Int J Legal Med (2009) 123:403–412
DOI 10.1007/s00414-009-0357-4

ORIGINAL ARTICLE

Evaluation of DXS9902, DXS7132, DXS6809, DXS7133, and DXS7423 in humans and chimpanzees: sequence variation, repeat structure, and nomenclature

Iva Gomes · Rui Pereira · Wolfgang R. Mayr · António Amorim · Angel Carracedo · Leonor Gusmão

Received: 10 March 2009 / Accepted: 12 May 2009 / Published online: 18 June 2009

Abstract Sequencing data obtained in this study provide information on the short tandem repeat allele structures of DXS9902, DXS7132, DXS6809, DXS7133, and DXS7423. Data were obtained from the three human major population groups, namely Africans, Caucasians, and Asians as well as from chimpanzees (*Pan troglodytes*). DXS7133 was found to be the most stable locus and DXS6809 seemed to have evolved from a simple array of CTAT units but currently reveals a highly complex and compound structure within and between humans and chimpanzees. DXS9902 results support a TAGA allele nomenclature, which increases in one repeat unit previously reported allele distributions at this locus. For DXS7132, human/chimpanzee comparisons performed in this study provided important evidence that the CTAT allele structure should be considered for allele nomenclature purposes. Also, possible population-specific intermediate type alleles (with Native American origin) were detected at this locus that could be useful for ethnic group differentiation. DXS7423 results revealed two different sequence structures and one of these structures seems to be restricted to a single allele class in just one population group (Africans).

Keywords X-STRs · Sequence structure · Nomenclature · Chimpanzee

I. Gomes · R. Pereira · A. Amorim · L. Gusmão (✉)
IPATIMUP, Institute of Molecular
Pathology and Immunology of the University of Porto,
Rua Dr. Roberto Frias, s/n,
4200-465 Porto, Portugal
e-mail: lgusmao@ipatimup.pt

I. Gomes · R. Pereira · A. Carracedo
Institute of Legal Medicine,
University of Santiago de Compostela,
15782 Santiago de Compostela, Spain

W. R. Mayr
Division of Blood Group Serology, Medical University of Vienna,
1090 Vienna, Austria

A. Amorim
Faculty of Sciences, University of Porto,
4050 Porto, Portugal

A. Carracedo
Genomics Medicine Group, CIBERER,
University of Santiago de Compostela,
15782 Santiago de Compostela, Spain

Introduction

X-chromosome short tandem repeats (X-STRs) are currently the object of study in many applications in population and forensic genetics (e.g., [1, 3–6, 8–11, 15–17, 20–28]). The main use of X-STRs is centered in scenarios of kinship analysis, particularly when the putative father is not available for testing and two possible half sisters are under investigation. In these cases, sisterhood can be excluded by X-STRs solely if they do not share the same paternal X chromosome alleles, which cannot be achieved with autosomal genetic markers [24]. Although numerous X-STRs have been evaluated individually or in multiplexes, studies have essentially focused on the genetic characterization of human population groups as well as on the forensic efficiency of X-chromosomal microsatellites (e.g., [1, 8–10, 16, 17, 20, 25, 28]). Only some of these studies have resulted in the analysis of allele and locus sequence structure but were essentially focused on the analysis of a single population group (e.g., [3–6, 11, 21, 26, 27]). The importance of a common nomenclature aims toward improved communication, data exchange, and comparison

among laboratories. During several studies, some unusual results called attention to the probable use of different nomenclatures by different laboratories for the same X-STR marker (e.g., [9, 11, 16, 20, 22, 23]). As additional research to help provide more accurate and secure genotyping results, the selection of X-chromosomal markers in the current and previous work [11] was based on those included in the X-linked decaplex recently developed by the Spanish–Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG) [15, 16]. This multiplex system has been applied in the genetic characterization of several populations [10, 16, 28] and is still a tool for current and future projects. Consequently, as an extension of preceding work [11], in this study, we aimed to investigate the sequence structural variation of the X-linked loci DXS9902, DXS7132, DXS6809, DXS7133, and DXS7423 in different human population groups, as well as in chimpanzees. Comparisons of *Homo sapiens* and *Pan troglodytes* genomes allowed a more precise analysis of the sequence structure of the loci studied in this work, as well as inferences on possible evolving mutational mechanisms.

Materials and methods

Samples

A total of 236 male samples from three different population groups were sequenced for five X-chromosomal loci. The groups studied in this work included samples from the three major human population groups: Africans (Uganda, Angola, and Mozambique), Caucasians (northern Portugal), and Asians (Macau). In addition, samples from nine to ten male chimpanzees (*P. troglodytes*) were sequenced for the same X-chromosome loci.

Allele selection

For the Caucasian group, alleles selected for sequencing were based on previously published genotyping data from northern Portugal [16, 20]. For the African group, sequenced alleles for DXS7132, DXS6809, and DXS7423 were selected from a genetic diversity study on three African groups [8] and for DXS9902 and DXS7133 from a Ugandan group genetic characterization [10]. Genotyping information for the Asian group was selected from a population from Macau (L. Gusmão, personal unpublished data). All observed allele classes were sequenced. In addition, the most common alleles were sequenced several times in each population group. The reference DNA samples, 9947A (female) and 9948 (male; from Promega commercial kits, Madison, WI, USA) were also sequenced for the five loci.

Primers, PCR, and sequencing conditions

Primer sequences used for amplification and sequencing were according to Gusmão et al. [16]. For DXS7132, a second forward sequence primer (F2; Fig. 1) was designed further away from the repeat region that includes the first forward primer (F1; Fig. 1) using the PRIMER3 software (http://frodo.wi.mit.edu). Polymerase chain reaction (PCR), thermocycling, and sequencing reaction conditions were the same as in Gomes et al. [11] except for annealing temperatures (T_m) used for amplification of chimpanzee samples which was 58°C for all loci. DXS7423 was the only locus that presented unspecific PCR products in chimpanzee using the human primer set. As these products would most likely interfere with sequencing reactions, DNA samples were re-amplified but using an increased T_m (62–64°C).

Alignment of human and chimpanzee sequences was performed using the sequence similarity search genome browsers BLAT (www.genome.ucsc.edu) and BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Human sequences were selected from the National Center for Biotechnology Information (NCBI) DNA database (http://www.ncbi.nlm.nih.gov/) with the following accession numbers: G27261 for DXS9902, Z83841 for DXS7132, AL365400 for DXS6809, NT_011651.16 for DXS7133, and Y15994 for DXS7423.

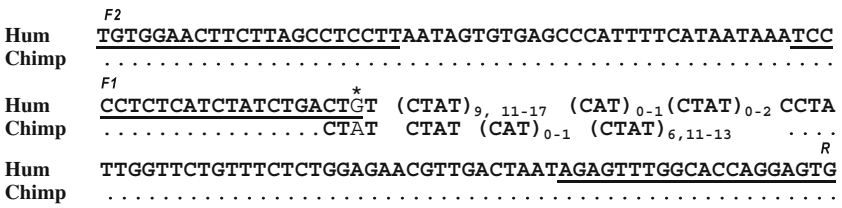

Fig. 1 DXS7132 sequence structure alignment for human and chimpanzee genomes obtained using BLAST (human sequence Z83841 and chimpanzee sequence NW_001251919). *Underlined* sequences represent primer binding sites: *F1* initial primer sequence, *F2* newly designed forward primer, *R* reverse primer. *Asterisk* *Nonbold letters* in F1 sequence (G in human and A in chimpanzee) represent the nucleotide mismatch at the first 3′ base of the initial forward primer hybridization site used for human and chimpanzee amplification

Results

Results were analyzed by comparing the sequence structure of X-STR alleles in the three human major population groups (Africans, Caucasians, and Asians) and between humans and chimpanzees. Sequencing results as well as chimpanzee and human genome alignments are displayed in Tables 1, 2, 3, 4, 5. Allele sequence structures were also obtained for the reference DNA samples 9947A and 9948 and are presented in Table 6. The variation studied within human populations and between humans and chimpanzees was restricted to the main repetitive region and to the flanking regions of the varying stretch. Results are discussed individually for each locus.

DXS6809

The polymorphic DXS6809 locus has been described as a complex repeat composed of three different motifs $(CTAT)_n$-$(ATCT)_3$-N_9-$(TATC)_n$-$(ATCT)_n$-N_{10}-$(ATCT)_n$ [5] and presents an allele distribution in diversified populations between 27 and 39 repeats (e.g., [5, 6, 8–10, 16, 20, 28]). In the present work, this complex and compound repeat sequence structure was also observed in the human samples from all three groups (Table 1). Intra-allelic variation was found in the majority of the allele classes with samples exhibiting equal fragment sizes and sequence differences at the repeat core (also observed by Edelmann et al. [5]).

Intermediate alleles revealed a thymine insertion in the repeat region (30.1, 31.1, 32.1, and 33.1) and were only observed in samples from the African population group, except for allele 36.1 (Table 1). In this allele, either a C insertion or an ATT deletion has reduced the ten-base nucleotide stretch attatctatc (that is not considered for allele nomenclature by the described structure [5]) to three bases and merged with the preceding ATCT tract increasing the number of repeats at this motif type (allele 36.1; Table 1). The position of the extra nucleotide base confers a different structure for these nonconsensus repeats (36.1): $(CTAT)_{11}$-$(ATCT)_3$-N_9-$(TATC)_4$-$(ATCT)_6$-atc-$(ATCT)_{14}$ when compared with the other intermediate alleles sequenced.

As in humans, DXS6809 also revealed to be polymorphic in chimpanzees. A highly heterogeneous sequence composition was detected as nine sequenced chromosomes showed five different structures (Table 7).

Comparison of chimpanzee and human sequences was relevant to the understanding of DXS6809 STR structure. What seems to have been a simple repeat composed of CTAT units has evolved to a more complex and compound repeat arrangement interrupted by several mutations that confer to both primates a complex repeat core at this locus and suggests a high instability of this region. We believe that the allele structure of DXS6809 was initially composed of CTAT varying units that were subsequently disrupted by several complex mutational events and have disguised the CTAT core in humans (Table 7). Considering this sequence organization, it was possible to divide and group the sequences obtained in this study into four different structures as shown in Table 7. When aligning the human and chimpanzee sequences deposited at NCBI (www.ncbi.nlm.nih.gov), two nucleotide mismatches were observed at bases 5 and 22 from the 5′ end of the forward primer (Table 1) but that did not inhibit primer annealing and subsequent amplification of the DXS6809 fragments when using the same human primers and PCR conditions. Two further single base variations between species were observed and are presented in Table 1.

DXS7133

The tetranucleotide locus DXS7133 has been reported to be a simple repeat of ATAG units and with an allele range of six to 14 copies (e.g., [1, 4, 6, 10, 16, 17, 20, 21, 25, 28]). This locus has been studied in many different populations and in particular has been one of the most frequently characterized X-chromosomal markers in Asian populations (e.g., [1, 17, 21]). Although DXS7133 is broadly used in population and forensic genetics, when compared with other commonly used X-STRs, it stands out by presenting low gene diversity values (e.g., [1, 16, 17, 25]).

In both *H. sapiens* and *P. troglodytes*, a simple ATAG varying unit was detected with no mutations found in the flanking upstream and downstream regions (Table 2). Variation was restricted to the different number of ATAG units. When aligning both genomes, no differences were seen outside the repeat structure region which supports a simple and stable nature (Table 2).

DXS9902

DXS9902 was first reported in a German population sample by Edelmann et al. [6] as a simple GATA repeat and described as presenting an allele range of seven to 16 repeats in several population studies (e.g., [6, 10, 16, 28]).

To our knowledge, intermediate alleles at this locus were first reported in a collaborative study by the GEP-ISFG characterizing 15 Iberian and Latin American populations [16]. Sequence analysis of the three groups involved in this work also confirmed the insertion of an extra adenine at a 9-mer poly-A sequence upstream of the repeat region of DXS9902 (Table 3). All sequenced alleles of this class showed the same single-base insertion (alleles 11.1, 12.1, and 13.1) and were detected only in Caucasian samples, being absent in the Asian group as well as the 393 chromosomes studied in an African sample from the Ugandan group used in this work [10]. As reported [16]

Table 1 STR and allele sequence structures for DXS6809

STR structure in humans and chimpanzees

Hum: *TCCATcTTTCTCTGAACCTTCCt*agctcaggaatactgagggcatgactagattatgtaggaatttgg*g*
Chim: .....t.............c.................a.....................a

Hum: (CTAT)$_{2,6-11}$(at)$_{0-1}$[at(CTAT)$_3$]$_{1-2}$ cat(CTAT)tat(CTAT)$_{0,2,3}$(cat)$_{0-1}$(CTAT)$_{3,5,6}$[tat(CTAT)]$_{0-1}$ cat (CTAT)$_{9-16}$
Chim: (CTAT)$_{4,11-13}$ cat(CTAT)$_{3,10-12}$ (cat)$_{0-1}$(CTAT)$_{0,1,3}$(cat)$_{0-1}$(CTAT)$_{0,3,9}$ at (CTAT)$_{2,8-12}$ [cat(CTAT)$_3$at]$_{0-1}$(CTAT)$_{0,6,9}$ [cat (CTAT)$_{10}$]$_{0-1}$

Hum: cctctatctctt*CCTCACATCAGCCTAAAGCA*
Chim: ....

Al.	Allele structure: $(CTAT)_n$ $(ATCT)_3$ N_9 $(TATC)_n$ $(ATCT)_n$ N_{10} $(ATCT)_n$	n	Sequenced in
27	Pr1 (22bp) 47bp **(CTAT)$_2$ATAT (CTAT)$_3$ (ATCT)$_3$ N$_9$(TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_9$** 14bp (20bp) Pr2	1	AFR
27	Pr1 (22bp) 47bp **(CTAT)$_7$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_3$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	2	CAU
28	Pr1 (22bp) 47bp **(CTAT)$_2$ ATAT (CTAT)$_3$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{10}$** 14bp (20bp) Pr2	1	AFR
28	Pr1 (22bp) 47bp **(CTAT)$_7$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_3$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	2	CAU
29	Pr1 (22bp) 47bp **(CTAT)$_2$ ATAT (CTAT)$_3$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	1	AFR
29	Pr1 (22bp) 47bp **(CTAT)$_6$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_6$ N$_{10}$ (ATCT)$_{10}$** 14bp (20bp) Pr2	1	AFR
29	Pr1 (22bp) 47bp **(CTAT)$_7$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_3$ N$_{10}$ (ATCT)$_{13}$** 14bp (20bp) Pr2	1	CAU
29	Pr1 (22bp) 47bp **(CTAT)$_8$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_9$** 14bp (20bp) Pr2	1	CAU
30	Pr1 (22bp) 47bp **(CTAT)$_2$ ATAT (CTAT)$_3$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	AFR
30	Pr1 (22bp) 47bp **(CTAT)$_7$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_3$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	AFR
30	Pr1 (22bp) 47bp **(CTAT)$_8$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{10}$** 14bp (20bp) Pr2	1	CAU
30	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{10}$** 14bp (20bp) Pr2	1	CAU
30.1	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_6$ T N$_{10}$ (ATCT)$_{12}$** 14bp (20bp) Pr2	2	AFR
31	Pr1 (22bp) 47bp **(CTAT)$_7$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	AFR
31	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_4$ N$_{10}$ (ATCT)$_{10}$** 14bp (20bp) Pr2	1	AFR
31	Pr1 (22bp) 47bp **(CTAT)$_8$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	1	CAU
31	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	4	AS, CAU
31.1	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_6$ T N$_{10}$ (ATCT)$_{13}$** 14bp (20bp) Pr2	1	AFR
32	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_4$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	1	AFR
32	Pr1 (22bp) 47bp **(CTAT)$_8$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{13}$** 14bp (20bp) Pr2	2	AFR, CAU
32	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	3	AS, CAU
32	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	1	AS
32.1	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_6$ T N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	AFR
33	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N9 (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	AFR
33	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{13}$** 14bp (20bp) Pr2	3	AS, CAU
33	Pr1 (22bp) 47bp **(CTAT)$_{11}$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{11}$** 14bp (20bp) Pr2	1	AS
33.1	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_6$ T N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	4	AFR
33.1	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_6$ T N$_{10}$ (ATCT)$_{15}$** 14bp (20bp) Pr2	1	AFR
34	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)14** 14bp (20bp) Pr2	6	AFR, CAU, AS
35	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)15** 14bp (20bp) Pr2	2	AFR, AS
35	Pr1 (22bp) 47bp **(CTAT)$_8$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_6$ N$_{10}$ (ATCT)$_{15}$** 14bp (20bp) Pr2	1	AFR
35	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{14}$** 14bp (20bp) Pr2	1	CAU
36	Pr1 (22bp) 47bp **(CTAT)$_9$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{16}$** 14bp (20bp) Pr2	3	AFR, CAU
36.1	Pr1 (22bp) 47bp **(CTAT)$_{11}$ (ATCT)$_3$ N$_9$ (TATC)$_4$ (ATCT)$_6$ ATC(ATCT)$_{14}$** 14bp (20bp) Pr2	2	AFR
37	Pr1 (22bp) 47bp **(CTAT)$_{10}$ (ATCT)$_3$ N$_9$ (TATC)$_3$ (ATCT)$_5$ N$_{10}$ (ATCT)$_{16}$** 14bp (20bp) Pr2	1	CAU

Allele structure variation obtained in 58 human and nine chimpanzee chromosomes according to the adopted allele nomenclature [5]. In STR sequence structure, italic uppercase bold letters represent primer hybridization sites

Hum human sequence, *Chim* chimpanzee sequence, *AFR* Africans, *AS* Asians, *CAU* Caucasians, N_9- atcatctat, N_{10}- attatctatc

Table 2 STR and allele sequence structures for DXS7133

STR structure in humans and chimpanzees

Hum: ***CACTTCCAAAAGGGGAAAAA***atattttaggtgtagcttccttagatggcattca**(ATAG)**$_{6-14}$ ataaaaataagcatgaacaccgtta
Chim:.....**(ATAG)**$_{9-12}$ .....

Hum: aaaataaagtaggagacttctgattcttggaagatgtagtagatgtatttttatctat***TCCTCCCACCAAGTACAAGT***
Chim:.....

Allele	n	Allele structure: $(ATAG)_{6-14}$	Sequenced in
6	1	Pr1 (20bp) 34 bp **(ATAG)**$_6$ 83bp (20bp) Pr2	AFR
9	9	Pr1 (20bp) 34 bp **(ATAG)**$_9$ 83bp (20bp) Pr2	AFR, AS, CAU
10	9	Pr1 (20bp) 34 bp **(ATAG)**$_{10}$ 83bp (20bp) Pr2	AFR, AS, CAU
11	7	Pr1 (20bp) 34 bp **(ATAG)**$_{11}$ 83bp (20bp) Pr2	AFR, AS, CAU
12	4	Pr1 (20bp) 34 bp **(ATAG)**$_{12}$ 83bp (20bp) Pr2	AFR, CAU
13	2	Pr1 (20bp) 34 bp **(ATAG)**$_{13}$ 83bp (20bp) Pr2	AFR
14	1	Pr1 (20bp) 34 bp **(ATAG)**$_{14}$ 83bp (20bp) Pr2	AFR

Allele structure variation obtained in 33 human and ten chimpanzee chromosomes according to the adopted allele nomenclature [6]. In STR sequence structure, italic uppercase bold letters represent primer hybridization sites

Hum human sequence, *Chim* chimpanzee sequence, *AFR* Africans, *AS* Asians, *CAU* Caucasians

and based on sequencing results of the three major human population groups, a TAGA repeat motif should be considered, corresponding to an increase in one extra unit when compared with the described GATA nomenclature [6] (Table 3). No other sequence composition differences were observed in or among Africans, Asians, or Caucasians.

Repeat variation was also observed among chimpanzees. The TAGA unit is present, as in humans, but interrupted by an additional repeat motif: a long varying tract of GA

Table 3 STR and allele sequence structures for DXS9902

STR structure in humans and chimpanzees

Hum: ***CTGGGTGAAGAGAAGCAGGA***attttttgtactgttcttgtaacttttctgtcggctgaaattatttccaaacattttaac
Chim: .....-.....

Hum: aagaaaaaaaaa(a)gat **(TAGA)**$_{9-14}$ atcactggtgatcccatac***TCCTGATAT***
Chim: ..... - ...**(TAGA)**$_{3-4}$ **(GA)**$_{11-14,16,18,20}$**(TAGA)**$_{3-6,9}$ **ATGA** .....

Hum: ***GAATGTGTATTGCT***
Chim: .....

*Allele	**New allele	n	Allele structure: $(TAGA)_{9-14}$	Sequenced in
8	**9**	4	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_9$ 19bp (23bp)Pr2	AFR, CAU
9	**10**	9	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_{10}$ 19bp (23bp)Pr2	AFR, CAU
10	**11**	11	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_{11}$ 19bp (23bp)Pr2	AFR, AS, CAU
10.1	**11.1**	1	Pr1(20bp) 64bp **A**aaaaaaaaagat **(TAGA)**$_{11}$ 19bp (23bp)Pr2	CAU
11	**12**	15	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_{12}$ 19bp (23bp)Pr2	AFR, AS, CAU
11.1	**12.1**	3	Pr1(20bp) 64bp **A**aaaaaaaaagat **(TAGA)**$_{12}$ 19bp (23bp)Pr2	CAU
12	**13**	9	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_{13}$ 19bp (23bp)Pr2	AFR, AS, CAU
12.1	**13.1**	1	Pr1(20bp) 64bp **A**aaaaaaaaagat **(TAGA)**$_{13}$ 19bp (23bp)Pr2	CAU
13	**14**	2	Pr1(20bp) 64bp **aaaaaaaaa**gat **(TAGA)**$_{14}$ 19bp (23bp)Pr2	AFR

Allele structure variation obtained in 55 human and ten chimpanzee chromosomes according to the recently described new allele nomenclature [16]. In STR sequence structure, italic uppercase bold letters represent primer hybridization sites. In allele structure, lowercase letters represent the upstream flanking sequence from the main repeat. Underlined uppercase A represents the extra adenine observed in the poly-A sequence tract

Hum human sequence, *Chim* chimpanzee sequence, *CAU* Caucasians, *AFR* Africans, *AS* Asians

[a] Number of repeats for the primary established nomenclature [6]

[b] Number of repeats for the recently described new allele nomenclature [16]

Table 4 STR and allele sequence structures for DXS7132

STR structure in humans and chimpanzees

Hum: ***TCCCCTCTCATCTATCTGACTgT*** **(CTAT)$_{9,11-17}$ (CAT)$_{0-1}$ (CTAT)$_{0-2}$** cctattggttctgtttctctggagaacgtt
Chim: .....**CTa*T*** **CTAT (CAT)$_{0-1}$ (CTAT)$_{6,11-13}$**

Hum: gactaat***AGAGTTTGGCACCAGGAGTG***
Chim: .....

Allele	n	Allele structure: (CTAT)$_{9,\ 11-17}$ (CAT)$_{0-1}$ (CTAT)$_{0-2}$	Sequenced in
9	**1**	Pr1(22bp) 1bp **(CTAT)$_9$** 37bp (20bp)Pr2	AFR
11	**5**	Pr1(22bp) 1bp **(CTAT)$_{11}$** 37bp (20bp)Pr2	AFR, AS, CAU
12	**7**	Pr1(22bp) 1bp **(CTAT)$_{12}$** 37bp (20bp)Pr2	AFR, AS, CAU
13	**6**	Pr1(22bp) 1bp **(CTAT)$_{13}$** 37bp (20bp)Pr2	AFR, AS, CAU
14	**8**	Pr1(22bp) 1bp **(CTAT)$_{14}$** 37bp (20bp)Pr2	AFR, AS, CAU
15	**6**	Pr1(22bp) 1bp **(CTAT)$_{15}$** 37bp (20bp)Pr2	AFR, AS, CAU
16	**8**	Pr1(22bp) 1bp **(CTAT)$_{16}$** 37bp (20bp)Pr2	AFR, AS, CAU
16.3	**1**	Pr1(22bp) 1bp **(CTAT)$_{14}$ CAT (CTAT)$_2$** 37bp (20bp)Pr2	*CAU
17	**4**	Pr1(22bp) 1bp **(CTAT)$_{17}$** 37bp (20bp)Pr2	AFR, AS, CAU
17.3	**1**	Pr1(22bp) 1bp **(CTAT)$_{15}$ CAT (CTAT)$_2$** 37bp (20bp)Pr2	*CAU
18.3	**1**	Pr1(22bp) 1bp **(CTAT)$_{16}$ CAT (CTAT)$_2$** 37bp (20bp)Pr2	*CAU

Allele structure variation obtained in 50 human and nine chimpanzee chromosomes according to the adopted allele nomenclature [6]. In STR sequence structure, italic uppercase bold letters represent primer hybridization sites

Hum human sequence, *Chim* chimpanzee sequence, *AFR* Africans, *AS* Asians, *CAU* Caucasians

[a] Three individuals sequenced for the observed intermediate alleles 16.3, 17.3, and 18.3 living in Northern Portugal but born in Brazil.

dinucleotides that is followed by an ATGA (Table 3). The 9-mer poly-A tract is also found in *P. troglodytes*; however, no additional adenines were observed in the ten chromosomes studied in contrast to that observed in *H. sapiens*. This fact as well as the observation of intermediate alleles only in Europeans supports a more recent emergence of the extra base insertion (ten poly-A sequence) that induces nonconsensus alleles. However, it is well known that during replication, these long nucleotide tracts are slippage-prone sequences and therefore, further analyses are needed to support this hypothesis since we cannot rule out the possibility of variation still occurring at this tract in other

Table 5 STR and allele sequence structures for DXS7423

STR structure in humans and chimpanzees

Hum: ***GTCTTCCTGTCATCTCCCAAC***ctgccctttatcacccagatttcctcccca**(TCCA)$_3$**[tctgtcct]$_{0-1}$ **(TCCA)$_{7,\ 9-14}$**
Chim: .....**(TCCA)$_3$** tctgtcct **(TCCA)$_{8-10}$**

Hum: catactcattcatttatttgtgtgacaaatatcttttgggcacctac***TATGTGCCAGGCGCTAAGCTA***
Chim: .....

Allele	n	Allele structure: (TCCA)$_3$ [tctgtcct]$_{0-1}$ (TCCA)$_{7,\ 9-14}$	Sequenced in
8	**4**	Pr1 (21bp) 30 bp **(TCCA)$_{10}$** 47bp (21bp) Pr2	AFR
10	**1**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_7$** 47bp (21bp) Pr2	CAU
12	**3**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_9$** 47bp (21bp) Pr2	AFR
13	**5**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_{10}$** 47bp (21bp) Pr2	AFR, CAU
14	**12**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_{11}$** 47bp (21bp) Pr2	AFR, AS, CAU
15	**6**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_{12}$** 47bp (21bp) Pr2	AFR, AS, CAU
16	**6**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_{13}$** 47bp (21bp) Pr2	AFR, CAU
17	**3**	Pr1 (21bp) 30 bp **(TCCA)$_3$** tctgtcct **(TCCA)$_{14}$** 47bp (21bp) Pr2	AFR, CAU

Allele structure variation obtained in 40 human and nine chimpanzee chromosomes according to the adopted allele nomenclature [26]. In STR sequence structure, italic uppercase bold letters represent primer hybridization sites. The lowercase eight-base nucleotides (tctgtcct) that interrupt the TCCA tandem are not considered for allele nomenclature

Hum human sequence, *Chim* chimpanzee sequence, *AFR* Africans, *AS* Asians, *CAU* Caucasians

Table 6 Sequencing results for reference DNA samples 9948 (male) and 9947A (female)

Locus	Allele	Reference DNA 9948	Allele	Reference DNA 9947A[a]
DXS9902	13	$(TAGA)_{13}$	12	$(TAGA)_{12}$
DXS7132	13	$(CTAT)_{13}$	12	$(CTAT)_{12}$
DXS6809	31	$(CTAT)_8$-$(ATCT)_3$ N_9 $(TATC)_4$ $(ATCT)_5$ $N_{10}(ATCT)_{11}$	–	–
DXS7133	11	$(ATAG)_{11}$	–	–
DXS7423	14	$(TCCA)_3$ tctgtcct $(TCCA)_{11}$	–	–

In DXS6809, N_9- atcatctat and N_{10}-attatctatc are not considered for allele nomenclature

[a] For reference sample 9947A, sequencing results were obtained only for the homozygous state

populations. Alignments of both genomes showed an extra thymine in another poly-T tract in humans also located upstream of the repeat region (Table 3). No additional sequence variations were observed at DXS9902 within or between the two species.

DXS7132

DXS7132 has been genetically characterized in different populations by displaying an allele distribution of nine to 18 copies (e.g., [4, 8–10, 16, 17, 20, 21, 25, 28]). This locus was first described by Edelmann et al. [6] as a simple tandem of CTAT repeats. In a later publication, the same authors presented DXS7132 as a repeat structure composed of TCTA units [4]. Irrespectively of the repeat motif considered, the number of repeats does not show alteration. However, when analyzing results for human and chimpanzee genome alignments, one base mismatch was observed between both primates (Table 4; Fig. 1). The base mismatch (G in *H. sapiens* and A in *P. troglodytes*) is located at the first base of the 3′ end of the initial forward primer used in this work (F1; Fig. 1) which is also part of the immediate flanking sequence of the repeat. This difference is responsible for two possible descriptions of allele structures in both species: TCTA in humans vs. CTAT in chimpanzees (Fig. 1). This observation led to an investigation on what the STR structure at DXS7132 should be. To confirm this sequence divergence between species, a new forward primer was designed (F2; Fig. 1) which includes the initial forward F1 sequence (F1; Fig. 1). Screening was performed by sequencing the new fragment in 24 out of the 50 human male individuals belonging to the three population groups, as well as for the nine chimpanzee chromosomes. Results confirmed the single nucleotide variance between species at this locus: G in *H. sapiens* and A in *P. troglodytes* (Fig. 1) as well as their conserved status among groups. The CTAT unit initially described seems to be the most accurate since this motif most probably evolved from a single tandem array of CTAT units that was disrupted in humans by an A→G transition at the first repeat (Fig. 1). In light of the human/chimpanzee comparisons, results support the original reported sequence CTAT [6] as the most accurate allele structure at DXS7132.

Nonconsensus alleles (alleles 16.3, 17.3, and 18.3) were also sequenced and revealed that a single-base deletion interrupts the perfect CTAT repeat allelic motif as follows $(CTAT)_n$-CAT-$(CTAT)_2$ (Table 4). This type of intermediate allele has been reported, to our knowledge, only in South American populations [16]. Additionally, when looking into the origin of our samples, we found that these alleles belong to individuals reported as living in Portugal, but born in Brazil. These findings suggest population specific-

Table 7 DXS6809 STR sequence structures in chimpanzee and human

Chimpanzee STR sequence structures

$(CTAT)_{11–13}$ CAT $(CTAT)_3$ AT $(CTAT)_{8,12}$

$(CTAT)_4$ CAT $(CTAT)_{11,12}$ CAT $(CTAT)_3$ AT $(CTAT)_{10,12}$

$(CTAT)_4$ CAT $(CTAT)_3$ AT $(CTAT)_9$ CAT $(CTAT)_3$ AT $(CTAT)_9$

$(CTAT)_4$ CAT $(CTAT)_{10}$ CAT (CTAT) CAT $(CTAT)_3$ AT $(CTAT)_{11}$

$(CTAT)_4$ CAT $(CTAT)_{11}$ CAT (CTAT) CAT $(CTAT)_9$ AT $(CTAT)_2$ CAT $(CTAT)_3$ AT $(CTAT)_6$ CAT $(CTAT)_{10}$

Human STR sequence structures

$(CTAT)_{10,11}$ AT $(CTAT)_3$ CAT (CTAT) TAT $(CTAT)_3$ CAT $(CTAT)_{5,6}$ CAT $(CTAT)_{14,16}$

$(CTAT)_{9,10}$ AT $(CTAT)_3$ CAT (CTAT) TAT $(CTAT)_6$ TAT (CTAT) CAT $(CTAT)_{12–15}$

$(CTAT)_{6–11}$ AT $(CTAT)_3$ CAT (CTAT) TAT $(CTAT)_{2–3}$ CAT $(CTAT)_{3–6}$ TAT (CTAT) CAT $(CTAT)_{9–16}$

$(CTAT)_2$ ATAT $(CTAT)_3$ AT $(CTAT)_3$ CAT (CTAT) TAT $(CTAT)_3$ CAT$(CTAT)_5$ TAT (CTAT) CAT $(CTAT)_{9–12}$

ity of these alleles, namely a possible Native American origin. Indeed, this is one of the South American population groups in addition to Europeans and Africans, and no intermediate alleles of this nature have been detected in the latter two groups.

Sequence data for chimpanzee X chromosomes also revealed that these primates are polymorphic at DXS7132 (Table 4). No variation outside the repeat area was detected among individuals of this species. Although with one mismatch at the 3′ end forward primer hybridization site, amplification of the chimpanzee chromosomes was achieved using the same human primer set and an annealing temperature of 58°C. Furthermore, no variation was found at DXS7132 among human groups or species outside the repetitive sequence (Table 4).

DXS7423

This tetranucleotide presents an allele distribution of eight to 18 repeats and has been one of the most extensively analyzed X-STRs (e.g., [1, 4, 6, 8–10, 16, 17, 20, 21, 23, 26–28]). It was initially described by Edelmann et al. [6] as a simple TCCA repeat. A different structure was reported in a Spanish population sample by Zarrabeitia and coauthors [26] who detected three extra nonvariable repeats with the same motif located upstream to the marker: $(TCCA)_3$-tctgtcct-$(TCCA)_n$. Both groups addressed these nomenclature discrepancies [23] to avoid possible confusion among the scientific community and the previously published nomenclature was revised by adding three extra repeats [6, 23].

In the present study, an allele with a fragment size corresponding to allele 8 was sequenced in four out of the 40 samples analyzed for this locus (Table 5). So far, this allele has only been detected in populations of African origin [8, 10] or with African admixture [16]. Such an observation indicates possible African specificity. This allele class presented a different structure composition where the TCCA was not interrupted by the eight-base nucleotide stretch which is present in the remaining allele classes but instead contains ten perfect TCCA repeats (Table 5). This finding could possibly suggest an ancestral state of the eight repeat alleles: a smaller allele class with a simpler structure and only observed in the African group. The STR could have evolved from a single tandem of TCCA units that was interrupted by a mutation giving rise to the observed common structure. However, chimpanzee results do not favor this argument since the eight bases are present in all sequenced samples just as in humans (Table 5) showing that the mutation responsible for the tract interruption emerged before the human/chimpanzee split. The detection of a different structure for allele 8 suggests that variation is still occurring at the eight nucleotide sequence tracts which are not considered for allele nomenclature (Table 5).

P. troglodytes also revealed to be polymorphic for DXS7423 and no sequence composition differences were seen when compared with *H. sapiens* (Table 5). No further structural variations were observed at this locus, within or between species, outside the repeat or flanking regions.

Discussion

The increased attention given to X-STRs in forensics and population genetics strengthens the importance of a well-established nomenclature. As with autosomal and Y-chromosomal markers (e.g., [2, 7, 12–14, 18, 19]), X-chromosome polymorphisms discrepancies have been observed in some of the published data (e.g., [9, 11, 16, 20, 22, 23]). For this reason, the present study intends to be an extension of the previous work [11] on the evaluation of the sequence structure variation in humans and chimpanzees that includes five additional X-linked markers. With both studies, we have covered the ten X-chromosomal STRs (DXS8378, DXS9902, DXS7132, DXS9898, DXS6809, DXS6789, DXS7133, GATA172D05, GATA31E08, and DXS7423) of the decaplex system developed by the GEP-ISFG [15, 16] which has been applied to the genetic profiling of various populations [10, 16, 28]. As this multiplex system is in current use and is under an ongoing collaborative exercise by the same working group

Table 8 Suggested STR consensus structures taking into consideration the locus variation observed in humans and chimpanzees and ISFG nomenclature guidelines [2, 7, 13]

X-STR	STR structure	Allele nomenclature	References
DXS9902	$(TAGA)_n$	$(TAGA)_n$	[16]
DXS7132	$(CTAT)_n$ $(CAT)_{0–1}$ $(CTAT)_{0–2}$	$(CTAT)_n$ $(CAT)_{0–1}$ $(CTAT)_{0–2}$	[6]
DXS6809	$(CTAT)_n$ $(AT)_{0–1}$ $[AT(CTAT)_3]_{1–2}$ CAT (CTAT) TAT $(CTAT)_{0,2,3}$ $(CAT)_{0–1}$ $(CTAT)_{3,5,6}$ $[TAT(CTAT)]_{0–1}$ CAT $(CTAT)_n$	$(CTAT)_n$ $(ATCT)_3$ atcatctat $(TATC)_n$ $(ATCT)_n$ attatctatc $(ATCT)_n$	[5]
DXS7133	$(ATAG)_n$	$(ATAG)_n$	[6]
DXS7423	$(TCCA)_3$ $(tctgtcct)_{0–1}$ $(TCCA)_n$	$(TCCA)_3$ $(tctgtcct)_{0–1}$ $(TCCA)_n$	[26]

(http://www.gep-isfg.org), a well-established nomenclature for these markers is mandatory.

Conclusions regarding nomenclature implications are discussed individually for each locus taking into consideration the ISFG guidelines [2, 7, 13]. Using the same approach as in Gomes et al. [11], the existing nomenclature recommendations on autosomal and Y STRs [2, 7, 13] were extended to X-STRs. A summary of STR structures and allele nomenclatures are presented in Table 8.

DXS6809

A highly complex compound structure was observed in humans and chimpanzees. This complexity does not suggest a straightforward allele nomenclature at this locus. However, alignments of both genomes would support a nomenclature change, with the gain of extra repeat units, based upon a single CTAT array (Tables 1 and 7) that underwent several mutations. Nevertheless, as this marker has been one of the most used X-STR markers and as recommended by the ISFG [2, 7, 13], a nomenclature change would generate unnecessary confusion (Table 8).

DXS7133

The STR structure stability observed both in *H. sapiens* and in *P. troglodytes* suggests unambiguous allele nomenclature based upon an ATAG repeat motif as first reported by Edelmann et al. [6] (Table 8).

DXS9902

Sequencing results obtained for all major human population groups are in agreement with the allele nomenclature recently reported by the GEP-ISFG [16]. A TAGA repeat motif increases in one unit the previous DXS9902 allele nomenclature and therefore should be the one considered as an alternative to the described GATA sequence structure [6] according to the ISFG guidelines [2, 7, 13]. Results of our study based on the chimpanzee and human genome analyses also support the TAGA varying unit as the most accurate allele nomenclature (Table 8).

DXS7132

Human and chimpanzee genome sequence alignments provided the detection of a single nucleotide base mismatch among the species (G in *H. sapiens* and A in *P. troglodytes*) located at the beginning of the repeat region (Fig. 1). This observation was crucial for the understanding of the most feasible STR structure at this locus. Chimpanzee and human comparisons strongly support the initially described CTAT repeat motif [6] as the most probable STR structure and consequently should be the basis for nomenclature (Table 8). Intermediate alleles, sequenced for the first time in this study, displayed a complex structure instead of the simple CTAT tandem described so far (Table 4).

DXS7423

The distinct structure found in allele 8 suggests an allele nomenclature change, which would include the eight bases that interrupt the tandem TCCA tracts that are not being considered for allele designation [$(TCCA)_3$ tctgtcct $(TCCA)_n$] hence increasing in two repeats the allele nomenclature as variation at this tract cannot be excluded. However, the wide use of DXS7423 as well as its inclusion in the commercial X-chromosomal Mentype Argus X-8 kit [3] does not favor a nomenclature change. Therefore, the previously established one should be maintained [27] (Table 8).

Final remarks

The scarcity of sequence data for the most commonly used X-STRs permits that in some cases, different nomenclatures for the same locus are in use. Furthermore, in general, nomenclatures are established based on sequencing variation observed in a single population group in contradiction to ISFG guidelines [2, 7, 13]. As examples are the nomenclatures in use for HPRTB and DXS7423 loci, which with the accumulation of data proved that a different nomenclature should have been proposed: HPRTB loses one repeat unit [22] and DXS7423 two repeats (detected in this study). However, changing allele nomenclatures would generate unnecessary confusion since these are two of the earliest described and used loci that are also included in the X-STR Mentype Argus X-8 kit [3]. As demonstrated in the previous study [11], the present research also argues in favor of screening different human populations as well as chimpanzees when establishing nomenclatures for a novel STR and therefore, we strongly emphasize the importance of following this particular ISFG recommendation [2, 7, 13].

Acknowledgments This work was partially supported by FCT through grants SFRH/BD/21647/2005, SFRH/BD/30039/2006, and POCI/AFR/62242/2004 and by "Programa Operacional Ciência, e Inovação 2010".

References

1. Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H (2006) MiniX-STR multiplex system population study in Japan and

application to degraded DNA analysis. Int J Legal Med 120:174–181
2. 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
3. Becker D, Rodig H, Augustin C et al (2008) Population genetic evaluation of eight X-chromosomal short tandem repeat loci using Mentype Argus X-8 PCR amplification kit. Forensic Sci Int Genet 2:69–74
4. 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
5. Edelmann J, Deichsel D, Plate I, Käser M, Szibor R (2003) Validation of the X-chromosomal STR DXS6809. Int J Legal Med 117(4):241–244
6. 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
7. Gill P, Brinkmann B, d'Aloja E et al (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87:185–192
8. Gomes I, Alves C, Maxzud K et al (2007) Analysis of 10 X-STRs in three African populations. Forensic Sci Int Genet 1:208–211
9. Gomes I, Prinz M, Pereira R et al (2007) Genetic analysis of 3 US population groups using an X-chromosomal STR decaplex. Int J Legal Med 121:198–203
10. Gomes I, Pereira V, Gomes V et al. (2008) The Karimojong from Uganda: genetic characterization using an X-STR decaplex system Forensic Sci Int Genet. doi:10.1016/j.fsigen.2008.10.009
11. Gomes I, Prinz M, Pereira R (2009) X chromosome STRs sequence variation, repeat structure and nomenclature in humans and chimpanzees. Int L Legal Med 123:143–149
12. González-Neira A, Elmoznino M, Lareu MV, Sánchez-Diz P, Gusmão L, Prinz M, Carracedo A (2001) Sequence structure of 12 novel Y chromosome microsatellites and PCR amplification strategies. Forensic Sci Int 122:19–26
13. Gusmão L, Butler JM, Carracedo A et al (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int J Legal Med 120:191–200
14. Gusmão L, González-Neira A, Alves C, Lareu M, Costa S, Amorim A, Carracedo A (2002) Chimpanzee homologous of human Y specific STRs. A comparative study and a proposal for nomenclature. Forensic Sci Int 126:129–136
15. Gusmão L, Sánchez-Diz P, Alves C et al (2008) Results of the GEP-ISFG collaborative study on an X-STR Decaplex. Forensic Sci Int Genet Suppl 1:677–679
16. Gusmão L, Sánchez-Diz P, Alves C et al (2008) A GEP-ISFG collaborative study on the optimization of an X-STR Decaplex: data on 15 Iberian and Latin American populations. Int J Legal Med. doi:10.1007/s00414-008-0309-4
17. Kang L, Li S (2006) X-chromosome STR polymorphism of Luoba ethnic group living in Tibet (SW China). Forensic Sci Int 156:88–90
18. Kayser M, Caglià A, Corach D et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110 (125–133):141–149
19. Lazaruk K, Wallin J, Holt C, Nguyen T, Walsh PS (2001) Sequence variation in humans and other primates at six short tandem repeat loci used in forensic identity testing. Forensic Sci Int 119:1–10
20. Pereira R, Gomes I, Amorim A, Gusmão L (2007) Genetic diversity of 10 X-chromosome STRs in northern Portugal. Int J Legal Med 121:192–197
21. Shin SH, Yu JS, Park SW, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41
22. Szibor R, Edelmann J, Hering S, Gomes I, Gusmão L (2009) Nomenclature discrepancies in the HPRTB short tandem repeat. Int J Legal Med 123:185–186
23. 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. Letter to the Editor Forensic Sci Int 134:72–73
24. 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
25. Turrina S, De Leo D (2004) Population genetic comparisons of three X-chromosomal STRs (DXS7132, DXS7133 and GATA172D05) in north and south Italy. Progress in Forensic Genetics 10. Int Congr Ser 1261:302–304
26. Zarrabeitia MT, Amigo T, Sanudo C, de Pancorbo MM, Riancho JA (2002) Sequence structure and population data of two X-linked markers: DXS7423 and DXS8377. Int J Legal Med 116:368–371
27. Zarrabeitia MT, Amigo T, Sanudo C, Zarrabeitia A, González-Lamuño D, Riancho JA (2002) A new pentaplex system to study short tandem repeat markers of forensic interest on X chromosome. Forensic Sci Int 129:85–89
28. Zarrabeitia MT, Pinheiro F, de Pancorbo MM, Cainé L, Cardoso S, Gusmão L, Riancho JA (2009) Analysis of 10 X-linked tetranucleotide markers in mixed and isolated populations. Forensic Sci Int Genet 3:63–66