Int J Legal Med (2011) 125:749–752
DOI 10.1007/s00414-011-0591-4

SHORT COMMUNICATION

Allele and haplotype diversity of X-chromosomal STRs in Ivory Coast

Serena Pasino · Stefano Caratti · Massimiliano Del Pero · Alfredo Santovito · Carlo Torre · Carlo Robino

Received: 21 March 2011 / Accepted: 16 June 2011 / Published online: 30 June 2011
© Springer-Verlag 2011

Abstract Twenty-one X-chromosomal short tandem repeat (STR) loci, including the six clusters of linked markers DXS10148–DXS10135–DXS8378 (Xp22), DXS7132–DXS10079–DXS10074 (Xq12), DXS6801–DXS6809–DXS6789 (Xq21), DXS7424–DXS101 (Xq22), DXS10103–HPRTB–DXS10101 (Xq26), DXS8377–DXS10146–DXS10134–DXS7423 (Xq28) and the loci DXS6800, GATA172D05 and DXS10011 were typed in a popula3tion sample from Ivory Coast (n =125; 51 men and 74 women). Allele and haplotype frequencies as well as linkage disequilibrium data for kinship calculations are provided. On the whole, no significant differences in the genetic variability of X-STR markers were observed between Ivorians and other sub-Saharan African populations belonging to the Niger–Kordofanian linguistic group.

Keywords X chromosome · Short tandem repeat · Ivory Coast · Haplotype · Linkage disequilibrium

Electronic supplementary material The online version of this article (doi:10.1007/s00414-011-0591-4) contains supplementary material, which is available to authorized users.

S. Pasino · S. Caratti · C. Torre · C. Robino (✉)
Department of Anatomy, Pharmacology and Legal Medicine, University of Turin,
Corso Galileo Galilei 22,
10126 Turin, Italy
e-mail: carlo.robino@unito.it

S. Pasino · M. Del Pero · A. Santovito
Department of Animal and Human Biology, University of Turin, Turin, Italy

Introduction

X chromosomal short tandem repeat (STR) loci combine the characteristics of uniparental and autosomal genetic markers and therefore possess highly desirable features for paternity testing, especially in deficiency paternity cases with female offspring [1]. Moreover, X-STR haplotype analysis can be used to detect kinship between alleged relatives in large and incomplete pedigrees [2, 3]. Because of the distinctive genetic properties of the X chromosome —low genetic diversity, strong effects of genetic drift, and long linkage disequilibrium (LD) intervals—the forensic application of X-STR markers requires precise knowledge of genetic variability in relevant populations [4]. At present, the number of studies describing X-STR allele and haplotype diversity in sub-Saharan Africa is very limited [5, 6]. In order to provide detailed population data for kinship calculations, we analyzed the genetic diversity of a large set of X-STR markers in a sample from Ivory Coast (northwest Africa). Markers included in the panel comprise six clusters of closely linked loci (<3 cM) for haplotype testing: DXS10148–DXS10135–DXS8378 (Xp22); DXS7132–DXS10079–DXS10074 (Xq12); DXS6801–DXS6809–DXS6789 (Xq21); DXS7424–DXS101 (Xq22); DXS10103–HPRTB–DXS10101 (Xq26); DXS8377–DXS10146–DXS10134–DXS7423 (Xq28). Their physical and genetic localization on the X chromosome is shown in Table S1.

Materials and methods

Blood samples (n=125) were obtained after informed consent from unrelated healthy adults (74 females and 51 males) residing in the rural village of Ouangolodougou,

located in northwestern Ivory Coast (9°58′N, 5°09′W). Sampling was anonymous in order to prevent linkage to the original donor. DNA was extracted using the ChargeSwithch gDNA Normalized Buccal Cell kit (Invitrogen, Carlsbad, CA) and the automatic magnetic separator KingFisher mL (Thermo Labsystems, Vantaa, Finland) following the manufacturer's instructions.

Amplification of X-STR markers was performed in four separate multiplex reactions. Loci DXS10135, DXS8378, DXS7132, DXS10074, HPRTB, DXS10101, DXS10134 and DXS7423 were amplified by means of the Mentype Argus X-8 kit (Biotype GE, Dresden Germany) according to the manufacturer's instructions. The remaining X-STRs were amplified with three in-house developed PCR assays as described by Bekada et al. [7]. Typing was done by capillary electrophoresis on the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) in comparison to allelic ladders and control DNA from K562 and 9947A cell lines (Promega, Madison, WI) using the GeneScan and Genotyper software version 3.7 (Applied Biosystems) [8]. Sequencing of DXS6801 and DXS10148 variant alleles was performed with the Big Dye Terminator Sequencing kit (Applied Biosystems), using the unlabelled multiplex PCR primers.

The following statistical parameters of forensic interest were calculated using the on-line functions provided by the ChrX-STR.org 2.0 database (http://www.chrx-str.org) [9]: expected heterozygosity (h), polymorphism information content (PIC), power of discrimination in females (PD^F) and males (PD^M), mean exclusion chance for autosomal markers in trios (MECI), X-STRs in trios with daughters (MECII), and X-STRs in father/daughter duos (MECIII). Haplotype diversity, exact test of differentiation (100,000 Markov steps) between male and female allele frequencies, test of Hardy–Weinberg equilibrium (HWE) in the female subsample, pairwise test of LD between pairs of markers within clusters of linked loci in the male subsample, and locus by locus population pairwise genetic distances (F_{ST}) between Ivorians and other populations were calculated with Arlequin software [10]. Principal component analysis (PCA) was performed by means of the R-package software v2.0.1 (http://www.r-project.org). Reduced median networks describing phylogenetic relationships between populations were obtained with the Network 4.1.1.2 software (www.fluxus-engineering.com), using the reduced median algorithm (r=2) followed by the median-joining algorithm (ε=0) [11, 12].

Results and discussion

No significant differences in allele distribution between the male and female subsamples were observed, based on the global test of population differentiation. Allele frequencies for the 21 X-STR markers were therefore combined over both sexes and are reported in Table S2. For ten loci, these are, to our knowledge, the first population data reported in the literature for sub-Saharan Africa [5, 6]. Among them, locus DXS6801 shows a relatively high occurrence of alleles <10 (7.5%), previously unobserved in worldwide populations [9]. At locus DXS10148, 11.5% of the analyzed chromosomes carry alleles with long repeat motifs (>36). Such variants were not detected in the original study describing DXS10148 variability in Europeans [13], but are present in Algeria (4.9%) [7]. They possibly represent a trace of the sub-Saharan African contribution to the genetic pool of northern African populations, already evidenced by Y chromosome and mitochondrial DNA studies [14, 15]. Rare variants observed in hemizygous males were confirmed by sequencing and a selection of electropherograms is shown in Fig. S1. It can be seen that, whereas DXS6801 short alleles display a conventional $[ATCT]_x–N_7–[ATCT)]_2$ motif [9], the repeat structure of DXS10148 long variants is complex and slightly different from the one originally described by Hundertmark et al. [7, 13].

Based on the observed and expected distribution of genotypes in the female subsample, and after Bonferroni correction for multiple testing, a single significant deviation from HWE (p<0.002) was detected at locus DXS10146, due to a reduction of observed heterozygosity (0.676) compared to expected heterozygosity (0.937). Interestingly, we have observed the same significant excess of homozygotes at locus DXS10146 in a similarly typed Italian population sample [16]. Allele drop-out due to low efficiency of the used PCR assay is an unlikely explanation for this result, since peak height ratio in DXS10146 heterozygous genotypes appeared extremely well balanced (0.855) and comparable to that observed in strictly validated STR typing commercial kits [17]. A possible alternative reason may be the presence, in both populations, of variants in the primer binding sites generating null alleles [18]. Concordance studies with different set of primers are therefore required to clarify the significance of these findings.

Statistical parameters of forensic interest calculated for the 21 X-STR loci are shown in Table S3. The p values of pairwise test of LD and the complete list of haplotypes observed in the male subsample are given in Table S4. Significant LD was observed exclusively between two pairs of markers, located in Xq26 (DXS10103–DXS10101, p=0.025) and Xq28 (DXS10146–DXS7423, p=0.001). The limited extent of LD found in the Ivorian sample, compared to Europeans [19], reflects the longer evolutionary history and relatively constant population size that shaped the genetic background of sub-Saharan African populations

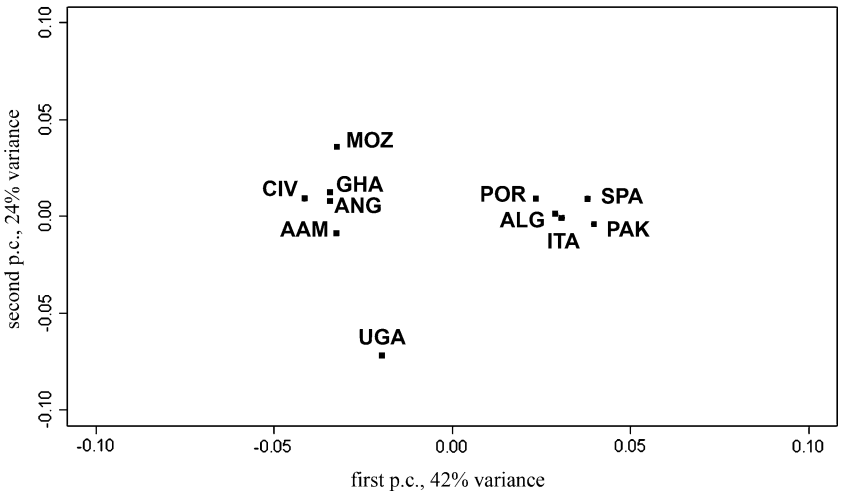

Fig. 1 PCA plot for Ivorians and ten reference populations: *ANG* Angola [5], *MOZ* Mozambique [5]; *UGA* Uganda [5], *GHA* Ghana [6], *ALG* Algeria [7], *ITA* Italy [16, 22], *AAM* African Americans [23], *POR* Portugal [24], *SPA* Spain [25], *PAK* Pakistan [26, 27]

[20], and should be taken into account in kinship testing calculations [21].

Locus by locus pairwise genetic distances (F_{ST}) between Ivorians and other relevant population samples from Angola, Mozambique, Uganda [5], Ghana [6], Algeria [7], Italy [16, 22], African Americans [23], Portugal [24], Spain [25] and Pakistan [26, 27] are shown in Table S5. Since at locus HPRTB two different nomenclatures exhibiting one repeat difference are currently used in forensic literature [28], population data were revised in order to make them consistent with the allele designation adopted in the present study [8]. It can be noticed that most of the statistically significant F_{ST} values were found in comparison with non-sub-Saharan African populations. The PCA plot summarizing the relationships between Ivorians and other populations, based on a subset of four unlinked loci (DXS8378, DXS7132, DXS101, and DXS8377), is shown in Fig. 1. As can be seen, Ivorians tightly cluster with sub-Saharan African populations belonging to the Niger–Kordofanian linguistic group, whereas the outlier position of Uganda can be explained by the more diversified genetic structure of middle-eastern Africa [29].

In Table 1, haplotype diversity calculated for the six clusters of linked loci in the Ivorian population sample is compared to that previously observed in Algerians and Italians [7, 16]. X-chromosomal variability in the sample from sub-Saharan Africa is generally greater than that found in the Mediterranean basin, as expected according to the Out-of-Africa evolutionary model for Eurasian populations [20]. The high genetic diversity of Ivorians is also evident in the networks depicted in Fig. S2 (the cluster in Xp22 was excluded from the analysis due to the complex repeat structure of locus DXS10148). It can be seen that the most common haplotypes were predominantly shared between Italian and Algerian individuals, whereas the Ivorian sample appears to be more diversified and located on peripheral ramifications of the networks. The observed pattern is coherent with the results of a previous study of the X chromosome, performed on a macro-geographic scale, confirming a recent out-of-Africa expansion for human populations [30].

The demand for genetic testing in families from sub-Saharan Africa is constantly increasing. This entails not only conventional paternity cases, but also complex kinship analysis in immigration procedures, where DNA can assist the reunification with undocumented relatives [31]. In such circumstances, the recourse to X-chromosomal markers can be highly effective in discriminating between alternative pedigrees [32]. The allele and haplotype frequencies reported here provide the necessary data for kinship calculations in cases involving subjects of western African descent.

Table 1 Haplotype diversity of X-STR clusters in Ivorians, Algerians and Italians

	Ivory Coast	Algeria	Italy
DXS10148–DXS10135–DXS8378	0.998	0.998	0.998
DXS7132–DXS10079–DXS10074	0.994	0.994	0.984
DXS6801–DXS6809–DXS6789	0.994	0.989	0.983
DXS7424–DXS101	0.981	0.979	0.977
DXS10103–HPRTB–DXS10101	0.986	0.989	0.985
DXS8377–DXS10146–DXS10134–DXS7423	1	0.999	1

Acknowledgements The continuing support of the Compagnia di San Paolo to C. Torre's laboratory is gratefully acknowledged.

References

1. Szibor R (2007) X-chromosomal markers, past, present and future. Forensic Sci Int Genet 1:93–99
2. 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
3. Hering S, Augustin C, Edelmann 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
4. Schaffner SF (2004) The X chromosome in population genetics. Nat Rev Genet 5:43–51
5. Gomes I, Alves C, Maxzud K et al (2007) Analysis of 10 XSTRs in three African populations. Forensic Sci Int Genet 1:208–211
6. Poetsch M, El-Mostaqim D, Tschentscher F, Browne EN, Timmann C, Horstmann RD, Von Wurmb-Schwark N (2009) Allele frequencies of 11 X-chromosomal loci in a population sample from Ghana. Int J Legal Med 123:81–83
7. Bekada A, Benhamamouch S, Boudjema A, Fodil M, Menegon S, Torre C, Robino C (2010) Analysis of 21 X-chromosomal STRs in an Algerian population sample. Int J Legal Med 124:287–294
8. Szibor R, Edelmann J, Hering S et al (2003) Cell line DNA typing in forensic genetics—the necessity of reliable standards. Forensic Sci Int 138:37–43
9. 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
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. Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753
12. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48
13. Hundertmark T, Hering S, Edelmann J, Augustin C, Plate I, Szibor R (2008) The STR cluster DXS10148–DXS8378–DXS10135 provides a powerful tool for X-chromosomal haplotyping at Xp22. Int J Legal Med 122:489–492
14. Robino C, Crobu F, Di Gaetano C et al (2008) Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample. Int J Legal Med 122:251–255
15. Plaza S, Calafell F, Helal A, Bouzerna N, Lefranc G, Bertranpetit J, Comas D (2003) Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. Ann Hum Genet 67:312–328
16. Inturri S, Menegon S, Amoroso A, Torre C, Robino C (2010) Linkage and linkage disequilibrium analysis of X-STRs in Italian families. Forensic Sci Int Genet. doi:10.1016/j.fsigen.2010.10.012
17. Hill CR, Duewer DL, Kline MC et al (2010) Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex((R)) ESX 17 and ESI 17 Systems. Forensic Sci Int Genet. doi:10.1016/j.fsigen.2010.03.014
18. Leibelt C, Budowle B, Collins P et al (2003) Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. Forensic Sci Int 133:220–227
19. Tillmar AO, Mostad P, Egeland T, Lindblom B, Holmlund G, Montelius K (2008) Analysis of linkage and linkage disequilibrium for eight X-STR markers. Forensic Sci Int Genet 3:37–41
20. Tishkoff SA, Williams SM (2002) Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet 3:611–621
21. Tillmar AO, Egeland T, Lindblom B, Holmlund G, Mostad P (2010) Using X-chromosomal markers in relationship testing: calculation of likelihood ratios taking both linkage and linkage disequilibrium into account. Forensic Sci Int Genet. doi:10.1016/j.fsigen.2010.11.004
22. 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
23. Gomes I, Prinz M, Pereira R et al (2007) Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int J Legal Med 121:198–203
24. 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
25. Aler M, Sánchez-Diz P, Gomes I et al (2007) Genetic data of 10 X-STRs in a Spanish population sample. Forensic Sci Int 173:193–196
26. Tariq MA, Ullah O, Riazuddin SA, Riazuddin S (2008) Allele frequency distribution of 13 X-chromosomal STR loci in Pakistani population. Int J Legal Med 122:525–528
27. Tariq MA, Sabir MF, Riazuddin SA, Riazuddin S (2009) Haplotype analysis of two X-chromosome STR clusters in the Pakistani population. Int J Legal Med 123:85–87
28. Szibor R, Edelmann J, Hering S, Gomes I, Gusmão L (2009) Nomenclature discrepancies in the HPRTB short tandem repeat. Int J Legal Med 123:185–186
29. Tishkoff SA, Reed FA, Friedlaender FR et al (2009) The genetic structure and history of Africans and African Americans. Science 324:1035–1044
30. Santos-Lopes SS, Pereira RW, Wilson IJ, Pena SD (2007) Aworldwide phylogeography for the human X chromosome. PLoS One 2:e557
31. Karlsson AO, Holmlund G, Egeland T, Mostad P (2007) DNA-testing for immigration cases: the risk of erroneous conclusions. Forensic Sci Int 172:144–149
32. Pinto N, Gusmão L, Amorim A (2011) X-chromosome markers in kinship testing: a generalisation of the IBD approach identifying situations where their contribution is crucial. Forensic Sci Int Genet 5:27–32