Int J Legal Med (2003) 117 : 241–244
DOI 10.1007/s00414-003-0369-4

SHORT COMMUNICATION

J. Edelmann · D. Deichsel · I. Plate · M. Käser · R. Szibor

Validation of the X-chromosomal STR DXS6809

Received: 2 October 2002 / Accepted: 24 February 2003 / Published online: 14 May 2003
© Springer-Verlag 2003

Abstract This paper presents sequence and population genetic data of the microsatellite marker DXS6809 (GDB 365492) obtained from a German population sample (n=725 chromosomes). DXS6809 is a highly polymorphic X-linked tetranucleotide polymorphism presenting 12 alleles in our population. Sequencing of 77 PCR products covering 12 alleles (by length), characterised DXS6809 as a marker with a complex repeat sequence structure. A polymorphism information content (PIC) of 0.825 and a mean exclusion chance (MEC) of 0.815 were obtained. A deviation from the Hardy-Weinberg equilibrium (HWE) could not be detected and male and female samples exhibited a similar allele distribution. Kinship testing revealed a typical X-linked inheritance and 2 mutations were found in 394 meioses. DXS6809 is located 90.18 Mb, i.e. 102.3 cM, from the Xp-telomere (Xp-tel), corresponding to Xq21.33. The presented data qualify DXS6809 as a useful supplement to the known forensic ChrX marker panel.

Keywords Short tandem repeats · X-chromosome · Population genetics · Repeat sequence · Mutation rate

Introduction

A large number of autosomal and Y-chromosomal DNA polymorphisms have been forensically evaluated and widely applied to stain analyses and kinship testing (Brinkmann 1998; Kayser et al. 1997). In addition, forensic interest has increasingly focused on X-chromosomal markers (Kishida and Tamaki 1997; Edelmann and Szibor 1999, 2001; Szibor et al. 2000; Hering et al. 2001; Edelmann et al. 2001, 2002; Zarrabeitia et al. 2002), which can effectively supplement autosomal and Y-chromosomal testing.

This paper is aimed at continuing the process of establishing a suitable panel of chromosome X (ChrX) STRs. Covering the whole ChrX with a tightly network of well established STRs bears the potential for X-chromosomal haplotyping which is a powerful tool in kinship testing.

J. Edelmann (✉) · D. Deichsel
Institut für Rechtsmedizin, Universität Leipzig,
Johannisallee 28, 04103 Leipzig, Germany
Fax: +49-341-9715109,
e-mail: edej@medizin.uni–leipzig.de

M. Käser
Laboratorium für Hämogenetik, Baden-Baden, Germany

I. Plate · R. Szibor
Institut für Rechtsmedizin,
Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany

Materials and methods

DNA was extracted from the blood of 725 unrelated Germans (306 females and 419 males) using the QIAamp DNA blood kit (Qiagen, Hilden, Germany). All sample donors were individuals born in Germany, which were involved in our local casework. A total of 197 family trios including female children were checked for regular X-chromosomal inheritance (394 meioses). The parental age structure was as follows:

- 15–20 years: 57 mothers, 20 fathers
- 21–25 years: 67 mothers, 73 fathers
- 26–30 years: 45 mothers, 57 fathers
- 31–35 years: 27 mothers, 29 fathers
- 36–40 years: 0 mothers, 9 fathers
- >40 years: 1 mother, 9 fathers

PCR amplification was performed using the following primer sequences based on gene bank information (http://www.gdb.org):

- Primer$^{F\text{-}I}$ ($P^{F\text{-}I}$): 5’-(6’-FAM)-TGA ACC TTC CTA GCT CAG GA-3’
- Primer$^{F\text{-}II}$ ($P^{F\text{-}II}$): 5’-(6’-FAM)-CTA GAT TAT GTA GGA ATT TGG-3’
- Primer^{R} (P^{R}): 5’-GCA AAA TTG GAT TCT CCA GA-3’

$P^{F\text{-}I}$ and P^{R} are established primers given in the gene bank document and $P^{F\text{-}II}$ is a newly created primer which yields PCR fragments that are 35 bp shorter compared to $P^{F\text{-}I}$.

Amplification was carried out in a 25 µl PCR reaction volume containing 10 ng DNA, 200 µM each dNTP, 1.5 mM $MgCl_2$, 0.5 µM each primer, 1 U Taq polymerase (Perkin-Elmer, Foster City, CA) and 1×PCR buffer, for 30 cycles using a thermocycler (Biometra, Göttingen, Germany) and the following conditions: 95°C – 10 min soak, 94°C – 60 s, 58°C – 60 s, 72°C – 90 s and 72°C – 10 min final extension. The resulting PCR products were analysed by capillary electrophoresis using an ABI Prism 310 sequencer with polymer POP 4 (Perkin-Elmer). Allele typing was based on sequenced alle-

Table 1 DXS6809 allele frequencies and standard deviations (SD) from 306 females and 419 males

Allele	Size[a] (bp)[c]	Size[b] (bp)[c]	Females		Males		Pooled alleles	
			n	Frequency±SD	n	Frequency±SD	n	Frequency±SD
27	235	200	10	0.016±0.005	10	0.024±0.007	20	0.019±0.004
28	239	204	23	0.038±0.007	18	0.043±0.009	41	0.040±0.006
29	243	208	7	0.011±0.004	11	0.026±0.007	18	0.017±0.004
30	247	212	31	0.051±0.008	10	0.024±0.007	41	0.040±0.006
31	251	216	101	0.165±0.015	49	0.117±0.015	150	0.145±0.011
32	255	220	93	0.152±0.014	58	0.138±0.016	151	0.146±0.011
33	259	224	164	0.268±0.017	122	0.291±0.022	286	0.277±0.013
34	263	228	108	0.176±0.015	86	0.205±0.019	194	0.188±0.012
35	267	232	47	0.077±0.010	32	0.076±0.013	79	0.077±0.008
36	271	236	23	0.038±0.007	16	0.038±0.009	39	0.038±0.005
37	275	240	4	0.007±0.003	4	0.010±0.004	8	0.008±0.002
38	279	244	1	0.002±0.001	3	0.007±0.004	4	0.004±0.001

PIC: 0.825, MEC: 0.815, HET: 0.808, PD^F: 0.953, PD^M: 0.835, PE: 0.815

[a]size using $P^{F\text{-}I}$ primer.
[b]size using $P^{F\text{-}II}$ primer.
[c] true sequenced allele sizes.
HET heterozygosity.
PE expected probability of exclusion.
PIC polymorphism information content.
MEC mean exclusion chance.
PD^F average power of discrimination in females.
PD^M PD in males.

Table 2 Allele nomenclature, sequence composition and allele size of 77 sequenced PCR products

Allele	Size (bp)[a]	Sequence composition	n
27	235	P_{F20}-N_{37}-(CTAT)$_7$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_3$-N_{10}-(ATCT)$_{11}$-N_{31}-P_{R20}	5
28	239	P_{F20}-N_{37}-(CTAT)$_7$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_3$-N_{10}-(ATCT)$_{12}$-N_{31}-P_{R20}	4
29	243	P_{F20}-N_{37}-(CTAT)$_7$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_3$-N_{10}-(ATCT)$_{13}$-N_{31}-P_{R20}	4
30	247	P_{F20}-N_{37}-(CTAT)$_8$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{11}$-N_{31}-P_{R20}	2
		P_{F20}-N_{37}-(CTAT)$_8$-(ATCT)$_3$-N_9-(TATC)$_4$-(ATCT)$_5$-N_{10}-(ATCT)$_{10}$-N_{31}-P_{R20}	2
		P_{F20}-N_{37}-(CTAT)$_7$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_3$-N_{10}-(ATCT)$_{14}$-N_{31}-P_{R20}	2
31	251	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{11}$-N_{31}-P_{R20}	8
		P_{F20}-N_{37}-(CTAT)$_8$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{12}$-N_{31}-P_{R20}	1
32	255	P_{F20}-N_{37}-(CTAT)$_8$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{13}$-N_{31}-P_{R20}	1
		P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{12}$-N_{31}-P_{R20}	5
		P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{11}$-N_{31}-P_{R20}	2
33	259	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{13}$-N_{31}-P_{R20}	7
		P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{12}$-N_{31}-P_{R20}	4
		P_{F20}-N_{37}-(CTAT)$_{11}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{11}$-N_{31}-P_{R20}	3
34	263	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{14}$-N_{31}-P_{R20}	4
		P_{F20}-N_{37}-(CTAT)$_{10}$.(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{13}$-N_{31}-P_{R20}	4
		P_{F20}-N_{37}-(CTAT)$_{11}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{12}$-N_{31}-P_{R20}	1
35	267	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{15}$-N_{31}-P_{R20}	4
		P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{14}$-N_{31}-P_{R20}	3
36	271	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{16}$-N_{31}-P_{R20}	3
		P_{F20}-N_{37}-(CTAT)$_{11}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{14}$-N_{31}-P_{R20}	1
		P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{15}$-N_{31}-P_{R20}	1
37	275	P_{F20}-N_{37}-(CTAT)$_9$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{17}$-N_{31}-P_{R20}	3
		P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{16}$-N_{31}-P_{R20}	2
38	279	P_{F20}-N_{37}-(CTAT)$_{10}$-(ATCT)$_3$-N_9-(TATC)$_3$-(ATCT)$_5$-N_{10}-(ATCT)$_{17}$-N_{31}-P_{R20}	1

[a]true sequenced allele sizes.
$Primer^{F\text{-}I}$: TGAACCTTCC-TAGCTCAGGA.
N_{37}: ATACTGAGGGCAT-GACTAGATTATGTAG-GAATTTGGG.
N_9: ATCATCTAT.
N_{10}: ATTATCTATC.
N_{31}:ATCCTCTATCTCTTCCT CACATCAGCCTAAA.
$Primer^R$: GCAAAATTG-GATTCTCCAGA.
Differences from GDB sequence are underlined.

les. For control DNA serving to calibrate allelic ladders, we used the cell line samples K562 and 9947A (Promega, Madison, WI).

Sequencing analysis

Male DNA samples were amplified using the primers $P^{F\text{-}I}$ and P^R for generating the sequencing templates. Direct Taq-cycle sequencing was carried out using the BigDyeTerminator kit (Perkin-Elmer) and the unlabelled PCR primer $P^{F\text{-}I}$ at a concentration of 3 pmol. Sequence data were obtained from 77 PCR products chosen for their fragment length.

Statistic analysis

Hardy-Weinberg equilibrium (HWE) was tested according to Guo and Thompson 1992. The parameters of forensic interest such as

Table 3 Sequence structure of the mutant alleles

Person (age)	Alleles	Sequence
Case 1		
Mother	32/33	
Child	32/**33**	$P_{F20}\text{-}N_{37}\text{-}\mathbf{(CTAT)_{10}}\text{-}(ATCT)_{3}\text{-}N_{9}\text{-}(TATC)_{3}\text{-}(ATCT)_{5}\text{-}N_{10}\text{-}(ATCT)_{12}\text{-}N_{31}\text{-}P_{R20}$
Father (49)	**34**	$P_{F20}\text{-}N_{37}\text{-}\mathbf{(CTAT)_{11}}\text{-}(ATCT)_{3}\text{-}N_{9}\text{-}(TATC)_{3}\text{-}(ATCT)_{5}\text{-}N_{10}\text{-}(ATCT)_{12}\text{-}N_{31}\text{-}P_{R20}$
Case 2		
Mother (23)	31/**34**	$P_{F20}\text{-}N_{37}\text{-}\mathbf{(CTAT)_{10}}\text{-}(ATCT)_{3}\text{-}N_{9}\text{-}(TATC)_{3}\text{-}(ATCT)_{5}\text{-}N_{10}\text{-}(ATCT)_{13}\text{-}N_{31}\text{-}P_{R20}$
Child	32/**33**	$P_{F20}\text{-}N_{37}\text{-}\mathbf{(CTAT)_{9}}\text{-}(ATCT)_{3}\text{-}N_{9}\text{-}(TATC)_{3}\text{-}(ATCT)_{5}\text{-}N_{10}\text{-}(ATCT)_{13}\text{-}N_{31}\text{-}P_{R20}$
Father	32	

polymorphism information content (PIC) (Botstein et al. 1980), mean exclusion chance in normal family trio tests (MEC) (Kishida et al. 1997), power of discrimination in females (PD^F), in males (PD^M) and paternity exclusion chance (PE) (Desmarais et al. 1998) were calculated. The 95% binomial confidence interval limit (CIL) of mutation rate was determined using http://www.swogstat.org/stat/public/binomial_conf.htm.

Mapping

The mapping information on DXS6809 and further linked markers was taken from international gene data bases (http://www.gdb.org, http://www.chlc.org, http://carbon.wi.mit.edu, and http://www.ncbi.nlm.nih.gov).

Results and discussion

Table 1 shows DXS6809 allele frequencies from a German population sample, calculated separately for males and females. Using the primers $P^{F\text{-}I}$ and P^R, a total of 12 alleles with length variation were identified, which increased in size by 4 bp increments ranging from 235 to 279 bp. Substitution of the forward primer $P^{F\text{-}I}$ by $P^{F\text{-}II}$ reduced the length of one of the non-polymorphic regions by 35 bp, resulting in PCR products of 200–244 bp in length. $P^{F\text{-}I}$ and $P^{F\text{-}II}$ can be used alternatively. The possibility to choose between two alternative primers is a good starting condition for incorporating DXS6809 into multiplex PCR mixtures. Our group has established a DXS6809-DXS6789-DXS6801 primer combination to create a tool for haplotyping the Xq21 region (data not shown).

Significant deviations from the HWE were not found and the exact test based on 46 different genotypes (306 females, data not shown) yielded a p-value of 0.32. Data describing the STR information content were calculated (Table 1) and parameters such as PIC (0.825) and MEC (0.815) indicate that DXS6809 is a highly informative marker. The following results refer to the use of the $P^{F\text{-}I}$ and P^R primer pair: sequencing of 77 PCR products covering 12 alleles (by length) revealed that DXS6809 is a marker with a complex repeat sequence structure (Table 2): $P_{F20}\text{-}N_{37}\text{-}(CTAT)_n\text{-}(ATCT)_3\text{-}N_9\text{-}(TATC)_n\text{-}(ATCT)_n\text{-}N_{10}\text{-}(ATCT)_n\text{-}N_{31}\text{-}P_{R20}$. Alleles were assigned in compliance with the recommendations of the ISFH Commission (Bär et al. 1997). The proposed designation of DXS6809 alleles ranges from 27 for a 235 bp allele to 38 for a 279 bp allele (i. e. 200 bp–244 bp in the shorter version). Alleles of identical size exhibited considerable structural variations. This sequence heterogeneity is not a disadvantage as it carries additional information that can be utilised by sequencing in special situations. The control cell line DNA K562 and 9947A displayed the alleles 34 and 31/34, respectively.

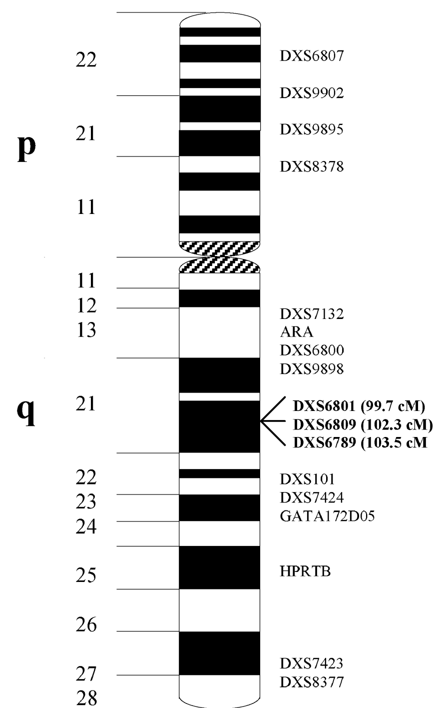

Fig. 1 ChrX marker localisation

Investigating 197 family trios (394 meiosis) with female children, we found 2 mutations and the alleles of the family members were sequenced. As a result we found single-step repeat losses of the first (CTAT) repeat in both cases (Table 3). The calculated mutation rate, including the 95% confidence interval limit (CIL), was 5.07×10^{-3} (0.0006–0.0182). This result is compatible with the more complex and extended repeat structure of DXS6809, compared to GATA172D05 (Edelmann et al. 2002) and some other STRs (DXS6800, DXS101, DXS9895, DXS9898) exhibiting simple repeats and no mutations in 440–754 meioses (Edelmann and Szibor 2001; Szibor et al. 2003). According to Brinkmann et al. (1998) the increase in length and complexity of the repeat region in microsatellites is correlated with a loss of stability.

DXS6809 is located 90.18 Mb, i.e. 102.3 cM, from the Xp-telomere (Xp-tel), corresponding to Xq21.33. DXS6789 (located at a distance of 90.7 Mb from the ChrX top) and DXS6801 (located at 87.8 Mb) are tightly linked with DXS6809 at a genetic distance of about 1 cM and 2.6 cM, respectively (Fig. 1). Whereas DXS6789 is already well characterised (Hering et al. 2001), DXS6801 is still under investigation. The three markers are situated close together within the X-chromosomal linkage group 2 as described by Szibor et al. 2003.

In summary we succeeded in adding a highly polymorphic STR to the known forensic ChrX marker panel and contributing to the development of a ChrX haplotyping procedure as an efficient tool in human kinship testing.

References

Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B (1997) DNA Commission of the ISFH DNA recommendations – Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Int J Legal Med 110:175–176

Botstein D, White RI, Skolnich M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Brinkmann B (1998) Overview of PCR-based systems in identity testing. Methods Mol Biol l98:105–119

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

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

Edelmann J, Szibor R (1999) Validation of the HumDXS6807 short tandem repeat polymorphism for forensic application. Electrophoresis 20:2844–2846

Edelmann J, Szibor R (2001) DXS101: a highly polymorphic X–linked STR. Int J Legal Med 114:301–304

Edelmann J, Hering S, Michael M, Lessig R, Deichsel D, Meier-Sundhausen G, Roewer L, Plate I, Szibor R (2001) 16 X-chromosome STR loci frequency data from a German population. Forensic Sci Int 124:215–218

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

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361–372

Hering S, Kuhlisch E, Szibor R (2001) Development of the X-linked tetrameric microsatellite marker DXS6789 for forensic purposes. Forensic Sci Int 119:42–46

Kayser M, Cagliá A, Corach D et al. (1997) Evaluation of Y–chromosomal STR's: a multicenter study. Int J Legal Med 110: 125–133

Kishida T, Tamaki Y (1997) Japanese population data on X-chromosomal STR locus AR. Nippon Hoigaku Zasshi 51:376–379

Kishida T, Wang W, Fukuda M, Tamaki Y (1997) Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. Nippon Hoigaku Zasshi 51: 67–69

Szibor R, Lautsch S, Plate I, Beck N (2000) Population data of X chromosomal STR HumHPRTB in two regions of Germany. J Forensic Sci 45:231–233

Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. Int J Legal Med (in press)

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