POPULATION GENETIC STUDIES OF SHORT TANDEM REPEAT LOCI (STRS): EFFICIENCY IN PATERNITY TESTING

A.P. Pai, T. Jackson, D.J. Pritchard, S.S. Papiha

Department of Human Genetics, University of Newcastle upon Tyne, UK.

Tandemly repeated DNA sequences are extensively polymorphic, due to variability in reiteration number (Jeffreys et al., 1985). There are two types of variable number of tandem repeats (VNTRs), long and short. Long VNTRs, also known as minisatellites, are *generally* detected by the technique of Southern hybridisation, but due to technical limitations the allele sizes of the long VNTR loci cannot be determined with precision, which complicates their application. In contrast, short trandem repeats (STRs), also known as microsatellites (Litt and Luty, 1989), are widely scattered throughout the human genome and provide a valuable source of polymorphic markers (Weber and May, 1989). These STRs can be precisely amplified by the polymerase chain reaction (PCR) and the variability of allele size can be determined faithfully in repeated experiments. Several STRs can simultaneously be studied by multiplex PCR (Chamberlain et al., 1988), making the STR analysis both rapid and less expensive. For both technical and economic reasons STR analysis provides an important and valuable tool for individual identification in paternity and forensic testing. This study presented here provides population frequency data for two tetrameric loci, one trimeric and one dimeric locus. The efficiency of these systems in paternity testing was also analysed by a retrospective analysis of ten paternity cases (trios) in which the exclusion had already been established with each of three long VNTR probes.

Materials and Methods

DNA was extracted by phenol/chloroform extractions, from 10ml EDTA blood samples, collected from 205 random individuals born and living in north-east England. For the timeric HUM*FABP*(AAT)n and tetrameric HUMH*PRTB*(AGAT)n and HUM*RENA4*(ACAG)n loci the primers and the conditions of PCR used were as described in Edwards et al. (1992). For the dimeric (CA)n locus D10S89, the primers were supplied by Professor S.S. Bhattacharya (London) and the amplification conditions were: denaturation for 1 minute at 94°C, annealing for 1 minute at 55°C, and extension at 72°C for 1 minute for 30 cycles. Phenotypic characterisation was carried out in DNA sequencing gels containing 6% acrylamide with 8M Urea.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

Results and Discussion

The STR loci investigated, their chromosomal locations, total number of alleles tested, allele frequencies, with their SE, heterozygosity, probability of exclusion of each system and the proportion of exclusions given by each autosomal locus are given in Table 1. The *PRTB* locus is on the X chromosome and a total of eight alleles were found with allele sizes between 263 and 299 bp. Allele 12 gave the highest frequency (34%) which is similar to the white Caucasian sample by Edwards et al. (1992) in the U.S. The frequency of allele 13 is second highest which also agrees with the Edwards study. For this allele the Asians investigated in the U.S. gave the highest value (51%). This locus therefore shows significant heterogeneity in different racial groups although the white Caucasians on either side of the Atlantic seem to be homogeneous. For loci RENA4 and FABP, alleles 8 and 19 respectively gave the highest frequencies (81% and 52%). Once again the patterns of allele frequencies for the two loci are similar to the white Caucasian population of the U.S. No data is available for the direct comparison of locus D10S89. In the present study of 148 individuals, 9 alleles were found and the allele sizes ranged from 140 to 156 bp. The highest values were found for alleles with 144 and 150 bp. Both showed similar frequencies, together constituting 68% of the total genetic variation at this locus.

The power of exclusion (PE) for the three autosomal loci was calculated (Garber and Morris, 1983). RENA4 with four alleles gave the lowest PE value (17%) while the dimeric loci D10S89 with nine alleles gave the highest value of power of exclusion (61%). The FABP locus showed an intermediate PE value. The cumulative PE value for the three autosomal loci is 81%

Paternity cases excluded by each of the three VNTR probes (g3, MS1 and MS43) were examined for exclusion with each of the three autosomal STRs investigated here. Sixty percent were excluded with the trimeric STR, while 43% (3 out of 7) and 30% of cases were independently excluded by the dimeric and tetrameric STRs. Three cases (43%) were excluded with two STR systems while the other four (57%) showed exclusion with only one system.

In conclusion, the magnitude of polymorphism for STR markers differs considerably, as does their power of exclusion. These systems are certainly useful for genetic differentiation studies, but the three autosomal STRs studied can not replace three long VNTRs for individual identification. These three STR systems could however be used as additional markers for confirmation of results.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

References

Chamberlain, J.W., Gibbs, R.A., Ranier, J.E., Nguyen, P. van. and Caskey, C.T. (1988) Deletion screening of Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acid Research*, **16**, 11141-56.

Edward, A., Hammond, H., Jiri, L., Caskey, C.T. and Chakraborty, R. (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, **12**, 241-253.

Garber, B.A. and Morris, J.W. (1983) General equation for average power of exclusion for genetic systems of n codominant allele in one parent and no parent case of disputed parentage. In: Walker, R.H. (ed.) *Inclusion probabilities in parentage testing.* American Association Blood Bank. pp. 277-280.

Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Hypervariable 'minisatellite' region in human DNA. *Nature* (London), **314**, 67-73.

Litt, M. and Luty, J.A. (1989) A hypervaribale microsatellite revealed by *in vitro* amplification of dinucleotide repeat within the cardic muscle actin gene. *American Journal of Human Genetics*, **44**, 397-401.

Weber, J.L. and May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-96.

Table 1: Allele frequencies with standard errors, chromosome location, heterozygosity, power of exclusion and the percentage of cases excluded by the STR loci

Loci	HUMHPRTB(AGAT)n (n = 182)		HUMRENA4(ACAG)n (n = 193)		HUMFABP(AAT)n (n = 205)		D10S89 (n = 148)	
	allele	frequency & standard error	allele	frequency & standard error	allele	frequency & standard error	allele	frequency & standard error
	9	0.003±0.003	8	0.808±0.020	10	0.524±0.025	140	0.003±0.003
	10	0.023±0.008	10	0.078±0.014	11	0.151±0.018	142	0.017±0.007
	11	0.160±0.018	11	0.104±0.016	12	0.029±0.008	144	0.358±0.028
	12	0.338±0.024	12	0.010±0.005	13	0.222±0.021	146	0.044±0.012
	13	0.282±0.023			14	0.049±0.011	148	0.095±0.017
	14	0.137±0.017			15	0.024±0.008	150	0.331±0.027
	15	0.053±0.011					152	0.125±0.019
	16*	0.005±0.004					154	0.014±0.007
							156	0.014±0.007
cho	Xq26		1q32		4q31		10p	
H	.6888**		.3419		.6439		.7568	
PE	-		0.1747		0.4138		0.6099	
%	-		30		60		43	

* new allele;
** for females;
Cho = Chromosomal location;
H = heterozygosity;
PE = power of exclusion;
% = percentage of exclusion

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg