VIA MEDICA

SHORT COMMUNICATION

Vol. 63, No. 2, pp. 249–252

ISSN 0015–5659
www.fm.viamedica.pl

Population genetics of 10 short tandem repeat (STR) loci in a population sample of the ethnic group of Polish Tatars living in the Podlasie area (Northeastern Poland)

Witold Pepiński[1], Jerzy Janica[1], Maria Aleksandrowicz-Bukin[2], Małgorzata Skawrońska[1], Ewa Koc-Zórawska[1], Anna Niemcunowicz-Janica[1]

[1]Department of Forensic Medicine, Medical University, Białystok, Poland
[2]Department of Human Anatomy, Medical University, Białystok, Poland

[Received 28 May 2003; Accepted 25 March 2004]

Laboratories worldwide are contributing to a large and growing database for different populations. This study provides a 10 STR database for a population sample of Polish Tatars living in the area of Podlasie for the use as a highly discriminatory system of genetic markers in the forensic community. The genotype frequency distributions showed no deviations from the Hardy-Weinberg equilibrium (HWE) except for D3S1358, FGA, D18S51 and D16S539, based on the Fisher Exact Test. Significant differences between the Polish Tatars and the native population of Podlasie were found in loci D3S1358, FGA, D2S1338, D21S11 and D19S433. The combined values of the Matching Probability and of the Power of Exclusion are 1 in 2.83 × 10^{-12} and 0.998, respectively.

Key words: STR, Multiplex PCR, AmpFlSTR SGM Plus

INTRODUCTION

The ancestors of the contemporary Polish Tatars (also known as Polish-Lithuanian Tatars) arrived in Poland in ancient times from the Golden Horde and Khanates of Kazan, Crimea, Volga and Astrakhan. They are descendants of the Nogai Ordu (Horde) to whom Grand Duke Vitautas applied for assistance in his struggle against the Teutonic Order. After the victory at Grunwald (1410) they were invited to settle in Lithuania. At that time they numbered about 200,000 and built around 260 mosques [3]. The distinctive culture of the Polish Tatars is a result of a few hundred years of contact between Muslims and Muslim Slavs. Living in separate closed communities, they long preserved their language, faith, and customs. Greater assimilation with the local population only began in the 17th century. The villages of that time, Kruszyniany and Bohoniki, together with the oldest mosques, still exist today. The autochtonous Tatars who live in Poland today are Sunni Muslim descendants of those settled from 1679 by the Polish King Jan III Sobieski. They number about 2,500 and all speak Polish. The Islamic centre of the Polish Tatars is Białystok in Podlasie (NE Poland). Despite several centuries of cultural and religious assimilation and separation from their Tatar-Turkic roots, the Polish Tatars have preserved their ethnic and religious identity [1].

Short tandem repeats (STRs) represent highly polymorphic microsatellite markers in the human

Address for correspondence: Witold Pepiński, MD, Department of Forensic Medicine, Medical University, ul. Waszyngtona 13, 15–230 Białystok, Poland, tel/fax: +48 85 748 59 85, e-mail: pepinski@amb.edu.pl

genome that have tandemly repetitive sequence elements of 2–7 bps in length, located approximately every 10–15 kbs. Multiplex PCR-based STR and kits with fluorescence detection technology have been validated to produce rapid and robust amplification of several DNA loci from biological samples [5, 7]. As a result of the different frequency distributions of the particular STR markers reported for different populations, population data are collected to make a frequency assessment for each observed allele and genotype so that the probability value describing a random genotype occurrence in a population can be determined [5]. The aim of this study was to provide a preliminary 10 STR database for a population sample of Polish Tatars living in the area of Podlasie (NE Poland).

MATERIAL AND METHODS

DNA was extracted with the use of the Chelex-100 and proteinase K method [9] from the buccal swabs of 120 unrelated Polish Tatars resident in Podlasie. All the persons gave their signed written consent for blood samples to be collected. The amount of DNA was determined spectrophotometrically. The commercially available kit AmpFlSTR SGM Plus (Applera, USA) was used according to the manufacturer's instructions. DNA samples were amplified using GeneAmp PCR System 9700 (Applera, USA). Genotyping was performed in a 310 ABI Genetic Analyzer (Applera, USA). Possible divergence from the Hardy-Weinberg equilibrium was tested using the exact test [2] based on 3200 shuffling experiments. Analyses were performed using the Genetic Data Analysis (GDA) software [4]. A comparison of allele frequency distributions was performed by means of a pairwise population comparison test (RxC contingency test; G. Carmody, Ottawa, Canada). The level of significance was 0.05 for all statistical tests (two-sided probability). For the calculations of the forensic efficiency parameters PowerStats software (Promega Corp.) was used. The Matching Probability (MP) displays the rate of coincidence that a random unrelated person would by chance have the same DNA profile as that obtained from the evidence or, alternatively, the number of individuals that may be surveyed before finding the same DNA profile in a randomly encountered individual. The combined MP for several loci is the product of the individual matching probability at each locus on the assumption that they are not linked. Mean Exclusion Probability, also termed the Power of Exclusion (PE), is the probability that a single genetic system will exclude a non-father from paternity prior to testing and is defined as the fraction of individuals with a DNA profile different from that of a randomly selected individual in a typical paternity case.

RESULTS AND DISCUSSION

The allele frequencies observed for the 10 short tandem repeat loci and the values of the forensic efficiency parameters determined for 120 unrelated individuals from the area of Podlasie are displayed in Table 1. The genotype frequency distributions showed no deviations from the Hardy-Weinberg equilibrium (HWE) except for D3S1358, FGA, D18S51 and D16S539, based on the Fisher Exact Test. Possible reasons for the disequilibrium include a possible sampling error, a population substructure or inbreeding [6]. Table 1 also shows the forensic efficiency parameters for the systems analysed. The calculated MP ranged from 0.117 (D3S1358) to 0.042 (D2S1338 and D18S51) and the a priori PE ranged from 0.400 (D3S1358) to 0.758 (D2S1338). The Combined PE calculated for the 10 loci was 0.998 and the Combined MP was 1 in 2.83×10^{-12}. Tatars living outside Tatarstan (about 70%) are in some danger of assimilation, although in Poland they remained isolated until the 18th century. A pairwise testing for heterogeneity using the χ^2-test according to Carmody revealed statistically significant differences between the Polish Tatar and Podlasie populations for loci D3S1358, FGA, D2S1338, D21S11 and D19S433 ($p < 0.05$) and similar allele distributions for VWA, TH01, D16S539, D18S51 and D8S1179 ($p > 0.05$) (Table 2). The 10 loci studied provide an extremely informative discriminatory system of genetic markers for applications in population genetics and in the forensic community.

REFERENCES

1. Aleksandrowicz-Bukin M (2002) Grupy odmienności kulturowej i religijnej na Podlasiu (Płn-Wsch. Polska) ze szczególnym uwzględnieniem Tatarów Polskich w przekazach Telewizji Regionalnej. UwB, Białystok.
2. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics, 48: 361–72.
3. Kalendarz Muzułmański Tatarów w Polsce (1939) Wilno.
4. Lewis PO, Zaykin D. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html

Table 1. Distribution of the observed allele frequencies and HWE exact test (P-values) for the 10 STR loci in the Polish Tatar population (NE Poland)

Allele	D3S1358	VWA	FGA	TH01	D16S539	D2S1338	D21S11	D18S51	D8S1179	D19S433
6	–	–	–	0.218	–	–	–	–	–	–
7	–	–	–	0.127	–	–	–	–	–	–
8	–	–	–	0.091	0.004	–	–	–	–	–
9	–	–	–	0.305	0.096	–	–	–	0.013	–
9.3	–	–	–	0.259	–	–	–	–	–	–
10	–	–	–	–	0.088	–	–	–	0.053	–
11	–	–	–	–	0.289	–	–	0.009	0.049	–
12	–	–	–	–	0.289	–	–	0.063	0.243	0.058
12.2	–	–	–	–	–	–	–	–	–	0.004
13	–	–	–	–	0.184	0.005	–	0.112	0.363	0.299
13.2	–	–	–	–	–	–	–	–	–	0.045
14	0.049	0.062	–	–	0.048	–	–	0.152	0.190	0.357
14.2	–	–	–	–	–	–	–	–	–	0.054
15	0.332	0.164	–	–	–	–	–	0.241	0.066	0.076
15.2	0.009	–	–	–	–	–	–	–	–	0.022
16	0.296	0.243	–	–	–	0.009	–	0.161	0.018	0.049
16.2	–	–	–	–	–	–	–	–	–	0.036
17	0.190	0.279	–	–	–	0.095	–	0.080	0.004	–
18	0.124	0.181	–	–	–	0.050	–	0.063	–	–
19	–	0.071	0.040	–	–	0.232	–	0.071	–	–
20	–	–	0.076	–	–	0.155	–	0.009	–	–
21	–	–	0.179	–	–	0.068	–	0.036	–	–
21.2	–	–	0.009	–	–	–	–	–	–	–
22	–	–	0.250	–	–	0.041	–	0.004	–	–
22.2	–	–	0.009	–	–	–	–	–	–	
23	–	–	0.129	–	–	0.127	–	–	–	–
23.2	–	–	0.018	–	–	–	–	–	–	–
24	–	–	0.147	–	–	0.155	–	–	–	–
24.2	–	–	0.013	–	–	–	–	–	–	–
25	–	–	0.085	–	–	0.064	–	–	–	–
26	–	–	0.022	–	–	–	–	–	–	–
27	–	–	0.014	–	–	–	0.090	–	–	–
28	–	–	–	–	–	–	0.080	–	–	–
29	–	–	–	–	–	–	0.212	–	–	–
30	–	–	0.022	–	–	–	0.301	–	–	–
30.2	–	–	–	–	–	–	0.031	–	–	–
31	–	–	–	–	–	–	0.088	–	–	–
31.2	–	–	–	–	–	–	0.062	–	–	–
32	–	–	–	–	–	–	0.022	–	–	–
32.2	–	–	–	–	–	–	0.124	–	–	–
33.2	–	–	–	–	–	–	0.062	–	–	–
34.2	–	–	–	–	–	–	0.009	–	–	–
Ho	0.681	0.779	0.875	0.727	0.702	0.882	0.779	0.804	0.708	0.777
He	0.751	0.798	0.885	0.772	0.783	0.864	0.829	0.864	0.766	0.768
P	0.006	0.515	0.027	0.243	0.003	0.705	0.128	0.024	0.538	0.413
MP	0.117	0.073	0.050	0.093	0.089	0.042	0.053	0.042	0.092	0.093
PE	0.400	0.560	0.745	0.472	0.431	0.758	0.560	0.606	0.441	0.557

Ho — heterozygosity observed, He — heterozygosity expected, P — exact test probability, MP — matching probability, PE — power of exclusion

Table 2. Pairwise comparison (RxC contingency table) between the Polish Tatar population and the population of Podlasie [8]

Locus	χ^2	P	G	P
D3S1358	15.3835	0.0100 ± 0.0031	17.7761	0.0050 ± 0.0022
VWA	10.1569	0.1740 ± 0.0120	12.3268	0.0960 ± 0.0093
FGA	33.7700	0.0060 ± 0.0024	25.8434	0.0100 ± 0.0031
TH01	7.8746	0.1660 ± 0.0118	8.3321	0.1490 ± 0.0113
D16S539	11.6615	0.1130 ± 0.0100	9.3856	0.2280 ± 0.0133
D2S1338	31.3029	0.0010 ± 0.0010	33.4794	0.0000 ± 0.0000
D21S11	19.7074	0.0320 ± 0.0056	19.8174	0.0350 ± 0.0058
D18S51	12.0162	0.5690 ± 0.0157	13.1162	0.5610 ± 0.0157
D8S1179	9.1799	0.2490 ± 0.0137	9.2609	0.2690 ± 0.0140
D19S433	23.2321	0.0100 ± 0.0031	26.7138	0.0020 ± 0.0014

χ^2 — Chi square test, G — G statistics, P — probability

5. Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B (2001) Validation of STR typing by capillary electrophoresis. J Forensic Sci, 46: 661–676.
6. National Research Council (1966) The Evaluation of Forensic DNA Evidence, National Academy Press, Washington, DC.
7. Parson W, Steinlechner M (2001) Efficient DNA database laboratory strategy for high throughput STR typing of reference samples. Forensic Sci Int, 22: 1–6.
8. Pepinski W, Janica J, Skawronska M, Niemcunowicz-Janica A, Soltyszewski I (2001) Population genetics of 15 STR loci in the population of Podlasie (NE Poland). Forensic Sci Int, 124: 226–227.
9. Wiegand P, Bajanowski T, Brinkmann B (1993) PCR typing of debris from fingernails Int J Legal Med, 106: 81–84.