Avens Publishing Group
J Forensic Investigation
October 2013 Vol.:1, Issue:2

Open Access Research Article

Journal of Forensic Investigation

Genetic Study of Three Closely Linked X chromosome STR Markers in an Argentinian Population

Pablo Noseda[1], Michael Hernandez[2], Brittney Gonzalez[2], Sheree Hughes-Stamm[2] and David Gangitano[2]*

[1]*Laboratorio de Huellas Digitales Geneticas. Secretaria Nacional de Niñez, Adolescencia y Familia (SENAF). Paz Soldan 5200, Buenos Aires, Argentina C1427DSJ*
[2]*Department of Forensic Science, College of Criminal Justice, Sam Houston State University, 1003 Bowers Blvd., Huntsville, Texas 77340, USA*

Address for Correspondence
David Gangitano, Ph.D., Department of Forensic Science, College of Criminal Justice, Sam Houston State University, 1003 Bowers Blvd., Office 221C, Huntsville, Texas 77340, USA, Tel: +1 936 294 4413; fax: +1 936-294-4905; E-mail : dag006@shsu.edu

Submission: 23 September 2013
Accepted: 22 October 2013
Published: 24 October 2013

Keywords: DNA typing; Short tandem repeats; X chromosome; DXS10079; DXS10074; DXS10075

Abstract

Three closely linked X chromosome markers (DXS10079, DXS10074 and DXS10075) were evaluated in a sample of 213 unrelated individuals (106 women and 107 men) from Buenos Aires (Argentina). Hardy-Weinberg equilibrium (HWE) was tested in the female sample and no significant deviations were observed. Homogeneity of allele frequencies of men and women was compared by Fisher's exact test and showed similar distributions. Overall linkage disequilibrium (LD) tests were performed in males for all pairs of loci and significant associations were detected only in the pair DXS10074-DXS10075. However, significant inter-allelic associations were detected among the three markers when the disequilibrium coefficient (D') was estimated. Comparison of allele frequencies among Argentinian, Asian, European, and Latin American populations was performed and parameters of forensic interest were also estimated.

Introduction

X-chromosome short tandem repeat (X-STR) markers have attracted increasing interest in the forensic community due to their advantage over autosomal and Y chromosome markers in kinship cases where the alleged father is absent and the child is female [1,2]. There are four linkage groups of STR markers distributed along the X chromosome [3]. Linkage group 2 contains a cluster composed of markers DXS10074, DXS10074 and DXS10075 spanning a 280-kb region at Xq12 that can provide stable haplotypes to solve complex kinship scenarios [4]. The use of closely located linked markers is of special interest for the reconstruction of the genetic structure of lineages [4,5]. Although several studies have documented allele frequencies of X-STR markers for different ethnic groups [6-8] further population genetic data is necessary in order to increase the knowledge and applicability of these markers in forensic genetics. The objective of the present study is to perform a population study of three closely linked X-STR markers –DXS10079, DXS10074 and DXS10075 in an Argentinian sample set, to estimate parameters of forensic interest and to compare the allele frequencies with those in Japanese, Chinese, German and Brazilian population sets.

Materials and Methods

Population

Blood samples were collected from 213 unrelated individuals from Buenos Aires city (106 women and 107 men) with informed consent. Samples were stripped of any personal identifiers and made anonymous prior to testing at our lab. The protocols used in this study were approved by the Institutional Review Board of Sam Houston State University.

DNA extraction and quantification

Genomic DNA (gDNA) was extracted from blood samples using a modified Chelex® method [9] suspending the resin in TE buffer (Tris-HCl 10mM, EDTA 0.1 mM, pH 8.0). DNA samples were quantified by real time PCR in a StepOne™ Real Time System (Life technologies, Carlsbard, CA) using 2 µL extracted gDNA, 400nM D21S11 primers (GenBank Accession number AP000433) (Integrated DNA Technologies, Coralville, IA), and 1X SYBR® Green PCR Master Mix (Life technologies). The following real time PCR cycling parameters were used: 10 min at 95°C and 40 cycles of 15 s at 95°C and 1 min at 60°C.

PCR amplification and genotyping

Three X-STRs were amplified in a multiplex amplification reaction including DXS10079, DXS10074 and DXS10075 according to previous studies [4]. Amplified PCR products were detected by capillary electrophoresis in either an ABI PRISM™ 310 or a 3500 Genetic Analyzer (Life Technologies). Allele calling was performed using the software GeneMapper v. 4.1 (Life Technologies) and compared with allelic ladders as recommended by the ISFG [10]. A custom designed bin set was implemented to allow automation of genotyping. A ladder containing all the observed alleles in the population was prepared according to Sajantila et al. [11]. Allelic ladders were designed for each X STR marker using a mixture of individual samples that represent each allele observed in the population data set. The samples were amplified separately for each locus accordingly to the previously described protocol. Each of the amplified products were diluted 1:1,000 with H_2O, mixed together in a total volume of 1 mL and re-amplified under the same conditions but using 20 cycles instead of 30. The ladder was calibrated with DNA control samples 9947A, 9948, and K562 (Promega, Madison, WI) s [12]. Microvariants 14.3 (DXS10074), 17.2, 18.2, and 19.2 (DXS10075) were sequenced. DXS10074 and DXS10075 primers were used in a direct *Taq*-cycle sequencing procedure using the BigDye® Direct Cycle Sequencing Kit and the 3500 Genetic Analyzer with Sequencing

ISSN: 2330-0396

Analysis software version 4.1 (Life Technologies). Sequences were assembled and proof-read using the Geneious Pro software v. 5.5.7 (Biomatters Ltd., Auckland, New Zealand). Sequenced alleles were used to calibrate the allelic ladder. Allele frequencies and parameters of forensic interest were generated with PowerStats v. 1.2 software [13]. Hardy-Weinberg equilibrium (HWE) in females, observed heterozygosity and gene diversity were carried out with the Genetic Data Analysis software [14]. Homogeneity of allele frequencies of men and women and overall linkage disequilibrium (LD) in the male sample was assessed using GENEPOP version 4.1 software package [15]. Inter-allelic linkage disequilibrium (ID) in males was estimated with the MIDAS software [16]. Power of exclusion calculations were estimated according to Desmarais et al. [17] Population comparisons were carried out by estimating genetic distances based on Fstusing the Arlequin v. 3.5 software [18]. Bonferroni correction for the number of loci analyzed in HWE and LD evaluation and the correction for the number of inter-population comparisons at adjusted significance level were performed [19].

Quality control

The laboratory of the first author successfully participated in proficiency testing of the Latin-American Society of Forensic Genetics 2007 (http://www.slagf.org.ar).

Results and Discussion

Combined allele frequencies for men and women and parameters of forensic interest are depicted in Table 1. Fisher's exact test did not reveal allele distribution differences between men and women, therefore their frequencies can be combined. No significant deviations from expectations of HWE were detected in any of the markers analyzed in women ($p > 0.015$; significant level after Bonferroni correction). Overall LD and ID are shown in Table 2. Analysis of the three X-STR markers in males showed no detectable evidence of LD except for pair DXS10074-DXS10075 ($p < 0.015$; significant level after Bonferroni correction). These results are consistent with the close genetic distance of these two markers (21 kb). Moreover, significant inter-allelic associations were detected among these three markers when the disequilibrium coefficient (D') was evaluated (D' > 0.8) (Table 2). Most studies in forensic X-STRs use pvalues of a Fisher's exact test as a measured variable to infer the value of gametic disequilibrium (GD); which is not a test to evaluate GD. In linkage studies, it is important not only to measure overall linkage disequilibrium between pairs of loci but also to detect specific inter-allelic associations between markers, which can be evaluated using D' [20]. A value of D' > 0.8 indicates the presence of strong linkage disequilibrium between two markers [21]. Haplotype frequencies for males are displayed in Table 3.

Microvariants observed in loci DXS10074 and DXS10075 were sequenced and submitted to Genbank (Accession numbers JX645356, JX645357, JX645358 and JX645359). Interestingly, the repeat structure of allele 14.3 in DXS10074 reflects the motif of "short" alleles (< 10) instead of "long" alleles (> 13), since it is composed of contiguous repetitions of the $(AAGA)_n$ unit, with a AGA downstream the core repeat unit [4]. The repeat structure of microvariants 17.2, 18.2, and 19.2 of marker DXS10075 is consistent with the repeat pattern of intermediate X.2 alleles described by Hering et al. [4].

Population sample comparisons by pairwise genetic distance analysis, based on Fst were performed with previously published data that contained the cluster DXS10079-DXS10074-DXS10075, when available. These included Brazil [22], China [23], Japan [24],and Germany [4] (Table 4). Results showed statistically

Table 1: Allele frequencies and parameters of forensic interest for three linked X-STR markers in an Argentinian population (106 women and 107 men; 319chromosomes tested).

Alleles	DXS10079	DXS10074	DXS10075
7	–	0.0392	–
8	–	0.0945	–
9	–	0.0323	–
12	–	0.0046	–
13	0.0047	0.0138	0.0386
14	0.0419	0.0276	0.0091
14.3	–	0.0023	–
15	0.0465	0.0991	0.0273
16	0.0465	0.1636	0.2523
16.2	–	–	0.0023
17	0.0372	0.2627	0.375
17.2	–	–	0.0864
18	0.1465	0.1866	0.175
18.2	–	–	0.0091
19	0.1977	0.0645	0.0182
19.2	–	–	0.0068
20	0.2442	0.0069	–
21	0.1488	0.0023	–
22	0.0605	–	–
23	0.0186	–	–
24	0.007	–	–
He_f	0.8558	0.8426	0.737
Ho_f	0.8679	0.8113	0.7
$HW_E f$	0.596	0.541	0.296
PIC	0.8354	0.8199	0.6932
PD_f	0.9511	0.9509	0.8922
PD_m	0.8361	0.8423	0.7737
PE_{trio}	0.8285	0.8251	0.7196
$PE_{motherless}$	0.7216	0.717	0.5846

He_f: gene diversity in females, Ho_f: observed heterozygosity in females, HWE_f: p-values of Hardy- Weinberg equilibrium test in females, PIC: polymorphic information content, PD_f: power of discrimination in females, PD_m: power of discrimination in males. PE: power of exclusion in parentage testing involving a daughter.

Table 2: Overall linkage disequilibrium (LD) and inter-allelic linkage disequilibrium (D') among three X-STR markers in Argentinean males (N =107).

Marker 1	Marker 2	LD[a]	Allele 1_Allele 2	D'[b]
DXS10079	DXS10074	0.086	19_7	1.0000
DXS10079	DXS10075	0.178	15_17	1.0000
			24_17	1.0000
			19_18.2	1.0000
			18_19.2	1.0000
DXS10074	DXS10075	0.000[c]	7_13	0.8237
			17_18.2	1.0000
			17_19.2	1.0000
			7_13	1.0000

[a] p-value of linkage disequilibrium exact test (3200 shufflings).
[b] D': disequilibrium coefficient. Reported only observed D' values > 0.8
[c] Statistically significant difference at $p < 0.015$ (after Bonferroni correction).

ISSN: 2330-0396

significant differences between all population pairs except for Alagoas-Rio, Buenos Aires-Germany and when the comparison involved the Argentinian and Brazilian population samples. This is not unexpected because the three South American populations (Buenos Aires, Alagoas and Rio) were subjected to a similar Amerindian and European (Portuguese for Brazil and Spaniard/

Table 3: Haplotype frequencies of three X-STR markers in males from Buenos Aires population (N = 107).

DXS10079	DXS10074	DXS10075	number	frequency
20	7	13	2	0.01869
18	8	16	3	0.02804
19	8	17	3	0.02804
15	17	17	3	0.02804
19	17	17	3	0.02804
19	17	18	3	0.02804
18	7	13	2	0.01869
18	16	16	2	0.01869
19	16	17	2	0.01869
20	15	17	2	0.01869
20	16	16	2	0.01869
21	17	16	2	0.01869
20	18	17	2	0.01869
19	18	18	2	0.01869
20	17	18	2	0.01869
20	18	17.2	2	0.01869
21	18	17	2	0.01869
21	18	17.2	2	0.01869
20	19	17.2	2	0.01869
22	17	18	2	0.01869
19	7	13	1	0.00935
20	7	14	1	0.00935
17	8	17	1	0.00935
18	8	17	1	0.00935
19	8	16	1	0.00935
18	8	18	1	0.00935
21	8	16	1	0.00935
19	8	18	1	0.00935
18	9	13	1	0.00935
17	9	16	1	0.00935
18	9	17	1	0.00935
19	9	17	1	0.00935
18	13	14	1	0.00935
16	13	17	1	0.00935
19	14	17	1	0.00935
21	14	16	1	0.00935
20	14	18	1	0.00935
19	15	16	1	0.00935
19	15	17	1	0.00935
20	15	16	1	0.00935
21	15	17	1	0.00935
19	15	19	1	0.00935
20	15	18	1	0.00935
14	16	16	1	0.00935
14	16	17	1	0.00935
16	16	16	1	0.00935
17	16	16	1	0.00935
16	16	17	1	0.00935
16	16	18	1	0.00935
20	16	15	1	0.00935
18	16	17	1	0.00935
21	16	15	1	0.00935
20	16	17	1	0.00935
21	16	17	1	0.00935
22	16	17	1	0.00935
23	16	17	1	0.00935
22	16	18	1	0.00935
14	17	16	1	0.00935
17	17	16	1	0.00935
20	17	16	1	0.00935
18	17	18	1	0.00935
18	17	19.2	1	0.00935
19	17	18.2	1	0.00935
21	17	17	1	0.00935
21	17	17.2	1	0.00935
21	17	18	1	0.00935
20	17	19	1	0.00935
22	17	17.2	1	0.00935
24	17	17	1	0.00935
16	18	15	1	0.00935
16	18	16	1	0.00935
15	18	17	1	0.00935
19	18	16	1	0.00935
18	18	17	1	0.00935
19	18	17	1	0.00935
20	18	16	1	0.00935
20	18	19	1	0.00935
23	18	17.2	1	0.00935
14	19	16	1	0.00935
20	19	17	1	0.00935
19	19	18	1	0.00935
20	19	18	1	0.00935

ISSN: 2330-0396

Table 4: Genetic distances based on pairwise comparisons (Fst) and associated p-values (in parentheses) among Buenos Aires, Asian, European and Latin-American populations.

	Buenos Aires	Han	Alagoas	Rio	Japan
Han (China)[20]	0.008				
	(0.000)[a]				
Alagoas (Brazil)[19]	0.003	0.018			
	(0.180)	(0.000)[a]			
Rio (Brazil)[19]	0.002	0.019	0.000		
	(0.225)	(0.000)[a]	(0.360)		
Japan[21]	0.018	0.009	0.022	0.025	
	(0.000)[a]	(0.000)[a]	(0.000)[a]	(0.000)[a]	
Germany[5]	0.005	0.012	0.010	0.013	0.015
	(0.036)	(0.000)[a]	(0.000)[a]	(0.000)[a]	(0.000)[a]

[a]Statistically significant differences at $p < 0.003$ (after Bonferroni correction).

Italian for Argentina) complex admixture process. Although X STR comparative data between Argentina and Germany were not found in the literature, two previous reports showed similarities in allele frequencies between Buenos Aires and other European sample sets when other X STR markers were used (i.e., Spain, Portugal and Italy) [25,26]. The genetic differences observed between Japanese and Chinese populationswere previously reported when autosomal single nucleotide polymorphisms were evaluated [27]. In conclusion, a population database for the cluster DXS10079-DXS10074-DXS10075 was developed for forensic and anthropological purposes. The use of these linked markers in conjunction with other X-STR markers should be considered especially in deficiency cases when a daughter is involved and complex kinship scenarios.

References

1. Gusmao L, Alves C, Gomes I, and Sanchez-Diz P (2012) Capillary electrophoresis of an X-chromosome STR decaplex for kinship deficiency cases. Methods Mol Biol 830: 57-71.
2. Szibor R (2007) X-chromosomal markers: past, present and future. Forensic Sci Int Genet 1: 93-99.
3. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, et al. (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117: 67–74.
4. Hering S, Augustin C, Edelmann J, Heidel M, Dressler J, et al. (2006) DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases. Int J Legal Med 120: 337-345.
5. Szibor R, Hering S, Kuhlisch E, Plate I, Demberger S, et al. (2005) Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 119: 363-369.
6. 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.
7. Hedman M, Palo JU, Sajantila A (2009) X-STR diversity patterns in the Finnish and the Somali population. Forensic Sci Int Genet 3: 173-178.
8. Ribeiro Rodrigues EM, Leite FP, Hutz MH, Palha Tde J, Ribeiro dos Santos AK, et al. (2008) A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. Forensic Sci Int Genet 2: 154-158.
9. Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10: 506-513.
10. Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, et al. (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. Sajantila A, Puomilahti S, Johnsson V, Ehnholm C (1992) Amplification of reproducible allele markers for amplified fragment length polymorphism analysis. Biotechniques 12: 16, 18, 20-22.
12. Szibor R, Edelmann J, Hering S, Plate I, Wittig H, et al. (2003) Cell line DNA typing in forensic genetics--the necessity of reliable standards. Forensic Sci Int 138: 37-43.
13. Tereba A, Tools for analysis of population statistics, in Profiles in DNA1999, Promega Corporation.
14. Lewis P, Zaykin D (2001) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0.
15. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.
16. Gaunt TR, Rodriguez S, Zapata C, and Day IN (2006) MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. BMC Bioinformatics 7: 227.
17. 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.
18. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1: 47-50.
19. Evett W, Weir BS (1998) Interpreting DNA evidence. Statistical genetics for forensic scientists. Sinauer Associates Inc, Sunderland, MA: 209.
20. Medina-Acosta E (2011) Evidence of partial and weak gametic disequilibrium across clusters of pericentromeric short tandem repeats loci on human X chromosome: Proceed with caution in forensic genetics. Forensic Sci Int Genet 5: 545-547.
21. Oki T, Hayashi T, Ota M, Asamura H (2012) Development of a multiplex assay with 16 SNPs on X chromosome for degraded samples. Leg Med 14: 11-16.
22. Ferreira da Silva IH, Barbosa AG, Azevedo DA, Sanchez-Diz P, Gusmao L, et al. (2010) An X-chromosome pentaplex in two linkage groups: haplotype data in Alagoas and Rio de Janeiro populations from Brazil. Forensic Sci Int Genet 4: e95-100.
23. Liu QL, Lu DJ, Quan L, Chen YF, Shen M, et al. (2012) Development of multiplex PCR system with 15 X-STR loci and genetic analysis in three nationality populations from China. Electrophoresis 33: 1299-1305.
24. Samejima M, Nakamura Y, Minaguchi K (2011) Population genetic study of six closely linked groups of X-STRs in a Japanese population. Int J Legal Med 125: 895-900.
25. Bobillo C, Sala A, Gusmao L, Corach D (2011) Genetic analysis of 10 X-STRs in Argentinian population. Forensic Sci Int Genet 5: e14-e16.
26. Noseda PA, Kenline J, Manning S, Gangitano DA (2013) Population data for DXS6800, DXS101 and DXS8377 loci from Buenos Aires (Argentina). J Forensic Leg Med 20: 522-524.
27. Abdulla MA, Ahmed I, Assawamakin A, Bhak J, Brahmachari SK, et al. (2009) Mapping human genetic diversity in Asia. Science 326: 1541-1545.