Acta Med. Okayama, 2003
Vol. 57, No. 2, pp. 59-71

Acta Medica Okayama
http://www.lib.okayama-u.ac.jp/www/acta/

Original Article

Analysis of Short Tandem Repeat (STR) Polymorphisms by the PowerPlex 16 System and Capillary Electrophoresis: Application to Forensic Practice

Osamu Okamoto, Yuji Yamamoto*, Sachiyo Inagaki, Kei Yoshitome, Takaki Ishikawa, Kiyomi Imabayashi, Satoru Miyaishi, and Hideo Ishizu

Department of Legal Medicine, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8558, Japan

Allele and genotype frequencies for 15 short tandem repeat (STR) polymorphisms — D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX and FGA — in a Japanese population were estimated. No deviations of the observed allele frequency from Hardy-Weinberg equilibrium expectations were found for any of the systems studied. Between 2 new pentanucleotide STR loci, Penta E and Penta D, for which there is only limited data regarding the allelic distribution in Japanese, the Penta E locus was found to be highly polymorphic and exhibited a tri- or tetra-modal distribution pattern having allelic peaks with 5, 11, 15 and 20 repeats. The distribution was significantly different from that of the other ethnic groups. Statistical parameters of forensic importance, the power of discrimination (PD), observed and expected heterozygosity values (H), polymorphism information content (PIC), power of discrimination (PD), matching probability (pM), power of exclusion (PE), and typical paternity index (PI), were calculated for the loci. These parameters indicated the usefulness of the loci in forensic personal identification and paternity testing among Japanese. The systems Penta E, FGA, D18S51 and D8S1179 were the most informative. This method was successfully applied to forensic personal identification and paternity testing among Japanese, thereby confirming its efficacy for forensic practice.

Key words: population data, DNA typing, short tandem repeats, personal identification, paternity testing

The short tandem repeats (STRs) are a rich source of highly polymorphic markers in the human genome, are relatively small in size, and can be studied with the polymerase chain reaction (PCR) in a multiplexed fashion [1]. Thus the STR polymorphisms are highly useful tools for personal identification and paternity testing in forensic practice [2-6]. The use of multiplex amplification of many STR loci in conjunction with semi-automated fluorescent typing has proven highly efficient and informative [5, 6]. In this study, we used a commercially available multiplex STR typing kit to study 15 STR systems (D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX, and FGA) in a Japanese population, and estimated the allele and genotype frequencies. These 15 STR loci include 2 new pentanucleotide repeat STR loci, Penta E and Penta D, for which there is only limited data in regard to the

Received August 30, 2002; accepted October 30, 2002.
*Corresponding author. Phone: +81-86-223-7151; Fax: +81-86-235-7201
E-mail: y-yamamo@md.okayama-u.ac.jp (Y. Yamamoto)

allelic distribution in Japanese. In addition, we evaluated the usefulness of these 15 STR systems in forensic investigations and paternity testing by calculating several statistical parameters of forensic interest and by examining materials in 20 paternity, 1 zygosity diagnosis, and 2 criminal cases.

Materials and Methods

Materials. Peripheral blood samples were taken after obtaining informed consent from 125 unrelated Japanese student or staff volunteers in 1999 and 2000 and from 52 subjects requesting paternity tests from 1997 to 2000. Peripheral blood samples were also obtained from 20 trios (60 Japanese subjects) requesting paternity tests and from a pair of twins in a zygosity test. Two criminal cases were also investigated. In the first, a cardiac blood sample from an infant corpse, a blood clot on the maternal side of the placenta, a piece of the umbilical cord attached to the placenta, and a peripheral blood sample of a woman suspected to be the mother of the infant were used for forensic identity tests. In the second, a tooth taken from the skeletonized remains of a 30–50 year old man and a peripheral blood sample taken from a man who was thought to be the father of the deceased were also tested.

DNA extraction. The peripheral blood samples, tissue sample and blood clot were treated with SDS and proteinase K, and DNAs were recovered by isopropanol precipitation after phenol-chloroform extraction. The hard tooth tissue was fragmented and pulverized, then decalcified with 0.5 M EDTA (pH 8.0) for 2 days. DNA was extracted in the manner described above, dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA pH 8.0), and stored at 4 °C or −20 °C until use.

PCR amplification. Multiplexed PCR amplification of 15 STR loci, D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX and FGA, was performed using the PowerPlex™ 16 System (Promega Corporation, Madison, WI, USA). The characteristics of the 15 STR loci are shown in Table 1 [7]. PCR was carried out in a 25 μl volume using 0.5–5 ng template DNA, 2.5 μl of 10X reaction buffer, 4 units of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA), 5 μg of bovine serum albumin and 2.5 μl of 10X primer pair mix following the procedure recommended by the manufacturer with slight modifications [7]. PCR was performed using a PC-700 thermal cycler (Astec, Fukuoka, Japan) with the following amplification conditions: an initial incubation at 95 °C for 11 min and 96 °C for 2 min; 10 cycles of 94 °C for 1 min, 60 °C for 1 min, and 70 °C for 1.5 min; and 23 cycles of 90 °C for 1 min, 60 °C for 1 min, 70 °C for 1.5 min; and a final extension of 60 °C for 25 min.

Table 1 The characteristics of the fifteen STR loci examined in this study

STR locus	Chromosome location	Repeat motif	GenBank locus or (Accession number)	Primer label
D3S1358	3p	TCTA complex	(AC066584)	FL
TH01	11p15.5	AATG	HUMTH01	FL
D21S11	21q11–21	TCTA complex	HUMD21LOC	FL
D18S51	18q21.3	AGAA	HUMUT574	FL
Penta E	15q	AAAGA	(AC027004)	FL
D5S818	5q23.3–32	AGAT	(G08446)	JOE
D13S317	13q22–31	TATC	(G09017)	JOE
D7S820	7q11.21–22	GATA	(G08616)	JOE
D16S539	16q24-qter	GATA	(G07925)	JOE
CSF1PO	5q33.3–34	AGAT	HUMCSF1PO	JOE
Penta D	21q	AAAGA	(AP001752)	JOE
vWA	12p12-pter	TCTA complex	HUMVWFA31	TMR
D8S1179	8q	TCTA complex	(G08710)	TMR
TPOX	2p23-pter	AATG	HUMTPOX	TMR
FGA	4q28	TTTC complex	HUMFIBRA	TMR

FL, fluorescein; JOE, 6-carboxy-4',5'-dichloro-2',7'-dimethoxyfluorescein; TMR, carboxy-tetramethylrhodamine.

DNA fragment analysis. PCR-amplified fragments were analyzed with an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Electrophoresis was performed using a 50 μm inner diameter, 47 cm length capillary and Performance Optimized Polymer 4 (POP 4) (Applied Biosystems) at 15 kV for 30 min at 60 °C. Then the fragments were typed based on allelic ladders contained in the kit described above and GeneScan® Analysis Software (version 2.1; Applied Biosystems).

Sequence analysis. After electrophoresing the amplified PCR products of the respective samples on 5% polyacrylamide gels, the targeted DNA bands were eluted from the gel, and the eluted DNAs were reamplified by PCR. After purification, the DNAs were subjected to sequence reaction with a BigDye Terminator Kit (Applied Biosystems). After terminating the reaction, nucleotide sequences were determined using the ABI PRISM 310 Genetic Analyzer and Sequence Analysis software (version 3.0; Applied Biosystems).

Statistical analysis. Possible divergence from Hardy-Weinberg equilibrium (HWE) was tested using the chi-square test, homozygosity test [8] and G-T's exact test [9]. When applying the chi-square test, alleles were pooled so that each allelic event could be under the "rule of five" [10]. For the exact test, alleles were also pooled with adjacent alleles in some loci so that no alleles would have less than 5 events [11] when the P-value based on the test without allele grouping was less than 0.05. To estimate the probability of the exact test, we used GENEPOP ON THE WEB software (Markov chain parameters: batchs = 100, iteration per batch = 1000) [12]. The allele frequencies calculated in this study were compared with those of Caucasian Americans and African Americans [13, 14] using the chi-square test after grouping of adjacent alleles so that no alleles would have less than 5 events. The allele frequencies other than that for the Penta D locus were also compared with those of a Japanese population living in Miyagi [6], a prefecture in the northeastern part of Japan, and the allele frequencies other than those for the Penta E and Penta D loci were compared with those of Chinese populations from Macau [16] on the southeastern coast of China, Shenyang [17] in the northeastern part of China, and Taiwan [18] using the chi-square test after grouping in the manner described above.

Population statistics for the loci, measures of the informativeness of genetic markers, the heterozygosity value (H) [19] and the polymorphism information content (PIC) [20] were calculated. From a forensic point of view, measures of the discriminatory power of genetic markers, the power of discrimination (PD) [21] and the matching probability or probability of match (pM) [22], a probability of random match of DNA types of 2 individuals, were computed. In addition, the power of exclusion (PE) [23], which is a measure of the ability of a genetic test to exclude non-fathers from paternity, and the typical paternity index (PIt) [23], which is a measure of the genetic odds in favor of paternity given the genotypes for the mother, child and alleged father, were calculated. Finally, the probability of paternity, a conditional probability of whether an alleged father is the actual father of a child, was calculated based on Essen-Möllor's equation, and the paternity index was determined followed by Brenner and Morris [23].

Results

1. Population studies

Allele frequencies in a Japanese population. Table 2 shows the estimated allele frequencies at the 15 STR loci in a Japanese population. Four of the alleles observed in the present study — allele 25 at the D18S51 locus, allele 25.2 at the FGA locus, and alleles 23 and 24 at the Penta E locus — were not observed in the Caucasian American or African American populations in previous studies [13–15, 24]. The sequence structures of these rare alleles were confirmed by sequencing. In the genotypic distribution of the 15 STR loci, a slight deviation from the HWE proportion was noted only in the D21S11 locus ($P = 0.0439$, SE = 0.0099) based on the exact test without allele grouping, but the test after grouping of alleles exhibited no deviations from HWE proportion ($P = 0.3548$, SE = 0.0380) (Table 3).

The Penta E locus on chromosome 15, one of the new pentanucleotide STR loci examined in this study, was highly polymorphic, and 19 alleles were found in this Japanese population. As shown in Fig. 1, the allele frequency of this locus displays a tri- or tetra-modal distribution having allelic peaks with 5, 11, 15 and 20 repeats, and the allele 15 was most frequent (0.150). The distribution was significantly different from those reported in the Caucasian American and African American populations [14, 15, 24]. As mentioned above, alleles 23 and 24, which had not been detected in Caucasians and Africans, were found in the present Japanese population. Allele 8, which is most frequent in African Americans

Table 2 Allele frequencies for 15 STR loci in 177 unrelated Japanese

Allele	D3S1358	TH01	D21S11	D18S51	PentaE	D5S818	D13S317	D7S820	D16S539	CSF1PO	PentaD	vWA	D8S1179	TPOX	FGA
5					0.088										
6		0.198													
7		0.285			0.003	0.003		0.003		0.011	0.003				
8		0.034			0.011	0.014	0.243	0.127	0.006		0.020			0.483	
9		0.415			0.023	0.082	0.161	0.073	0.280	0.068	0.297			0.150	
9.3		0.054													
10		0.014			0.059	0.172	0.121	0.229	0.172	0.184	0.220		0.116	0.017	
11				0.003	0.121	0.319	0.243	0.316	0.243	0.203	0.167		0.110	0.308	
12	0.003			0.028	0.102	0.254	0.175	0.212	0.192	0.427	0.141		0.073	0.042	
13				0.195	0.011	0.144	0.042	0.037	0.099	0.073	0.110		0.251		
14	0.042			0.218	0.051	0.008	0.014	0.003	0.008	0.023	0.034	0.206	0.195		
15	0.384			0.181	0.150	0.003				0.011	0.003	0.031	0.147		
16	0.305			0.153	0.121							0.175	0.088		
17	0.189			0.085	0.073						0.006	0.277	0.017		0.008
18	0.071			0.056	0.045							0.209	0.003		0.023
19	0.006			0.025	0.023							0.085			0.085
20				0.014	0.051							0.014			0.099
21				0.014	0.042							0.003			0.096
22				0.017	0.020										0.240
22.2															0.006
23				0.003	0.003										0.147
23.2															0.003
24				0.003	0.003										0.178
24.2															0.003
25				0.006											0.071
25.2															0.003
26			0.006												0.034
27															0.006
28			0.042												
29			0.226												
29.2			0.008												
30			0.384												
30.2			0.006												
31			0.082												
31.2			0.079												
32			0.023												
32.2			0.116												
33			0.008												
33.2			0.020												
H_{obs}	0.7345	0.6893	0.7401	0.8362	0.9040	0.7514	0.7966	0.7740	0.8136	0.7458	0.7684	0.7966	0.8757	0.6780	0.8249
H_{exp}	0.7167	0.7028	0.7720	0.8462	0.9105	0.7760	0.8086	0.7797	0.7863	0.7322	0.8020	0.7981	0.8383	0.6474	0.8561
PIC	0.6682	0.6524	0.7447	0.8279	0.9036	0.7418	0.7808	0.7463	0.7526	0.6965	0.7742	0.7679	0.8185	0.5869	0.8405
PD	0.8711	0.8597	0.9168	0.9552	0.9807	0.9181	0.9312	0.9132	0.9108	0.8891	0.9322	0.9252	0.9454	0.8000	0.9607
pM	0.1289	0.1403	0.0832	0.0448	0.0193	0.0819	0.0688	0.0868	0.0892	0.1109	0.0678	0.0748	0.0546	0.2000	0.0393
PE	0.4836	0.4119	0.4930	0.6678	0.8035	0.5122	0.5928	0.5517	0.6244	0.5025	0.5417	0.5928	0.7461	0.3950	0.6460
PIt	1.8830	1.6091	1.9239	3.0517	5.2059	2.0114	2.4583	2.2125	2.6818	1.9667	2.1585	2.4583	4.0227	1.5526	2.8548

Hobs, observed heterozygosity; Hexp, expected heterozygosity; PIC, polymorphism information content; PD, power of discrimination; pM, matching probability; PE, power of exclusion; PIt, typical paternity index.

Table 3 Tests for Hardy-Weinberg equilibrium

Locus	Homozygosity test			χ^2 test	G-T's Exact test	
	P-value	Observed homozygosity	Expected homozygosity	After allele grouping[a] *P*-value	No allele grouping *P*-value ± SE	After allele grouping[b] *P*-value ± SE
D3S1358	0.5125	0.2655	0.2833	0.9292 (3)	0.8097 ± 0.0132	
TH01	0.6161	0.3107	0.2972	0.5869 (4)	0.9328 ± 0.0067	
D21S11	0.2017	0.2599	0.2280	0.8865 (4)	0.0439 ± 0.0099[c]	0.3548 ± 0.0380 (6)
D18S51	0.6781	0.1638	0.1538	0.7038 (4)	0.6778 ± 0.0292	
Penta E	0.7461	0.0960	0.0895	0.9533 (5)	0.6010 ± 0.0315	
D5S818	0.3538	0.2486	0.2240	0.8754 (4)	0.7254 ± 0.0195	
D13S317	0.6448	0.2034	0.1914	0.8180 (4)	0.4051 ± 0.0156	
D7S820	0.8286	0.2260	0.2203	0.7218 (4)	0.3332 ± 0.0194	
D16S539	0.3086	0.1864	0.2137	0.2915 (5)	0.0873 ± 0.0099	
CSF1PO	0.5996	0.2542	0.2678	0.9573 (3)	0.5987 ± 0.0178	
Penta D	0.1926	0.2316	0.1980	0.6530 (4)	0.2575 ± 0.0236	
vWA	0.9545	0.2034	0.2019	0.7238 (5)	0.2949 ± 0.0173	
D8S1179	0.1286	0.1243	0.1617	0.1184 (4)	0.0533 ± 0.0085	
TPOX	0.2511	0.3220	0.3526	0.3957 (3)	0.3385 ± 0.0118	
FGA	0.1840	0.1751	011439	0.7613 (4)	0.2721 ± 0.0280	

[a], Alleles were pooled so that the expected values of all genotypes are more than one, and that the number of the genotypes of which the expected value are less than 5 should be less than 20% ("rule of five"). The figure in the brackets indicates the number of the allele groups after grouping. [b], Alleles containing less than 5 entries were pooled with adjacent alleles so that no alleles have less than 5 entries. The figure in the brackets indicates the number of the allele groups after grouping. [c], Significant departures from HWE, $P < 0.05$.

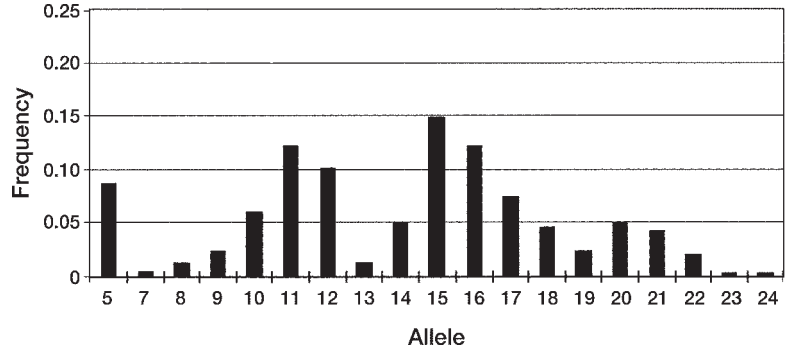

Fig. 1 Allele frequency distribution at the Penta E locus. 177 Japanese individuals were examined by the present method.

(0.14–0.18) but very rare in Caucasians and Mongoloids, including Japanese [6, 14, 15, 24, 25], was found in 3 subjects of the present population. Sequence analysis of the samples typed as having this allele, *i.e.*, types 8/14, 8/16 and 8/17, revealed that the amplified fragments of allele 8 in these samples had the same repeat structures as those reported previously on the STRbase web site (http://www.cstl.nist.gov/biotech/strbase/str_pene.htm). Another new pentanucleotide STR locus, Penta D on chromosome 21, showed a unimodal allelic distribution, and allele 9 was the most frequent at 0.297 (Fig. 2).

The allelic distributions for 14 STR loci were compared between 2 Japanese populations, *i.e.*, the inhabitants of Okayama in western Japan (present data), and those of Miyagi [6] in northeastern Japan (Table 4). Significant differences were noted in the TH01, CSF1PO,

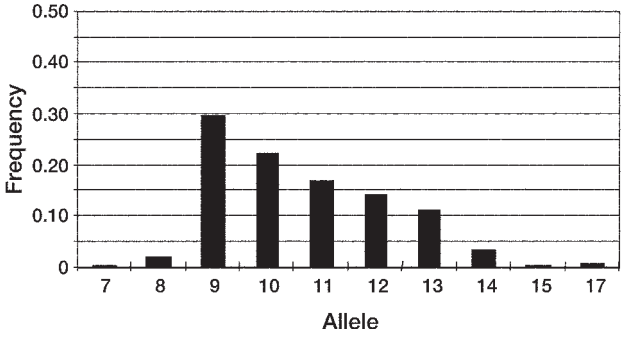

Fig. 2 Allele frequency distribution at the Penta D locus. 177 Japanese individuals were examined by the present method.

Table 4 χ^2 tests for allele frequencies of 14 STR loci between 2 Japanese populations

Locus	No. of individuals		*P*-value	χ^2	df
	Okayama (present data)	Miyagi			
D3S1358	177	323	0.0776	8.412	4
TH01	177	594	$<0.001^a$	25.639	4
D21S11	177	215	0.2314	9.306	7
D18S51	177	207	0.1223	14.001	9
Penta E	177	197	0.0908	20.179	13
D5S818	177	451	0.3215	4.682	4
D13S317	177	451	0.3036	6.027	5
D7S820	177	450	0.4176	4.986	5
D16S539	177	350	0.0946	7.918	4
CSF1PO	177	375	0.0157^b	13.990	5
vWA	177	589	0.9746	0.838	5
D8S1179	177	205	0.0036^a	19.335	6
TPOX	177	583	0.0204^b	11.617	4
FGA	177	306	$<0.001^a$	90.230	7

[a], Significant difference was noted, $P<0.01$. [b], Significant difference was noted, $0.01<P<0.05$.

D8S1179, TPOX and FGA loci; however, the differences were not highly significant in the CSF1PO and TPOX loci ($0.01 < P < 0.05$). The allele frequencies for the 15 STR loci were also compared with those for Caucasian Americans and African Americans (data not shown). Significant differences were noted in all 15 loci ($P < 0.01$). The allele frequencies for 13 out of 15 STR loci were compared with those of Chinese populations, and significant differences were noted in the TH01, D18S51, D7S820, CSF1PO, TPOX and FGA loci (Table 5).

Statistical parameters of forensic importance. Several statistical parameters of forensic importance, such as the observed allele numbers, expected and observed heterozygosity (H), polymorphism information content (PIC), power of discrimination (PD), matching probability (pM), power of exclusion (PE), and typical paternity index (Plt) values of the 15 STR loci are shown in Table 2. The Penta E locus had the highest number of alleles (19) and the TPOX locus had the lowest number of alleles (5). The observed H was highest in Penta E (0.904) and lowest in TPOX (0.678). Moreover, the PIC and PD were also highest in Penta E (PIC = 0.9036; PD = 0.9807) and lowest in TPOX, while the pM was highest in TPOX and lowest in Penta E. The 15 loci showed a combined pM of 1.3729×10^{-17} and a combined PD of $1-1.3729 \times 10^{-17}$. The cumulative power of exclusion and the typical paternity index values of the 15 loci were very high at 0.9999983 and 505,587.8, respectively.

Table 5 χ^2 tests for allele freqencies of 13 STR loci between Japanese and Chinese populations

Locus	No. of individuals			*P*-value	χ^2	df
	Okayama (present data)	Chinese	Reference			
D3S1358	177	500	18)	0.2714	5.158	4
TH01	177	500	18)	$<0.001^a$	54.293	5
D21S11	177	87	16)	0.0526	9.365	4
D18S51	177	87	16)	0.0180^b	16.914	7
D5S818	177	500	18)	0.6455	3.355	5
D13S317	177	500	18)	0.0593	10.627	5
D7S820	177	500	18)	0.0041^a	17.196	5
D16S539	177	100	17)	0.2833	5.039	4
CSF1PO	177	500	18)	$<0.001^a$	32.447	5
vWA	177	500	18)	0.6683	3.988	6
D8S1179	177	87	16)	0.5690	4.805	6
TPOX	177	500	18)	0.0167^b	12.083	4
FGA	177	500	18)	$<0.001^a$	31.273	8

[a], Significant difference was noted, $P<0.01$. [b], Signifcant difference was noted, $0.01<P<0.05$.

2. Paternity cases and zygosity diagnosis

STR typing in paternity cases. We next applied the present method of STR typing to 20 paternity cases. The results were consistent with the conclusion obtained from conventional blood group systems (ABO, MNSs, Rh-Hr, P, Kidd, Duffy, Diego, Se blood systems, a PGM1 red cell isozyme, Tf, Hp, Gc, alpha-2-HS glycoprotein serum types, and/or Acp, EsD

red cell isozymes, HLA-A, B, C and/or an X-linked Xg blood system), and 8 conventional STR systems (TH01, D5S818, D13S317, D7S820, D16S539, CSF1PO, vWA and TPOX). In the 12 non-excluded cases, the combined paternity index (PI) for the 15 STR systems ranged from 575,847 to 7,142,856,552, and these values were far superior to those obtained by the conventional systems (Table 6). In the 8 excluded cases, paternity was excluded in 6 to 10 of the present 15 STR loci, but in only 1 to 5 of the 8 conventional STR loci and in 1 to 6 of the conventional blood group systems (Table 7).

Fig. 3 shows the results of electrophoretic analysis of a non-excluded paternity case. In this case, no genetic discrepancy for the father-child relationship was observed in any of the 15 systems, and the probability of paternity and the paternity index were calculated to be 0.999999973 and 37,313,432.0, respectively. Fig. 4 shows the results for an excluded paternity case. In this case, genetic discrepancies for the father-child relationship were observed in 9 out of 15 STR loci, and the paternity was excluded decisively.

STR typing for a zygosity diagnosis of twins. The STR typing of twin sisters was performed for a zygosity determination (Fig. 5). Both twins had the same STR types in all 15 systems, and the probability of monozygosity [26] was calculated to be 99.9996 %. We could thus conclude that the sisters were monozygotic.

3. Criminal cases

STR typing from an abandoned infant and placenta. An infant corpse abandoned in a plastic container was discovered at a vacant lot beside a national road in September 199X. The infant was found to have been mature and alive at birth by autopsy. The umbilical cord was cut sharply. A placenta in a corrugated cardboard box was discovered near the infant. STR typing from blood of the infant and a piece of the umbilical cord attached to the placenta revealed the identity of the infant and the placenta, and moreover the STR types of a suspected mother were obtained from a putrefied blood clot on the maternal side of the placenta. After several days, a woman suspected to be the mother of the infant was investigated and her blood was taken for DNA typing. As shown in Fig. 6, all of the STR types of the woman were identical to those obtained from the blood

Table 6 Comparison of diagnostic potential among the examinations of the present 15 STR loci, previous 8 STR loci and conventional blood groups in 12 non-excluded paternity cases

Case	15 STR systems		8 STR systems		Blood groups	
	PRO	PI	PRO	PI	PRO	PI
090609	0.9999990541	10,571,941	0.99866765	750	0.999964	28,089
090982	0.99999999555	224,719,094	0.9974605	393	0.962995	26
091105	0.99999975475	4,077,471	0.99963938	2,772	0.820960	5
101207	0.9999997894	3,557,958	0.997117	346	0.988670	87
110809	0.99999992900	14,084,506	0.99985948	7,115	0.951180	19
111220	0.99999963459	2,736,651	0.9963779	275	0.992496	132
111221	0.99999999986	7,142,856,552	0.99999149	117,508	0.991271	114
120201	0.99999997750	44,444,443	0.99983477	6,051	0.966166	29
120330	0.99999826343	575,847	0.99929959	1,427	0.974420	38
120626a	0.9999999591	244,498,771	0.99944286	1,794	0.980150	49
120626b	0.99999994563	18,392,495	0.99802972	507	0.923470	12
130213	0.99999997320	37,313,432	0.99948924	1,957	0.977788	44

PRO, probability of paternity; PI, paternity index.

Table 7 STR typing in the samples of 8 excluded paternity cases and comparison of the number of excluded loci or systems

Case		D3S1358	THO1	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA	No. of excluded systems		
																	15 STRs[a]	8 STRs[b]	Blood groups[c]
	M	15/16	7/9	30/30	12/17	18/20	9/11	10/11	11/13	9/10	11/12	10/11	16/18	15/16	8/8	23/23			
090707	Ch	15/15	9/9	30/32.2	12/17	11/20	11/13	10/12	11/13	9/12	11/12	10/10	16/18	13/15	8/11	23/24	9	4	6/17
	AF	16/17	6/7	29/32.2	16/16	9/10	10/12	10/12	10/12	9/12	10/12	9/13	14/14	11/14	11/11	22/24			
	M	15/16	9/9.3	32/32.2	13/14	11/15	11/12	8/9	11/12	11/12	10/12	11/12	17/18	10/13	11/12	22/22			
090611	Ch	16/17	6/9	29/32	12/13	11/12	10/12	8/10	11/12	11/12	10/10	9/11	17/17	12/13	8/12	22/24	9	4	2/18
	AF	15/16	7/7	30/30	15/20	12/14	11/13	8/12	11/12	10/11	10/10	9/9	14/17	15/16	9/12	21/23			
	M	15/17	9/9	29/30	18/22	11/13	12/13	10/13	8/11	9/10	12/12	11/11	14/16	10/15	8/11	20/23			
100630	Ch	15/16	9/9	29/30	17/18	13/19	11/12	10/12	10/11	9/10	10/12	10/11	14/16	13/15	8/8	23/23	6	3	1/14
	AF	15/17	9/9	31/32.2	16/17	10/15	10/13	12/12	8/9	11/12	10/11	10/13	14/19	10/13	8/11	22/23			
	M	15/17	7/9	29/30	14/15	14/15	10/12	11/12	8/12	12/12	10/11	9/11	14/17	10/14	8/11	22/25			
110824	Ch	15/15	8/9	29/31.2	13/15	15/21	10/11	9/11	9/12	12/13	10/10	11/13	14/14	14/14	11/11	22/24	8	4	4/14
	AF	15/17	7/7	30/32	18/18	10/12	11/11	9/10	10/11	12/12	12/13	10/13	14/14	13/14	8/11	19/20			
	M	14/15	9/9	30/32	12/18	15/21	11/11	8/11	8/9	9/12	12/12	10/11	16/17	13/13	8/11	22/26			
111214	Ch	15/17	6/9	28/30	12/16	5/21	11/14	8/8	8/12	9/9	7/12	11/15	17/17	13/13	8/9	22/22	10	5	3/15
	AF	14/17	6/9	30/30	14/23	11/12	11/12	8/8	8/10	9/12	10/10	11/13	14/16	10/13	8/10	20/24			
	M	15/17	9/9	29/30	14/14	5/15	10/12	11/12	8/11	9/11	9/11	9/11	17/18	13/14	8/9	21/23			
120424	Ch	16/17	9/9	29/30	14/14	15/17	12/12	11/11	11/12	11/12	9/12	9/9	17/18	10/14	8/11	18/23	6	1	2/14
	AF	16/17	7/9	30/31	13/13	5/15	12/12	9/11	9/10	10/12	12/12	12/15	14/18	14/14	11/11	24/24			
	M	14/17	6/6	30/30	12/16	17/20	9/13	9/12	10/11	11/11	12/12	9/13	17/17	12/15	8/8	23.2/24			
120626	Ch	14/16	6/7	29/30	13/16	14/20	11/13	9/11	10/11	11/11	12/12	9/12	17/18	15/16	8/8	20/23.2	10	4	2/15
	AF	15/18	7/9	28/30	15/16	10/21	8/11	5/12	8/11	10/12	11/13	12/13	15/17	14/15	8/9	25/25			
	M	15/16	8/9.3	29/30	16/17	9/15	10/12	11/12	11/11	9/11	11/12	10/12	18/19	10/11	8/11	20/24			
121031	Ch	15/16	6/9.3	29/31	15/16	9/15	10/11	11/11	11/12	10/11	11/12	9/10	18/18	10/10	11/11	20/22	9	5	1/14
	AF	15/15	7/9	30/33.2	13/14	11/16	9/11	8/11	10/11	12/13	10/12	9/10	18/18	13/16	8/9	23/23			

Shadowed box shows the typing result which indicates the paternity exclusion. M, mother; Ch, child; AF, alleged father. [a], Number of excluded loci in the present 15 STRs. [b], Number of excluded loci in the previous 8 STRs. [c], Number of excluded systems per number of test systems in the conventional blood group systems.

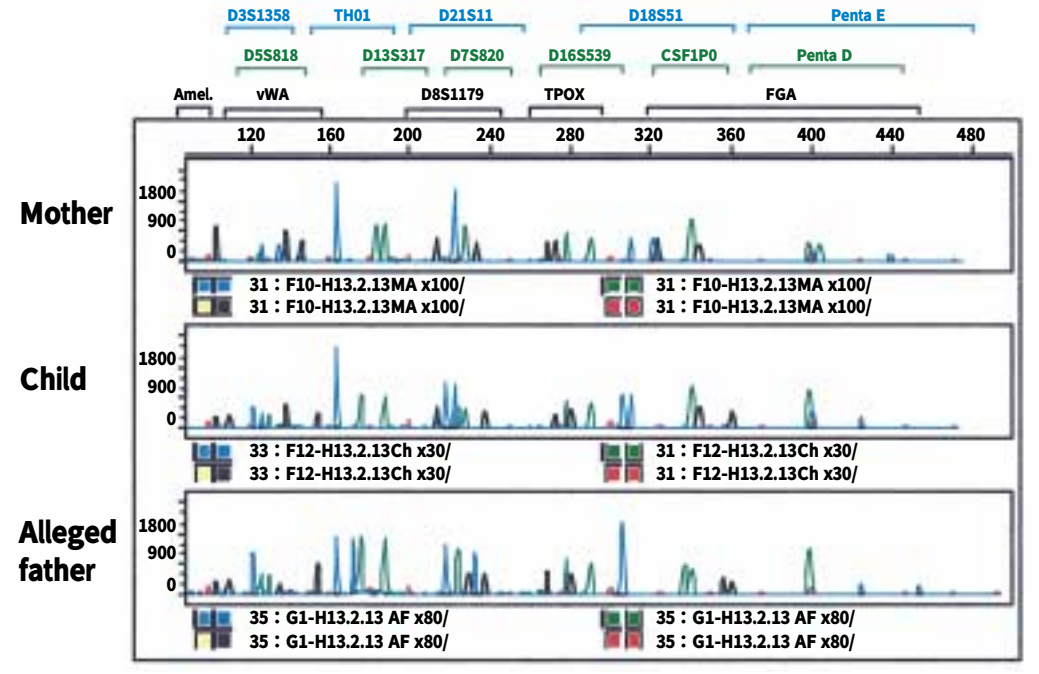

Fig. 3 STR typing for a non-excluded paternity case. PCR-amplified products from DNA samples of a child, the mother and the alleged father were analyzed by capillary electrophoresis. The X-axis indicates the base pair size of the fragments, and the Y-axis indicates relative fluorescence units (RFU).

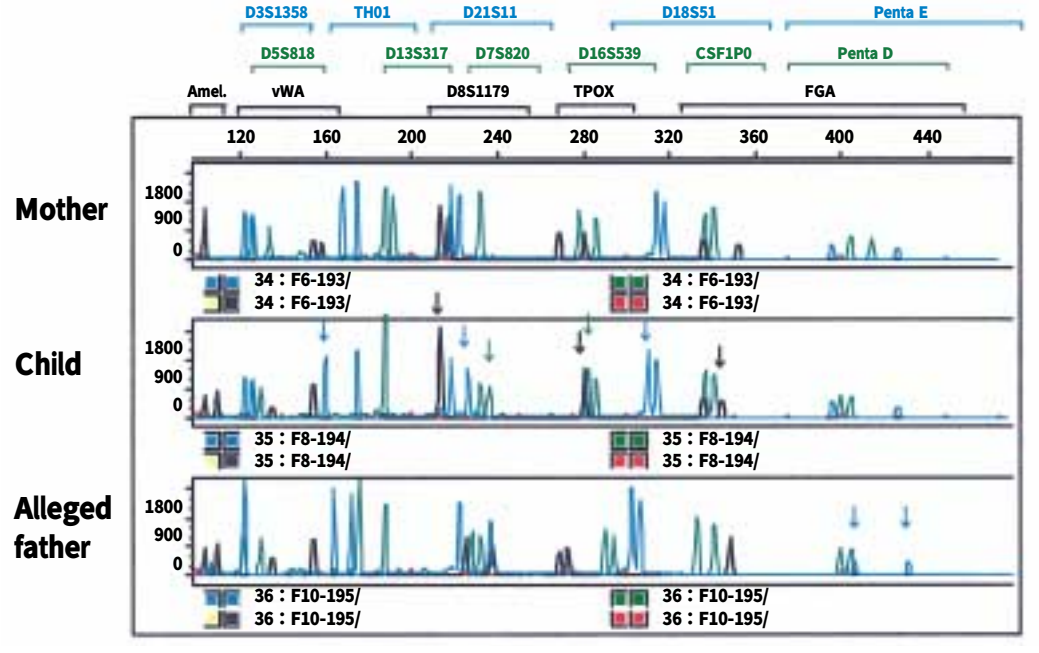

Fig. 4 STR typing for an excluded paternity case. PCR-amplified products from DNA samples of a child, the mother and the alleged father were analyzed by capillary electrophoresis. An arrow indicates the allele peak that shows the paternity exclusion. The X-axis indicates the base pair size of the fragments, and the Y-axis indicates relative fluorescence units (RFU).

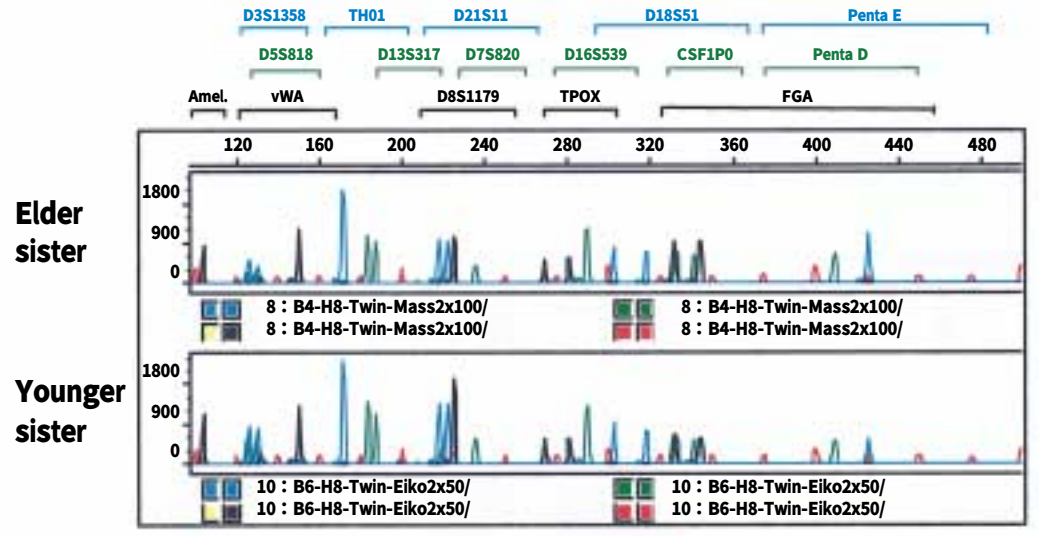

Fig. 5 STR typing for a zygosity diagnosis of twin sisters. PCR-amplified products from DNA samples of twin sisters were analyzed by capillary electrophoresis. The X-axis indicates the base pair size of the fragments, and the Y-axis indicates relative fluorescence units (RFU).

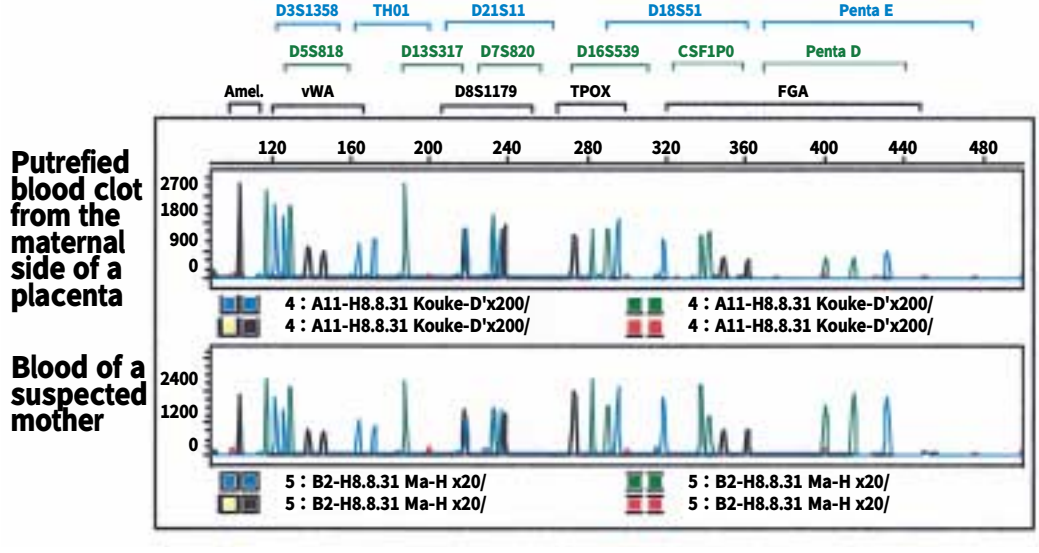

Fig. 6 STR typing of a case of an abandoned infant and placenta. PCR-amplified products from DNA samples of a putrefied blood clot from the maternal side of the placenta and the blood of a woman suspected to be the mother of the infant were analyzed by capillary electrophoresis. The X-axis indicates the base pair size of the fragments, and the Y-axis indicates relative fluorescence units (RFU).

clot of the maternal side of the placenta. The matching probability was calculated to be 1 in 7.3×10^{25}, and thus we could conclude that the woman was the mother of the infant by the present method.

STR typing of a skull immersed in water for one year. A human skull immersed in a small irrigation channel was discovered in September 199X. The skull was almost completely devoid of remaining flesh, although small masses that appeared to be decomposed cerebral tissue remained in the cranial cavity. The lower jaw was missing. From morphological findings of the skull and the teeth, it was surmised that the deceased was a 30–50 year-old male. A review of police investigations pointed to the case of a 44-year-old man who had been missing for one year. The superimposition of the skull and a photograph of the missing man supported the identity, but there was need for more conclusive evidence. DNA analysis was carried out from a molar of the upper jaw of the skull and the blood of the father of the missing man. The STR types detected in these specimens strongly suggested that the skull was that of the missing man (Fig. 7). The probability of paternity calculated between the skeletonized person and the father of the missing man was 0.9995294.

Discussion

The use of multiplex PCR in conjunction with semi-automated fluorescent typing has yielded both high productivity and high discriminatory power. In this study, we examined 15 STR systems (D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX, and FGA) in a Japanese population and estimated the allele and genotype frequencies for each. In addition, we evaluated the usefulness of these 15 STR systems in cases of forensic investigations and paternity testing.

The estimated allelic distribution for all 15 STR loci studied in this work did not deviate from the HWE proportion based on the chi-square test, the homozygosity test or G-T's exact test after the grouping of alleles so that all the grouped alleles had more than 5 events.

In all 15 loci, the allele frequency distributions of the Japanese population in the present study were significantly different from those of Caucasian American, African American or Hispanic populations ($P < 0.05$) [13–15, 24, 25] based on the chi-square test after the allele grouping (data not shown). In the loci TH01, D8S1179 and FGA, a significant difference ($P < 0.01$) was noted between the present and the Miyagi Japanese populations

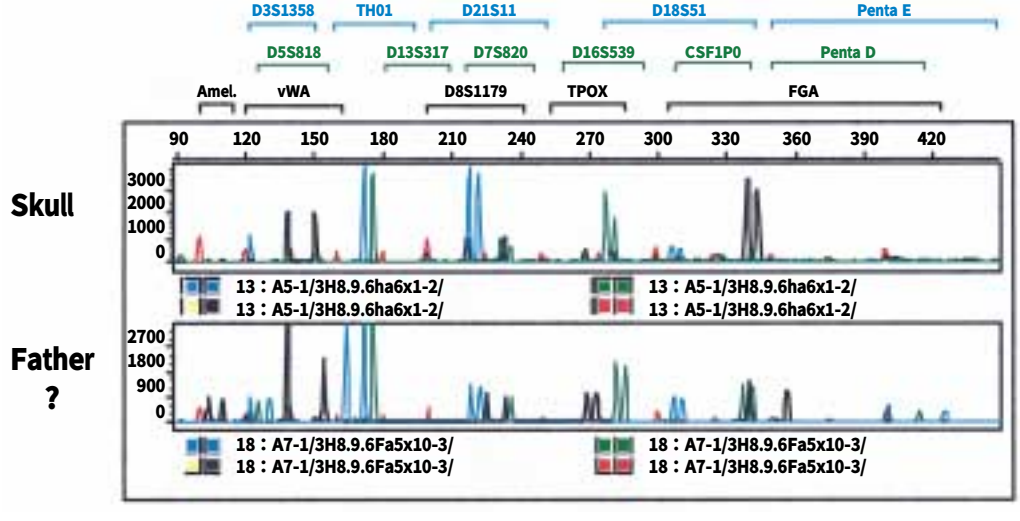

Fig. 7 STR typing for personal identification of a skull. PCR-amplified products of DNA samples from a tooth of the skull and from blood of the father of a missing man were analyzed by capillary electrophoresis. The X-axis indicates the base pair size of the fragments, and the Y-axis indicates relative fluorescence units (RFU).

[6]. From these findings, it was considered that in paternity and identity testing the allele frequency database obtained from the population to which the person in question belongs should be used in the calculations for forensic evaluations.

Statistical parameters such as the observed and expected heterozygosity values (H_{obs}, H_{exp}), the polymorphism information content (PIC), the power of discrimination (PD), and the matching probability (pM), which are of importance in applying the DNA typing method to forensic identity testing, demonstrated the usefulness of the present method in forensic practice. The most informative locus among the 15 STR systems was Penta E (PD = 0.980), while the least informative was TPOX (PD = 0.800). The systems FGA, D18S51 and D8S1179 were also highly informative. The 15 loci showed a combined PD of $1-1.3729\times10^{-17}$ and a combined pM of 1.3729×10^{-17}. Practically, this means no other individuals with the same types for all 15 loci exist in the world, except the monozygote multiples or the same clone. Moreover, both the cumulative power of exclusion and the paternity index values of the 15 loci, which are measures of diagnostic potential for paternity testing by this method, were high enough to obtain practically conclusive results.

This method was applied to the samples of paternity tests, and no deviations were detected among the diagnoses obtained by this method, those obtained by the previous 8 STR systems, and those obtained by the conventional blood groups in any of the 12 non-excluded or 8 excluded cases. Moreover, the calculated probability of paternity and the paternity index values calculated, which are of importance in applying the DNA typing method to parentage testing, were extremely high in the tests with the present 15 loci compared to the tests with the previous 8 STR systems and the conventional blood groups, and more conclusive scientific evidence was obtained by the present method.

In criminal cases, the STR analysis was also very useful for personal identification, because of its high discriminatory power. Moreover, the STR loci detected by the present method are relatively small in size (100–450 bp), and can easily be amplified by PCR even in a multiplex PCR fashion. This is an advantage when analyzing forensic specimens, which often contain severely damaged and minute DNA. We applied this method to forensic personal identification from putrefied specimens and skeletal remains. In the case of an abandoned infant and placenta, the specimens were putrefied to some extent, but DNA which could serve as the template for PCR could be recovered, and all 15 STR types of the specimens could be detected by this method. Moreover, in an examination of a skull immersed in water for one year, the soft tissues were almost completely decayed and lost, but the STR types could be detected from DNA extracted from a molar in the skull. This may have been because the DNA contained in the tooth was protected from degradation by the hard tooth tissue, and intact DNA fragments available as a template for PCR detection of the 15 STR loci remained even under these wet conditions. In conclusion, analysis of the 15 STR loci by the present multiplex PCR method was shown to be highly discriminating, and could detect STR types even from degraded specimens, suggesting it would be very useful in forensic practice.

Part of this work was presented at the 85th Congress of the Japanese Society of Legal Medicine in Kurume, April 2001, and at the 7th Indo Pacific Congress on Legal Medicine and Forensic Sciences (INPALMS-2001) in Melbourne, September 2001.

References

1. Edwards A, Civitello A, Hammond HA and Caskey CT: DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet (1991) 49: 746–756.
2. Edwards A, Hammond HA, Jin L, Caskey CT and Chakraborty R: Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics (1992) 12: 241–253.
3. Hammond HA, Jin L, Zhong Y, Caskey CT and Chakraborty R: Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet (1994) 55: 175–189.
4. Santos SM, Budowle B, Smerick JB, Keys KM and Moretti TR: Portuguese population data on the six short tandem repeat loci-CSF1PO, TPOX, THO1, D3S1358, VWA and FGA. Forensic Sci Int (1996) 83: 229–235.
5. Sparkes R, Kimpton C, Gilbard S, Carne P, Andersen J, Oldroyd N, Thomas D, Urquhart A and Gill P: The validation of a 7-locus multiplex STR test for use in forensic casework. (II), Artefacts, casework studies and success rates. Int J Legal Med (1996) 109: 195–204.
6. Hashiyada M: Short tandem repeat analysis in Japanese population. Electrophoresis (2000) 21: 347–350.
7. Technical manual of the PowerPlex™ 16 System. Promega Corporation, Madison (2001).
8. Weir BS: Independence of VNTR alleles defined as floating bins. Am J Hum Genet (1992) 51: 992–997.
9. Guo SW and Thompson EA: Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics (1992) 48: 361–372.
10. Cochran WG: Some methods for strengthening the common chi 2 test. Biometrics (1954) 110: 417–451.
11. National Research Council II Report: The Evaluation of Forensic DNA Evidence, Statistical issues -Alleles with low frequency; in National

Research Council ed, National Academy Press, Washington DC (1996) pp 148.
12. Raymond M and Rousset F: GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Heredity (1995) 86: 248–249.
13. Technical manual of the *GenePrint*® PowerPlex™ 1.2 System. Promega Corporation, Madison (1998).
14. Technical manual of the *GenePrint*® PowerPlex™ 2.1 System. Promega Corporation, Madison (1999).
15. Budowle B and Sprecher CJ: Concordance study on population database samples using the PowerPlex™ 16 kit and AmpFlSTR Profiler Plus kit and AmpFlSTR COfiler kit. J Forensic Sci (2001) 46: 637–641.
16. Gusmao L, Prata MJ, Sanchez-Diz P, Lareu MV, Carracedo A, Alves C, Martins N and Amorim A: STR data for the AmpFlSTR profiler plus loci from Macau (China). Forensic Sci Int (2001) 123: 74–75.
17. Wang B, Pang H, Koda Y, Soejima M and Kimura H: Polymorphisms of eight STR loci in Chinese and African (Xhosa) populations. Forensic Sci Int (2002) 125: 279–280.
18. Pu CE, Wu FC, Cheng CL, Wu KC, Chao CH and Li JM: DNA short tandem repeat profiling of Chinese population in Taiwan determined by using an automated sequencer. Forensic Sci Int (1998) 97: 47–51.
19. Nei M: Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics (1978) 89: 583–590.
20. Botstein D, White RL, Skolnick M and Davis RW: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet (1980) 32: 314–331.
21. Fisher RA: Standard calculations for evaluating a blood-group system. Heredity (1951) 5: 95–102.
22. Jones DA: Bloodsamples: Probability of discrimination. J Forensic Sci Soc (1972) 12: 355–359.
23. Brenner C and Morris JW: Paternity index calculations in single locus hypervariable DNA probes: varidetion and other studies; in Proceedings from the International Symposium on Human Identification 1989, Promega Corporation ed, Promega Corporation, Wisconsin (1990) pp 21–53.
24. Levadokou EN, Freeman DA, Budzynski MJ, Early BE, McElfresh KC, Schumm JW, Amin AS, Kim YK, Sprecher CJ, Krenke BE, Silva DA, McIntosh TM, Grubb JC, Johnston LJ, Sailus JS, Ban JD, Crouse CA and Nelson MS: Allele frequencies for fourteen STR loci of the PowerPlex 1.1 and 2.1 multiplex systems and Penta D locus in Caucasians, African-Americans, Hispanics, and other populations of the United States of America and Brazil. J Forensic Sci Soc (2001) 46: 736–761.
25. Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S, Brown BL, Cramer J, DeGroot GA, Douglas D, Duceman B, Eastman A, Giles R, Hamill J, Haase DJ, Janssen DW, Kupferschmid TD, Lawton T, Lemire C, Llewellyn B, Moretti T, Neves J, Palaski C, Schueler S, Sgueglia J, Sprecher C, Tomsey C and Yet D: STR primer concordance study. Forensic Sci Int (2001) 124: 47–54.
26. Akane A, Matsubara K, Shiono H, Yamada M and Nakagome Y: Diagnosis of twin zygosity by hypervariable RFLP markers. Am J Med Genet (1991) 41: 96–98.