

Multiplex single nucleotide polymorphism genotyping by adapter ligation-mediated allele-specific amplification

Wei-peng Wang^{a,b}, Kun-yi Ni^b, Guo-hua Zhou^{a,c,*}

^a Huadong Research Institute for Medicine and Biotechnics, Nanjing 210002, People's Republic of China

^b Department of Analytical Chemistry, China Pharmaceutical University, Nanjing 210038, People's Republic of China

^c Medical School, Nanjing University, Nanjing 210093, People's Republic of China

Received 28 February 2006

Available online 5 May 2006

Abstract

An improved approach for increasing the multiplex level of single nucleotide polymorphism (SNP) typing by adapter ligation-mediated allele-specific amplification (ALM–ASA) has been developed. Based on an adapter ligation, each reaction requires n allele-specific primers plus an adapter-specific primer that is common for all SNPs. Thus, only $n + 1$ primers are used for an n -plex PCR amplification. The specificity of ALM–ASA was increased by a special design of the adapter structure and PCR suppression. Given that the genetic polymorphisms in the liver enzyme cytochrome P450 *CYP2D6* (debrisoquine 4-hydroxylase) have profound effects on responses of individuals to a particular drug, we selected 17 SNPs in the *CYP2D6* gene as an example for the multiplex SNP typing. Without extensive optimization, we successfully typed 17-plex SNPs in the *CYP2D6* gene by ALM–ASA. The results for genotyping 70 different genome samples by the 17-plex ALM–ASA were completely consistent with those obtained by both Sanger's sequencing and PCR restriction fragment length polymorphism (PCR–RFLP) analysis. ALM–ASA is a potential method for SNP typing at an ultra-low cost because of a high multiplex level and a simple optimization step for PCR. High-throughput SNP typing could be readily realized by coupling ALM–ASA with a well-developed automation device for sample processing.

© 2006 Published by Elsevier Inc.

Keywords: SNP; Genotyping; Adapter; ALM–ASA; *CYP2D6*

Accompanied by the completion of the Human Genome Project (HGP),¹ more than 2 million sequence-verified single nucleotide polymorphisms (SNPs) have been clarified [1,2]. In the field of detection of linkage disequilibrium between SNP markers and potential alleles of disease or environment susceptibility, genotyping of known SNPs with high throughput is becoming more and more important. Therefore, there is an increasing demand for develop-

ing efficient, accurate, and inexpensive SNP typing methods. To this point, many approaches based on various principles have been developed for genotyping such as sequence-specific hybridization [3–5], allele-specific extension [6,7] or ligation [8–10], pyrosequencing [11,12], and structure-specific cleavage [13,14]. Most of them are based on single SNP typing, but some of them can be used as a multiplex format by coupling multiplex PCR with microarray [15,16], cytometry [17], or mass spectrometry [18,19] platforms. Because the running cost and the amount of required genomic DNAs for typing an SNP decrease dramatically along with the increase of a multiplex level, a high level of multiplex SNP typing is required for inexpensive and high-throughput genotyping.

Because most of the methods for multiplex SNP typing rely on a multiplex PCR, the capability of multiplex PCR is

* Corresponding author. Fax: +86 25 8451 4223.

E-mail address: zhouguohua@tsinghua.org.cn (G. Zhou).

¹ Abbreviations used: HGP, Human Genome Project; SNP, single nucleotide polymorphism; PS, PCR suppression; DMSO, dimethyl sulfoxide; EtBr, ethidium bromide; PCR–RFLP, PCR restriction fragment length polymorphism; ARMS, amplification refractory mutation system; MADGE, microplate array diagonal gel electrophoresis.

of great importance to the accuracy and throughput of SNP typing. The bottlenecks of multiplex PCR are mainly primer–primer interactions and nonspecific amplifications [20]. Several strategies have been developed to improve multiplex PCR amplification, including a two-step amplification (one step for preamplification of a long DNA and the other step for reamplification of short amplicons) [21], a two-round amplification (the first round for a gene-specific amplification by anchored primers and the next round for a uniform amplification by universal primers with the same sequence as the anchored part in the anchored primers) [22], and a PCR suppression (PS)-based amplification [23]. Using PS-based multiplex PCR, the multiplex level is increased and the optimization procedure is simplified significantly. However, the amplification efficiency for longer targets is decreased greatly because the hairpin loops of large size form stable secondary structures that inhibit the annealing of the universal primer [23]. In addition, a non-specific amplification by only the universal primer may occur in PS-based PCR because the targets (hairpin loops) contain the sequence complementary to the universal primer, although there is no allele-specific extension product. Moreover, the adapter in PS-based PCR has a very high GC content, and this may increase the chance of non-specific amplification. Although PS-based PCR has been used for 30-plex PCR, only 9-plex PCR using nine allele-specific primers were described [24]. In the current article, we demonstrate an improved procedure for increasing the level of multiplex allele-specific PCR termed adapter ligation-mediated allele-specific amplification (ALM-ASA), which is based on adapter ligation combined with allele-specific amplification. The liver enzyme cytochrome P450 *CYP2D6* (debrisoquine 4-hydroxylase) has a role in the metabolism of many different types of pharmacotherapeutic drugs. It shows a very high degree of interindividual and interethnic variability due primarily to the extensive genetic polymorphism that influences expression and function. Hence, we selected 17 SNPs in the *CYP2D6* gene as an example for evaluating ALM-ASA.

Materials and methods

Reagents and genomic samples

LA Taq DNA polymerase, *Taq* DNA polymerase, T4 DNA ligase, and restriction endonucleases *Mbo*I were obtained from TaKaRa BioTech (Dalian, China). Other chemicals were of a commercially extra-pure grade. All solutions were prepared with deionized and sterilized water.

A total of 70 Chinese volunteers participated in this research after giving written informed consent. They all were healthy individuals of Han nationality. Genomic DNA was extracted according to the standard phenol/chloroform method from whole blood collected in tubes containing anticoagulant EDTA. The purity and concentration of the extracted DNA were determined by a UV-Vis spectrophotometer (Naka Instruments, Japan). The extracted

DNA was stored at 4°C in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

Target sequences and primers

The nucleotide sequences and 17 SNPs in the *CYP2D6* gene (GenBank Accession No. AY545216) were used for the method validation. Based on the database in the public domain (www.imm.ki.se/CYPalleles/cyp2d6.htm), these 17 SNPs are as follows: 100C>T, 138G>insT, 883G>C, 957C>T, 997C>G, 1023C>T, 1661G>C, 1707T>del, 1846G>A, 1973A>insG, 2470T>C, 2549A>del, 2613AGA>del, 2950G>C, 3198C>G, 3277T>C, and 3828G>A; however, they were renamed as those shown in Table 1. All of the oligonucleotides were synthesized and purified by Invitrogen (Shanghai, China).

Preparation of adapter-ligated DNA

To increase the sensitivity and specificity of SNP typing, initial amplification of the entire *CYP2D6* gene with 4682 bp was carried out using the primers complementary to unique intronic sequences [25]. PCR was performed in a reaction volume of 10 µl that contained 2.5 mM MgCl₂, 0.4 mM of each dNTP, 0.4 µM of each outer primer, 5% dimethyl sulfoxide (DMSO), 0.5 U of *LA Taq* DNA polymerase, and 10 to 20 ng of genomic DNA. Thermal cycling was performed in a PTC-225 Thermal Cycler (MJ Research, Watertown, MA, USA) with an initial denaturation of 5 min at 94°C, followed by 30 cycles of 1 min at 94°C, 30 s at 60°C, 5 min at 68°C, and a terminal extension of 10 min at 68°C. Then 1 µl of long PCR products (~100–500 ng) was digested in 5 µl with 5 U of *Mbo*I restriction enzyme for 1 h at 37°C. After heat inactivation for 15 min at 70°C, 1 µl of the digested DNA was incubated in 5 µl with 1.0 µM of adapter for 1 h at 16°C containing 5 U of T4 DNA ligase. The adapter was prepared by annealing oligonucleotides of 5'-CCCCACTTCTTGTCTCTCATCAG GCG-CATCACTCG-3' and 5'-GATCCGAGTGATGC GCTAAG-3'. The ligation was terminated by heating the solution at 70°C for 15 min.

Allele-specific amplification and detection

The adapter-ligated DNA was amplified in a 10-µl PCR mixture consisting of 10 mM of Tris-HCl (pH 8.0), 50 mM of KCl, 1.5 mM of MgCl₂, 0.2 mM of each dNTP, 1 µl of adapter-ligated template solution, 1.0 µM of universal primer (5'-CCCCACTTCTTGTCTCTCAT-3'), approximately 2.0 µM of allele-specific primers mixture (the final concentrations are listed in Table 1), and 1 U of *Taq* DNA polymerase. The PCR cycling was carried out in a PTC-225 Thermal Cycler as follows: 94°C for 3 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min. Final extension was completed at 72°C for 10 min. After amplification, 3 µl of allele-specific PCR products was analyzed by electrophoresis on a 2% agarose gel consisting of

Table 1
SNPs used for the method evaluation

SNP code	SNP identifier ^a	Primer code ^b	Primer sequence (5' → 3') ^c	Melting temperature (°C) ^d	Product size (bp)	Concentration in 17-plex PCR (μmol/L)
SNP-1	4300 C>T	P1W(M)	CGCTGGGCTGCACGCTtCC(T)	66.6 (64.5)	704	0.050
SNP-2	4338G>insT	P2W(M)	TGCCCCGGGCTGGGCAACgTG(T)	68.6 (66.6)	667 (668)	0.050
SNP-3	5082G>C	P3W(M)	CCTGACCCTCCCTCTGtAG(C)	64.5 (64.5)	202	0.250
SNP-4	5156C>T	P4W(M)	TCGTGCTCAATGGGCTcGC(T)	62.3 (60.2)	128	0.250
SNP-5	5196C>G	P5W(M)	TGACCCACGtCGAGGACtCC(G)	66.6 (66.6)	89	0.250
SNP-6	5222C>T	P6W(M)	GCCCGCCTGTGCCCATgAC(T)	66.6 (64.5)	62	0.250
SNP-7	5862G>C	P7W(M)	GCAGAGGCGCTTCTCCcTG(C)	64.5 (64.5)	293	0.057
SNP-8	5908del T	P8W(M)	AAGAAGTCGCTGGAGCtGT(G)	59.9 (62.3)	247 (246)	0.057
SNP-9	6047G>A	P9W(M)	CCGCATCTCCCACCCcAG(A)	66.6 (64.5)	108	0.286
SNP-10	6174A>insG	P10W(M)	TGCTGGACCTAGCTCAcGA(G)	60.2 (60.2)	1012 (1013)	0.010
SNP-11	6671T>C	P11W(M)	TGTCCCCGTCCTCCTGgAT(C)	62.3 (64.5)	515	0.050
SNP-12	6750del A	P12W(M)	GAGCTGCTAACTGAGCtCA(G)	60.2 (62.3)	436 (435)	0.050
SNP-13	6814del AGA	P13W(M)	TTCCTGGCAGAGATGGtGA(G)	62.5 (62.3)	371 (368)	0.100
SNP-14	7151G>C	P14W(M)	TACATCCGGATGTGCAcCG(C)	60.2 (60.2)	164	0.100
SNP-15	7399C>G	P15W(M)	GACGTGATAGGtCAGGaGC(G)	62.3 (62.3)	1182	0.025
SNP-16	7478T>C	P16W(M)	TGCAGCGCTTTGGGGAaAT(C)	62.3 (62.3)	1103	0.050
SNP-17	8029G>A	P17W(M)	CATCACCAACCTGTCAaCG(A)	60.2 (60.2)	552	0.125

^a All SNPs are from the *CYP2D6* gene with the Accession No. AY545216 in GenBank. The nucleotide residue 1 is located at the start of the reference sequence.

^b The letter W represents the primer specific to the wild-type, and the letter M in parentheses represents the primer specific to the mutant.

^c The uppercase letters (in bold) in the 3 terminus of the primer represent the base specific to the SNP type, and the lowercase letters represent the artificially mismatched base.

^d Melting temperatures were calculated according to the software of Primer Premier version 5.0 (www.premierbiosoft.com). Melting temperatures of mutant-specific primers are in parentheses.

0.5 μg/ml of ethidium bromide (EtBr) at 100 V for 25 min in 1× TAE buffer. Images of the gels were taken using Gene-Genius BioImaging Systems (Syngene, UK).

Results

ALM-ASA

The procedure of multiplex SNP typing by ALM-ASA is illustrated in Fig. 1, which is based on preamplification, enzymatic digestion, an adapter ligation, and an allele-specific amplification. To increase the sensitivity and specificity of SNP typing, first, a preamplification is carried out for producing a long target containing all SNPs of interest. Second, the preamplified DNA fragments are digested by a restriction endonuclease to form sticky ends. Third, the adapter is ligated to either end of the digested fragment. The adapter is designed as a “Y” shape to prevent the extension reaction from the 3' ends of the ligated fragments. Usually, two or three uncomplementary bases in the 3' ends are enough to block the primer extension. Fourth, an allele-specific amplification is performed by allele-specific primers and a universal primer. To type a biallelic SNP, in general, two allele-specific primers are required. The amplification only occurs when the 3' end of the allele-specific primer matches the target sequence. There are two ways to discriminate the allele-specific amplification products. One way is to label two allele-specific primers with different dyes and then to detect the signals at different wavelengths; the other way is to perform each of the allele-specific amplifications separately without the use of dye-

labeled primers. To lower the detection cost, we used two reaction tubes to type SNPs. For single SNP typing, two primers (a universal primer and an allele-specific primer) are needed in each of two reaction tubes (tube W and tube M). However, only $n + 1$ primers are required in each tube to type n -plex SNPs because the universal primer is common for all SNPs. Finally, the allele-specific amplification

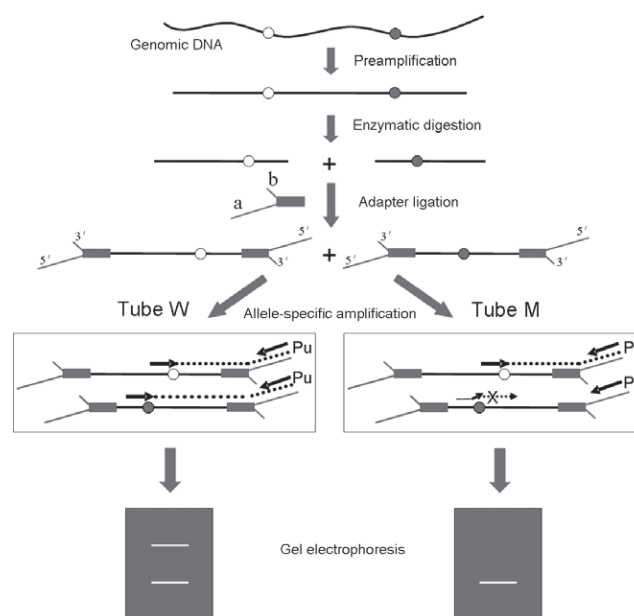


Fig. 1. Schematic of SNP genotyping by ALM-ASA. Two SNPs of heterozygote and mutant homozygote were used as an example for the illustration. W, wild-type; M, mutant; Pu, universal primer.

products are separated by electrophoresis such as agarose gel electrophoresis, PAGE, and chip-based electrophoresis. The type of each SNP is obtained by comparing the electrophoresis profile from each of two tubes.

As shown in Fig. 1, the sequence of the universal primer corresponds to that of flap “a” (5′-terminal region of the adapter-ligated fragments). Because the length of flap a (uncomplementary to flap “b”) is longer than the length of the universal primer, no template complementary to the universal primer exists in the PCR mixture when no allele-specific extension occurs. If the 3′ end of the allele-specific primer is complementary to the target, an extension reaction occurs and the template complementary to the universal primer is produced. Therefore, only allele-specific amplification takes place in the current ALM-ASA system.

Adapter and allele-specific primers

In multiplex PCR, preferential amplification of certain targets may occur because of the impaired rates of annealing and extension of different primers. If the difference between each primer pair used for allele-specific amplifications is reduced, the preferential amplification will be minimized. In this study, we used an adapter to ligate either end of the enzymatically digested fragment before allele-specific amplification for reducing the number of primers in multiplex PCR (Fig. 1). To type a biallelic SNP, in general, two allele-specific primers are required. Because every PCR shared a common primer whose sequence is the same as that in the 5′-terminal region of the adapter-ligated fragments, only $n + 1$ primers were employed for n -plex PCR. As a result, the preferential amplification caused by primers was greatly minimized.

Another problem in multiplex PCR is the spurious amplification that causes the consumption of reaction components as well as nonspecific bands in an electrophoresis profile. To suppress the spurious amplification, the adapter is designed as a “Y” shape to prevent the production of templates complementary to the universal primer. Because the 3′ end of the adapter-ligated fragment is blocked for the extension, no template complementary to the universal primer exists in the PCR mixture when no allele-specific extension occurs. If the 3′ end of the allele-specific primer is complementary to the target, an extension reaction occurs and the template complementary to the universal primer is produced. Therefore, only allele-specific amplification occurred in the current ALM-ASA system and there was no linear amplification by the universal primer. In consequence, nonspecific amplification was reduced by the structure of the adapter. To form the Y-shaped adapter for preventing the nonspecific extension, four uncomplementary bases were put in the 3′ end of the adapter (flap b in Fig. 1). Except to form a Y shape for the adapter, we can modify the oligonucleotide at the 3′ end of the adapter to block the nonspecific extension [26]. In the current study, we employed the Y-shaped adapter because it is inexpensive.

In ALM-ASA, the universal primer is adapter specific and is common for all SNPs. Because there are two choices for selecting the target DNA strand for the hybridization, two sets of genotyping primers could be designed. The sequences of the allele-specific primers were selected according to the lengths of the amplified fragments in this study that are listed in Table 1. A pair of allele-specific primers is complementary to the target sequence and was designed to differ by a 3′-terminal nucleotide that is specific to one of the alleles of an SNP. Therefore, the location of an allele-specific primer is fixed, and only the length of an allele-specific primer can be regulated to get an optimal melting temperature. In case allele discrimination by a single base in the 3′ end of an allele-specific primer is not enough to perfectly control the allele-specific extension, an artificial mismatch base is introduced in the 3′-terminal region to increase the allele specificity. For simplifying the primer design, we only put the same base as that in the target in the third position from the 3′ end of the primer [27].

To check the performance of the adapter and the specificity of allele-specific primers in ALM-ASA, each of the allele-specific amplifications was carried out separately using the universal primer and an allele-specific primer. For the rare alleles, “mock” templates were generated artificially by PCR as described previously by Roberts and coworkers [25]. All of the allele-specific amplifications gave a single band with the expected size. The length of the PCR products ranges from 62 to 1182 bp.

Optimization of the proportion of allele-specific primers

For typical multiplex PCRs, one of the problems is the impaired amplification efficiency of different targets. Even if the equimolar allele-specific primers were employed, amplification efficiency might change at a different locus due primarily to the different annealing efficiency of allele-specific primers [28].

The following simple protocol for optimizing the proportions of each allele-specific primer was proposed. First, perform a uniplex PCR amplification for a single locus using the universal primer and an allele-specific primer. Second, perform multiplex PCR using equimolar allele-specific primers of interest. Third, increase the amount of primers for the weak loci and decrease the amount of primers for the strong loci simultaneously. Fourth, change the position of the artificial mismatch base in the 3′ end of an allele-specific primer to adjust the extension efficiency if necessary. Finally, lengthen or shorten allele-specific primers to adjust the melting temperature if all of the above strategies are not effective. As an example, we used this protocol to optimize the primer concentration in a 7-plex PCR for typing SNP-1, -3, -7, -8, -9, -11, and -12. When pooling all 7 wild-type-specific primers and the universal primer together with the same concentration for multiplex amplification, an uneven strength of each band was observed. Based on the concentration of each amplified product in multiplex PCR, the concentration of allele-specific primers

were adjusted until uniform band strength was achieved. It is concluded that only an adjustment in the concentration of the related primers is possible to get a satisfactory result in many cases.

Optimization of PCR conditions for multiplex amplification

It is well known that the optimization of PCR conditions, including annealing temperature, extension time, the concentration of dNTP and Mg^{2+} , DNA polymerase, and additives, can result in improvement in the specificity or yield of PCR amplification. When a universal primer is used in ALM-ASA, only the melting temperature of the allele-specific primer needs to be optimized through the

selection of a suitable primer length. In this study, the annealing temperature varied from 50 to 62 °C, and the temperature of 55 °C was compromised for all allele-specific amplifications.

By trying various concentrations of dNTP and $MgCl_2$ for 17-plex PCR, the optimal concentrations of dNTP and $MgCl_2$ were determined as 0.2 and 1.5 mM, respectively, which are similar to the concentrations suggested in the protocol of a PCR reagent kit. We also tested various *Taq* DNA polymerases, including *Taq* DNA polymerase (GeneBase, Shanghai, China), *Taq* DNA polymerase (Promega, Madison, WI, USA), AccuPrime *Taq* DNA polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA), and Platinum *Taq* DNA polymerase (Invitrogen Life Technologies), for the

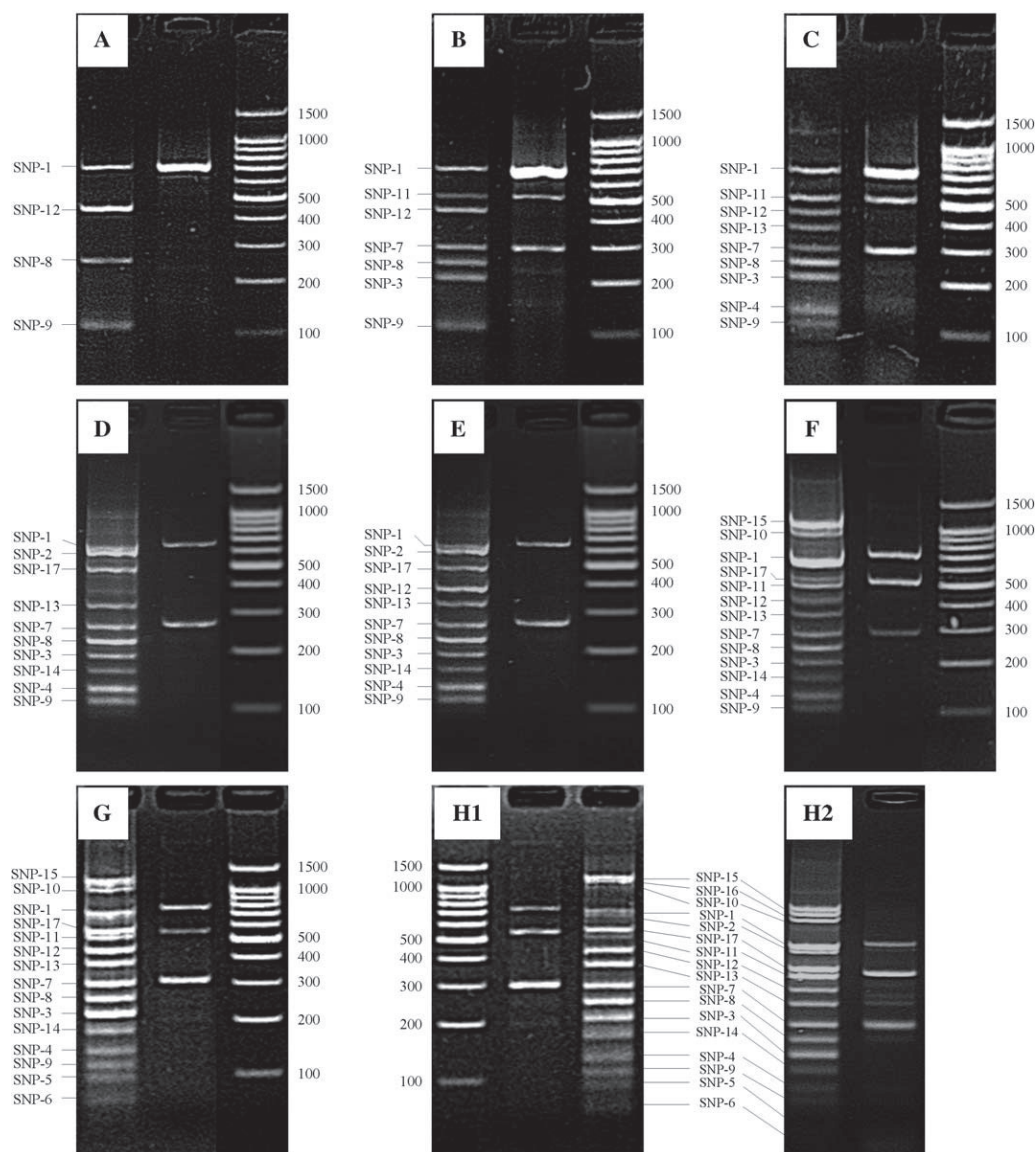


Fig. 2. Electrophoretogram of multiplex amplicons from 4-plex SNP typing (A), 7-plex SNP typing (B), 9-plex SNP typing (C), 10-plex SNP typing (D), 11-plex SNP typing (E), 13-plex SNP typing (F), 15-plex SNP typing (G) and 17-plex SNP typing (H1, H2). In each of panels A, B, C, D, E, F, and G, the left lane represents wild-types, the middle lane represents mutants, and the right lane represents the DNA marker. In panel H1, the left lane represents the DNA marker, the middle lane represents mutants, and the right lane represents wild-types. In panel H2, the left lane represents wild-types and the right lane represents mutants. The genotypes of the sample are wild-type homozygotes for SNP-2, -3, -4, -5, -6, -8, -9, -10, -12, -13, -14, -15, -16, and -17 and heterozygotes for SNP-1, -7, and -11.

ALM-ASA and found that all of the tested *Taq* DNA polymerases except for GeneBase *Taq* polymerase produced reliable and reproducible results under the same PCR conditions.

Multiplex allele-specific amplification

Multiplexing of allele-specific amplification was achieved by using the universal primer and allele-specific primers. Each allele-specific primer corresponded to a PCR product with a specific length that was used to identify the polymorphic locus. To type n SNPs by allele-specific amplifications, two separate n -plex PCR reactions were required to type wild-type and mutant, respectively. In either reaction, a mixture of primers corresponding to either wild-type or mutant together with the universal primer specific to the adapter was added.

Fig. 2 shows the gel electrophoretic profiles of allele-specific amplicons of 4-, 7-, 9-, 10-, 11-, 13-, 15-, and 17-plex PCRs from both wild-type-specific reaction and mutant-specific reaction. All of the bands are completely resolved over a range between 62 and 1182 bp. It is found that the band intensity for different targets is uniform except for several shorter fragments (Fig. 2H2). If some short bands are not recognizable, we can perform an additional short DNA-specific gel electrophoresis for achieving a good resolution of shorter fragments (Fig. 2H1).

Because the determination of an allele is based on the appearance of an allele-specific band in the electrophoretogram,

the readout is easy when the distance between two close bands is large enough (e.g., 4-plex SNP typing in Fig. 2). On the other hand, the readout becomes difficult when an electrophoretogram contains bands with similar lengths (e.g., 17-plex SNP typing in Fig. 2). Therefore, the multiplex level in ALM-ASA is limited by the capability and resolution of gel electrophoresis. To increase the multiplex level while keeping an accurate readout, a new strategy is proposed by assigning different allele-specific primer combinations to each allele-specific amplification reaction. For example, we can mix primers having high allele frequency (usually wild-type) with primers having low allele frequency (usually mutant) in each reaction instead of adding solely wild-type-specific primer or mutant-specific primer in one reaction. As shown in Fig. 3, the gel electrophoretograms of amplicons by 15- and 17-plex PCR become much simpler than those in Fig. 2.

Because microchip-based electrophoresis has the advantages of high speed and low consumption of DNA samples, the separation of allele-specific amplicons with different multiplex levels was also carried out on an Agilent BioAnalyzer 2100 with commercially available DNA 1000 assay sizing reagent kits. Microchip-based electrophoretograms of 4- and 7-plex PCR products are shown in Fig. 4. It is very easy to determine the genotype by comparing the peaks in the pair of electrophoretograms. All four wild-type-specific peaks were observed in electrophoretogram 1a in Fig. 4, and only one peak corresponding to the mutant of SNP-1 was observed in electrophoretogram 1b in Fig. 4, so the genotypes of SNP-1, -8, -9, and -12 were discriminated as CT, TT, GG, and AA, respectively. Based on the peaks in electrophoretograms 2a and 2b in Fig. 4, the genotypes of SNP-1, -3, -7, -8, -9, -11, and -12 were identified as CT, GG, GC, TT, GG, TC, and AA, respectively. These results coin-

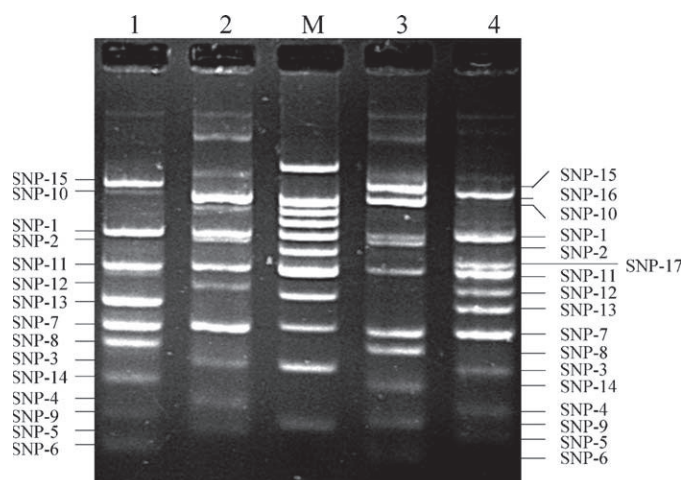


Fig. 3. Electrophoretograms of amplicons from 15-plex SNP typing (lanes 1 and 2) and 17-plex SNP typing (lanes 3 and 4) using allele-specific primer mixtures consisting of wild-type-specific primers with high allele frequency and mutant-specific primers with low allele frequency in either reaction. Lane 1: amplicons of 15-plex PCR using wild-type-specific primers of SNP-1, -6, -8, -9, -11, -13, -14, and -15 and mutant-specific primers of SNP-2, -3, -4, -5, -7, -10, and -12; lane 2: amplicons of 15-plex PCR using wild-type-specific primers of SNP-2, -3, -4, -5, -7, -10, and -12 and mutant-specific primers of SNP-1, -6, -8, -9, -11, -13, -14, and -15; lane 3: amplicons of 17-plex PCR using wild-type-specific primers of SNP-2, -6, -7, -8, -9, -10, -11, -14, and -15 and mutant-specific primers of SNP-1, -3, -4, -5, -12, -13, -16, and -17; lane 4: amplicons of 17-plex PCR using wild-type-specific primers of SNP-2, -6, -7, -8, -9, -10, -11, -14, and -15. The genotypes of the sample are the same as those in Fig. 2.

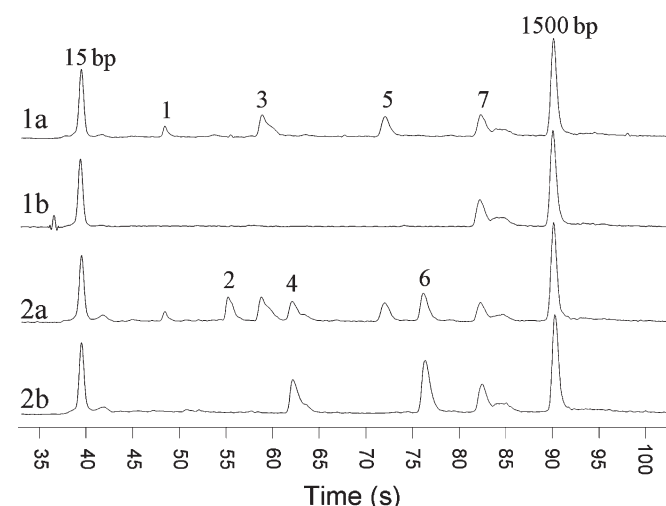


Fig. 4. Microchip electrophoretograms of amplicons from 4-plex SNP typing (1a, 1b) and 7-plex SNP typing (2a, 2b). Electrophoretograms 1a and 2a represent the wild-type-specific amplicons, and electrophoretograms 1b and 2b represent the mutant-specific amplicons. Peaks 1, 2, 3, 4, 5, 6, and 7 correspond to amplicons from SNP-9, -3, -8, -7, -12, -11, and -1, respectively. The lengths of the lower internal standard and the upper internal standard are 15 and 1500 bp, respectively.

Table 2
Allele frequencies (%) of 17 SNPs in the *CYP2D6* gene in a Chinese population

SNP code	Garcia-Barcelo et al. [29]	Garcia-Barcelo et al. [30]	Ji et al. [31]	Ji et al. [32]	Wang et al. [33]	In the current study
SNP-1	64.71	48	51.6	51.3	70.0	57.1
SNP-2	0	NA	NA	NA	NA	0
SNP-7	7.98	26	NA	2.0	NA	35.0
SNP-8	NA	NA	0	0	NA	0
SNP-9	0	NA	0.2	0.2	0.8	0.7
SNP-11	NA	NA	NA	NA	NA	13.6
SNP-12	NA	NA	0	0	NA	0

Note. NA, not available. SNP-2, -3, -4, -5, -6, -10, -13, -14, -15, -16, and -17 were not available, and the allele frequencies of them were zero in this study.

cide with those obtained by the PCR restriction fragment length polymorphism (PCR-RFLP) method.

Genotyping of different genome samples with ALM-ASA assay

To demonstrate the applicability of the ALM-ASA method in clinical diagnosis, DNA samples from 70 individuals were employed for the test. All 17 SNPs were typed in two reactions. As described previously, we can select allele-specific primers in either reaction to simplify electrophoretograms. However, for the large scale of genotyping on genomic samples with many rare alleles, it is much more convenient to analyze the results by two lanes: wild-type and mutant. If no band appears in a mutant lane, all of the SNPs are of the wild-type, and if some bands appear in the mutant lane, we only need to check whether or not there is another band existing in the corresponding position of the wild-type lane. A heterozygote sample should have two bands with the same size in both lanes. If the frequencies of most mutants are much lower, an electrophoretogram with a wild-type lane and a mutant lane can be rapidly interpreted without any difficulties. Frequencies of some alleles in a Chinese population are similar to those reported in Refs. [29–33], as shown in Table 2.

To examine the specificity and reliability of ALM-ASA in genotyping, the same samples were also genotyped by both Sanger's sequencing and PCR-RFLP. The typing results of the 17 SNPs by Sanger's sequencing and PCR-RFLP (see Table 1-s and Fig. 1-s in the supplementary material) for 20 samples accorded with those by ALM-ASA. In addition, the reproducibility of ALM-ASA was investigated by determining 17 SNPs in 10 different genomic samples in triplicate, and the reproducibility was as good as 100%.

Discussion

In this study, an improved method (ALM-ASA) with $n + 1$ primers for n -plex allele-specific PCRs in a single tube was developed for multiplex SNP genotyping. We found that ALM-ASA is very accurate and inexpensive for SNP typing. At the moment, preamplification was used to increase the sensitivity and specificity. If preamplification is not performed, the sensitivity becomes very low and the specificity of allele-specific amplification is

very poor, although a large amount of genomic DNA was used for SNP detection. In most cases, the sample amount is very limited and low genome consumption is preferable. Preamplification gave a significant decrease in the amount of genomic DNA used for SNP typing. However, preamplification limits the selection of SNPs on different genes because a preamplified PCR product has a size restriction. Usually one gene contains many SNPs, and it will not be a problem for ALM-ASA to type a large number of SNPs on one gene. If SNPs are on different genes, we can simply employ allele-specific PCR (amplification refractory mutation system [ARMS]) for singleplex SNP typing directly. Therefore ALM-ASA is very suitable for typing multiple SNPs in several disease-related genes.

Unlike uniplex PCR, a competitive amplification may occur in multiplex PCR when a single template is shared for multiple allele-specific amplifications. In the current study, we did not find a serious interaction between allele-specific primers complementary to a common template. By adjusting the concentration of allele-specific primers in PCR mixtures, all of the loci sharing a digested fragment as the template of PCR were successfully discriminated by the size of the amplified products.

Because the readout of ALM-ASA is based on a gel-based size separation of allele-specific PCR products on electrophoresis, a prerequisite for multiplex SNP typing is that the difference in length between each two SNP-specific amplicons should be larger than the resolution of gel electrophoresis. The length of an allele-specific amplicon is dependent on the distance between the location of an SNP and the cutting site proximal to the SNP. Because the location of an SNP is fixed, the only way to alter the length of an amplicon is to select a different cutting site created by a restriction enzyme. For an SNP, there are two proximal cutting sites (upstream and downstream), so there are two possible amplicons (usually with different lengths) specific to an SNP. If an allele-specific amplicon has a length similar to others, we can select the other one. For n -plex SNP typing, it is easy to pick out n allele-specific amplicons, each with appropriate size, from $2n$ candidates. If it is still impossible for all allele-specific primers to yield appropriately sized PCR products, another endonuclease can be employed because the location of a cutting site relies on the type of a restriction enzyme. Usually we can use a four-base class II restriction enzyme, *Mbo*I,

for a default because this enzyme recognizes the 5'-GATC-3' sequence that has a moderate frequency in the human genome. In this study, all of the 17 SNPs of interest can be discriminated by the length of allele-specific PCR products with the use of *Mbo*I for the digestion. As a general protocol, software may be beneficial for the selection of a suitable restriction enzyme for a random target sequence.

The switching performance of the 3' end in the allele-specific primer is perfect by coupling the introduction of an artificial mismatch base in the 3'-terminal region of an allele-specific primer with the adapter-specific universal primer. It is relatively flexible to customize the ALM-ASA system by selecting allele-specific oligonucleotides according to the SNPs of interest and by simply mixing various allele-specific primers with a universal primer to create a multiplexed assay. SNP typing by ALM-ASA does not need any special optimization step for PCR, and a default condition can be applied for any SNP typing. Also, it is not necessary to spend much time on primer design. The only parameter that should be optimized is the concentration of each allele-specific primer in the multiplex amplification reaction.

The time-consuming steps in ALM-ASA are preamplification of long fragments containing SNPs of interest, enzymatic digestion, and adapter ligation. However, all of these steps can be operated in parallel. It is an easy job for one person to type 17-plex SNPs in 96 genomic samples within the working hours of 1 day. The throughput can be facilitated by adapting the ALM-ASA system to 384-well plate, microplate array diagonal gel electrophoresis (MADGE) [34], and automated sample handling devices. The bottleneck of the typing throughput by the method is the capability of gel electrophoresis. When a 768- or 384-well microplate gel [35] is coupled with ALM-ASA for separating amplicons, the throughput will be increased significantly. Compared with chip-based platforms, such as the Affymetrix Gene Chip System [36] and the Illumina Golden Gate Assay [37], the throughput of ALM-ASA is much lower, but it is not necessary to use such expensive platforms for routine diagnostic tests in hospitals. Therefore, ALM-ASA is complementary to these high-throughput systems. Moreover, ALM-ASA can be extended to DNA diagnosis by simultaneously amplifying DNA fragments specific to different types of a microbe.

Acknowledgments

Financial support for this study was provided by the National Natural Science Foundation of China (30270368) and the High-Technology Program of Jiangsu Province in China (BG2003005).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ab.2006.04.022.

References

- [1] R. Sachidanandam, D. Weissman, S.C. Schmidt, J.M. Kakol, L.D. Stein, G. Marth, S. Sherry, J.C. Mullikin, B.J. Mortimore, D.L. Willey, S.E. Hunt, C.G. Cole, P.C. Coggill, C.M. Rice, Z. Ning, J. Rogers, D.R. Bentley, P.Y. Kwok, E.R. Mardis, R.T. Yeh, B. Schultz, L. Cook, R. Davenport, M. Dante, L. Fulton, L. Hillier, R.H. Waterston, J.D. McPherson, B. Gilman, S. Schaffner, W.J. Van Etten, D. Reich, J. Higgins, M.J. Daly, B. Blumenstiel, J. Baldwin, N. Stange-Thomann, M.C. Zody, L. Linton, E.S. Lander, D. Altshuler, A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms, *Nature* 409 (2001) 928–933.
- [2] Z. Zhao, E. Boerwinkle, Neighboring-nucleotide effects on single nucleotide polymorphisms: a study of 2.6 million polymorphisms across the human genome, *Genome Res.* 12 (2002) 1679–1686.
- [3] S. Bourgeois, D. Labuda, Dynamic allele-specific oligonucleotide hybridization on solid support, *Anal. Biochem.* 324 (2004) 309–311.
- [4] W.M. Howell, M. Jobs, U. Gyllenstein, A.J. Brookes, Dynamic allele-specific hybridization: a new method for scoring single nucleotide polymorphisms, *Nat. Biotechnol.* 17 (1999) 87–88.
- [5] S. Tyagi, F.R. Kramer, Molecular beacons: probes that fluoresce upon hybridization, *Nat. Biotechnol.* 14 (1996) 303–308.
- [6] C.R. Newton, A. Graham, L.E. Heptinstall, S.J. Powell, C. Summers, N. Kalsheker, J.C. Smith, A.F. Markham, Analysis of any point mutation in DNA: the amplification refractory mutation system (ARMS), *Nucleic Acids Res.* 17 (1989) 2503–2516.
- [7] G.H. Zhou, H. Shirakura, M. Kamahori, K. Okano, K. Nagai, H. Kambara, A gel-free SNP genotyping method: bioluminometric assay coupled with modified primer extension reactions (BAMPER) directly from double-stranded PCR products, *Hum. Mutat.* 24 (2004) 155–163.
- [8] X. Chen, K.J. Livak, P.Y. Kwok, A homogeneous, ligase-mediated DNA diagnostic test, *Genome Res.* 8 (1998) 549–556.
- [9] H. Baron, S. Fung, A. Aydin, S. Bahring, F.C. Luft, H. Schuster, Oligonucleotide ligation assay (OLA) for the diagnosis of familial hypercholesterolemia, *Nat. Biotechnol.* 14 (1996) 1279–1282.
- [10] S. Bakht, X. Qi, Ligation-mediated rolling-circle amplification-based approaches to single nucleotide polymorphism detection, *Expert Rev. Mol. Diagn.* 5 (2005) 111–116.
- [11] M. Ronaghi, M. Uhlen, P. Nyren, A sequencing method based on real-time pyrophosphate, *Science* 281 (1998) 363–365.
- [12] M. Ronaghi, Pyrosequencing sheds light on DNA sequencing, *Genome Res.* 11 (2001) 3–11.
- [13] V. Lyamichev, A.L. Mast, J.G. Hall, J.R. Prudent, M.W. Kaiser, T. Takova, R.W. Kwiatkowski, T.J. Sander, M. de Arruda, D.A. Arco, B.P. Neri, M.A. Brow, Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes, *Nat. Biotechnol.* 17 (1999) 292–296.
- [14] M. Olivier, The Invader assay for SNP genotyping, *Mutat. Res.* 573 (2005) 103–110.
- [15] A.C. Syvanen, Toward genome-wide SNP genotyping, *Nat. Genet.* 37 (Suppl.) (2005) S5–S10.
- [16] P. Hardenbol, J. Baner, M. Jain, M. Nilsson, E.A. Namsaraev, G.A. Karlin-Neumann, H. Fakhrai-Rad, M. Ronaghi, T.D. Willis, U. Landegren, R.W. Davis, Multiplexed genotyping with sequence-tagged molecular inversion probes, *Nat. Biotechnol.* 21 (2003) 673–678.
- [17] J.D. Taylor, D. Briley, Q. Nguyen, K. Long, M.A. Iannone, M.S. Li, F. Ye, A. Afshari, E. Lai, M. Wagner, J. Chen, M.P. Weiner, Flow cytometric platform for high-throughput single nucleotide polymorphism analysis, *BioTechniques* 30 (2001) 661–669.
- [18] I.G. Gut, DNA analysis by MALDI-TOF mass spectrometry, *Hum. Mutat.* 23 (2004) 437–441.
- [19] P. Ross, L. Hall, I. Smirnov, L. Haff, High level multiplex genotyping by MALDI-TOF mass spectrometry, *Nat. Biotechnol.* 16 (1998) 1347–1351.

- [20] E.M. Elnifro, A.M. Ashshi, R.J. Cooper, P.E. Klapper, Multiplex PCR: optimization and application in diagnostic virology, *Clin. Microbiol. Rev.* 13 (2000) 559–570.
- [21] D. Li, J. Vijg, Multiplex co-amplification of 24 retinoblastoma gene exons after pre-amplification by long-distance PCR, *Nucleic Acids Res.* 24 (1996) 538–539.
- [22] A.P. Shuber, V.J. Grondin, K.W. Klinger, A simplified procedure for developing multiplex PCRs, *Genome Res.* 5 (1995) 488–493.
- [23] N.E. Broude, L. Zhang, K. Woodward, D. Englert, C.R. Cantor, Multiplex allele-specific target amplification based on PCR suppression, *Proc. Natl. Acad. Sci. USA* 98 (2001) 206–211.
- [24] N.E. Broude, K. Driscoll, C.R. Cantor, High-level multiplex DNA amplification, *Antisense Nucleic Acid Drug Dev.* 11 (2001) 327–332.
- [25] R. Roberts, P. Joyce, M.A. Kennedy, Rapid and comprehensive determination of cytochrome P450 *CYP2D6* poor metabolizer genotypes by multiplex polymerase chain reaction, *Hum. Mutat.* 16 (2000) 77–85.
- [26] M.H. Shaperro, J. Zhang, A. Loraine, W. Liu, X. Di, G. Liu, K.W. Jones, MARA: a novel approach for highly multiplexed locus-specific SNP genotyping using high-density DNA oligonucleotide arrays, *Nucleic Acids Res.* 32 (2004) e181.
- [27] G. Zhou, M. Kamahori, K. Okano, G. Chuan, K. Harada, H. Kambara, Quantitative detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER), *Nucleic Acids Res.* 29 (2001) e93.
- [28] M.T. Suzuki, S.J. Giovannoni, Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR, *Appl. Environ. Microbiol.* 62 (1996) 625–630.
- [29] M. Garcia-Barcelo, L.Y. Chow, H.F. Chiu, Y.K. Wing, D.T. Lee, K.L. Lam, M.M. Waye, Genetic analysis of the *CYP2D6* locus in a Hong Kong Chinese population, *Clin. Chem.* 46 (2000) 18–23.
- [30] M. Garcia-Barcelo, L.Y. Chow, K.L. Lam, H.F. Chiu, Y.K. Wing, M.M. Waye, Occurrence of *CYP2D6* gene duplication in Hong Kong Chinese, *Clin. Chem.* 46 (2000) 1411–1413.
- [31] L. Ji, S. Pan, J. Marti-Jaun, E. Hanseler, K. Rentsch, M. Hersberger, Single-step assays to analyze *CYP2D6* gene polymorphisms in Asians: Allele frequencies and a novel *14B allele in mainland Chinese, *Clin. Chem.* 48 (2002) 983–988.
- [32] L. Ji, S. Pan, J. Wu, J. Marti-Jaun, M. Hersberger, Genetic polymorphisms of *CYP2D6* in Chinese mainland, *Chin. Med. J. (Engl.)* 115 (2002) 1780–1784.
- [33] S.L. Wang, M.D. Lai, J.D. Huang, *G169R* mutation diminishes the metabolic activity of *CYP2D6* in Chinese, *Drug Metab. Dispos.* 27 (1999) 385–388.
- [34] S.D. O'Dell, T.R. Gaunt, I.N. Day, SNP genotyping by combination of 192-well MADGE, ARMS, and computerized gel image analysis, *BioTechniques* 29 (2000) 500–506.
- [35] T.R. Gaunt, L.J. Hinks, H. Rassouljian, I.N. Day, Manual 768 or 384 well microplate gel “dry” electrophoresis for PCR checking and SNP genotyping, *Nucleic Acids Res.* 31 (2003) e48.
- [36] J.B. Fan, X. Chen, M.K. Halushka, A. Berno, X. Huang, T. Ryder, R.J. Lipshutz, D.J. Lockhart, A. Chakravarti, Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays, *Genome Res.* 10 (2000) 853–860.
- [37] K.L. Gunderson, F.J. Steemers, G. Lee, L.G. Mendoza, M.S. Chee, A genome-wide scalable SNP genotyping assay using microarray technology, *Nat. Genet.* 37 (2005) 549–554.