Legal Medicine 12 (2010) 256–258

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

Announcement of Population Data

Genetic variability of 15 autosomal STR loci in Russian populations

Vadim A. Stepanov [a,*], Alexander V. Melnikov [b], Andrey Yu. Lash-Zavada [b], Vladimir N. Kharkov [a], Svetlana A. Borinskaya [c], Tatiana V. Tyazhelova [c], Olga V. Zhukova [c], Yuri V. Schneider [c], Irina N. Shil'nikova [c], Valery P. Puzyrev [a], Anna A. Rybakova [b], Nikolai K. Yankovsky [c]

[a] Institute for Medical Genetics, Russian Academy of Medical Sciences, Nab. Ushayky 10, 634050, Tomsk, Russia
[b] Forensic Center, Ministry of Interior of Russian Federation, ul. A. i Z. Kosmodemyanskikh 5, 125130 Moscow, Russia
[c] Vavilov Institute of General Genetics, Russian Academy of Sciences, ul. Gubkina 3, 119991 Moscow, Russia

ARTICLE INFO

Article history:
Received 27 February 2010
Received in revised form 30 April 2010
Accepted 12 May 2010
Available online 13 July 2010

Keywords:
Short tandem repeats (STR)
Population data
PowerPlex® 16 System
Russian population

ABSTRACT

Allele frequencies for 15 STRs (CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, THO1, TPOX, and vWA) in the PowerPlex 16 System (Promega Corporation) were assessed in 386 individuals from five Russian urban populations. No significant between-population differences in frequencies and molecular variance of 15 microsatellites were revealed. For all 15 loci, the combined matching probability is 3.19×10^{-18} and the power of exclusion is 99.99989%.

Population: Blood samples were obtained from 386 unrelated Russian individuals from 5 cities located in different part of Russian Federation (Fig. 1) and representing the general Russian urban population. The locations are Belgorod ($N = 50$) and Orel ($N = 51$), situated in the southern part of European Russia; Yaroslavl ($N = 50$) located in the northern part; Orenburg ($N = 50$) in the Ural region; and Tomsk ($N = 185$) in Siberia. Voluntary informed consent was obtained from all participants in the study, as well as, the information of their ethnic background.

Population of the cities under study varies from 333600 in Orel to 613200 in Yaroslavl according to the last population census of 2002 [1]. The majority of population is represented by Russians: 80% in Orenburg and 91–95% in other cities. The distinctive demographic feature of Tomsk and Belgorod is the intensive gene flow from outside of the region. Thus, in Tomsk only 32% of the current residents were born in the city [2]. Marriage pattern in these two cities is characterized by the high portion of marriages between individuals born outside the city of their current residence. Only about 45% of the marriages in Tomsk and about 69% in Belgorod were contracted between locally born spouses [2,3].

DNA extraction: DNA was extracted form peripheral blood lymphocytes by sodium perchlorate or by salt-based extraction methods [4,5].

PCR: Amplification was performed using a PowerPlex® 16 System Kit (Promega Corporation) following the conditions recommended by the manufacturer.

Typing: Multiplex PCR samples were separated by capillary gel electrophoresis on ABI Prism 3130 and 310 Genetic Analyzers (Applied Biosystems). Genotyping was performed using GeneMapper v. 3.7 software (Applied Biosystems).

Results: Allelic frequency and statistical parameters of the 15 STR loci in the overall sample of 386 Russian individuals are presented in Table 1. Distributions of allelic frequency with statistical parameters of forensic interest in each of 5 urban populations are provided in Supplementary Tables 1–5.

Quality control: Laboratory internal control standards and kit controls.

Data analysis: Correspondence of genotype distributions to Hardy–Weinberg equilibrium was estimated by exact test of Guo and Thompson [6] implemented in the Arlequin v. 3.11 software [7]. Genetic structure analysis and pairwise population comparisons were performed by analysis of molecular variance (AMOVA) method. Matrix of genetic distances between populations based on sum of squared differences in repeat numbers (R_{ST}) was used in AMOVA. Allelic and genotype frequency, matching probability (MP), power of discrimination (PD), probability of paternity exclusion (PE), paternity index (PI) and polymorphic information content (PIC) were calculated with the PowerStats v. 1.2 software [8].

Access to the data: The data are available upon request.

* Corresponding author.
E-mail address: vadim.stepanov@medgenetics.ru (V.A. Stepanov).

1344-6223/$ - see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.legalmed.2010.05.006

Fig. 1. Map showing the location of the 5 population samples analyzed in this study.

Table 1
Allele frequencies for 15 STR loci in the overall Russian population ($N = 386$).

Allele	CSF1PO	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51	D21S11	FGA	PentaD	PentaE	TH01	TPOX	vWA
4	–	–	–	–	–	–	–	–	–	–	–	–	0.004	–	–
5	–	–	–	0.001	–	–	–	–	–	–	–	0.056	0.001	–	–
6	–	–	–	0.001	–	–	–	–	–	–	0.003	–	0.226	–	–
7	–	–	0.013	0.012	–	–	–	–	–	–	0.001	0.150	0.152	0.001	–
8	–	–	0.003	0.161	0.005	0.149	0.006	0.001	–	–	0.009	0.006	0.097	0.558	–
9	0.052	–	0.044	0.137	0.005	0.084	0.073	0.003	–	–	0.256	0.012	0.186	0.100	–
9.3	–	–	–	–	–	–	–	–	–	–	–	–	0.329	–	–
10	0.260	–	0.101	0.303	0.063	0.079	0.070	0.005	–	–	0.153	0.124	0.005	0.071	–
11	0.278	–	0.316	0.227	0.080	0.345	0.277	0.018	–	–	0.154	0.082	–	0.237	–
11.2	–	–	–	–	–	–	–	–	–	–	–	0.144	–	–	–
12	0.341	–	0.354	0.124	0.193	0.205	0.316	0.082	–	–	0.185	–	–	0.031	–
13	0.053	0.010	0.152	0.027	0.332	0.080	0.214	0.109	–	–	0.157	0.092	–	–	0.003
13.2	–	–	–	–	–	–	–	0.003	–	–	–	0.071	–	0.001	–
14	0.013	0.115	0.017	0.005	0.192	0.041	0.040	0.149	–	–	0.066	–	–	–	0.074
15	0.003	0.297	0.001	0.001	0.096	0.016	0.004	0.177	–	–	0.012	0.071	–	–	0.100
16	–	0.247	–	–	0.028	0.001	–	0.163	–	0.001	0.004	0.060	–	–	0.193
17	–	0.199	–	–	0.005	–	–	0.124	–	0.017	–	0.048	–	–	0.293
18	–	0.124	–	–	–	–	–	0.083	–	0.063	–	0.047	–	–	0.241
19	–	0.005	–	–	–	–	–	0.040	–	0.123	–	0.021	–	–	0.082
20	–	0.001	–	–	–	–	–	0.027	–	0.165	–	0.010	–	–	0.014
21	–	–	–	–	–	–	–	0.006	–	0.005	–	0.005	–	–	0.001
21.2	–	–	–	–	–	–	–	–	–	0.247	–	–	–	–	–
22	–	–	–	–	–	–	–	0.006	–	0.006	–	0.001	–	–	–
22.2	–	–	–	–	–	–	–	–	–	0.136	–	–	–	–	–
23	–	–	–	–	–	–	–	0.003	–	0.008	–	–	–	–	–
23.2	–	–	–	–	–	–	–	–	–	0.122	–	–	–	–	–
24	–	–	–	–	–	–	–	–	–	0.083	–	–	–	–	–
25	–	–	–	–	–	–	–	–	–	0.019	–	–	–	–	–
26	–	–	–	–	–	–	–	–	0.004	0.003	–	–	–	–	–
27	–	–	–	–	–	–	–	–	0.019	0.001	–	–	–	–	–
28	–	–	–	–	–	–	–	–	0.175	0.001	–	–	–	–	–
28.2	–	–	–	–	–	–	–	–	0.001	–	–	–	–	–	–
29	–	–	–	–	–	–	–	–	0.181	–	–	–	–	–	–
29.2	–	–	–	–	–	–	–	–	0.010	–	–	–	–	–	–
30	–	–	–	–	–	–	–	–	0.242	–	–	–	–	–	–
30.2	–	–	–	–	–	–	–	–	0.079	–	–	–	–	–	–
31	–	–	–	–	–	–	–	–	0.067	–	–	–	–	–	–
31.2	–	–	–	–	–	–	–	–	0.070	–	–	–	–	–	–
32	–	–	–	–	–	–	–	–	0.009	–	–	–	–	–	–

(*continued on next page*)

Table 1 (*continued*)

Allele	CSF1PO	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51	D21S11	FGA	PentaD	PentaE	TH01	TPOX	vWA
32.2	–	–	–	–	–	–	–	–	0.104	–	–	–	–	–	–
33	–	–	–	–	–	–	–	–	0.003	–	–	–	–	–	–
33.2	–	–	–	–	–	–	–	–	0.028	–	–	–	–	–	–
34	–	–	–	–	–	–	–	–	0.004	–	–	–	–	–	–
34.2	–	–	–	–	–	–	–	–	0.003	–	–	–	–	–	–
H_0	0.736	0.790	0.744	0.834	0.746	0.808	0.744	0.876	0.834	0.855	0.788	0.899	0.792	0.601	0.767
H_E	0.734	0.783	0.744	0.797	0.796	0.796	0.767	0.877	0.851	0.853	0.825	0.905	0.775	0.617	0.798
PD	0.876	0.916	0.891	0.920	0.928	0.931	0.907	0.970	0.956	0.958	0.945	0.980	0.909	0.798	0.932
PE	0.486	0.581	0.499	0.664	0.503	0.615	0.499	0.746	0.664	0.705	0.576	0.793	0.585	0.292	0.539
P	0.100	0.070	0.625	0.030	$<10^{-3}$	0.301	0.176	0.643	$<10^{-3}$	0.062	0.859	0.092	0.007	0.038	0.052

H_0, observed heterozygocity; H_E, expected heterozygocity; PD, power of discrimination; PE, power of exclusion; P, *p*-values for Hardy–Weinberg expectations (exact tests, 10,000 runs).

Table 2
Population pairwise F_{ST} indexes.

	Belgorod	Orel	Yaroslavl	Orenburg	Tomsk
Belgorod	x	0.48649	0.85586	0.63063	0.54955
Orel	−0.00235	x	0.36036	0.64865	0.44144
Yaroslavl	−0.00423	−0.00084	x	0.77477	0.27928
Orenburg	−0.00170	−0.00293	−0.00423	x	0.36937
Tomsk	−0.00129	−0.00055	0.00136	−0.00020	x

F_{ST} values are shown in the lower triangle of the matrix, and corresponding *p*-values in the upper.

Other remarks: Among 75 genotype distributions (15 loci per 5 populations), 9 significant deviations from Hardy–Weinberg equilibrium were observed versus 4 that would be expected in case of significance level of 0.05. It is important to note that 5 out of 9 significant deviations were found in population of Tomsk, whereas only one significant deviation from Hardy–Weinberg proportions was observed in each of other 4 Russian urban populations. The possible explanation of such a high level of possible disequilibrium in population of Tomsk is the "mixed" nature of this population which is characterized by high genetic diversity and continuous gene flow from other part of Russia according to earlier genetic demographic research [2,9]. This argument is supported indirectly by the fact that majority of Hardy—Weinberg disequilibrium in Tomsk population is due to the excess of heterozygotes. None of the 9 deviations remained significant after conservative Bonferroni correction, which reduces the corrected *p*-value threshold for this dataset to 0.0007.

The average level of genetic diversity (expected heterozygosity) for 15 STRs in overall population under study was 0.794. Most variable loci were D21S11, D18S51, PentaE, and FGA, having 15 or more alleles each and expected heterozygosity more than 85%. TPOX proved to be the least polymorphic marker (7 alleles, $H_E = 0.616$). The similar level of genetic diversity was shown for all 5 populations: average population heterozygosity varied from 77.9% in Orel to 79.7% in Tomsk.

The AMOVA analysis reveals no significant between-population differences in frequencies and molecular variance of 15 microsatellites in Russian populations (Table 2). Moreover, the total F_{ST} index has small negative value (−0.00088, $p = 0.6099$) indicating that on average, individuals within particular population differs rather more from each other than individuals from different populations. Most pairwise F_{ST} also are negative.

Taking into account the total absence of genetic subdivision of Russian urban populations we combined allele frequencies observed in 5 studied cities. In the pooled sample of 386 individuals, 5 loci were in Hardy–Weinberg disequilibrium, and deviations from Hardy—Weinberg expectations in D8S1179 and D21S11 loci remained significant even after Bonferroni correction (corrected *p*-value is 0.05/15 = 0.0033). However, departures from Hardy—Weinberg expectations are expected, particularly as the dataset combines populations from various geographic regions.

Very high values of discriminating power of the PowerPlex 16 system were obtained in all populations under study. In the overall Russian sample the combined power of discrimination (*PD*) and combined power of exclusion (*PE*) in the 15 loci studied were 0.999999999999999997 and 0.999998, respectively.

In summary, the database for 15 forensic STR loci for Russian population was established in this research. This paper follows the guidelines for publication of population data proposed by the journal [10].

Acknowledgements

This study was supported by the grant 07-04-01749 from the Russian Foundation of Basic Research. The authors are grateful to Promega Corporation for the gift of PowerPlex16 kit used in the study and to Dr. Olga Kurbatova for helpful remarks.

References

[1] All-Russian population census 2002 official website http://www.perepis2002.ru/.
[2] Kucher AN, Ivanova OF, Puzyrev VP, Tsymbaliuk IA, Trotsenko VA. Genetic demographic characterization of a modern Siberian town (exemplified by Tomsk). Genetika (Moscow) 1994;30:276–81.
[3] Artamentova LA, Filiptsova OV. Spatial characteristics of marriage migration in the Belgorod Population. Rus J Genetics 2005;41:686–96.
[4] Johns MB, Pauls–Thomas JE. Purification of human genomic DNA from whole blood using sodium perchlorate in place of phenol. Anal. Biochem. 1989;180:276–8.
[5] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16(1215).
[6] Guo S, Thompson E. Performing the exact test of Hardy–Weinberg proportions for multiple alleles. Biometrics 1992;48:361–72.
[7] Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 2005;1:47–50.
[8] Powerstats Version 1.2, Promega Corporation Website: http://www.promega.com/geneticidtools/powerstats/.
[9] Stepanov VA, Puzyrev VP, Spiridonova MG, Kitrinskaya IYu. Analysis of the Alu insertion polymorphism in urban and rural Russian populations of Siberia. Rus J Genet 1999;35:979–83.
[10] Lincoln P, Carracedo A. Publication of population data of human polymorphisms. Forensic Sci Int 2000;110:3–5.