Int J Legal Med (2002) 116:368–371
DOI 10.1007/s00414-002-0316-9

SHORT COMMUNICATION

María T. Zarrabeitia · Teresa Amigo · Carolina Sañudo
Marian Martínez de Pancorbo · José A. Riancho

Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377

Received: 11 December 2001 / Accepted: 14 May 2002 / Published online: 26 June 2002

Abstract DXS7423 and DXS8377 are two microsatellite markers located in the q28 band of chromosome X. We developed a protocol to amplify both markers in a single reaction, sequenced the most common alleles and studied allele frequencies in a Spanish population sample. DXS7423 allele variability was due to different numbers of (TCCA) repeats and five different alleles were found with apparent sizes between 181 and 197 bp. The probability of discrimination (PD) was 87% for female samples, and the expected probability of exclusion (PE) was 71%. DXS8377 appeared as a highly polymorphic marker with variable numbers of (CTC), (TCC) and (TTC) repeats. We found 18 alleles of different sizes (204–258 bp) and the PD and PE were 99% and 93%, respectively. These data suggest that DXS7423 and DXS8377 can be very useful markers for genetic forensic studies.

Keywords X-chromosome · Short tandem repeat · Paternity testing · DNA

Introduction

In recent years the analysis of Y chromosome markers has been the subject of considerable interest. They are very useful in paternity cases when the alleged father is not available for study. Since all males in the paternal lineage share the same Y chromosome haplotype, the haplotype of the alleged father can be inferred by studying other males in the family lineage. Of course, such a study only has utility in cases involving males. As fathers transmit their Y chromosome to all their sons, they also transmit the same X chromosome to all their daughters. Therefore, the analysis of X-STRs could be of great usefulness in deficiency paternity cases with a female alleged child.

However, studies of X-linked STRs are scanty in forensic medicine. Only one STR (HPRTB) can be analysed with a commercial method, and only a few other markers have been proposed as useful for forensic purposes [1, 2, 3, 4, 5, 6, 7, 8]. While studying the localisation of candidate genes causing X-linked myotubular myopathy, Hu et al. [9] reported two other polymorphic STRs, but their sequence structure has not been well established. Therefore the aim of this study was to characterise the structure of those STRs and explore their potential utility for forensic genetic testing.

M.T. Zarrabeitia (✉) · C. Sañudo
Unit of Legal Medicine, Faculty of Medicine,
University of Cantabria, Av. Herrera Oria s/n,
39011 Santander, Spain
e-mail: zarrabet@unican.es,
Tel.: +34-942-201984, Fax: +34-942-201903

J.A. Riancho
Unit of Internal Medicine, Faculty of Medicine,
University of Cantabria, Santander, Spain

T. Amigo
Unit of Pediatrics, Faculty of Medicine, University of Cantabria,
Santander, Spain

M.M. de Pancorbo
Department of Zoology, Faculty of Pharmacy,
University of Basque Country, Vitoria, Spain

Materials and methods

DNA amplification

Genomic DNA was extracted from peripheral blood by a commercial method, according to the manufacturer's instructions (Qiagen, Hilden, Germany), and quantified by light absorbance (Genequant, Pharmacia). An aliquot containing 10–30 ng DNA was used to amplify the regions of interest by PCR, with the primers reported by Hu et al. [9]:

- DXS7423
  - primer 1: TAGCTTAGCGCCTGGCACATA
  - primer 2: GTCTTCCTGTCATCTCCCAAC
- DXS8377
  - primer 1: CACTTCATGGCTTACCACAG
  - primer 2: GACCTTTGGAAAGCTAGTGT)

Primer 1 for each locus was HEX-labelled.

After optimisation of various parameters, the final reaction mix contained 4 pmol each DXS7423 primer and 10 pmol each DXS8873 primer, 1.5 mM $MgCl_2$, 5 nmol each dNTP, and 2 U Taq polymerase (AmpliTaq gold, Applied Biosystems; Foster City, Calif.)

in a 25 µl volume. Both markers were amplified in the same reaction tube with a touch-down protocol consisting of an initial denaturation at 94°C for 10 min, followed by 8 cycles with denaturation at 94°C for 1 min, annealing at decreasing temperature between 62°C and 59°C (1°C decrease every 2 cycles), and extension at 72°C for 1 min. Then 24 cycles at 94°C for 1 min, 58°C for 1 min, 72°C for 1 min; followed by a final extension at 72°C for 30 min.

Fragment analysis

Aliquots containing 1.2 µl of PCR product were mixed with 24 µl formamide and 1 µl TAMRA as a internal size standard, heated for 5 min at 95°C, quenched in an ice bath for 10 min and injected into an ABI prism 310 capillary electrophoresis system (Applied Biosystems). Fragment size was determined with the local Southern method implemented in Genescan software.

Sequencing analysis and allelic ladders

DNA samples from male subjects were taken for sequencing several common alleles. PCR products were purified (Concert PCR, Gibco-BRL Life Technologies, Inchinnan, UK) and eluted in Tris-EDTA buffer and then amplified in a dye-terminator based sequencing PCR reaction using unlabelled primer 2 and a commercial kit (dRhodamine Terminator, Applied Biosystems). Products of sequencing reactions were precipitated with ethanol/$MgCl_2$, washed with chilled ethanol, re-suspended in the template suppression reagent, denatured and injected into an ABI prism 310. Allelic ladders were constructed with the sequenced alleles to help in typing unknown samples (the size of alleles not sequenced was determined by interpolation).

Population samples and data analysis

Samples from 125 caucasoid individuals (60 men and 65 women) living in Cantabria, a region in northern Spain, were genotyped and allele frequencies were determined by counting. The results in samples of male and female origin were compared using an extension of the Fisher exact test implemented in SPSS software. Hardy-Weinberg equilibrium in female samples and the possibility of linkage disequilibrium in both male and female samples were examined with GDA software (Lewis PO and Zaykin D. Genetic Data Analysis, v.1.1. http://lewis.eeb.uconn.edu/lewishome/software.html), which implements the formulae published by Weir [10]. The average power of discrimination (PD) and the expected probability of exclusion (PE) were estimated with the following formulae proposed by Desmarais et al. [4]:

$$\text{PD (female sample)} = 1 - 2\left(\sum p_i^2\right)^2 + \sum p_i^4$$

$$\text{PD (male sample)} = 1 - \sum p_i^2$$

$$\text{PE (trio)} = 1 - \sum p_i^2 + \sum p_i^4 - \left(\sum p_i^2\right)^2$$

$$\text{PD (motherless)} = 1 - 2\left(\sum p_i^2\right) + \sum p_i^3$$

Results

Sequencing data and nomenclature

PCR with DXS7423 primers resulted in fragments of 181–197 bp, including 5 alleles with different numbers of the repeat sequence (TCCA; Table 1). Besides the polymorphic region, an invariant region with three (TCCA) repeats was identified. Allele designation included the total number of repeats, variant and invariant, as recommended by the ISFH [11].

Table 1 DXS7423 allele repeats and designation

Allele	Structure	Apparent size (bp)
13	Not sequenced	181
14	...CCA $(TCCA)_3$ TCT GTC CT $(TCCA)_{11}$ CAT...	185
15	...CCA $(TCCA)_3$ TCT GTC CT $(TCCA)_{12}$ CAT...	189
16	...CCA $(TCCA)_3$ TCT GTC CT $(TCCA)_{13}$ CAT...	193
17	...CCA $(TCCA)_3$ TCT GTC CT $(TCCA)_{14}$ CAT...	197

Table 2 Structure of allele 37 at the DXS8377 locus (the polymorphic region is underlined)

Primer 2-TTTTGCTCTTCGTTCCTGTC$(TTC)_6$
CTC $(TTC)_3$ CTC TTC CTC TTC CTC TTC CTC $(TTC)_{21}$
TCC TTC TCC TCC TTC
TCC$(TTC)_2$ TCC TC-Complementary to Primer1

Table 3 DXS8377 alleles. Structure of the polymorphic region and designation of sequenced alleles

Allele	Structure	Apparent size (bp)
37	... $(YTC)_{11}$ $(TTC)_{21}$ $(TYC)_5$ ...	204
42	... $(YTC)_{13}$ $(TTC)_{24}$ $(TYM)_5$ ...	219
44	... $(YTC)_{13}$ $(TTC)_{26}$ ($TYC)_5$ ...	225
46	... $(YTC)_{15}$ $(TTC)_{26}$ ($TYM)_5$ ...	231
47	... $(YTC)_{13}$ $(TTC)_{29}$ ($TYC)_5$ ...	234
50	... $(YTC)_{17}$ $(TTC)_{28}$ $(TYC)_5$ ...	243
51	... $(YTC)_{15}$ $(TTC)_{33}$ $(TYC)_3$ ...	246

According to the IUB codes, Y designates either T or C and M designates either A or C

Unlike DXS7423, the sequencing of DXS8377 revealed a complex repetitive structure (Table 2) and two short fragments adjacent to the primers were the only constant regions. The rest of the DNA sequence was comprised of a long region with a variable number of (TTC) repeats, and other smaller regions with variable numbers of interspersed (TTC), (CTC) and (TCC) triplets. In some cases, (TCA) was substituted for (TCC). In order to maintain the coherence between allele designation and size [11], the nomenclature proposed includes the total number of triplets in the variable region, independent of their base composition (Table 3). The size of amplified fragments varied between 204 and 255 bp, with 18 different alleles actually observed in the subjects studied. It seems possible to design other primers resulting in smaller amplicons. Hence, in order to keep the same allele nomenclature in possible future studies, the number of invariant repeats situated near the current primers was not included in allele designation.

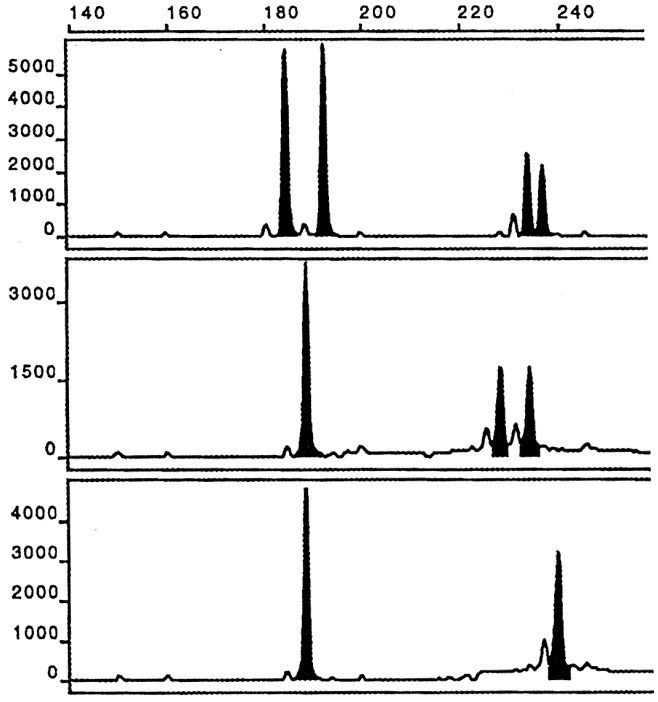

Fig. 1 Electrophoresis profiles. *Top* bi-heterozygotic female (alleles 14–16 and 47–48), *middle* a female homozygote for DXS7423 (alleles 15–15) and heterozygote for DXS8377 (alleles 45–47) and *bottom* male (alleles 15 and 49)

Duplex performance

The markers were amplified conveniently under the above reported conditions and repeated analysis of random DNA samples gave consistent results. Although we routinely used about 20 ng DNA, 1 ng was usually enough to type both markers. With smaller samples, DXS7423, but not DXS8377 alleles, could still be amplified. DXS7423 peaks usually had heights and areas higher than those of DXS8377 peaks (Fig. 1).

DXS7423 amplification sometimes resulted in small stutter bands with size n–4 bp (i.e., repeat units minus 1) which were usually less than 5% of the true band height. On the other hand, n–3 bp stutter bands up to 20% of the height of the true band were frequent in the case of DXS8377 amplification.

Population data

Since allele frequencies for both markers were similar in males and females (p=0.9 by the Fisher test), data from both sexes were combined. Allele frequencies are shown in Tables 4 and 5. Alleles in samples of female origin were in Hardy-Weinberg equilibrium (p=0.3 and 0.9, for DXS7423 and DXS8377, respectively). There was no evidence of significant linkage disequilibrium between genotypes at both loci either in female (p=0.4) or male samples (p=0.9).

Table 4 DXS7423 allele frequencies

Allele	Frequency	Standard error
13	0.073	0.019
14	0.272	0.032
15	0.403	0.035
16	0.209	0.029
17	0.041	0.014

Table 5 DXS8377 allele frequencies

Allele	Frequency	Standard error
37	0.011	0.008
38	0.027	0.012
39	0.022	0.011
40	0.038	0.014
41	0.065	0.018
42	0.054	0.017
43	0.059	0.017
44	0.065	0.018
45	0.086	0.021
46	0.091	0.021
47	0.081	0.020
48	0.086	0.021
49	0.118	0.024
50	0.065	0.018
51	0.054	0.017
52	0.038	0.014
53	0.022	0.011
54	0.022	0.011

Table 6 Forensic efficiency of DXS7423 and DXS8377 markers in identification and paternity analysis (*PD* power of discrimination in identification of samples of male or female origin, *PE* expected power of exclusion in paternity testing of a female child for standard trio cases and cases with unavailable mother)

Forensic efficiency values	DXS7423	DXS8377
PD female sample	0.868	0.991
PD male sample	0.712	0.929
PE trio cases	0.663	0.925
PE motherless cases	0.520	0.865

As shown in Table 6 the markers, particularly DXS8377, were highly informative with a probability of discrimination of female samples above 0.99.

Among 107 meioses we found no evidence of mutations at the DXS7423 locus, and found 1 mutation in DXS8377. The father had allele 46, the mother had alleles 41 and 47 and the daughter alleles 41 and 45, suggesting a loss of 1 repeat unit from the paternal allele. The father was 29 years old at conception, the mother 30.

Discussion

DXS7423 and DXS8377 loci are situated near the terminal end of the X chromosome in the Xq28 band. They

were described several years ago and used in candidate gene localisation studies [9], but their utility in legal medicine has remained almost unexplored. The results of this study show that they are polymorphic loci with great potential utility in forensic cases. DXS8377 in particular is very informative, with a very high probability of discrimination and expected exclusion power in paternity analysis. The alleles can be conveniently determined by conventional PCR and fragment analysis by capillary electrophoresis. Moreover, since the fragment sizes do not overlap, their primers can be labelled with a single fluorochrome, thus allowing the possibility of using other labels in multiplex PCR reactions.

The allele frequency distribution of DXS8377 in our population is similar to that reported recently by Athanasiadou et al. [12], who did not find any mutations in 93 meioses. Among 107 meioses, we had no evidence of mutations in DXS7423 (95% confidence interval: 0–3.4%), but we found 1 mutation in DXS8377 (0.9%; 95% confidence interval: 0.02–5.1%). Microsatellites with an interrupted repeat structure, as those studied here, are usually quite stable [13] and mutation rates of about 0.4% have been reported in other X-linked STRs [14]. The analysis of a larger number of meioses would be needed to determine more precisely the mutation rate of these systems, which is certainly important in forensic analysis. Although both loci are separated by only 150 kb [9], we did not find evidence of linkage disequilibrium between them. Disequilibrium is not a monotonic function of the distance between markers, as has been recently demonstrated by analysing a set of single nucleotide polymorphisms on the X chromosome [15]. Nevertheless, care should be taken in the concomitant use of both markers in forensic cases originating from small isolated populations.

In summary, DXS7423 and DXS8377 are polymorphic STRs that can be very useful in forensic cases. As other markers on the X chromosome, they could be particularly interesting in deficient cases of paternity, when the alleged child is a female and a sample of the alleged father is not available. They could also be useful in some cases of identification through kinship testing. Nevertheless, the tendency of DXS8377 to produce stutter bands may sometimes make the interpretation of the results difficult if the possibility of mixed samples exists.

Acknowledgements The authors thank Dr. Domingo Gonzalez-Lamuño for his support and Dr. P. Schneider for sharing his results. This work was supported in part by a grant from the Fundación Marqués de Valdecilla.

References

1. Alford RL, Hammond HA, Coto I, Caskey CT (1994) Rapid and efficient resolution of parentage by amplification of short tandem repeats. Am J Hum Genet 55:190–195
2. Edelman J, Szibor R (2001) DXS101: a highly polymorphic X-linked STR. Int J Legal Med 114:301–304
3. 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
4. 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
5. Dupuy BM, Gedde-Dahl T, Olaisen B (2000) DXY267:DYS393 and its X chromosome counterpart. Forensic Sci Int 112:111–121
6. Hering S, Szibor R (2000) Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. J Forensic Sci 45:929–931
7. Mertens G, Gielis M, Mommers N, Mularoni A, Lamartine J, Heylen H, Muylle L, Vandenberghe A (1999) Mutation of the repeat number of HPRTB locus and structure of rare intermediate alleles. Int J Legal Med 112:192–194
8. Watanabe G, Umetsu K, Yuasa I, Suzuki T (2000) DXS10011: a hypervariable tetranucleotide STR polymorphism on the X chromosome. Int J Legal Med 113:249–250
9. Hu LJ, Laporte J, Kioschis P, Heyberger S, Kretz C, Poustka A, Mandel JL, Dahl N (1996) X-linked myotubular myopathy: refinement of the gene to a 280-kb region with new and highly informative microsatellite markers. Hum Genet 98:178–181
10. Weir BS (1996) Genetic data analysis II. Sinauer, Sunderland
11. DNA Commission of the International Society of Forensic Haemogenetics (1997) DNA recommendations – Further report of the DNA commission of the ISFH regarding the use of short tandem repeat systems. Forensic Sci Int 87:179–184
12. Athanasiadou D, Stradmann-Bellinghausen B, Rittner C, Alt KW, Schneider PM (2002) Development of a quadruplex PCR system for the genetic analysis of x-chromosomal STR loci. In: Brinkmann B et al. (ed) Progress in forensic genetics 9. Elsevier Science, Amsterdam (in press)
13. Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62:1408–1415
14. Szibor R, Lautsch S, Plate I, Beck N (2000) Population data on the X chromosome short tandem repeat locus HumHPRTB in two regions of Germany. J Forensic Sci 45:231–233
15. Taillon-Miller P, Bauer-Sardiña I, Saccone NL, et al, (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. Nat Genet 25:324–328