

Sandra Hering · Nicola Brundirs · Eberhard Kuhlisch ·
Jeanett Edelmann · Ines Plate · Mark Benecke ·
Pham Hung Van · Matthias Michael · Reinhard Szibor

DXS10011: studies on structure, allele distribution in three populations and genetic linkage to further q-telomeric chromosome X markers

Received: 2 March 2004 / Accepted: 16 April 2004 / Published online: 10 July 2004
© Springer-Verlag 2004

Abstract The hypervariable tetranucleotide STR polymorphism DXS10011 is a powerful marker for forensic purposes. Investigation of this STR led to an allele nomenclature which is in consensus with the ISFG recommendations. DXS10011 is located at Xq28 and genetically closely linked to DXS7423 and DXS8377 but is unlinked to HPRTB and more distant X-chromosomal STRs. DXS10011 is a very complex marker exhibiting some structural variants within alleles of identical length. Two types of repeat structure (regular and inter-alleles) are

known and described as types A and B. Two SNPs which are in strong linkage disequilibrium to the different sequence types were found in the repeat flanking region. The type A sequence consists of a long stretch of uninterrupted homogenous repeats which is highly susceptible to slippage mutation during male meiosis.

Keywords X-Chromosome · DXS10011 · Population study · Sequencing

S. Hering
Institut für Rechtsmedizin, Technische Universität Dresden,
Fetscherstrasse 74,
01307 Dresden, Germany

N. Brundirs · I. Plate · R. Szibor (✉)
Institut für Rechtsmedizin, Otto-von-Guericke-Universität
Magdeburg,
Leipziger Strasse 44,
39120 Magdeburg, Germany
e-mail: reinhard.szibor@medizin.uni-magdeburg.de
Tel.: +49-391-6715812
Fax: +49-391-6715810

E. Kuhlisch
Institut für Medizinische Informatik und Biometrie, Technische
Universität Dresden,
Fetscherstraße 74,
01307 Dresden, Germany

J. Edelmann
Institut für Rechtsmedizin, Universität Leipzig,
Johannisallee 28,
04103 Leipzig, Germany

M. Benecke
International Forensic Research & Consulting,
Postfach 250411, 50520 Köln, Germany

P. H. Van
Institute of Medicine, University of Hanoi,
Vietnam

M. Michael
Institut für Rechtsmedizin, Friedrich-Schiller-Universität Jena,
Fürstengraben 23,
07743 Jena, Germany

Introduction

The hypervariable tetranucleotide STR polymorphism DXS10011, first described by Gerken et al. [1] as DYS384, is also known as HUMUT413. Allele frequencies were published for Japanese and German population samples by Matsuki et al. [2, 3], Watanabe et al. [4] and Koyama et al. [5]. Two types of the repeat structure (regular and inter-alleles) were described and designed as type A and type B. DXS10011 is located at Xq28 and is physically closely linked to DXS7423 and DXS8377.

The aim of this study was to implement this highly polymorphic marker into our routine work using an STR panel of the X-chromosome (ChrX).

Since in earlier papers on DXS10011 no reliable DNA typing results for cell line DNA were communicated, we were not able to adjust our allelic ladder to the data of the previous papers. Hence, we could hardly compare our results to the allele distributions published [2, 3, 4, 5].

Thus, we started an extensive reinvestigation of DXS10011 leading to an allele nomenclature which is in consensus with the ISFG recommendations [6].

In addition, genetic linkage to forensically relevant STR markers localised on the Xq had to be tested by LOD-score analysis. Investigation of Germans and some other small population samples, e.g. Amerindians from Peru and Vietnamese populations, showed that interpopulation diversity seems to be low.

Materials and methods

Blood samples and buccal swabs for population studies were collected from 1,328 unrelated Germans (745 male and 583 female), 116 Vietnamese (19 male and 97 female) and 81 Peruvians (40 male and 41 female). Cases of routine kinship testing were also involved. In this context, 409 trios with female children were investigated (paternity index 1,000 or more). DNA from 187 mothers with 2 or more sons were DNA typed to estimate the genetic distance from DXS10011 to DXS7423, DXS8377, HPRTB, GATA172D05, DXS7133, DXS101, DXS7424, DXS6789, DXS6809, DXS9898, DXS6800, HumARA, DXS7132, DXS9902, DXS8378, DXS9895 and DXS6807 [7, 8, 9, 10]. To carry out this genetic study samples were predominantly collected from students and their families. Therefore we had to anonymise the samples before investigation and do not have information on the mothers age at the time of conception.

DNA extraction was carried out using the QIAamp DNA blood kit (Qiagen, Hilden, Germany).

Primers for amplifying long PCR products which were suitable as sequencing templates, were synthesised according to Gerken et al. [1] and the amplicon length for allele 31.2 was 363 bp.

The following primers were used in a direct Taq-cycle sequencing procedure using the BigDye-Terminator kit v1.1 and the ABI Prism 310 sequencer, Sequencing Analysis v3.7 (Perkin Elmer, Foster City, CA) [11]:

- Primer 1F: 5'-TTAGCTGAACGTGGTGGCT-3'
- Primer 1R: 5'-ATGCTGGACAATTGCCAGG-3'

To determine the amplicon length, the primer pairs 2, 3 and 4 were used.

The primer 2F sequence was adopted unchanged from Matsuki et al. [3], the reverse primer was modified

according to our sequencing data (allele 31.2, 227 bp) as follows:

- Primer 2F: 5'-FAM-CCAGCCAGGGCAACAA-GAGTGAA-3'
- Primer 2R: 5'-CAGCGTGGGAGAACCGTTT-GAAG-3'.

The sequence of our primer pair 3 is close to the suggestion of Watanabe et al. [4] but was slightly changed according to our sequencing data (allele 31.2, 167 bp).

- Primer 3F: 5'-FAM-GAGTGAAACTCTGAAAA-3'
- Primer 3R: 5'-TGAAATCATCTATCTTTCTTTC-3'.

New primers with the potential to amplify low copy numbers of templates were designed using the Primer3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi), here the amplicon length for allele 31.2 was 227 bp.

- Primer 4F: 5'-FAM-CTGAGATTGCACCATTGCAC-3'
- Primer 4R: 5'-TGGGAGAACCGTTTGAAGTT-3'.

PCR set-up and fragment analysis were carried out following standard protocols [11]. The amplification was carried out in a 25 µl PCR reaction volume containing 0.1–1 ng DNA, 200 µM each dNTP, 1.5 mM MgCl₂, 0.5 µM of each primer, 1 U Taq polymerase (Perkin-Elmer, Foster City, CA) and 1×PCR buffer. For primer pair 3 the optimal MgCl₂ concentration was 2.5 mM.

The cyclor program on a Biometra thermocycler (Goettingen, Germany) was set to the following conditions: 95°C for 3 min soak, 94°C for 60 s, 56°C (primer pairs 2 and 3) or 60°C (primer pairs 1 and 4) for 60 s, 72°C for 90 s and 72°C for 10 min final extension (30 cycles). The resulting PCR products were analysed by capillary electrophoresis using an ABI Prism 310 sequencer with POP4 polymer.

```

1   ttagctgaac gtggtggctg tcgctgtaa tcccagctac ttgggagact
51  aagtcaggag aatcgcttga acctgggagg cagaggtttc agtgagctga
101 gattgcacca ttgcactcca gccagggcaa caagagtgaa actctgaaaa
151 aaaaCagaaG aaaGaGaaaG aaaGaaaGaa aGaaaGaaaG agaGaaaGaa
201 aGaaaGaaaG aaaGaaaGaa aGaaaGaaaG aaaGaaaGaa aGaaaGaaaG
251 aagGaaaGga aGgaaGgaaG gaaAgaaAga aAgaaagata gatgatttca
301 ttcaaaatgc atgattatta acttcaaaTg gttctcccacc gcLgcctggc
351 aattctccag cata

```

Fig. 1 Type B sequence of the DXS10011 inter-allele 31.2 which can be found in the cell line DNA NA9948. The binding regions of primer pair 4 are underlined. Capital letters indicate the first nucleotide of a repeat which determines the allele nomenclature. Grey letters indicate differences between our findings and the DDBJ release (accession AB024611). We found nt A in position 147

instead of nt T. We did not observe nt T in position 339, but additionally, nt 330 and nt 343 were found. SNPs (*shown as letters with accent*): the type B sequence is characterised by nt C in position 155 and nt T in position 339, whereas the type A sequence exhibits nt A and nt C, respectively

The following control DNAs were typed to serve as standards for ladder calibrating: K562 (GIBCO and Promega), 9947A (Promega), NA 9948 and NA 03567 (Coriell Institute for Medical Research, Camden, NJ).

The homogeneity of the allele frequencies of males and females was compared by Fisher's exact test using Monte Carlo estimated *p*-values.

Formulae for calculation of parameters of forensic interest such as polymorphic information content etc. were used as given in Szibor et al. [8]. Statistical analysis of the DXS10011 population data to check the Hardy-Weinberg equilibrium (HWE) was performed according to Nei and

Roychoudhury [12]. The allele frequencies for different countries were arranged into very large contingency tables (e.g. 144 cells) and to test the hypothesis of homogeneity, the *p*-values were estimated using the Monte Carlo method. Adjusted standardised residuals were considered to localise differences between allele frequencies. All the tests were corrected for multiple comparisons. Mutation rates are given with an exact 95% confidence interval (Blyth-Still-Casella).

Table 1 Allele nomenclature, fragment length and repeat composition of 125 PCR fragments (alleles 26–50)

Allele	Type	bps	Repeat structure	Number sequenced
26	A	221	(GAAA) ₁₇ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
27	A	225	(GAAA) ₁₈ GAAG GAAA (GGAA) ₄ (AGAA) ₃	3
28	A	229	(GAAA) ₁₉ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
29	A	233	(GAAA) ₂₀ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
29.2	B	235	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₂ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
30	A	237	(GAAA) ₂₁ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
30.2	B	239	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₃ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
31	A	241	(GAAA) ₂₂ GAAG GAAA (GGAA) ₄ (AGAA) ₃	3
31.2	B	243	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₄ GAAG GAAA (GGAA) ₄ (AGAA) ₃	5
32	A	245	(GAAA) ₂₃ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
32.2	B	247	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₅ GAAG GAAA (GGAA) ₄ (AGAA) ₃	10
32.2*	B	247	GAAA GA (GAAA) ₇ GAGA (GAAA) ₁₄ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
33	A	249	(GAAA) ₂₄ GAAG GAAA (GGAA) ₄ (AGAA) ₃	3
33.2	B	251	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₆ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
33.2*	B	251	GAAA GA (GAAA) ₇ GAGA (GAAA) ₁₅ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
34	A	253	(GAAA) ₂₅ GAAG GAAA (GGAA) ₄ (AGAA) ₃	6
34.2	B	255	GAAA GA (GAAA) ₇ GAGA (GAAA) ₁₆ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
34.2*	B	255	GAAA GA (GAAA) ₈ GAGA (GAAA) ₁₅ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
35	A	257	(GAAA) ₂₆ GAAG GAAA (GGAA) ₄ (AGAA) ₃	6
35.2	B	259	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₈ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
36	A	261	(GAAA) ₂₇ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
36.2	B	263	GAAA GA (GAAA) ₆ GAGA (GAAA) ₁₉ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
37	A	265	(GAAA) ₂₈ GAAG GAAA (GGAA) ₄ (AGAA) ₃	6
38	A	269	(GAAA) ₂₉ GAAG GAAA (GGAA) ₄ (AGAA) ₃	5
39	A	273	(GAAA) ₃₀ GAAG GAAA (GGAA) ₄ (AGAA) ₃	6
40	A	277	(GAAA) ₃₁ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
40*	A	277	(GAAA) ₃₃ (GGAA) ₅ (AGAA) ₂	1
41	A	281	(GAAA) ₃₂ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
42	A	285	(GAAA) ₃₃ GAAG GAAA (GGAA) ₄ (AGAA) ₃	6
43	A	289	(GAAA) ₃₄ GAAG GAAA (GGAA) ₄ (AGAA) ₃	9
43.3	A	292	(GAAA) ₃₁ GAA (GAAA) ₃ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
44	A	293	(GAAA) ₃₅ GAAG GAAA (GGAA) ₄ (AGAA) ₃	3
45	A	297	(GAAA) ₃₆ GAAG GAAA (GGAA) ₄ (AGAA) ₃	4
46	A	301	(GAAA) ₃₇ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
47	A	305	(GAAA) ₃₈ GAAG GAAA (GGAA) ₄ (AGAA) ₃	3
48	A	305	(GAAA) ₃₉ GAAG GAAA (GGAA) ₄ (AGAA) ₃	2
49	A	313	(GAAA) ₄₀ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1
50	A	317	(GAAA) ₄₁ GAAG GAAA (GGAA) ₄ (AGAA) ₃	1

The fragment lengths (bps) are given for PCR products amplified with the primer pair 4. All sample with the exception of 40*, which originated from a Macedonian male, were taken from the German population.

Results

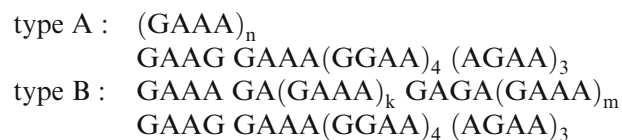
Flanking region sequence and primer design

The DXS10011 sequence was reported by Matsuki [2] in two versions (type A and type B). Comparing our data with the database release (DDBJ/EMBL/GenBank <http://srs.ddbj.nig.ac.jp> March 1999), we found some differences as shown in Fig. 1.

Two SNPs are located in the flanking region of the repeat region and one of them is located in the site of the reverse primer of the pairs 2 and 4. We have created the reverse primer of pair 4 with an appropriate wobble and did not observe any advantage with regard to PCR yield. This observation confirms experiences that one mismatch located far from the 3' end does not usually disturb the amplification process. The best amplification reliability was found for primer pair 4, which constantly detected a minimum of 10 pg DNA per 25 µl assay (diluted control male DNA). In contrast to this primer pair 3 may fail when less than 250 pg DNA is applied.

Repeat region and nomenclature

The allele designation proposed here is in compliance with the recommendations of the ISFH [6] and thus does not match with the alleles as given by Matsuki et al. [3] in all points exactly. The repeat region starts with a variable number of GAAA motifs and ends with 7 additional tetranucleotide repeats, which contain only the nucleotides G and A. The last 7 repeats and the intermediate single repeats GAAG and GAAA were used in our counting procedure. The intermediary single repeats can change into the neighbouring motifs and therefore contribute to the number of variable repeats (e.g. allele 40* in Table 1). The main difference between types A and B as shown in Fig. 1 is the additional occurrence of one GA and one GAGA repeat in type B which is counted as 1.2 repeats when the nomenclature for tetranucleotide STRs is used. Thus, type B is characterised by 0.2 interalleles. Common formulae for the repeat structure for type A and B can be given as follows:



The repeat composition of the 123 alleles shown in Table 1 is in compliance with the common repeat formulae. Only two sequences (alleles 40* and 43.3) fall out of this range.

PCR product sizing

Exact genotyping of three populations and routine kinship cases were performed on the background of sizing data of 123 sequenced alleles. When GAAA strands were labelled with 5-FAM during the PCR, the ABI310 data of these products had values that were up to 9 bp shorter than expected. Results of cell line DNA typing are shown in Table 2. These values can be used as intralaboratory and interlaboratory standards. Sequence analysis revealed a common repeat structure as listed in Table 1.

Population data

The DXS10011 STR is highly polymorphic in all population samples investigated here. Allele frequencies and all parameters of forensic interest calculated from our German population sample are listed in Table 3. We found 25 regular alleles classified as type A and 8 intermediate alleles, which corresponded to the type B structure. In addition, there was one irregular type A allele which showed a 0.3 repeat. Fisher's exact test did not reveal allele distribution differences between males and females ($p=0.187$) and therefore the samples can be cumulated.

Statistical examination of the observed heterozygosity did not reveal any deviation from the expectation for all three populations (Tables 3 and 4). However, our non-European sample sizes might be too small for definitive statements.

Mutations

Results of checking the meiotic stability are shown in Table 5. We found 5 incompatibilities in 782 mother-child pairs. The mutations appeared as changes of one repeat, i.e. a loss in two cases and a gain in two cases. In one case the mode of the mutation could not be identified. Hence, the female mutability was calculated as 0.0025–0.0147 (mean=0.0064) and the paternal mutation rate as 0.0207–0.0583 (mean=0.0367). We detected 6 single repeat gains in 409 father-daughter pairs, 8 single repeat losses and 1 two repeat loss. In our sample, only the sequence type A was affected by mutations. The range of alleles was from numbers 33 to 45. The age of the males at the time of conception of the offspring spanned from 19 to 37 years old.

Table 2 DXS10011 alleles of control DNAs which can be used for ladder calibration

Cell line DNA	Alleles
K562 (Promega, Gibco)	29
NA 9948 (Coriell)	31.2
NA 3567(Coriell)	38
9947A (Promega)	36–38

Table 3 DXS10011 allele frequencies and parameters of forensic interest in 745 male and 583 female Germans

Allele designation	Male number	Male proportion	Female number	Female proportion	Cumulated requery	
26	1	0.0013±0.0013	0	0	0.0005±0.0005	
27	2	0.0027±0.0019	8	0.0069±0.0024	0.0052±0.0017	
28	8	0.0107±0.0038	20	0.0172±0.0038	0.0147±0.0027	
29	16	0.0215±0.0053	25	0.0214±0.0042	0.0215±0.0033	
29.2	5	0.0067±0.0030	3	0.0026±0.0015	0.0042±0.0015	
30	9	0.0121±0.0040	13	0.0111±0.0031	0.0115±0.0024	
30.2	17	0.0228±0.0055	33	0.0283±0.0049	0.0262±0.0037	
31	8	0.0107±0.0038	18	0.0154±0.0036	0.0136±0.0027	
31.2	35	0.0470±0.0078	89	0.0763±0.0078	0.0649±0.0056	
32	18	0.0242±0.0056	32	0.0274±0.0048	0.0262±0.0037	
32.2	69	0.0926±0.0106	92	0.0789±0.0079	0.0842±0.0064	
33	12	0.0161±0.0046	29	0.0249±0.0046	0.0215±0.0033	
33.2	31	0.0416±0.0073	42	0.0360±0.0055	0.0382±0.0044	
34	30	0.0403±0.0072	52	0.0446±0.0060	0.0429±0.0046	
34.2	23	0.0309±0.0063	16	0.0137±0.0034	0.0204±0.0032	
35	46	0.0617±0.0088	65	0.0557±0.0067	0.0581±0.0054	
35.2	4	0.0054±0.0027	5	0.0043±0.0019	0.0047±0.0016	
36	55	0.0738±0.0096	80	0.0686±0.0074	0.0706±0.0059	
36.2	5	0.0067±0.0030	0	0	0.0026±0.0012	
37	53	0.0711±0.0094	88	0.0755±0.0077	0.0738±0.0060	
38	62	0.0832±0.0101	94	0.0806±0.0080	0.0816±0.0063	
39	43	0.0577±0.0085	69	0.0592±0.0069	0.0586±0.0054	
40	54	0.0725±0.0095	80	0.0686±0.0074	0.0701±0.0058	
41	35	0.0470±0.0078	61	0.0523±0.0065	0.0502±0.0050	
42	36	0.0483±0.0079	56	0.0480±0.0063	0.0481±0.0049	
43	22	0.0295±0.0062	39	0.0334±0.0053	0.0319±0.0040	
43.3	0	0	1	0.0009±0.0009	0.0005±0.0005	
44	16	0.0215±0.0053	26	0.0223±0.0043	0.0220±0.0034	
45	18	0.0242±0.0056	15	0.0129±0.0033	0.0173±0.0030	
46	6	0.0081±0.0033	6	0.0051±0.0021	0.0063±0.0018	
<i>PIC</i> Polymorphism information content.	47	5	0.0067±0.0030	5	0.0043±0.0019	0.0052±0.0017
<i>PD</i> Power of discrimination.	48	1	0.0013±0.0013	1	0.0009±0.0009	0.0010±0.0007
<i>MEC</i> Mean exclusion chance for chromosome X markers.	49	0	0	2	0.0017±0.0012	0.0010±0.0007
<i>Het exp.</i> Expected heterozygosity.	50	0	0	1	0.0009±0.0009	0.0005±0.0005
<i>Het obs.</i> Observed heterozygosity.		<i>PIC</i>	0.945	<i>Het exp.</i>	0.947±0.018	
		<i>MEC</i>	0.944	<i>Het obs.</i>	0.964	
		<i>PD (male)</i>	0.947	<i>PD (female)</i>	0.995	

Linkage analysis

As has been established by the investigation of 203 and 334 informative meioses, DXS10011 lies 6.8 cM (4.3–11.2) and 8.4 cM (4.5–15.1) distal to DXS8377 and DXS7423, respectively (Table 6). Linkage analyses to HPRTB and more distant markers (e.g. GATA172D05, DXS7133, DXS101, DXS7424, DXS6789, DXS6809, DXS9898, DXS6800, HumARA, DXS7132, DXS9902, DXS8378, DXS9895 and DXS6807) revealed LOD scores below 2.0 which indicate that non-genetic linkage exists (data not shown).

Discussion

Highly variable STR markers such as DXS10011 should be extensively explored when used in forensic genetics. Hence, it is justified to add further research results to earlier papers, which already described DXS10011 as a promising marker for forensic purpose. Our results are broadly in accordance with Matsuki et al. [3], but in contrast to these authors, the allele designation proposed here involves 9 more repeats which are located at the 3' end of the repeat region. Our nomenclature proposal is in compliance with the recommendations of the International Society of Forensic Haemogenetics (ISFH) [6]. Hence, Matsuki et al.'s population data can be compared with our results. The alleles 30, 31 and 33.2 were more frequently found in Japanese than in German populations, whereas an

Table 4 DXS10011 alleles found in Vietnamese (19 male and 97 female) and Peruvians (40 male and 41 female)

Allele designation	Vietnam male Number	Vietnam female number	Peru male number	Peru female number
<26				1
26		1		3
27		1	3	4
28		3		5
29	1	5	2	6
30		11	1	2
30.2		2		
31		9		1
31.2		2	1	
32	1	2	1	1
32.2	2	7	1	
33		5	1	5
33.2	1	15	2	
34		8	1	6
34.2	2	8		
35	2	8	3	2
35.2		2		
36	1	13	4	4
36.2		2		
37		12	3	6
38	2	13	1	4
39	1	16	3	6
40	2	13	6	9
41	1	7	3	8
42		5	1	6
42.2		1		
43		6	1	1
44	1	3	1	
45	2	5		1
46		4	1	
47		2		1
48		3		
Het exp.		0.956 ±0.041		0.945 ±0.070
Het obs.		0.9278		0.951

opposite situation was observed with regard to the allele 31.2. Similarities in the allele distribution seem to exist between Japanese and Vietnamese populations, however, this is based on a relatively small data base.

As shown by Matsuki et al. [3] and Watanabe et al. [4], DXS10011 is a very complex marker and sequencing a high number of PCR products revealed some structural variants within alleles of identical length. The rare allele variant 43.3 containing a 0.3 repeat, seems to be the result of a loss of a single base in the regular repeat region (type A). We also introduced a few sequence corrections and addressed two SNPs located in the repeat flanking region associated only with sequence type B. The sequence corrections for type A are in accordance with NCBI sequence information from SNP ss85177997 (entry

Table 5 Number and mode of DXS10011 mutations in relation to the age distribution of the parents at the time of conception

Age at conception (years)	Number of mothers	Mutation mode	Number of fathers	Mutation mode
Unknown or anonymised group	375	1U	3	-
15–19	81	1L	30	1L
20–24	145		103	2L, 1G
25–29	106	2G, 1L	113	2L, 1*L, 4G
30–34	55		94	2L, 1G
35–39	16		33	1L
40–44	4		18	
45–49	0		9	
50–54	0		0	
55–59	0		5	
60–64	0		1	

If not otherwise indicated the change of the repeat number was 1. G Gain, L loss, U unidentified mutation mode, *2 repeat changes.

Table 6 Genetic mapping results of DXS10011 in relation to DXS7423 and DXS8377

Marker	Number of informative meioses investigated	Max. LOD score (Z_{max})	Genetic distance (cM Kosambi)
DXS7423	203	10.6	8.4 (4.5–15.1)
DXS8377	334	24.5	6.8 (4.3–11.2)

The genetic distance is given with the $Z_{max}-1$ support interval.

05/23/03) and ss10557653 (entry 06/29/03). One of the SNPs is located in the binding site of the reverse primer of pair 4 which causes systematic mismatches. Nevertheless, it does not disturb the PCR success, and compared with the other primers described here, pair 4 is highly effective. Using the Primer3 software it was not possible to design an alternative primer exhibiting high specificity and sensitivity outside the mismatch region.

We did not observe an advantage when primers with an appropriate wobble were used. However, in general there is the risk that individuals exists which have an unexpected single nucleotide mutation in a primer binding site [13]. At the SNP discussed here, such an event would most likely result in a null allele. Therefore, using a wobbled reverse primer for pair 4 would give this system more robustness.

Alleles with incomplete repeats showed differences of 1 bp between some alleles. Therefore, typing of DXS10011 requires a high measurement accuracy and therefore highly discriminating electrophoresis systems such as gene scanner systems should be used. The phenomenon of large differences between the real length measured by sequencing and the virtual fragment length

computed by the ABI310 system may be caused by the complementary combination of a heavy chain (AAAG strand) and a light chain (TCCC strand). Labelling of the TCCC strand and/or the usage of alternative dyes may dramatically change this situation. Typing the world-wide available control DNAs which are shown in Table 2, may help to adjust the measurement conditions and establish allelic ladders [14].

Population studies show that in the German population none of the DXS10011 alleles exceeded a frequency value of 0.09. Due to the very high number of alleles, a comparison of the observed and expected genotypes (exact-test as suggested by Guo and Thompson [15]) is not practicable. However, comparison of the estimated and the expected heterozygosity gave no hint towards a possible deviation of the allele distribution from the Hardy-Weinberg equilibrium.

Concerning mutations, our findings are in accordance with the observation of Brinkmann et al. [16] that a positive exponential correlation exists between mutation rate and the geometric mean of the number of uninterrupted repeats. The long homogenous repeat structure of DXS10011 alleles (type A) is conducive to mutation especially in males who have a high number of cell divisions in spermatogenesis. Furthermore, it seems that during phylogenesis, the high mutability of the type A sequences led to a higher number of type A alleles whereas the low mutating B type sequences created much less alleles. Using DXS1001 in kinship testing, one has to consider the high mutation rate and it may be justified to assess mutations in type A and type B differently.

Forensic use of ChrX markers requires knowledge on the linkage situation. DXS10011 which is located at Xq28, approximately 149.806 Mb from the top of the X chromosome, is physically close to DXS7423 (148.351 Mb) and DXS8377 (148.207 Mb, http://www.ensembl.org/Homo_sapiens/contigview). According to an earlier suggestion (Szibor et al. [8]), these q-telomeric markers can be ordered in the linkage group 4. We estimated the genetic distance from DXS10011 to DXS8377 and DXS7423 by LOD score analysis with 6.8 and 8.4 cM, respectively. These findings are in relative concordance with the expected range. Earlier findings suggest that near the end of the long arm Xq27.2-qter, 1 cM corresponds to 340–800 kb (Kenwick and Gitschier [17]). When markers are closely linked, linkage disequilibrium (LDE) may occur. Due to the multiple haplotype combinations caused by the very high number of DXS10011 alleles, our present study is not sufficient to carry out a LDE evaluation. On the other hand, LDE checking between markers distanced more than 5 cM is not strictly required since a LDE manifestation is unlikely in such cases. Certainly, LDE is not simply a monotonic function of the distance but there are also other influences present which are as yet poorly understood. However, it has been shown that the LDE between SNPs at Xp28 drops against zero when their distance exceeds 250 kb (Taillon-Miller et al. [18]). Due to their much higher mutability, STRs are normally much less affected by LDE than SNPs. This may justify the use of the three STRs in combination even if a LDE test was not performed. Haplotyping of these three STRs as a tool for solving

complicated deficiency kinship cases may be very successful.

Summarising the results of our investigation we agree with earlier findings that DXS10011 is a powerful marker for forensic purpose. Further practical experience is required to ensure that the DNA typing technique suggested here leads to the highest degree of reliability.

References

- Gerken SC, Matsunami N, Plaetke R et al. (1994) HUMUT413: DDBJ/GenBank accession #L29968
- Matsuki T (1999) DXS10011: accession #AB024610 and accession #AB024611. DDBJ/GenBank/EMBL <http://www.ncbi.nlm.nih.gov>
- Matsuki T, Sawazaki K, Tsubota E, Iida R (2003) DXS10011: a hypervariable TTC/GAAA repeat marker on human chromosome Xq27-q28. In: Brinkmann B, Carracedo A (eds) *Progress in forensic genetics 9*. International Congress Series 1239, pp 363–366
- 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
- Koyama H, Iwasa M, Tsuchimochi T et al. (2002) Y-STR haplotype data and allele frequency of the DXS10011 locus in a Japanese population sample. *Forensic Sci Int* 125:273–276
- DNA Commission of the International Society of Forensic Haemogenetics (1994) DNA recommendations—1994 report concerning further recommendations of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. *Int J Legal Med* 107:159–160
- 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
- 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* 117:67–74
- 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
- Szibor R, Edelmann J, Zarrabeitia MT, Riancho JA (2003) Sequence structure and population data of the X-linked markers DXS7423 and DXS8377—clarification of conflicting statements published by two working groups. *Forensic Sci Int* 134:72–73
- Hering S, Szibor R (2000) Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. *J Forensic Sci* 45:929–931
- Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics* 76:379–390
- Hering S, Edelmann J, Dreßler J (2002) Sequence variations in the primer binding regions of the highly polymorphic STR system SE33. *Int J Legal Med* 116:365–367
- 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
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361–372
- 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
- Kenwick S, Gitschier J (1989) A contiguous, 3-Mb physical map of Xq28 extending from the color blindness locus to DXS15. *Am J Hum Genet* 45:873–882
- Taillon-Miller P, Bauer-Sardina I, Saccone NL et al. (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nat Genet* 25:324–328