Int J Legal Med (2000) 113:146–149

ORIGINAL ARTICLE

S. Barral · M. V. Lareu · A. Salas · A. Carracedo

Sequence variation of two hypervariable short tandem repeats at the D22S683 and D6S477 loci

Received: 17 March 1999 / Received in revised form: 19 May 1999

Abstract For two short tandem repeats at the D22S683 and D6S477 loci, 30 and 22 selected alleles, respectively were sequenced. A total of 20 different alleles were found for the D22S683 locus and 12 alleles for the D6S477 locus. In both systems the alleles were designated according to the total number of repeats. D22S683 is a hypervariable STR consisting of blocks of (TATC) repeats with a basic sequence structure $(TATATC)_n$ $(TATC)_n$ $(ATC)_{0–1}$ $(TATC)_n$. The D6S477 locus consists of blocks of (TCTA) repeats with a basic sequence structure $(TCTA)_n$ $(TA)_1$ $(TCTA)_{0–2}$ $(TA)_{0–1}$ $(TCTA)_n$. Population data showed a heterozygosity of 0.89 for D22S683 and 0.75 for D6S477. These STRs are promising markers for forensic genetics as they are robust and can be easily included in multiplexes.

Key words D22S683 · D6S477 · Short tandem repeats · Sequence structure · DNA polymorphisms

Introduction

STRs are currently the markers which are most commonly used for forensic purposes and they are likely to continue to be widely used in the forthcoming years.

General criteria for selecting STRs for forensic use include characteristics such as hypervariability, low mutations, robustness, ease of multiplexing, low stutter characteristics and small fragment size.

Using a GATA probe short tandem repeats were identified at the loci D22S683 (CHLC–GATA–11B12.32029) and D6S477 (CHLC–GATA–3H05.73). Preliminary population data indicated the potential usefulness of these systems for forensic purposes (Lareu et al. 1998). The nucleotide sequences of all the common alleles of these systems have not yet been determined. Here we report the sequences of 20 different alleles for the D22S683 locus and 12 alleles for the D6S477 system. In addition a nomenclature for these systems is proposed and population data from Galicia (NW Spain) is reported for both systems.

S. Barral · M. V. Lareu · A. Salas · A. Carracedo (✉)
Institute of Legal Medicine, Faculty of Medicine,
c/San Francisco s/n, E-15705, Santiago de Compostela, Spain
e-mail: apimllar@usc.es,
Tel. +34-981-582327, Fax +34-981-580336

Materials and methods

Genomic DNA was isolated from human blood as described previously (Valverde et al. 1993). The primers used for amplification were as follows:

D22S683.–forward 5′ AACAAAACAAAACAAAACAAACA, reverse *5′ GGTGGAAATGCCTCATGTAG (* fluorescein labeled at 5′ end)

D6S477.–forward *5′ GATTTGCCATGATAGATGGC (* fluorescein labeled at 5′ end), reverse 5′ GGGGGATATCTCAAACAACC

The PCR reaction for both systems was performed using 5 ng of genomic DNA in a 50 µl reaction volume with 10 mM Tris-HCl (pH8.3), 50 mM KCl, 0.01% gelatin, 1.5 mM $MgCl_2$, 200 µM each dNTP, 0.25 µM each primer and 1.25 U AmpliTaq DNA polymerase (Cetus, Emerville, Calif.). PCR conditions were 30 cycles of denaturation at 94 °C for 45 s, annealing at 58 °C for 60 s and extension at 72 °C for 60 s in a Perkin Elmer thermocycler (Perkin-Elmer, Foster City, Calif.). The size of the PCR products were first determined in a 6% PAGE gel in an automated sequencer (A.L.F., APB, Uppsala, Sweden). The PCR products were purified from a PAGE gel after silver staining and DNA sequences were obtained using the PCR Fentomol sequencing kit (Promega, Madison, Wis.). The conditions for cycle sequencing were the same as previously described except for the annealing temperature which in both cases was 50 °C. The resulting PCR products were denatured and run on a 6% PAGE DNA sequencing gel.

Results and discussion

Nomenclature

Allelic designation was according to the recommendations of the DNA Commission of the International Society for Forensic Haemogenetics (Bär et al. 1997) and based on the number of repeats. Since a variation more complex than just tetranucleotide repeats was observed in both sys-

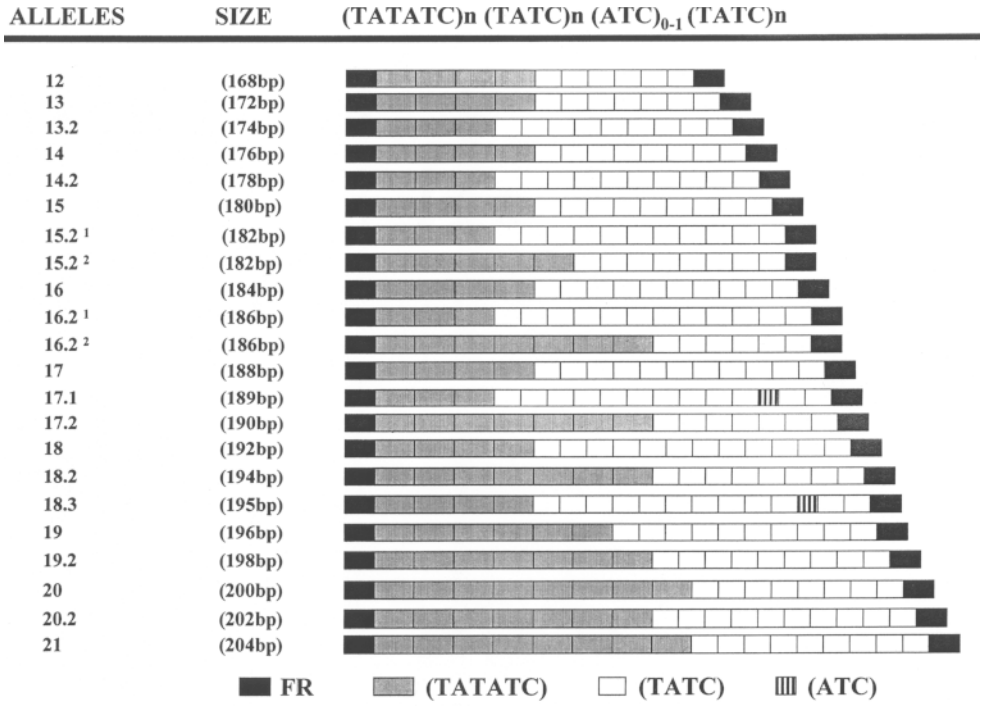

Fig. 1 Sequence structure of the different D22S683 alleles found in this study

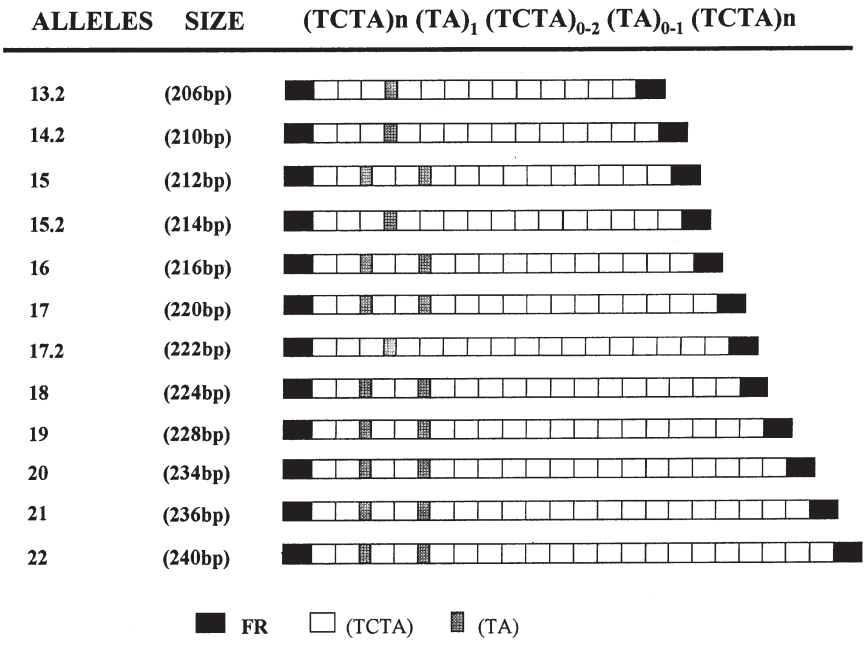

Fig. 2 Sequence structure of the different D6S477 alleles found in this study

tems (i.e. dinucleotide repeats), the alleles were named according to the number of bp of the variable tract divided by 4 bp (i.e. two dinucleotide repeats were considered as a single tetranucleotide repeat for nomenclature purposes).

Sequencing variation

The STR in the D22S683 locus was sequenced in 30 arbitrarily chosen alleles, including at least 2 individuals from the most common allelic groups and 20 different alleles were found.

The sequence composition of the D22S683 alleles is displayed in Fig. 1 and no variation was found in the constant regions. D22S683 is a hypervariable STR consisting of blocks of (TATC) repeats with a basic sequence structure $(TATATC)_n\,(TATC)_n\,(ATC)_{0-1}\,(TATC)_n$.

A variable number of (TA)n repeats inserted within the (TATC) units gives rise to regular and .2 alleles in most of the allelic classes. In addition some rare .1 alleles and .3 alleles were observed due to the insertion of an (ATC) in the variable (TATC)n sequence. D22S683 has not only a length polymorphism but also some structural variation in the alleles 15.2 and 16.2 (Fig. 1).

Table 1 Allele frequencies of D22S683 and D6S477 in the Galician population and statistical parameters for the two systems (*h* heterozygosity value, *Pd* power of discrimination, *CE* chance of exclusion)

Allele	D22S683 $n = 102$	D6S477 $n = 101$
12	0.1324	
13	0.0539	
13.2	0.0588	0.0198
14	0.1225	
14.2	0.1373	0.0792
15	0.1422	0.0099
15.2	0.0196	0.0149
16	0.0490	0.0792
16.2	0.0196	
17	0.0343	0.2178
17.1	0.0098	
17.2	0.0049	0.0099
18	0.0441	0.0792
18.2	0.0245	
18.3	0.0049	
19	0.0049	0.3069
19.2	0.0833	
20	0.0049	0.1733
20.2	0.0148	
21	0.0343	0.005
22		0.005
HWE: Exact test	$p = 0.56$	$P = 0.086$

System	h	Pd	CE
D22S683	0.88	0.98	0.81
D6S477	0.75	0.94	0.63

h = Heterocygosity value
Pd = power of discrimination
CE = chance of exclusion

The STR at the D6S477 locus was sequenced in 22 arbitrarily chosen alleles and 12 different alleles were found. The sequence composition of the D6S477 alleles is displayed in Fig. 2 and no variation was found in the constant regions. This STR consists of blocks of (TCTA) repeats with a basic sequence structure (TCTA)n. Intermediate alleles showed only one (TA) repeat instead of the two (TA) repeats observed in regular alleles.

Population data

For the D22S683 locus, 20 different allelic groups were found in a total of 102 unrelated healthy individuals from Galicia (NW Spain). Allele and genotype frequencies are given in Table 1. Frequencies range from 0.004 (alleles 17.2, 18.3, 19 and 20) to 0.142 (allele 15). The system showed no deviation from Hardy-Weinberg equilibrium and the exact test (Guo and Thompson 1992) gave a *P* value of 0.56. The discrimination power (Fisher 1951) was found to be 0.98, the heterozygosity (Nei and Roychoudhury 1974) 0.88.

For the D6S477 locus 12 different allelic groups were found in a total of 101 healthy unrelated individuals from Galicia (NW Spain). Allele and genotype frequencies are given in Table 1. Frequencies ranged from 0.005 (alleles 21 and 22) to 0.30 (allele 19). The system showed no deviation from Hardy-Weinberg equilibrium and the exact test gave a *P* value of 0.086. The discrimination power was 0.94 and the heterozygosity 0.75.

Additional information

Allelic ladders were constructed for D22S683 and D6S477. For D22S683 the allelic ladder was constructed with the following allele types: 12, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2, 18, 18.2, 19, 19.2, 20, 20.2, 21 (Fig. 3). For D6S477 the allelic ladder was constructed from the following alleles: 13.2, 14.2, 15, 15.2, 16, 17, 17.2, 18, 19, 20, 21, 22 (Fig. 4). Both allelic ladders are freely available from the authors.

D22S683 showed a lower percent stutter (< 1%) than any other STR used in our laboratory, even lower than other tetranucleotides such as TH01 and FES/FPS which

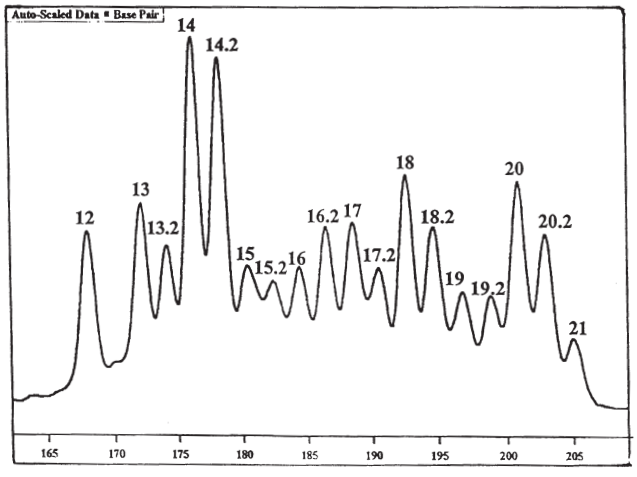

Fig. 3 Representation of the D22S683 allelic ladder composed of 18 sequenced alleles

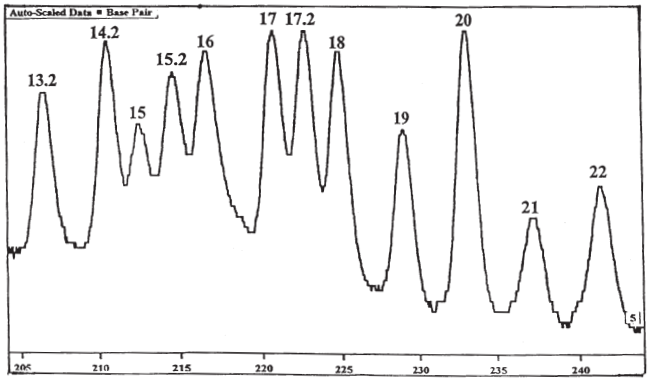

Fig. 4 Representation of the D6S477 allelic ladder composed of 12 sequenced alleles

are known to have good stuttering characteristics. D6S477 has medium stutter characteristics similar to FIBRA/FGA or D21S11.

Both STRs at the D22S683 and D6S477 loci are highly polymorphic and show a relatively simple structure which is uncommon since highly polymorphic STRs are usually complex in structure.

The characteristics of these systems, including easy amplification, high heterozygosity, low stutter characteristics and sequence simplicity, make these STRs very interesting DNA polymorphisms for forensic purposes.

The D6S477 locus is included in the commercially available kit SGM Plus (Perkin Elmer) but no sequence data has yet been reported.

Acknowledgements This work was supported by the grants XUGA 20806B97 and XUGA 20816B96 (Xunta de Galicia). The technical assistance of Meli Rodríguez is also acknowledged with appreciation.

References

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

Fisher RA (1951) Standard calculations for evaluating a blood group system. Heredity 5 : 95–102

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

Lareu MV, Barral S, Salas A, Rodriguez–Calvo MS, Pestoni C, Carracedo A (1998) Further exploring of new STRs of interest for forensic genetic analysis. In: Olaisen B, Brinkmann B, Lincoln P (eds) Progress in forensic genetics 7. Elsevier Science, Amsterdam, pp 198–200

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76 : 379–390

Valverde E, Cabrero C, Cao R, Rodriguez–Calvo M, Díez A, Barros F, Alemany J, Carracedo A (1993) Population genetics of the three VNTR polymorphisms in two different Spanish populations. Int J Legal Med 105 : 251–256