Int J Legal Med (2002) 116:317–321
DOI 10.1007/s00414-002-0307-x

ORIGINAL ARTICLE

Jin Young Son · Yoon Seong Lee · Chong Min Choung
Soong Deok Lee

Polymorphism of nine X chromosomal STR loci in Koreans

Received: 12 November 2001 / Accepted: 20 March 2002 / Published online: 28 August 2002

Abstract This study describes the polymorphism of the nine STR loci on the X chromosome, DXS6803, DXS8378, GATA164A09, DXS7132, DXS7133, DXS9895, DXS9898, DXS6789, and DXS6795 in Koreans. In each locus, 4–10 alleles were noted and the allelic distribution patterns were the same for males and females. Heterozygosity in females ranged from 0.42 to 0.84. Among the 303 father-daughter or mother-child pairs examined 29 cases of mutation were found, 13 at the DXS6803 locus, 2 at DXS8378, 4 at DXS164A09, 3 at DXS7132, 1 at DXS7133, 2 at DXS9895, 2 at DXS9898 , 1 at DXS6789 and 1 at DXS6795. In 208 families including 180 fathers and 177 mothers, 530 different haplotypes were found. Unlike the STR loci on the Y chromosome, cases showing recombination were frequent, and in combination with mutation this made it difficult to discriminate the exclusion cases from those with mutation or recombination based on the haplotype. Details of X chromosomal STRs in Koreans which would be useful for a future large scale database are described.

Keywords Polymorphism · X-chromosome · STR · Koreans · Haplotype · Recombination

Introduction

Many autosomal STRs have been introduced for individual identification in forensic applications, and have become invaluable tools (Urquhart et al. 1995; Brinkmann 1998). Several STR loci on the Y chromosome have been recently studied, and their usefulness, for example in crimes of violence such as rape or in terms of phylogenic analysis through paternal lineage, seems to be increasing (de Knijff et al. 1997). However, the number of X chromosome markers is limited (Kishida and Tamaki 1997; Kishida et al. 1997; Szibor et al. 2000).

The usefulness of the STR loci on the X chromosome for forensic purposes seems to be restricted, but may be valuable in several specific situations, for example to confirm a grandmother-grandchild relationship. As males are hemizygous for the X chromosome, the MEC (mean exclusion chance) may be higher than diploid autosomal STRs (Hering and Szibor 2000). In paternity testing the X chromosomal STRs can also be used when the disputed child is a girl. In this article nine STR loci located on the X chromosome, DXS6803 (accession ID; GDB366184), DXS8378 (accession ID; GDB683439), CHLC-GATA164A09, DXS7132 (accession ID; GDB456721), DXS7133 (accession ID; GDB456765), DXS9895 (accession ID; GDB9799097), DXS9898 (accession ID; GDB1298173), DXS6789 (accession ID; GDB364689) and DXS6795 (accession ID; GDB365041) were studied to validate their usefulness in forensic applications.

Materials and methods

DNA was extracted from blood samples taken from 208 unrelated Korean families by phenol-chloroform extraction. The families included were mainly trios or deficient father-child or mother-child cases, and 180 fathers and 177 mothers were studied. A total of 303 meioses, 103 father to daughter transmissions, 91 mother to daughter transmissions and 109 mother to son transmissions were examined. The majority of parents ,110 out of 180 fathers (61.1%) and 171 out of 177 mothers (96.66%), screened were between 20 and 40 years of age. Family relationships were confirmed using autosomal STR loci. After determining the DNA concentration using a DyNA Quant 200 (Hoefer Scientific Instruments, San Francisco, Calif.), each loci was amplified. DXS6803, DXS8378, GATA164A09, and DXS7132 were amplified together in one quadruplex PCR (X-multi I), DXS7133, DXS9895 and DXS9898 were amplified together in one triplex PCR (X-multi II), and DXS6789 and DXS6795 were amplified in one duplex PCR (X-multi III) reaction. No overlap between different loci within a same multiplex reaction was observed. PCR amplification were performed using sequences based on Genbank information, with

J.Y. Son · Y.S. Lee · C.M. Choung · S.D. Lee (✉)
Department of Forensic Medicine,
Seoul National University College of Medicine,
28 Yeongon-dong, Chongno-gu, Seoul 110-799 Korea
e-mail: sdlee@snu.ac.kr,
Tel.: +82-2-7408353, Fax: +82-2-7648340

Table 1 Allele frequencies observed for the nine STR loci on the X chromosome in the Korean population (The number within parenthesis in the bottom column of each locus represents the largest allele noted and its size in bp)

Allele	Locus																	
	DXS6803		DXS8378		GATA164A9		DXS7132		DXS7133		DXS9895		DXS9898		DXS6789		DXS6795	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
6.3	–	–	–	–	–	–	–	–	–	–	–	–	0.04	0.02	–	–	–	–
8	–	–	0.02	0.02	–	–	–	–	–	–	–	–	–	0.01	–	–	–	–
9	–	–	0.56	0.59	–	–	–	–	0.72	0.71	–	–	0.06	0.07	–	–	0.03	0.02
10	–	0.00	0.28	0.24	–	0.01	–	–	0.26	0.26	0.3	0.3	0.55	0.52	–	–	0.15	0.13
11	0.12	0.11	0.12	0.13	0.2	0.17	0.01	0.00	0.02	0.04	0.39	0.28	0.27	0.32	–	–	0.35	0.31
11.3	0.13	0.16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
12	0.1	0.1	0.02	0.02	0.12	0.18	0.05	0.09	–	–	0.17	0.22	0.09	0.05	–	–	0.04	0.06
12.1	0.01	–	(12; 212)		–	–	–	–	–	–	–	–	–	–	–	–	–	–
12.3	0.56	0.56	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13	–	–	–	–	0.06	0.06	0.16	0.17	0.01	–	0.14	0.18	0.01	0.01	–	–	0.4	0.47
13.3	0.08	0.07	–	–	0.01	0.01	–	–	(13; 140)		–	–	–	–	–	–	–	–
14	–	0.01	–	–	–	–	0.39	0.36	–	–	0.01	0.02	–	0.00	0.01	0.01	0.04	0.01
14.3	(14; 128)		–	–	0.12	0.08	–	–	–	–	(14; 159)		(14; 218)		–	–	–	–
15	–	–	–	–	–	–	0.31	0.28	–	–	–	–	–	–	0.16	0.16	–	0.01
15.3	–	–	–	–	0.36	0.38	–	–	–	–	–	–	–	–	–	–	(15; 294)	
16	–	–	–	–	–	–	0.08	0.09	–	–	–	–	–	–	0.26	0.23	–	–
16.3	–	–	–	–	0.12	0.1	–	–	–	–	–	–	–	–	–	–	–	–
17	–	–	–	–	–	–	0.01	0.01	–	–	–	–	–	–	0.05	0.04	–	–
17.3	–	–	–	–	0.01	0.01	(17; 300)		–	–	–	–	–	–	–	–	–	–
18	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
18.3	–	–	–	–	–	0.00	–	–	–	–	–	–	–	–	–	–	–	–
19	–	–	–	–	(18.3, 262)		–	–	–	–	–	–	–	–	0.05	0.03	–	–
20	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.21	0.2	–	–
21	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.19	0.23	–	–
22	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.07	0.06	–	–
23	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.01	0.03	–	–
24	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.01	–	–
	–	–	–	–	–	–	–	–	–	–	–	–	–	–	(24; 153)		–	–

the exception of the GATA164A09 locus. The primer sequences used for this locus were as follows:

- Primer I: 5' – AGA GCC CAT GGA GCT ATC TT – 3'
- Primer II: 5' – CTC AGT TCA AGG AAA TGG GA – 3'

The reaction conditions used were 30 cycles of 95°C for 30 s, 56°C for 30 s and 72°C for 30 s for X-multi I, 30 cycles of 95°C for 30 s, 58°C for 30 s and 72°C for 30 s for X-multi II and X-multi III. PCR products were separated in 8 M urea, 5% denaturing PAGE gels and DNA fragments were visualized by silver staining and typed using home-made reference allelic ladders. Analysis of the exact size and sequence of the various alleles was performed by direct sequencing of the PCR products as follows: after amplification, the products were confirmed using 1.2% agarose gels and the gel fragments containing the exact band were excised. The DNA was then extracted using a QIAEX II gel extraction kit (Qiagen, Germany) and the eluates were used as sequencing templates. Direct sequencing was performed using the Dye Terminator Cycle sequencing kit (PE Biosystems, Foster City, Calif.) following the manufacturer's recommendations. At least two examples of each allele found were sequenced.

Results and discussion

For the DXS6803 locus, alleles with only one base pair difference needed repeated electrophoresis for exact typing. Initially the X-multi II reaction was designed as a quadruplex PCR, which additionally contained DXS6807 (GDB366257), but it proved difficult to amplify DXS6807 simultaneously and this locus was excluded. The X-multi III reaction was also designed as a quadruplex PCR containing DXS6806 and GATA144D04 (GDB9799178) or GATA31E08. Multiple bands were amplified in cases of GATA144D04 and GATA31E08 loci, and therefore these loci were excluded. DXS6806 was found to be composed of a complex repeat unit structure – $[(C)TAT]_n$ with different locations of the trinucleotide repeat unit –, which made the exact sizing difficult (data not shown), therefore this locus was also excluded from the analysis.

For each locus 5–11 alleles were noted (Table 1), and DXS6789 proved to be the most polymorphic locus with 11 alleles. No differences in the allelic distribution patterns for male and female subjects were found. Among females 4–35 different genotypes, the number of which was locus dependent, were noted with a heterozygosity ranging from 0.43 to 0.82. Several statistical parameters for each locus are listed in Table 2. Some reports are available on X chromosomal STRs studied in this report (Duo et al. 1998; Hering and Szibor 2000; Hering et al. 2001). No re-

Table 2 Several statistical parameters for each locus (*H* observed homozygosity, *h* observed heterozygosity, *PE* paternity exclusion chance according to Desmarais et al. 1998, *p*-value for Hardy-Weinberg equilibrium with exact test from Guo and Thompson 1992 , *PD* power of discrimination from Jones 1972)

Locus	Statistical parameters				
	H	h	PE	*p*-value	PD
DXS6803	0.38	0.62	0.34	0.14	0.83
DXS8378	0.43	0.57	0.15	0.17	0.76
GATA164A09	0.23	0.77	0.33	0.02	0.92
DXS7132	0.31	0.69	0.29	0.37	0.89
DXS7133	0.42	0.58	0.11	0.56	0.61
DXS9895	0.27	0.73	0.29	<0.01	0.89
DXS9898	0.33	0.67	0.17	0.99	0.79
DXS6789	0.16	0.84	0.41	0.01	0.84
DXS6795	0.36	0.64	0.20	0.99	0.83

markable population differences were found between Koreans and other populations studied.

In total 29 cases of mutation were identified in 25 families covering all loci examined (Table 3) and the DXS6803 locus showed the highest mutation rate. Mutations usually occurred at one locus in a family, but in two families mutations were noted in three loci simultaneously. In two cases, both paternal and maternal alleles were transmitted abnormally. Most of the mutated cases were maternally derived and deletion-type mutations were slightly more prevalent than the insertion type. The number of mutations found was 13 in DXS6803, 2 in DXS8378, 4 in GATA164A09, 3 in DXS7132, 2 in DXS7133, 1 in DXS9895, 2 in DXS9898, 1 in DXS6789, and 1 in DXS6795.

The nine loci are on the same chromosome and it is safe to use haplotype rather than the usual multiplication rule for statistical analysis. A total of 531 different haplotypes were noted, but few individuals shared the same haplotype. Many cases of recombination were also noted (Table 4), i.e. parents transmitted one haplotype to their first child and another haplotype – a haplotype with a different combination of alleles, but not a mutation – to the second child. This fits well with the previous reports that some of the loci are not linked and are transmitted independently (Edelmann and Szibor 2001). These were noted in all the families with multiple children. Recombination occurred for every locus and no hot spot within different loci was apparent. It would appear that discrimination between the exclusion case and the inclusion case with mutation or recombination is very difficult with X-chromosomal haplotypes only.

Although the total number of available polymorphic loci on the X chromosome may be low, the polymorphism of an individual STR locus seems comparable to that of autosomal STRs and for 1 locus up to 36 different alleles were reported (Watanabe et al. 2000). Y chromosomal STRs are transmitted rather strictly as haplotypes and a supportive statistical analysis method for mutations was suggested (Rolf et al.2001). Contrary to this, the high mutation or recombination rates in the X chromosomal STRs may be an obstacle to statistical analysis when using X chromosomal STRs. Free recombination between two X chromosomes as occurs between autosomes during cell division is probably the cause. As several STRs on the X chromosome are closely linked within a short distance, the concept of haplotypes should be used for the statistical considerations, although it looks conservative. More detailed information, such as the recombination rate within different loci sets would be necessary in order to use individual statistical parameters instead of haplotypes.

In some deficiency cases, tests for the X chromosome may be the only method available. Therefore it is important that a validation study of the STR loci on the X chromosome be conducted before the method is used practi-

Table 3 Mutation patterns noted in this study (The sex of the child and the age of the parents are shown in parentheses of Family No. and Origin columns. If the origin of the mutation could not be determined, the age of both father and mother are shown in order, *f* female, *m* male, *P* paternal, *M* maternal, *D* decrease of repetition number, *I* increase of repetition number, *U* undetermined)

Family No.	Mutated locus	Origin	Type	Family No	Mutated locus	Origin	Type
F506 (f)	DXS7132	P (31)	D (–1)	F758 (f)	DXS6803	M (23)	D (–1)
F507 (m)	164A09	M	I (+1)	F846 (f)	DXS6803	P&M (27&26)	D (–1)
	DXS7132	M (30)	I (+2)	F998 (f)	DXS6789	M (32)	I (+4)
	DXS9895	M	I (+1)	F1010 (f)	DXS7133	U (27&26)	D (–1)
F538 (m)	164A09	M (21)	I (+1)	F1022 (m)	DXS6803	M	I (+1)
F541 (m)	164A09	M (32)	I (+1)		DXS8378	M (23)	D (–1)
F551 (f)	DXS6803	M (26)	D (–1)		164A09	M	I (+1)
F557 (f)	DXS6803	U (28&25)	D (–1)	F1075 (m)	DXS6803	M (33)	U
F569 (f)	DXS9898	M (33)	D (–1)	F1078 (m)	DXS6803	M (29)	D (–0.1)
F576 (m)	DXS6803	M (27)	U	F1089 (f)	DXS6803	P&M (32&28)	D (–0.1)
F601 (f)	DXS9898	M (26)	D (–1)	F1109 (f)	DXS6795	M (26)	D (–1)
F650 (f)	DXS6803	P (30)	I (+0.1)	F1151 (f)	DXS6803	M (37)	D (–0.1)
F686 (f)	DXS6803	M (27)	I (+0.1)	F1162 (f)	DXS9895	M (24)	D (–4)
F702 (f)	DXS7132	U (28&25)	D (–2)	F1199 (f)	DXS6803	P (22)	U
F719 (m)	DXS8378	M (25)	D (–1)				

Table 4 Typing results for the families with multiple children in which families with mutations were excluded (*C1* was regarded as a reference in deciding the recombined loci, which are represented with bold and underlined letter)

Family		Locus								
		6803	8378	164A09	7132	7133	9895	9898	6789	28C05
F600	F	12.3	9	16.3	15	9	10	10	17	11
	M	12.3	11–9	13–12	14–13	9	12	10	22–21	13–11
	C1(m)	12.3	11	13	14	9	12	10	22	11
	C2(m)	12.3	9	**13**	13	9	12	10	**22**	13
F629	F	12.3	10	15.3	16	9	10	10	16	11
	M	11.3–11	10–9	15.3–13	14–13	11–9	12–10	12–10	21–16	13
	C1(m)	11	10	13	13	11	12	10	16	13
	C2(m)	11	10	13	13	11	12	10	16	13
	C3(m)	11.3	**10**	15.3	**13**	9	10	12	21	13
F673	F	12.3	11	11	15	9	14	10	16	11
	M	12.3–12	9	15.3	15–14	9	10	11–10	20–15	13–11
	C1(f)	12.3	11–9	15.3–11	15–14	9	14–10	11–10	20–16	13–11
	C2(f)	12.3	11–9	15.3–11	15–14	9	14–10	11–10	16–**15**	**11**
F679	F	12.3	9	15.3	13	9	10	9	21	10
	M	12.3–12	10	15.3–12	14–13	10–9	12–11	11–10	16–15	11
	C1(m)	12	10	15.3	14	10	11	10	15	11
	C2(f)	**12.3**	10–9	15.3	14–13	10–9	11–10	**11**–9	21–**16**	11–10
F697	F	11	10	12	14	9	10	11	16	11
	M	12.3	9	14.3–11	14	11–9	11–10	11–10	21–15	15–13
	C1(f)	12.3–11	10–9	14.3–12	14	9	10	11	16–15	15–11
	C2(m)	12.3	9	11	14	11	**10**	10	21	13
F699	F	11	9	15.3	15	9	11	11	20	11
	M	12.3	10–8	15.3	13–12	9	10	11	21–16	11–10
	C1(f)	12.3–11	10–9	15.3	15–12	9	11–10	11	21–20	11–10
	C2(f)	12.3–11	10–9	15.3	15–**13**	9	11–10	11	21–20	**11**
F721	F	12.3	9	11	14	10	10	10	16	14
	M	12.3–11	12–9	13.3–12	16–14	9	13–12	10	16–15	13–11
	C1(f)	12.3	9	13.3–11	16–14	10–9	13–10	10	16–15	14–11
	C2(m)	12.3	9	**12**	**14**	9	13	10	**16**	11
F841	F	12.3	9	13	15	9	13	10	15	11
	M	12.3	12–10	12	14	10–9	13–11	11–10	17–15	11
	C1(f)	12.3	12–9	13–12	15–14	10–9	13–11	11–10	15	11
	C2(m)	12.3	**12**	12	14	9	**11**	10	17	11
F852	F	12.3	9	12	13	9	13	12	15	11
	M	12.3–11	10–9	15.3	14–12	10–9	13–11	11–10	20–15	10–9
	C1(f)	12.3–11	9	15.3–12	13–12	10–9	13–11	12–11	15	11–9
	C2(m)	12.3	10	15.3	**12**	9	13	10	20	10
F916	F	11.3	10	14.3	15	9	11	11	20	13
	M	12.3–11	11–10	14.3–13	15–13	10–9	13–10	10–9	22–20	13
	C1(m)	12.3	11	13	13	9	13	10	20	13
	C2(m)	12.3	**10**	**14.3**	13	**10**	**10**	10	20	13
F942	F	12.3	9	15.3	15	9	12	10	21	13
	M	12.3–12	11–10	15.3	14	9	13–10	11–10	22–15	13–11
	C1(f)	12.3–12	10–9	15.3	15–14	9	12–10	10	21–15	13
	C2(m)	12	**11**	15.3	14	9	10	10	15	**11**
F1034	F	12.3	10	15.3	14	9	10	12	23	11
	M	12.3–11	10–9	15.3	15	9	13–11	11–10	20–15	13–10
	C1(f)	12.3–11	10–9	15.3	15–14	9	11–10	12–11	23–15	11–10
	C2(f)	12.3	10–**9**	15.3	15–14	9	13–10	12–10	23–20	11–**10**
	C3(m)	11	**10**	15.3	15	9	**13**	11	15	**13**
F1142	F	12	10	11	14	10	12	10	16	9
	M	12.3–11	9	15.3	15–13	9	12–11	10	21–20	11
	C1(m)	11	9	15.3	13	9	12	10	21	11
	C2(m)	11	9	15.3	13	9	12	10	**20**	11
F1182	F	12.3	10	15.3	17	9	11	10	21	9
	M	12.3	11–9	15.3–14.3	15	10	11	11–9	21	11–10
	C1(m)	12.3	9	15.3	15	10	11	9	21	11
	C2(f)	12.3	10–9	15.3	17–15	10–9	11	10–9	21	**10**–9

cally, and this provides valuable population data. The three loci GATA164A09, DXS9895 and DXS6789 did not conform to the Hardy-Weinberg equilibrium based on the exact test in females. There may be several explanations for this. First, there may exist some subgroups in the Korean population and the female subjects studied in the present study could have been selected from such subgroups. Second, the number of female subjects screened in this study may have been insufficient to test the Hardy-Weinberg equilibrium compared to the allelic number in these loci. More data are required to provide an answer.

Acknowledgements This paper was supported by the Seoul National University Hospital Research Fund.

References

1. Brinkmann B (1998) Overview of PCR-based systems in identity testing. Methods Mol Biol 198:105–119
2. Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). J Forensic Sci 43:1046–1049
3. Dou X, Guo C, Liu X, Zou Y, Gone Y (1998) Polymorphism analysis of 4 loci of X-chromosome Chinese population of the Han nationality. Chin J Med Genet 15:92–94
4. Edelmann J, Szibor R (2001) DXS101 – a highly polymorphic X-linked STR. Int J Legal Med 114:301–304
5. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372
6. Hering S, Szibor R (2000) Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. J Forensic Sci 45:929–931
7. Hering S, Kuhlisch E, Szibor R (2001) Development of the X-linked tetrameric microsatellite marker HumDXS6789 for forensic purposes. Forensic Sci Int 119:42–46
8. Jones DA (1972) Blood samples: probabilities and discriminations. J Forensic Sci Soc 12:355–359
9. Kishida T, Tamaki Y (1997) Japanese population data on X-chromosomal STR locus AR. Nippon Hoigaku Zasshi 51:376–379
10. Kishida T, Wang W, Fukuda M, Tamaki Y (1997) Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. Nippon Hoigaku Zasshi 51: 67–69
11. De Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig P, et al. (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110: 134–140
12. Rolf B, Keil W, Brinkmann B, Roewer L, Fimmers R (2001) Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations. Int J Legal Med 115: 12–15
13. Szibor R, Lautsch S, Plate I, Beck N (2000) Population data on the X chromosome short tandem repeat locus HumHPRTB in two regions of Germany. J Forensic Sci 45:231–233
14. Urquhart A, Oldroyd NJ, Kimpton CP, Gill P (1995) Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. Biotechniques 18:116–118
15. Watanabe G, Umetsu K, Yuasa I, Suzuki T (2000) DXS10011; a hypervariable tetranucleotide STR polymorphism on the X chromosome. Int J Legal Med 113:249–250