Int J Legal Med (1998) 111 : 343–345

SHORT COMMUNICATION

Feng-Xia Xiao · A. Gilissen · Xiao-Xiao Gu
J.-J. Cassiman · R. Decorte

Genetic data obtained for two Chinese Han populations with a quadruplex fluorescent STR typing system (HUMVWA, HUMTH01, D21S11 and HPRT)

Received: 22 December 1997 / Accepted: 23 April 1998

Abstract DNA typing of four tetrameric repeat loci (HUMVWA, HUMTH01, D21S11 and HPRT) was carried out in a Chinese Han population from Shanghai (East China) and one from Guangzhou (South-East China) using a quadruplex PCR amplification and detection of the fluorescent-labeled alleles on the ALF DNA sequencer. All loci were in accordance with Hardy-Weinberg equilibrium except for D21S11 in the Guangzhou population. A test for population differentiation showed no statistical difference in the allele frequency distribution between the two populations. Comparison of the allele frequency data with other Chinese Han populations from North and South-West China for the STR loci HUMVWA and HUMTH01 revealed heterogeneity between Northern Chinese Han and Southern Chinese Han, which is in accordance with previous studies on the basis of protein markers.

Key words DNA · STR · Multiplex PCR · Forensic identification · Population genetics

Introduction

Tetrameric short tandem repeat (STR) typing systems are a reliable, rapid and sensitive method for forensic identification, paternity testing (Hammond et al. 1994; Alford et al. 1994; Füredi et al. 1996; Rousselet et al. 1996) and population genetic studies (Wall et al. 1993; Brinkmann et al. 1996).

In this study the allele frequencies for four tetrameric STR loci (HUMVWA, HUMTH01, D21S11 and HPRT) were determined in two Chinese Han populations from East and South-East China. Some statistical parameters of forensic interest, such as expected heterozygosity, matching probability and chance of paternity exclusion were determined. In addition, the allele frequencies were compared with three other Chinese Han populations (Hou and Walter 1996).

F.-X. Xiao · A. Gilissen · X.-X. Gu · J.-J. Cassiman
R. Decorte (✉)
Center for Human Genetics, University of Leuven,
Campus Gasthuisberg O&N6, Herestraat 49,
B-3000 Leuven, Belgium
Tel. (32-16)-34.60.77; Fax (32-16)-34.59.97
e-mail: Ronny.Decorte@med.kuleuven.ac.be

Materials and methods

Population samples

Saliva samples from 98 unrelated Chinese Han individuals were collected in Shanghai located at the East coast of China. These samples were collected with a sterile cheekbrush (Integrated Genetics, Framington, MA) by brushing the inner side of the cheek for about 20–30 s. The brush was placed into 600 µl of denaturant solution (50 mM NaOH). Samples were boiled in a waterbath (95°C) for 5 min and neutralized with 60 µl of 1 M Tris-HCl (pH 8.0). DNA samples from 82 unrelated Chinese Han individuals from the Guangzhou area in South-East China were obtained from the Department of Biochemistry of the Sun Yat-Sen University of Medical Sciences (Guangzhou, China).

Typing protocol

Three autosomal STR loci (HUMVWA, HUMTH01, D21S11) and one X-linked STR locus (HPRT) were investigated in this study. The primer sequences and chromosomal locations have been described by Decorte and Cassiman (1996).

A quadruplex PCR for the four STR loci was used for amplification in a 25 µl PCR reaction containing 0.1–5 ng of genomic DNA and the fluorescently-labeled PCR products were resolved by electrophoresis in a denaturing gel on the ALF DNA Sequencer (Pharmacia-Biotech,Uppsala, Sweden) as described by Xiao et al. (in press). Identification of the amplified alleles was done using locus specific allelic ladders (Xiao et al. in press) with Fragment Manager Software V1.2.

Statistical analysis

The allele frequencies for the two populations at the four loci were calculated from the observed number of genotypes. The expected degree of heterozygosity was computed using Levene's correction (1949). The exact probability without bias for Hardy-Weinberg equilibrium (HWE), population differentiation and genotypic disequilibrium among pairs of loci was estimated using the GENEPOP Software V3.1 (Raymond and Rousset 1995) with a Markov chain

method described by Guo and Thompson (1992). The matching probability (pM) was calculated from the genotype frequency according to Fisher (1951). The chance of paternity exclusion (CE) was estimated by the method of Ohno et al. (1982).

Table 1 Allele frequencies for HUMVWA, HUMTH01 and D21S11 in the two Chinese populations investigated

Locus	Allele size (bp)	Repeat number	Shanghai (n = 196)	Guangzhou (n = 164)
HUMVWA	134	13	0.005	0.000
	138	14	0.240	0.305
	142	15	0.051	0.031
	146	16	0.179	0.152
	150	17	0.301	0.281
	154	18	0.138	0.128
	158	19	0.076	0.079
	162	20	0.010	0.018
	166	21	0.000	0.006
HUMTH01	175	4	0.005	0.000
	183	6	0.122	0.140
	187	7	0.245	0.238
	191	8	0.056	0.024
	195	9	0.510	0.488
	198	9.3	0.036	0.043
	199	10	0.026	0.067
D21S11	213	27	0.000	0.006
	217	28	0.031	0.043
	219	28.2	0.010	0.000
	221	29	0.301	0.299
	225	30	0.260	0.220
	227	30.2	0.010	0.018
	228*	30.3	0.010	0.000
	229	31	0.117	0.098
	231	31.2	0.072	0.079
	233	32	0.015	0.030
	235	32.2	0.138	0.122
	237	33	0.000	0.024
	239	33.2	0.026	0.055
	243	34.2	0.005	0.006
	247	35.2	0.005	0.000

n = the number of chromosomes
* new allele, not reported before

Results and discussion

The allele frequencies observed for the four loci in the two Chinese Han populations are summarized in Tables 1 and 2. Efficient co-amplification of the four loci was obtained on all blood DNA samples (DNA template > 5 ng) with 28 cycles. Saliva samples (template DNA between 0.5 and 2 ng) could only be co-amplified when the number of cycles were increased to 40. A few saliva DNA samples failed to be amplified due to a low amount of template DNA (< 0.5 ng), which was resolved by single locus amplification for 40 cycles. The fact that 40 cycles were necessary was probably due to the high salt concentration present in the DNA samples (see materials and methods). This inhibition was also observed by typing Y chromosomal STRs but could be overcome by desalting on a Microcon 100 device (Xiao et al., manuscript in preparation).

At the D21S11 locus, one allele (228 bp) was observed in the Shanghai population, that has not yet been reported. Sequence analysis of the 228 bp allele revealed the nonconsensus sequence structure:

$(TCTA)_6(TCTG)_5(TCTA)_3TA(TCTA)_3TCA(TCTA)_2$
$TCCATA(TCTA)_5\mathbf{TCA}(TCTA)_6$

At the 3′-end, the sixth repeat of the variable region (TCTA) showed a T residue deletion. This new allele was designated as allele 30.3 according to the recommendations of the DNA Commission of the ISFH (1997).

The exact test by the Markov chain method showed no significant deviation from HWE for all loci except for D21S11 in the Guangzhou population (P = 0.010). This was further confirmed by a χ^2-test based on the observed (60) and expected (68.07) number of heterozygotes (P = 0.018). The exact test for genotypic disequilibrium revealed a significant association between HPRT and D21S11 in the Guangzhou population (P = 0.025) and no significant association between all pairs of loci in the Shanghai population.

The exact test for population differentiation showed no significant difference for the allele frequencies at the four loci between the two populations. This homogeneity demonstrates that the population data obtained in this study may be valid for all Southern Chinese Han populations. In addition, the forensic efficiency values indicated that the

Table 2 Allele frequency for males, females and for the total population at the HPRT locus

Allele size (bp)	Repeat number	Shanghai			Guangzhou		
		Total (n = 147)	Male (n = 49)	Female (n = 98)	Total (n = 117)	Male (n = 47)	Female (n = 70)
279	12	0.054	0.041	0.061	0.077	0.106	0.057
283	13	0.258	0.286	0.245	0.282	0.298	0.271
287	14	0.531	0.469	0.561	0.436	0.362	0.486
291	15	0.116	0.184	0.082	0.137	0.170	0.114
295	16	0.027	0.000	0.041	0.060	0.043	0.072
299	17	0.007	0.020	0.000	0.000	0.000	0.000
303	18	0.007	0.000	0.010	0.008	0.021	0.000

n = the number of chromosomes

Table 3 Statistical parameters of forensic interest

Population	Locus	H (%)	pM	CE	Frequency of the most common genotype
Shanghai	HUMVWA	79.65	0.074	0.595	0.133
	HUMTH01	66.30	0.160	0.427	0.286
	D21S11	80.57	0.065	0.618	0.133
	HPRT	61.92	0.184	0.293	0.306
	Combined	–	1.416×10^{-4}	0.937	1.548×10^{-3}
Guangzhou	HUMVWA	78.60	0.080	0.577	0.146
	HUMTH01	68.30	0.146	0.447	0.244
	D21S11	83.01	0.051	0.663	0.134
	HPRT	67.88	0.135	0.329	0.314
	Combined	–	8.042×10^{-5}	0.947	1.499×10^{-3}

H: expected heterozygosity; pM: matching probability; CE: chance of paternity exclusion. HPRT: heterozygosity only applies to females

four tetrameric STR loci have a high discrimination power for identity and paternity testing in Chinese populations (Table 3).

Some STR data for three other Chinese Han populations [Guangzhou, Chengdu (South-West China) and Changchun (Northern China)] were compared with our data (Hou and Walter 1996). A significant difference in allele frequencies for the HUMVWA locus between the Chengdu population and the Changchun population ($P = 0.011$), as well as for the HUMTH01 locus between our Guangzhou population and the Changchun population ($P = 0.015$) was observed. Genetic studies between different Chinese populations on the basis of protein and physical markers have demonstrated that modern Chinese could be divided into two distinctive clusters, Northern and Southern Chinese (Zhao and Lee 1989; Chen et al. 1993). They also provided evidence for the hypothesis that modern Chinese populations originated from two distinct groups of Mongoloids. Our results indicate that there is heterogeneity in the nuclear gene pool between Southern Chinese Han and Northern Chinese Han, which confirms the hypothesis on the origin of modern Chinese at the DNA level.

Acknowledgements We wish to thank Dr. Sun Chen-Guang (Department of Biochemistry, Sun Yat-Sen University of Medical Sciences, Guangzhou, China) for his contribution of the Guangzhou population samples. This work was supported by a grant (nr G.0241.98) from the Fund for Scientific Research - Flanders (1998–1999)

References

Alford RL, Hammond HA, Coto I, Caskey CT (1994) Rapid and efficient resolution of parentage by amplification of short tandem repeats. Am J Hum Genet 55:190–195

Brinkmann B, Sajantila A, Goedde HW, Matsumoto H, Nishi K, Wiegand P (1996) Population genetic comparisons among eight populations using allele frequency and sequence data from three microsatellite loci. Eur J Hum Genet 4:175–182

Chen R, Ye G, Geng Z, Wang Z, Kong F, Tian D, Bao P, Liu R et al, (1993) Revelations of the origin of Chinese Nation from clustering analysis and frequency distribution of HLA polymorphism in major minority nationalities in mainland China. I Chuan Hsueh Pao 20:389–398

Decorte R, Cassiman JJ (1996) Evaluation of the ALF sequencer for high-speed sizing of short tandem repeat alleles. Electrophoresis 17:1542–1549

DNA recommendations – further report of the DNA commission of the ISFH regarding the use of short tandem repeat systems (1997). Forensic Sci Int 87:179–184

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

Füredi S, Budowle B, Woller J, Pádár Z (1996) Hungarian population data on six STR loci – HUMVWAA31, HUMTHO1, HUMCSF1PO, HUMFES/FPS, HUMTPOX, and HPRTB - derived using multiple PCR amplification and manual typing. Int J Legal Med 109:100–101

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

Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 55:175–189

Hou Y, Walter H (1996) Genetic substructure at the STR loci HUMTH01 and HUMVWA in Han populations, China. In: Carracedo A, Brinkmann B, Bär W (eds) Advances in forensic haemogenetics 6. Springer, Berlin Heidelberg New York, pp 468–470

Levene H (1949) On a matching problem arising in genetics. Ann Math Stat 20:91–94

Ohno Y, Sebetan IM, Akaishi S (1982) A simple method for calculating the probability of excluding paternity with any number of codominant alleles. Forensic Sci Int 19:93–98

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rousselet F, Pfitzinger H, Mangin P (1996) French Caucasian population data obtained from fluorescently detected HUMVWAA31/A and HumF13AO1 short tandem repeat loci. Int J Legal Med 109:5–9

Wall WJ, Williamson R, Petrou M, Papaioannou D, Parkin BH (1993) Variation of short tandem repeats within and between populations. Hum Mol Genet 2:1023–1029

Xiao FX, Gilissen A, Cassiman JJ, Decorte R (1998) A quadruplex fluorescent STR typing system (HUMVWA, HUMTH01, D21S11 and HPRT) with sequence-defined allelic ladders: identification of a new allele at D21S11. Forensic Sci Int, in press

Zhao T, Lee TD (1989) Gm and Km allotypes in 74 Chinese populations: a hypothesis of the origin of the Chinese nation. Hum Genet 83:101–110