Full Length Research Paper

Punjabi population data for seven X-chromosome short tandem repeat (X-STR) loci using a new miniplex system

Muhammad Israr, Ahamad Ali Shahid*, Ziaur Rahman, Muhammad Saqib Shahzad, Obaid Ullah and Tayyab Husnain

National Centre of Excellence in Molecular Biology, University of the Punjab, 87-West Canal Bank Road, Lahore-53700, Pakistan.

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. This study presents development of a new X-chromosomal STR multiplex kit providing short-amplicon (<200 bp) fragments. About 200 samples from Punjabi population were typed for seven X-STR markers: DXS101, DXS6789, DXS7132, DXS7423, DXS8378, GATA172D05 and GATA31E08. A total of 51 alleles were detected in the range of five to 10 alleles for each marker. The data can be used as reference database for Punjabi population along with the current battery of autosomal STR for forensic case work to increase the discrimination capacity and strengthen the existing system.

Key words: X-chromosome short tandem repeats (X-STRs), mini-STRs, Punjabi population.

INTRODUCTION

Chromosome X short tandem repeats (X-STRs) analysis has recently attracted attention of the forensic community because of its usefulness in complex kinship testing. It is worth-while including X-STRs with autosomal markers for the cases when father/daughter relationships are to be tested. Forensic X-STR markers, as per size of amplicons, are just like autosomal and Y chromosomal STRs. Mini-STRs for autosomal, X and Y chromosomal can be efficiently used for degraded DNA. X-STRs have also advantage over autosomal STRs for paternity cases involving close blood relatives as alternative putative fathers. The major advantage of X-STRs is proven in deficiency paternity cases, that is, when the DNA sample from putative father is not available and DNA from paternal relative has to be analyzed instead. Further, X-linked STRs can be used to solve sibling ship status of

two females having the same biological father, even without father's DNA. X-STRs can determine the relation-ship of grandmother/granddaughter as granddaughter theoretically has to carry at least one allele in common with the grandmother. In forensic analysis of mixed stains, X-STRs are helpful to identify the female DNA (Diegoli and Coble, 2011; Gomes et al., 2007; Shin et al., 2005).

To alleviate the problems associated with analyzing DNA from degraded samples, new sets of STR primers known as Miniplexes have been recently designed. The primers were created by moving the primer binding sites as close as possible to the repeat region (Hill et al., 2008). A number of studies have demonstrated that successful analysis of degraded DNA specimens from mass disasters or forensic evidence improves with smaller sized polymerase chain reaction (PCR) products (Drabek et al., 2004).

The Punjab is the largest province of Pakistan consisting about 55% of the country's population (http://www.statpak.gov.pk). This study reports the successful optimization and testing of miniX multiplex

---
*Corresponding author. E-mail: ahmadali.shahid@gmail.com.

Abbreviations: STR, Short tandem repeat; X-STR, X-chromosome STR.

which is capable of parallel amplification of multiple X-miniSTRs. The markers included in multiplex have been DXS101, DXS6789, DXS7132, DXS7423, DXS8378, GATA172D05, and GATA31E08. After the optimization of miniX PCR, it was used to analyze Punjabi population, using an ABI PRISM 3730 genetic analyzer (Applied Biosystems). The samples were collected from a large group of individuals representing all areas of the Punjab. A total of five to 10 alleles were observed for each locus and altogether 51 alleles for all seven X-STR loci. Heterozygosity in females ranged from 0.38 to 0.74. No significant deviation was observed from Hardy–Weinberg equilibrium for all seven microsatellites. No new alleles were reported that were not previously reported for any Pakistani population (Tariq et al., 2008).

MATERIALS AND METHODS

Population

Two hundred (200) blood samples from randomly selected, unrelated healthy males (109) and females (91) individuals were collected from the Punjab province of Pakistan. Informed consent was obtained from every individual before his/her sample was taken. The samples were collected in tubes containing 0.5 M ethylenediaminetetraacetic acid (EDTA) and were immediately shifted to -4°C until the time of DNA extraction.

DNA extraction and quantification

Genomic DNA was isolated from the whole blood by organic extraction method using phenol, chloroform and isoamyl alcohol (PCI) (Maniatis et al., 1982). DNA was quantified using a spectrophotometer (SpectraMax Plus384 Microplate Reader, Molecular Devices. CA, USA).

Amplification and genotyping

DNA was amplified for seven microsatellite loci. As shown in Supplementary Table S1, primer sequences for six loci, DXS101, DXS6789, DXS7423, DXS8378, GATA31E08 (Asamura et al., 2006) and GATA172D05 (Edelmann et al., 2002) were already reported while one new primer for locus DXS7132 was designed using primer 3 (Rozen and Skaletsky, 2000). Multiplex polymerase chain reaction (PCR) was performed in a single reaction using GeneAmp PCR Thermal Cycler System 2700 (Applied Biosystems (ABI), Foster City, CA). PCR was performed in 15 µL reaction volume containing 5 ng of genomic DNA, 75 mM Tris-HCl, 20 mM $(NH)_2SO_4$, 2 mM $MgCl_2$, 200 µM of each deoxyribonucleotide triphosphate (dNTP), 1U of AmpliTaq Gold polymerase (Applied Biosystems) and forward and reverse primers for each X-STR locus. The primer concentrations are shown in Supplementary Table S1. Human control DNA 9947A (ABI) was used as a positive control. The capillary electrophoresis was performed using 1 µL of PCR product, 13.5 µL Hi-Di Formamide, and 0.5 µL GeneScan LIZ-500 size standard (Applied Biosystems). The amplified products were separated on an ABI 3730 genetic analyzer (Applied Biosystems). Results were analyzed using GeneMapper software v3.7 (Applied Biosystems). Allele nomenclature and allele ranges were according to already reported results (Allen and Belmont,

1993; Hering et al., 2001; Edelmann et al., 2002; Zarrabeitia et al., 2002).

Statistical analysis

Allele frequencies for each locus were calculated for both males and females collectively by hand. Observed heterozygosities (HET), polymorphism information content (PIC), power of exclusion (PE) and paternity index (PI) were calculated with PowerStats v12 software (http://www.promega.com). Hardy–Weinberg equilibrium (HWE) was calculated by an exact test with Arlequin v3.5 software (Excoffier and Lischer, 2010). Power of discrimination in females (PDf) and power of discrimination in males (PDm) were calculated with chromosome X web version (Szibor et al., 2006).

RESULTS AND DISCUSSION

We reported a new multiplex consisting of reduced size amplicons for seven X-chromosomal microsatellites (DXS101, DXS6789, DXS7132, DXS7423, DXS8378, GATA172D05 and GATA31E08). These markers span the entire X chromosome except for linkage group 3 (Szibor et al., 2006). Punjabi population of Pakistan was genotyped with this new miniplex kit. Allele frequency data for the Punjabi population and forensic efficiency parameters are presented in Table 1. A total of 51 alleles were detected for all seven X-STR markers. For each locus, about five to 10 alleles were observed. Heterozygosity in females ranged between 0.38 to 0.74. Based on these results locus GATA31E08 seemed to be the most informative marker with heterozygosity at 0.74, while locus GATA172D05 seemed to be the least informative, with heterozygosity at 0.38. The X-STR markers showed no significant deviation from the Hardy–Weinberg equilibrium (p>0.01) except for locus DXS101. After Bonferroni correction, the locus DXS101 showed no significant deviation (Table 1). Genotypes for the reference DNA sample (NA9947A) are in accordance with the recommendations of International Society for Forensic Genetics (ISFG) (Bar et al., 1997; Diegoli and Coble, 2011; Gomes et al., 2009).

Conclusion

Our results suggest that all seven 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. This miniX system is highly recommended to be used along with the current STR locus set of forensic markers as a supplement in certain cases. Human population charac-teristics at the genetic level are integral to both forensic biology and population genetics. Although, a study conducted on the whole Pakistani population has been done which include the markers used in this study (Tariq et al., 2008) while a study of five X-STR was done exclusively on Punjabi population, but that did not include

Table 1. Allele frequencies and Forensic Parameters of different STR loci in Punjabi population (n=200).

Allele	Frequency						
	DXS101	DXS6789	DXS7132	DXS7423	DXS8378	GATA172D05	GATA31E08
6	-	-	-	-	-	0.060	-
7	-	-	-	-	-	0.044	0.011
8	-	-	-	-	-	0.132	0.264
9	-	-	-	-	0.033	0.049	0.016
10	-	-	-	-	0.275	0.440	0.159
11	-	-	0.011	-	0.390	0.170	0.198
12	-	-	0.126	0.082	0.264	0.104	0.225
13	-	-	0.286	0.407	0.038	-	0.121
14	-	0.022	0.346	0.368	-	-	0.005
15	-	0.192	0.159	0.126	-	-	
16	-	0.077	0.060	0.016	-	-	
17	-	0.027	0.011	-	-	-	-
18	-		-	-	-	-	-
19	-	0.022	-	-	-	-	-
20	-	0.385	-	-	-	-	-
21	0.022	0.170	-	-	-	-	-
22	0.033	0.055	-	-	-	-	-
23	0.280	0.044	-	-	-	-	-
24	0.264	0.005	-	-	-	-	-
25	0.203	-	-	-	-	-	-
26	0.071	-	-	-	-	-	-
27	0.066	-	-	-	-	-	-
28	0.044	-	-	-	-	-	-
29	0.016	-	-	-	-	-	-
PIC[a]	0.77	0.75	0.71	0.62	0.64	0.71	0.77
PE[a]	0.165	0.4	0.353	0.235	0.283	0.105	0.505
PI[a]	0.95	1.57	1.42	1.11	1.23	0.81	1.98
HET[a]	0.473	0.681	0.648	0.549	0.593	0.385	0.747
PD_f[a]	0.912	0.919	0.902	0.822	0.864	0.837	0.925
PD_m[b]	0.816	0.763	0.753	0.688	0.689	0.5081	0.796
HWE[a]	0.03	0.09	0.646	0.646	0.477	0.617	0.165

HWE, Hardy–Weinberg equilibrium; PIC, polymorphism information content; HET, heterozygosity; PD_f, power of discrimination in females; PD_m, power of discrimination in males; PE, power of exclusion; PI, paternity index. [a]Only female's data is used for the analysis. [b]Only male data is used for the analysis

the markers used in our study (Nadeem et al., 2009) and hence, to the best of our knowledge, this is the first study done exclusively on the Punjabi population with these seven markers. Although this study lacks any comparison between the mini-STRs and their larger-amplicon counterparts, but still it is an improvement on the published work of Asamura et al. (2006) from which some of the primers were used in this study.

ACKNOWLEDGEMENT

REFERENCES

Allen RC, Belmont JW (1993). Trinucleotide repeat polymorphism at DXS101. Hum. Mol. Genet. 2: p. 1508.

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.

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.

Diegoli TM, Coble MD (2011). Development and characterization of two mini-X chromosomal short tandem repeat multiplexes. Forensic Sci. Int. Genet. 5: 415-421.

Drabek J, Chung DT, Butler JM, McCord BR (2004). Concordance study between Miniplex assays and a commercial STR typing kit. J. Forensic Sci. 49: 859-860.

Edelmann J, Deichsel D, Hering S, Plate I, Szibor R (2002). Sequence variation and allele nomenclature for the X-linked STRs DXS9895, DXS8378, DXS7132, DXS6800, DXS7133, GATA172D05, DXS7423 and DXS8377. Forensic Sci. Int. 129: 99-103.

Excoffier L, Lischer HE (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10: 564-567.

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.

Gomes I, Prinz M, Pereira R, Bieschke E, Mayr WR, Amorim A, Carracedo A, Gusmao L (2009). X-chromosome STR sequence variation, repeat structure, and nomenclature in humans and chimpanzees. Int. J. Legal Med. 123: 143-149.

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.

Hill CR, Kline MC, Coble MD, Butler JM (2008). Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. J. Forensic Sci. 53: 73-80.

Maniatis T, Fritsch EF, Sambrook J (1982). Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Nadeem A, Babar M, Hussain M, Tahir M (2009). Development of pentaplex PCR and genetic analysis of X chromosomal STRs in Punjabi population of Pakistan. Mol. Biol. Rep. 36: 1671-1675.

Rozen S, Skaletsky H (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132: 365-386.

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.

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.

Tariq M, Ullah O, Riazuddin S, Riazuddin S (2008). Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. Int. J. Legal Med. 122: 525-528.

Zarrabeitia MT, Amigo T, Sanudo C, de Pancorbo MM, Riancho JA (2002). Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. Int. J. Legal Med. 116: 368-371.

Supplementary Table S1. Mini-X STR primers sequences and concentrations used in this study.

Marker name	Primer sequence (5′ to 3′)	Reference	Final primer concentration (μM)	Amplicon size range (bp)
DXS101	NED-TCTCCCTTCAAAAACAAAGATAAGTGCATATTCTGCGCATGT	Asamura et al., 2006	0.3	142 to 166
DXS6789	VIC-CCTCGTGATCATGTAAGTTGGGCAGAACCAATAGGAGATAGATGGT	Asamura et al., 2006	0.12	122 to 162
DXS7132	FAM-ATAAATCCCCTCTCATCTATCTGACACTCCTGGTGCCAAACTCTA	This study	0.1	124 to 148
DXS7423	FAM-AGATTTCCTCCCCATCCATCGTTGTCACACAAATAAATGAATGAGT	Asamura et al., 2006	0.09	99 to 115
DXS8378	VIC-GCTCCTGGCAGGTCACTATCGCGACAAGAGCGAAACTCCA	Asamura et al., 2006	0.05	95 to 111
GATA172D05	NED-TAGTGGTGATGGTTGCACAGATAATTGAAAGCCCGGATTC	Edelmann et al., 2002	0.07	108 to 132
GATA31E08	PET-CAGAGCTGGTGATGATAGATGAGCTCACTTTTATGTGTGTATGTATCTCC	Asamura et al., 2006	0.09	101 to 129