Int J Legal Med (2006) 120: 337–345
DOI 10.1007/s00414-005-0061-y

ORIGINAL ARTICLE

Sandra Hering · Christa Augustin · Jeanett Edelmann · Micaela Heidel · Jan Dressler · Heike Rodig · Eberhard Kuhlisch · Reinhard Szibor

DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases

Received: 18 July 2005 / Accepted: 25 October 2005 / Published online: 13 December 2005

Abstract The evaluation of the short tandem repeat (STR) markers DXS10079, DXS10074 and DXS10075 was amended to establish a STR cluster spanning a genetic distance<1 cM. These three STRs are located within a 280-kb region at Xq12 and provide stable haplotypes useful for solving complex kinship cases. Theoretically, this cluster could give rise to 2,548 different haplotypes in the German population and the genotyping of 781 men revealed the presence of 172 haplotypes. Since the three STRs were shown to be in strong linkage disequilibrium (LD), haplotype frequencies cannot be computed on the basis of a single locus allele frequency alone but have to be estimated directly. Here, we present data on linkage, haplotype frequencies and LD in a German population. Further clusters from other regions of the X chromosome will be published in the future to cover the chromosome with a well-structured network of highly informative sites.

Keywords Kinship testing · X chromosome · STR cluster · Haplotyping · Xq12

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00414-005-0061-y

S. Hering · M. Heidel · J. Dressler
Institut für Rechtsmedizin,
Technische Universität Dresden,
Fetscherstrasse 74,
01307 Dresden, Germany

C. Augustin
Institut für Rechtsmedizin,
Universitätsklinikum Hamburg-Eppendorf,
Butenfeld 34,
22529 Hamburg, Germany

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

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

H. Rodig
Biotype AG,
Moritzburger Weg 67,
01109 Dresden, Germany

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

Introduction

In kinship testing not all pedigree constellations are suitable for X chromosome (ChrX) analysis. However, when X-chromosomal transmission lines are present, typing of ChrX short tandem repeats (STRs) can significantly contribute to solving complex deficiency cases [17]. The number of established ChrX markers suitable for forensic usage has risen continually during recent years. One of the advantages of ChrX typing is that in men, ChrX typing reveals the haplotype directly. Investigation of second or third degree kinships such as the relationship between cousins or between aunts and nieces would require extremely polymorphic markers. However, the degree of polymorphism is limited in STRs. Furthermore, extremely polymorphic STRs such as DXS1001 are highly prone to mutations [3, 13]. Therefore, our group is engaged in the investigation of ChrX STR clusters, which can provide stable haplotypes for forensic use. When the linkage between the ChrX STRs is close enough, haplotypes can be used in kinship testing instead of a single STR. We previously combined STR clusters from known markers [9, 19] and were able to establish clusters that span 2–5 cM. In another case, we established a haplotyping procedure for two very closely linked pairs of STRs [16]. We have conducted an amended GenBank search to find STR clusters that provide highly stable haplotypes and

present three newly established markers, DXS10079, DXS10074 and DXS10075, which can potentially produce 2,548 haplotypes. This number was calculated from the allele numbers of each STR, i.e., 13, 14 and 14 for DXS10079, DXS10074 and DXS10075, respectively. However, due to a considerable linkage disequilibrium (LD) the real number of haplotypes is much lower. The recombination frequency of this cluster is in the same magnitude as, or lower than, the average mutation frequency of STR markers. The STR clusters subjected to investigation were selected according to their localization and are located within a 280-kb region of Xq12. Further, closely-linked STR combinations from other regions will be published in the future to cover the ChrX with a well-structured network of highly informative sites.

Materials and methods

We imported several ChrX contig sequences from http://www.ncbi.nlm.nih.gov into the GeneRunner program (http://www.generunner.com) and examined the Xq12 region for all tri-, tetra- and pentanucleotide repeats. STRs exhibiting ten or more repeat units were regarded as loci of interest and some were submitted for experimental evaluation. As a result we chose three tetranucleotide microsatellites which are now registered in a gene database (http://www.gdb.org) as DXS10079, DXS10074 and DXS10075.

Robustness in the amplification procedure, low tendency to stutter, moderate or high polymorphism and homogeneity of the repeat structure were of prime importance for selection. Primer binding regions were defined avoiding loci of known single nucleotide polymorphisms. The three STRs are located in the contig NT_011669, DXS10079 in component AL049564 and DXS10074 and DXS10075 in component AL356358. The physical distance from DXS10079 to DXS10075 is approximately 260 kb, and 280 kb from DXS10079 to DXS10075.

Blood samples and buccal swabs for population studies were collected from unrelated Germans (693 men and 328 women). Cases of routine kinship testing were also done and 237 trios with female children with a paternity index of 1,000 or more were investigated. To estimate the stability of the 3-locus haplotypes in female meioses, the DNA of 99 (grand) father–daughter–grandson trios were typed. The samples were predominantly taken from students and their families, and were made anonymous before they were processed.

The DNA extraction was carried out using the commercially available QIAamp DNA Blood Kit (Qiagen GmbH, Hilden, Germany). The primers were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Amplification was carried out in a 25-µl polymerase chain reaction (PCR) volume containing approximately 0.1–1 ng DNA, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 1 U *Taq* polymerase (Perkin–Elmer, Foster City, CA) and 1xPCR buffer. The cycle conditions in a PTC-200 cycler (MJ Research, Watertown, MASS) were as follows: (1) 95°C for 3 min soaking; (2) 94°C for 30 s; (3) 58°C for 1 min (primer pair 4) or 60°C (primer pairs 1–3 and 5–6); (4) 72°C for 1 min, for 30 cycles and (5) 72°C for 10 min final extension. For multiplexing, primer concentrations were reduced to 0.2 µM for pair 1 and 0.15 µM for pairs 2 and 3. Following our initial experiments the annealing temperature suitable for all primer pairs was adjusted at 60°C.

The primers for amplifying long amplicons suitable as sequencing templates were synthesized as follows:

- DXS10079: Amplicon length for allele 17 was 556 bp.
- Primer 1 F: 5'-AGG AGA ATG GCT TGA ACC TG-3'
- Primer 1 R: 5'-TGG GTA TCC TTG TGT TCC AA-3'
- DXS10074: Amplicon length for allele 14 was 498 bp.
- Primer 2 F: 5'-TAG GCG CTT CCT AGA CCT CA-3'
- Primer 2 R: 5'-CCT TCC TTC CCA TGT TCT CA-3'
- DXS10075: Amplicon length for allele 17 was 462 bp.
- Primer 3 F: 5'-GGC TTC AGA AGG CAG AAA TG-3'
- Primer 3 R: 5'-CCA TGG ATC CCC TAA CAG AA-3'.

These primers were used in a direct *Taq*-cycle sequencing procedure using the BigDye-Terminator Kit version 1.1 and the ABI Prism 310 sequencer, Sequencing Analysis version 3.7 (Perkin–Elmer, Foster City, CA) as described [12].

To determine the alleles by the amplicon length, the primer pairs 4, 5 and 6 were used in a triplex PCR.

- DXS10079
- Primer 4 F: 5'-TET-AGA TTG TGC CAA TGC TCT CC-3'
- Primer 4 R: 5'-GTT TGC CTG TGT TGT AAC ATC CTT-3'
- A G was added to the 5'-end of the reverse primer to obtain preferentially adenylated PCR products ("pig-tailing") [4].
- DXS10074
- Primer 5 F: 5'-HEX-ACT TCC TAC TGC CCC ACC TT-3'
- Primer 5 R: 5'-GTT TCC CCT CAG AGA GCT GAC ACA-3'
- The reverse primer was GTTT-tailed.
- DXS10075
- Primer 6 F: 5'-FAM-AGG AGG GGC CTA GAC AAG TG-3'
- Primer 6 R: 5'-CAG ATT ATG CTT GGG CCT GT-3'

Additional X-chromosomal STRs were typed using the Mentype Argus X Pentaplex-UL PCR amplification kit (Biotype, Dresden, Germany) or as described [8, 10, 13].

The resulting PCR products were resolved and detected by capillary electrophoresis in the denaturing polymer POP4 (Perkin–Elmer, Foster City, CA) in the ABI 310 sequencer (Perkin–Elmer, Foster City, CA, USA) following standard protocols. Amplicon sizing was supported using the GeneScan 500 TAMRA size standard (Perkin–Elmer, Foster City, CA).

As standards for ladder calibration we typed the cell lines K562 (GIBCO and Promega), 9947A (Promega), NA 9948 and NA 03567 (Coriell Institute for Medical Research, Camden, NJ).

Homogeneity of the allele frequencies of men and women was compared by Fisher's exact test using Monte Carlo estimated p values (99% confidence interval). For

Table 1 DXS10079: allele nomenclature, amplicon length and repeat composition of 31 PCR amplicons (alleles 14–25)

Allele	bps[a]	Repeat structure			Number sequenced
14	271	$(AGAA)_{10}$		AGAG $(AGAA)_3$	2
15	275	$(AGAA)_{11}$		AGAG $(AGAA)_3$	2
16	279	$(AGAA)_{12}$		AGAG $(AGAA)_3$	2
17	283	$(AGAA)_{13}$		AGAG $(AGAA)_3$	2
18	287	$(AGAA)_{14}$		AGAG $(AGAA)_3$	2
19	291	$(AGAA)_{15}$		AGAG $(AGAA)_3$	8
19.3	294	$(AGAA)_5$	AGA $(AGAA)_{10}$	AGAG $(AGAA)_3$	1
20	295	$(AGAA)_{16}$		AGAG $(AGAA)_3$	4
21	299	$(AGAA)_{17}$		AGAG $(AGAA)_3$	1
22	303	$(AGAA)_{18}$		AGAG $(AGAA)_3$	2
23	307	$(AGAA)_{19}$		AGAG $(AGAA)_3$	3
24	311	$(AGAA)_{20}$		AGAG $(AGAA)_3$	1
25	315	$(AGAA)_{21}$		AGAG $(AGAA)_3$	1

[a]The amplicon lengths (bps) are given for PCR products amplified with the primer pair 4

calculation of parameters of forensic interest such as polymorphism information content (PIC), we used formulas which we have reviewed recently [17]. Statistical analyses of the population data to detect any deviation from the Hardy–Weinberg equilibrium (HWE) were performed by an exact test [11] using the DNAVIEW software package version 21.06 (Brenner C. 1998, Berkeley, CA). The observed heterozygosity was compared with the unbiased estimate of the expected heterozygosity according to Nei and Roychoudhury [15]. Mutation rates are given with an exact 95% confidence interval [2, 5].

To test for association of STR alleles in 781 haplotypes of men, we performed an exact χ^2-test based on a two-sided Monte Carlo estimated p value. The Monte Carlo p value was calculated using the program StatXact-4 (Cytel Software Corporation, Cambridge, MA).

Results

Sequencing analysis and nomenclature

The allele designation proposed here is in compliance with the recommendations of the International Society for Forensic Genetics [1].

DXS10079

The structure of this tandem tetranucleotide repeat region can be summarized as follows: $(AGAG)_3$ TGAA AGAG **$(AGAA)_n$AGAG $(AGAA)_3$**. This STR contains a long homogenous repeat structure $(AGAA)_n$ in which n varies from 10 to 21. In the German population investigated, only

Table 2 DXS10074: Allele nomenclature, amplicon length and repeat composition of 30 PCR amplicons (alleles 7–21)

Allele	bps[a]	Repeat structure			Number sequenced
7	165	$(AAGA)_7$			4
8	169	$(AAGA)_8$			2
9	173	$(AAGA)_9$			2
10	177	$(AAGA)_{10}$			1
13	195	$(AAGA)_{10}$	AAGG	$(AAGA)_2$	1
14	199	$(AAGA)_{11}$	AAGG	$(AAGA)_2$	1
15	203	$(AAGA)_{12}$	AAGG	$(AAGA)_2$	2
16	207	$(AAGA)_{13}$	AAGG	$(AAGA)_2$	1
16.2	209	AA———	$(AAGA)_{13}$	AAGG $(AAGA)_2$	1
17	211	$(AAGA)_{14}$	AAGG	$(AAGA)_2$	5
18	215	$(AAGA)_{15}$	AAGG	$(AAGA)_2$	3
19	219	$(AAGA)_{16}$	AAGG	$(AAGA)_2$	3
20	223	$(AAGA)_{17}$	AAGG	$(AAGA)_2$	1
21	227	$(AAGA)_{18}$	AAGG	$(AAGA)_2$	3

[a]The amplicon lengths (bps) are given for PCR products amplified with the primer pair 5

Table 3 DXS10075: allele nomenclature, amplicon length and repeat composition of 26 PCR amplicons (alleles 12–21.3)

Allele	bps[a]	Repeat structure					Number sequenced
12	218	$(TAGA)_5$	*TGA*			$(TAGA)_7$	2
13	222	$(TAGA)_4$	*TGA*			$(TAGA)_9$	3
14	226	$(TAGA)_4$	*TGA*			$(TAGA)_{10}$	1
15	230	$(TAGA)_4$	*TGA*			$(TAGA)_{11}$	1
16	234	$(TAGA)_5$	*TGA*			$(TAGA)_{11}$	2
16.2	236	$(TAGA)_5$	*TGA*	TAGA	TA	$(TAGA)_{10}$	1
16.3	237	$(TAGA)_5$	*TGA*	$(TAGA)_9$	TGA	$(TAGA)_2$	2
17	238	$(TAGA)_5$	*TGA*			$(TAGA)_{12}$	4
18	242	$(TAGA)_5$	*TGA*			$(TAGA)_{13}$	5
19	246	$(TAGA)_5$	*TGA*			$(TAGA)_{14}$	1
20	250	$(TAGA)_5$	*TGA*			$(TAGA)_{15}$	1
20.3	253	$(TAGA)_5$	*TGA*	$(TAGA)_3$	TGA	$(TAGA)_{12}$	1
21	254	$(TAGA)_5$	*TGA*			$(TAGA)_{16}$	1
21.3	257	$(TAGA)_5$	*TGA*	$(TAGA)_3$	TGA	$(TAGA)_{13}$	1

[a]The amplicon lengths (bps) are given for PCR products amplified with the primer pair 6

one intermediate allele containing the incomplete motif AGA was found as a 19.3 allele. The repeat flanking region did not show any variations. Details are shown in Table 1.

DXS10074

This STR is characterized by two different types of alleles (Table 2):

1. Short alleles are in agreement with the GenBank sequence and contain the repeat block $(AAGA)_n$ in which *n* varies between 7 and 10: $(ACAC)_2$ $(AGAG)_4$ AA AAAG **$(AAGA)_n$** AAGG AAGA.
2. Longer alleles contain the repeat block $(AAGA)_n$ in which *n* varies from 10 to 18. This allele type is characterized by a 12 base pair insertion, which is underlined in the following common formula: $(ACAC)_2$ $(AGAG)_4$ AA AAAG **$(AAGA)_n$**<u>AAGG $(AAGA)_2$</u> AAGG AAGA.

We propose to count the number of the repeats as (*n*+3). The only intermediate allele we have found originates from an AA-insertion (in bold italics) upstream of the variable (AAGA)-repeats and in our nomenclature is designated as a 16.2 allele: $(ACAC)_2$ $(AGAG)_4$ ***AA*** AA AAAG $(AAGA)_{13}$ AAGG $(AAGA)_2$ AAGG AAGA.

DXS10075

DXS10075 shows two variable repeat blocks and can be described with the following common formula, in which *n* is taken as 4 or 5 and *m* as 7–15: TATC **$(TAGA)_n$***TGA* **$(TAGA)_m$**. The complete data are given in Table 3. The conserved TGA motif (in italics) was not included in the allele nomenclature. Intermediate X.3 alleles originated from a second TGA motif disrupting the $(TAGA)_m$-repeats. A further type of intermediate allele (16.2) is caused by an incomplete repeat unit in the block **$(TAGA)_m$**.

Amplicon analysis

All repeat structures of the three microsatellites revealed by sequence analysis are listed in Tables 1, 2 and 3. DNA typing patterns of cell lines are shown in Table 4. These results can be used as intra- and interlaboratory standards.

Allele frequencies of a German sample are listed in Table 5. Fisher's exact test did not reveal allele distribution differences between men and women (DXS10079: p=0.0599, 0.0538–0.0660; DXS10074: p=0.6668 and DXS10075: p=0.6544) and therefore results from male and female samples can be combined. Parameters of forensic interest are shown in Table 6. The examination of the observed heterozygosity did not reveal any deviation from the expectation for all three loci. However, the exact test failed for DXS10075 (p=0.025, 0.023–0.027).

The evaluation of each STR justifies the statement that these three microsatellites which exhibit PIC values of 0.811, 0.778 and 0.645 are highly or moderately variable.

Testing an eastern German population of 781 unrelated men revealed 172 different haplotypes of which 72% showed frequencies<0.02 (Table 7). The complete haplotype distribution of our sample will be published as Electronic Supplementary Material.

Table 4 Xq12 alleles of control DNAs which can be used for ladder calibration

Cell line DNA	DXS10079	DXS10074	DXS10075
K562 (Promega, Gibco)	17	17	18
NA 9948 (Coriell)	19	18	18
NA 3567(Coriell)	19	7	13
9947A (Promega)	20–23	16–19	17–18

Table 5 Allele frequencies in 693 male and 328 female Germans

Allele designation	DXS10079 male proportion	DXS10079 female proportion	DXS10079 cumulated frequency	DXS10074 male proportion	DXS10074 female proportion	DXS10074 cumulated frequency	DXS10075 male proportion	DXS10075 female proportion	DXS10075 cumulated frequency
7				0.0620±0.0092	0.0701±0.0100	0.0660±0.0068			
8				0.1299±0.0128	0.1402±0.0136	0.1349±0.0093			
9				0.0130±0.0043	0.0152±0.0048	0.0141±0.0032			
10				0.0014±0.0014	0	0.0007±0.0007			
12							0.0029±0.0020	0.0015±0.0015	0.0022±0.0013
13				0.0058±0.0029	0.0046±0.0026	0.0052±0.0020	0.0620±0.0092	0.0640±0.0096	0.0630±0.0066
14	0.0043±0.0025	0.0046±0.0026	0.0044±0.0018	0.0101±0.0038	0.0076±0.0034	0.0089±0.0026	0.0058±0.0029	0.0122±0.0043	0.0089±0.0026
15	0.0361±0.0071	0.0244±0.0060	0.0304±0.0047	0.0779±0.0102	0.0625±0.0095	0.0704±0.0070	0.0043±0.0025	0.0107±0.0040	0.0074±0.0023
15.3									
16	0.0418±0.0076	0.0213±0.0056	0.0319±0.0048	0.2006±0.0152	0.2363±0.0166	0.2179±0.0112	0.1544±0.0137	0.1372±0.0134	0.1460±0.0096
16.2				0.0029±0.0020	0.0015±0.0015	0.0022±0.0013	0	0.0015±0.0015	0.0007±0.0007
16.3							0	0.0030±0.0022	0.0015±0.0010
17	0.0635±0.0093	0.0564±0.0090	0.0600±0.0065	0.2496±0.0164	0.2485±0.0169	0.2491±0.0118	0.4444±0.0189	0.4497±0.0194	0.4470±0.0135
18	0.1385±0.0131	0.1235±0.0128	0.1312±0.0092	0.1631±0.0140	0.1341±0.0133	0.1490±0.0097	0.3016±0.0174	0.2896±0.0177	0.2958±0.0124
19	0.2367±0.0161	0.2713±0.0174	0.2535±0.0118	0.0722±0.0098	0.0579±0.0091	0.0652±0.0067	0.0188±0.0052	0.0229±0.0058	0.0208±0.0039
19.3	0.0014±0.0014	0	0.0007±0.0007						
20	0.2944±0.0173	0.2759±0.0175	0.2854±0.0123	0.0101±0.0038	0.0168±0.0050	0.0133±0.0031	0.0043±0.0025	0.0046±0.0026	0.0044±0.0018
20.3							0	0.0015±0.0015	0.0007±0.0007
21	0.1300±0.0132	0.1433±0.0137	0.1416±0.0095	0.0014±0.0014	0.0046±0.0026	0.0030±0.0015	0	0.0015±0.0015	0.0007±0.0007
21.3							0.0014±0.0014	0	0.0007±0.0007
22	0.0332±0.0068	0.0625±0.0095	0.0474±0.0058						
23	0.0087±0.0035	0.0152±0.0048	0.0119±0.0029						
24	0.0014±0.0014	0	0.0007±0.0007						
25	0	0.0015±0.0015	0.0007±0.0007						

Table 6 Parameters of forensic interest in 693 male and 328 female Germans

Parameter	DXS10079	DXS10074	DXS10075
Het obs.	0.774	0.851	0.668
Het exp.	0.807 (0.765–0.850)	0.833 (0.793–0.874)	0.691 (0.641–0.741)
HWE exact test	p=0.161 (0.157–0.165)	p=0.446 (0.441–0.451)	p=0.025 (0.023–0.027)
PIC	0.778	0.811	0.645
MEC	0.780	0.811	0.642
PD (male)	0.806	0.832	0.690
PD (female)	0.937	0.951	0.856

PIC Polymorphism information content, *PD* power of discrimination, *MEC* mean exclusion chance for chromosome X markers, *Het exp.* expected heterozygosity, *Het obs.* observed heterozygosity

The test for association indicated a LD between DXS10079 and DXS10074 (p=0.025, 0.021–0.029), DXS10079 and DXS10075 (p=0.006, 0.004–0.008) and between DXS10074 and DXS10075 (p=0.002, 0.001–0.003).

Table 7 Haplotype analysis in German men of the DXS10079–DXS10074–DXS10075 cluster on Xq12

Haplotype			*N*	Frequency	
DXS10079	DXS10074	DXS10075		MLE	UCL
20	16	17	40	0.05122	0.06615
19	17	17	29	0.03713	0.05029
20	17	17	28	0.03585	0.04883
20	17	18	25	0.03201	0.04442
19	17	18	19	0.02433	0.03549
18	8	17	18	0.02305	0.03399
18	17	17	18	0.02305	0.03399
20	18	18	18	0.02305	0.03399
21	17	18	17	0.02177	0.03247
19	18	18	15	0.01921	0.02942
20	16	18	15	0.01921	0.02942
19	16	17	14	0.01793	0.02788
19	7	13	13	0.01665	0.02633
20	7	13	13	0.01665	0.02633
17	8	17	12	0.01536	0.02478
18	16	17	12	0.01536	0.02478
19	16	18	12	0.01536	0.02478
19	18	17	12	0.01536	0.02478
20	18	17	12	0.01536	0.02478
15	18	16	11	0.01408	0.02321
16	16	16	11	0.01408	0.02321
19	8	17	11	0.01408	0.02321
19	19	17	11	0.01408	0.02321
21	16	17	11	0.01408	0.02321
18	18	17	10	0.01280	0.02162
20	19	17	10	0.01280	0.02162
21	16	18	10	0.01280	0.02162
18	7	13	9	0.01152	0.02002
19	8	16	8	0.01024	0.01841
19	8	18	8	0.01024	0.01841
19	15	17	8	0.01024	0.01841
20	19	18	8	0.01024	0.01841
21	17	17	8	0.01024	0.01841
21	18	18	8	0.01024	0.01841

Only haplotypes with a MLE>0.01 are displayed here
MLE Maximum likelihood estimate, *UCL* upper 95% confidence limit

Mutations

Checking the meiotic stability of these three loci, we found no mother–child exclusions in 388 meioses. In 333 paternal ChrX transfers, we detected a total of 12 incompatibilities involving all three loci. As shown in Table 8, the mutations appeared as the gain or loss of one repeat in each case. The highest mutation rate was found for DXS10079 with six inconsistencies in paternal meioses. Furthermore, in three families we observed evidence of a duplication of this locus (Table 9), where two of the female children exhibited a 3-allele pattern. In the third family the father showed a biallelic pattern. As a consequence, his daughter carried a double dose of one allele, which could

Table 8 Number and mode of mutations in relation to the age distribution of the 333 fathers and 388 mothers at the time of conception

Age at conception (years)	Total number (fathers)	Total number (mothers)	DXS10079 number of fathers	DXS10074 number of fathers	DXS10075 number of fathers
Unknown	25	42			1 L
14–19	32	70	1 L		
20–24	69	124	1 G, 1 L		
25–29	104	89	1 G	2 L, 1 G	
30–34	66	52			1 G
35–39	23	11	1 L		
40–44	9	0	1 G		
45–49	5	0		1 L	
>50	0	0			
Paternal mutation rate[a]			0.0180 (0.0066–0.0388)	0.0120 (0.0033–0.0305)	0.0060 (0.0007–0.0215)
Maternal mutation rate[a]			0 (0–0.0095)	0 (0–0.0095)	0 (0–0.0095)

The change of the repeat number was 1. No mutations were found in 388 maternal meioses
G Gain, *L* loss
[a]Maximum likelihood estimation of mutation rates with 95% confidence limit

Table 9 Three families with presumed duplications at locus DXS10079

	Mother	Daughter	Father
Family 1	20/21/22	18/21/22	18
Family 2	20/22	19/20/22	19
Family 3	19/19	19/20	19/20

Underlined alleles exhibited the twofold peak height at amplicon analysis

be recognized by a twofold peak height. No duplications were observed at the other loci investigated.

In DXS10079 we found one man with an additional allele much smaller in intensity compared with the regular alleles with a peak height approximately one-third of the main allele. Likewise, in another man, an additional allele was found in DXS10074. The peak height was about 15% compared to the main allele. In each of the two cases, the minor allele was one repeat longer than the main allele, which rules out the possibility that this occurrence can be regarded as a stutter phenomenon. In these present cases both men transmitted only the main alleles to their daughters. We have never observed aberrant alleles of this type at other ChrX loci.

Investigation of haplotype stability and segregation

Recombination analysis of 152 (grand)father–daughter–grandson trios did not provide any evidence of a cross-over within the 280-kb region between DXS10079 and DXS10075. However, only 57 out of 106 women were fully informative, i.e., heterozygous in all the three loci; 35 women were heterozygous in two STRs. No information was obtainable regarding women who were homozygous in two (n=12) or in all three (n=2) loci.

The same (grand)father–daughter–grandson trios were checked for segregation. In 150 informative constellations the ratio of grandmaternal is to grandpaternal haplotypes in the grandsons was 76:74. This result exactly meets the statistical expectation. Testing constellations are depicted in Fig. 1.

Discussion

We have shown previously that for closely linked locus clusters such as DXS7424 and DXS101 [9] and DXS6801, DXS6809 and DXS6789 [19], X-chromosomal haplotyping can greatly facilitate complex kinship testing. The chance of success in X-chromosomal kinship testing critically depends upon the availability of a dense range of markers and marker clusters. Since the segregation of chromosomal segments in a family occurs at random, even a panel of well-established markers may be completely uninformative in a given case. Hence, our current efforts to evaluate the Xq12 cluster are aimed at improving the density of X-chromosomal markers that are readily formatted for use in forensic practice. Our finding that about 74% of haplotypes of this newly established cluster occur at a frequency of less than 2% in the German population illustrates that our efforts have been worthwhile, and that the cluster provides a powerful tool for kinship testing. Since this STR cluster lies in a 280-kb segment, it provides relatively stable haplotypes. Before X-chromosomal STRs or STR clusters can be routinely used in forensic practice, it is necessary to investigate and confirm the intermarker recombination rates expected on the basis of their physical localization. As observed in the present study, genetic linkage between the markers

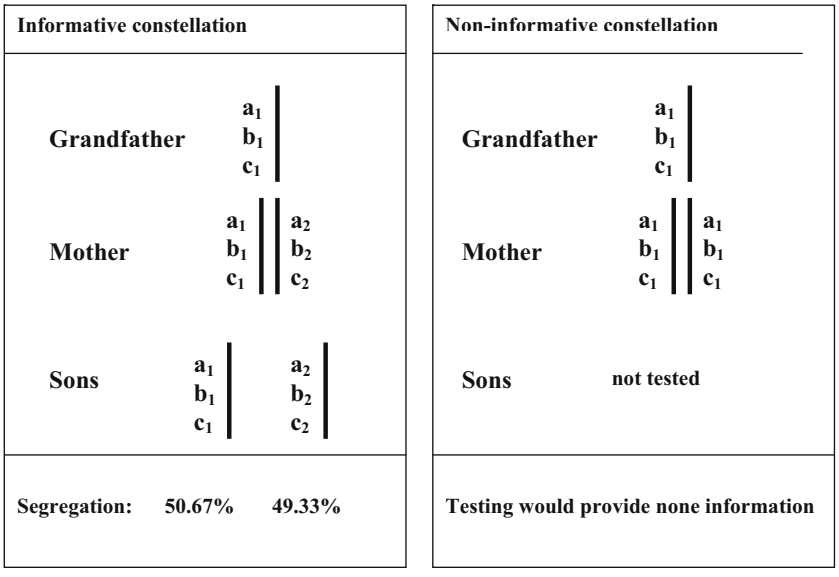

Fig. 1 Informative and noninformative testing constellations in grandfather–mother–son trios and the testing result in a segregation study including 150 informative constellation (*left*). Letters *a*, *b* and *c* indicate the loci DXS10079, DXS10074 and DXS10075, respectively and the numbers *1* and *2* symbolize different alleles

DXS10079, DXS10074 and DXS10075 appears to be very tight indeed. If the rule of thumb is true that 1 cM equals about 1,000 kb, the recombination rate between these markers should not exceed 0.3 cM. Taking into account that such a recombination rate is very low, exact figures could be established only by the investigation of several thousands of meioses. We were not able to carry out such extensive investigations and assume that our postulation that this cluster spans a distance<0.5 cM is sufficient to introduce it into forensic practice.

As was demonstrated before [3, 14], STR mutations mainly occur as single-repeat gains or losses and predominantly happen in male meioses. Furthermore, the mutation frequencies seem to be positively correlated to the length of the uninterrupted block of repeat units. Our observations in the three STRs evaluated here seem to conform to those findings. Additionally, we speculate that the AAAG-repeats DXS10079 and DXS10074 are more prone to mutations than repeats containing a mixture of purines and pyrimidines. Recently, it was shown that DXS10011 [13] and D21S1245 [20], which also contain AAAG repeats, show very high mutation rates. In the future, investigations on further markers may subject this aspect to closer examination.

When haplotypes contain several STRs, the mutation rates of the single STRs must be added to receive a cumulated haplotype mutation risk. In the case presented here we calculate for a male meiosis an average mutation risk of 1-[(1−0.018)(1−0.012)(1−0.006)]=0.036. Because we did not find any mutation in female meioses, an average value cannot be calculated here. Based on the upper confidence limit (see Table 7), the upper risk calculation for a female meiosis is: 1−[(1−0.0095)(1−0.0095)(1−0.0095)]=0.0282. The fact that in female meioses the risk of recombination may be up to 0.005 has to be included in the error risk calculation. Summarizing all the individual aspects, the calculated figures for the risk that the use of the haplotypes reported here may provide misleading results is 0.036 in male and up to 0.0332 in female gamete production. The latter results from the addition of mutation and recombination risk.

At locus DXS10079 we demonstrated the segregation of additional alleles in three families, confirming a duplication of a small chromosomal region. In those families we have checked further markers in the Xq12–13.1 region (DXS7132, HumARA, DXS981) and revealed only the normal gender-related pattern. These findings in connection with normal fertility exclude the possibility of the existence of a numerical X-chromosomal aberration or a duplication of a large ChrX region. Triallelic patterns that are most likely present in equal copy number are observed for several autosomal loci, e.g., TPOX, CSF1PO and D16S539 [7, 21]. Duplications are also known for Y-chromosomal markers (www.yhrd.org). The genetic basis for anomalous band patterns in STR profiling has been discussed by Clayton et al. [6].

Due to several rare alleles at locus DXS10075, a comparison of the observed and expected genotypes (exact test as suggested by Guo and Thomson [11]) is not practicable. However, comparison of the estimated and the expected heterozygosity gave no indication of a possible deviation in the allele distribution from the HWE.

Owing to the strong LD between all STRs investigated here, the number of haplotypes is much lower than initially expected. When LD exists, haplotype frequencies cannot be calculated from allele frequencies but have to be estimated directly from appropriate population samples. Since the corresponding databases need to be large for the estimates to be sufficiently accurate, a wider practical application of newly characterized X-chromosomal makers may be difficult. However, nationwide collaborations and the exchange of data between forensic genetics laboratories should overcome this problem. Since the amount of data regarding these topics may be very voluminous, basic research published in journals may be complemented online (http://www.chrx-str.org/) [18]. Furthermore, due to the strong LD, the use of haplotype population data as a reference may be problematic when no appropriate populations can be compared. In this regard we speculate that these problems are more serious than in common STR testing and that the situation could be similar to the ChrY population genetics. Therefore, it would be of interest to obtain haplotype population data from other German regions and other countries in the future.

References

1. Bär W, Brinkmann B, Budowle B et al (1997) 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
2. Blyth C, Still H (1983) Binomial confidence intervals. J Am Stat Asso 78:108–116
3. Brinkmann B, Klintschar M, Neuhuber F, Huhne 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
4. Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1006, 1008–1010
5. Casella G (1986) Refining binomial confidence intervals. Can J Stat 14:113–129
6. Clayton TM, Guest JL, Urquhart AJ, Gill PD (2004) A genetic basis for anomalous band patterns encountered during DNA STR profiling. J Forensic Sci 49:1207–1214
7. Crouse CA, Rogers S, Amiott E, Gibson S, Masibay A (1999) Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. J Forensic Sci 44:87–94
8. 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
9. Edelmann J, Hering S, Kuhlisch E, Szibor R (2002) Validation of the STR DXS7424 and the linkage situation on the X-chromosome. Forensic Sci Int 125:217–222
10. Edelmann J, Deichsel D, Plate I, Kaser M, Szibor R (2003) Validation of the X-chromosomal STR DXS6809. Int J Legal Med 117:241–244
11. Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372
12. Hering S, Szibor R (2000) Development of the X-linked tetrameric microsatellite marker DXS9898 for forensic purposes. J Forensic Sci 45:929–931

13. Hering S, Brundirs N, Kuhlisch E et al (2004) DXS10011: studies on structure, allele distribution in three populations and genetic linkage to further q-telomeric chromosome X markers. Int J Legal Med 118:313–319
14. Kayser M, Roewer L, Hedman M et al (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66:1580–1588
15. Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390
16. Poetsch M, Repenning A, Lignitz E, Kuhlisch E, Szibor R (2005) DXS6797 contains two STRs which can be easily haplotyped in both sexes. Int J Legal Med. DOI 10.1007/s00414-005-0003-8:1–6
17. 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
18. Szibor R, Hering S, Edelmann J (2005) A new Web-site compiling forensic chromosome X research is now online. Int J Legal Med. DOI 10.1007/s00414-005-0029-y
19. 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(6):363–369
20. Talbot CC Jr, Avramopoulos D, Gerken S, Chakravarti A, Armour JA, Matsunami N, White R, Antonarakis SE (1995) The tetranucleotide repeat polymorphism D21S1245 demonstrates hypermutability in germline and somatic cells. Hum Mol Genet 4:1193–1199
21. Zamir A, Shpitzen M, Oz C, Motro U, Meiner V, Gafny R (2002) Presentation of a three-banded allele pattern-analysis and interpretation. J Forensic Sci 47:824–826