RESEARCH NOTE

Allele frequencies of ten short tandem repeats loci in the central Tunisian human population

YOUSRA SOUIDEN*, KAMEL CHAIEB, OLFA CHEHAB and KACEM MAHDOUANI

Laboratoire de Biologie Moléculaire, Hôpital Ibn El Jazzar, Avenue Ibn El Jazzar, Kairouan 3140, Tunisia

Introduction

Different methodologies have been used in studies on genetic variability in human populations to assess their genetic composition and the evolutionary factors to which they are subjected, as well as for forensic purposes. Among these, the short tandem repeats (STRs) have been increasingly used in the last decade because of their high level of informativeness (Tautz 1989; Edwards *et al*. 1991; Gill *et al*. 1995; Brinkmann 1996; Pérez-Lezaun *et al*. 2000; Tomàs *et al*. 2000). The purpose of the present study was to report data on the STR frequencies in the population living in central Tunisia, for the application of these markers in forensic investigations and in population studies.

Allele frequencies for 10 short tandem repeats (STR) loci: *D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, THO1 and FGA* included in the AmpFℓSTR® SGM Plus^{TM} (Applied Biosystems, Foster City, USA) were estimated from a sample of unrelated individuals from central Tunisia. A deviation of the observed allele frequency from Hardy–Weinberg equilibrium (HWE) expectations was found only in loci *D18S51* and *D19S433*. Statistical parameters of forensic importance, the power of discrimination (PD), observed and expected heterozygosity values (H), polymorphism information content (PIC), matching probability (pM), power of exclusion (PE), and typical paternity index (PI), were calculated for the loci. These parameters indicated the usefulness of the loci in forensic personal identification and paternity testing among Tunisian.

Materials and methods

Population

Blood samples of 297 healthy unrelated individuals living in central Tunisia were collected after obtaining their informed consent (figure 1).

*For correspondence. E-mail: souidenyousra@yahoo.fr.

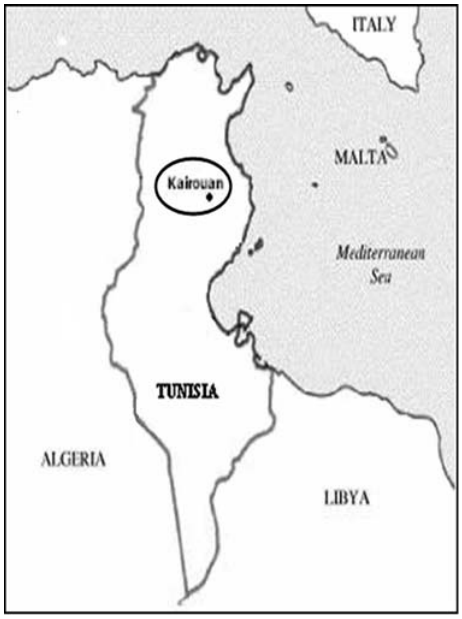

Figure 1. Region from which the population samples were collected.

DNA extraction

Genomic DNA was extracted from 300 μl of whole blood using the Wizard Genomic Purification A1120 kit (Promega, Madison, USA).

PCR amplification

Multiplexed PCR amplification of 10 STR loci: *D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, THO1* and *FGA* was performed using the AmpFℓSTR® SGM Plus^{TM} kit (Applied Biosystems, Foster City, USA).

Keywords. short tandem repeats; population data; DNA polymorphism; allele frequencies; human genetics.

PCR was carried out in a 25 μl volume using 2 ng template DNA; and performed according to manufacturer instructions using a Gene Amp PCR System 9700 thermocycler (Applied Biosystems, Foster City, USA) with the following amplification conditions: an initial incubation at 95°C for 11 min, 28 cycles of denaturation at 94°C for 1 min, annealing at 59°C for 1 min and extension at 72°C for 1 min; and a final extension at 60°C for 45 min.

Typing

PCR-amplified fragments were analysed with an ABI PRISM 310 Genetic Analyser (Applied Biosystems, Foster City, USA). Electrophoresis was performed using a 50 μm inner diameter, 47-cm length capillary and performance optimized polymer 4 (POP4) (Applied Biosystems, Foster City, USA) at 15 kV for 28 min at 60°C. Then the fragments were typed based on allelic ladders contained in the kit described above by GeneScan and Genotyper® Analysis Software packages (version 3.7; Applied Biosystems, Foster City, USA).

Statistical analysis

Statistical analysis was performed using the genetic data analysis GENEPOP (Raymond and Rousset 1995, http://wbiomed.curtin.edu.au/genepop/index.html) and PowerStat v12 (http://www.promega.com/geneticidtools) software packages. Results were considered to be significant at $P < 0.05$.

Results and discussion

In this study, we investigated the polymorphisms of 10 STR loci in the human population from the central province of Tunisia. As shown in table 2, in the genotypic distribution of these loci, a deviation from the HWE due to the lack of heterozygosity was noted only in loci *D18S51* (corrected $P = 0.0010$) and *D19S433* ($P = 0.0001$) after Bonferroni correction, based on the exact test. In fact, the studied population presents typical features of traditional Arabo–Islamic societies characterized by a high rate of endogamy and consanguinity ($F_{D18S51} = 7.2\%$; $F_{D19S433} = 6.1\%$).

Table 1 contains the allele frequency distributions of the 10 STR loci. The frequencies ranged from 0.002 to 0.330. Table 2 shows the polymorphism value of these STRs for forensic analysis and paternity testing as expressed by various statistical parameters. The observed heterozygosity (H_{obs}) ranged from 0.747 (*D3S1358*) to 0.838 (*D19S433*). The highest and the lowest values of PIC and PD were respectively observed in *D19S433* (PIC = 0.89; PD = 0.976) and *D16S539* (PIC = 0.73; PD = 0.904) contrary to those of pM (*D16S539* (pM = 0.096), *D19S433* (pM = 0.024)).

Table 1. Allele frequencies of 10 STR of central Tunisian population ($n = 297$).

Alleles	*D3S1358*	*vWA*	*D16S539*	*D2S1338*	*D8S1179*	*D21S11*	*D18S51*	*D19S433*	*THO1*	*FGA*
4									0.002	
5									0.007	
6									0.160	
7									0.229	
8			0.015		0.005				0.178	
9			0.113		0.002			0.002	0.326	
9.2					0.005			0.003		
9.3									0.076	
10			0.067		0.052		0.007	0.027	0.022	
11			0.291		0.118		0.008	0.052		
11.2					0.003			0.022		
12	0.003	0.002	0.330		0.141		0.086	0.078		
12.2					0.007			0.118		
12.3								0.002		
13	0.012	0.002	0.162		0.279		0.188	0.126		
13.2	0.007						0.002	0.120		
14	0.113	0.138	0.017		0.170		0.100	0.162		
14.2		0.002						0.081		
15	0.190	0.120		0.002	0.153		0.140	0.093		
15.2		0.003						0.054		
16	0.270	0.244		0.044	0.044		0.148	0.035		
16.2					0.002			0.015		
17	0.263	0.252	0.002	0.276	0.007		0.158	0.005		0.003
17.2	0.002							0.005		
18	0.118	0.133	0.003	0.116	0.005		0.072			0.008
18.2	0.002						0.003			0.002
19	0.020	0.087		0.150	0.007		0.049			0.037

Table 1 *(contd.)*

Alleles	*D3S1358*	*vWA*	*D16S539*	*D2S1338*	*D8S1179*	*D21S11*	*D18S51*	*D19S433*	*THO1*	*FGA*
19.2							0.003			0.002
20		0.017		0.167			0.018			0.128
20.2							0.002			
21				0.050			0.003			0.160
21.2										0.002
22				0.027			0.012			0.167
22.2										0.002
23				0.072			0.002			0.195
23.2				0.002						
24				0.056						0.155
24.2				0.003		0.002				
25				0.030						0.074
25.2										0.002
26				0.003						0.042
27				0.002		0.042				0.012
27.2						0.003				
28						0.094				0.005
29						0.214				0.003
29.2						0.002				
30						0.225				0.002
30.2						0.012				
31						0.034				
31.2						0.124				
32						0.002				
32.2						0.143				
33.2						0.069				
34						0.005				
34.2						0.010				
35						0.012				
35.2						0.005				
36						0.002				

Table 2. Statistical parameters of the 10 STR loci for forensic interest (n = 297).

	D3S1358	*vWA*	*D16S539*	*D2S1338*	*D8S1179*	*D21S11*	*D18S51*	*D19S433*	*THO1*	*FGA*
H_{obs}	0.747	0.757	0.754	0.791	0.801	0.831	0.801	0.838	0.754	0.828
H_{exp}	0.795	0.817	0.762	0.846	0.831	0.850	0.873	0.899	0.777	0.859
PD	0.927	0.941	0.904	0.960	0.949	0.957	0.968	0.976	0.918	0.961
pM	0.073	0.059	0.096	0.040	0.051	0.043	0.032	0.024	0.082	0.039
PIC	0.76	0.79	0.73	0.83	0.81	0.83	0.86	0.89	0.74	0.84
PE	0.505	0.523	0.517	0.582	0.602	0.658	0.595	0.671	0.517	0.653
PIT	1.98	2.06	2.03	2.39	2.52	2.96	2.48	3.08	2.03	2.91
P	0.0764	0.0090	0.5189	0.1897	0.0134	0.0414	0.0010	0.0000	0.3144	0.0116
P_{comp1}	0.0000484	0.01522	0.00044	0.000089	8.9×10^{-5}	2.6×10^{-7}	6×10^{-7}	10^{-7}	2×10^{-7}	0.08558
P_{comp2}	7.99×10^{-6}	0.01902	0.05107	0.01379	0.000062	8×10^{-7}	0.00002	10^{-7}	4×10^{-7}	0.62845

H_{obs}, observed heterozygosity; H_{exp}, expected heterozygosity; PD, power of discrimination; pM, matching probability; PIC, polymorphism information content; PE, power of exclusion; PIT, typical paternity index; P, Hardy–Weinberg equilibrium (significant difference is noted for $0.01 < P < 0.05$); P_{comp1}, exact test of population stratification between five populations; P_{comp2}, exact test of population stratification between three Tunisian populations.

The 10 loci showed a combined pM and PD of $8.908086126 \times 10^{-14}$ and 0.99999999999878 respectively, and a cumulative power of exclusion at 0.99986. Data comparison between other Arabian (Syria and Morocco; Abdin *et al*. 2003) and Tunisian populations (Mahfoudh-Lahiani *et al*. 2006; Brandt-Casadevall *et al*. 2002), revealed significant differences in all markers (table 2).

However, when comparing the three Tunisian samples (our data and those of (Mahfoudh-Lahiani *et al*. 2006) and Brandt-Casadevall *et al*. (2002), we found significant differences in all markers except *FGA*($P = 0.62845$) (table 2). These discrepancies in frequency were observed mainly for alleles 14 of *D3S1358* (0.113 versus 0.064) and *vWA* (0.138 versus 0.085), 8 (0.015 versus 0.035) of *D16S539*, 10 (0.052 versus 0.105) and 12 (0.141 versus 0.085) of *D8S1178*, 31.2 (0.124 versus 0.077) of *D21S11*, 12 (0.086 versus 0.136) and 17 (0.158 versus 0.079) of *D18S51* and 9.3 (0.076 versus 0.149) of *THO1*.

In the studied population, we note the presence of specific alleles for the majority of the markers: *D3S1358* (12, 13.2, 17.2, 18.2), *vWA* (12, 13, 14.2, 15.2), *D16S539* (17, 18), *D8S1179* (9.2, 11.2, 12.2, 16.2, 19), *D21S11* (24.2, 27.2, 29.2, 35.2), *D18S51* (13.2, 18.2, 19.2, 20.2), *THO1* (4, 5) and *FGA* (18.2, 25.2, 29). This probably reflects a specificity of this continental area economically based on agriculture, compared to other neighbouring areas and could indicate alleles characterizing of the particular zones or even genetic isolates. In conclusion, the present data can be a good preparation for the preliminary construction of the local DNA database of Tunisian population.

Acknowledgements

This work is supported by the Tunisian Ministry of Health. We are grateful to Samia Fatnassi and Mustapha Romdhani for their help.

References

Abdin L., Shimada I., Brinkmann B. and Hohoff C. 2003 Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria. *Legal Med*. **5**, S150–S155.

Brandt-Casadevall C., Ben Dhiab M., Taroni F., Gehrig C., Dimo-Simonin N., Zemni M. and Mangin P. 2002 Tunisian population allele frequencies for 15 PCR-based loci. *Forensic Sci. Int*. **126**, 272–274.

Brinkmann B. 1996 The STR approach. In *Advances in forensic haemogenetics 6* (ed. A. Carracedo, B. Brinkmann and W. Bar), pp. 41–51. Springer, Berlin.

Edwards A., Citivello A., Hammond H. A. and Caskey C. T. 1991 DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet*. **49**, 746–756.

Gill P., Kimpton C. P., Urquhuart A., Oldroyd N., Millican E. S., Watson S. K. and Downer T. J. 1995 Automated short tandem repeat (STR) analysis in forensic casework: a strategy for the future. *Electrophoresis* **16**, 1543–1552.

Mahfoudh-Lahiani N., Rebaï A. and Makni H. 2006 Data for 10 autosomal STR markers in south Tunisian population. *Forensic Sci. Int*. **164**, 254–256.

Pérez-Lezaun A., Calafell F., Clarimón J., Bosch E., Mateu E., Gusmão L. *et al*. 2000 Allele frequencies of 13 short tandem repeats in population samples of the Iberian Peninsula and Northern Africa. *Int. J. Legal Med*. **113**, 208–214.

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

Tautz D. 1989 Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res*. **17**, 6464–6471.

Tomàs C., Picornell A., Castro J. A., Ramon M. M., Gusmão L., Lareu M. V. and Carracedo A. 2000 Genetic variability at nine STR loci in the Chueta (Majorcan Jews) and the Balearic populations investigated by a single multiplex reaction. *Int. J. Legal Med*. **113**, 263–267.

Received 1 September 2008, in revised form 26 November 2008; accepted 30 November 2008
Published on the Web: 3 April 2009