Int J Legal Med (2008) 122:525–528
DOI 10.1007/s00414-008-0263-1

SHORT COMMUNICATION

Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population

Muhammad Akram Tariq · Obaid Ullah · S. Amer Riazuddin · Sheikh Riazuddin

Received: 8 January 2008 / Accepted: 9 June 2008 / Published online: 16 July 2008

Abstract Short tandem repeat (STR) markers are extensively being used for human identification as well as paternity and forensic case work. X-chromosome STR (X-STR) markers are a powerful complementary system especially in deficiency paternity testing. Many X-linked microsatellites have been evaluated but further studies are required to determine population specific statistics. Here, we report allele frequencies of 13 X-linked microsatellites (DXS8378, DXS9902, DXS6810, DXS7132, DXS981, DXS6793, DXS6801, DXS6789, GATA172D05, HPRTB, GATA31E08, DXS8377, and DXS7423) in the Pakistani population. Blood samples were collected from individuals representing all major ethnic groups of the Pakistan population. A total of 5–18 alleles were observed for each locus and altogether 109 alleles for all 13 X-STR loci. Heterozygosity in females ranged from 0.524 to 0.884. No significant deviation was observed from Hardy–Weinberg equilibrium for all 13 microsatellites. In addition, there was no evidence of linkage disequilibrium in any pairs of these markers. These results strongly suggest that the X-linked microsatellites described here can potentially serve as an extension to autosomal systems currently used in parentage analysis and forensic case work.

Keywords X-chromosome short tandem repeats (X-STRs) · Pakistani population

Electronic supplementary material The online version of this article (doi:10.1007/s00414-008-0263-1) contains supplementary material, which is available to authorized users.

M. A. Tariq · O. Ullah · S. A. Riazuddin (✉) · S. Riazuddin (✉)
National Centre of Excellence in Molecular Biology,
University of the Punjab,
87 West Canal Bank Road,
Lahore 53700, Pakistan
e-mail: amerriazuddin@ncemb.org
e-mail: riaz@lhr.comsats.net.pk

Introduction

Short tandem repeat (STR) markers are extensively being used for human identification as well as paternity and forensic case work. Recently, X-chromosomal STR (X-STR) markers have attracted attention because of their usefulness in complex kinship and deficiency paternity, in particular, when the disputed child is female [1–3]. These markers can also be used in sibship analysis of two females having the same biological father and in relationship testing, especially for grandmother/granddaughter relationship, as both share a common X-chromosome allele. Additionally, X-STR markers have proven to be extremely helpful to investigate stains to identify the female DNA profile [4, 5].

Although many X-STRs microsatellites have been validated for paternity testing, additional population studies would enhance our understanding of the polymorphism and allelic distribution associated with these loci. Since there is correlation between degree of polymorphism and the mutation rate at a given locus [6], the estimation of mutation rates would be beneficial to both parentage and forensic case work as they are important for correct genetic profile interpretation. Furthermore, linkage disequilibrium (LD) between X-STR loci has proven to be population-specific [7]. Therefore, LD studies of pertinent populations would be essential for X-STRs use in forensic case work. In short, with additional population statistics, X-STRs could potentially serve as an extension to autosomal systems currently used in parentage analysis and forensic case work.

Here, we report allele frequencies of 13 X-linked microsatellites (DXS8378, DXS9902, DXS6810, DXS7132, DXS981, DXS6793, DXS6801, DXS6789, GATA172D05, HPRTB, GATA31E08, DXS8377, and DXS7423) in a Pakistani population. The samples were collected from a large group of individuals representing all major ethnic groups of

the Pakistani population. A total of 5–18 alleles were observed for each locus and altogether 109 alleles for all 13 X-STR loci. Heterozygosity in females ranged from 0.524 to 0.884. No significant deviation was observed from Hardy–Weinberg equilibrium for all 13 microsatellites. In addition, there was no evidence of linkage disequilibrium in any pairs of these markers.

Materials and methods

According to the recent census, the Pakistani population consists of five major ethnic groups. The Punjabis make up 55 %; the Pakhtuns, 15%; the Sindhis, 14%; the Muhajirs, 8 %; and the Balochs, 4% of the total population (Statistics Division, Ministry of Economic Affairs and Statistics, Government of Pakistan, http://www.statpak.gov.pk). Blood samples were collected from 432 unrelated individuals (285 males and 147 females) representing all five major ethnic groups. In addition 50 "true trios" families, i.e., father, mother, and daughter were collected to investigate X-chromosomal inheritance and stability of these markers. Genomic DNA was extracted from whole blood using the phenol-chloroform method [8]. The quantity of DNA was determined using a spectrophotometer.

Polymerase chain reactions were carried out in a 20 µl reaction volume containing 10 ng of genomic DNA, 0.125 µM of forward and reverse primers, 75 mM Tris HCl, 20 mM $(NH)_2SO_4$, 2 mM $MgCl_2$, 200 µM of each deoxyribonucleotide triphosphate (dNTP), 1 U of Taq Gold polymerase. Each STR marker was amplified individually. The amplification was performed in a GeneAmp PCR Thermal Cycler System 2700 (Applied Biosystems). The polymerase chain reaction (PCR) consisted of a denaturation at 95°C for 5 min, followed by 30 cycles consisting of denaturation at 94°C for 1 min, annealing at 57°C for 1 min, extension at 72°C for 1 min and a final extension at 72°C for 30 min. The capillary electrophoresis was performed using 1 µl of amplified product, 12 µl formamide, and 0.5 µl GeneScan LIZ-500 size standard (Applied Biosystems). The amplified products were separated on an ABI 3100 Genetic Analyzer (Applied Biosystems). Results were analyzed using GeneScan Analysis software 3.7 (Applied Biosystems). Human DNA sample from Promega was used as reference for genotyping (Promega) [9].

Primer sequences for DXS981, DXS6789, HPRTB, DXS8377, and DXS7423 were redesigned to adjust fragment length to have similar annealing temperatures in multiplex PCR (Supplementary Table 1 online). The primer sequences for DXS8378, DXS9902, DXS6810, DXS7132, DXS6793, DXS6801, GATA172D05, and GATA31E08 were obtained from the Genome Database [http://www.gdb.org]. Alleles were assigned according to the recommendations of the International Society of Forensic Haemogenetics commission [10].

The presence of new alleles was confirmed by bidirectional sequencing. Primers were designed using the primer3 program (http://primer3.sourceforge.net). Amplifications were performed in 25 µl reactions containing 50 ng of genomic DNA, 2.5 µl 10X GeneAmp PCR buffer II, eight pmoles of each primer, 2.5 mM dNTP, 2.5 mM $MgCl_2$ and 1 unit of *Taq* DNA polymerase. The PCR consisted of a denaturation step at 96°C for 5 min, followed by 40 cycles, each consisting of 96°C for 45 s followed by 57°C for 45 s and at 72°C for 1 min. The PCR products were used for bidirectional sequencing using BigDye Terminator Ready reaction mix according to the manufacturer's instructions (Applied Biosystems). Sequencing products were re-suspended in 10 µl of formamide and denatured at 95°C for 5 min. Sequencing was performed on an ABI PRISM 3100 automated sequencer (Applied Biosystems). Sequencing results were assembled using ABI PRISM sequencing analysis software version 3.7 (Applied Biosystems) and analyzed with Chromas software (http://www.technelysium.com.au/chromas.html).

Table 1 Statistical parameters for X-chromosomal STRs in Pakistani population (n=432)

	DXS8378	DXS9902	DXS6810	DXS7132	DXS981	DXS6793	DXS6801	DXS6789	GATA172D05	HPRTB	GATA31E08	DXS8377	DXS7423
HWE[a]	0.477	0.759	0.43	0.646	0.047	0.028	0.217	0.09	0.617	0.275	0.165	0.976	0.646
PIC[a]	0.654	0.647	0.568	0.715	0.75	0.511	0.613	0.737	0.738	0.708	0.748	0.882	0.578
MECI[a]	0.653	0.648	0.567	0.715	0.75	0.511	0.613	0.737	0.738	0.708	0.748	0.882	0.578
MECII[a]	0.511	0.503	0.42	0.578	0.622	0.364	0.465	0.606	0.606	0.569	0.617	0.797	0.433
H_{obs}[a]	0.653	0.687	0.66	0.741	0.599	0.524	0.66	0.741	0.721	0.714	0.719	0.884	0.667
PD_f[a]	0.862	0.857	0.801	0.9	0.924	0.758	0.837	0.918	0.916	0.895	0.918	0.978	0.807
PD_m[b]	0.687	0.73	0.621	0.726	0.808	0.613	0.707	0.758	0.751	0.719	0.766	0.905	0.643

HWE Hardy–Weinberg equilibrium, *PIC* polymorphism information content, *MEC* mean exclusion chance (*I* for X chromosomal markers in trios, *II* for X chromosomal markers in duos [22], H_{obs} observed heterozygosity, PD^f power of discrimination in females, PD^m power of discrimination in males

[a] Only female's data is used for the analysis

[b] Only male data is used for the analysis

Allele frequencies for each locus were calculated for both males and females collectively. Observed heterozygosities (Hobs) were calculated with PowerStatsV12 software (http://www.promega.com). Hardy–Weinberg equilibrium (HWE) was calculated by an exact test with GENEPOP software [11]. The polymorphism information content (PIC), mean exclusion chance (MEC), power of discrimination in females (PDf) and power of discrimination in males (PDm) were calculated with chromosome X web software [12]. Linkage disequilibrium test was performed for all the pairs of markers with ARLEQUIN software [13].

Results and discussion

Here, we report allele frequencies of 13 X-linked microsatellites (DXS8378, DXS9902, DXS6810, DXS7132, DXS981, DXS6793, DXS6801, DXS6789, GATA172D05, HPRTB, GATA31E08, DXS8377, and DXS7423) in the Pakistani population. The samples were collected from a large group of individuals (n=432) including 285 males and 147 females, representing all major ethnic groups of Pakistan population. These X-STR microsatellites span the entire X-chromosome and represent all 4-linkage groups previously categorized by Szibor et al. [5]. Allelic frequencies for 13 STR markers are shown in Supplementary Table 2 online and other statistical parameters are specified in Table 1.

A total of 5–18 alleles for each locus were observed and altogether 109 alleles for all 13 X-STR loci. Heterozygosity in females ranged from 0.524–0.884. These results showed that DXS8377 is the most informative marker with heterozygosity at 0.844, where as DXS6793 is the least informative, with heterozygosity at 0.524. The X-STR markers show no significant deviation from the Hardy–Weinberg equilibrium (p>0.01) (Table 1). Genotypes for the reference DNA sample (NA9947A) are in accordance with the recommendations of Szibor et al., except for marker HPRTB the recommendations of Gomes et al. were followed [9, 14]. Genotype of (10, 10) were observed for the reference DNA sample (NA9947A) with marker DXS6793, which has not been reported in earlier studies. Alleles 14, 15, and 16 for marker DXS6793 were absent in Chinese but are present in the Pakistani population [15]. Similarly, alleles 10 and 18 for marker DXS981 were absent in Korean population but were observed in the Pakistani population [4]. The allele 7 of GATA172D05 was reported in German, Korean, and Colombian population but is absent in the Pakistani population, consistent with the Italian population [4, 16–19]. The allele 10 of marker HPRTB was reported in German, Hungarian, Korean, African Americans, Hispanics, and Portuguese populations but were not observed in the Pakistani population [1, 4, 15, 17]. Presence of new alleles was confirmed by sequencing the PCR products. We did not detect any mutations in 100 meiotic events (50 maternal transfers, 50 paternal transfers) at any locus.

Since X-chromosome recombines only in female germline, X-linked STR markers have higher a degree of linkage disequilibrium compared with autosomal STR markers [20]. The exact test for LD was performed for all the markers described here in the Pakistani population. No evidence of LD was detected in any pairs of these markers. Previously markers DXS6789 and DXS6801 have been reported to be in strong linkage disequilibrium [21]. Although an estimated p value of 0.020 was observed in the Pakistani population for markers DXS6789 and DXS6801. Nevertheless, this p value is above the significance level of 0.000641 (obtained after Bonferroni correction) suggesting no evidence of association between the above mentioned pair of markers. Therefore, all the markers of this study can be considered to be independent for parentage analysis and forensic case work in the Pakistani population. Haplotype frequencies for these two markers (DXS6789 and DXS6801) are shown in Supplementary Table 3 online.

Conclusion

Our results strongly suggest that all 13 X-STRs described here can efficiently be used in parentage analysis and provide a powerful tool in forensic case work, in particular, to identify the female DNA profile in mixture analysis. Human population characteristics at the genetic level are integral to both forensic biology and population genetics and to the best of our knowledge, this is the first study involving all major ethnic groups of Pakistani population. A study for development of a multiplex PCR kit based on the X-STR markers described here is in progress.

Acknowledgements The authors are grateful to all individuals for their participation in this study. This work was supported in part by Higher Education Commission, Islamabad, Pakistan and Ministry of Science and Technology, Islamabad, Pakistan.

References

1. Pereira R, Gomes I, Amorim A, Gusmao L (2007) Genetic diversity of 10 X chromosome STRs in northern Portugal. Int J Legal Med 121:192–197
2. Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H (2006) MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. Int J Legal Med 120:174–181
3. Shin KJ, Kwon BK, Lee SS, Yoo JE, Park MJ, Chung U, Lee HY, Han GR, Choi JH, Kim CY (2004) Five highly informative X-chromosomal STRs in Koreans. Int J Legal Med 118:37–40

4. Shin SH, Yu JS, Park SW, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41
5. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117:67–74
6. Lee HY, Park MJ, Chung U, Lee HY, Yang WI, Cho SH, Shin KJ (2007) Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father-son pairs. Int J Legal Med 121:128–135
7. Lee HY, Park MJ, Jeong CK, Lee SY, Yoo JE, Chung U, Choi JH, Kim CY, Shin KJ (2004) Genetic characteristics and population study of 4 X-chromosomal STRs in Koreans: evidence for a null allele at DXS9898. Int J Legal Med 118:355–360
8. Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
9. Szibor R, Edelmann J, Hering S, Plate I, Wittig H, Roewer L, Wiegand P, Cali F, Romano V, Michael M (2003) Cell line DNA typing in forensic genetics—the necessity of reliable standards. Forensic Sci Int 138:37–43
10. Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B (1997) DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. Int J Legal Med 110:175–176
11. Raymond M, Rousset F (1995) GENEPOP version 1.2: population genetics software for exact tests and ecumenicism. J Heredity 86:248–249
12. Szibor R, Hering S, Edelmann J (2006) A new Web site compiling forensic chromosome X research is now online. Int J Legal Med 120:252–254
13. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47–50
14. Gomes I, Prinz M, Pereira R, Meyers C, Mikulasovich RS, Amorim A, Carracedo A, Gusmao L (2007) Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int J Legal Med 121:198–203
15. Jia Y, Zhou B, Liang WB, Lv ML, Zhang L (2004) Two X-chromosome STR loci DXS6803 and XS6793 frequency data in Chinese population. J Forensic Sci 49:845–846
16. Edelman J, Hering S, Michael M, Lessig R, Deischel D, Meier-Sundhausen G, Roewer L, Plate I, Szibor R (2001) 16 X-chromosome STR loci frequency data from a German population. Forensic Sci Int 124:215–218
17. Pico A, Castillo A, Vargas C, Amorim A, Gusmao L (2008) Genetic profile characterization and segregation analysis of 10 X-STRs in a sample from Santander, Colombia. Int J Legal Med
18. Turrina S, De LD (2003) Population data of three X-chromosomal STRs: DXS7132, DXS7133 and GATA172D05 in North Italy. J Forensic Sci 48:1428–1429
19. Zalan A, Volgyi A, Jung M, Peterman O, Pamjav H (2007) Hungarian population data of four X-linked markers: DXS8378, DXS7132, HPRTB, and DXS7423. Int J Legal Med 121:74–77
20. Robino C, Giolitti A, Gino S, Torre C (2006) Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. Int J Legal Med 120:315–318
21. Szibor R, Hering S, Kuhlisch E, Plate I, Demberger S, Krawczak M, Edelmann J (2005) Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 119:363–369
22. 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