AMPFLP-TYPING OF THE D21S11 MICROSATELLITE POLYMORPHISM: ALLELE FREQUENCIES AND SEQUENCING DATA IN THE AUSTRIAN POPULATION

D.W.M.Schwartz, E.M.Dauber, B.Glock, W.R.Mayr

Clinical Department for Blood Group Serology,

University of Vienna

Waehringer Guertel 18-20/4I, A-1090 Vienna, Austria

INTRODUCTION

Human DNA microsatellite polymorphisms (Variable Number of Tandem Repeats, VNTR) with short repeat sizes (2-5bp) have become very useful markers in forensic haemogenetics in the last few years. For a number of reasons, typing for these markers is more and more preferred to conventional techniques in forensic casework:

- The large number, high polymorphism and thus the high information content of microsatellite markers
- The use of just one technique (PCR-Amplification Fragment Length Polymorphism, AMPFLP) for all markers
- The possibility to successfully employ these markers even for very low amounts of highly degraded stains

However, it must be stated that there are still some problems and restrictions with these systems. Typing by AMPFLP on polyacrylamide gels is not always very clear-cut, especially with certain markers (e.g. APOB, SE33) because of interalleles (variants with incomplete repeats or sequence variants with distinct electrophoretic mobility) and/or the limited resolution capability of electrophoresis. The allele frequency distribution in different populations has shown significant differences for certain systems, whereas other loci showed a more uniform distribution, at least in the populations studied so far. Since the amount of data is limited, especially for the microsatellites with short core repeat length (Short Tandem Repeats, STR) population studies are still required. Furthermore, detailed sequencing studies should also be performed for all markers in forensic use since only with this technique the polymorphism can be determined to the ultimate level. For routine use the AMPFLP technique will still stay the standard, but it is required that a marker consisting of alleles with known sequence is used with this technique [DNA comission of the ISFH, 1994]. In this work, we established optimized conditions for AMPFLP-typing of the tetranucleotide polymorphism at the D21S11 locus (Genbank M84567, [Sharma 1992]) and determined allele and genotype frequencies in our local (Austrian) population based on the alleles that could by resolved by AMPFLP. Sequencing studies were undertaken to define the alleles included in the marker and to identify sequence variants which have been previously reported for this locus [Moeller 1994].

MATERIAL AND METHODS

Population sample

A total of 200 healthy, unrelated indivduals (caucasians) of both sexes from Vienna were included in this study. Genomic DNA was extracted by standard techniques.

Primers

According to Sharma [1992] (Tab.2).

PCR

40ng template DNA, 0.4µM each primer, 2U polymerase (Dynazyme™, Finn Zymes Oy), 1x PCR buffer (50mM KCl, 10mM TrisCl pH=9.0 at 25°C, 0.1% Triton-X-100 and 1.5mM MgCl2) and 200µM of each dNTP, final volume 50µl overlaid with 50µl paraffine oil;
Hybaid Omnigene thermocycler;

first cycle	98°C 5min	62°C 10min	
26 cycles	94°C 60sec	62°C 60sec	72°C 75sec
last cycle	72°C 75sec		

Electrophoresis

PAGE was carried out as previously described [Schwartz 1994] on 6% native polyacrylamide gels (C=3%) in 112mM Tris-Acetic Acid rehydration buffer and 200mM Tris-Tricine electrode buffer.
running conditions: ramp: 1200V 12mA 10min
final 1200V 20mA
Subsequent silver staining was applied to visualize DNA [Bassam et al. 1991]. Typing was done by side-to-side comparison with the allelic ladder.

Allelic ladder

Single bands of heterozygous population samples, corresponding to distinct alleles, were eluted from the gel. DNA was purified using Wizard PCR Preps DNA Purification System (Promega, technical bulletin), diluted and reamplified. Equal concentrations of reamplification products were pooled to construct an allelic ladder.

Sequencing

Single strand sequence determination of the different alleles was performed using an automatic DNA-Sequencer (ALF™, Pharmacia LKB Technology) according to the protocol of the Pharmacia AutoRead™ Sequencing Kit (dye primers, T7 polymerase) on a 6% sequencing gel. A total of 37 alleles was sequenced including at least one sample for each distinct allele (as judged by AMPFLP). For each sample the sense and the antisense-strand were sequencend.

RESULTS AND DISCUSSION

AMPFLP typing

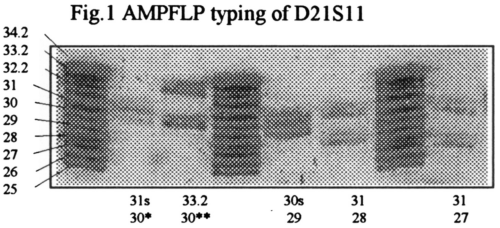
A total number of 12 alleles could be identified by PCR-AMPFLP in the population sample (Tab.1). These alleles were compared to the allelic ladder kindly provided by B. Brinkmann and designated according to the proposed nomenclature [Moeller 1994]. From the electrophoretic moblilty, it was not possible to clearly assign alleles 30.2, 31.2 and 32 (Fig.1). So two provisional designations were made: 30s and 31s (s=slow). Allele frequencies, mean paternity exclusion chance (MEC), mean exclusion probability (MEP) and polymorphism information content (PIC) are given in Tab.1. No significant deviation from Hardy-Weinberg expectations was observed.

Table 1: D21S11 allele frequencies (n=200)

allele designation	frequency (%)
25	0.3
26	0.3
27	2.8
28	21.5
29	19.8
30	23.8
30s	9.0
31	11.8
31s	0.8
32.2	5.5
33.2	4.3
34.2	0.5

Mean paternity exclusion chance (MEC) 0.665
Mean exclusion probability (MEP) 0.661
Polymorphism information content (PIC) 0.809

Fig.1 AMPFLP typing of D21S11



* variant 6-5-11
** variant 4-6-12 according to Tab.2

Sequencing:

As already described [Moeller 1994], at the molecular level D21S11 actually contains 3 polymorphic regions (vrI-III, Tab.2). vrI consists of a (tcta)-tetranucleotide that shows a variation from 4-7 repeats. Just adjacent, vrII 2 shows either 3, 5 or 6 (tctg)-blocks. The most polymorphic part is vrIII that shows from 8-14 (tcta)-repeats. Additionally, in some alleles a constant block (ta tcta) is inserted at the 3'-end of vrIII which possibly represents an incomplete (tcta)2 repeat resulting from a dinucleotide deletion.

Table 2: D21S11 microsatellite sequence polymorphism

allele designation based on AMPFLP	sequence	size (bp)	constant region 1 (28bp)	constant region 2 (43bp)	insert	constant region 3 (66bp)	no.of alleles sequenced
24.2*	4-6-del-8	203	5'.....(tcta)4(tctg)6.....	del	(tcta)8	.....3'	0
C25	4-3-10	205	.....(tcta)4(tctg)3.....(tcta)10			.....	1
C26	4-6-8	209	.....(tcta)4(tctg)6.....(tcta)8			.....	1
27+	4-6-9	213	.....(tcta)4(tctg)6.....(tcta)9			.....	2
C27+	6-5-8	213	.....(tcta)6(tctg)5.....(tcta)8			.....	1
C28	4-6-10	217	.....(tcta)4(tctg)6.....(tcta)10			.....	3
C29**,+	4-6-11	221	.....(tcta)4(tctg)6.....(tcta)11			.....	2
29**,+	6-5-10	221	.....(tcta)6(tctg)5.....(tcta)10			.....	1
30+	6-5-11	225	.....(tcta)6(tctg)5.....(tcta)11			.....	3
C30+	5-6-11	225	.....(tcta)5(tctg)6.....(tcta)11			.....	1
30+	4-6-12	225	.....(tcta)4(tctg)6.....(tcta)12			.....	1
30.2+	5-6-10-ins	227	.....(tcta)5(tctg)6.....(tcta)10		ta tcta	.....	2
30.2+	5-5-11-ins	227	.....(tcta)5(tctg)5.....(tcta)11		ta tcta	.....	1
31**	5-6-12	229	.....(tcta)5(tctg)6.....(tcta)12			.....	2
31+	6-5-12	229	.....(tcta)6(tctg)5.....(tcta)12			.....	0
C31+	7-5-11	229	.....(tcta)7(tctg)5.....(tcta)11			.....	1
31.2	5-6-11-ins	231	.....(tcta)5(tctg)6.....(tcta)11		ta tcta	.....	5
32+	6-5-13	233	.....(tcta)6(tctg)5.....(tcta)13			.....	1
32+	5-6-13	233	.....(tcta)5(tctg)6.....(tcta)13			.....	1
C32.2	5-6-12-ins	235	.....(tcta)5(tctg)6.....(tcta)12		ta tcta	.....	3
C33.2	5-6-13-ins	239	.....(tcta)5(tctg)6.....(tcta)13		ta tcta	.....	3
C34.2	5-6-14-ins	243	.....(tcta)5(tctg)6.....(tcta)14		ta tcta	.....	2

constant region 1: 5'- [gtg agt caa ttc ccc aag] tga att gcc t -3'

constant region 2: 5'- (tcta)3 ta (tcta)3 tca (tcta)2 tcca ta -3'

constant region 3 : 5'- tcg (tcta)2 tccag (tcta)2 cct cct att agt ctg tct ctg [gag aac att gac taa tac aac] - 3'

[] denotes primer target sequences

C25,26... Alleles included in allelic cocktail for AMPFLP typing

* Genbank sequence M 84567 shows a deletion of a 6 bp sequence (tcca ta) on the 3'-end of constant region 2

** alleles and variants described by Moeller et al.(1994)

\+ alleles and variants described by Schwartz et al.(1995)

Several new sequence variants were identified (Tab.2). Remarkably, the Genbank M84567 sequence [Sharma 1992] shows another polymorphism so far only described on this occasion, namely a hexanucleotide deletion at the 5'-flank of vrIII. We suggest a new nomenclature for D21S11 alleles at the nucleotide level which allows the discrimination of all the so far described polymorphisms (Tab.2). Taking into account the relatively small number of alleles sequenced so far (compared to the observed polymorphism) it must be expected that even more variants will exist in the population. So, at the nucleotide level D21S11 represents one of the most polymorphic STR so far described and is therefore of high informative value in forensic haemogenetics when using DNA-sequencing. However, since the occurence of incomplete repeats is relatively common, this makes this locus difficult to type for with the AMPFLP-technique and the results should be interpreted with caution.

REFERENCES

Bassam BJ, Caetano Anolles G, Gresshoff PM: Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal.Biochem.: 196: 80-83 (1991)

DNA comission of the ISFH: DNA recommendations - 1994 report concerning further recommendations of the DNA comission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Int.J.Leg.Med 107: 159-160 (1994)

Moeller A, Meyer E, Brinkmann B: Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. Int.J.Leg.Med 106: 319-323 (1994)

Schwartz DWM, Jungl EM, Krenek OR, Mayr WR: Simple and rapid typing for VNTR polymorphisms using high resolution electrophoresis of PCR products on rehydratable polyacrylamide gels. Advances in Forensic Haemogenetics 5: 170-172, Springer, Berlin, FRG (1994)

Sharma V, Litt M: Tetranucleotide repeat polymorphism at the D21S11 locus. Hum.Mol.Genet. 1: 67 (1992)

ACKNOWLEDGEMENTS

The authors are grateful to B. Brinkmann for providing an allelic ladder for standardisation and to H.Gnauer for his excellent technical assistance.