Single nucleotide polymorphism (SNP) detection by polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP)

Masao Ota[1*], Hirofumi Fukushima[1], Jerzy K. Kulski[2,3], Hidetoshi Inoko[3]

1) Department of Legal Medicine, Shinshu University School of Medicine, Nagano, 390-8621, Japan.
2) Centre for Comparative Genomics, School of Information Technology, Murdoch University, Murdoch, Western Australia 6150, Australia
3) Department of Molecular Life Science, Division of Molecular Medical Science and Molecular Medicine, Tokai University School of Medicine, Kanagawa, 259-1143, Japan

* Corresponding Author: Masao Ota: e-mail; otamasao@sch.md.shinshu-u.ac.jp, Tel:81-263-37-3217

ABSTRACT

The accurate analysis of DNA sequence variation in not only humans and animals but also other organisms has played a significant role in expanding our knowledge about genetic variety and diversity in a number of different biological areas. The search for an understanding of the causes of genetic variants and mutations has resulted in the development of a simple laboratory technique, known as the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method, for the detection of single nucleotide polymorphisms (SNPs). PCR-RFLP allows rapid

detection of point mutations after the genomic sequences are amplified by PCR. The mutation is discriminated by digestion with specific restriction endonucleases and is identified by gel electrophoresis after staining with ethidium bromide. This convenient and simple method is inexpensive and accurate for SNP genotyping and especially useful in small basic research studies of complex genetic diseases. The whole protocol takes only a day to carry out.

Introduction

The application and refinement of the PCR method since its invention in the 1980s[1-3] has dramatically progressed molecular genetic research. This simple and sensitive enzymatic technique for the amplification of DNA fragments has been used for various purposes, but especially for the detection of nucleic acid polymorphisms to find biological meaning in genetic variation and molecular divergence in living organisms. The majority of polymorphisms in the human genome are single nucleotide polymorphisms (SNPs) that account for more than 90% of sequence variation[4] and are utilized as an important tool for the study of the structure and history of our genome. There are more than 27 million SNPs in the human genome that have been recorded and stored in the SNPs database (http://www.ncbi.nlm.nih.gov/SNP). This immense amount of data provides valuable information for SNP-based studies[5], identification of candidate genes involved with complex genetic diseases[6-8], pharmacogenetic analysis[9-10], drug development[11], population genetics[12-13], evolutionary studies[14] and forensic investigations[15].

Over the past twenty years, many different methods have been developed for SNP genotyping by PCR, including hybridization[16], allele-specific PCR[17], primer extension[18],

oligonucleotide ligation[19], direct DNA sequencing[20] and endonuclease cleavage after amplification of the subjected genomic region by PCR[21-22]. Recent technologies for SNP genotyping (Taq Man method[23-24], Invader method[25-26], MALDI-TOF method[27], GeneChips[28]) have the potential for high-throughput, but they require the purchase of expensive equipment. Often, however, the throughput needs for SNP detection may be relatively low as this will depend mostly on the aim of the study.

One inexpensive, simple and convenient method for SNP genotyping is PCR-RFLP and its utility in molecular genetic studies has been demonstrated in a wide range of fields. Since 1988 (see ref. 21), more than 3700 reports related to PCR-RFLP have been cited in the PubMed bibliographic information database at NCBI (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi).

PCR-RFLP is based on endonuclease digestion of PCR-amplified DNA. The specific restriction endonuclease recognizes and cleaves the DNA in the region of the point mutation of the PCR products. The SNP type is easily identified by using gel electrophoresis to confirm and separate the sizes of the smaller DNA fragments generated by the endonuclease digestion. The PCR-RFLP is a simple, sensitive and reliable method and requires minimal investment in instrumentation. The limitations of this protocol are in the cases such that the sequences of target SNPs are not suitable for commercial restriction enzymes, and inversely, that the sequences have too many recognition sites for a single restriction enzyme. This method is available for genotyping not only a SNP, but also insertion/deletion polymorphisms[29-31] and multiple mutations such as Human Leukocyte Antigen (HLA) alleles[22,32-33] (**Figure 1**), which consist of multiple assemblies of SNPs. The ABO blood group polymorphism was the first molecular polymorphism to have been observed and analyzed in humans. This

antigenic variation for red blood cells is due to a few DNA single-base substitutions that produce the A, B, and O alleles that can be identified as the ABO genotypes by the PCR-RFLP method[34] (**Figure 2**).

Experimental Design

Choice of Restriction Enzymes and design of PCR Primers for PCR-RFLP.

PCR-RFLP always starts with the need to design an optimum primer pair and to find the restriction enzymes that will identify the SNPs in the PCR-amplified product. To obtain high quality results, it is important to design specific amplification primers and select an appropriate restriction enzyme from the REBASE Restriction Enzyme Database[35] (http://rebase.neb.com). To assist with this potentially difficult and tedious step, several comprehensive and automated search tools have been provided by web interfaces that are available at http://cedar.genetics.soton.ac.uk/public_html/primer (see ref. 36) or http://bioinfo.bsd.uchicago.edu/SNP_cutter.htm (see ref. 37). These programs are capable of designing primers for either natural PCR-RFLP or mismatch PCR-RFLP, depending on the sequence data of the targeted genomic region; natural PCR-RFLP can be used if discriminative restriction sites are found at the SNP site, otherwise, additional mismatched bases are introduced to adjacent sites in order to provide artificial restriction sites (known as mismatch PCR-RFLP); even if the target SNP has no cleavage site for available restriction enzymes, it is possible to use the mismatch PCR-RFLP method[38] (**Figure 3**). This method uses a primer containing an artificial restriction site (an additional mismatched base) adjacent to the SNP site. PCR primer sets should be designed to amplify a given DNA region but not to amplify an orthologous non-specific region or a paralogous (duplicated) region. Of course, it is

alternatively possible to manually find an appropriate restriction enzyme and to design the corresponding primers using either commercial software (e.g. Genetyx, Genetyx Co, Japan; DNASIS Pro, Hitachi Software Engineering Co. Ltd., Japan; Gene Construction Kit, Textco BioSoftware, Inc., USA; DNADynamo, Blue Tractor Software, Ltd., UK) or free software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (see ref. 39), http://www2.eur.nl/fgg/kgen/primer/SNP_Primers.html). To find the appropriate restriction sites and amplification primers for producing PCR products, a comprehensive software tool has been created for PCR-RFLP assay design[37, 40].

Template preparation for PCR-RFLP. Template DNA isolation methods for PCR-RFLP vary considerably in the starting material (e.g., fresh or frozen animal tissues and cells, blood, body fluids, buccal swab, paraffin-embedded tissue, formalin-fixed tissue, insects, yeasts or bacteria). The most commonly used method for genomic DNA preparation is based on phenol extraction and ethanol precipitation[41]. Alternatively, many DNA preparation kits are now commercially available and can be useful to obtain high quality and quantity template DNA. The purity of template DNA is determined by calculating the ratio of absorbance at 260 nm to absorbance at 280 nm. An A260/A280 ratio of 1.7-1.9 means pure template DNA and is better suited for PCR procedure.

Controls. Prepare at least one positive control for amplifying each of the alleles and one negative control with no DNA. Amplification for a positive control and a negative control is performed in parallel with sample in a different tube. Each gel should also include the controls for checking enzyme function. When we use single enzyme for single sample or multiple samples, we need to prepare

internal control DNA which has a restriction site with each enzyme to check the activity of each enzyme or slip of enzyme addition to a reaction tube. Meanwhile, using multiple enzymes for single sample or multiple samples, it is best to prepare same numbers of different internal control DNA as numbers of enzyme. Each internal control DNA is consisted of an appropriate length PCR products from plasmid DNA which has a restriction site by each enzyme.

MATERIALS

REAGENTS

*10xPCR buffer (usually supplied by the manufacturer with the Taq polymerase)

*Taq Polymerase (ABI: N8080160; N8880240, QIAGEN: 201203; 203203, Promega: M7122; M5661, TaKaRa: R001A, R007A, etc.; use hot-start PCR depending on PCR specificity)

*dNTP Mixture (dATP, dCTP, dGTP, dTTP: usually supplied by the manufacturer with the Taq polymerase, Promega: U1511)

*PCR primers (forward and reverse: 20 μM in TE buffer; primers are usually obtained from company (SIGMA Genosys, Invitrogen)).

*Sterile water (autoclaved)

*Template DNA (see Experimental Design for further information)

CRITICAL: 10ng to 50ng of good quality of template DNA, extracted from various samples as described in the Experimental Design can be used for PCR-RFLP.

*Restriction Enzymes (New England Biolabs, Promega, TaKaRa)

CRITICAL: Enzyme activity drops quickly if restriction enzymes are not kept at -20 °C. They should be kept on ice during handling and never be left on ice after use.

*Restriction Digestion Buffer (usually supplied by the manufacturer with the restriction enzymes)

*Agarose (NuSieve 3:1 agarose) (Lonza, 50091)

CRITICAL: Usually discriminates DNA fragments less than 1,000bp. Within a given range of differently sized DNA fragments, the smaller DNA less than 1,000 bp are best separated in TBE buffer.

*Tris base (Sigma, T6066)

*Boric acid (Sigma, B7901)

*0.5M EDTA, pH8.0 (Sigma, E7889)

*10% (wt/vol) AP (ammonium persulfate; Sigma, A3678) prepared freshly each time.

*Acrylamide (Sigma; A3663, Wako;011-08015)

Caution: Acrylamide is highly neurotoxic and readily absorbed through the skin. Mask and gloves should be worn when handling this chemical.

*N,N'-Methylene bisacrylamide (Sigma;M7279, Wako;138-0632)

Caution: Bisacrylamide is highly neurotoxic and readily absorbed through the skin. Mask and gloves should be worn when handling this chemical.

*TEMED (N,N,N',N'-tetramethylethylenediamine; Sigma; 87689) stored at 4 ºC.

*Ethidium bromide (0.5 μg/ml) (Sigma, E8751)**.** Stock solution is prepared as 10 mg/ml in water and stocked in light-resistant bottle at room temperature.

CAUTION: Ethidium bromide is a mutagen with potential carcinogenicity; should wear gloves and take care in handling, storage, and disposal.

*Bromophenol Blue (BPB: Sigma, B0126)

*Xylene cyanol FF (XC: Sigma, X4126)

*Glycerol (Sigma, G5516)

*An optimal DNA size marker used in analysis for PCR products.

EQUIPMENT

*PCR tubes, 96-well multiwell plates

*PCR Thermal Cycler

*Pipettors and autoclaved tips (filter tips are better for avoiding cross-contamination)

* Horizontal minigel electrophoresis apparatus (e.g. Mupid, ADVANCE Co. Ltd., Tokyo, or Scie-Plas Mini Horizontal Gel Unit, Harvard Apparatus Co. Massachusetts).

*Vertical polyacrylamide gel electrophoresis apparatus (e.g. Atto Co. Tokyo, Bio-Rad Lab. Inc. California).

*Dry block heater

REAGENT SETUP

***TE buffer, pH8.0 (Tris/EDTA):** 10mM Tris-Cl, 1mM EDTA

***10 x Stock solution electrophoresis buffer (TBE), 1 liter:** 108 g Tris base, 55 g Boric acid, 40 ml 0.5M EDTA, pH8.0

CRITICAL: Store at room temperature and remake if major residues appear in solution.

***Gel-loading dye:** 0.25% Bromophenol Blue (BPB: Sigma), 0.25% Xylene cyanol FF (XC: Sigma), 50% glycerol (Sigma), 1mM EDTA (pH8.0).

***Acrylamide-bisacrylamide stock solution (29:1):** prepared as an unpolymerized 30% solution (wt/vol) in H_2O as a stock solution, as follows: 29g Acrylamide, 1g N,N'-Methylenbisacrylamide, water to 100 ml.

CRITICAL: Stock solution is protected from light and stored in the refrigerator.

Caution: Acrylamide and bisacrylamide are highly neurotoxic and readily absorbed

through the skin. Mask and gloves should be worn when handling these chemicals.

***Polyacrylamide gels**: the stock acrylamide-bis solution is mixed with water and 10 x TBE to yield the appropriate percentage, volume, and buffer concentration (1 x) (e.g. for 100 ml of 12% polyacrylamide gel; 40ml of 30 %acrylamide-bis solution + 10 ml of 10 x TBE + 49 ml of water + 1 ml of 10% AP). This solution is gently deaerated by swirling to make a well polymerized gel. Then, 60 µl of TEMED is added. After mixing well for a moment, pour polyacrylamide gel solution into the glass plate of the electorphoresis apparatus.

Procedure

PCR-amplification of DNA samples • TIMING 1.5 - 2.5 h

1. PCR reactions can be set up to amplify a single DNA sample for digestion with a single restriction enzyme (see option A), or to amplify more than one DNA sample for digestion with multiple restriction enzymes (option B):

A. PCR set up to amplify a single DNA sample for digestion with a single restriction enzyme.

i) To PCR-amplify a single DNA sample, set up a 20µl reaction mixture in a 0.2 ml micro amplification tube, as follows:

component	volume (µl)	final concentration
10 x PCR buffer ($Mg2^{+}$)	2.0	1 xPCR buffer
dNTP mix (each 2.5 mM)	1.6	0.2mM of each dNTP
Primer (forward) (100pmol/µl)	0.2	1µM forward primer
Primer (reverse) (100pmol/µl)	0.2	1µM reverse primer
Taq polymerase (5U/µl)	0.2	1 unit Taq polymerase
Sample DNA (10-50 ng/µl)	1.0	10-50 ng
DNase-free water	14.8	
total	20.0	

B. PCR set up to amplify more than one DNA sample for digestion with multiple

restriction enzymes.

i) To PCR-amplify more than one DNA sample, prepare a master reaction mixture in a 1.5ml tube, as follows, adding the template DNA individually to PCR tubes in advance:

component	n=10* volume (μl)
10 x PCR buffer ($Mg2^{+}$)	22.0
dNTP mix (each 2.5 mM)	17.6
Primer (forward) (100pmol/μl)	2.2
Primer (reverse) (100pmol/μl)	2.2
Taq polymerase (5U/μl)	2.2
Sample DNA (10-50 ng/μl)	11.0
DNase-free water	162.8
total	220.0

dispense DNA sample in advance

*: add extra 10% more samples, finally make up 11 samples

CRITICAL STEP: Calculate the total volume of the master reaction mixture as the volume required for one DNA sample multiplied by the number of DNA samples plus an extra 10 % to account for loss of liquid during pipetting steps.

CRITICAL STEP: If the PCR products are to be digested by more than one restriction enzyme, increase the total volume proportional to the number of enzymes that will be used.

2. Perform PCR-amplification of template DNA using an automated thermal cycler, with a standard amplification program, as follows, and stop PCR reactions by chilling at 4°C:

	Standard PCR		Hot start PCR	
Initial activating step	5 min	95°C	15min	95°C
3-step cycling				
Deanaturation	0.5-1 min	95°C	0.5-1 min	95°C
Annealing	0.5-1 min	55-65 °C		50-70 °C
Extension	1 min	72°C		72°C
Number of cycles	25-30		30-40	
Final extension	5 min	72°C	10min	72°C

CRITICAL STEP: The program is generally designed according to manufacturer's instructions of *Taq* polymerase and primers used in PCR. Temperatures and cycling times are optimized for individual templates and primer pairs.

Analysis of PCR Products by Gel Electrophoresis • TIMING 0.5 - 1 h

3. To separate and determine the PCR product sizes, the PCR yield and nonspecific background, load 6 μl (5 μl of each PCR reaction product plus 1μl of gel-loading dye) onto a 2% (w/v) agarose gel. Load an optimal DNA size marker (e.g. 100 bp ladder, pBR322-HaeIII, ϕx174-Hae III, or ϕx174-Hinf I) in a separate lane that is parallel with the samples on the gel.

4. Run the gel for appropriate times at constant voltage.

CRITICAL STEP: Running time depends on the expected size of the amplified product and concentration of agarose gel. Usually electrophoresis is performed 20 min at 100 V in 2 % agarose when the size of PCR fragment is ranging from 100 to 500 bp.

5. Stain the gel with ethidium bromide at 0.5 μg/ml for 10 minutes.

CRITICAL STEP: Alternatively, Ethidium bromide may be included in the gel and running buffer at 0.5 μg/ml.

6. Visualize PCR products using a UV-transilluminator and record the results by photography. Check whether the amplified product with the expected size is observed on the gel by comparing with an optimal DNA size marker. Amplicons of the correct size is preferable to be greater than 25 ng per band for use in the next digestion step. (An approximate mass of PCR products is figured comparably from the intensity of the bands with the size marker used, e.g. 100 bp ladder NEB; N3231L)

TROUBLESHOOTING

CAUTION: Safety glasses and preferably a face mask should be worn around UV light sources.

Restriction Enzyme Digestion of PCR products • TIMING 2 - 3 h

7. The general method for digestion of PCR-amplified DNA is the addition of a restriction enzyme directly to a reaction tube containing an aliquot of the PCR product. Restriction enzyme digests can be set up to digest PCR products with a single enzyme (option A) or to digest PCR products with multiple enzymes (option B):

A. Restriction enzyme digest set up to digest PCR products with a single restriction enzyme.

i) Prepare a master-mix solution to digest a number of PCR products with a selected restriction enzyme, as follows; each sample is reacted in a total of 10 μl of mixture containing 5-10 units of enzyme, 1x restriction digestion buffer and 6 μl of PCR product:

	n=1	n=50*
	μl	μl
10 x enzyme buffer	1.0	55.0
restriction enzyme (10U /ml)	0.5	27.5
internal control DNA$ (20ng/ml)	0.5	27.5
DNase-free water	2.0	110.0
total	4.0	220.0

n: sample number
*: add extra 10% more samples, finally make up 55 samples
$: purified PCR products after amplification of plasmid DNA which has a cleavage site with subjected restriction enzyme.

ii) Transfer 5 μl of PCR products to a PCR tube or a 96-well multiwell plate.

iii) Dispense 5μl of digestion master-mix to each well containing the PCR product.

B. **Restriction enzyme digest set up to digest PCR products with multiple restriction enzymes.**

i) Prepare a master-mix solution to digest PCR products with multiple restriction enzymes (e.g. typing HLA alleles; **Figure 1**), as follows:

	n=1	n=10*
	μl	μl
10 x enzyme buffer	1.0	11.0
PCR product	6.0	66.0
DNase-free water	2.0	22.0
total	9.0	99.0

n: number of enzyme for analysis
*: add extra 10% more samples, finally make up 11 samples

CRITICAL STEP: A 10 μl volume of mastermix should be made for each restriction enzyme and sample combination. For example, if 9 restriction enzymes are employed to determine HLA alleles, then we would prepare a tenfold quantity of reaction mixture based on the total volume (10 μl), for each sample

ii) Transfer 9 μl of the master-mix containing the PCR product to a new PCR tube or a 96-well multiwell plate.

CRITICAL STEP: The volume of the master-mix may vary according to an activity of restriction enzyme; units/μl). One unit of restriction enzyme will completely digest 1 μg of substrate DNA in a 50 μl reaction in 60 minutes by definition. Enzyme activity is usually dictated in the statement of commercially available restriction enzymes.

iii) Add 1 μl of restriction enzyme (5-10 units/μl) to the master-mix and mix the reaction solution gently.

CRITICAL STEP: An important factor to achieve complete digestion with a restriction enzyme is gentle mixing (pipetting the reaction mixture up and down or tapping the reaction tube gently).

CRITICAL STEP: Enzymes should be stored at -20°C. While in use, enzymes should be carefully maintained on ice.

8. Incubate reaction mixture (restriction enzyme and PCR product) on a PCR Thermal Cycler or a Dry block heater at the optimal temperature (depending on restriction enzyme) for appropriate times (2 hrs to overnight).

CRITICAL STEP: Generally 1 U of restriction enzyme digests 1 μg of DNA in 1 hour. If an excess of enzyme is added to the reaction, the length of incubation time may be decreased to save time. Conversely, a longer incubation time might be expected to conserve on the amount of an expensive enzyme by using fewer units for the reaction.

Analysis of PCR-RFLP products by Gel Electrophoresis• TIMING 1 - 1.5 h

9. Load 6 μl (5 μl of each sample plus 1μl of gel-loading dye) of the amplified DNAs digested by restriction enzymes (from step 8) onto either a 6%-12% polyacrylamide gel or a 4% agarose (NuSieve 3:1) gel. Also load an optimal DNA size marker (e.g. 100 bp ladder, pBR322-HaeIII, ϕx174-Hae III, or ϕx174-Hinf I) in a separate lane and in parallel with samples on the gel.

CRITICAL STEP: Gel choice depends on user’s choice. Polyacrylamide gel is less expensive than agarose gel, however requires more time and is more complicated to make up.

10. Run the gel for appropriate times at constant voltage.

Running conditions depend on the respective gel. When electrophoresis is performed in

horizontal minigel electrophoresis apparatus (e.g. Mupid, ADVANCE Co. Ltd., Tokyo), experimental conditions are 1 hr at 100 V for 12% polyacrylamide gel and 45 minutes at 100 V for 4% agarose gel.

11. Stain the gel with ethidium bromide at 0.5 μg/ml for 10 minutes.

12. Visualize products using a UV-transilluminator and a record the results by photography. Check the cleaved fragments or undigested PCR product on the gel by comparing with an optimal DNA size marker.

TIMING

Preparation DNA samples: 1- 24 h (depend on DNA extraction methods and samples)

PCR amplification of DNA samples (steps1-2): 1.5-2.5 h

Analysis of PCR products by gel electrophoresis (steps 3-6): 0.5-1 h

Restriction enzyme digestion of PCR products (steps7-8): 2 - 3 h

Analysis of PCR-RFLP products by gel electrophoresis (steps 9-12): 1-1.5 h

?Troubleshooting

If RFLP fragments are not observed as expected, then the unexpected outcome might be due to many different reasons, such as e.g. (1) the digestion of PCR products is inefficient, (2) the efficiency of the restriction enzyme reaction is poor or inactive, (3) some restriction enzymes are inhibited by methylation[35, 42] of nucleotides within their recognition sequences, (4) star activity[43] of some restriction enzymes, which cleave similar but not identical sequences at other sites besides those containing the defined recognition site. This altered or relaxed specificity is termed “star” activity. The star activity is induced under certain extreme conditions (including high endonuclease

concentrations, high glycerol concentrations, low ionic strength, and high pH, and so on). The following restriction endonucleases are reported to exhibit star activity, for example, BamHI[44], EcoRI[45], MboII[46], and SalI[47].

When PCR products are incompletely digested or not digested at all, purification of the PCR products with a commercial DNA extraction kit is sometimes useful for improving or recovering the efficiency of the restriction enzyme activity. Increasing the number of enzyme units added to the reaction mixture may also overcome the decreased reaction efficiency associated with impure DNA preparation.

When a reaction of restriction enzyme appears to be incomplete, confirm the inefficiency of the reaction by adding an internal control DNA which has cleavage site with same restriction enzyme into the same tube.

Troubleshooting advice can be found in Table 1.

Table 1. Troubleshooting information.

Problems	Possible Causes	Solutions
Step 6 Presence of extra amplification products	Incorrect thermal cycling program was used.	Repeat amplification and ensure the cycling program. Change annealing temperature. Change extension temperature. Change cycling numbers.
	Amplification mixture was not optimized.	Ensure that primer concentration is optimized. Decrease volume of the template DNA. Change Mg^{2+} concentration.
	Template DNA or Work area was contaminated.	Eliminate all sources of contamination.
Step 6 Absence of PCR products	Taq polymerase was not added to the amplification mix.	Repeat amplification paying attention to the addition and mixing of Taq with the amplification mix.
	Taq polymerase was inactivated.	Use new Taq or another Taq from different company.
	Amplification mixture was not optimized.	Check whether all reagents were optimal, especially template DNA concentration was optimal.
	Thermal cycler failure	Check the cycler run history and ensure workings of the thermal cycler.
Step 12 Unexpected fragments	Digestion was incomplete	Use new enzyme or another enzyme from different company. Increase volume of enzyme. Dilute PCR product. Perform digestion with longer incubation time. Check reaction temperature. remove PCR additives such as DMSO or glycerol.
	PCR products were contaminated	Purify PCR products with the DNA extraction kit.

ANTICIPATED RESULTS

PCR-RFLP is a relatively simple and inexpensive method of genotyping mainly single nucleotide polymorphisms (SNPs). This method enables not only the discrimination of homozygous and heterozygous samples at the target point mutation, but also the genotyping of multiple mutations such as Human Leukocyte Antigen (HLA) alleles[22,32-33] (**Figure 1**). This method is especially convenient to confirm nucleotide sequences at particular sites of interest without complex typing methods such as hybridization, nucleotide sequencing DNA chips etc.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

References

1. Saiki, R.K., et al. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**, 1350-1354 (1985).
2. Saiki, R.K., Bugawan T.L., Hron, G.T., Mullis, K.B. & Erlich, H.A. Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. Nature **324**, 163-166 (1986).
3. Mullis, K.B. & Faloona, F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. **155**, 335-350 (1987).
4. Collins, F.S., Brooks, L.D. & Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. Genome Res. **8**, 1229-1231, (1998).
5. The International HapMap Consortium. The international HapMap

project. Nature **426**, 789-796 (2003).

6. Martin, E.R., et al. SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. Am. J. Hum. Genet. **67**, 383-394, (2000).
7. de Bakker P.I.W., et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat. Gent. **38**, 1166-1172, (2006).
8. Pearson, J.V., et al. Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism associated studies. Am. J. Hum. Genet. **80**, 126-139, (2007).
9. Goldstein, D.B., Ahmadi, K.R., Weale, M.E. & Wood, N.W. Genome Scans and candidate gene approaches in the study of common diseases and variable drug responses. Trends Genet. 19, 615-622 (2003)
10. Ahmadi, K.R., et al. A single-nucleotide polymorphism tagging set for human drug metabolism and transport. Nat. Genet. 37, 84-89 (2005)
11. Delrieu, O. & Bowman,C. Visualizing gene determinants of disease in drug discovery. Pharmacogenomics, **7**, 311-329 (2006).
12. Syvanen, A.C. Accessing genetic variation: genotyping single nucleotide polymorphisms. Nat. Rev. Genet., **2**, 930-942 (2001).
13. Shriver, M.D., et al. Large-scale SNP analysis reveals clustered and continuous patterns of human genetic variation. Hum. Genomics, **2**, 81-89 (2005).
14. Hacia, J.G., et al. Determination of ancestral alleles for human

single-nuclotide polymorphisms using high-density oligonucleotide arrays. Nat. Genet. **22**, 164-167 (1999).

15. Kidd, K.K., et al. Developing a SNP panel for forensic identification of individuals. Forensic Sci. Int. **164**, 20-32 (2006).
16. Iwasaki, H., et al. Accuracy of genotyping for single nucleotide polymorphisms by a microarray-based single nucleotide polymorphim typing method involving hybridization of short allele-specific oligonucleotides. DNA Res. **9**, 59-62 (2002).
17. Papp, AC., Pinsonneault, JK. Cooke, G. & Sadee W. Single nucleotide polymorphism genotyping using allele-specific PCR and fluorescence melting curves. Biotechniques **34**, 1068-1072 (2003).
18. O'Meara, D., Ahmadian, A., Odeberg, J. & Lundeberg, J. SNP typing by apyrase-mediated allele-specific primer extension on DNA microarrays. Nucleic Acids Res. **30**, e75 (2002).
19. Pickering, J. Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms. Nucleic Acids Res. **30**, e60 (2002).
20. Chatterjee, P.D. Direct sequencing of bacterial and P1 artificial chromosome-nested deletions for identifying position-specific single-nucleotide polymorphisms. Proc. Natl. Acad. Sci. U.S.A **.96**, 13276-13281 (1999).
21. Deng, G.R. A sensitive non-radioactive PCR-RFLP analysis for detecting point mutations at 12th codon of oncogene c-Ha-ras in DNAs of gastric cancer. Nucleic Acids Res. **16**, 6231 (1988).

22. Ota, M., et al. Modified PCR-RFLP method for HLA-DPB1 and –DQA1 genotyping. Tissue Antigens 38, 60-71 (1991).
23. Livak, K.J., Flood S.J., Marmaro, J., Giusti, W. & Deetz, K. Ologonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl. **4**, 357-362 (1995).
24. Livak, K.J. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet. Anal.**14**, 143-149 (1999).
25. Kwiatkowski, R.W., Lyamichev, V., de Arruda, M. & Neri, B. Clinical, genetic, and pharmacogenetic applications of the invader assay. Mol. Diag. **4**, 353-364 (1999).
26. Fors, L, Lieder, K.W., Vavra, S.H. & Kwiatkowski, R.W. Large-scale SNP scoring from unamplified genomic DNA. Pharmacogenomics **1**, 219-229 (2000).
27. Haff, L.A. & Smirnov, I.P. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. Genome Res. **7**, 378-388 (1997).
28. Gunderson, K.L., Steemers, F.L., Lee, G., Mendoza, L.G. & Chee, M. A genome-wide scalable SNP genotyping assay using microarray technology. Nat. Genet. **37**, 549-554 (2005).
29. Abbud, Z.A., Wilson, A.C., Cosgrove, N.M. & Kostis, J.B. Angiotensin-converting enzyme gene polymorphism in systemic hypertension. Am. J. Cardiol. **81**, 244-246 (1998).
30. Simsek, M., Al-Wardy, N., Al-Khayat, A., Al-Khabory, M. A PCR-RFLP

test for simultaneous detection of two single-nucleotide insertions in the Connexin-26 gene promoter. Genet. Test. **6**, 225-228 (2002).

31. Bu, H., Rosdahl, I., Sun, Z.-F. & Zhang, H. Importance of polymorphisms in NF-κBI and NF-κBIα genes for melanoma risk, clinicopathological features and tumor progression in Swedish melanoma patients. J. Cancer Res. Clin. Oncol. DOI 10.1007/s00432-007-0228-7 (2007).
32. Ota, M., Seki, T., Fukushima, H., Tsuji, K. & Inoko, H. HLA-DRB1 genotyping by modified PCR-RFLP method combined with group-specific primers. Tissue Antigens **39**, 187-202 (1992).
33. Inoko, H. & Ota, M. PCR-RFLP. In Handbook of HLA typing techniques. (eds. Hui, K.M., & Bidwell, J.L.) 9-70 (CRC Press, Boca Raton, Ann Arbor, London, Tokyo, 1993).
34. Lee, J.C.-I. & Chang, J.-G. ABO genotyping by polymerase chain reaction. J. Forensic Sci. **37**, 1269-1275 (1992).
35. Roberts, R.J., Vincze, T., Posfai, J. & Macelis, D. REBASE-enzymes and genes for DNA restriction and modification. Nucleic Acids Res. **35**, D269-D270 (2007).
36. Ke, X., Collins, A. & Ye, S. PCR designer for restriction analysis of various types of sequence mutation. Bioinformatics **18**, 1688-1689 (2002).
37. Zhang, R., et al. SNP cutter: a comprehensive tool for SNP PCR-RFLP assay design. Nucleic Acids Res. **33**, W489-W492, (2005).
38. Love-Gregory, L.D., Dyer, J.A., Grasela, J., Hillman, R.E. & Phillips, C.L. Carrier detection and rapid newborn diagnostic test for the common Y393N maple syrup urine disease allele by PCR-RFLP: Culturally

permissible testing in the Mennonite community. J. Inherit. Metab. Dis. **24**, 393-403, (2001).

39. van Baren, M.J. & Heutink, P. The PCR suite. Bioinformatics **20**, 591-593 (2004).
40. Gardner, S.N. & Wagner, M.C. Software for optimization of SNP and PCR-RFLP genotyping to discriminate many genomes with the fewest assays. BMC Genomics **6**, 73 (2005).
41. Moore, D.D. Purification and concentration of DNA from aqueous solutions. In Current Protocols in Molecular Biology. (eds. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.) 2.1.1-2.1.9 (John Wiley & Sons, Inc., USA, 1994)
42. Roberts, R.J. Vincze, T., Posfai, J. & Macelis, D. REBASE-restriction enzymes and DNA methyltransferases. Nucleic Acids Res. **35**, D230-D232 (2005).
43. Polisky, B. et al. Specificity of substrate recognition by the EcoRI restriction endonuclease. Proc. Nat. Acad. Sci. USA. **72**, 3310-3314 (1975).
44. George, J., Chirikjian, J.G., Sequence-specific endonuclease BamHI: Relaxation of sequence recognition, Proc. Natl. Acad. Sci. USA **79**, 2432-2436 (1982).
45. Woodbury, C.P., et al., DNA site recognition and reduced specificity of the EcoRI endonuclease, J. Biol. Chem. **255**, 11534-11546 (1980).
46. Sektas, M., Kaczorowski, T. & Podhajska, A.J. Purification and properties of the MboII, a class-IIS restriction endonuclease. Nucleic Acids Res. **20**, 433-438 (1992).

47. Conlan, L.H., Jose, T.J. Thornton, K.C. & Dupureur, C.M. Modulating restriction endonuclease activities and specificities using neutral detergents. Biotechniques **27**, 955-960 (1990).

Figure 1

PCR-RFLP typing of HLA-DRB1 (see Ref. 32)

A. HLA alleles are determined with the combined patterns of restriction enzymes, which have a single recognition site in some alleles but none in the amplified regions of other alleles. Assignment of alleles is based on the combination of band patterns produced from multiple restriction enzymes[32,33]. The gel picture shows the typical restriction pattern obtained in the case of the HLA-DRB1*0803 allele. The enzymes used in each lane, the sizes of the products obtained and whether the enzyme cuts the product or not are detailed in the box on the right.
B. The nucleotide sequence of the PCR-amplified exon 2 region of the DRB1*0803 allele is shown, with the cleavage sites for the restriction endonucleases detailed. Primer sequences are highlighted in red.

Figure 2

ABO genotyping by PCR-RFLP (see ref. 34).

A. RFLP patterns of 5 DNA samples after digestion of the PCR-amplified products by Kpn I and Alu I enzymes. Lanes 1-5 are samples with genotype OO, BO, AO, AO, and BO, respectively. Their types are determined according to the restriction patterns observed, as explained in **Fig. 2b**.
B. ABO genotype is assigned by intersection of possible genotypes from patterns of

DNA fragments by Kpn I and Alu I digestion (upper table). ABO genotypes can be determined for the 5 DNA samples shown in **Fig. 2a** (lower table).

C. ABO alleles can be determined by amplifying the ABO glycosyl transferase locus. Exons 6 and 7 are amplified by PCR with 2 pairs of primers highlighted on the figure by labeled coloured lines, primer1/primer 2 and primer 3/primer4, respectively and digesting with KpnI an AluI enzymes; Nucleotide 258 bp (G) is deleted in O allele (represented by a green triangle on the right hand diagram) which creates the Kpn I site. Allele B is discriminated by the Alu I enzyme which recognizes the G to A transition at 700 bp. The partial nucleotide sequence of the 'A' allele of the ABO gycosyl transferase locus is shown , highlighting the bases at 258 and 700 bp (in brackets) which allow discrimination of the 'O' and 'B' alleles, respectively. The cleavage sites for the restriction endonucleases are also detailed on the right.

Figure 3

The mismatch PCR-RFLP method for the detection of Y393N in the exon 9 of the *BCKDHA* gene (see ref. 38).

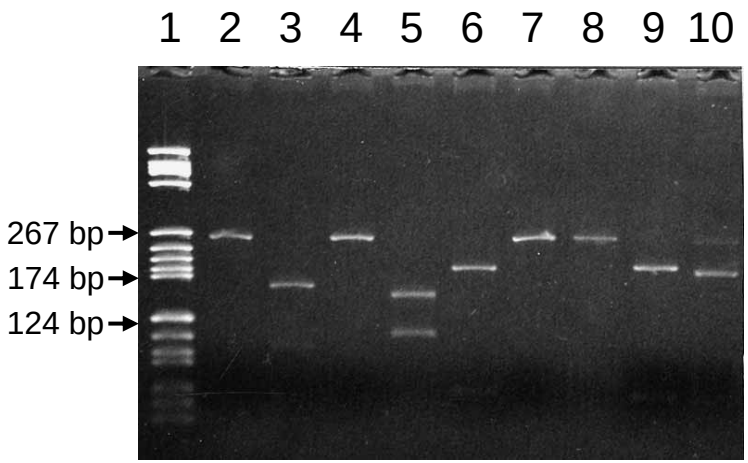
A. A single nucleotide change (a T to A transversion at position 1324 in BCKDHA gene, shown in pink in a yellow box in the sequence) causes a tyrosine to asparagine substitution (Y393N). This substitution is discriminated using a mismatch PCR-RFLP method; a PCR primer (5'-primer) was designed to incorporate one mismatch with the original template sequence (shown in red) to identify the Y393N substitution by creating a Sca I site (5'AGTACT3') in the absence of the substitution. As an internal control, a Sca I restriction site was

incorporated in the 3’ primer. Therefore, all PCR products yield 170bp fragment after digestion with Sca I.

B. Schematic electrophoretic pattern of RFLP after digestion with Sca I enzyme. The Y393N *BCKDHA* allele is identified by RFLP pattern.

Figure 1PCR-RFLP typing of HLA-DRB1

1-A



HLA-DRB1*0803

1: marker,pBR322/HaeIII digestion

2: AvaII, uncut, 265 bp

3: Fok I, cut, 170 bp, 95 bp

4: Kpn I, uncut, 265 bp

5: Hae II, cut, 159 bp, 106 bp

6: Cfr13I, cut, 201 bp, 63 bp, 1 bp

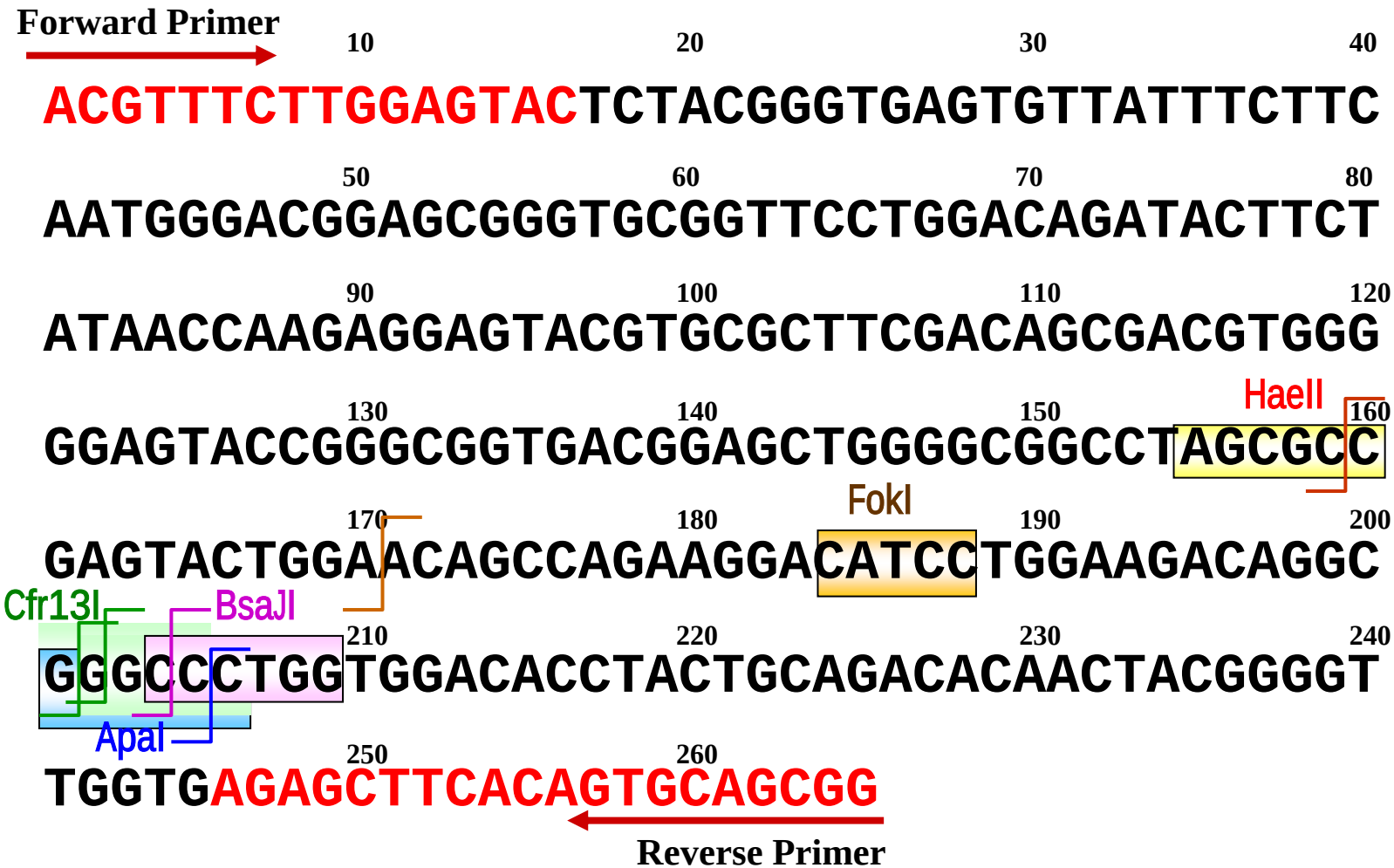
7: SfaNI, uncut, 265 bp

8: SacII, uncut, 265 bp

9: BsaJI, cut, 204 bp, 61 bp

10: ApaI, cut, 205 bp, 60 bp

1-B



Cleavage sites of restriction enzymes

HaeII : 5’···R▼GCGCY···3’

Cfr13I : 5’···G▼GNCC···3’

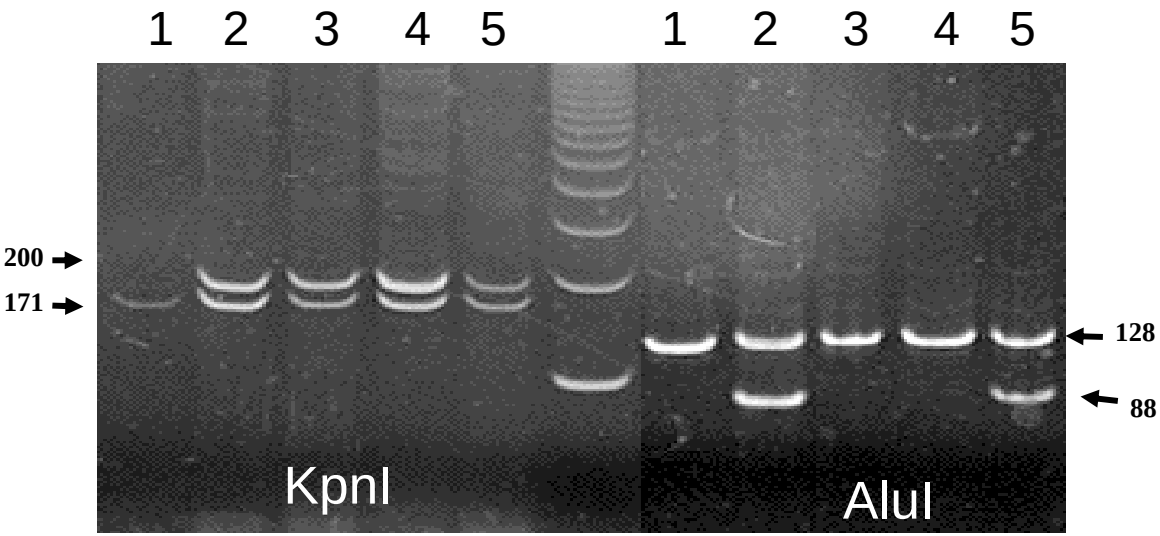
FokI : 3’···CCTAC(N)$_{13}$▼···5’

BsaJI : 5’···C▼CNNGG···3’

ApaI : 5’···GGGCC▼C···3’

Figure 2 ABO genotyping by PCR-RFLP

2-A



2-B

KpnI(Exon6)		Genotype	AluI(Exon7)		Genotype
200bp	171bp		128bp	88bp	
+	-	AA,AB,BB	+	-	AA,AO,OO
+	+	AO,BO	+	+	AB,BO
-	+	OO	-	+	BB

KpnI(Exon6)		No	AluI(Exon7)		Genotype
200bp	171bp		128bp	88bp	
-	+	1	+	-	OO
+	+	2	+	+	BO
+	+	3	+	-	AO
+	+	4	+	-	AO
+	+	5	+	+	BO

2-C

Primer 1 (258 bp)

ACA CCG TGG AAG GAT GTC CTC GTG GTG ACC CCT TGG GTG GCT CCC
ATT GTC TGG GAC GGC ACA TTC AAC ATC GAC ATC CTC AAC GAG CAG
TTC AGG CTC CAG AAC ACC ACC ATT GGG TTA ACT GAG TTT GCC ATC
AAG AAA | gtaagtcagtgaggtggccgagggtagagacccaggcagtggcgagtgactgtggacattg

exon | intron

Primer 2

(661) Primer 3

(700 bp)

GTG GAG ATC CTG ACT CCG CTG TTC GGC ACC CTG CAC CCC GGC TTG
TAC GGA AGC AGC CGG GAG GCC TTC ACC TAC GAG CGC CGG CCC CAG
TCC CAG GCC TAC ATC CCC AAG GAC GAG GGC GAT TTC TAC (789)

Primer 4

KpnI

G T A C | C
C | C A T G G

A,B:GGTGACC
O :GGT▲ACC

AluI

A G | C T
T C | G A

A,O : GGCT
B : AGCT

Figure 3 The mismatch PCR-RFLP for the detection of Y393N in the B*CKDHA* gene