



Evaluation of sensitivity and specificity of sibship determination in the Caucasian population of the Russian Federation using the 23 STR loci VeriFiler panel

V. Zavarin*, V. Ilina, Y. Krassotkin, T. Makarova, D. Sutiagina, A. Semikhodskii

LLC Medical Genomics, 48 Zhelyabova St, Tver, 170100, Russian Federation

ARTICLE INFO

Keywords:

Kinship testing
Russian Federation
Likelihood ratio

ABSTRACT

We performed a simulation study to evaluate sensitivity and specificity of sibship determination in the Caucasian population of the Russian Federation using 23 autosomal STR loci included in the VeriFiler Express (Applied Biosystems) kit. Population genetic data were obtained from 1609 unrelated Caucasian individuals from different regions of the Russian Federation. Several scenarios were analyzed: half siblings duo vs unrelated; half siblings vs unrelated (mother(s) will be genotyped); full siblings duo vs half siblings; full siblings vs half siblings (mother will be genotyped); full siblings duo vs unrelated.

1. Introduction

The 13th edition of Standards for Relationship Testing Laboratories [1] contains a new requirement to report “the estimate of the percentage of individuals of known relationship that may have a combined likelihood ratio that is inconclusive, or supportive, or not supportive of the tested relationship” for two-party comparisons of full sibling, half sibling, avuncular, and single grandparent.

We performed a simulation study to evaluate sensitivity and specificity of sibship determination in the Caucasian population of the Russian Federation using 23 autosomal STR loci included in the VeriFiler Express (Applied Biosystems) kit (D3S1358, vWA, D16S539, CSF1PO, TPOX, D8S1179, D21S11, D18S51, D2S441, D19S433, TH01, FGA, D22S1045, D5S818, D13S317, D7S820, D10S1248, D1S1656, D12S391, D2S1338, D6S1043, Penta D, Penta E).

2. Materials and methods

Population genetic data were obtained from 1609 unrelated Caucasian individuals from different regions of the Russian Federation (admixed urban population). Allelic frequencies, forensic and population genetic parameters were determined using the STRAF v1.0.5 software package [2]. For F-statistics, 8 regional subpopulations were defined: Altai Area (198 samples), Amur Region (116 samples), Belgorod Region (130 samples), Bryansk Region (276 samples), Vladimir Region (114 samples), Volgograd Region (424 samples), Ivanovo

Region (257 samples), and Kursk Region (94 samples). All selected subgroups included at least 90 individuals each.

The simulation study was performed using the Familias v3.2.7 software package [3]. Different scenarios were analyzed: half siblings duo vs unrelated; half siblings vs unrelated (mother(s) will be genotyped); full siblings duo vs unrelated; full siblings duo vs half siblings; full siblings vs half siblings (mother will be genotyped). Genotypic configurations were simulated assuming silent allele frequency 0.005 at each locus. Mutations were not considered. Percentage of true and false positives was determined using the likelihood ratio (LR) threshold of 10, percentage of true and false negatives was determined using the LR threshold of 0.1 [1].

3. Results and discussion

Population genetic parameters of our dataset relevant to forensics and kinship testing are given in Table 1.

All loci are in the Hardy-Weinberg equilibrium.

Table 2 presents results of simulations for each scenario tested (100,000 simulations for each combination of the true hypothesis and the number of testing participants).

4. Conclusions

For half siblings duo, simulating the true relationship, 73.576% of the 100,000 simulations were equal or above the LR limit, 1.355% of

* Corresponding author at: LLC Medical Genomics, 48 Zhelyabova St, Tver, 170100, Russian Federation.

E-mail address: v.zavarin@medicalgenomics.ru (V. Zavarin).

<https://doi.org/10.1016/j.fsigss.2019.09.023>

Received 29 August 2019; Accepted 21 September 2019

Available online 23 September 2019

1875-1768/ © 2019 Elsevier B.V. All rights reserved.

Table 1

Population genetic data and forensic parameters of the 23 STR loci Verifiler panel in the Caucasian population of the Russian Federation (n = 1609). N_{all} – number of different alleles observed; GD – gene diversity (expected heterozygosity); PIC – polymorphism information content, PM – match probability; PD – power of discrimination; H_{obs} – observed heterozygosity; PE – power of exclusion; TPI – typical paternity index; F_{ST} – subpopulation correction; F_{IS} – coefficient of inbreeding; pHW – p-values for deviation from the Hardy-Weinberg equilibrium.

Locus	N _{all}	GD	PIC	PM	PD	H _{obs}	PE	TPI	F _{ST}	F _{IS}	pHW
CSF1PO	10	0.741	0.695	0.114	0.886	0.738	0.489	1.906	1.7E-05	0.0159	0.216
D10S1248	9	0.763	0.724	0.095	0.905	0.756	0.520	2.047	-1.2E-03	0.0233	0.157
D12S391	16	0.886	0.875	0.025	0.975	0.880	0.755	4.168	-7.9E-04	0.0013	0.121
D13S317	10	0.780	0.751	0.078	0.922	0.791	0.582	2.387	-5.0E-05	-0.0241	0.663
D16S539	8	0.765	0.729	0.095	0.905	0.775	0.554	2.222	-5.7E-04	-0.0061	0.493
D18S51	18	0.877	0.865	0.028	0.972	0.881	0.756	4.190	-3.7E-04	-0.012	0.100
D19S433	16	0.785	0.758	0.075	0.925	0.789	0.578	2.366	4.7E-04	-0.0193	0.595
D15I656	21	0.901	0.892	0.019	0.981	0.897	0.789	4.846	-4.6E-04	0.0054	0.150
D21S11	18	0.852	0.835	0.039	0.961	0.857	0.709	3.498	1.1E-04	-0.0091	0.197
D22S1045	10	0.743	0.701	0.111	0.889	0.748	0.506	1.982	2.7E-04	-0.0088	0.913
D2S1338	15	0.882	0.870	0.026	0.974	0.881	0.756	4.190	6.0E-04	-0.002	0.568
D2S441	12	0.748	0.707	0.104	0.896	0.744	0.499	1.953	-2.0E-04	0.0207	0.794
D3S1358	10	0.786	0.751	0.080	0.920	0.781	0.564	2.279	3.8E-04	-0.0014	0.661
D5S818	10	0.731	0.687	0.117	0.883	0.727	0.470	1.828	-6.8E-04	0.0013	0.909
D6S1043	15	0.830	0.811	0.049	0.951	0.814	0.626	2.691	4.8E-04	0.0202	0.266
D7S820	10	0.801	0.773	0.068	0.932	0.787	0.576	2.352	3.2E-04	0.0112	0.292
D8S1179	11	0.803	0.776	0.066	0.934	0.807	0.612	2.587	-3.5E-04	1.00E-04	0.249
FGA	22	0.862	0.847	0.035	0.965	0.865	0.724	3.690	-4.2E-06	8.00E-04	0.174
Penta D	18	0.826	0.802	0.055	0.945	0.836	0.667	3.047	-7.7E-04	-0.005	0.910
Penta E	19	0.903	0.894	0.018	0.982	0.892	0.780	4.650	3.9E-04	0.0069	0.201
TH01	9	0.778	0.742	0.085	0.915	0.777	0.557	2.241	-3.8E-04	-0.0028	0.922
TPOX	8	0.605	0.555	0.208	0.792	0.605	0.297	1.267	3.2E-05	0.0067	0.512
vWA	9	0.807	0.779	0.066	0.934	0.810	0.617	2.629	1.9E-04	-0.007	0.344

Table 2

Simulation results. HS – half sibling, FS – full sibling.

Scenario	True hypothesis	LR				Percentage of simulations with:					
		Median	95%	5%	Standard deviation	LR > 10 (supporting H1), %	0.1 < LR < 10 (inconclusive), %	LR < 0.1 (supporting H2), %			
HS (H1) vs. Unrelated (H2)	H1	68.29	1.528e+004	0.5015	2.595e+006	73.576	25.069	1.355			
	H2	0.02095	1.722	0.0003351	26.78	1.209	26.449	72.342			
HS (H1) vs. Unrelated (H2)	H1	751.5	7.263e+005	1.369	3.5e+009	86.789	12.339	0.872			
	H2	0.003411	0.6427	3.48e-005	1165	0.716	13.309	85.975			
Data one mother											
HS (H1) vs. Unrelated (H2)	H1	1.8e+004	9.585e+007	6.591	7.623e+011	94.021	5.461	0.518			
	H2	0.0004914	0.163	4.167e-006	83.17	0.361	6.118	93.521			
Data both mothers											
FS (H1) vs. HS (H2)	H1	132.2	5.891e+004	0.5256	6.199e+006	77.121	21.435	1.444			
	H2	0.01899	1.398	0.0005725	32.1	1.015	24.101	74.884			
FS (H1) vs. HS (H2)	H1	2.455e+004	1.432e+008	7.861	7.567e+011	94.456	5.071	0.473			
	H2	0.0003956	0.1406	3.203e-006	33.65	0.343	5.582	94.075			
FS (H1) vs. Unrelated (H2)	H1	6.11e+006	2.832e+011	288.2	5.936e+017	98.627	1.275	0.098			
	H2	4.06e-006	0.00674	4.675e-009	78.8	0.072	1.293	98.635			

false negatives; simulating the alternative hypothesis yielded 72.342% of true negatives and 1.209% of false positives. For half siblings (one mother will be genotyped), simulating the true relationship, 86.789% of the 100,000 simulations were equal or above the LR limit, 0.872% of false negatives; simulating the alternative hypothesis yielded 85.975% of true negatives and 0.716% of false positives. For half siblings (mothers will be genotyped), simulating the true relationship, 94.021% of the 100,000 simulations were equal or above the LR limit, 0.518% of false negatives; simulating the alternative hypothesis yielded 93.521% of true negatives and 0.361% of false positives. For full siblings duo vs. half siblings, simulating the true relationship, 77.121% of the 100,000

simulations were equal or above the LR limit, 1.444% of false negatives; simulating the alternative hypothesis yielded 74.884% of true negatives and 1.015% of false positives. For full siblings vs. half siblings (mother will be genotyped), simulating the true relationship, 94.456% of the 100,000 simulations were equal or above the LR limit, 0.473% of false negatives; simulating the alternative hypothesis yielded 94.075% of true negatives and 0.343% of false positives. For full siblings vs. unrelated, simulating the true relationship, 98.627% of the 100,000 simulations were equal or above the LR limit, 0.098% of false negatives; simulating the alternative hypothesis yielded 98.635% of true negatives and 0.072% of false positives.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank G. Kostinyuk, I. Kalambet, and N. Kalambet for laboratory assistance with DNA analysis.

References

- [1] Standards for Relationship Testing Laboratories, 13th ed., AABB, 2017.
- [2] A. Gouy, M. Zieger, STRAF - A convenient online tool for STR data evaluation in forensic genetics, *Forensic Sci. Int. Genet.* 30 (2017) 148–151.
- [3] D. Kling, A.O. Tillmar, T. Egeland, Familias 3 - Extensions and new functionality, *Forensic Sci. Int. Genet.* 13 (2014) 121–127.