Arch Med Sadowej Kryminol 2022; vol. 72 (3): 120-137 DOI: 10.4467/16891716AMSIK.22.015.17393

archiwum medycyny sądowej i kryminologii

Praca oryginalna
Original paper

Karolina Hołub[1], Boris A. Malyarchuk[2], Miroslava V. Derenko[2], Nataša Kovačević-Grujičić[3], Milena Stevanović[3,4,5], Danijela Drakulić[3], Slobodan Davidović[6], Tomasz Grzybowski[1]

Verification of insertion-deletion markers (InDels) and microsatellites (STRs) as subsidiary tools for inferring Slavic population ancestry

Weryfikacja markerów insercyjno-delecyjnych (InDels) i mikrosatelitarnych (STR) jako narzędzi pomocniczych do wnioskowania o pochodzeniu populacji słowiańskiej

1 Department of Forensic Medicine, The Ludwik Rydygier Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Toruń, Bydgoszcz, Poland
2 Institute of Biological Problems of the North, Russian Academy of Sciences, Magadan, Russia
3 Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia
4 Faculty of Biology, University of Belgrade, Belgrade, Serbia
5 Serbian Academy of Sciences and Arts, Belgrade, Serbia
6 Institute for Biological Research "Siniša Stanković", National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia

Abstract

Genetic markers for the prediction of biogeographical ancestry have proved to be effective tools for law enforcement agencies for many years now. In this study, we attempted to assess the potential of insertion-deletion markers (InDel) and microsatellites (STRs) as subsidiary polymorphisms for inference of Slavic population ancestry. For that purpose, we genotyped Slavic-speaking populations samples from Belarus, the Czech Republic, Poland, Serbia, Ukraine and Russia in 46 InDels and 15 STRs by PCR and capillary electrophoresis and analyzed for between-population differentiation with the use of distance-based methods (F_{ST}, principal component analysis and multidimensional scaling). Additionally, we studied a sample from a Polish individual of well-documented genealogy whose biogeographic ancestry had previously been inferred by commercial genomic services using autosomal single nucleotide polymorphisms (SNPs), mitochondrial DNA and Y-SNP markers. For comparative purposes, we used genotype data collected in the "forInDel" browser and allele frequencies from previously published papers. The results obtained for InDels and STRs show that the Slavic populations constitute a genetically homogeneous group, with the exception of the Czechs differing clearly from the other tested populations. The analysis of the known Polish sample in the Snipper application proves the usefulness of the InDel markers on the continental level only. Conversely, microsatellites not only improve prediction, but are also informative if considered as an independent set of ancestry markers.

Key words: insertion-deletion polymorphisms, microsatellites, the Slavs

Submitted: 2022-12-31
Accepted: 2023-02-07

Streszczenie

Markery genetyczne do przewidywania pochodzenia biogeograficznego od wielu lat okazują się skutecznymi narzędziami dla organów ścigania. W tym badaniu podjęliśmy próbę oceny potencjału markerów insercyjno-delecyjnych (InDel) i mikrosatelitarnych (STR) jako pomocniczych polimorfizmów do wnioskowania o pochodzeniu populacji słowiańskiej. W tym celu genotypowaliśmy próbki populacji słowiańskojęzycznych z Białorusi, Czech, Polski, Serbii, Ukrainy i Rosji w w zakresie 46 markerów InDel oraz 15 loci STR za pomocą PCR i elektroforezy kapilarnej oraz analizowaliśmy pod kątem różnicowania między populacjami za pomocą metod bazujących na dystansach genetycznych (F_{ST}, analiza głównych składowych i skalowanie wielowymiarowe). Dodatkowo zbadaliśmy próbkę mężczyzny z populacji polskiej o dobrze udokumentowanej genealogii, którego pochodzenie biogeograficzne zostało wcześniej ustalone przez komercyjne usługi genomiczne przy użyciu autosomalnych polimorfizmów pojedynczych nukleotydów (SNP), mitochondrialnego DNA i markerów Y-SNP. Do celów porównawczych wykorzystaliśmy dane genotypowe zebrane w przeglądarce „forInDel" i częstości alleli z wcześniej opublikowanych artykułów. Uzyskane wyniki dla InDels i STR wskazują, że populacje słowiańskie stanowią grupę genetycznie jednorodną, z wyjątkiem Czechów wyraźnie różniących się od pozostałych badanych populacji. Analiza znanej polskiej próbki w aplikacji Snipper dowodzi przydatności markerów InDel jedynie na poziomie kontynentalnym. Z kolei, mikrosatelity nie tylko poprawiają wyniki predykcji, ale są informatywne jako niezależny zestaw markerów pochodzenia biogeograficznego.

Słowa kluczowe: polimorfizm insercyjno-delecyjny, mikrosatelity, Słowianie

Introduction

Ancestry Informative Markers (AIMs) are genetic polymorphisms characterized by large differences in allele frequencies between populations [1] and employed in both population and forensic genetics. In the latter, they are most frequently applied for genetic prediction of ancestry of the donors of biological samples or in the identification of human remains when a reference material is unavailable [2]. There are well-established methods of determining biogeographical origin based on the analysis of the mitochondrial DNA molecule and the non-recombining fragment of the Y chromosome. However, these markers, due to their haploid nature, reflect the genealogy of small portions of the human genome, which can be misleading in reliable predicting the ancestry of the person with an admixed background. There are also several panels useful for predicting ancestry with continental accuracy, most of them being based on autosomal SNPs [3-6]. One of the panels, developed by Pereira et al. [7] contains 46 autosomal insertion-deletion markers (InDel), typed by multiplex PCR and capillary electrophoresis. Additionally, the "forInDel" database was developed (available at http://spsmart.cesga.es/forindel.php?dataSet=forindel46) to facilitate the handling of population based InDel data [7]. The obtained genotypes can also be analyzed using the Snipper application (available at http://mathgene.usc.es/index.php) based on the Bayesian statistics, which allows a studied sample to be assigned to the population with the best match [8]. The Snipper also allows for the analysis of microsatellite markers (STR), used predominantly for human identification, but also being capable of improving the prediction of biogeographic ancestry [9].

In this study, we aimed at verifying two sets of genetic markers, i.e. InDels described by Pereira et al. [7] and STRs from the commercially available AmpFLSTR™ Identifiler™ Plus kit as potential subsidiary tools for prediction of ancestry among the Slavic-speaking populations of Europe. Previous studies of 46 InDel have not included Slavic-speaking populations, and their predictive power has been proved for the continental populations of Africans, Europeans, East Asians, and Native Americans. Microsatellite markers were included in the analysis due to their potential ability to improve the resolution of prediction. For this purpose, we genotyped different population samples representing Slavic populations and verified the degree of their differentiation using the distance-based methods. Additionally, we predicted the ancestry of one actual sample from a Polish individual of the well-documented genealogy using the Snipper application.

Material and methods

Population samples

The research material consisted of 409 DNA extracts from unrelated Slavic-speaking individuals including 61 samples from Belarus, 43 from the Czech Republic, 151 from Poland, 60 from Serbia, 57 from Ukraine and 37 samples obtained from persons of European descent (mostly Russian) from Magadan in Russia. Both maternal and paternal ancestry was confirmed for all sample donors for at least two generations. The study was approved by the Bioethics Committee of the Collegium Medicum, NCU (statement no. KB 423/2017). The part of the study concerning Serbian samples was approved by the Ethics Committee of the Institute of Molecular Genetics and Genetic Engineering, University of Belgrade (decision no. O-EO-005/2018/1).

Laboratory analyses

Concentrations of DNA isolates were measured using a Qubit® fluorimeter (Thermo Fisher Life Technologies). The amplification reaction for the 46 InDels was performed with 1 x Qiagen multiplex PCR master Mix and up to 5 ng of template DNA in a final reaction volume of 10 µl. PCR reactions were performed in GeneAmp® PCR System 9700 thermal cyclers. Amplification conditions were 95°C for 15 min, 28 cycles: 94°C for 30 s, 60°C for 90 s, 72°C for 45 s and 72°C for 60 minutes [10]. Concentrations of the PCR primers described by Pereira et al. [7] were experimentally optimized and presented in Table I. Microsatellite markers were tested with the commercially available multiplex amplification kit AmpFLSTR™ Identifiler™ Plus according to the manufacturer's recommendations (AmpFℓSTR™ Identifiler™ Plus PCR Amplification Kit USER GUIDE, number 4440211, revision H). Spectral calibration of the ABI PRISM® 3130 xl Genetic Analyzer (Applied Biosystems) capillary analyzer was carried out using PCR products obtained as a result of the InDel loci amplification (for the purposes of the analysis of insertion-deletion markers). Capillary electrophoresis was performed with 1 µl of product of PCR, 9 µl of Hi-Di Formamide (Applied Biosystems) and 0.6 µl of GeneScan™ 600 LIZ™ Size Standard. The GeneMapper® ID v.3.2 software was used to visualize the results (Applied Biosystems).

Statistical and bioinformatic analyses

Allele frequencies of individual genetic markers, F_{ST} pairwise genetic distances between populations, deviations from the Hardy-Weinberg equilibrium and linkage disequilibrium (LD) were estimated at the statistical significance level of 0.05 with the Bonferroni correction [11] using the Arlequin 3.5.2.2 program [12]. Negative F_{ST} values were changed to 0 [13]. Multidimensional scaling (MDS) and a principal component analysis (PCA) were performed using allele frequencies of the tested markers (NIPALS algorithm, 5-fold cross-validation) in the Statistica 13 (StatSoft) program. For comparative purposes, genotype InDel data were retrieved using the "forInDel" browser. As far as microsatellite markers are concerned, allele frequencies from the published papers were used [14-40]. Three analyses were conducted: a binary classification at the continental level, i.e., into one of the groups: Europe, East Asia, Africa, America, Oceania; a classification using our own genotype results and the data retrieved from the "forInDel" browser using a set of 46 InDels; and a classification using frequencies of STR data (our own published data).

The analysis of a sample with a well-documented genealogy

Additionally, the InDel and STR genotyping of a sample from a Polish male was performed, followed by the Snipper classification. The sample had previously been analyzed by the commercial genomic companies MyHeritage and 23andMe, using the Infinium® Global Screening Array-24 v1.0 (Illumina, 642 824 autosomal markers) and the Illumina OmniExpress platform (730 295 autosomal SNPs), respectively. In both cases, the sample had been reported as of mostly Eastern European ancestry (43% and 97.7%, respectively). The individual's self-inferred family genealogy also pointed to the Eastern European origins, with the most probable population ancestry of the present-day Ukraine.

Table I. Sequences and concentrations of primers (μM) used for amplification of a multiplex set of 46 insertion-deletion polymorphisms.

Marker	Product length (del/ins)	Labeled primer (5′→3′)	Unlabeled primer (5′→3′)	Concentration
1470	65/70	6FAM-GAGTCTGACCCTTCATAAGC	GCCATGGTGATATTACGTCCC	0.0475
777	75/78	6FAM-TGGAAGACACGTCCTAAGAG	GTATTCCTCCAGGCTCTTTGC	0.055
196	83/86	6FAM-CCAAGTTCTAGCCATATGGA	GTTTCTTGACTATCTTCTCTGACCATC	0.145
881	92/96	6FAM-TTGGCTCCCTATGATAATCC	GTTTCTTGTGTTCCCAAAGTTCTCC	0.1525
3122	104/108	6FAM-TCACAAGTCCGGAATACCAG	GAGTTATGGGATGGGAAGGAG	0.05
548	111/113	6FAM-AGTCAGGACTGAAGAAACCC	GTTTCAGTAAAACAAGAGCCCGTG	0.04
659	120/122	6FAM-CACTGCATCAGACTGACTTC	GTTTCTTGGCTGCTTTGCTTTGAATTG	0.06
2011	129/134	6FAM-TGAGAAACTAGGAGCTCTGG	GTTTCCTAAGAGCCACTGACAT	0.075
2929	150/152	6FAM-TGTGATGTGGATAGGCAAGG	GAGGCTCCATTTGTTAAGAGG	0.05
593	159/161	6FAM-TGCTCACTTTAGTGGAGACC	GTTTGCGTTTAGGTCCCTTCTG	0.05
798	169/175	6FAM-ACGACAGTGTTCACAAAGAG	GCTGTTGTCTGACCTGTGAAG	0.055
1193	182/184	6FAM-GCTGGGTAGTTTTTTCCTCC	GTTCCACCATCTACCTTCTATG	0.095
1871	190/192	6FAM-TTGTAGTCAGAGAGTGTGCC	GAGCCTTTTCCCTAACGTCAC	0.0375
17	200/204	6FAM-AGAAACTGCAACCCTCCAAG	GATCCCAGACACTGAAGATG	0.0725
2538	212/216	6FAM-CTCGCAAAGTAGGCAAGTTC	GACACCAACAATCTTGGCACC	0.05
1644	223/225	6FAM-ACACCACTGAAGATCTGACC	GGTCTAAAGTCAGTGCACAG	0.055
3854	58/62	HEX-TCACCCTTATTCAGGGTTGC	GCCAGGGATTTAGTGTAGAG	0.13
2275	72/79	HEX-CTACCTGACTACCACCTATG	GACCCAGCCTATCTGACTTTG	0.075
94	92/95	HEX-TGGTGGCTCATGCACTTTTG	GTTTACAGGGTCTCGCTATGATGC	0.0475
3072	110/117	HEX-AGCTTTTTCCGGCAACTCTC	GTTTGGATGTGTCTGAGCTCAAC	0.0475
772	120/123	HEX-GTCTCRTTTTCCTGCAGTAG	GTTTCTTATCCTTCTGCTCACTCTACC	0.12
2313	130/139	HEX-GCACACATGCAGAAATGCAG	GTTGTAACATCTGTGAGGTC	0.085
397	163/167	HEX-TGGGCTTCTTCTGGGAAAAC	GCCACATTCAGGCGTTTTGTC	0.0725
1636	174/176	HEX-TTAGGAAGAGGTGCTATGGG	GCCTCCTTTGAAGACACACAG	0.0675
51	183/188	HEX-AAGATTGGAGGGAAAAGTGC	GCGTCCTCCACCTTCTTTTTC	0.08
2431	209/213	HEX-AGGAGGAGCTGATAGACTTC	GCAGTGTTGCAACTGATACG	0.0775
2264	224/228	HEX-CTTTGGCTATCCTGTCTCAC	GTAGGAGACCACTCACATTC	0.095
2256	61/64	TAMRA-ATCGAACCGTTCCTAAGGAC	GCAAGAAAAGGAATCCAGGC	0.05
128	70/73	TAMRA-ATCAGGAGACAATCCAGCAG	GTCCAGCCATTCAGACAAAGG	0.045
15	81/91	TAMRA-GGGTTATTTGCCTCATCTCC	GTTTCTAGGTATTCTCTGTTCCCACG	0.115
2241	110/118	TAMRA-ACATACACGTGGAAGACTGC	GTTACTGTCGACTGATCCAATAG	0.0575
419	124/131	TAMRA-CAGGAAAGTATGGCCCATTC	GTCCATGTTTTCTTTGAGCATC	0.065
943	156/160	TAMRA-TCTTCCTACCCCTGTTAGTG	GACAAGATCACTAGCTTGAC	0.135
159	168/173	TAMRA-ACCAGAGCACTACAGCCTTT	GCAAGGYAGTAACAATGAGGG	0.1
2005	185/191	TAMRA-TGTAGCGGCAATATAGGCAG	GAAAGTTGTGGCTTAACTGG	0.2775
250	202/204	TAMRA-ATGGAGCAGTAAAGCAGCAC	GTCACTTTGGTTTTTGCAGG	0.21
1802	213/216	TAMRA-ACGGTCAACTTTGTAGCTCC	GCCAGTTGAGAATCACTGCAC	0.1375
1607	221/223	TAMRA-TGTTGCAGAAGAACTCAACC	GATAAGCACCTAACTCCCAG	0.275
1734	56/60	ROX-TTCGTGTTCTCACACTGTCC	GTGCATCCCATACAACTGAC	0.1
406	63/65	ROX-TGGCTGCTGTAGATTGTAGG	GACAAATGGACAACGGCCAAG	0.1075

Marker	Product length (del/ins)	Labeled primer (5′→3′)	Unlabeled primer (5′→3′)	Concentration
1386	72/93	ROX-AGAGGATCATGGAGACCAAC	GTTTATGTTCCAAGTCAGCAGCAC	0.15
1726	106/119	ROX-GGTCCAAATGCACCACAATC	GCTCTGCTATTTTGGTTTGC	0.225
3626	142/157	ROX-TGTTGGTTCTCTCCTTTTCC	GGTGACCCCTTCTTTATCTC	0.405
360	167/169	ROX-AGATCAACTGCCAATCTGGG	GCTCAAGTGACCAACCCACCT	0.1275
1603	214/218	ROX-TTACAATTTCAAGCCTCCGC	GGAGCTGTTAGTCTGAGTAG	0.275
2719	225/229	ROX-GTCAGGAGTCTAGAAACTTC	GGGTGATGAAATGTTCCGAA	0.425

Table II. Allele frequencies of insertion-deletion polymorphisms.

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
15	**deletion**	0.385	0.36	0.354	0.311	0.4	0.333
	insertion	0.615	0.64	0.646	0.689	0.6	0.667
17	**deletion**	0.32	0.384	0.341	0.405	0.317	0.351
	insertion	0.68	0.616	0.659	0.595	0.683	0.649
51	**deletion**	0.68	0.628	0.728	0.716	0.642	0.693
	insertion	0.32	0.372	0.272	0.284	0.358	0.307
94	**deletion**	0.164	0.186	0.146	0.203	0.25	0.193
	insertion	0.836	0.814	0.854	0.797	0.75	0.807
128	**deletion**	0.41	0.535	0.477	0.446	0.525	0.526
	insertion	0.59	0.465	0.523	0.554	0.475	0.474
159	**deletion**	0.484	0.453	0.553	0.486	0.533	0.509
	insertion	0.516	0.547	0.447	0.514	0.467	0.491
196	**deletion**	0.434	0.585	0.52	0.459	0.5	0.561
	insertion	0.566	0.415	0.48	0.541	0.5	0.439
250	**deletion**	0.73	0.721	0.722	0.662	0.667	0.789
	insertion	0.27	0.279	0.278	0.338	0.333	0.211
360	**deletion**	0.803	0.767	0.851	0.892	0.842	0.746
	insertion	0.197	0.233	0.149	0.108	0.158	0.254
397	**deletion**	0.852	0.779	0.834	0.716	0.8	0.728
	insertion	0.148	0.221	0.166	0.284	0.2	0.272
406	**deletion**	0.885	0.791	0.858	0.878	0.775	0.781
	insertion	0.115	0.209	0.142	0.122	0.225	0.219
419	**deletion**	0.803	0.814	0.825	0.797	0.8	0.807
	insertion	0.197	0.186	0.176	0.203	0.2	0.193
548	**deletion**	0.23	0.209	0.212	0.257	0.25	0.237
	insertion	0.77	0.791	0.788	0.743	0.75	0.763
593	**deletion**	0	0	0.003	0	0	0.009
	insertion	1	1	0.997	1	1	0.991
659	**deletion**	0.098	0.093	0.086	0.108	0.142	0.123
	insertion	0.902	0.907	0.914	0.892	0.858	0.877
772	**deletion**	0.967	0.93	0.987	0.946	0.958	0.974
	insertion	0.033	0.07	0.013	0.054	0.042	0.026
777	**deletion**	0.402	0.36	0.377	0.351	0.267	0.307

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
	insertion	0.598	0.64	0.623	0.649	0.733	0.693
798	**deletion**	0.639	0.628	0.659	0.649	0.65	0.675
	insertion	0.361	0.372	0.341	0.351	0.35	0.325
881	**deletion**	0.885	0.919	0.911	0.932	0.875	0.939
	insertion	0.115	0.081	0.089	0.068	0.125	0.061
943	**deletion**	0.721	0.791	0.781	0.757	0.825	0.754
	insertion	0.279	0.209	0.219	0.243	0.175	0.246
1193	**deletion**	0.23	0.163	0.156	0.203	0.175	0.123
	insertion	0.77	0.837	0.844	0.797	0.825	0.877
1386	**deletion**	0.27	0.314	0.189	0.203	0.183	0.281
	insertion	0.73	0.686	0.811	0.797	0.817	0.719
1470	**deletion**	0.631	0.477	0.682	0.595	0.653	0.623
	insertion	0.369	0.523	0.318	0.405	0.347	0.377
1603	**deletion**	0.311	0.314	0.325	0.365	0.325	0.36
	insertion	0.689	0.686	0.676	0.635	0.675	0.64
1607	**deletion**	0.23	0.163	0.281	0.284	0.233	0.272
	insertion	0.77	0.837	0.719	0.716	0.767	0.728
1636	**deletion**	0.73	0.733	0.735	0.743	0.842	0.711
	insertion	0.27	0.267	0.265	0.257	0.158	0.289
1644	**deletion**	0.967	0.942	0.94	0.946	0.975	0.982
	insertion	0.033	0.058	0.06	0.054	0.025	0.018
1726	**deletion**	0.762	0.721	0.692	0.649	0.708	0.702
	insertion	0.238	0.279	0.308	0.351	0.292	0.298
1734	**deletion**	0.861	0.791	0.894	0.892	0.9	0.842
	insertion	0.139	0.209	0.106	0.108	0.1	0.158
1802	**deletion**	0.008	0	0.003	0.014	0.008	0.009
	insertion	0.992	1	0.997	0.986	0.992	0.991
1871	**deletion**	0.41	0.326	0.404	0.419	0.442	0.465
	insertion	0.59	0.674	0.596	0.581	0.558	0.535
2005	**deletion**	0.672	0.663	0.606	0.608	0.717	0.763
	insertion	0.328	0.337	0.394	0.392	0.283	0.237
2011	**deletion**	0.762	0.721	0.719	0.703	0.692	0.737
	insertion	0.238	0.279	0.281	0.297	0.308	0.263
2241	**deletion**	0.23	0.372	0.315	0.284	0.225	0.246
	insertion	0.77	0.628	0.685	0.716	0.775	0.754
2256	**deletion**	0.213	0.174	0.149	0.162	0.167	0.246
	insertion	0.787	0.826	0.851	0.838	0.833	0.754
2264	**deletion**	0.484	0.477	0.47	0.459	0.525	0.465
	insertion*	0.066	0.035	0.043	0.027	0.058	0.044
	insertion	0.451	0.488	0.487	0.514	0.417	0.491
2275	**deletion**	0.123	0.058	0.103	0.081	0.117	0.132
	insertion	0.877	0.942	0.897	0.919	0.883	0.868
2313	**deletion**	0.246	0.291	0.268	0.324	0.35	0.237
	insertion	0.754	0.709	0.732	0.676	0.65	0.763
2431	**deletion**	0.098	0.058	0.126	0.122	0.05	0.088

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
	insertion	0.902	0.942	0.874	0.878	0.95	0.912
2538	**deletion**	0.5	0.453	0.454	0.432	0.483	0.491
	insertion	0.5	0.547	0.546	0.568	0.517	0.509
2719	**deletion**	0.393	0.314	0.444	0.486	0.483	0.43
	insertion	0.607	0.686	0.556	0.514	0.517	0.57
2929	**deletion**	0.762	0.756	0.725	0.851	0.767	0.781
	insertion	0.238	0.244	0.275	0.149	0.233	0.219
3072	**deletion**	0.943	0.942	0.944	0.946	0.933	0.939
	insertion	0.057	0.058	0.056	0.054	0.067	0.061
3122	**deletion**	0.992	0.977	0.987	1	0.958	0.965
	insertion	0.008	0.023	0.013	0	0.042	0.035
3626	**deletion**	0.615	0.674	0.629	0.662	0.625	0.675
	insertion	0.385	0.326	0.371	0.338	0.375	0.325
3854	**deletion**	0.025	0.035	0.01	0.014	0.008	0.018
	insertion	0.975	0.965	0.99	0.986	0.992	0.982

Table III. Allele frequencies of microsatellite polymorphisms.

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
D8S1179	**8**	0.016	0.023	0.013	0	0.025	0
	9	0.008	0	0.003	0	0.008	0.018
	10	0.066	0.081	0.066	0.081	0.075	0.079
	11	0.033	0.07	0.06	0.122	0.058	0.07
	12	0.164	0.233	0.192	0.108	0.142	0.211
	13	0.328	0.233	0.341	0.392	0.35	0.316
	14	0.246	0.151	0.212	0.189	0.217	0.184
	15	0.115	0.186	0.083	0.095	0.108	0.088
	16	0.025	0.023	0.026	0	0.017	0.035
	17	0	0	0.003	0.014	0	0
D21S11	**26**	0.008	0.012	0	0	0	0
	27	0.008	0.07	0.017	0.014	0.042	0.035
	28	0.205	0.128	0.195	0.176	0.133	0.123
	29	0.246	0.174	0.219	0.243	0.217	0.228
	29.2	0	0	0.007	0	0.008	0
	30	0.23	0.244	0.185	0.176	0.192	0.254
	30.2	0.066	0.07	0.05	0.068	0.033	0.044
	31	0.033	0.047	0.05	0.014	0.05	0.053
	31.2	0.074	0.07	0.089	0.122	0.125	0.105
	32	0.008	0.035	0.013	0.027	0.008	0
	32.2	0.107	0.093	0.129	0.162	0.142	0.096
	33.2	0.016	0.047	0.04	0	0.042	0.061
	34.2	0	0.012	0.007	0	0.008	0
D7S820	**7**	0.008	0.012	0.01	0.041	0.025	0.018
	8	0.123	0.093	0.176	0.135	0.133	0.167
	9	0.115	0.256	0.162	0.135	0.117	0.167

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
	10	0.311	0.221	0.275	0.243	0.308	0.263
	11	0.303	0.174	0.202	0.297	0.258	0.193
	12	0.107	0.198	0.142	0.122	0.142	0.158
	13	0.033	0.035	0.026	0.014	0.017	0.035
	14	0	0.012	0.007	0.014	0	0
CSF1PO	**8**	0	0	0.003	0	0	0
	9	0.049	0.047	0.046	0.081	0.017	0.053
	10	0.279	0.326	0.272	0.27	0.225	0.333
	11	0.295	0.244	0.301	0.176	0.367	0.228
	12	0.328	0.326	0.311	0.365	0.35	0.325
	13	0.049	0.047	0.053	0.081	0.025	0.061
	14	0	0.012	0.01	0	0.017	0
	15	0	0	0	0.027	0	0
	16	0	0	0.003	0	0	0
D3S1358	**13**	0	0.012	0	0	0	0
	14	0.115	0.07	0.199	0.162	0.125	0.184
	15	0.254	0.221	0.209	0.23	0.275	0.237
	15.2	0	0	0.003	0	0	0
	16	0.246	0.279	0.272	0.23	0.242	0.211
	16.2	0	0	0.003	0	0	0
	17	0.254	0.244	0.179	0.23	0.258	0.184
	18	0.131	0.163	0.123	0.149	0.092	0.167
	19	0	0.012	0.013	0	0.008	0.018
TH01	**5**	0	0	0.003	0	0	0
	6	0.18	0.279	0.245	0.27	0.317	0.202
	7	0.148	0.163	0.109	0.081	0.092	0.14
	8	0.082	0.07	0.07	0.162	0.158	0.114
	8.3	0	0.023	0	0	0	0
	9	0.213	0.105	0.202	0.149	0.183	0.219
	9.3	0.369	0.36	0.358	0.311	0.233	0.325
	10	0	0	0.01	0.027	0.017	0
	10.3	0.008	0	0	0	0	0
	11	0	0	0.003	0	0	0
D13S317	**8**	0.139	0.07	0.172	0.23	0.158	0.175
	9	0.09	0.07	0.063	0.054	0.117	0.07
	10	0.041	0.07	0.02	0.027	0.067	0.018
	11	0.402	0.337	0.391	0.432	0.308	0.368
	12	0.197	0.314	0.228	0.122	0.233	0.219
	13	0.09	0.047	0.076	0.054	0.075	0.105
	14	0.041	0.093	0.05	0.068	0.042	0.044
	15	0	0	0	0.014	0	0
D16S539	**8**	0.008	0	0.007	0	0.017	0.018
	9	0.115	0.174	0.06	0.068	0.092	0.088
	10	0.008	0.058	0.033	0.054	0.058	0.061
	11	0.23	0.244	0.285	0.203	0.275	0.228
	12	0.352	0.302	0.384	0.351	0.283	0.377

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
	12.2	0	0	0	0.014	0	0
	13	0.246	0.198	0.182	0.257	0.225	0.211
	14	0.041	0.023	0.05	0.041	0.05	0.018
	15	0	0	0	0.014	0	0
D2S1338	**15**	0	0	0	0	0.008	0
	16	0.025	0.07	0.073	0.068	0.058	0.026
	17	0.254	0.174	0.205	0.257	0.258	0.193
	18	0.074	0.081	0.076	0.081	0.083	0.088
	19	0.139	0.128	0.116	0.041	0.083	0.123
	20	0.139	0.14	0.132	0.216	0.175	0.167
	21	0.057	0.035	0.02	0.014	0.033	0.044
	22	0.008	0.047	0.017	0.014	0.042	0.035
	23	0.09	0.14	0.096	0.068	0.1	0.114
	24	0.066	0.116	0.116	0.135	0.058	0.096
	25	0.115	0.07	0.126	0.081	0.083	0.105
	26	0.033	0	0.02	0	0.008	0.009
	27	0	0	0.003	0.027	0.008	0
D19S433	**10.2**	0	0	0.003	0	0	0
	11	0	0	0.007	0	0.017	0
	12	0.057	0.058	0.096	0.122	0.067	0.132
	12.2	0	0	0.003	0	0	0
	13	0.189	0.291	0.212	0.243	0.233	0.289
	13.2	0.049	0.012	0.007	0.041	0.008	0.018
	14	0.418	0.349	0.361	0.324	0.317	0.281
	14.2	0.041	0	0.017	0.027	0.008	0.009
	15	0.156	0.163	0.169	0.189	0.167	0.149
	15.2	0.016	0.023	0.026	0.014	0.067	0.035
	16	0.041	0.081	0.046	0	0.092	0.026
	16.2	0.016	0.012	0.046	0.027	0.017	0.035
	17	0.008	0	0	0.014	0.008	0
	17.2	0.008	0	0.003	0	0	0.018
	18.2	0	0.012	0.003	0	0	0.009
vWA	**13**	0	0	0.003	0	0.017	0
	14	0.139	0.07	0.076	0.081	0.125	0.079
	15	0.139	0.093	0.086	0.189	0.108	0.123
	16	0.197	0.163	0.219	0.176	0.158	0.237
	17	0.213	0.326	0.278	0.311	0.308	0.325
	18	0.254	0.267	0.222	0.189	0.183	0.175
	19	0.049	0.07	0.096	0.054	0.083	0.044
	20	0.008	0.012	0.017	0	0.017	0.018
	21	0	0	0.003	0	0	0
TPOX	**7**	0.008	0	0	0	0	0
	8	0.615	0.43	0.563	0.608	0.592	0.535
	9	0.066	0.14	0.116	0.068	0.083	0.088
	10	0.033	0.081	0.056	0.041	0.058	0.088
	11	0.254	0.291	0.222	0.23	0.233	0.272

Marker	Allele	BLR	CZE	POL	RUS	SRB	UKR
	12	0.025	0.058	0.04	0.054	0.033	0.018
	13	0	0	0.003	0	0	0
D18S51	**10**	0.008	0.012	0.01	0	0.025	0
	11	0.016	0	0.026	0.027	0.025	0.009
	12	0.09	0.128	0.076	0.054	0.108	0.061
	13	0.123	0.128	0.099	0.081	0.167	0.14
	14	0.131	0.174	0.152	0.122	0.133	0.175
	15	0.197	0.14	0.142	0.149	0.15	0.193
	16	0.197	0.163	0.152	0.176	0.2	0.158
	17	0.082	0.081	0.156	0.189	0.067	0.053
	18	0.057	0.093	0.086	0.081	0.083	0.061
	19	0.049	0.023	0.04	0.054	0	0.105
	20	0.033	0.023	0.043	0.041	0.025	0.035
	21	0.008	0.035	0.013	0.014	0.008	0.009
	22	0.008	0	0.003	0.014	0.008	0
D5S818	**7**	0.025	0.012	0.003	0	0.008	0
	9	0.074	0.047	0.04	0.041	0.025	0.044
	10	0.09	0.116	0.06	0.122	0.058	0.079
	11	0.303	0.291	0.344	0.365	0.292	0.351
	12	0.344	0.326	0.401	0.324	0.367	0.368
	13	0.156	0.198	0.149	0.135	0.25	0.132
	14	0	0.012	0.003	0.014	0	0.026
	16	0.008	0	0	0	0	0
FGA	**16**	0	0	0	0	0.008	0
	17	0	0	0.003	0	0	0.009
	18	0	0	0.023	0	0.008	0.018
	19	0.066	0.093	0.086	0.122	0.058	0.079
	20	0.213	0.093	0.113	0.095	0.133	0.14
	21	0.123	0.233	0.185	0.122	0.158	0.211
	21.2	0.016	0	0.007	0.027	0	0
	22	0.139	0.244	0.179	0.27	0.133	0.167
	22.2	0.025	0	0.007	0	0	0.009
	23	0.107	0.116	0.126	0.122	0.192	0.123
	23.2	0	0.012	0.007	0	0.008	0.009
	24	0.18	0.081	0.132	0.149	0.167	0.114
	24.2	0	0	0	0.014	0	0
	25	0.123	0.105	0.093	0.054	0.092	0.123
	26	0.008	0.023	0.036	0.027	0.042	0
	27	0	0	0.003	0	0	0

Results and discussion

Population data analyzed with the use of distance-based methods

The allele frequencies of the 46 InDel markers and 15 STRs are presented in Tables II and III, respectively. No significant results were observed in the linkage disequilibrium and Hardy-Weinberg tests after applying the Bonferroni correction. Matrices of pairwise genetic distances between populations are depicted in Table IV. Graphs of multidimensional scaling are shown in Figures I, II and III. The "stress" values in all MDS analyses indicate a very good fit between the distance matrices of the reconstructed distances and the distance matrices observed. Scatter plot of self-loads is presented in Figure IV. Snipper classification results for a set of 46 InDel markers are presented in Tables V, VI and Figure V, while Table VII shows the classification based on STRs.

The obtained results of genetic distances (greater than zero) reflect low genetic differentiation, which means high homogeneity of the studied Slavic-speaking populations (Table IV), the condition being proved previously for European gene pool in general with the use of different genetic markers. However, both MDS and PC analyses indicate that the population of the Czech Republic is the most genetically distant from the other studied Slavic populations (absence of genetic differentiation was observed relative to the population of Ukraine only, Table IV). Comparing both sets of markers while setting aside the subtle Czech outlier, it can be seen that 46 InDels gave similar results to the set of STR markers commonly used for human identification in forensics. The exceptions constitute positive (but not significant) InDel genetic distances observed between Poland and Ukraine, Serbia and Belarus, which were not noted using microsatellites. Conversely, positive distances between Poland and Belarus, Serbia and Russia, Serbia and Ukraine were not obtained in the study of InDels but were observed in STRs (Table IV).

The case of individual prediction

According to the Snipper classification based on 46 InDel markers at the continental level, the studied sample of a Polish individual was assigned to the European population (probabilities are presented on a negative logarithmic scale, so a lower value indicates a better match). The profile of the tested person 25 893 434 times more likely belongs to the population of EUROPE than OCEANIA (Table V, Figure V). While at the continental level the classification result is correct, after further subdivision of Europe into subpopulations, the Polish sample was assigned with the highest probability to the population of France (Table VI). It should be noted here that the differences between the probabilities for successive matches are not large – the profile of the studied Pole only 2 times more likely belongs to the French population than for the Adygei and 355 times more likely belongs to the French population than for the Polish population. Considering the above, the set of 46 insertion-deletion markers is characterized by continental rather than population resolution. However, the Snipper microsatellite classification (Table VII) gave a much more plausible result in terms of the probability of assigning the sample to a particular European population. In this case, the sample was most probably assigned to the Belarusian population; the result being consistent with both previous high-resolution SNPs' inference and the individuals' self-inferred family genealogy. Simultaneously, small differences between the assignment probabilities indicate very small genetic differentiation between the populations on the STR level.

In conclusion, STRs but not InDels appear useful supplementary markers for inferring biogeographic ancestry within the Slavic-speaking population of Europe.

Table IV. Matrices of F_{ST} pairwise genetic distances between populations for insertion-deletion markers (InDel), microsatellite markers (IF) and both sets together (InDel + IF). Statistically significant distances are written in bold.

	BLR	CZE	POL	RUS	SRB	UKR
InDel						
BLR	0					
CZE	0.00089	0				
POL	0	**0.00413**	0			
RUS	0	0.00025	0	0		
SRB	0.00122	0.00376	0.00206	0	0	
UKR	0	0	0.00231	0	0	0
IF						
BLR	0					
CZE	**0.00459**	0				
POL	0.00051	**0.00292**	0			
RUS	0	**0.00499**	0	0		
SRB	0	**0.00306**	0.00162	0.00047	0	
UKR	0	0	0	0	0.00087	0
InDel + IF						
BLR	0					
CZE	0.00251	0				
POL	0.00007	**0.00360**	0			
RUS	0	0.00233	0	0		
SRB	0.00066	**0.00346**	**0.00187**	0	0	
UKR	0	0	0.00068	0	0.00036	0

Table V. Bayesian classification of the studied Polish sample using the Snipper application for a set of 46 insertion-deletion markers at the continental level.

Population	-ln(likelihood)
EUROPE	36.4906
OCEANIA	53.5601
EAST ASIA	64.3729
AMERICA	78.5671
AFRICA	87.5719

Table VI. Bayesian classification of the studied Polish sample using the Snipper application for a set of 46 insertion-deletions. Both our own genotypic data and those downloaded from the "forInDel" browser were used in the analysis*.

Population	-ln(likelihood)
France	35.9454
Adygei, Russia	36.6048
Italy	36.9797
Russia	38.0448
Basque, France	38.9981
Belarus	39.8187
Middle East	39.8794
Central-south Asia	40.3405
Serbia	40.4148
Ukraine	40.557
Poland	41.8174
Magadan, Russia	41.893
Orkney Islands	41.9066
Czech	42.215
Oceania	53.7207
East Asia	64.3847
America	78.5061
Africa	87.5988

Table VII. Bayesian classification in the Snipper application for a set of microsatellite markers using our own data (population names in capital letters) and allele frequencies from published papers.

Population	Likelihood	-ln(likelihood)
BELARUS	3.55431E-24	53.9939
Italy [14]	2.51676E-24	54.3391
Spain [15]	1.1528E-24	55.1198
Spain [16]	8.65509E-25	55.4065
Basque Country [17]	8.44144E-25	55.4315
Sweden [18]	6.9279E-25	55.6291
Greece [19]	5.00861E-25	55.9535
Croatia [20]	4.07136E-25	56.1607
Romania [21]	3.87226E-25	56.2108
Galicia [22]	2.48123E-25	56.6559
Ukraine [23]	2.16913E-25	56.7903
Poland- Pomorze Zachodnie [24]	1.404E-25	57.2253
Denmark [25]	9.73306E-26	57.5917
Sweden [26]	9.57944E-26	57.6076
Serbia and Montenegro-Vojvodina [27]	9.57835E-26	57.6077
Serbia- Kosovo- Albanians [28]	8.48903E-26	57.7284
Serbia [29]	7.88233E-26	57.8026
Czech [30]	6.46332E-26	58.0011
Poland [31]	6.21621E-26	58.0401
Poland- Łodź [32]	5.77245E-26	58.1141
Hungary [33]	5.09134E-26	58.2397
Croatia [34]	4.24412E-26	58.4217
The United Kingdom [35]	3.9333E-26	58.4977
SERBIA	3.77698E-26	58.5383
POLAND	2.82265E-26	58.8295
Russia- Saratov [36]	2.12107E-26	59.1153
The Netherlands [37]	2.02562E-26	59.1613
UKRAINE	1.06236E-26	59.8067
CZECH	3.93937E-27	60.7988
Russia- Min-Vody [36]	3.59655E-27	60.8898
Estonia [38]	3.12041E-27	61.0318
Russia- Orel [36]	2.62156E-27	61.206
Eastern Slovakia- Spis [39]	2.03594E-28	63.7614
RUSSIA	1.4934E-28	64.0713
Eastern Slovakia- Abov-Gemer [39]	1.30146E-28	64.2089
Eastern Slovakia- Saris [39]	3.66531E-29	65.4761
Italy-Tuscany [40]	1.54905E-29	66.3373

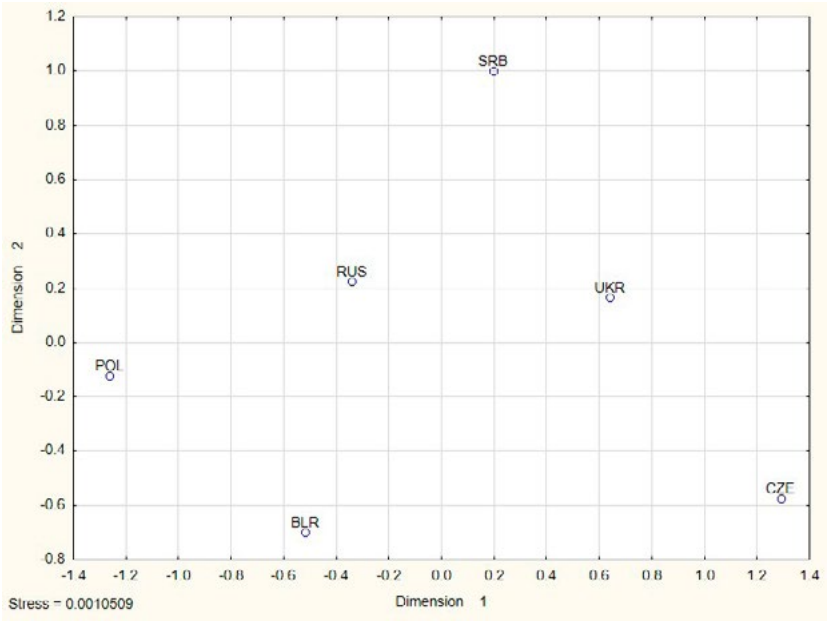

Figure I. Multidimensional scaling graph of studied populations regarding InDel markers.
Population abbreviations:
BLR – Belarus,
CZE – Czech Republic,
POL – Poland,
RUS – Russia,
SRB – Serbia,
UKR – Ukraine.

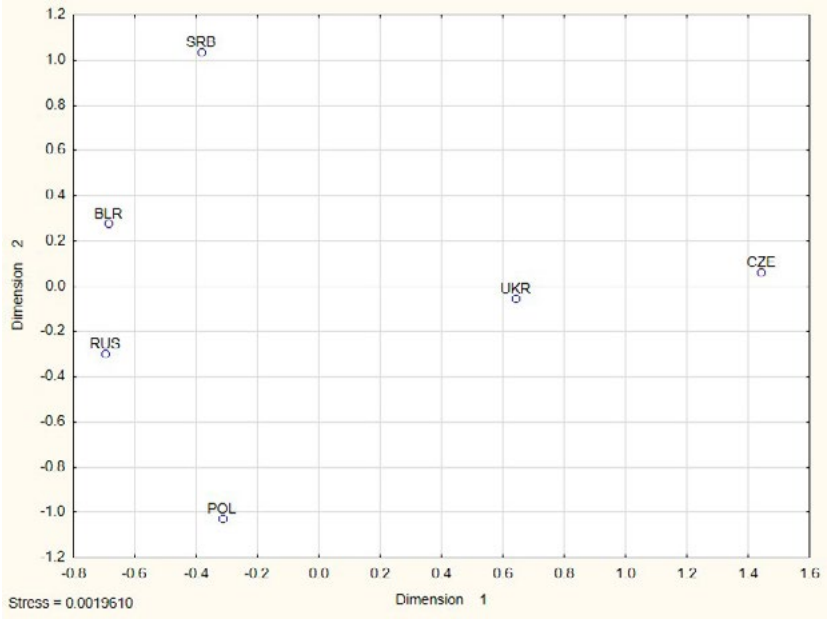

Figure II. Multidimensional scaling graph of studied populations regarding STR markers.
Population abbreviations:
BLR – Belarus,
CZE – Czech Republic,
POL – Poland,
RUS – Russia,
SRB – Serbia,
UKR – Ukraine.

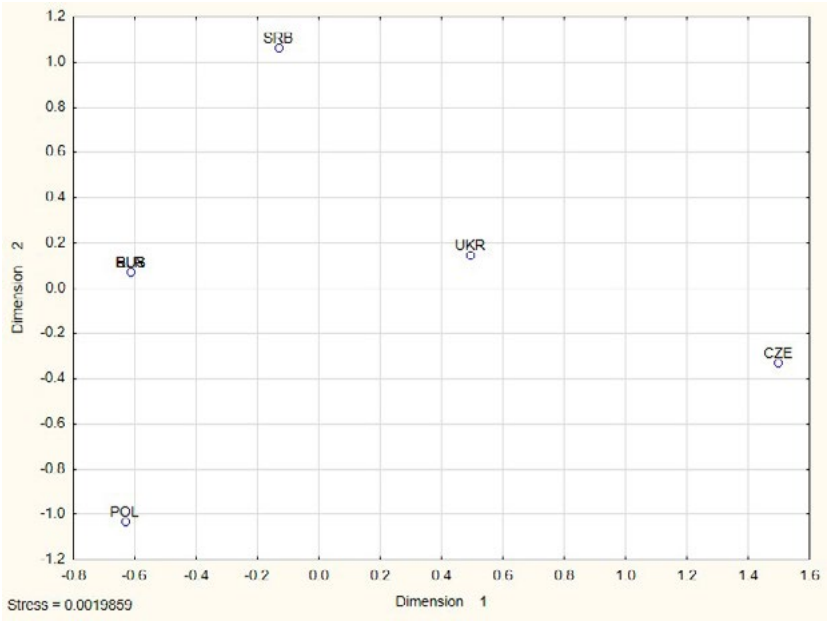

Figure III. Multidimensional scaling graph of studied populations regarding both InDel and STR markers. Population abbreviations:
BLR – Belarus,
CZE – Czech Republic,
POL – Poland,
RUS – Russia,
SRB – Serbia,
UKR – Ukraine.

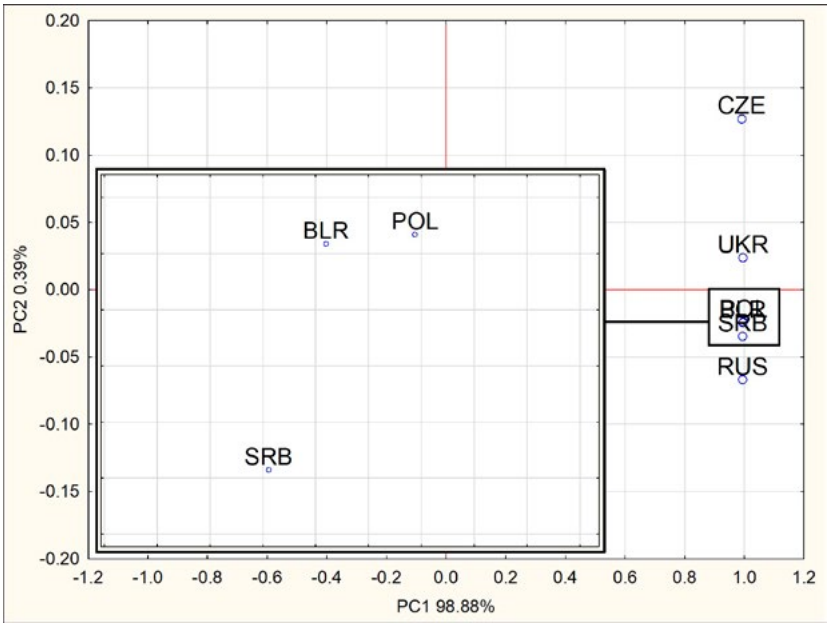

Figure IV. Scatter plot of self--loads of studied populations regarding all studied markers. Population abbreviations:
BLR – Belarus,
CZE – Czech Republic,
POL – Poland,
RUS – Russia,
SRB – Serbia,
UKR – Ukraine.

PC1 – the first principal component that explains 98.88% of the observed genetic variation;

PC2 – the second principal component that explains 0.39% of the observed genetic variation.

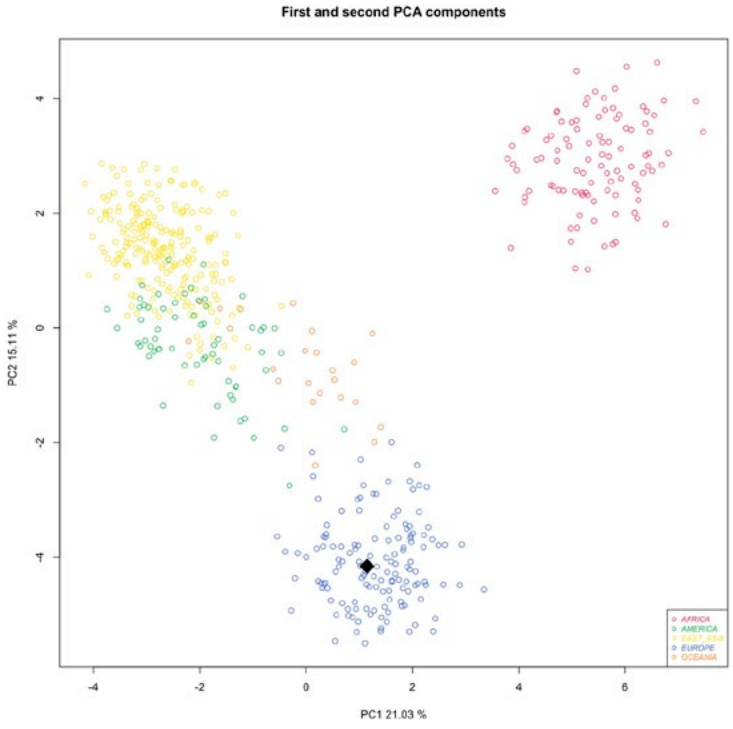

Figure V. Scatter plot of self-loads generated in the Snipper application in continental level analysis for a set of 46 InDel markers. The black square on the graph represents a sample from the studied Polish individual.

PC1 – the first principal component;
PC2 – the second principal component.

Funding

The study was supported by grants for young scientists and doctoral students from the Faculty of Medicine, Collegium Medicum in Bydgoszcz of the Nicolaus Copernicus University in Toruń (nos. MN-SDL-1/WL/2017 and MN-SDL-1/WL/2019) as well as the Ministry of Education, Science and Technological Development of the Republic of Serbia (Agreement no. 451-03-68/2022-14/200042).

Supplementary data in the electronic version: The component table for PCA analysis (Suppl. Table I).

References

1. Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM, Kittles RA. Skin pigmentation, biogeographical ancestry and admixture mapping. Hum Genet 2003; 112(4): 387-99.
2. Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, Alvarez-Dios J, Alonso A, Blanco-Verea A, Brión M, Montesino M, Carracedo A, Lareu MV. Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 2009; 4(8): e6583.
3. Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, Madbouly A, Maiers M, Middha M, Friedlaender FR, Kidd JR. Progress toward an efficient panel of SNPs for ancestry inference. Forensic Sci Int Genet 2014; 10: 23-32.
4. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 2009; 30(1): 69-78.
5. Phillips C, Parson W, Lundsberg B, Santos C, Freire-Aradas A, Torres M, Eduardoff M, Børsting C, Johansen P, Fondevila M, Morling N, Schneider P; EUROFORGEN-NoE Consortium, Carracedo A, Lareu MV. Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set. Forensic Sci Int Genet 2014; 11: 13-25.
6. Pereira V, Freire-Aradas A, Ballard D, Børsting C, Diez V, Pruszkowska-Przybylska P, Ribeiro J, Achakzai NM, Aliferi A, Bulbul O, Carceles MDP, Triki-Fendri S, Rebai A, Court DS, Morling N, Lareu MV, Carracedo Á; EUROFORGEN-NoE Consortium,

Phillips C. Development and validation of the EUROFORGEN NAME (North African and Middle Eastern) ancestry panel. Forensic Sci Int Genet 2019; 42: 260-267.
7. Pereira R, Phillips C, Pinto N, Santos C, dos Santos SE, Amorim A, Carracedo Á, Gusmão L. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. PLoS One 2012; 7(1): e29684.
8. Phillips C, Salas A, Sánchez JJ, Fondevila M, Gómez-Tato A, Alvarez-Dios J, Calaza M, de Cal MC, Ballard D, Lareu MV, Carracedo A; SNPforID Consortium. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet 2007; 1: 273-80.
9. Phillips C, Fernandez-Formoso L, Garcia-Magariños M, Porras L, Tvedebrink T, Amigo J, Fondevila M, Gomez-Tato A, Alvarez-Dios J, Freire-Aradas A, Gomez-Carballa A, Mosquera-Miguel A, Carracedo A, Lareu MV. Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel. Forensic Sci Int Genet 2011; 5(3): 155-69.
10. Santos C, Phillips C, Oldoni F, Amigo J, Fondevila M, Pereira R, Carracedo Á, Lareu MV. Completion of a worldwide reference panel of samples for an ancestry informative Indel assay. Forensic Sci Int Genet 2015; 17: 75-80.
11. Rice WR. Analyzing tables of statistical tests. Evolution 1989; 1: 223-225.
12. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010; 10(3): 564-7.
13. Nei M. Molecular evolutionary genetics. New York. Columbia University Press; 1987.
14. Barbaro A, Phillips C, Formoso LF, Lareu MV, Carracedo Á. Distribution of allele frequencies of 20 STRs loci in a population sample from Calabria, Southern Italy. Forensic Sci Int Genet 2012; 6(5): e137-8.
15. Hill CR, Duewer DL, Kline MC, Coble MD, Butler JM. U.S. population data for 29 autosomal STR loci. Forensic Sci Int Genet 2013; 7(3): e82-3. Erratum in: Forensic Sci Int Genet 2017; Nov; 31: e36-e40.
16. García O, Alonso J, Cano JA, García R, Luque GM, Martín P, de Yuso IM, Maulini S, Parra D, Yurrebaso I. Population genetic data and concordance study for the kits Identifiler, NGM, PowerPlex ESX 17 System and Investigator ESSplex in Spain. Forensic Sci Int Genet 2012; Mar; 6(2): e78-9. Erratum in: Forensic Sci Int Genet 2014; 9: 192.
17. Yurrebaso I, Ajuriagerra JA, Alday A, Lezama I, Pérez JA, Romón E, Uriarte I, García O. Allele frequencies and concordance study between the Identifiler and the PowerPlex ESX17 systems in the Basque Country population. Forensic Sci Int Genet 2011; 5(3): e79-80.
18. Montelius K, Karlsson AO, Holmlund G. STR data for the AmpFlSTR Identifiler loci from Swedish population in comparison to European, as well as with non-European population. Forensic Sci Int Genet 2008; 2(3): e49-52.
19. Sánchez-Diz P, Menounos PG, Carracedo A, Skitsa I. 16 STR data of a Greek population. Forensic Sci Int Genet 2008; 2(4): e71-2.
20. Barbarić L, Ozretić P, Horjan I, Korolija M, Mršić G. Forensic evaluation of the 20 STR loci in the population of Croatia. Forensic Sci Int Genet 2017; 28: e49-e50.
21. Stanciu F, Popescu OR, Stoian IM. Allele frequencies of 15 STR loci in Moldavia region (NE Romania). Forensic Sci Int Genet 2009; 4(1): e39-40.
22. Fernandez-Formoso L, Phillips C, Rodriguez A, Calvo R, Barbaro A, Lareu MV, Carracedo Á. Allele frequencies of 20 STRs from Northwest Spain (Galicia). Forensic Sci Int Genet 2012; 6(5): e149-50.
23. Serga SV, Dombrovskyi IV, Maistrenko OM, Ostapchenko LI, Demydov SV, Krivda RG, Kozeretska IA. Allele frequencies for 15 STR loci in the Ukrainian population. Forensic Sci Int Genet 2017; 29: e40-e41.
24. Piatek J, Jacewicz R, Ossowski A, Parafiniuk M, Berent J. Population genetics of 15 autosomal STR loci in the population of Pomorze Zachodnie (NW Poland). Forensic Sci Int Genet 2008; 2(3): e41-3.
25. Hussing C, Bytyci R, Huber C, Morling N, Børsting C. The Danish STR sequence database: duplicate typing of 363 Danes with the ForenSeq™ DNA Signature Prep Kit. Int J Legal Med 2019; 133(2): 325-334.
26. Staadig A, Tillmar A. An overall limited effect on the weight-of-evidence when taking STR DNA sequence polymorphism into account in kinship analysis. Forensic Sci Int Genet 2019; 39: 44-49.
27. Veselinović I, Kubat M, Furac I, Skavić J, Martinović Klarić I, Tasić M. Allele frequencies of the 15 AmpF lSTR Identifiler loci in the population of Vojvodina Province, Serbia and Montenegro. Int J Legal Med 2004; 118(3): 184-6.
28. Kubat M, Skavić J, Behluli I, Nuraj B, Bekteshi T, Behluli M, Klarić IM, Pericić M. Population genetics of the 15 AmpF lSTR Identifiler loci in Kosovo Albanians. Int J Legal Med 2004; 118(2): 115-8.
29. Novković T, Panić B, Banjac A, Dekić TK, Tomisić-Kosić I, Vucetić-Dragović A, Stamenković G, Blagojević J, Marjanović D, Pojskić N. Genetic polymorphisms of 15 AmpFlSTR Identifiler loci in a Serbian population. Forensic Sci Int Genet 2010; 4(5): e149-50.
30. Simková H, Faltus V, Marvan R, Pexa T, Stenzl V, Broucek J, Horínek A, Mazura I, Zvárová J. Allele frequency data for 17 short tandem repeats in a Czech population sample. Forensic Sci Int Genet 2009; 4(1): e15-7.
31. Ossowski A, Diepenbroek M, Szargut M, Zielińska G, Jędrzejczyk M, Berent J, Jacewicz R. Population analysis and forensic evaluation of 21 autosomal loci included in GlobalFiler™ PCR Kit in Poland. Forensic Sci Int Genet 2017; 29: e38-e39.
32. Jacewicz R, Jedrzejczyk M, Ludwikowska M, Berent J. Population database on 15 autosomal STR loci in 1000 unrelated individuals from the Lodz region of Poland. Forensic Sci Int Genet 2008; 2(1): e1-3.
33. Petrić G, Drašković D, Zgonjanin-Bosić D, Budakov B, Veselinović I. Genetic variation at 15 autosomal STR loci in the Hungarian population of Vojvodina Province, Republic of Serbia. Forensic Sci Int Genet 2012; 6(6): e163-5.

34. Haliti N, Carapina M, Masić M, Strinović D, Klarić IM, Kubat M. Evaluation of population variation at 17 autosomal STR and 16 Y-STR haplotype loci in Croatians. Forensic Sci Int Genet 2009; 3(4): e137-8.
35. Devesse L, Ballard D, Davenport L, Riethorst I, Mason-Buck G, Syndercombe Court D. Concordance of the ForenSeq™ system and characterisation of sequence-specific autosomal STR alleles across two major population groups. Forensic Sci Int Genet 2018; 34: 57-61.
36. Zhivotovsky LA, Malyarchuk BA, Derenko MV, Wozniak M, Grzybowski T. Developing STR databases on structured populations: the native South Siberian population versus the Russian population. Forensic Sci Int Genet 2009; 3(4): e111-6.
37. Westen AA, Kraaijenbrink T, Robles de Medina EA, Harteveld J, Willemse P, Zuniga SB, van der Gaag KJ, Weiler NEC, Warnaar J, Kayser M, Sijen T, de Knijff P. Comparing six commercial autosomal STR kits in a large Dutch population sample. Forensic Sci Int Genet 2014; 10: 55-63.
38. Sadam M, Tasa G, Tiidla A, Lang A, Axelsson EP, Pajnič IZ. Population data for 22 autosomal STR loci from Estonia. Int J Legal Med 2015; 129(6): 1219-20.
39. Soták M, Petrejčíková E, Bôžiková A, Bernasovská J, Bernasovský I, Sovičová A, Boroňová I, Svičková P, Gabriková D, Mačeková S, Carnogurská J, Rębała K, Vlček D. Population database of 17 autosomal STR loci from the four predominant Eastern Slovakia regions. Forensic Sci Int Genet 2011; 5(3): 262-3.
40. Carboni I, Nutini AL, Porfirio B, Genuardi M, Ricci U. Genetic STRs variation in a large population from Tuscany (Italy). Forensic Sci Int Genet 2007; 1(3-4): e10-1.

The study is partly based on the data presented in Karolina Hołub's PhD thesis entitled "The application of molecular markers of biogeographic ancestry within Europe in population and forensic genetics" (Bydgoszcz, 2021).

ORCID

Karolina Hołub
Boris A. Malyarchuk 0000-0002-0304-0652
Miroslava V. Derenko 0000-0002-1849-784X
Nataša Kovačević-Grujičić 0000-0002-3837-1283
Milena Stevanović 0000-0003-4286-7334
Danijela Drakulić 0000-0001-6790-6673
Slobodan Davidović 0000-0002-1519-8828
Tomasz Grzybowski 0000-0001-6228-6460

CORRESPONDING AUTHOR

Karolina Hołub
Department of Forensic Medicine,
The Ludwik Rydygier Collegium Medicum in Bydgoszcz,
Nicolaus Copernicus University in Toruń, Bydgoszcz, Poland
e-mail: karolina.holub@proton.me