A Validation Study of Four Single Locus Probes (MS1, MS31, MS43 and g3) in a Korean Population — Further Evaluation for Paternity Testing —

Kyung-Ok Lee, Kyu-Pum Lee, M.D.

Department of Molecular Genetics and Biochemistry Seoul Medical Science Institute,
Seoul Clinical Laboratories, Seoul, Korea

Hypervariable tandem repetitive regions in human DNA are proving to be increasingly useful for genetic analysis in humans. We chose four single locus probes(SLP; MS1, MS43, MS8 and g3) for a validation test among Koreans. The specimens were from 216 unrelated individuals and 33 paternity inclusion families. Extracted DNA from EDTA blood was restricted by HinfI and electrophoresed in 0.7% agarose gel, transferred and hybridized with chemiluminescent probes. Heterozygosity was over 90% by all of the probes. Total numbers of unassignable mutant bands from 33 paternity inclusion cases were 5, and the highest mutation rate was determined in probe MS1(0.045). The probability of having the same DNA band between two unrelated individuals was 5.7×10^{-10} when four SLPs were used at the same time. The data presented here on allele frequencies and mutation rates provide preliminary data supporting the validity of these probes in paternity analysis and forensic investigations in the Korean population.

Key Words : Single locus probe, Allele frequency, Mutation rate

INTRODUCTION

Tandem repetitive minisatellite regions of human DNA frequently show extensive allelic variation in repeat unit copy number and provide highly informative genetic markers. Since the highly polymorphic locus in human DNA was first reported(Whyman and White, 1980), a number of additional hypervariable loci have been discovered, such as the minisatellite 5' to the insulin gene(Bell et al., 1982) and the sequ-

ences 3' to the α -globin gene(Jarman et al., 1986). These loci have subsequently been shown to contain tandem repetitive sequences. If random dispersal is assumed, it would provide sufficient hypervariable probes to saturate the human linkage map with informative markers. A tandem repetitive sequence was discovered in the first intron of the human myoglobin gene(Weller et al., 1984) and used to isolate four minisatellite single locus probes(SLPs; MS1, MS31, MS43 and g3) containing tandem repeats of a common core sequence(Wong et al., 1987). Because of the sensitivity and ease of use, these SLPs have a great potential for individual identification in forensic analysis and paternity testing (Wong et al., 1987). However, before they can be adopted for routine use, the allele frequency distribution and mutation rate at these loci must be deter-

Address for correspondence : Kyung-Ok Lee, Department of Seoul Medical Science Institute, Seoul Clinical Laboratories, 7-14, 39, Dongbinggo-dong, Yongsan-gu, Seoul, 140-230, Korea.
Tel : (02) 7906-500, Fax : (02) 7906-509.

mined(Baird et al., 1986, Jeffreys and Royle, 1988, Cotton et al., 1990). In this report, we present the data on the frequency distribution of alleles, heterozygosity, probability of identity and mutation rate detected by four SLPs in a Korean population. The results provide the basis for the use of four hypervariable SLPs for individual identification in forensic analysis and in paternity determination.

MATERIALS AND METHODS

Proteinase K(20 units/mg), HinfI(20 units/ul) and DNA polymerase(klenow fragment) were obtained from Boehringer, Germany. Standard DNA(K562 DNA/HinfI, DG1961) was obtained from Promega, U.S.A. DNA probes including MS1, MS31, MS43 and g3 were purchased from Cell Mark Co, U.K. Blood

samples were collected from 216 unrelated Koreans. Samples from 33 Korean family groups were obtained from the Seoul Medical Science Institute.

DNA Preparation and Processing

Genomic DNA was isolated by phenol-chloroform-isoamyl alcohol extraction and ethanol precipitation from peripheral blood(Laborca and Paigen, 1980). The DNA pellet(10 ug) was digested with 4 ul of HinfI(20 U/ul) for 16 hours at 37°C and electrophoresed as described by Jeffreys et al(1986). Southern blotting and fixation of DNA to nylon membrane was performed according to the manufacturer's instructions. Filters were prehybridized in 500 ml of 0.5M Na₂HPO₄(pH7.2), 10% SDS for 20 min at 50°C. Membranes were hybridized with probes in the same

Fig.1.(A) Allele distribution of SLP MS1

Fig.1.(B) Allele distribution of SLP MS31

Fig.1.(C) Allele distribution of SLP MS43

Fig.1.(D) Allele distribution of SLP g3

Fig. 1. Distribution of allele sizes of four SLPs. (A) probe MS1, (B) probe MS31, (C) probe MS43, (D) probe g3, DNAs from 216 unrelated Koreans were digested with Hinf I and fractionated by electrophoresis in a 0.7% agarose gel. After transfer to nylon membranes, hybridization to probe and allele length measurement was carried out as described in the methods section.

solution. Luminography was performed by adding 0.3mM MgOAC Lumi-phos to the membrane and exposed for 4 hrs at 30°C in the presence of X-ray sensitive film. The luminograms were scanned with a Laser scanner system(Bio-Med Co, U.S.A). The size of the alleles was calculated from the internal calibration curve produced by RFLP gene analysis program (Bio-Med Co, U.S.A) with the DNA molecular weight markers on the same luminogram. To have a conservative estimation of the allele frequency, the alleles were classified into 34 bins which are demarcated by the more or less evenly spaced molecular weight markers.

RESULTS

Allele Frequency and Probability of Identity

Allele sizes and the allele frequency distributions were determined for four hypervariable loci detected by probes(MS1, MS31, MS43 and g3) among 216 unrelated Koreans. The results were summarized in Fig. 1 and 2. The range of frequent alleles among four SLPs was from 4.0 kb to 6.0 kb(Fig.1). Most alleles of MS31 were distributed to the relatively narrow range of 2.7 kb~13.3 kb, whereas alleles of MS1 varied in size from 2.5 kb~18.9 kb(Table 1). The most frequent alleles of probe MS1, MS31, MS43 and g3 were 4.0 kb, 5.8 kb, 4.5 kb and 6.2 kb

Table 1. Summary of four hypervariable locus specific probes.

Probe	Chromosome*	Chromosomal localization	Size range of alleles detected
MS1	1p33-p35	D1S7	2.5kb~18.9kb
MS31	7p22-pter	D7S21	2.7kb~13.3kb
MS43	12q24.3-qter	D12S11	3.0kb~18.2kb
g3	7p36 qter	D7S22	2.6kb~15.2kb

*Wong Z et al., 1987

Table 2. Allele size and frequency distribution determined in 216 unrelated Korean individuals.

Probe	Hetero-zygosity	No. of allele resolved		most frequent allele		
		0~16kb	>16kb	Size(kb)	Frequency(qm)	$q^2(2-q)^*$
MS1	94.3	402	6	4.0	0.057	6.3×10^{-3}
MS31	96.9	418	1	5.8	0.031	1.9×10^{-3}
MS43	95.7	411	3	4.5	0.044	3.7×10^{-3}
g3	91.7	398	1	6.2	0.083	1.3×10^{-2}

* $q^2(2-q)$ =The probability of 2 unrelated individuals possessing the same genotypes, assuming equal allele frequencies where q =the mean allele frequency=(1-heterozygosity)(Jeffrey et al., 1988).

Fig. 2. Highly polymorphic allele patterns in 15 persons (Sample No.1~15) by using four single locus probes.

M: Molecular weight marker(λ DNA/HindIII)(A) probe MS1, (B) probe MS31, (C) probe MS43, (D) probe g3.

respectively(Table 2). Accurate values for fragment lengths were obtained only in the range 1.0 kb~16.0 kb. Alleles of length >16kb were combined in a single group, which represented less than 2% of the alleles detected with all of the four SLPs(Table 2). All loci were highly variable, with heterozygosities ranging from 92% for g3 to 97% for MS31(Table 2). The probability(P) of two individuals possessing the same genotype ranges from 6.3×10^{-3} for MS1 to 1.3×10^{-2} for g3(Table 2). When the above four polymorphic probes are used at the same time, the probability of chance association is decreased to 5.7×10^{-10} . Table 3 showed the binned allele frequency distribution at four single locus probes.

Table 3. Binned allele frequency distribution at four single locus probes.

Bin	Range (Kbp)	MS1		MS31		MS43		g3	
		No.of Alleles	Allele Frequency	No.of Alleles	Allele Frequency	No.of Alleles	Allele Frequency	No.of Alleles	Allele Frequency
1	2.1~2.5	2	0.005	0	0.000	0	0.000	0	0.000
2	2.6~3.0	30	0.075	4	0.010	3	0.007	7	0.018
3	3.1~3.5	48	0.119	29	0.069	2	0.005	27	0.068
4	3.6~4.0	58	0.145	63	0.151	18	0.044	42	0.106
5	4.1~4.5	35	0.087	13	0.031	90	0.219	36	0.090
6	4.6~5.0	48	0.119	29	0.069	36	0.188	32	0.080
7	5.1~5.5	56	0.139	57	0.136	23	0.056	33	0.083
8	5.6~6.0	17	0.042	77	0.184	4	0.010	42	0.106
9	6.1~6.5	18	0.045	52	0.124	13	0.032	47	0.118
10	6.6~7.0	18	0.045	27	0.065	15	0.036	30	0.075
11	7.1~7.5	15	0.037	19	0.045	21	0.051	18	0.045
12	7.6~8.0	16	0.040	13	0.031	12	0.029	20	0.050
13	8.1~8.5	5	0.012	14	0.033	27	0.066	14	0.035
14	8.6~9.0	4	0.010	9	0.022	15	0.036	15	0.038
15	9.1~9.5	8	0.020	5	0.012	17	0.041	5	0.013
16	9.6~10.0	2	0.005	3	0.007	9	0.022	6	0.015
17	10.1~10.5	5	0.012	1	0.002	9	0.022	4	0.010
18	10.6~11.0	2	0.005	1	0.002	8	0.019	4	0.010
19	11.1~11.5	2	0.005	0	0.000	8	0.019	5	0.013
20	11.6~12.0	1	0.002	0	0.000	21	0.051	3	0.008
21	12.1~12.5	4	0.010	0	0.000	9	0.022	3	0.008
22	12.6~13.0	1	0.002	1	0.002	18	0.044	2	0.005
23	13.1~13.5	0	0.000	1	0.002	13	0.032	0	0.000
24	13.6~14.0	2	0.005	0	0.000	4	0.010	0	0.000
25	14.1~14.5	0	0.000	0	0.000	4	0.010	0	0.000
26	14.6~15.0	1	0.002	0	0.000	7	0.017	0	0.000
27	15.1~15.5	0	0.000	0	0.000	0	0.000	2	0.005
28	15.6~16.0	0	0.000	0	0.000	3	0.007	1	0.003
29	16.1~16.5	0	0.000	0	0.000	0	0.000	0	0.000
30	16.6~17.0	2	0.005	0	0.000	0	0.000	0	0.000
31	17.1~17.5	0	0.000	0	0.000	0	0.000	0	0.000
32	17.6~18.0	0	0.000	0	0.000	0	0.000	0	0.000
33	18.1~18.5	0	0.000	0	0.000	2	0.000	0	0.000
34	18.6~19.0	2	0.005	0	0.000	0	0.000	0	0.000

Mutation rate

Prior to applying in the analysis of family relationships, the mutation rate of these polymorphic loci must be determined. We have studied the mutation rate in 33 paternity inclusion families. The observed mutation rate is shown in Table 3. The highest mutation rate was shown by MS1(0.045), and probe MS31 did not have any mutant bands(Table 4A). The frequency of mutant bands among 33 cases of paternity inclusion was 15.2% (Table 4B) and the 5 mutations detected by SLPs were distributed between those of paternal and those of maternal origin(Table 4C).

Application to paternity

In this study, 33 paternity inclusion cases correlated well with multilocus probes procedures reported previously(Lee et al., 1994). Examples of paternity analysis were shown in Fig. 3 using SLP mix A (MS1, MS31 and g3). Family e and f (Fig.3) were the cases of paternity exclusion, where none of the non-maternal bands in the child matched those of the alleged father.

Table 4. Incidence of mutant bands in offspring in the paternity inclusion 33 cases(66 offsprings).

A. Observed mutation rate				
SLP probes	No of mutant alleles observed	Mutation rate /gamete	Previous estimate of mutation rate	
MS1	3/66	0.045	0.034(10/298) ^{a)}	
MS31	0/66	0	0.007(2/298)	
MS43	1/66	0.016	0 (0/298)	
g3	1/66	0.016	0 (0/298)	
B. Frequency of mutant band				
No of mutant bands in offspring	No. of cases	% of cases		
0	28	84.8%		
1	5	15.2%		
C. Origin of Mutant band				
	Parental origin of mutant band			Total
	Maternal	Paternal	Not determinable ^{b)}	
Total no. of mutant bands				5
No. detected by MS1	2	1	0	3
No. detected by MS31	0	0	0	0
No. detected by MS43	1	0	0	1
No. detected by g3	1	0	0	1

Note a) : Smith et al., 1990b

b) : Mother and father shared an indistinguishable allele which was transmitted to the child.

In part B, data for probes MS1, MS31, MS43 and g3 are combined. Part C describes the parental origin of mutant band.

Fig. 3. Use of SLP mix(MS1, MS31 and g3) in paternity testing among Koreans.

Case a,b,c,d : Paternity inclusion

Case e,f : Paternity exclusion

M : mother, C : child, F : father, S : Standard

When paternity tests were examined by using SLPs, the bands are examined to determine whether the alleged father can be excluded as the biological father of the child. First, the mother and child bands are examined to determine which bands they share. When one band in the adjacent lanes of the mother and child is aligned at the same kb position, then the maternal band in the child has been identified. The other band in the child, therefore, must be the paternal band, and inspection of the bands in the lane of the alleged father will reveal whether the alleged father has a band of the appropriate size. When the lane of the alleged father does not have a band that lines up with the paternal band in the lane of the child, then the alleged father is excluded as being the biological father of the child.

DISCUSSION

Tandem repeats(TRs), also known as satellite DNAs tend to appear at single loci on the chromosome(Armour et al., 1989). These regions in vertebrate DNA frequently show high levels of allelic variability in the number of repeat units(Bell et al., 1982). The TRs may be short(microsatellite) or long(minisatellite). The number of TRs found at a given locus also varies among ethnic groups and is genetically determined on each of the two homologous chromosomes of the individual(Walker, 1992). These highly polymorphic DNA regions show restriction fragment length polymorphism(RFLP) due to allelic variation in the number of repeated units. Because of the large number of alleles, yet it's low frequency of identity, these repetitive DNA areas provide highly informative genetic markers(Helminen et al., 1988, Logtenberg and Bakker, 1988). The four hypervariable loci in human DNA have been cloned and characterized(Wong et al., 1986). All of these four cloned minisatellites act as single locus probes-MS1, MS31, MS43 and g3. The chromosomal localization and repeat unit sequences of the above four probes have been described(Wong et al., 1987)(Table 1). Allele frequencies at the hypervariable loci detected by these probes were already

Table 5. Heterozygosity(%) compared with four ethnic groups.

Probe	Korean (n=216)	Japan ^{a)} (n=306)	Caucasian ^{b)} (n=200)	Afro-Caribbean ^{c)} (n=300)	Asian ^{c)} (n=300)
MS1	94.3	96.6	99.0	96.7	95.0
MS31	96.9	92.1	97.0	94.9	95.4
MS43	95.7	92.6	96.0	96.3	90.7
g3	91.7	93.6	96.0	95.8	94.5

a) : Kamiyama, 1993

b) : Smith et al., 1990b

c) : Helminen et al., 1988

determined in several ethnic groups and proved to be clinical usefulness for individual identification and paternity testing (Smiths et al., 1990a, Kamiyama, 1993, Lee et al., 1993). In this experiment, we determined population genetic data, including allele frequency distribution, heterozygosity, probability of identity and mutation rate, for four SLPs before using them in routine clinical work among Koreans. The reproducibility of measurements was determined by including an aliquot of HinfI-digested DNA from a single individual on each gel. The heterozygosities of four SLPs are compared with several ethnic groups in Table 5. Probe MS31 showed the highest heterozygosity (96.89%) among Koreans, whereas probe MS1 had the highest value in Caucasians (99%) and Japanese (96.6%) (Kamiyama, 1993) (Table 5). The degree of individual specificity of genotypes determined by SLPs can be estimated from the heterozygosity H at each locus. The mean allele frequency q at a locus is given by $(1-H)$, and, assuming that all alleles share the same population frequency q , then the probability that two randomly selected individuals share the same genotype can be estimated as $q^2(2-q)$ (Wong et al., 1987). The probability of four SLPs varies from 1.0×10^{-2} to 6×10^{-3} (Table 2). The cumulative probability for all four SLPs is 5.7×10^{-10} , which is higher than the combined value obtained with the multilocus probes among Koreans (4.4×10^{-15}) (Lee et al., 1994). We have determined the mutation rate in 33 paternity inclusion families. The highest mutation rate was shown by MS1 (0.045) (Table 4), it was slightly higher than previous reports (0.034) (Smith et al., 1990b). The frequency of mutant bands (15.6%) among 33 cases of paternity inclusion was lower than that of the result analyzed by multilocus probes (20.02%) (Jeffreys et al., 1986). The data on allele frequencies and mutation rates presented in this study provide a foundation for the use of the

panel of hypervariable probes in a Korean populations. The panel of SLPs described in this study has been used to establish paternity in 33 Korean families where paternity had already been determined by multilocus DNA fingerprinting. The results of that comparison correlated well with the previous data (Lee et al., 1994). In short, SLPs are hypervariable, have low mutation rates and have high heterozygosity. Therefore we believe these four SLPs can be applied to bone marrow transplantation analysis (Knowlton et al., 1986), zygosity testing (Motomura et al., 1987), detection of tumors (Thein et al., 1987), individual identification in forensic medicine (Gill et al., 1985) as well as determination of paternity (Helmin et al., 1988) among Koreans.

ACKNOWLEDGMENTS

This work was supported by a grant from the Seoul Medical Science Institute. The authors would like to thank Moo-Joo Lee for his technical assistance.

REFERENCES

- Armour JAL, Wong Z, Wilson V, Royle NJ, Jeffreys AJ. Sequence flanking the repeat arrays of human minisatellites: association with tandem and dispersed repeat elements. *Nucleic Acid Res* 1989; 17: 4925-35.
- Baird M, Balaz I, Giusti A, Miyasaki L, Nicholas L, Wexier K, Kanter E, Glassberg J, Allen F, Rubinstein P, Sussman L. Allele frequency distribution of two highly polymorphic DNA sequences in ethnic groups and its application to the determination of paternity. *Am J Hum Genet* 1986; 39: 489-501.
- Bell GI, Selby MJ, Rutter WJ. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 1982; 295: 31-5.
- Cotton RW, Anderson MB, Herrin G, Corey AC, Sheridan KT, Tonelli LA, Waskowski CA, Garner DD. Current case

- experience with single locus hypervariable probes. *J Forensic Sci*: in press. 1990.
- Gill P, Jeffreys AJ, Werrett DJ. Forensic application of DNA fingerprints. *Nature* 1985; 318: 577-9.
- Helminen P, Ehnholm C, Likki ML, Jeffreys AJ, Peltonen L. Application of DNA fingerprinting to paternity determination. *Lancet* 1988; 574-6.
- Jarman A, Nicholls RD, Weatherall DJ, Clegg JB, Higgs DR. Molecular characterization of a hypervariable region downstream of the human α globin gene cluster. *J Embryol Exp Morphol* 1986; 5: 1857-63.
- Jeffreys AJ, Wilson V, Thein SL. DNA fingerprints and segregation analysis of multiple markers in Human pedigrees. *Am J Hum Genet* 1986; 39: 11-24.
- Jeffreys AJ, Royle NJ. Spontaneous mutation rates to new length alleles at tandem repetitive hypervariable loci in human DNA. *Nature* 1988; 332: 278-81.
- Kamiyama. The allele frequency using single locus probes in Japanese population. *The 76th Jap Soci of forensic sci (abs)* 1993; 253.
- Knowlton RG, Brown VA, Braman JC, Barker D, Schumm JW, Takvorian T, Ritz J, Donnis-Keller H. Use of highly polymorphic DNA probes for genotype analysis following bone marrow transplantation. *Blood* 1986; 68: 378-85.
- Laborca C, Paigen K. A simple rapid and sensitive DNA assay procedure. *Anal Biochem* 1980; 102: 344-52.
- Lee KP, Lee KO, Choi CS, Lee MJ. Validation study of multilocus DNA fingerprint probes for individualization and paternity test in Koreans. *Korean J Clin Pathol* 1994; 14: 123-33.
- Lee KO, Lee MJ, Park YS, Choi CS, Lee KP. A study on paternity test using DNA fingerprinting by four single locus probes in Koreans. *Korean Soci Med Genet(abst)* 1993; 289.
- Logtenberg H, Bakker E. The DNA fingerprint. *Nature* 1988; 12: 28-33.
- Motomura K, Tateishi H, Nishisho I, Okazaki M, Miki T, Tonomura A, Takai S, Mori T, Jeffreys AJ. Zygosity determination of Japanese twins using a minisatellite core probe. *Jap J Hum Genet* 1987; 32: 9-14.
- Smith JC, Newton CR, Alves A, Anwar R, Jenner D, Markham AF. Highly polymorphic minisatellite DNA probes. *J Forensic Sci* 1990(a); 30: 3-18.
- Smith JC, Anwar R, Riley J, Jenner D, Markham AF, Jeffreys AJ. Highly polymorphic minisatellite sequences. *J Forensic Sci* 1990(b); 30: 19-32.
- Thein SL, Jeffreys AJ, Good HC, Cotter F, Flint J, O'Conner NT, Wainscoat JS. Detection of somatic changes in human cancer DNA by DNA fingerprint analysis. *Br J Cancer* 1987; 55: 353-6.
- Walker HR. Molecular biology in paternity testing. *Lab Med* 1992; 23: 752-6.
- Weller P, Jeffreys AJ, Wilson V, Blanchetot A. Organization of the human immunoglobulin gene. *J Embryol Exp Morphol* 1984; 3: 439-46.
- Whyman AR, White R. A highly polymorphic locus in human DNA. *Proc Natl Acad Sci* 1980; 77: 6754-8.
- Wong Z, Wilson V, Jeffreys AJ, Thein SL. Cloning a selected fragment from a human DNA "fingerprint". Isolation of an extremely polymorphic minisatellite. *Nucleic Acids Res* 1986; 14: 4605-46.
- Wong Z, Wilson V, Patel I, Povey S, Jeffreys AJ. Characterisation of a panel of highly variable minisatellites cloned from human DNA. *Ann Hum Genet* 1987; 51: 269-88.