Int J Legal Med (2008) 122:347–351
DOI 10.1007/s00414-007-0215-1

SHORT COMMUNICATION

Genetic profile characterization and segregation analysis of 10 X-STRs in a sample from Santander, Colombia

Adriana Pico · Adriana Castillo · Clara Vargas · António Amorim · Leonor Gusmão

Received: 17 August 2007 / Accepted: 17 October 2007 / Published online: 8 March 2008

Abstract Ten X-chromosome short tandem repeats (X-STRs: DXS8378, DXS7132, DXS9898, DXS6809, DXS6789, DXS101, GATA172D05, HPRTB, DXS8377, and DXS7423) were analyzed in a sample of unrelated individuals (108 males and 110 females) from the Santander Department in Colombia. In this sample, gene diversities varied between 63.56%, for DXS8378, and 91.41%, for DXS8377. For this set of 10 X-STRs, a high discrimination power was obtained for both male (1 in 3×10^6) and female (1 in 9×10^{10}) samples and a high mean exclusion chance in father/daughter duos (99.993%) and in father/mother/daughter trios (99.9999%), demonstrating the usefulness of this set of markers in forensic and kinship analysis. Hardy–Weinberg equilibrium was tested in the female sample and no significant deviations were found. Pairwise analysis showed significant differences in the comparison with samples from Spain, Peru, and Argentina and with African American and Hispanic samples from New York. This same set of X-STRs was also typed in 51 mother/father/daughter trios, 43 mother/son duos, and in a single father/daughter pair. In total, four mutations were observed; one at DXS7132 and at DXS6809, and two at DXS8377. Two mutations were paternal and one maternal; and to a fourth mutation, it was not possible to define its origin.

Keywords X-chromosome · STRs · Santander Department · Colombia · Mutation rate

Electronic supplementary material The online version of this article (doi:10.1007/s00414-007-0215-1) contains supplementary material, which is available to authorized users.

A. Pico (✉) · A. Castillo · C. Vargas
Genetic Laboratory, UIS–Industrial University of Santander,
Carrera 32, No. 29–31,
Bucaramanga, Colombia
e-mail: labgen2@uis.edu.co

A. Amorim · L. Gusmão
IPATIMUP, Institute of Molecular Pathology and Immunology
of the University of Porto,
4200-465 Porto, Portugal

A. Amorim
Faculty of Sciences, University of Porto,
4050 Porto, Portugal

Introduction

The study of X-chromosome short tandem repeats (X-STRs) is, at the present time, of great interest in population genetic studies, specially in the forensic field and in kinship testing [8, 21]. Many studies revealed their efficiency in supplementing autosomal and Y-chromosomal STR analyses [7, 13, 16].

Several X-STRs have been recently validated for forensic use. Nevertheless, further studies are still needed on allele frequency distributions in different populations, mutation rates, and linkage disequilibrium to establish reference population databases [2, 12, 15]. In the case of X-STRs, it is particularly important to screen possible associations between loci (linkage disequilibrium). This can be specially important in admixed and possibly substructured populations, such as in South American populations with Amerindian, African, and Caucasian ancestry. Both historical and genetic data have shown that this is the case in Colombia [5, 18–20].

The aim of the present study was to type 10 X-STR loci, using a single multiplex PCR-based system, to establish the genetic profile of a Colombian (Santander Department) population and also to compare our results with those from historically related populations.

Reliable estimates of mutation rates are important in kinship analysis, however, few X-STRs segregation studies have been reported for these markers. For this reason, we also report the results of the analysis of meiotic transfers in a sample where biological kinship was controlled with autosomal markers.

Materials and methods

DNA samples

Samples were selected from individuals living in Santander, Colombia that have been submitted for routine paternity tests. Unrelated father and mother samples (108 males and 110 females) were selected for allele frequency estimations. In cases where the biological relationship had been previously confirmed by using autosomal markers, the offspring was also typed for segregation analysis (in total 51 trios involving a daughter, 43 trios involving a son, and a single father/daughter pair). The biological relationship was previously confirmed by using autosomal STRs with paternity and maternity index values above 10,000. Autosomal STRs were typed using the Powerplex 16 kit (Promega Corporation, Madison, WI, USA), following the manufacturer's instructions.

Genomic DNA was extracted from peripheral blood samples using the salting out method according to Miller et al. [14].

Markers genotyping

The 10 X-STR loci (DXS8378, DXS7132, DXS9898, DXS6809, DXS6789, DXS101, GATA172D05, HPRTB, DXS8377, and DXS7423) were amplified in one PCR multiplex reaction using the protocol described by Gomes et al. [12]. Amplicons were separated in an ABI 310 Sequencer (AB Applied Biosystems, Foster City, CA, USA) and typed using the Genescan 2.1.1 Analysis software. Allele designations were based on comparison with reference DNA control sample 9947A (female) from Promega commercial kits (Promega Corporation) and following the same nomenclature as in Gomes et al. [12].

Statistical analysis

Allelic frequencies, gene diversities, and population pairwise genetic distances (R_{ST} and F_{ST}) were calculated using the ARLEQUIN ver 3.0 software [10]. The same software was used to perform exact test of population differentiation and linkage disequilibrium test. Genetic distance estimations were based on the sum of squared size differences (R_{ST}), except in the case of DXS9898 where a high number of nonconsensus 8.3 alleles prevented the use of this method. In this case, we used the method based on the number of different alleles (F_{ST}) in accordance with Pereira et al. [17].

Statistics for forensic efficiency evaluation of each loci, namely, expected probability of exclusion (PE) in trios involving daughters (PE_T) and in father/daughter duos (PE_D), power of discrimination in females (PD_F) and in males (PD_M) were calculated according to the method of Desmarais et al. [6].

Results and discussion

A Colombian population sample of 218 genetically unrelated individuals (108 males and 110 females) living in Santander Department was studied for the 10 X-chromosome markers. Allele frequency distributions observed for the 10 loci in male and female samples were compared by an exact test of differentiation. Exact test p values varied between 0.01338± 0.0018 for DXS7423 and 0.85762±0.0103 for HPRTB (data not shown). After applying Bonferroni correction for multiple tests (significance level, 0.005), no significant differences were found for any loci between the male and female subgroups. In view of this, samples were pooled and allele frequencies are presented in Table 1. Alleles 13 and 14 not previously described in other populations were found in the GATA172D05 loci.

Hardy–Weinberg equilibrium was tested in the female sample and no deviations were detected when the 10 loci were analyzed (exact p values ≥0.08237; Table 2).

Linkage disequilibrium analysis

Linkage disequilibrium was tested for all pairs of loci in the male sample. Exact p values below 5% were only found in 3 out of the 45 pairwise comparisons (p=0.0069, between DXS8378 and DXS101 loci; p=0.0060, between DXS9898 and DXS101; and p=0.0264, between DXS7132 and DXS6809). Nevertheless, all tests released values above the significance level of 0.0011 (obtained after Bonferroni correction) providing no evidence of association between the studied markers Thus, genetic substructure in the present admixed population is not reflected in our sample for these markers. These results are in accordance with the distances between the selected X-chromosome loci (>5 Mb), except for DXS6809 and DXS6789 that are closely located and have been shown to be in strong linkage disequilibrium in a German population sample [22]. Therefore, although no evidences of association were found between these two loci, to allow future comparisons and sample size enlargements, haplotype frequencies for these two loci were included in Table S2.

Table 1 Allele frequencies for ten X-STRs in a Santander, Colombia population sample

Allele	X-STR loci	SD
DXS8378		
10	0.4817	0.0276
11	0.2805	0.0248
12	0.2317	0.0233
13	0.0031	0.0030
14	0.0031	0.0030
DXS7132		
11	0.0061	0.0043
12	0.0671	0.0138
13	0.2317	0.0233
14	0.3232	0.0258
15	0.1951	0.0219
16	0.0823	0.0151
17	0.0701	0.0141
18	0.0152	0.0067
19	0.0092	0.0052
DXS9898		
7	0.0031	0.0030
8.3	0.1098	0.0172
10	0.0122	0.0060
11	0.0732	0.0144
12	0.3293	0.0259
13	0.3933	0.0270
14	0.0762	0.0146
15	0.0031	0.0030
DXS6809		
28	0.0122	0.0060
29	0.0061	0.0043
30	0.0366	0.0103
31	0.0915	0.0159
32	0.1738	0.0209
33	0.3293	0.0259
34	0.2713	0.0245
35	0.0457	0.0115
36	0.0213	0.0079
37	0.0122	0.0060
DXS6789		
15	0.0366	0.0103
16	0.1250	0.0182
17	0.0061	0.0043
18	0.0031	0.0030
19	0.0518	0.0122
20	0.4756	0.0276
21	0.1768	0.0210
22	0.0945	0.0161
23	0.0244	0.0085
24	0.0031	0.0030
25	0.0031	0.0030
DXS101		
15	0.0031	0.0030
16	0.0061	0.0043
18	0.0183	0.0074
19	0.0243	0.0085
20	0.0061	0.0043
21	0.0305	0.0095
22	0.0183	0.0074
23	0.0762	0.0146
24	0.3079	0.0255
25	0.1585	0.0201
26	0.1738	0.0209
27	0.0976	0.0164
28	0.0488	0.0119
29	0.0244	0.0085
30	0.0061	0.0043
GATA172D05		
6	0.1098	0.0172
7	0.0031	0.0030
8	0.1768	0.0210
9	0.0640	0.0135
10	0.3293	0.0259
11	0.1799	0.0212
12	0.1281	0.0184
13	0.0061	0.0043
14	0.0031	0.0030
HPRTB		
11	0.0061	0.0043
12	0.0915	0.0159
13	0.2317	0.0233
14	0.3415	0.0262
15	0.2348	0.0234
16	0.0762	0.0146
17	0.0092	0.0052
18	0.0092	0.0052
DXS8377		
39	0.0061	0.0043
40	0.0092	0.0052
41	0.0335	0.0099
42	0.0488	0.0119
43	0.0457	0.0115
44	0.0915	0.0159
45	0.0854	0.0154
46	0.1342	0.0188
47	0.1189	0.0178
48	0.0884	0.0157
49	0.1067	0.0170
50	0.0915	0.0159
51	0.0518	0.0122
52	0.0213	0.0079
53	0.0305	0.0095
54	0.0274	0.0090
55	0.0031	0.0030
56	0.0031	0.0030
58	0.0031	0.0030
DXS7423		
13	0.0427	0.0111
14	0.2896	0.0250
15	0.4543	0.0275
16	0.0793	0.0149
17	0.1342	0.0188

SD: standard deviation

Table 2 Parameters of forensic interest in 108 male and 110 female of Santander Department, Colombia

	HET obs	HET esp	HWE exact test	PE_T	PE_D	PD_M	PD_F
DXS8378	0.60909	0.63130	0.82847±0.00050	0.5657	0.4174	0.6356	0.7973
DXS7132	0.72727	0.78402	0.36990±0.00061	0.7573	0.6293	0.7873	0.9248
DXS9898	0.69091	0.69971	0.58655±0.00480	0.6673	0.5257	0.7135	0.8718
DXS6809	0.73636	0.77011	0.50753±0.00062	0.7428	0.6122	0.7752	0.9171
DXS6789	0.66364	0.73196	0.08237±0.00036	0.6836	0.5427	0.7133	0.8881
DXS101	0.76364	0.82740	0.17466±0.00029	0.8107	0.6985	0.8292	0.9523
GATA172D05	0.81818	0.81233	0.37713±0.00049	0.7677	0.6414	0.7953	0.9304
HPRTB	0.74545	0.76916	0.42834±0.00052	0.7224	0.5868	0.7602	0.9046
DXS8377	0.96364	0.92026	0.75715±0.00046	0.9076	0.8368	0.9141	0.9862
DXS7423	0.60909	0.67277	0.41779±0.00070	0.6336	0.4883	0.6837	0.8498
Combined %				99.99990	99.993	99.99997	99.9999999990

HET obs: observed heterozygosity, *HET esp*: expected heterozygosity, *HWE exact test*: *p* value HWE analysis, PE_T: expected probability of exclusion in trios involving daughters, PE_D: expected probability of exclusion in mother/son duos, PD_M: power of discrimination in males, PD_F: power of discrimination in females.

Forensic efficiency parameters

Based on frequency distributions estimated in the global sample (Table 1), forensic efficiency statistical parameters were calculated for each locus in the Santander sample, and are presented in Table 2. In our population sample, DXS8378 was the least polymorphic marker whereas the most diverse was DXS8377. Overall values obtained for the power of discrimination where high in females (PD_F= 0.99999999999) and males (PD_M=0.9999997), and combined mean exclusion chances in trios (PE_T=0.999999) and duos (PE_D=0.99993).

Population comparisons

The allele frequency distributions in our sample were compared with data available for the same markers studied in other populations from the same continent [9, 11, 12, 15] and also with data available for the same 10 STRs in a Spanish sample [1] because Spain was one of the main sources of emigration that contributed to the Colombian gene pool. Although available data on other South American populations are scarce, in the few possible comparisons, low, nonsignificant genetic distances were observed (see Table S1). Nevertheless, in comparisons with a Spanish sample, we found high genetic distances for these 10 X-STRs. When performing comparisons with data available for African American and Hispanic samples, significant differences were also observed in 9 out of the 10 comparison with African Americans, decreasing to 2 out of 10 comparisons with the Hispanic sample.

This contrasts with what was previously observed for Y-STRs where urban populations from Colombia cluster with European origin samples and stands far apart from the non-Caucasian ones [3, 4]. However, this is not a surprising result taking into account the highly asymmetric pattern of mating that has been observed in most South American countries with maternal lineages predominantly from Amerindian origin and a high proportion of male contribution of European origin in Colombia [5].

Segregation analysis

The same set of 10 X-STR loci was typed in a total of 51 mother/father/daughter trios, 43 mother/son duos, and in a single father/daughter pair. Four mutations were observed; one at DXS7132 and at DXS6809, and two at DXS8377 (Table 3). Two of the mutations were paternal and one maternal, and for a fourth mutation, it was not possible to define its origin. In three cases, the mutations were single step and in one case, for DXS8377, it was found a two-step mutation (allele 53 to 55). Because this STR presents a complex structure with trinucleotide and hexanucleotide repeat motifs, sequence analysis was carried out. The result was compatible with the gain of two repeats occurred in the

Table 3 Observed mutations in a total of 51 mother/father/daughter trios, 43 mother/son duos, and in a single father/daughter for X-chromosome STR markers

Locus	Case		Genotype	Age[a]
DXS7132	1	Father	13	32
		Mother	13–15	23
		Daughter	13–14	–
DXS6809	2	Father	34	48
		Mother	33–34	39
		Daughter	33–35	–
DXS8377	3	Father	53	30
		Mother	45–46	30
		Daughter	46–55	–
	4	Mother	46–58	14
		Son	59	–

The paternity and maternity indexes for the four cases are $>1.5\times10^6$.
[a] Father's and mother's age at time of birth (years).

trinucleotide motif from father [$(AGA)_{33}$-(GGA-AGA)$_5$-$(AGA)_2$-GGA-$(AGA)_6$] to son [$(AGA)_{35}$-(GGA-AGA)$_5$-$(AGA)_2$-GGA-$(AGA)_6$]. It is worth mentioning that the observed allele structures correspond to one repeat less than that determined by fragment size analysis. This result is explained by a 3-bp insertion (AGG) at the 5′ repeat flanking region in both the father and son alleles, giving rise to the following sequence:

Forward primer—AGAGAGGAGAAGAAGGAGAAG GAGGAGA $(AGG)_2$—Repeat.

In conclusion, the 10 X-STRs studied in this work demonstrate to be highly informative and present a high discrimination power. The differences obtained in the allele frequency distributions between our sample and those on other populations confirms the need for using population-specific databases for forensic casework and kinship analysis in Santander.

Acknowledgements The authors would like to acknowledge Iva Gomes for her collaboration. IPATIMUP is partially supported by Fundação para a Ciência e a Tecnologia, through Programa Operacional Ciência e Inovação 2010 (POCI 2010).

References

1. Aler M, Sánchez-Diz P, Gomes I, Gisbert M, Carracedo A, Amorim A, Gusmão L (2007) Genetic data of 10 X-STRs in a Spanish population sample. Forensic Sci Int 173:193–196
2. Bini C, Ceccardi S, Ferri G, Pelotti S, Alú M, Rocaglia E, Beduschi G, Caenazzo L, Ponzano E, Tasinato P, Turchi C, Mazzanti M, Tagliabracci A, Toni C, Spinetti I, Domenico R, Presciuttini S (2005) Development of a heptaplex PCR system to analyse X-chromosome STR loci from five Italian population simples. A collaborative study. Forensic Sci Int 153:231–236
3. Builes JJ, Bravo ML, Gomez C et al (2006) Y-chromosome STRs in an Antioquian (Colombia) population sample. Forensic Sci Int 164:79–86
4. Builes JJ, Martinez B, Gomez A et al (2007) Y chromosome STR haplotypes in the Caribbean city of Cartagena (Colombia). Forensic Sci Int 167:62–69
5. Carvajal-Carmona LG, Soto ID, Pineda N et al (2000) Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in Northwest Colombia. Am J Hum Genet 67:1287–1295
6. 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
7. Edelmann J, Hering S, Michael M et al (2001) 16 X-chromosome STR loci frequency data from a German population. Forensic Sci Int 124:215–218
8. Edelmann J, Deichsel D, Hering S et al (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
9. Edelmann J, Hering S, Brundirs N, Lessig R, Kuhlisch E, Szibor R (2004) Allele frequencies for X-chromosomal microsatellites in different populations. In: von Doutrempuich C, Morling N (eds) Progress in forensic genetics, 10. Elsevier, Amsterdam, pp 263–265
10. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50
11. Gangitano DA, Garofalo MG, Juvenal GJ, Budowle B, Padula RA (2001) Distribution of HumHPRTB and HumF13A01 alleles in Buenos Aires population (Argentina). J Forensic Sci 46:418
12. Gomes I, Prinz M, Pereira R, Meyers C, Mikulasovich RS, Amorim A, Carracedo A, Gusmão L (2006) Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int J Legal Med 121(3):198–203. DOI 10.1007/s00414-006-0146-2
13. Hering S, Augustin C, Edelmann J et al (2005) 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 116:144–149
14. Miller SA, Dykes DD, Polesky HF (1988) A single salting out procedure for extracting DNA from human nucleated cells. Nucl Acids Res 16:1215
15. Moreno MA, Builes JJ, Jaramillo P, Espinal C, Aguirre D, Pancorbo MM, Gusmão L, Bravo MLJ (2006) Validation of five X-chromosomal STRs DXS6800, DXS6807, DXS6798, DXS8377, and DXS7423 in an Antioquian population simple. In: Progress in Forensic Genetics, 11. International Congress Series 1288. Elsevier, Amsterdam, pp 295–297
16. Oguzturun C, Thacker CR, Syndercombe D (2006) Population study of four X-chromosomal STR loci in the UK and Irish population. In: Progress in Forensic Genetics, 11. International Congress Series 1288. Elsevier, Amsterdam, pp 283–285
17. Pereira R, Gomes I, Amorim A, Gusmão L (2006) Genetic diversity of 10 X-chromosome STRs in northern Portugal. Int J Legal Med 121:192–197. DOI 10.1007/s00414-006-0144-4
18. Rodas C, Gelvez N, Keyeux G (2003) Mithocondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and Mestizo populations. Hum Biol 75:13–30
19. Sánchez-Albornoz N (1977) La población de America Latina: desde los tiempos precolombinos al año 2000, 2nd edn. Alianza Editorial, Madrid
20. Sandoval C, de la Hoz A, Yunnis E (1993) Estrutura genética de la poblacion Colombiana. Rev Fac Med Univ Nac Colombiana 41:3–14
21. Shin SH, Yu JS, Park WS, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41
22. 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