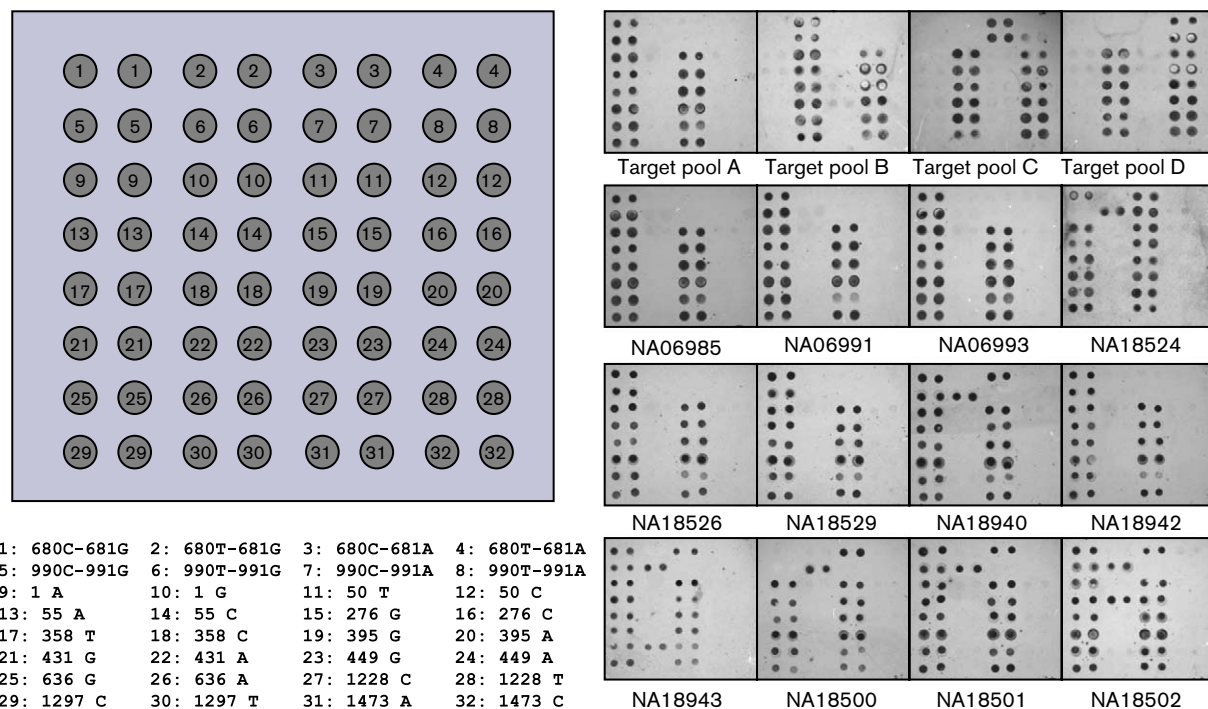
black–white digital camera and is shown in the right panel of Fig. 3. A 100% concordance of the genotypes of the five SNPs was obtained by the two methods, indicating that high genotyping accuracy is obtained by the ligation-based genotyping platform on the thin-film biosensor chip. It is worth pointing out that in sample NA18524 (a Chinese), three markers, rs4244285, rs3758580, and rs3758581, are listed as heterozygous in the HapMap database, but it is not possible to distinguish what haplotypes of the two-adjacent SNPs at rs3758580 and rs3758581 (990–991) exist in this sample based on the genotyping results. Our chip image, however, clearly shows haplotype status, indicating 990T-991G on one chromosome and 990C-991A on the other chromosome. In addition to heterozygous rs4244285 and rs3758580 in sample NA18502 (a mother in a family trio of Yoruba), a heterozygous 55A/C is also identified by the thin-film biosensor chip. This set of data shows linkage disequilibrium between 681G and 990C and between 681A and 990T. To further validate genotyping accuracy on thin-

film biosensor chips, we genotyped the 16 SNPs in 70 Mexican–American samples that have been genotyped previously for CYP2C19\*2 and CYP2C19\*3 using the restriction fragment length polymorphism method [46]. Again, there was 100% concordance by the two methods.

#### Haplotyping of CYP2C19 in an European–American population

To further demonstrate application of the thin-film biosensor chip to analyze haplotype structure of CYP2C19 exonic variants in a population, we genotyped the 16 CYP2C19 exonic variants in a sample of 85 European–Americans. The genotype results were uploaded into HAPLOT and the genotyping data are summarized in Table 3 with allele frequencies and genotype frequencies. All allele frequencies are in Hardy–Weinberg equilibrium. A strong LD block was found that covers the entire CYP2C19 gene using Kidd  $r^2$  block partition algorithm. Six haplotypes were identified by the EM algorithm and presented in Fig. 4a with blue

Fig. 3



Genotyping of the 16 CYP2C19 exonic variants on thin-film biosensor chips. Left panel: positions of the capture probes for each single nucleotide polymorphism (SNP) arrayed in duplicate. Right panel top row: targets are a pool of synthetic single-stranded oligonucleotides. Pool a: common alleles for all SNPs; pool b: less common alleles for the 12 single SNP together with 680T-680G and 990T-991G; pool c: less common alleles for the 12 single SNP together with 680C-680A and 990C-991A; pool d: less common alleles for the 12 single SNP together with 680T-680A and 990T-991A. Right panel row 2-4: targets are a pool of polymerase chain reaction products from each HapMap sample.

Table 2 Genotypes of five CYP2C19 SNPs in the 12 selected HapMap samples downloaded from the HapMap database

| HapMap sample | rs4986893 636 G→A | rs6413438 680 C→T | rs4244285 681 G→A | rs3758580 990 C→T | rs3758581 991 G→A |
|---------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| NA06985       | GG                | CC                | GG                | CC                | GG                |
| NA06991       | GG                | CC                | GG                | CC                | GG                |
| NA06993       | GG                | CC                | GG                | CC                | GG                |
| NA18524       | GG                | CC                | GA                | CT                | GA                |
| NA18526       | GG                | CC                | GG                | CC                | GG                |
| NA18529       | GG                | CC                | GG                | CC                | GG                |
| NA18940       | GG                | CC                | GA                | CT                | GG                |
| NA18942       | GG                | CC                | GG                | CC                | GG                |
| NA18943       | GA                | CC                | GA                | CT                | GG                |
| NA18500       | GG                | CC                | AA                | TT                | GG                |
| NA18501       | GG                | CC                | GA                | CT                | GG                |
| NA18502       | GG                | CC                | GA                | CT                | GG                |

SNP, single nucleotide polymorphism.

boxes, and red boxes representing common alleles and less common alleles, respectively. The dominant haplotype is haplotype 1 (72.3%) with the common variants at all loci, which corresponds to CYP2C19\*1 in CYP allele nomenclature and has normal CYP2C19 enzyme activity. Haplotype 2 (5.9%), corresponding to CYP2C19\*1C, has one nucleotide change from G to A at 991. Although this variant results in an amino-acid change from Val to Ile at 331, in-vivo experiments showed that this change does not alter CYP2C19 enzyme activity [19]. Haplotypes 3

(12.2%), 4 (4.7%), and 5 (2.6%) all have the splice variant caused by 681A and thus should have no enzyme activity. A synonymous substitution at 990 (T allele) is strongly associated with the 681A allele. Two nonsynonymous alleles are also associated with the 681A allele in haplotype 4 (276 C) and haplotype 5 (449 A). The close approximations to the alleles listed by the CYP Allele Nomenclature Committee are given in Fig. 4 of the haplotypes. Haplotype 3 corresponds to CYP2C19\*2A and haplotype 4 corresponds to CYP2C19\*2B, but

**Table 3 Genotypes of the CYP2C19 exonic alleles in European-American samples**

| Position in cDNA sequence                   | 1    | 50  | 55  | 276  | 358 | 395 | 431  | 449 | 636 | 680 | 681  | 990  | 991  | 1228 | 1297 | 1473 |
|---|------|-----|-----|------|-----|-----|------|-----|-----|-----|------|------|------|------|------|------|
| Common allele                               | A    | T   | A   | G    | T   | G   | G    | G   | G   | C   | G    | C    | G    | C    | C    | C    |
| Less common allele                          | G    | C   | C   | C    | C   | A   | A    | A   | A   | T   | A    | T    | A    | T    | T    | T    |
| Allele frequency (%)                        |      |     |     |      |     |     |      |     |     |     |      |      |      |      |      |      |
| Common allele                               | 98.8 | 100 | 100 | 95.3 | 100 | 100 | 97.1 | 100 | 100 | 100 | 79.4 | 79.4 | 92.9 | 100  | 100  | 100  |
| Less common allele                          | 1.2  | 0   | 0   | 4.7  | 0   | 0   | 2.9  | 0   | 0   | 0   | 20.6 | 20.6 | 7.1  | 0    | 0    | 0    |
| Genotype                                    |      |     |     |      |     |     |      |     |     |     |      |      |      |      |      |      |
| Homo common allele ( <i>n</i> )             | 84   | 85  | 85  | 77   | 85  | 85  | 80   | 85  | 85  | 85  | 55   | 55   | 73   | 85   | 85   | 85   |
| Homo less common allele ( <i>n</i> )        | 1    | 0   | 0   | 0    | 0   | 0   | 0    | 0   | 0   | 0   | 5    | 5    | 0    | 0    | 0    | 0    |
| Heterozygous ( <i>n</i> )                   | 0    | 0   | 0   | 8    | 0   | 0   | 5    | 0   | 0   | 0   | 25   | 25   | 12   | 0    | 0    | 0    |
| Observed heterozygosity (%)                 | 0    | 0   | 0   | 9.4  | 0   | 0   | 5.9  | 0   | 0   | 0   | 29.4 | 29.4 | 14.1 | 0    | 0    | 0    |
| Predict heterozygosity (%)                  | 2.3  | 0   | 0   | 9.0  | 0   | 0   | 5.7  | 0   | 0   | 0   | 32.7 | 32.7 | 13.1 | 0    | 0    | 0    |
| Hardy–Weinberg <i>P</i> -value <sup>a</sup> | 0.14 |     |     | 0.67 |     |     | 1.0  |     |     |     | 0.19 | 0.19 | 0.40 |      |      |      |

<sup>a</sup>Probability based on an exact test with 10 000 simulations.

**Fig. 4**


(a) Haplotypes of the 16 CYP2C19 exonic variants in a European-American population and the corresponding CYP nomenclature alleles. (b) Groups of genotypes explained as the possible combination of two haplotypes for the European-American individuals and predicted CYP2C19 phenotypes. EM, extensive metabolizers; IM, intermediate metabolizers; PM, poor metabolizers.

haplotype 5 has no allele nomenclature yet. Only one sample was found in this population, which carries homozygous haplotypes 6 (CYP2C19\*4). With a sample

size of 85 unrelated individuals, the common haplotypes of these SNPs should be correct and their frequencies are reasonably accurate for an European-American sample.

The less frequent haplotypes and their frequencies, however, need to be verified in additional samples, especially because of likely heterogeneity among the ancestral European populations. On the basis of the most probable combination of two of the estimated haplotypes, the 85 European-Americans can be assigned to 10 genotypes (Fig. 4b). These genotypes predict 63.5% EM, 29.4% IM, and 7.1% PM, similar to other reports [11,12].

## Discussion

We report here a rapid, accurate, and inexpensive method using optical thin-film biosensor chips for the detection of 16 CYP2C19 exonic variants, including both single and two adjacent SNPs. In addition to their presence at 680C→T-681G→A and 990C→T-991G→A in the CYP2C19 gene, two adjacent SNPs also occur in other CYP450 genes, for example, 85C→A-86G→C in CYP2B6 (\*17) [59], 1075A→C-1076T→C in CYP2C9 (\*3/\*4) [60], 1757C→T-1758G→T, 1846G→A-1847G→A, 1978C→T-1979T→C in CYP2D6 (\*2L/\*8, \*4/rs11568728, and \*2C/\*20) [61,62]. Therefore, a simple procedure that can detect both single and two adjacent SNPs is necessary for genotyping these CYP alleles. By using a ligation strategy with a highly selective thermostable ligase (Ampligase), we have demonstrated that the thin-film biosensor chip assay can detect both single and two adjacent SNPs in a single step with good specificity.

Pharmacogenomics is likely to change our current practice of medicine into personalized medicine. Drug therapy will be optimized to provide better efficacy, greater safety, and fewer side effects for each individual. This new generation of medicine is expected to become routine in the next 10–20 years, but it will require breakthroughs in research, education, and clinical practice. One of the most important requirements is simple, inexpensive, and easy-to-use pharmacogenomic assays for clinical laboratories in hospitals and point-of-care doctor's offices. EU and US Food and Drug Administration have approved the first pharmacogenomic test 'AmpliChip P450 test' produced by Roche Diagnostics (Basel, Switzerland), which is a microarray-based platform analyzing 29 polymorphisms and mutations for the CYP2D6 gene and two exonic variants for the CYP2C19 gene [63]. This platform requires instruments for PCR amplification, DNA hybridization, and fluorescent-image scanning. Costs for the chip and reagents per sample are in the range of \$400 to \$500. Another commercially available pharmacogenomic platform is TaqMan Drug Metabolism Genotyping Assays provided by Applied Biosystems (Foster City, California, USA). This assay uses the TaqMan platform and can detect about 2000 variants in more than 220 drug metabolizing genes, including 17 variants for the CYP2C19 gene. A TaqMan instrument system is required. The cost of reagents per

sample per variant is about \$0.40. Here, we have demonstrated an alternative platform for genotyping functional P450 alleles on thin-film biosensor chips. This assay is extremely robust, exhibits high flexibility, can be adapted for low throughput or moderate-to-high throughput requirements, and most importantly, is less expensive than the AmpliChip P450 test, TaqMan Drug Metabolism Genotyping Assays, and other high throughput alternatives. On the basis of an analysis for 1536 genotypes of the 16 CYP2C19 SNPs in 96 samples (one microtiter plate), cost for reagents and materials is ~15 cents per sample per SNP (a total \$216/1536 genotypes), which includes \$107 for isolation of genomic DNA from 96 samples using Invitrogen ChargeSwitch gDNA Mini Tissue Kit; \$15 for 576 PCR reactions (six single or multiplex PCR per sample for 96 samples) using Go Taq PCR kit from Promega; \$3 for oligonucleotide synthesis of P1 and P2 by Invitrogen (\$80 for one set, enough for more than 10 000 reactions); \$21 for thin-film biosensor chips from Inverness Medical-Biostar; \$36 for Ampligase Thermostable DNA ligase from Epicentre Technologies; \$10 for Anti-biotin-HRP from Jackson Immuno Research Laboratory and HRP substrate 3,3',5,5' tetramethyl benzidine from BioFox; \$24 for PCR plates, 96-well microtiter plate, tips, and tubes. Genotyping results can be obtained within a single day from sample collection. The results are visible to the unaided human eye and do not require any image-reading instrument. Once the chips are printed, this assay can be applied for pharmacogenomic tests in any research or clinical laboratory with basic molecular biology facilities (PCR machines and water bath). We have demonstrated that this platform is applicable in a field study in Africa (Zambia) on macrophage migration inhibitory factor genotypes with malaria-infected patients [52], indicating this is an ideal platform in developing countries that lack modern high-throughput genotyping instruments.

Using European-American DNA samples as an example, we have demonstrated how to use the thin-film biosensor chip to genotype the 16 CYP2C19 exonic variants in this population and how to construct haplotypes or alleles of the exonic variants using statistical tools. This strategy allows one to determine how genetic variants organize as haplotypes or alleles in a population, and to determine phase of adjacent two SNP haplotypes in a single step. If more than one functional variant is identified in an individual, this strategy allows us to determine whether the variants are located on the same chromosome or on different chromosomes. This is very important because different combinations of two haplotypes can have different effects on CYP2C19 function. For example in genotype group 4 (Fig. 4b), eight individuals have been identified as heterozygous at three loci (276 G/C, 681 G/A, and 990 C/T). The 276 C variant results in an amino-acid change reducing CYP2C19 activity, the 681 A

variant causes a splicing variant and produces a truncated protein with no CYP2C19 activity, and the 990T is a synonymous polymorphism. If 276 C and 681 A are located on the same chromosome, these individuals have one chromosome with normal sequences and another chromosome with two mutant sequences. Therefore, they should be CYP2C19 IM. If 276 C and 681 A are on different chromosomes, both copies of the CYP2C19 gene will produce a CYP2C19 enzyme with reduced or no activity. These individuals should be CYP2C19 PM. On the basis of the most likely combination of the two haplotypes, we can assign these individuals into the group of Genotype 4, heterozygous for haplotype 1 and haplotype 4, and predict that they have an intermediate metabolizer phenotype. This finding is a clear example of how haplotype data can have more predictive power than genotype data on individual SNPs.

Not only do exonic variants affect protein structure and function, variants in 5'-upstream (promoter and 5'-UTR), 3'-downstream (3'-UTR), and introns can affect transcription, translation, and protein processing. Therefore, the combination of all variants ultimately determines the phenotypes. Recently, a C→T SNP at -806 in the CYP2C19 promoter region (\*17) was identified [64]. The less common allele T increases transcription activity of the CYP2C19 enzyme. Individuals with homozygous T/T allele have higher metabolic ratio of omeprazole than those with C/C or C/T. Future studies should include this allele and other genetic variants in all different portions of the CYP2C19 gene, and determine functional effects of the genetic variants on CYP2C19 phenotype. With such information, a comprehensive pharmacogenomic test for genotyping and haplotyping of multiple genetic variations in various portions of the CYP2C19 gene will provide prediction of the effect of a combination of genetic alterations on drug metabolism, guiding physicians in the selection of appropriate dosages. The thin-film biosensor chip for genotyping CYP2C19 variants will be one of the platforms for such pharmacogenomic tests.

## Acknowledgement

We thank Dr Eifm Glob for providing mutant Ampligase for SNP genotyping.

## References

- 1 Wilkinson GR, Guengerich FP, Branch RA. Genetic polymorphism of S-mephenytoin hydroxylation. *Pharmacol Ther* 1989; **43**:53-76.
- 2 Ono S, Hatanaka T, Miyazawa S, Tsutsui M, Aoyama T, Gonzalez FJ, Satoh T. Human liver microsomal diazepam metabolism using cDNA-expressed cytochrome P450s: role of CYP2B6, 2C19 and the 3A subfamily. *Xenobiotica* 1996; **26**:1155-1166.
- 3 Kupfer A, Branch RA. Stero selective mephobarbital hydroxylation cosegregates with mephenytoin hydroxylation. *Clin Pharmacol Ther* 1985; **38**:414-418.
- 4 Sindrup SH, Brosen K, Hansen MGJ, Aaesjorgensen T, Overo KF, Gram LF. Pharmacokinetics of citalopram in relation to the sparteine and the mephenytoin oxidation polymorphisms. *Ther Drug Monit* 1993; **15**:11-17.
- 5 Baumann P, Jonzier-Perey M, Koeb L, Kupfer A, Tinguely D, Schopf J. Amitriptyline pharmacokinetics and clinical response: II. Metabolic polymorphism assessed by hydroxylation of debrisoquine and mephenytoin. *Int Clin Psychopharmacol* 1986; **1**:102-112.
- 6 Ward SA, Helsby NA, Skjelbo E, Brosen K, Gram LF, Breckenridge AM. The activation of the biguanide antimalarial proguanil co-segregates with the mephenytoin oxidation polymorphism - a panel study. *Br J Clin Pharmacol* 1991; **31**:689-692.
- 7 Wright JD, Helsby NA, Ward SA. The role of S-mephenytoin hydroxylase (CYP2C19) in the metabolism of the antimalarial biguanides. *Br J Clin Pharmacol* 1995; **39**:441-444.
- 8 Andersson T, Regardh CG, Lou YC, Zhang Y, Dahl ML, Bertilsson L. Polymorphic hydroxylation of S-mephenytoin and omeprazole metabolism in Caucasian and Chinese subjects. *Pharmacogenetics* 1992; **2**:25-31.
- 9 Ward SA, Walle T, Walle UK, Wilkinson GR, Branch RA. Propranolol's metabolism is determined by both mephenytoin and debrisoquin hydroxylase activities. *Clin Pharmacol Ther* 1989; **45**:72-79.
- 10 Khaliq Y, Gallicano K, Seguin I, Fyke K, Carignan G, Bulman D, *et al.* Single and multiple dose pharmacokinetics of nelfinavir and CYP2C19 activity in human immunodeficiency virus-infected patients with chronic liver disease. *Br J Clin Pharmacol* 2000; **50**:108-115.
- 11 de Morais SMF, Wilkinson GR, Blaisdell J, Nakamura K, Meyer UA, Goldstein JA. The major genetic defect responsible for the polymorphism of S-mephenytoin in humans. *J Biol Chem* 1994; **269**:15419-15422.
- 12 Goldstein JA, Ishizaki T, Chiba K, de Morais SM, Bell D, Krahn PM, Evans DA. Frequencies of the defective CYP2C19 alleles responsible for the mephenytoin poor metabolizer phenotype in various Oriental, Caucasian, Saudi Arabian and American black populations. *Pharmacogenetics* 1997; **7**:59-64.
- 13 Herrlin K, Massele AY, Jande M, Alm C, Tybring G, Abdi YA, *et al.* Bantu Tanzanians have a decreased capacity to metabolize omeprazole and mephenytoin in relation to their CYP2C19 genotype. *Clin Pharmacol Ther* 1998; **64**:391-401.
- 14 Edeki TI, Goldstein JA, deMorais SMF, Hajiloo L, Butler M, Chapdelaine P, *et al.* Genetic polymorphism of S-mephenytoin 4'-hydroxylation in African-Americans. *Pharmacogenetics* 1996; **6**:357-360.
- 15 de Morais SMF, Goldstein JA, Xie H-G, Huang S-L, Lu Y-Q, Xia H, *et al.* Genetic analysis of the S-mephenytoin polymorphism in a Chinese population. *Clin Pharmacol Ther* 1995; **58**:404-412.
- 16 Kaneko A, Lum JK, Yaviong L, Takahashi N, Ishizaki T, Bertilsson L, *et al.* High and variable frequencies of CYP2C19 mutations: medical consequences of poor drug metabolism in Vanuatu and other Pacific islands. *Pharmacogenetics* 1999; **9**:581-590.
- 17 De Morais SM, Wilkinson GR, Blaisdell J, Meyer UA, Nakamura K, Goldstein JA. Identification of a new genetic defect responsible for the polymorphism of (S)-mephenytoin metabolism in Japanese. *Mol Pharmacol* 1994; **46**:594-598.
- 18 Ferguson RJ, De Morais SM, Benhamou S, Bouchardy C, Blaisdell J, Ibeanu G, *et al.* A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. *J Pharmacol Exp Ther* 1998; **284**:356-361.
- 19 Blaisdell J, Mohrenweiser H, Jackson J, Ferguson S, Coulter S, Chanas B, *et al.* Identification and functional characterization of new potentially defective alleles of human CYP2C19. *Pharmacogenetics* 2002; **12**:703-711.
- 20 Ibeanu GC, Blaisdell J, Ghanayem BI, Beyeler C, Benhamou S, Bouchardy C, *et al.* An additional defective allele, CYP2C19\*5, contributes to the S-mephenytoin poor metabolizer phenotype in Caucasians. *Pharmacogenetics* 1998; **8**:129-135.
- 21 Ibeanu GC, Goldstein JA, Meyer U, Benhamou S, Bouchardy C, Dayer P, *et al.* Identification of new human CYP2C19 alleles (CYP2C19\*6 and CYP2C19\*2B) in a Caucasian poor metabolizer of mephenytoin. *J Pharmacol Exp Ther* 1998; **286**:1490-1495.
- 22 Richardson TH, Jung F, Griffin KJ, Wester M, Raucy JL, Kemper B, *et al.* A universal approach to the expression of human and rabbit cytochrome P450s of the 2C subfamily in *Escherichia coli*. *Arch Biochem Biophys* 1995; **323**:87-96.
- 23 Persson I, Akillu E, Rodrigues F, Bertilsson L, Ingelman-Sundberg M. S-mephenytoin hydroxylation phenotype and CYP2C19 genotype among Ethiopians. *Pharmacogenetics* 1996; **6**:521-526.
- 24 Roh HK, Dahl ML, Johansson I, Ingelman-Sundberg M, Cha YN, Bertilsson L. Debrisoquine and S-mephenytoin hydroxylation phenotypes and genotypes in a Korean population. *Pharmacogenetics* 1996; **6**:441-447.
- 25 Marandi T, Dahl ML, Rago L, Kiivet R, Sjoqvist F. Debrisoquine and S-mephenytoin hydroxylation polymorphisms in a Russian population living in Estonia. *Eur J Clin Pharmacol* 1997; **53**:257-260.
- 26 Roh HK, Dahl ML, Tybring G, Yamada H, Cha YN, Bertilsson L. CYP2C19 genotype and phenotype determined by omeprazole in a Korean population. *Pharmacogenetics* 1996; **6**:547-551.

- 27 Hoskins JM, Shenfield GM, Gross AS. Relationship between proguanil metabolic ratio and CYP2C19 genotype in a Caucasian population. *Br J Clin Pharmacol* 1998; **46**:499–504.
- 28 Kaneko A, Bergqvist Y, Taleo G, Kobayakawa T, Bjorkman A. Proguanil disposition and toxicity in malaria patients from Vanuatu with high frequencies of CYP2C19 mutations. *Pharmacogenetics* 1999; **9**:317–326.
- 29 Hoskins JM, Shenfield GM, Gross AS. Concordance between proguanil phenotype and CYP2C19 genotype in Chinese. *Eur J Clin Pharmacol* 2003; **59**:611–614.
- 30 Bathum L, Skjelbo E, Mutabingwa TK, Madsen H, Horder M, Brosen K. Phenotypes and genotypes for CYP2D6 and CYP2C19 in a black Tanzanian population. *Br J Clin Pharmacol* 1999; **48**:395–401.
- 31 Hamdy SI, Hiratsuka M, Narahara K, El-Enany M, Moursi N, Ahmed MS, Mizugaki M. Allele and genotype frequencies of polymorphic cytochromes P450 (CYP2C9, CYP2C19, CYP2E1) and dihydropyrimidine dehydrogenase (DPYD) in the Egyptian population. *Br J Clin Pharmacol* 2002; **53**:596–603.
- 32 Dandara C, Masimirembwa CM, Magimba A, Sayi J, Kaaya S, Sommers DK, et al. Genetic polymorphism of CYP2D6 and CYP2C19 in east- and southern African populations including psychiatric patients. *Eur J Clin Pharmacol* 2001; **57**:11–17.
- 33 Sviri S, Shpizen S, Leitersdorf E, Levy M, Caraco Y. Phenotypic-genotypic analysis of CYP2C19 in the Jewish Israeli population. *Clin Pharmacol Ther* 1999; **65**:275–282.
- 34 Luo HR, Aloumanis V, Lin KM, Gurwitz D, Wan YJ. Polymorphisms of CYP2C19 and CYP2D6 in Israeli ethnic groups. *Am J Pharmacogenomics* 2004; **4**:395–401.
- 35 Aynacioglu AS, Sachse C, Bozkurt A, Kortunay S, Sacak M, Schroder T, et al. Low frequency of defective alleles of cytochrome P450 enzymes 2C19 and 2D6 in the Turkish population. *Clin Pharmacol Ther* 1999; **66**:185–192.
- 36 Scordo MG, Caputi AP, D'Arrijo C, Fava G, Spina E. Allele and genotype frequencies of CYP2C9, CYP2C19 and CYP2D6 in an Italian population. *Pharmacol Res* 2004; **50**:195–200.
- 37 Gaikovitch EA, Cascorbi I, Mrozikiewicz PM, Brockmoller J, Frotschl R, Kopke K, et al. Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. *Eur J Clin Pharmacol* 2003; **59**:303–312.
- 38 Lamba JK, Dhiman RK, Kohli KK. CYP2C19 genetic mutations in North Indians. *Clin Pharmacol Ther* 2000; **68**:328–335.
- 39 Jose R, Chandrasekaran A, Sam SS, Gerard N, Chanolean S, Abraham BK, et al. CYP2C9 and CYP2C19 genetic polymorphisms: frequencies in the south Indian population. *Fundam Clin Pharmacol* 2005; **19**:101–105.
- 40 Tassaneeyakul W, Tawalee A, Tassaneeyakul W, Kukongviriyapan V, Blaisdell J, Goldstein JA, Gaysomsiri D. Analysis of the CYP2C19 polymorphism in a North-eastern Thai population. *Pharmacogenetics* 2002; **12**:221–225.
- 41 Zhang S, Dong Z, Tang L, Zhou Q, Wu Z. Cytochrome P450 2C19 gene polymorphism in four Chinese nationality populations. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2002; **19**:52–54.
- 42 He N, Yan FX, Huang SL, Wang W, Xiao ZS, Liu ZQ, Zhou HH. CYP2C19 genotype and S-mephenytoin 4'-hydroxylation phenotype in a Chinese Dai population. *Eur J Clin Pharmacol* 2002; **58**:15–18.
- 43 Kimura M, Ieiri I, Mamiya K, Urae A, Higuchi S. Genetic polymorphism of cytochrome P450s, CYP2C19, and CYP2C9 in a Japanese population. *Ther Drug Monit* 1998; **20**:243–247.
- 44 Masta A, Lum JK, Tsukahara T, Hwaihwanje I, Kaneko A, Panu MM, et al. Analysis of Sepik populations of Papua New Guinea suggests an increase of CYP2C19 null allele frequencies during the colonization of Melanesia. *Pharmacogenetics* 2003; **13**:697–700.
- 45 Mizutani T. PM frequencies of major CYPs in Asians and Caucasians. *Drug Metab Rev* 2003; **35**:99–106.
- 46 Luo HR, Poland RE, Lin KM, Wan YJ. Genetic polymorphism of CYP2C19 in Mexican Americans: a cross-ethnic comparative study. *Clin Pharmacol Ther* 2006 (in press).
- 47 Bravo-Villalta HV, Yamamoto K, Nakamura K, Baya A, Okada Y, Horiuchi R. Genetic polymorphism of CYP2C9 and CYP2C19 in a Bolivian population: an investigative and comparative study. *Eur J Clin Pharmacol* 2005; **61**:179–184.
- 48 Hindiye M, Goulding C, Morgan H, Kenyon B, Langer J, Fox L, et al. Evaluation of BioStar Flu oia assay for rapid detection of influenza A and B viruses in respiratory specimens. *J Clin Virol* 2000; **17**:119–126.
- 49 Pate M, Dixon P, Hardy K, Crosby M, Hook EW. Evaluation of the biostar chlamydia oia assay with specimens from women attending a sexually transmitted disease clinic. *J Clin Microbiol* 1998; **36**:2183–2186.
- 50 Carroll KC, Ballou D, Varner M, Chun H. Rapid detection of group B streptococcal colonization of the genital tract by a commercial optical immunoassay. *Eur J Clin Microbiol Infect Dis* 1996; **15**:206–210.
- 51 Zhong XB, Reynolds R, Kidd JR, Kidd KK, Jenison R, Marlar RA, Ward DC. Single-nucleotide polymorphism genotyping on optical thin-film biosensor chips. *Proc Natl Acad Sci USA* 2003; **100**:11559–11564.
- 52 Zhong XB, Leng L, Beitin A, Chen R, McDonald C, Hsiao B, et al. Simultaneous detection of microsatellite repeats and SNPs in the macrophage migration inhibitory factor (MIF) gene by thin-film biosensor chips and application to rural field studies. *Nucleic Acids Res* 2005; **33**:e121.
- 53 Gu S, Pakstis AJ, Kidd KK. HAPLOT: a graphical comparison of haplotype blocks, tagSNP sets and SNP variation for multiple populations. *Bioinformatics* 2005; **21**:3938–3939.
- 54 Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science* 2002; **296**:2225–2229.
- 55 Wang N, Deng M, Chen T, Waterman MS, Sun F. A dynamic programming algorithm for haplotype partitioning. *Proc Natl Acad Sci USA* 2002; **99**:7335–7339.
- 56 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**:263–265.
- 57 Hawley ME, Kidd KK. HAPLO: a program using the EM algorithm to estimate the frequencies of multi-site haplotypes. *J Hered* 1995; **86**:409–411.
- 58 Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via EM algorithm. *J Roy Statist Soc Ser B* 1977; **39**:1–22.
- 59 Klein K, Lang T, Saussele T, Barbosa-Sicard E, Schunck WH, Eichelbaum M, et al. Genetic variability of CYP2B6 in populations of African and Asian origin: allele frequencies, novel functional variants, and possible implications for anti-HIV therapy with efavirenz. *Pharmacogenet Genomics* 2005; **15**:861–873.
- 60 Imai J, Ieiri I, Mamiya K, Miyahara S, Furuumi H, Nanba E, et al. Polymorphism of the cytochrome P450 (CYP) 2C9 gene in Japanese epileptic patients: genetic analysis of the CYP2C9 locus. *Pharmacogenetics* 2000; **10**:85–89.
- 61 Gaedigk A, Bhatena A, Ndjountche L, Pearce RE, Abdel-Rahman SM, Alander SW, et al. Identification and characterization of novel sequence variations in the cytochrome P4502D6 (CYP2D6) gene in African Americans. *Pharmacogenomics J* 2005; **5**:173–182.
- 62 Marez-Allorge D, Ellis SW, Lo Guidice JM, Tucker GT, Broly F. A rare G2061 insertion affecting the open reading frame of CYP2D6 and responsible for the poor metabolizer phenotype. *Pharmacogenetics* 1999; **9**:393–396.
- 63 Jain KK. Applications of AmpliChip CYP450. *Mol Diagn* 2005; **9**: 119–127.
- 64 Sim SC, Risinger C, Dahl ML, Akililu E, Christensen M, Bertilsson L, Ingelman-Sundberg M. A common novel CYP2C19 gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. *Clin Pharmacol Ther* 2006; **79**:103–113.