Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International 162 (2006) 74–79
Forensic Science International
www.elsevier.com/locate/forsciint

SNP genotyping by multiplex amplification and microarrays assay for forensic application

Li Li [a], Cheng-Tao Li [a], Rong-Yu Li [b,*], Yan Liu [a], Yuan Lin [a], Ting-Zhi Que [a], Mei-Qian Sun [b], Yao Li [b]

[a] *Institute of Forensic Sciences, Ministry of Justice, Shanghai 200063, PR China*
[b] *Shanghai Biostar Genechip Research Institute, Shanghai 200092, PR China*

Available online 1 August 2006

Abstract

Objective: Research on the application feasibility of SNP genotyping for forensic identification by microarrays.
Methods: Oligonucleotide microarrays which could detect 34 different SNPs were used. After hybridization and washing, the arrays were scanned and fluorescence intensities analyzed using Microarray software. Population studies on 34 SNP loci were carried out in a sample of 109 unrelated Chinese Han individuals using oligonucleotide microarrays for genotype detection. The method was also applied to cases.
Results: According to the results of population studies, no deviations from Hardy–Weinberg equilibrium could be found. Among the 34 loci, 3 SNPs were low informative, 4 were medium informative and 27 were high informative. The combination discrimination power (CDP) of the 31 optimal polymorphic SNPs was 0.9999999999979. The matching probability was 2.13×10^{-12}. The average exclusion probability in paternity testing for duos was 0.9609. The average exclusion probability in paternity testing for trios was 0.9970.
Conclusion: The data and case application demonstrated that SNP typing by oligonucleotide probe microarrays was a useful technique for paternity testing and individual identification. Combined with the 28 SNPs loci distributed on HLA-DRB1 and ABO genes, the combination discrimination power (CDP) was 0.99999999999999010. The matching probability was 9.02×10^{-15}. The average exclusion probabilities in duos and in trios were 0.9894 and 0.9992, respectively. It may be concluded that the 59 SNPs loci yield the same power in forensic identification as CODIS STRs currently used.
© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Forensic science; Genotyping; SNP; Microarray

1. Introduction

Single nucleotide polymorphisms (SNPs) are sites in the genome that have at least two different bases at the same location. These polymorphisms are highly abundant, on average, one in 1000 nucleotide positions [1]. In forensic and medical practice, SNPs may provide an important alternative to conventional markers such as STRs. First, the availability of highly abundant SNPs should facilitate the selection of genetic markers [2]. Second, being biallelic markers, SNPs may supersede STRs since SNPs are more robust in terms of laboratory handling, data interpretation, stability of inheritance, and population genetic analysis [3–5]. Finally, the combination of many SNPs could reach a high discrimination power as STRs, although each SNP is less informative than a STR. For these reasons, SNPs may become a key component in future studies of forensic sciences.

SNPs are inferior to microsatellites in polymorphic information content, since SNPs have relatively few alleles. So it is necessary that multiple markers are screened for genotyping. In the present paper we consider 34 SNP markers that are located on autosomal 20 chromosomes (2#–21#). Our study sample consists of 109 individuals from Chinese Han population.

In the present SNP genotyping protocol, 34 SNPs are amplified by a 12-plex PCR amplification, a 7-plex PCR amplification, a 10-plex PCR amplification and a 5-plex PCR amplification. The Cy5-labelled PCR products were applied to the array and hybridized to the oligonucleotide probes. The signals representing the SNP genotypes are detected by Scanarray 4000.

* Corresponding author.
E-mail address: qzjdzx@163.com (R.-Y. Li).

0379-0738/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2006.06.010

2. Materials and methods

2.1. Preparations of oligonucleotide probes on glass supports

2.1.1. Preparation and silanization of glass substrates

Clean glass slides were immersed in 1% 3-aminopropyl-trimethoxy-silane (APS)/95% acetone solution for 2 min and then washed in acetone for 5 min, six times. The slides were roasted at 110 °C for 45 min. Immersed in 0.2% 1,4-phenylene diisothiocyanate (PDC)/10% pyridine/DMF solution for 2 h. Following incubation, the slides were rinsed with acetone once, dried by airing, and kept at 4 °C for use.

2.1.2. Selection of SNP loci and design of typing probes

To ensure that there is enough resolution for forensic application, we selected 34 SNP loci by using the distribution information about the East Asia population via SNP database: http://www.ncbi.nlm.nih.gov. The criteria are as followings: there are only two kinds of genotypes in every SNP locus and the frequency of each genotype is more than 0.3; there are no short tandem repeats in hybridization sequence to avoid the cross-hybridization. According to the specific sequence of each SNP locus (shown in Table 1), 34-pair probes with perfect match and mismatch are designed (shown in Table 2). Each probe contained a 5′-amino group for immobilization chemistry, a 10_mer poly-dT spacer, followed by the nucleotide hybridization sequence. The polymorphic sequence was situated near the center of each hybridization sequence.

2.1.3. Immobilization of amino-modified oligonucleotides onto silanized glass

The 5′-amino-modified oligonucleotide probes were diluted to a concentration of 10 μM in 500 mM Na_2CO_3/$NaHCO_3$ buffer (pH 9.0) and arrayed onto the silanized slide surface by Cartesian Pixsys 7500. Arrayed probes were incubated in humid chamber for 1–2 h, followed by washing once with 0.2% SDS, three times with distilled water, dried by airing and stored at 4 °C for use in hybridization experiments.

2.2. DNA extraction and amplification

To determine the allele frequencies of 34 SNPs, 109 unrelated blood samples were collected from Chinese Han population. DNA was extracted by the Chelex extraction method [6]. Thirty-four SNP loci divided into four groups were amplified in four multiplex PCR reactions as followings: a 12-plex PCR reaction of SPF22, SPF26, SPF29, SPF30, SPF31, SPF32, SPF34, SPF36, SPF49, SPF71, SPF94, SPF99; a 7-plex PCR reaction of SPF1, SPF53, SPF58, SPF63, SPF64, SPF84, SPF90; a 10-plex PCR reaction of SPF16, SPF35, SPF52, SPF54, SPF59, SPF60, SPF68, SPF69, SPF81, SPF97 and a 5-plex PCR reaction of SPF55, SPF62, SPF82, SPF89, SPF91. The groups and the primers were listed in Table 3. All PCR amplifications were performed on a GeneAmp PCR System 9600 (APPLIED BIOSYSTEMS, Norwalk, CT) in 25 μl reactions containing 1× PCR buffer, 1.5 mmol/l $MgCI_2$, 200 μmol/l dNTPs, 400 nmol/l of each primer, 2 U AmpliTaq Gold polymerase and 5 ng template DNA. All the reactions consisted of one cycle at 94 °C for 12 min, 30 cycles at 94 °C for 30 s, 56 °C for 45 s and 72 °C for 30 s, followed by a final extension at 72 °C for 5 min.

Table 1
Details of the 34 SNP loci

Locus	ss#	rs#	Chr	GenBank accession	Variation
SPF1	458585	361519	6	AL139095	AGG/–
SPF16	4576	3243	10	AK095787	C/T
SPF22	4554	2663	11	AP000867	A/G
SPF26	18768	16344	4	AC097459	TTG/–
SPF29	18853	16428	18	AP000845	AGAG/–
SPF30	19068	16663	21	AL163301	AGTG/–
SPF31	19076	16671	20	AL031652	CC/–
SPF32	28266	25544	17	AC004757	TCCTT/–
SPF34	1566163	1160910	9	AL157936	AAAG/–
SPF35	1566210	1305059	5	AC027323	GGTCA/–
SPF36	2418598	1610863	16	AC006112	GAA/–
SPF49	2978717	2067204	16	AC092122	TGGTCA-AAGGCA/–
SPF52	2978839	2067326	9	AL390962	AAC/–
SPF53	2978886	2067373	13	AL139328	ATATAA/–
SPF54	2978889	2067376	16	AC126773	TCCCTG/–
SPF55	3255370	2307519	15	AC100840	TTTCAA/–
SPF58	3255404	2307553	14	AC008056	TGAC/–
SPF59	3255410	2307559	13	AL157878	TAAC/–
SPF60	3255422	2307571	7	AC018728	TACTT/–
SPF62	3255483	2307632	17	AC004797	GAA/–
SPF63	3255506	2307655	12	AC078925	CAT/–
SPF64	3255531	2307680	6	AL031122	CAAA/–
SPF68	3255551	2307700	22	AL078460	TCAC/–
SPF69	3255559	2307708	2	AC068195	GA/–
SPF71	3255608	2307757	2	AC008274	CTTC/–
SPF81	3255971	2308116	6	AL022724	GGA/–
SPF82	3256036	2308181	5	AC079262	AACTACGGC-ACGCCC/–
SPF84	3256110	2308255	3	AC112907	GAGTCGT/–
SPF89	1566205	1160945	19	AC010492	–/AGG
SPF90	1566214	1160953	5	AC018752	TCTA/–
SPF91	1566215	1160954	5	AC008509	–/ACA
SPF94	1566227	1160963	3	AC092040	–/AT
SPF97	1566243	1160979	5	AC005217	–/AAG
SPF99	1566245	1160981	5	AC034252	–/CAG

2.3. Hybridization and fluorescence detection

After multiplex PCR, the products of four reactions (i.e. 12-plex, 7-plex, 10-plex and 5-plex) were mixed and precipitated by alcohol method. Dissolved in 14 μl pre-heated (45 °C) hybridization solution, the PCR products were heat denatured for 3 min at 95 °C and quick cooled on ice for 3 min and then applied to the microarray. The hybridization reaction was kept at 45 °C for 1 h. The slide was then put into 0.2% SDS for 5 min followed by two washes with distilled water. After drying at room temperature, the slide was scanned at 649 nm using Scanarray 4000 (Tatara Inc.). The acquired images were saved in TIFF format and analyzed with the software GenePix Pro3.0.

Tecan Ex. 2131, Page 2 of 6
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Table 2
Oligonucleotide probes used for SNP genotyping

Locus	Sequence	
	Probe 1	Probe 2
SPF1	5′-NH2-(T)10-TTCAAGCAGGAGGTTAGACC-3′	5′-NH2-(T)10-GTTCAAGCAGGTTAGACCAG-3′
SPF16	5′-NH2-(T)10-TGAACATGACAAAACAA-3′	5′-NH2-(T)10-TGAACATGGCAAAACA-3′
SPF22	5′-NH2-(T)10-TTCAGATTCAGAAGGG-3′	5′-NH2-(T)10-TTCAGATCCAGAAGGG-3′
SPF26	5′-NH2-(T)10-TTGTAGTAACAACAAGTGAATT-3′	5′-NH2-(T)10-TGTTGTAGTAACAAGTGAATTA-3′
SPF29	5′-NH2-(T)10-GTGTCTCTCTCTGATGTACTG-3′	5′-NH2-(T)10-TTATTGTGTCTCTGATGTACT-3′
SPF30	5′-NH2-(T)10-ACCCTCCTCACTAACTGTCC-3′	5′-NH2-(T)10-TGTCCCTCCTAACTGTCC-3′
SPF31	5′-NH2-(T)10-TTCAGGTACCCCCTTCCCTCC-3′	5′-NH2-(T)10-CCTTCAGGTACCCTTCCCTCC-3′
SPF32	5′-NH2-(T)10-CTCTCTCCTTTCCTATCCTTC-3′	5′-NH2-(T)10-GTTCTCTCTCTCCTATCCTTC-3′
SPF34	5′-NH2-(T)10-AGCAAAAAGAAAGGACCTTA-3′	5′-NH2-(T)10-GATGCAAAAAGGACCTTAA-3′
SPF35	5′-NH2-(T)10-TGTTAGTGGTCAGGTCACATG-3′	5′-NH2-(T)10-AGTTAGTGGTCACATGTGCATC-3′
SPF36	5′-NH2-(T)10-GGGAAATTCTTCTTGGT-3′	5′-NH2-(T)10-GGGAAATTCTTGGTTTT-3′
SPF49	5′-NH2-(T)10-CAACATGGTCAAAGGCACAG-3′	5′-NH2-(T)10-AAGCATCAACACAGAGGTGG-3′
SPF52	5′-NH2-(T)10-AAGAGAGTTGTTGTGAG-3′	5′-NH2-(T)10-AGAGAGTTGTGAGCTGA-3′
SPF53	5′-NH2-(T)10-TCAAATTAATATAATTTAAATC-3′	5′-NH2-(T)10-AAGTCAAATTATTTAAATCTA-3′
SPF54	5′-NH2-(T)10-CC AGGTT TCCCTG TCCTGC-3′	5′-NH2-(T)10-TCTCC AGGTT TCCTGCATGC-3′
SPF55	5′-NH2-(T)10-GACCTT CAA TTTCAA GTACACA-3′	5′-NH2-(T)10-TAGACCTT CAAGTACACAGCT-3′
SPF58	5′-NH2-(T)10-GGGCATTTGAC TGACAGAGT-3′	5′-NH2-(T)10-TGGGCATTTGACAGAGTAGG-3′
SPF59	5′-NH2-(T)10-CATTTATCTTTAACAAACAAA-3′	5′-NH2-(T)10-CATTTATCTTAAACAAAGAGT-3′
SPF60	5′-NH2-(T)10-CTTCATCACTTAAGTAAAGAAGATTT-3′	5′-NH2-(T)10-TTCATCACTTAAGAAGATTTT-3′
SPF62	5′-NH2-(T)10-GAAGGAATGGAAGAAATGCTGGA-3′	5′-NH2-(T)10-G GAAGGAATGGAAATGCTGGA G-3′
SPF63	5′-NH2-(T)10-CTGTTTGCATGATTTTTTC-3′	5′-NH2-(T)10-CTGTTTGCATTTTTTTCC-3′
SPF64	5′-NH2-(T)10-ATGATTGTTTGACCAAAA-3′	5′-NH2-(T)10-CATACATGATTGACCAAA-3′
SPF68	5′-NH2-(T)10-AAACTGACTGAGTGACAAGCC-3′	5′-NH2-(T)10-CTTTTACTGACTGACAAGCC-3′
SPF69	5′-NH2-(T)10-AATTTTGTTCTTTTTTATT-3′	5′-NH2-(T)10-AAATTTTGTTTTTTTATTA-3′
SPF71	5′-NH2-(T)10-TTTCTTTGGAAGAAGAAAG-3′	5′-NH2-(T)10-CAGTTTCTTTGAAGAAAGAG-3′
SPF81	5′-NH2-(T)10-G GAATGGAGGAGGA-GAAGTT-3′	5′-NH2-(T)10-CTG GAATGGAGGA-GAAGTTTT-3′
SPF82	5′-NH2-(T)10-GGGCGTGCCGTAGTTGATTT-3′	5′-NH2-(T)10-CCCACAGAAGGATTTCACCC-3′
SPF84	5′-NH2-(T)10-ATATATCAGCCCTGAGTCGTTAT-3′	5′-NH2-(T)10-ATATATCAGCCCTTATTTTTCTGC-3′
SPF89	5′-NH2-(T)10-GGTC TCGA AGG AGAAGGGC-3′	5′-NH2-(T)10-GGTGGTC TCGA AGAAGGGC-3′
SPF90	5′-NH2-(T)10-TAAGTACTTAGATAGATATATGAGAG-3′	5′-NH2-(T)10-GCTTAAGTACTTAGATATATGAGAG-3′
SPF91	5′-NH2-(T)10-GGCTCAACAACAGAACTAAGACT-3′	5′-NH2-(T)10-CTTGGCTCAACAGAACTAAGACT-3′
SPF94	5′-NH2-(T)10-TG TGCACACACATGATACACT-3′	5′-NH2-(T)10-TG TGCACACACGATACACTG-3′
SPF97	5′-NH2-(T)10-GGCAGGAGAAGGAGGGCC-3′	5′-NH2-(T)10-GAGGCAGGAGGAGGGCCT-3′
SPF99	5′-NH2-(T)10-GAACAAAAACTGCTGGTGATAG-3′	5′-NH2-(T)10-G GAACAAAAACTGGTGATAGGA-3′

The genotype for each SNP loci was determined according to the ratio of signal intensities between two corresponding probes.

2.4. Statistical analysis

Basic statistical computations including allele frequency, observed heterozygosity, expected heterozygosity, average power of paternity exclusion in duos (PE(1)) and trios (PE(2)), polymorphism information contents (PIC) and Hardy–Weinberg equilibrium (HWE) tests were performed using the Cervus program [7–8]. The discrimination power (DP) was computed according to established methods [9–12].

3. Results

3.1. Validation

Validating carried out on DNA extracts prepared from eight blood samples with known SNP genotypes at the 34 loci by sequencing revealed fully consistent and reproducible results. The loci were also genotyped in 106 families. The SNP genotypes were fully consistent with Mendelian segregation.

To evaluate the minimum quantity of DNA needed to obtain an optimal result, genomic DNA of the cell line 9947A (Promega, USA), in ascending amounts from 40 pg to 20 ng were used for amplification. Results showed that not all SNP loci tested could be clearly genotyped simultaneously with DNA amounts of 40 pg in each PCR system. With an increasing DNA concentration, the signal was improved step by step, until all SNPs were detectable at a DNA amount of 1 ng. Due to background noise, unspecific amplification, reliable typing of all 34 SNP loci was possible at DNA template amounts above 4 ng for all four PCRs.

We applied the genotyping method to different types of biological specimen including bloodstains, muscle tissues, saliva stains and urines in forensic cases. All the types of specimen could be given the unambiguous genotypes across 34 SNPs except SPF35 locus. At SPF35 one-third samples showed

Tecan Ex. 2131, Page 3 of 6
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Table 3
Groups of 34 SNP loci and primers used for four multiplex PCR

No.	Locus	Forward primer	Reverse primer	Size (bp)
Group 1 (nos. 1–12, 12 SNPs)				
1	SPF22	5′-cy5-GCTCTAGTGGGAAACCTCAGG-3′	5′-GAATTCCAGGCTCTTGCTTG-3′	275
2	SPF26	5′-cy5-GCACAGGTGTTTTAGAGGC-3′	5′-GTTCCAGTTGTTGGTGTGAC-3′	116
3	SPF29	5′-cy5-CAAAGCCATTCCAAAGTTAA-3′	5′-TATTTGTCCAGCCTATTGCA-3′	135
4	SPF30	5′-cy5-GAACAGTGACATGGACACCT-3′	5′-AAGACACAGCAAGGTGATGT-3′	106
5	SPF31	5′-CAGACTGAGAGACAGCCTTG-3′	5′-cy5-ATATCCTTAAGCTCCAGGGA-3′	99
6	SPF32	5′-GGAAAATTTCCAGCCATTAT-3′	5′-cy5-TGTAATTGGAGTCCCTGAAG-3′	71
7	SPF34	5′-GTCACATGCAAGTTCAGTCA-3′	5′-cy5-AGCACTTCATGGAATTAAGGT-3′	123
8	SPF36	5′-cy5-TGGAAGACACGTCCTAAGAG-3′	5′-AAATGGGTGATTATTCCTCC-3′	86
9	SPF49	5′-GTAAGGGTGTGTCTGTTTGG-3′	5′-cy5-CAGGGTTAGTTTGGTGAATG-3′	113
10	SPF71	5′-cy5-CTGTGAGATAAAGGGCTCAG-3′	5′-GCCATGCATATACATGATCA-3′	124
11	SPF94	5′-cy5-TCCTTACAAATGTGTAGCCAG-3′	5′-CTGGATCCTGGGAACATAAT-3′	124
12	SPF99	5′-cy5-CCAGGTTACCAGTTTCCTTT-3′	5′-AGAACAATAATGGCAGCAAC-3′	94
Group 2 (nos. 13–19, 7 SNPs)				
13	SPF1	5′-TCAGCTTTAATTTAAGGGACATG-3′	5′-cy5-GGGTATCAGTTTGTAAGTGTGGT-3′	150
14	SPF53	5′-AAAGGGGAAAAACATAAATGA-3′	5′-cy5-CACTCTCAGCCTCAATAGATTT-3′	73
15	SPF58	5′-cy5-GAGCATGCATTTCAAAGTTT-3′	5′-AGCAAACAGAACACAAGGAG-3′	110
16	SPF63	5′-cy5-CCTTGTCATCAAAATTCACC-3′	5′-CCTGTGAACCAAACTCTTTG-3′	146
17	SPF64	5′-cy5-CACCATGGAGACAGAGAAAC-3′	5′-TGAATTTGCCAGCTACTGTT-3′	147
18	SPF84	5′-TGTTACCTAACAAAGGCCTG-3′	5′-cy5-GGGTTCTGTTCTGGGTATTT-3′	131
19	SPF90	5′-TGTTCAGAAAATGCATCTCA-3′	5′-cy5-AGGAAGAGGCTTTAAAGCAG-3′	142
Group 3 (nos. 20–29, 10 SNPs)				
20	SPF16	5′-cy5-GACGTGGACAAAGGAGGTTT-3′	5′-CTGCCCTGGTGGTGTTTTAT-3′	271
21	SPF35	5′-AGATTCTCAGTCACCTGCTG-3′	5′-cy5-ACTGATTCACCACATTCCAG-3′	98
22	SPF52	5′-cy5-GCACATTTCATTAAGTCTGTCA-3′	5′-GTCTGTGAACTCAGTCCAGAA-3′	76
23	SPF54	5′-GTTTTCGCTTCCTTCTGAGT-3′	5′-cy5-GAGGAGAATGACTTGTCCCT-3′	96
24	SPF59	5′-CAAAAAGTTTGACTTCTTGGTT-3′	5′-cy5-CAGGACAAGTCTGTTTGGTC-3′	131
25	SPF60	5′-cy5-GGAAGCCATGTCTTTGTTAA-3′	5′-AGGGCATGTACAGTGTGAGT-3′	119
26	SPF68	5′-cy5-GAGCTCAGCTGATCATTCAT-3′	5′-GGAGTGACCTTTCCATCTTT-3′	100
27	SPF69	5′-cy5-TTTGAAATATGTCCCAAGGA-3′	5′-TCAATTTTATCTCCAATTTCGT-3′	107
28	SPF81	5′-TACCAAGCCAGAACTACCAG-3′	5′-cy5-TTGTTGCCAAATACTGTCAA-3′	80
29	SPF97	5′-CATTACCAGTAGAGTGGGGA-3′	5′-cy5-CTATGCCCTACTGGATCTAGG-3′	74
Group 4 (nos. 30–34, 5 SNPs)				
30	SPF55	5′-CCCCACTCCCATTCTAATAT-3′	5′-cy5-TTCTTTTCTTTGGTGACAGG-3′	120
31	SPF62	5′-ATCAACACAGACCTGCCTTA-3′	5′-cy5-TTGGCCATCTGACTATGAAT-3′	124
32	SPF82	5′-cy5-ACCAAGAATGTGACAGGAAA-3′	5′-AAAAAGCGATTCTTCTCCTT-3′	123
33	SPF89	5′-TCTCAGTAGGAGAGGGGTGT-3′	5′-cy5-CTTGTGAAACCAGATATCCG-3′	113
34	SPF91	5′-cy5-TTTCACAAATGCTTTCACAA-3′	5′-AAGTTTCAAGGTGACTATGTCAA-3′	117

weak signals which were usually caused by low amounts of PCR products.

Fig. 1 and Table 4 showed the SNP typing results of a muscle tissue and a bloodstain in a case. In order to know the truth of a traffic accident, we need to detect the bloodstain found at the wheel of a suspected truck and muscle sample from the dead. According to CODIS STR genotyping results the two samples could not be excluded from the same person. SNP genotypes of every locus also matched between the two samples. The random matching probability of the SNPs in Chinese Han population was calculated as 4.79×10^{-12}.

3.2. Allelic distributions and tests of Hardy–Weinberg and genotypic disequilibrium

A total of 109 related bloodstain samples were genotyped at the 34 SNP loci among Chinese Han population. The allele frequencies were shown in Table 5. These SNP loci were in Hardy–Weinberg equilibrium.

3.3. Calculation for 'forensic statistics'

The 'forensic statistics' were summarized in Table 6. According to the forensic values, 27 loci (i.e. SPF1, SPF16, SPF22, SPF26, SPF29, SPF30, SPF31, SPF36, SPF49, SPF52, SPF53, SPF54, SPF55, SPF58, SPF59, SPF60, SPF62, SPF68, SPF69, SPF71, SPF82, SPF84, SPF89, SPF90, SPF94, SPF97 and SPF99) were high informative markers: the distribution frequency of minor allele of each SNP was above 0.30; the observed heterozygosity (H(O)) was above 0.5; the discrimination power (DP) was above 0.5; the exclusion probability in duos (PE(1)) and trios (PE(2)) was above 0.10 and 0.17, respectively; the polymorphic information content (PIC) was above 0.30. No deviations from Hardy–Weinberg expectations were observed in the 27 loci ($P > 0.5$). Among all the 34 SNPs,

Tecan Ex. 2131, Page 4 of 6
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

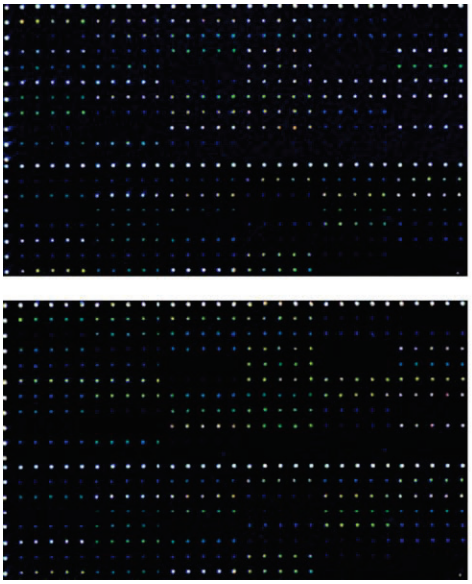


Fig. 1. SNP genotyping of two samples from the same person.

Table 5
Distribution of genotypes and allele frequencies for 34 SNP loci in 109 related Han Chinese

No.	Locus	Allele 1	Allele 2	No.	Locus	Allele 1	Allele 2
1	SPF1	0.4907	0.5093	18	SPF59	0.6606	0.3394
2	SPF16	0.4450	0.5550	19	SPF60	0.3073	0.6927
3	SPF22	0.5000	0.5000	20	SPF62	0.6055	0.3945
4	SPF26	0.2569	0.7431	21	SPF63	0.1560	0.8440
5	SPF29	0.7110	0.2890	22	SPF64	0.8165	0.1835
6	SPF30	0.3991	0.6009	23	SPF68	0.6651	0.3349
7	SPF31	0.5917	0.4083	24	SPF69	0.4312	0.5688
8	SPF32	0.1972	0.8028	25	SPF71	0.3119	0.6881
9	SPF34	0.0917	0.9083	26	SPF81	0.1193	0.8807
10	SPF35	0.0714	0.9286	27	SPF82	0.5596	0.4404
11	SPF36	0.4037	0.5963	28	SPF84	0.7315	0.2685
12	SPF49	0.4633	0.5367	29	SPF89	0.3945	0.6055
13	SPF52	0.3119	0.6881	30	SPF90	0.4722	0.5278
14	SPF53	0.3532	0.6468	31	SPF91	0.8426	0.1574
15	SPF54	0.6881	0.3119	32	SPF94	0.5421	0.4579
16	SPF55	0.7018	0.2982	33	SPF97	0.5046	0.4954
17	SPF58	0.3486	0.6514	34	SPF99	0.5780	0.4220

Table 4
SNP genotyping of two samples from the same person[a]

SNP locus	Muscle sample	Blood sample
SPF1	n1–n2	n1–n2
SPF16	n1–n2	n1–n2
SPF22	n1	n1
SPF26	n2	n2
SPF29	n1–n2	n1–n2
SPF30	n1–n2	n1–n2
SPF31	n1	n1
SPF32	n2	n2
SPF34	n2	n2
SPF35	n2	n2
SPF36	n1–n2	n1–n2
SPF49	n2	n2
SPF52	n2	n2
SPF53	n2	n2
SPF54	n1	n1
SPF55	n1	n1
SPF58	n2	n2
SPF59	n1–n2	n1–n2
SPF60	n2	n2
SPF62	n1	n1
SPF63	n2	n2
SPF64	n1	n1
SPF68	n1	n1
SPF69	n1	n1
SPF71	n2	n2
SPF81	n2	n2
SPF82	n1–n2	n1–n2
SPF84	n1	n1
SPF89	n2	n2
SPF90	n1–n2	n1–n2
SPF91	n1	n1
SPF94	n1–n2	n1–n2
SPF97	n1	n1
SPF99	n1	n1

[a] n1 and n2 indicate alleles 1 and alleles 2, respectively.

Table 6
Forensic efficiency values of 34 SNPs in the Han Chinese population sample

Locus	H(O)	H(E)	PIC	PE(1)	PE(2)	DP
SPF1	0.519	0.502	0.375	0.125	0.187	0.6151
SPF16	0.615	0.496	0.372	0.122	0.186	0.5419
SPF22	0.413	0.502	0.375	0.125	0.188	0.6572
SPF26	0.367	0.384	0.309	0.073	0.154	0.5468
SPF29	0.468	0.413	0.327	0.084	0.163	0.5505
SPF30	0.486	0.482	0.365	0.115	0.182	0.6112
SPF31	0.541	0.485	0.366	0.117	0.183	0.5850
SPF32	0.303	0.318	0.267	0.050	0.133	0.4819
SPF34	0.147	0.167	0.153	0.014	0.076	0.2811
SPFF35	0.057	0.134	0.124	0.009	0.062	0.1849
SPF36	0.367	0.484	0.366	0.116	0.183	0.6464
SPF49	0.450	0.500	0.374	0.124	0.187	0.6437
SPF52	0.459	0.431	0.337	0.092	0.169	0.5723
SPF53	0.358	0.459	0.353	0.104	0.176	0.6227
SPF54	0.422	0.431	0.337	0.092	0.169	0.5841
SPF55	0.431	0.420	0.331	0.088	0.165	0.5708
SPF58	0.422	0.456	0.351	0.103	0.176	0.6090
SPF59	0.459	0.451	0.348	0.101	0.174	0.5915
SPF60	0.413	0.428	0.335	0.091	0.168	0.5829
SPF62	0.477	0.480	0.364	0.114	0.182	0.6134
SPF63	0.183	0.264	0.229	0.035	0.114	0.3963
SPF64	0.275	0.301	0.255	0.045	0.127	0.4612
SPF68	0.560	0.448	0.346	0.099	0.173	0.5353
SPF69	0.440	0.493	0.370	0.120	0.185	0.6400
SPF71	0.330	0.431	0.337	0.092	0.169	0.5959
SPF81	0.165	0.211	0.188	0.022	0.094	0.3343
SPF82	0.477	0.495	0.371	0.121	0.186	0.6286
SPF84	0.444	0.395	0.316	0.077	0.158	0.5410
SPF89	0.569	0.480	0.364	0.114	0.182	0.5612
SPF90	0.519	0.501	0.374	0.124	0.187	0.6137
SPF91	0.259	0.266	0.230	0.035	0.115	0.4237
SPF94	0.542	0.499	0.373	0.123	0.187	0.5978
SPF97	0.587	0.502	0.375	0.125	0.187	0.5700
SPF99	0.385	0.490	0.369	0.119	0.184	0.6505

Tecan Ex. 2131, Page 5 of 6
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

four loci (i.e. SP32, SP63, SP64, SP91) were medium informative, three loci (i.e. SPF34, SPF35 and SPF81) were low informative. The combination discrimination power (CDP) of the 31 optimal polymorphic SNPs above was 0.9999999999979. The matching probability (PM) was 2.13×10^{-12}. The average exclusion probability in duos (PE(1)) was 0.9609. The average exclusion probability in trios (PE(2)) was 0.9970.

4. Discussion and conclusion

We selected 34 SNPs from the SNP database for study. Each SNP locus is biallelic. The most common forms of sequence variation are insertions/deletions.

Currently employed methods for SNP typing include discrimination-by-hybridization, chemical mismatch detection, DHPLC and minisequencing [2,13]. High-density oligonucleotide microarrays are typing systems based on hybridization assays [14,15]. To genotype the 34 SNP markers, 34 pairs of cy5-labelled primers and oligonucleotide probes are employed. PCR reactions were performed in four groups (i.e. 12-plex, 7-plex, 10-plex and 5-plex). The PCR products sizing from 71 to 275 bp were genotyped simultaneously on the microarrays across all the 34 SNP loci. The results showed that the typing method would be reliable for forensic practices.

The population studies and tests for HWE showed no deviation at 34 loci among Chinese Han ethnic groups. Among these SNPs, 31 loci were optimal (i.e. 27 loci were high informative and 4 loci were medium informative), the other 3 loci were low informative. The combined DP of the optimal SNPs was estimated to be 0.9999999999979 and cumulative PE in duos and trios was 0.9609 and 0.9970, respectively. The matching probability (PM) was 2.13×10^{-12}. We have genotyped different types of forensic samples. The results show that multiplex amplification and oligonucleotide arrays assay can rapidly and accurately score the genotypes. It was suggested that 31 loci among the 34 SNPs would be very suitable and useful for application in personal identification cases.

The population studies showed low informative at three loci (i.e. SPF34, SPF35 and SPF81) among Chinese Han ethnic groups. It might mean that the loci were not so valuable for application in forensic sciences.

The data and case application demonstrated that SNP typing by oligonucleotide probe microarrays was a useful technique for paternity testing and individual identification. In addition to this method, we have developed a typing system for HLA-DRB1 and ABO by hybridization of fluorescently labeled PCR produces to oligonucleotide probes immobilized on microarrays [16–18]. Combined HLA-DRB1, ABO with the 31 optimal SNPs loci, the combination discrimination power (CDP) was 0.9999999999999910. The matching probability was 9.02×10^{-15}. The average exclusion probabilities in duos (PE(1)) and trios (PE(2)) were 0.9894 and 0.9992, respectively. It may be concluded that the loci (i.e. HLA-DRB1, ABO and 31 informative SNPs) yield the same power in forensic identification as CODIS STRs currently used.

Acknowledgements

The authors would like to thank Mr. Tsui Pui for reviewing the manuscript. This study was supported by the Ministry of Scientific Technology, PR China (Grant no. 2001BA801B03).

References

[1] D.G. Wang, J.B. Fan, C.J. Siao, et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, Science 280 (1998) 1077–1082.
[2] M. Krawczak, Informativity assessment for biallelic single nucleotide polymorphisms, Electrophoresis 20 (1999) 1676–1681.
[3] J.L. Weber, C. Wong, Mutation of human short tandem repeats, Hum. Mol. Genet. 2 (1993) 1123–1128.
[4] L. Kruglyak, The use of genetic map of biallelic markers in linkage studies, Nat. Genet. 17 (1997) 21–24.
[5] A.M. Valdes, M. Slatkin, N.B. Freimer, Allele frequencies at microsatellite loci: the stepwise mutation model revisited, Genetics 133 (1993) 737–749.
[6] P.S. Walsh, D.A. Metzger, R. Higuchi, Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques 10 (1991) 506–518.
[7] S. Guo, E.A. Thompson, Performing the exact test of Hardy–Weinberg proportion for multiple alleles, Biometrics 48 (1992) 361–372.
[8] http://helios.bto.ed.ac.uk/evolgen.
[9] P. Bostein, R.L. White, M. Skolnick, Construction of a genetic linkage map in man using restriction length polymorphism, Am. J. Hum. Genet. 32 (1980) 314.
[10] S.J. Odelberg, R. White, Repetitive DNA: molecular structure, polymorphism and forensic application, in: DNA and Other Polymorphisms in Forensic Science, 1st ed., Year Book Medical Publishers Inc., Chicago, 1990, p. 26.
[11] A. Edwards, H.A. Hammond, J. Jin, Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups, Genomics 12 (1992) 241–253.
[12] B.S. Weir, Genetic Data Analysis II, Sinauer Associates, Sunderland, MA, 1996, pp. 161–201.
[13] S. Inagaki, Y. Yamamoto, Y. Doi, et al., A new 39-plex analysis method for SNPs including 15 blood group loci, Forensic Sci. Int. 144 (2004) 45–57.
[14] M.H. Shapero, J. Zhang, A. Loraine, et al., MARA: a novel approach for highly multiplexed locus-specific SNP genotyping using high-density DNA oligonucleotide arrays, Nucl. Acids Res. 32 (22) (2004) e181.
[15] J. Huang, W. Wei, J. Zhang, et al., Whole genome DNA copy number changes identified by high density oligonucleotide arrays, Hum. Genomics 1 (4) (2004) 287–299.
[16] C.-T. Li, L. Li, Y. Li, et al., HLA-DRB1 and ABO genotyping by DNA chip, Chin. J. Forensic Sci. 2 (5) (2002) 23–27 (in Chinese).
[17] L. Li, C.-T. Li, Y. Liu, et al., Research on the application feasibility of HLA-DRB1 genotyping for forensic identification by oligonucleotide chip, J. Forensic Med. 20 (2) (2004), pp. 81–84, 87 (in Chinese).
[18] L. Li, C.-T. Li, R.-Y. Li, et al., ABO genotyping by duplex amplification and oligonucleotide arrays assay, J. Forensic Med. 20 (4) (2004) 193–196 (in Chinese).

Tecan Ex. 2131, Page 6 of 6
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028