



A modified simple RFLP-PCR method for single nucleotide polymorphism (SNP) typing

Junhua Xiao¹, Xiujuan Xin¹, Xiaohui Luan¹, Dongzhi Wei¹, Shengli Yang²

¹State Key Laboratory of Bioreactor Engineering, Institute of Biochemistry, East China University of Science and Technology, Shanghai, P.R. China.

²Shanghai Bioengineering Research Center, Shanghai, China.

Abstract

We describe a modified single nucleotide polymorphism (SNP) typing method based on the restriction fragment length polymorphism polymerase chain reaction (RFLP-PCR). This is a simple, economical method without the need for special equipment. For most SNP loci, a common restriction endonuclease (*Hind* III, *Eco*R I or *Bam*H I) recognizing site (RER) can be introduced into one allelic form, but not the other by two rounds of mismatched PCR. The flanking regions can be changed by as many as five bases after PCR amplification with specially designed mismatching primers so the genotypes can be distinguished after digestion of the PCR products with corresponding endonucleases.

Key words: SNP genotyping, PCR-RFLP, RER site.

Received: May 14, 2005; Accepted: December 16, 2005.

With the completion of the Human Genome Project, about 5.3 million single nucleotide polymorphisms (SNPs) have been found in the human genome, which means that every 600 base pairs there exists one SNP (Patil *et al.*, 2001). Thus, by studying SNP typing information, researchers can investigate disease-associated genes and analyze the genetic structure of a population. Compared with many other methods, such as Taqman assay, the ligase detection reaction and D-HPLC (Kirk *et al.*, 2002), developed in recent years, RFLP-PCR is relatively simple and economical (Haliassos *et al.*, 1989) but the original RFLP-PCR method has several limitations, including the fact that the flanking regions of most SNP loci do not have appropriate restriction endonuclease recognizing site (RER) sites for typing. Some researchers have improved the method by changing one or two bases of the sequence adjacent to the SNP with mismatching primers to create a RER site for typing (Hiesh *et al.*, 2001), but the substitution of one or two bases is not competent enough to introduce a RER sequence, usually 4-6 bp long, to most of the anticipated regions. Therefore, the improved method is still not applicable for most SNP typing and sometimes the created RER sites can only be recognized

by rare restriction endonucleases that are expensive or with low digestion efficiency.

In this study we used specially designed mismatching PCR primers to introduce a common *Hind* III, *Eco*R I or *Bam*H I RER site into the flanking regions of the SNP loci by changing 4 to 5 bases after two rounds of PCR amplification for typing. The first round reaction system consisted of 1-5 ng of genomic DNA isolated from Chinese population (described below) peripheral leukocytes using phenol-chloroform methods (Joseph *et al.*, 2001), 200 μ M dNTPs mixture, 2.0 mM MgCl₂, 1.0 μ M of each of the forward and reverse primers and 0.5 u of *Taq* polymerase (TaKaRa) in a final volume of 5 μ L of reaction mixture. The first round PCR reaction was performed using an initial denaturation of 5 min at 94 °C followed by 14 denaturation cycles of 30 s at 94 °C, a first annealing step of 30 s at 63 °C in the first cycle decreased by 0.5 °C per cycle, a second annealing step of 10 s at 45 °C and extension for 50 s at 72 °C. This was followed by 30 cycles of 30 s at 94 °C, 5 s at 65 °C, 30 s at 56 °C, 10 s at 45 °C and 50 s at 72 °C, and a final extension of 7 min at 72 °C. The first round products were diluted 10 times and used as templates for the second round reactions in which 1 μ L of the diluted PCR products was added to a final reaction volume of 30 μ L containing 200 μ M dNTPs mixture, 1.5 mM MgCl₂, 3.0 μ M of each of the specially designed forward and reverse primers and 2 u of *Taq* polymerase (TaKaRa). The reaction program was

Send correspondence to Dongzhi Wei. State Key Laboratory of Bioreactor Engineering, Institute of Biochemistry, East China University of Science and Technology, 130 Meilong Road, 311 P.O. Box, 200237 Shanghai, China. E-mail: dzhwei@ecust.edu.cn. The first two authors contributed equally to the article.

set to an initial denaturation of 5 min at 94 °C followed by 45 cycles of 30 s at 94 °C, 40 s at 56 °C, 10 s at 45 °C and 50 s at 72 °C, and then a final extension of 5 min at 72 °C.

After two rounds of amplification, a common RER site was introduced into the product of one allele (usually the wild type) of the SNP (see Figure 1) and 2 µL of the PCR products was digested by the corresponding endonuclease before electrophoresis on 3% (w/v) agarose gel and the genotypes separated.

We used the methodology described above to type three CYP3A4 gene SNPs (CYP3A4*1B, CYP3A4*6 and rs#2242480) in a female population consisting of 196 apparently healthy Chinese volunteers who are freshman of East China University of Science and Technology (ECUST) come from everywhere of china, and they are with an age range of from 17 to 19 years. This study was approved by the Medical Ethics Committee of the ECUST and informed consent was obtained from each volunteer.

For genotyping, blood samples (5 mL) were taken from each individual and genomic DNA extracted by a standard method (Joseph *et al.*, 2001) and used in the reac-

tions as described above. The PCR reactions was carried out as the method described above, the primers used are shown in Table 1 with substituted bases in boldface type, the amplified products and their enzymatic digestion being shown in Figure 2. The CYP3A4 gene encodes a member of Cytochrome P450 family, which play important roles in the oxidative and reductive metabolism of a variety of endogenous and exogenous compounds and contains some SNPs which have been thought to be associated with drug metabolism and the development of puberty (Kitada *et al.*, 1987; Waxman *et al.*, 1988). With the method described above, three RER sites (*Hind* III, *Eco*R I and *Bam*H I) were introduced successfully for typing the three selected SNPs on *CYP3A4* gene.

In the Chinese population investigated by us the RFLP-PCR allele frequencies were zero for the CYP3A4*1B allele, 0.0052 for the CYP3A4*6 allele and 0.219 for the rs#2242480 allele, which are similar to the results obtained by Heish *et al.* (2001) for Taiwanese Chinese population (CYP3A4*1B = 0, CYP3A4*6 = 0.004902 and rs#2242480 = 0.248).

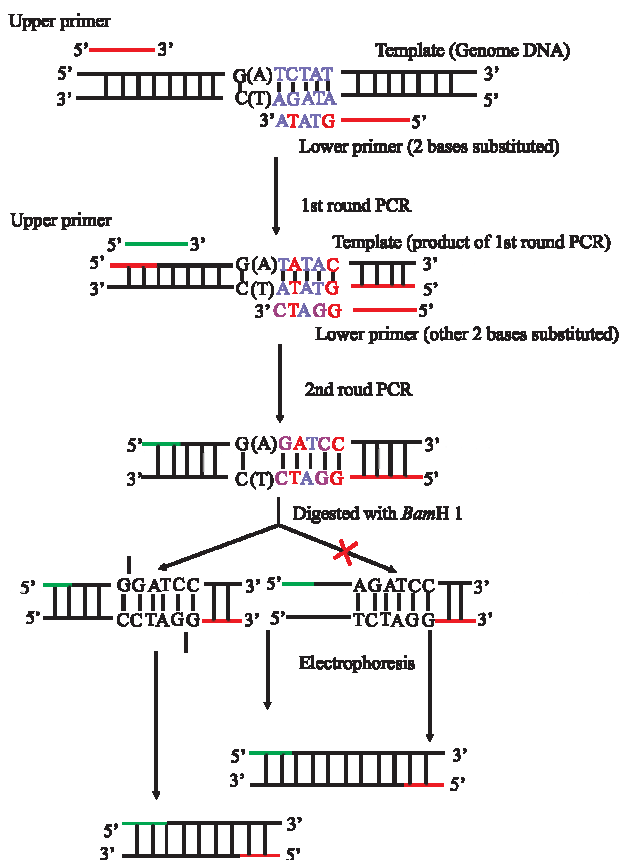


Figure 1 - Schematic of the method described in this article. The SNP used as an example is a G to A substitution in the CYP3A4 gene. After two rounds of amplification with mismatched primers the sequence around the SNP locus is changed and a RER site is introduced into the product of one genotype only. The 'G' genotypes with can be digested with *Bam*H I but the 'A' genotype cannot, so the two genotypes can be distinguished using agarose gel electrophoreses.

Table 1 - Primer Design of the three single nucleotide polymorphisms (SNPs).

| SNP and methodology | Primers, restriction enzymes, product and fragment lengths |
|--------------------------------|--|
| CYP3A4*1B | |
| Sequence around the SNP | ATAGAGACAAGGCCAA(G)GAGAGAGCGCATT |
| 1 st Round primer F | 5' ATGACCTAAGAAGTCACAAGAAAGTCAGAA 3' |
| 1 st Round primer R | 5' TGGCATAAAATCTATTAATTCGCCTCACAC 3' |
| 2 nd Round primer F | 5' GGAAAGACTCTAAGAGAACGCAGTCTGT 3' |
| 2 nd Round primer R | 5' TGGCATAAAATCTATTAATTCGCCTCAAGC 3' |
| Length of PCR product | 197 |
| Restriction endonuclease used | <i>Hind</i> III |
| Fragment length of alleles | G 197 bp A 166+31bp |
| CYP3A4*6 | |
| Sequence around the SNP | CTGATGATTGA(insertion in mutation type)CTCTCAGATT |
| 1 st Round primer F | 5' ACACATCTTCAAATGAAAGTCCCTATCAGG 3' |
| 1 st Round primer R | 5' TGGGACTCAGTTTCTTTTGAACCTCTGAAAG 3' |
| 2 nd Round primer F | 5' GACCACGCTGTGATTACTTCTGACTTCAGG 3' |
| 2 nd Round primer R | 5' TGG GACTCAGTTTCTTTTCAACTCTGGAAT 3' |
| Length of PCR product | 132 bp |
| Restriction endonuclease used | <i>Eco</i> R I |
| Fragments length of alleles | No insertion 132 bp A insertion 106+26 bp |
| rs# 2242480 | |
| Sequence around the SNP | AATAAGGTGAGTGGATGG(A)TACATGGAGAAG |
| 1 st Round primer F | 5' CACATCAGA ATGTAACGACCCCCAGTGTAC 3' |
| 1 st Round primer R | 5' GGTTCACCTCCTCCCTCTTCTCCGTATA 3' |
| 2 nd Round primer F | 5' CGTGGCCCAATCAATTATCTTTATTTTTCG 3' |
| 2 nd Round primer R | 5' GGTTCACCTCCTCCCTCTTCTCCGGATC 3' |
| Length of PCR product | 154 bp |
| Restriction endonuclease used | <i>Bam</i> H I |
| Fragments length of alleles | A 154 bp G 128+26bp |

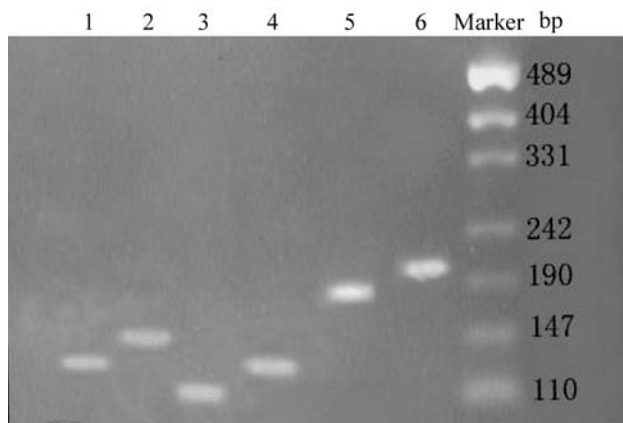


Figure 2 - Electrophoresis of the PCR products (after enzymatic digestion) on 3% (w/v) agarose gel. Lanes: 1 = rs# 2242480 Bam HI digestion product; 2 = undigested rs# 2242480; 3 = CYP3A4*6 EcoR I digestion product; lane 4 = undigested CYP3A4*6; 5 = CYP3A4*1B Hind III digestion product; 6 = undigested CYP3A4*1B.

Most researchers believed that primers with 3' mismatching nucleotide sequences would not function well in PCR reactions (Lawyer FC *et al.*, 1989). To investigate the effects of the primers on the PCR reaction, we synthesized three reversal primers complementary to the rs# 2242480 sequence and which had a different number of mismatching bases at their 3' ends, R1 with no 3' mismatched bases, R2 with two 3' mismatches and R3 with three mismatches (Table 2, substituted bases in boldface type). We performed PCR reactions using different thermocycling regimes (described above) and found that during the first 25 cycles the amplification efficiency decreased with the number of 3' mismatched bases increased but that after 35 cycles different primers rarely resulted in any differences and enough PCR products can be obtained for RFLP typing. The amplifying efficiency of primers with two mismatching bases at different positions at the 3' end (R4-R9, also show in table 2) was also studied and the results showed that there was little difference among the primers after 40 cycles.

Table 2 - Primers used to study the effects of primer design on amplification efficiency.

| Primer | Sequence |
|--------------------|---|
| 1 st R1 | 5' GGTTTCACCTCCTCCCTCCTTCTCCATGTA 3' |
| 1 st R2 | 5' GGTTTCACCTCCTCCCTCCTTCTCCATGCC 3' |
| 1 st R3 | 5' GGTTTCACCTCCTCCCTCCTTCTCCATACC 3' |
| 1 st R4 | 5' GGTTTCACCTCCTCCCTCCTTCTCCATATC 3' |
| 1 st R5 | 5' GGTTTCACCTCCTCCCTCCTTCTCC GT ATA 3' |
| 1 st R6 | 5' GGTTTCACCTCCTCCCTCCTTCTCC AG GTC 3' |
| 1 st R7 | 5' GGTTTCACCTCCTCCCTCCTTCTCC GT GTC 3' |
| 1 st R8 | 5' GGTTTCACCTCCTCCCTCCTTCTCC AG ATA 3' |
| 1 st R9 | 5' GGTTTCACCTCCTCCCTCCTTCTCC GG GTA 3' |
| 2 nd R | 5' GGTTTCACCTCCTCCCTCCTTCTCC GG ATC 3' |

Though the method is applicable for most sequences containing a SNP, in some instances the primers must be carefully designed. An anticipated *Hind*III recognition site can not be introduced to the sequence 5'- G₁A₂A₃A₄A₅G₆T₇G₈C₉C₁₀**A**₁₁(**C**)T₁₂C₁₃T₁₄C₁₅T₁₆A₁₇T₁₈A₁₉G₂₀C₂₁T₂₂G₂₃A₂₄G₂₅-3' (NCBI GenBank no. AF209389) in which one A(C) polymorphism is at the 11th position (indicated by boldface and italic type). However, if the underlined bases were designed to be changed to A(C)AGCTT, the nucleotide at the 11th or 12th position would be missed after PCR amplification. Further experiments showed that when two adjoining adenine (A) and thymine (T) bases exist in the sequence (*e.g.* ATCG) either the A or T base is missing from the PCR product if we try to substitute A with T or T with A to get a new sequence such as AACG or TTCG. We suppose that either of the two adjoining bases might have been skipped when the mismatching primer anneals with the template in the reaction. However, a similar phenomenon has not been found in the case of adjoining C's and G's.

Currently, SNP typing is becoming more and more popular in genetic research and some typing methods have been developed which need no special instrumentation, including allele specific PCR (AS-PCR) and mutagenically separated PCR (MS-PCR) (Underhill *et al.*, 1996; Rust, *et al.*, 1993) but unspecific amplification is inevitable when using these techniques and this introduces some difficulties into these SNP typing methods. This means that RFLP-PCR is a simple and reliable method for SNP typing and our methodology outlined in this paper has greatly expanded its application. For most SNPs, a common RER site can be easily introduced into the one allelic form of the SNP and the RFLP typing can be carried out with a cheap DNA sample and without any special instrumentation. In most instances, even five mismatching bases could be introduced into the DNA after two rounds of PCR (*i.e.* two in the first round and three in the second round) using this modified method. The result also proves that PCR amplification has only relative specificity, and implies that previous typing methods (such as MS-PCR and AS-PCR) based on specific amplification may lead to unreliable results if carried out without extensive optimization.

Acknowledgments

This research is supported by National Nature Science Foundation of China (30300198). The authors deeply appreciate Dr. Wenjun Guo for his kind help in revising the English of this paper.

References

- Kirk BW, Feinsod M, Favis R, Kliman RM and Barany F (2002) Single nucleotide polymorphism seeking long term association with complex disease. *Nucleic Acids Res* 30:3295-3311.
- Haliassos A, Chomel JC, Tesson L, Baudis M, Kruh J, Kaplan JC and Kitzis A (1989) Modification of enzymatically ampli-

- fied DNA for the detection of point mutations. *Nucleic Acids Res* 17:3606.
- Hiesh KP, Lin YY, Cheng CL, Lai ML, Lin MS, Siest JP and Huang JD (2001) Novel Mutations of CYP3A4 In Chinese. *Drug Metab Dispos* 29:268-273.
- Joseph S and David WR (2001) *Molecular Cloning: A Laboratory Manual*. 3rd edition. Cold spring Harbor Laboratory press, Cold Spring Harbor, 463-470 pp.
- Kitada M, Kamataki T, Itahashi K, Rikihisa T and Kanakubo Y (1987) Significance of cytochrome P-450 (P-450 HFLa) of human fetal livers in the steroid and drug oxidations. *Biochem Pharmacol* 36:453-456.
- Lawyer FC, Stoffel S, Saiki RK, Myambo K, Drummond R and Gelfand DH (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* 264:6427-6436.
- Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BTN, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SPA and Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 294:1719-1723.
- Underhill PA, Jin L, Zemans R, Oefner PJ and Cavalli-Sforza LL (1996) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93:196-200.
- Rust S, Funke H and Assmann G (1993) Mutagenically separated PCR (MS-PCR): A highly specific one step procedure for easy mutation detection. *Nucleic Acids Res* 21:3623-3629.
- Waxman DJ, Attisano C, Guengerich FP and Lapenson DP (1988) Human liver microsomal steroid metabolism: Identification of the major microsomal steroid hormone 6 beta-hydroxylase cytochrome P-450 enzyme. *Arch Biochem Biophys* 263:424-436.

Associate Editor: Carlos C.F. Menck