Automated fluorescent PCR based analysis of the STR polymorphism at locus D8S639 and at the CYP19 gene

C.Seidl, O.Jäger, M.Kilp, E.Seifried

Institute for Transfusion Medicine and Immunohematology, Red Cross Blood Donor Service Hessen, Sandhofstrasse 1, 60528 Frankfurt, Germany

INTRODUCTION

Short tandem repeat (STR) loci are polymorphic markers that can be used for discrimination between individuals in paternity and forensic testing. The feasability to analyse these markers by amplification using the polymerase chain reaction (PCR) has improved the sensitivity of DNA analysis from small amounts of DNA. Furthermore, due to the allele distribution of most of these markers between 100 - 500 bp, the analysis of individual allele patterns is very robust to DNA degradation. Unfortunately, the analysis of the common dinucleotide repeat polymorphism can be problematic due to the occurence of 'shadow bands'. The less common tri-, tetra- and pentanucleotide repeats in contrast are less sensible for these PCR artefacts and are more preferable for PCR analysis in paternity or forensic casework. We have studied the allele distribution and sequence structure of two tetranucleotide repeat polymorphism, one at locus D8S639 (Nelson 1994) and the second starting at base pair 682 of the human aromatase cytochrome P-450 gene (CYP19) (Polymeropoulos 1992).

Systems and locus: HUMD8S639 (8p21-p11)
HUMCYP19 (15q21.1)

Population and sample size: Hesse (Germany)
N: 110 (HUMD8S639)
N: 132 (HUMCYP19)

METHODS

Locus D8S639: Primersequences were choosen according to Nelson (1994). The 5' primer was labelled with HEX (6-carboxy-2',4',7',4,7-Hexachlorofluorescein, Perkin Elmer - ABD)
PCR amplification conditions: 95 C - 5 min; 28 cyles 95 C - 30 sec , 54 C 40 sec, 72 C 30 sec; 72 C 5 min. (Perkin Elmer 9600).
PCR reaction conditions: 10-100ng DNA, 5pmol 5'and 3' primer, 200 µmol dNTP's, 0,5 Units Taq DNA polymerase (Appligene) and the corresponding buffer (Appligene) in a final volume of 50 µl.

Locus CYP19: Primersequences were choosen according to Polymeropoulos (1992). The 5' primer was labelled with FAM (5-carboxylfluorescein, Perkin Elmer - ABD).
PCR amplification conditions: 95 C - 5 min; 28 cyles 95 C - 30 sec , 54 C 40 sec, 72 C 30 sec; 72 C 5 min. (Perkin Elmer 9600).
PCR reaction conditions: 10-100ng DNA, 10pmol 5'and 3' primer, 200 µmol dNTP's, 0,5 Units Taq DNA polymerase (Appligene) and the corresponding buffer (Appligene) in a final volume of 50 µl.

Electrophoretic methods: 6% polyacrylamide denaturing gel electrophoresis. The gels were run for 8h at constant power (30W) 1600V and 28mA on an ABD automated DNA seqencer 373A. Typing was performed by comparison with an ROX (6-carboxy-X-rhodamine) labelled internal

standard generated from the vector pGL-2-Basic (Promega) using the Southern local method for fragment size assignement (Genescan software, Perkin Elmer - ABD).

Sequence analysis: Individual alleles were sequenced using the solid-phase sequence strategy. Biotin labelled PCR primers were included in the amplification reaction and PCR fragments were subsequently separated by streptavidin coated magnetic beads (Dynal). Sequence reaction were performed using the T7 DNA polymerase and Dye terminators (Prism T7 sequence kit, Perkin Elmer - ABD). Analysis of sequence reaction was conducted on an ABD 373A automated DNA sequences using the Sequence Navigator software. Allele assignement was conducted according to the number of repeat units of the individual alleles.

Statistical analysis: The polymorphic information content (PIC) was calculated using the formula of Botstein (1980). The discrimination index (DI) and the matching probability (pM) were calculated by the method of Jones (1972). The sample gene diversity (geneD) (frequency of heterozygotes expected under Hardy Weinberg equilibrium was calculated as described by Kimpton (1993).

RESULTS

HUMCYP19 - Observed genotypes

Gen.	Obs. (N)	Gen.	Obs.(N)	Gen.	Obs.(N)	Gen.	Obs.(N)
7-3 - 7-3	16	**7-3 - 12**	9	**7 - 12**	1	**9 - 11**	1
7-3 - 8	13	**7 - 7**	1	**8 - 8**	1	**10 - 11**	2
7-3 - 9	7	**7 - 8**	4	**8 - 9**	1	**11 - 11**	18
7-3 - 10	1	**7 - 10**	2	**8 - 10**	1		
7-3 - 11	28	**7 - 11**	18	**8 - 11**	9		

HUMCYP19 - Allele frequencies

Allele (bp*)	Frequency	Allele (bp*)	Frequency	Allele (bp*)	Frequency
7-3** (168)	0,341	**9** (179)	0,004	**12** (191)	0,038
7 (171)	0,152	**10** (183)	0,023		
8 (175)	0,087	**11** (187)	0,356		

* lenght as detected by Genescan software

** allele 7-3 has a 3bp deletion in the 5'flanking region

HUMD8S639 - Observed genotypes

Gen.	Obs. (N)	Gen.	Obs. (N)	Gen.	Obs. (N)	Gen.	Obs. (N)
22 - 26	2	**25 - 30-1**	2	**27 - 28**	9	**28 - 32-1**	2
23 - 27	2	**25 - 31-1**	1	**27 - 29**	5	**28 - 34-1**	2
24 - 25	2	**25 - 33-1**	1	**27 - 30**	2	**28 - 36-1**	1
24 - 27	1	**26 - 26**	2	**27 - 31-3**	1	**29 - 29**	1
24 - 28	3	**26 - 27**	12	**27 - 31**	1	**29 - 30**	2
25 - 25	1	**26 - 28**	5	**27 - 33-1**	1	**29 - 31-1**	1
25 - 26	2	**26 - 29**	5	**28 - 28**	3	**29 - 34-1**	1
25 - 27	7	**26 - 30**	1	**28 - 29**	4	**30-1 - 31**	2
25 - 28	8	**26 - 33-1**	1	**28 - 30**	4		
25 - 29	2	**27 - 27**	8	**28 - 31**	1		

HUMD8S639 - Allele frequencies

Allele (bp*)	Frequency	Allele (bp*)	Frequency	Allele (bp*)	Frequency
22 (317)	0,009	28 (341)	0,200	32-1 (356)	0,009
23 (321)	0,009	29 (345)	0,100	33-1 (360)	0,014
24 (325)	0,027	30-1 (348)	0,018	34-2 (363)	0,005
25 (329)	0,123	30 (349)	0,041	34-1 (364)	0,005
26 (333)	0,145	31-1 (352)	0,014	36-1 (372)	0,005
27 (337)	0,259	31 (353)	0,018		

* lenght in base pairs as detected by Genescan software
** number of trinucleotide repeats are indicated by (-n).

COMMENTS

STR locus	HR	gene D	PIC	DI	pM	mean exclusion change
HUMCYP19	0,727	0,724	0,719	0,88	0,12	48,9
HUMD8S639	0,857	0,838	0,819	0,95	0,05	67,5

Hardy Weinberg equilibrium was observed for both STR loci.
The sequence analysis of alleles at the HUMCYP19 locus revealed a regular tetranucleotide repeat structure. Alleles 7-3 (168bp) and allele 7 (171bp) exhibit the same repeat unit structure containing 7 tetranucleotide repeat units, but allele 7-3 has a 3bp deletion in the 5' flanking region.
Alleles at locus D8S639 contain tri- and tetranucleotide repeat units leading to allele sizes that differ only by one base pair. In addition, several alleles at this locus contain structural differences in their repeat order but have the same number of repeat units leading to PCR fragments of identical sizes.

REFERENCES

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment lenght polymorphisms. Am.J.Hum.Genetics 32: 182-190
Jones DA (1972) Blood samples: Probability of discrimination. J.Forens.Sci.Soc.12: 355-359
Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR methods and applications 3: 13-22
Nelson L, Lu J, Petterson M, Riley R, Ward K (1994) Tetranucleotide repeat polymorphism at the D8S639 locus. Human Molecular Genetics Vol. 3 (7): 1209
Polymeropoulos MH, Xiao H, Rath DS, Merril CR (1992) Tetranucleotide repeat polymorphism at the human aromatase cytochrome P-450 gene (CYP19). Nucleic Acids Research Vol.19(1): 195