

The Uniparental Genetic Landscape of Modern Slavic Speaking Populations

Sena Karachanak-Yankova^{1*}, Desislava Nesheva¹, Draga Toncheva¹, Angel S. Galabov^{2*}

¹Department of Medical Genetics, Medical University-Sofia, Sofia, Bulgaria

²The Stephan Angeloff Institute of Microbiology, Bulgarian Academy of Sciences, Sofia, Bulgaria

Email: *senulica@yahoo.com, *galabov@microbio.bas.bg

How to cite this paper: Karachanak-Yankova, S., Nesheva, D., Toncheva, D., & Galabov, A. S. (2017). The Uniparental Genetic Landscape of Modern Slavic Speaking Populations. *Advances in Anthropology*, 7, 318-332.

<https://doi.org/10.4236/aa.2017.74018>

Received: August 14, 2017

Accepted: November 12, 2017

Published: November 15, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Slavic speaking populations are the most numerous Indo-European ethnolinguistic group in Europe. They show great variety and fall into three groups: West, East and South Slavic populations. In order to contribute to the understanding of the correlation between linguistic and genetic affiliation of Slavic populations, we have analyzed for the first time their matrilineal and patrilineal relationships and we have also illustrated their position in the European uniparental genetic landscape. For the purpose, we have collected previously published data for the frequencies of mitochondrial DNA (mtDNA) and Y-chromosome haplogroups in Slavic and other European populations and compared them by Principal Component Analysis (PCA). In the inter-Slavic population comparisons, West and East Slavs are in a closer position, whereas South Slavic populations are rather grouped on their own. In the European context, South Slavic populations are positioned more close to neighboring Balkan non-Slavic and North Italian populations, than to other Slavic populations. When considering the uniparental diversity of Slavic speaking populations, one should also take into account the prevalence of Y-chromosome haplogroup N among East Slavs (comprising almost half of the paternal gene pool in instances), which is almost absent among the other groups (not exceeding 2% - 3%). In conclusion, the data in the present study point that West-East and South Slavic speaking populations, behave as separate groups based on their uniparental genetic structure, which shows that they do not share substantial common genetic ancestry and that there is great genetic variety in the Slavic linguistic unity.

Keywords

Slavic Speaking Populations, Mitochondrial DNA, Y-Chromosome, Haplogroup, Principal Component Analysis

1. Introduction

Slavs are the most numerous Indo-European ethno-linguistic group in Europe (Kipfer, 2000). Their proposed homeland is in the middle Dnieper basin: the area North-East of the Carpathians, the upper stream of rivers Bug and Dniester, mainly around Pripet river (Mutafchiev, 1943; Rebala et al., 2007). They spread from the 6th century AD to inhabit whole Eastern Europe and parts of Central and South-Eastern Europe.

Roman writers mentioned Slavs as Venethi or Venedi (Curta, 2001). The Southern European Slavs are first named as Sclavenes by East Roman (Byzantine) authors in the 6th century AD (Džino, 2010; Smith, 2005).

Extant Slavic languages show great variety and fall into three groups: West (Czech, Slovakian and Polish), East (Russian, Belorussian and Ukrainian) and South (Serbian, Croatian, Bulgarian and Slovenian). The Bulgarian literary language differs from other Slavic languages by the almost complete loss of grammatical case; the creation of definite article of nouns (appearing in the form of a suffix, added to the stem); analytical comparative and superlative (by word-particles); and a complex tense system where the infinitive is completely lost (Aeppli, von Waldenfels, & Samardzic, 2014; Kushniarevich et al., 2015; Raykov, 2005).

Orthodox Christian Slavs use the Cyrillic alphabet, while Roman Catholic Slavs and Bosniaks use the Latin alphabet.

Genetic and genomic analysis of Slavs from different countries have been the object of many previously published studies; however, comparative genetic investigations among different Slavic groups are only few (Grzybowski et al., 2007; Malyarchuk et al., 2008; Mielnik-Sikorska et al., 2013; Rebala et al., 2007). To the best of our knowledge, the most comprehensive study of Slavic genetic heritage to date (Kushniarevich et al., 2015) reveals that the Slavic genetic diversity was formed through assimilation of preexisting regional genetic components and *in situ* gene pool shaping; as it also identifies an apparent genetic homogeneity of the majority of West and East Slavs and a substantial genetic difference between them and South Slavs.

In order to further contribute to the understanding of the correlation between language and genetic origin in Slavs, the present study analyzes for the first time the matrilineal and patrilineal relationships among European Slavic-speaking countries' populations and also illustrates their position in the European uniparental genetic landscape.

2. Materials and Methods

We have collected previously published data for the frequencies of mitochondrial DNA (mtDNA) and Y-chromosome haplogroups in Slavic speaking and other European populations. In these studies, the mtDNA haplogroup assignment was based on partial or entire control region sequences and/or coding region markers; and the Y-chromosome haplogroup classification was performed by genotyping of informative biallelic markers. To analyze comparable results for larger

number of populations, the data were normalized to the highest possible level of phylogenetic resolution.

We have included 18 and 27 (sub-) populations in the mtDNA and Y-chromosome haplogroup analyses of Slavic speaking populations (**Supplementary Table S1** and **Table S2**); and 41 and 42 (sub-) populations in the mtDNA and Y-chromosome haplogroup analyses of European populations (**Supplementary Table S3** and **Table S4**), respectively.

The mtDNA and Y-chromosome relationships among the populations were depicted by Principal Component Analysis (PCA) performed using XLSTAT.

3. Results and Discussion

The plot of the PCA performed on mtDNA haplogroup frequencies in Slavic speaking populations is presented in **Figure 1**. In the Slavic mtDNA landscape, the East Slavic populations (European Russians, Ukrainians) predominantly occupy areas covering negative values of PC2. It should be marked that European Russians are separated in two distant groups—regions of the Central (Vladimir, Yaroslavl, Tula and Kaluga) and of the North-Western part (Pskov, Velikii Novgorod and Volot) of the country. Ukrainians and Belarusians are interspersed with West Slavic populations (Poles, Slovaks and Czechs). South Slavic speaking populations (Serbians, Bulgarians, Slovenians, Croatians and the populations of Bosnia and Herzegovina) are quite dispersed despite the comparatively small geographic area of the Balkan Peninsula. Nevertheless, they are positioned in an area of positive values of PC2.

Patrilineal relationships of Slavic speaking populations based on Y-chromosome haplogroup frequencies are depicted in the PCA plot in **Figure 2**. Unlike in the case of mtDNA, in the PC analysis of Y-chromosome haplogroup frequencies there are clear-cut dispositions: South Slavic speaking populations (Bulgarians, Serbs, Bosnians, Bosnia-Serbs, Bosnia-Croats, Croats and Slovenians) are again much dispersed, but all of them are located in the PC1 positive values area. Most East Slavic speaking populations are residing in the negative part of PC2. West Slavic speaking populations (Slovaks, Czechs and Poles) are in between South and East Slavs, but are closer to East Slavs, being interspersed among them. Certain Russian populations (Vologda, Arkhangel and Orel) are markedly distant.

In the PC analyses of both mtDNA and Y-chromosome haplogroup frequencies of Slavic speaking populations, South Slavs are quite dispersed. This is probably due to their different pasts. It is established that the ancestors of Croatians and Serbians migrated to the Balkan Peninsula in the 8th century AD from territories in Central Europe (White Croats and White Serbs) (**Borri, 2011; Chadwick, 2014**). The area of their dispersal is in the territory of present-day West Slavic populations' countries. Very different is the story of the contemporary Bulgarians. South Slavic tribes (Sclavenes) and Proto-Bulgarians arrived almost simultaneously on the Balkan Peninsula in the 7th century AD, when the

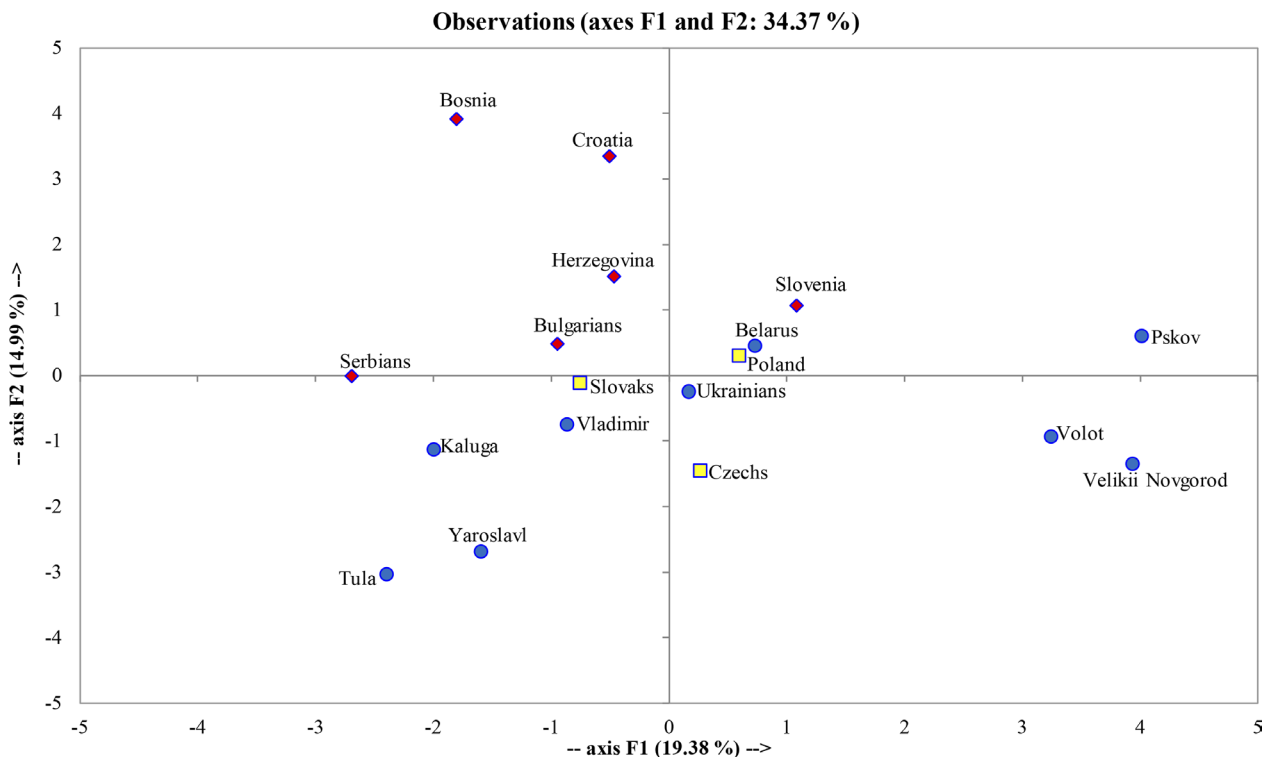


Figure 1. PCA plot of Slavic speaking populations based on mtDNA haplogroup frequencies. The variance of the first and second principal components (F1 and F2, respectively) is given in brackets. ■—West, ●—East and ◆—South Slavic speaking populations.

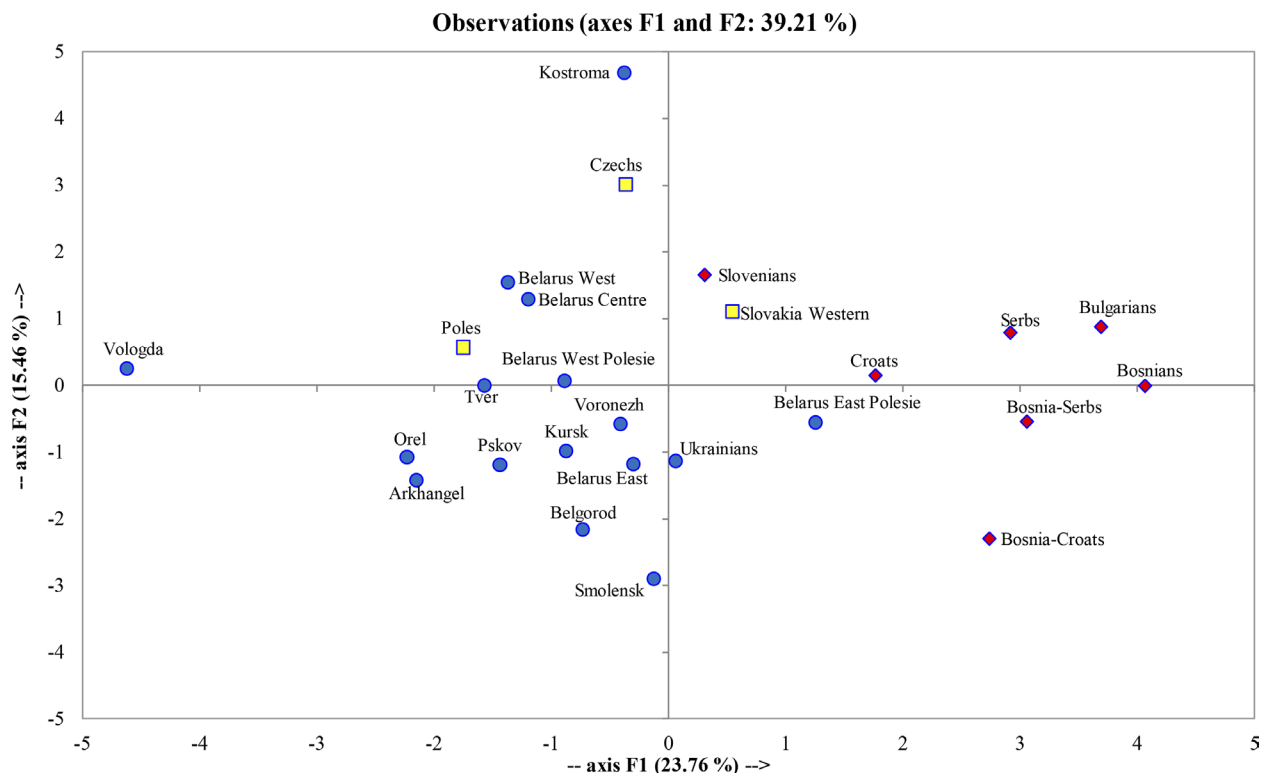


Figure 2. PCA plot of Slavic speaking populations carried out on Y-chromosome haplogroup frequencies. The variance retained by the first and second principal components (F1 and F2, respectively) is shown in brackets. ■—West, ●—East and ◆—South Slavic speaking populations.

Proto-Bulgarians established the Danubian Bulgarian state (Dobrev, 2016; Mutafovich, 1943).

The position of the mtDNA haplogroup frequency profiles of Slavic speaking populations in the European context is represented in **Figure 3**. In this PCA plot, South Slavs (Bulgarians, Slovenians, Serbians, Croats and populations of Bosnia and Herzegovina) are grouped alongside Balkan (Northern Greeks and Romanians) and Northern Italian populations. West Slavic populations (Slovaks, Czechs and most of the Poles) are adjacent to some North European non-Slavic populations (from Finland and Sweden). The majority of East Slavic populations (from European Russia and Belarus) are scattered as some European Russians (from Vladimir and Yaroslavl) and Ukrainians are close to West Slavic populations. Furthermore, Germanic and Romance speaking populations (from Germany and Austria; and Iberia, France and Italy; respectively) are located separately, being situated in the positive part of PC1.

The comparison of the Y-chromosome haplogroup frequencies in Slavic speaking and remaining European populations performed by PC analysis is given in **Figure 4**. Compared to the PCA of mtDNA haplogroup frequencies, it again shows a more clear-cut grouping of most of South Slavs (Serbs, Bulgarians, Croats, Bosnia-Croats, Bosnia-Serbs, Bosnians) with neighboring populations from the Balkans (Romanians, Greeks and Macedonian Greeks). On the other hand, West (Czechs, Poles and Western Slovaks) and East Slavs (Ukrainians, Belarus and European Russians) are located separately. From the non-Slavic populations Swedish Saami and Finns are almost outliers; the two populations from Germany are almost overlapping, whereas Italian populations form a cluster which embraces Catalonia.

In general, the obtained results show that based on the distribution of mtDNA and Y-chromosome haplogroups West and East Slavic speaking populations locate separately from South Slavic populations. Furthermore, in the European uniparental landscape South Slavic speaking populations are positioned more close to neighboring Balkan non-Slavic populations and North Italian populations, than to other Slavic populations. This hints that the linguistic resemblance of South Slavic speaking populations with East and West Slavic groups is not paralleled to a similar extent by a genetic one, which is in line with previous findings demonstrating that the basis of the gene pool of West-East and South Slavic speaking populations is different (Kushniarevich et al., 2015).

When considering the uniparental diversity of Slavic speaking populations, one should highlight a peculiarity in the East Slavs' gene pool, namely the prevalence of Y-chromosome haplogroup N, which is typical for certain Asian populations: China-Naxi—25%, China-Oroqen—28.5%, China-Tu—28.5%, Cambodians—16.7% (Sengupta et al., 2006), Mongols—8.7% (Derenko et al., 2007; Hammer et al., 2005), Manchurian—4.9% (Hammer et al., 2005), Koreans—4.3% (Hammer et al., 2005; Zhong et al., 2010), Vietnamese—2.8% (Hammer et al., 2005); Central Asian Turkic populations, such as Khakassians—50% and

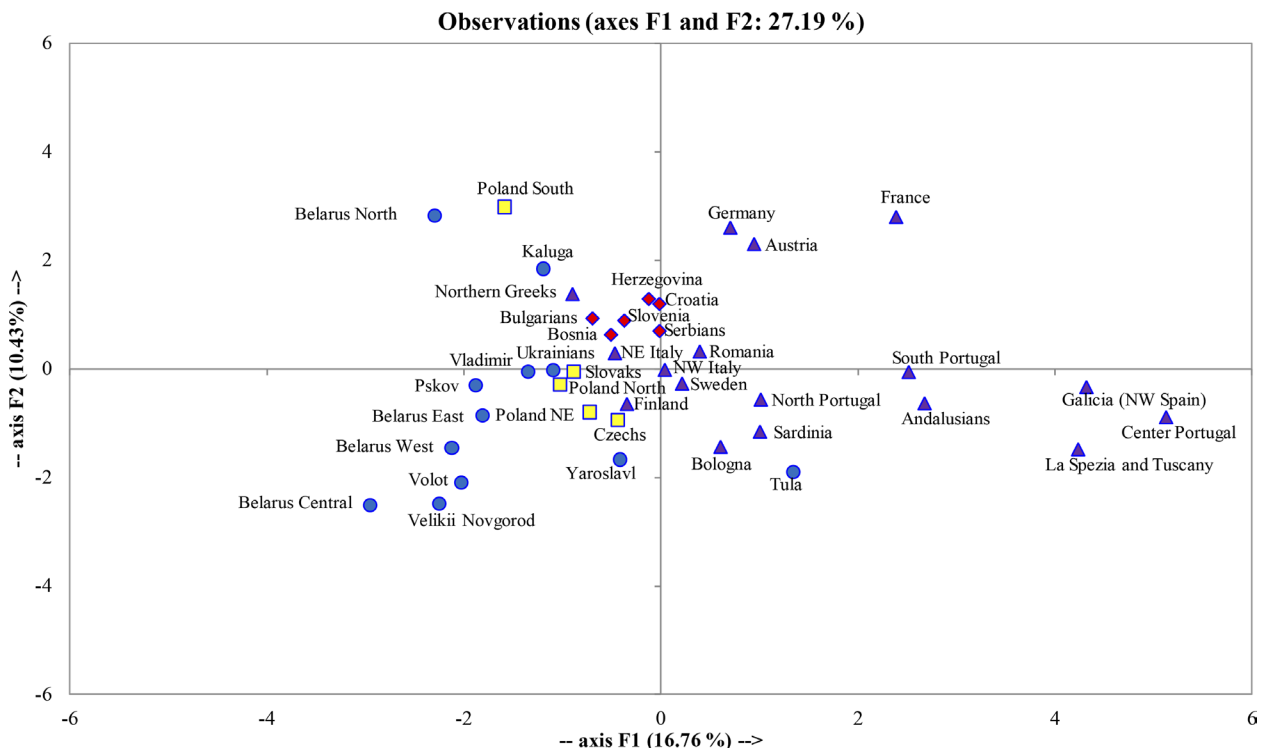


Figure 3. PCA plot of European populations based on mtDNA haplogroup frequencies. The variance captured by the first and second principal components (F1 and F2, respectively) is written in brackets. NE—northeast, NW—northwest, ■—West, ●—East Slavic, ◆—South Slavic, ▲—non-Slavic populations.

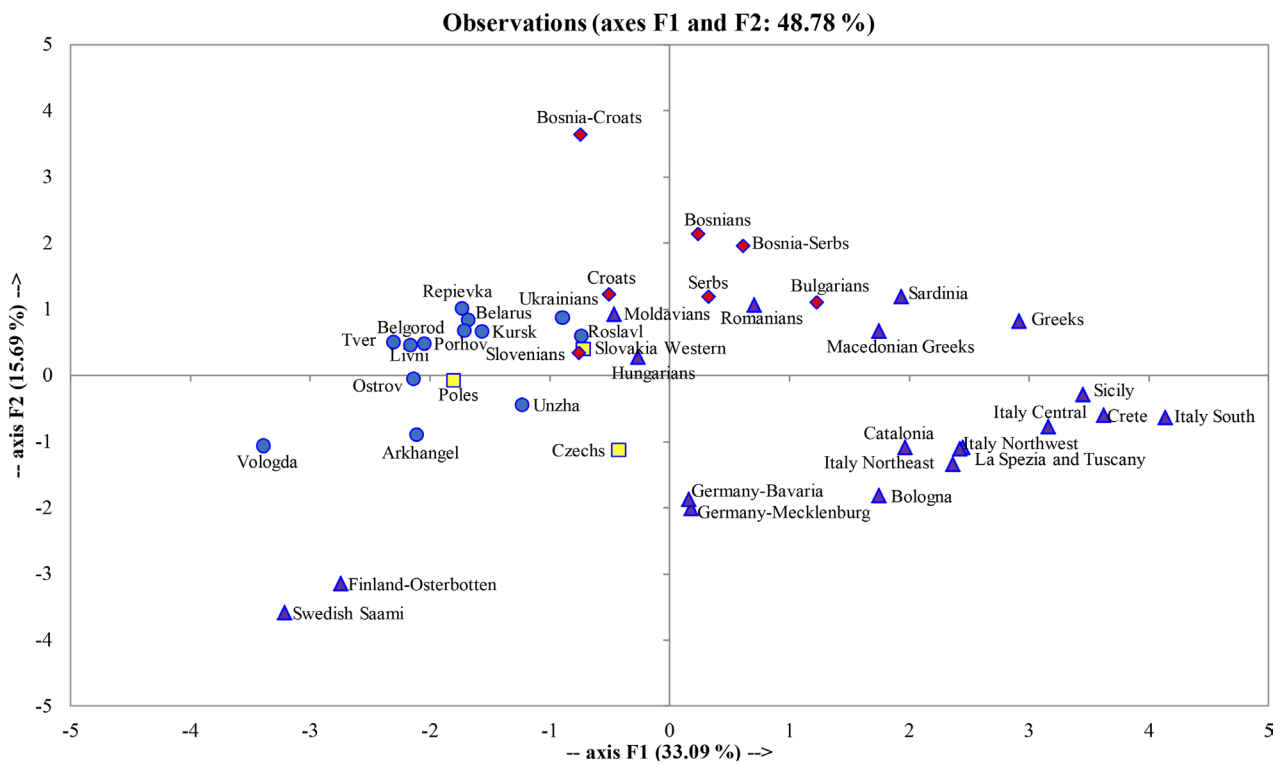


Figure 4. Map of Principal Component Analysis of European populations based on their Y-chromosome haplogroup frequencies. Variance explained by the first and second principal components (F1 and F2, respectively) is shown in brackets. ■—West, ●—East, ◆—South Slavic speaking, ▲—non-Slavic populations.

Altaians—9.2% (Derenko et al., 2007); and several Siberian populations: Khanty—81.5% (Mirabal et al., 2009), Yakuts—80.0% (Derenko et al., 2007), Buryats—30.9% (Hammer et al., 2005), Evenks—29.3%, Tuvinians—29.1% and Koryaks—25% (Derenko et al., 2007).

In Europe Y-chromosome haplogroup N shows high frequency in Saami—44.7% (Karlsson et al., 2006). The frequency values for European Russians are striking: Arkhangel—41.1% (Balanovsky et al., 2008; Mirabal et al., 2009), Vologda—38.8%, Pskov—24.2%, Smolensk—14.0%, Kostroma—13.5% (Balanovsky et al., 2008), Kursk—12.9% (Balanovsky et al., 2008; Mirabal et al., 2009), Belgorod—12.6% (Balanovsky et al., 2008), Tver—11.7% (Balanovsky et al., 2008; Mirabal et al., 2009), Voronezh—6.3%, Orel province—5.5% (Balanovsky et al., 2008) and Kalmyks—3.3% (Derenko et al., 2007). The data for Belarussians are 9.6% (Kushniarevich et al., 2013; Kushniarevich et al., 2015) and for Ukrainians—5.7% (Battaglia et al., 2009; Kushniarevich et al., 2015).

In the Y-chromosome gene pool of any West or South Slavic speaking population haplogroup N does not exceed 2-3% in frequency: Poles—2.7% (Battaglia et al., 2009; Rebalá et al., 2013), Czechs—2.7%, Croats—1.7% (Battaglia et al., 2009), Bulgarians—0.5% (Karachanak et al., 2013), Serbs 1.9% (Regueiro et al., 2012), Slovenians—0% and Bosniaks—0% (Battaglia et al., 2009). The same pattern is observed in the remaining populations in Europe, the Caucasus: Balkarians—0% and Georgians—0% (Battaglia et al., 2009) and in South Asia (Sengupta, et al., 2006). Obviously, this characteristic in the patrilineal gene pool of East Slavs can be explained by the Mongol invasion and the presence of Mongols in Russia (called Tatars by Russians) between 1237 and 1480 AD (Karamzin, 1811/2016; Deynichenko, 2003).

4. Conclusion

In conclusion, as illustrated by the PC analysis of mtDNA and Y-chromosome haplogroup frequencies, West-East and South Slavic speaking populations, traditionally called “Slavs” (a term introduced in the 16th century AD) (Šafařík, 1848) are heterogeneous based on the uniparental genetic diversity, which shows that they do not share substantial common genetic ancestry and that there is great genetic variety in the Slavic linguistic unity.

Acknowledgements

We would like to thank prof. Antonio Torroni, prof. Ornella Semino and their collaborators (Department of Biology and Biotechnology, University of Pavia, Italy) for their contribution for the development of the uniparental genetic study in Bulgaria.

References

Aeppli, N., von Waldenfels, R., & Samardzic, T. (2014). Part-of-Speech Tag Disambiguation by Cross-Linguistic Majority Vote. *COLING*, 76.

- <https://doi.org/10.3115/v1/W14-5309>
- Balanovsky, O., Rootsi, S., Pshenichnov, A., Kivisild, T., Churnosov, M., Evseeva, I. et al. (2008). Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context. *American Journal of Human Genetics*, 82, 236-250. <https://doi.org/10.1016/j.ajhg.2007.09.019>
- Battaglia, V., Fornarino, S., Al-Zahery, N., Olivieri, A., Pala, M., Myres, N. M. et al. (2009). Y-Chromosomal Evidence of the Cultural Diffusion of Agriculture in Southeast Europe. *European Journal of Human Genetics*, 17, 820-830. <https://doi.org/10.1038/ejhg.2008.249>
- Borri, F. (2011). White Croatia and the Arrival of the Croats: an Interpretation of Constantine Porphyrogenitus on the Oldest Dalmatian History. *Early Medieval Europe*, 19, 204-231. <https://doi.org/10.1111/j.1468-0254.2011.00318.x>
- Chadwick, H. M. (2014). *The Nationalities of Europe and the Growth of National Ideologies*. Cambridge: Cambridge University Press.
- Curta, F. (2001). *The Making of the Slavs: History and Archaeology of the Lower Danube Region, c. 500-700* (Vol. 52). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511496295>
- Derenko, M., Malyarchuk, B., Denisova, G., Wozniak, M., Grzybowski, T., Dambueva, I. et al. (2007). Y-Chromosome Haplogroup N Dispersals from South Siberia to Europe. *Journal of Human Genetics*, 52, 763-770. <https://doi.org/10.1007/s10038-007-0179-5>
- Deynichenko, P. G. (2003). *Russia. Full Encyclopedic Illustrated Book of Reference*. Moscow: OLMA Press.
- Dobrev, P. (2016). *The Spirit of Bulgaria. Bulgarians from 1st until 21st Century AD According to Most Novel Data*. Sofia: Galik.
- Džino, D. (2010). *Becoming Slav, Becoming Croat: Identity Transformations in Post-Roman and Early Medieval Dalmatia* (Vol. 12). Leiden/Boston: Brill. <https://doi.org/10.1163/ej.9789004186460.i-272>
- Grzybowski, T., Malyarchuk, B. A., Derenko, M. V., Perkova, M. A., Bednarek, J., & Woźniak, M. (2007). Complex Interactions of the Eastern and Western Slavic Populations with Other European Groups as Revealed by Mitochondrial DNA Analysis. *Forensic Science International: Genetics*, 1, 141-147. <https://doi.org/10.1016/j.fsigen.2007.01.010>
- Hammer, M. F., Karafet, T. M., Park, H., Omoto, K., Harihara, S., Stoneking, M. et al. (2005). Dual Origins of the Japanese: Common Ground for Hunter-Gatherer and Farmer Y Chromosomes. *Journal of Human Genetics*, 51, 47-58. <https://doi.org/10.1007/s10038-005-0322-0>
- Karachanak, S., Grugni, V., Fornarino, S., Nesheva, D., Al-Zahery, N., Battaglia, V. et al. (2013). Y-Chromosome Diversity in Modern Bulgarians: New Clues about Their Ancestry. *PLoS ONE*, 8, e56779. <https://doi.org/10.1371/journal.pone.0056779>
- Karamzin, N. M. (1811). *History of the Russian State*. St. Petersburg: Russian Imperial Library; Moscow: Editorial House "E", 2016.
- Karlsson, A. O., Wallerstrom, T., Gotherstrom, A., & Holmlund, G. (2006). Y-Chromosome Diversity in Sweden—A Long-Time Perspective. *European Journal of Human Genetics*, 14, 963-970. <https://doi.org/10.1038/sj.ejhg.5201651>
- Kipfer, B. A. (2000). *Encyclopedic Dictionary of Archaeology*. New York, NY: Springer Science & Business Media. <https://doi.org/10.1007/978-1-4757-5133-8>
- Kushniarevich, A., Sivitskaya, L., Danilenko, N., Novogrodskii, T., Tsybovsky, I., Kiseleva, A. et al. (2013). Uniparental Genetic Heritage of Belarusians: Encounter of Rare Middle

- Eastern Matrilineages with a Central European Mitochondrial DNA Pool. *PLoS ONE*, 8, e66499. <https://doi.org/10.1371/journal.pone.0066499>
- Kushniarevich, A., Utevska, O., Chuhryaeva, M., Agdzhoyan, A., Dibirova, K., Uktveryte, I. et al. (2015). Genetic Heritage of the Balto-Slavic Speaking Populations: A Synthesis of Autosomal, Mitochondrial and Y-Chromosomal Data. *PLoS ONE*, 10, e0135820. <https://doi.org/10.1371/journal.pone.0135820>
- Malyarchuk, B., Grzybowski, T., Derenko, M., Perkova, M., Vanecsek, T., Lazur, J. et al. (2008). Mitochondrial DNA Phylogeny in Eastern and Western Slavs. *Molecular Biology and Evolution*, 25, 1651-1658. <https://doi.org/10.1093/molbev/msn114>
- Mielnik-Sikorska, M., Daca, P., Malyarchuk, B., Derenko, M., Skonieczna, K., Perkova, M. et al. (2013). The History of Slavs Inferred from Complete Mitochondrial Genome Sequences. *PLoS ONE*, 8, e54360. <https://doi.org/10.1371/journal.pone.0054360>
- Mirabal, S., Regueiro, M., Cadenas, A. M., Cavalli-Sforza, L. L., Underhill, P. A., Verbenko, D. A. et al. (2009). Y-Chromosome Distribution within the Geo-Linguistic Landscape of Northwestern Russia. *European Journal of Human Genetics*, 17, 1260-1273. <https://doi.org/10.1038/ejhg.2009.6>
- Mutafchiev, P. (1943). *History of Bulgarian People* (Vol. 1). Sofia: Bulgarian House.
- Raykov, M. (2005). *Beginner's Bulgarian*. New York, NY: Hippocrene Books.
- Rebala, K., Martinez-Cruz, B., Tonjes, A., Kovacs, P., Stumvoll, M., Lindner, I. et al. (2013). Contemporary Paternal Genetic Landscape of Polish and German Populations: From Early Medieval Slavic Expansion to Post-World War II Resettlements. *European Journal of Human Genetics*, 21, 415-422. <https://doi.org/10.1038/ejhg.2012.190>
- Rebala, K., Mikulich, A. I., Tsybovsky, I. S., Sivakova, D., Dzapinkova, Z., Szczerkowska-Dobosz, A. et al. (2007). Y-STR Variation among Slavs: Evidence for the Slavic Homeland in the Middle Dnieper Basin. *Journal of Human Genetics*, 52, 406-414. <https://doi.org/10.1007/s10038-007-0125-6>
- Regueiro, M., Rivera, L., Damnjanovic, T., Lukovic, L., Milasin, J., & Herrera, R. J. (2012). High Levels of Paleolithic Y-Chromosome Lineages Characterize Serbia. *Gene*, 498, 59-67. <https://doi.org/10.1016/j.gene.2012.01.030>
- Šafařík, P. J. (1848). *Slavonic Antiquity*. Moscow: Universitetskoi Tipografii.
- Sengupta, S., Zhivotovsky, L. A., King, R., Mehdi, S. Q., Edmonds, C. A., Chow, C.-E. T. et al. (2006). Polarity and Temporality of High-Resolution Y-Chromosome Distributions in India Identify Both Indigenous and Exogenous Expansions and Reveal Minor Genetic Influence of Central Asian Pastoralists. *American Journal of Human Genetics*, 78, 202-221. <https://doi.org/10.1086/499411>
- Smith, J. M. (2005). *Europe after Rome: A New Cultural History 500-1000*. Oxford: Oxford University Press.
- Zhong, H., Shi, H., Qi, X.-B., Duan, Z.-Y., Tan, P.-P., Jin, L. et al. (2010). Extended Y Chromosome Investigation Suggests Postglacial Migrations of Modern Humans into East Asia via the Northern Route. *Molecular Biology and Evolution*, 28, 717-727. <https://doi.org/10.1093/molbev/msq247>

Supplementary

Table S1. Number of individuals belonging to each mtDNA haplogroup in European Slavic speaking populations.

Population	Total	H	HV	J	T	U1	U2	U3	U4	U5a	U5b	U6	U7	U8	U*	K	N1	N2	X	M	L
Belarus ^a	267	99	25	24	20	1	4	6	10	17	15	0	1	0	0	10	15	9	3	7	0
Bosnia ^{b,c}	383	171	37	28	15	3	2	3	18	22	14	1	4	1	4	22	12	11	7	2	1
Bulgarians ^d	885	359	65	67	91	12	12	16	33	37	20	0	4	4	0	50	22	24	18	8	3
Croatia ^c	569	255	43	58	41	6	19	9	13	33	18	1	2	1	5	26	17	10	9	0	1
Czechs ^e	179	79	7	21	22	0	1	4	5	9	6	0	0	1	0	7	8	1	3	3	0
Herzegovina ^c	130	55	14	11	8	2	3	0	6	3	3	0	0	0	0	12	3	3	1	2	0
Kaluga ^f	71	27	4	3	10	1	2	1	6	6	2	0	0	0	0	3	1	0	2	2	1
Poland ^g	413	174	25	32	31	2	7	4	21	31	20	0	0	5	0	18	15	15	8	2	0
Pskov ^f	68	29	4	8	4	0	2	0	2	12	2	0	0	1	0	0	0	2	0	1	0
Serbians ^h	139	66	9	9	7	1	1	2	12	5	4	0	1	0	0	11	3	3	2	1	1
Slovaks ^e	207	94	14	19	22	1	2	0	11	4	12	0	1	0	0	3	11	6	3	2	2
Slovenia ^{b,c}	201	86	10	24	21	1	5	3	7	14	10	0	0	1	0	9	2	6	2	0	0
Tula ^f	73	43	3	4	7	0	0	2	3	2	2	0	0	0	0	2	4	0	0	0	1
Ukrainians ⁱ	607	233	49	50	72	0	9	6	35	36	24	0	0	1	0	30	25	16	8	7	1
Velikii Novgorod ^g	79	33	6	4	5	0	2	1	0	12	3	0	0	3	0	0	3	1	1	2	0
Vladimir ^f	72	25	1	6	14	0	1	1	2	5	0	0	1	0	0	5	6	3	1	1	0
Volot ^g	78	32	10	7	7	0	0	0	1	3	1	0	0	2	0	3	2	2	0	5	0
Yaroslavl ^f	41	20	3	2	1	0	0	0	2	4	0	0	0	0	0	3	3	0	2	1	0

^aKushniarevich et al., 2013; ^bMalyarchuk et al., 2003; ^cSarac et al., 2014; ^dKarachanak et al., 2012; ^eMalyarchuk et al., 2008; ^fMalyarchuk et al., 2004; ^gGrzybowski et al., 2007; ^hDavidovic et al., 2015; ⁱPshenichnov et al., 2013.

References

- Davidovic, S., Malyarchuk, B., Aleksic, J. M., Derenko, M., Topalovic, V., Litvinov, A. et al. (2015). Mitochondrial DNA Perspective of Serbian Genetic Diversity. *American Journal of Physical Anthropology*, 156, 449-465. <https://doi.org/10.1002/ajpa.22670>
- Karachanak, S., Carossa, V., Nesheva, D., Olivieri, A., Pala, M., HooshiarKashani, B. et al. (2012). Bulgarians vs. the Other European Populations: A Mitochondrial DNA Perspective. *International Journal of Legal Medicine*, 126, 497-503. <https://doi.org/10.1007/s00414-011-0589-y>
- Malyarchuk, B., Derenko, M., Grzybowski, T., Lunkina, A., Czarny, J., Rychkov, S. et al. (2004). Differentiation of Mitochondrial DNA and Y Chromosomes in Russian Populations. *Human Biology*, 76, 877-900. <https://doi.org/10.1353/hub.2005.0021>
- Malyarchuk, B. A., Grzybowski, T., Derenko, M. V., Czarny, J., Drobic, K., & Miscicka-Sliwka, D. (2003). Mitochondrial DNA Variability in Bosnians and Slovenians. *Annals of Human Genetics*, 67, 412-425. <https://doi.org/10.1046/j.1469-1809.2003.00042.x>
- Pshenichnov, A., Balanovsky, O., Utevska, O., Metspalu, E., Zaporozhchenko, V., Agdzhoyan, A. et al. (2013). Genetic Affinities of Ukrainians from the Maternal Perspective. *American Journal of Physical Anthropology*, 152, 543-550. <https://doi.org/10.1002/ajpa.22371>
- Sarac, J., Saric, T., Augustin, D. H., Jeran, N., Kovacevic, L., Cvjetan, S. et al. (2014). Maternal Genetic Heritage of Southeastern Europe Reveals a New Croatian Isolate and a Novel, Local Sub-Branching in the x2 Haplogroup. *Annals of Human Genetics*, 78, 178-194. <https://doi.org/10.1111/ahg.12056>

Table S2. Absolute frequencies of Y-chromosome haplogroups in Slavic speaking populations.

Population	Total	C-M217	E-M78	G-P15	G-M285	I-M253	I-P37	I-M223	I-M170 (x M253, P37, M223)	J-M304 (x M172)	J-M172	N-M231	T-M70	Q-M242	R-SRY1532	R-M343 (x M269)	R-M269	R-M124
Arkhangel ^{a,b}	287	0	1	2	0	16	14	0	13	1	5	118	0	1	92	0	22	0
Belarus Centre ^c	88	0	1	1	0	6	16	3	0	1	0	7	0	0	44	0	7	1
Belarus East ^{c,d}	182	0	7	1	0	2	29	1	1	2	3	13	1	0	106	0	14	0
Belarus East Polesie ^c	96	0	1	6	0	4	25	0	0	2	4	8	1	0	41	0	2	0
Belarus North ^{c,d}	201	0	6	0	0	3	26	4	0	1	5	25	0	0	119	0	12	0
Belarus West ^c	73	0	4	0	0	5	2	1	0	1	1	10	0	0	42	1	6	0
Belarus West Polesie ^c	121	0	2	1	0	10	31	0	0	1	4	10	1	1	54	1	5	0
Belgorod ^d	143	1	1	1	0	5	18	1	0	0	6	18	2	0	85	0	4	0
Bosnia-Croats ^e	90	0	8	1	0	0	66	0	0	0	2	0	0	0	11	0	2	0
Bosnians ^e	84	0	11	3	0	4	38	0	0	2	10	0	0	0	13	0	3	0
Bosnia-Serbs ^e	81	0	16	1	0	2	28	1	2	1	7	5	0	0	11	0	5	0
Bulgarians ^f	808	4	158	40	0	35	163	14	3	27	85	4	13	4	138	2	86	1
Croats ^e	118	0	11	5	0	8	38	1	0	0	7	2	0	0	35	0	11	0
Czechs ^e	75	0	0	3	0	7	0	1	0	0	4	2	1	0	31	1	25	0
Kostroma ^a	52	0	3	0	0	6	6	2	0	2	2	7	0	0	17	1	5	0
Kursk ^{a,b}	85	0	5	0	0	4	11	1	0	0	2	11	1	0	46	0	4	0
Orel ^a	110	0	1	0	0	9	15	1	0	0	1	6	3	1	69	0	4	0
Poles ^{a,g}	619	0	17	7	1	56	28	9	14	2	12	17	0	3	371	0	80	0
Pskov ^a	132	0	4	0	0	7	13	0	2	1	1	32	0	0	64	0	5	0
Serbs ^h	103	0	18	6	0	8	30	1	0	1	7	2	0	0	21	0	8	0
Slovakia Western ^g	164	0	6	8	0	9	25	3	0	3	3	5	1	0	74	0	25	0
Slovenians ^e	75	0	2	2	0	5	15	2	1	1	2	0	0	0	29	0	16	0
Smolensk ^a	107	1	8	0	0	2	11	0	0	0	3	15	3	0	49	0	12	0
Tver ^{a,b}	111	0	3	0	0	5	10	3	0	0	5	13	0	1	63	0	6	0
Ukrainians ^{d,e}	748	1	49	21	0	38	145	2	17	3	24	43	9	3	338	1	49	0
Vologda ^a	121	0	1	1	2	14	10	2	0	0	1	47	0	2	40	0	1	0
Voronezh ^a	96	0	1	2	0	3	16	0	1	2	1	6	0	0	57	0	5	1

^aBalanovsky et al., 2008; ^bMirabal et al., 2009; ^cKushniarevich et al., 2013; ^dKushniarevich et al., 2015; ^eBattaglia et al., 2009; ^fKarachanak et al., 2013; ^gRebala et al., 2013; ^hReguiero et al., 2012.

Table S3. Number of individuals per mtDNA haplogroup in European populations.

Population	Total	H	HV	J	T	U1	U2	U3	U4	U5a	U5b	U6	U7	U8	U*	K	N1	N2	X	M	L
Andalusians ^a	279	113	17	19	21	0	2	8	3	4	21	16	1	0	1	20	4	8	6	1	10
Austria ^b	273	119	7	23	36	2	6	3	13	16	9	0	0	0	3	23	4	4	4	1	0
Belarus Central ^c	47	20	6	4	1	0	0	1	2	6	1	0	1	0	0	0	3	1	1	0	0
Belarus East ^c	87	31	9	5	8	0	3	2	2	5	6	0	0	0	0	4	6	3	1	2	0
Belarus North ^c	55	20	1	9	5	1	0	0	3	4	2	0	0	0	0	2	1	3	1	2	0
Belarus West ^c	78	28	9	6	6	0	1	3	3	2	6	0	0	0	0	4	5	2	0	3	0
Bologna ^d	35	16	1	5	3	0	0	1	0	2	0	0	0	1	0	3	1	0	1	0	0
Bosnia ^{e,f}	383	171	37	28	15	3	2	3	18	22	14	1	4	1	4	22	12	11	7	2	1
Bulgarians ^g	885	359	65	67	91	12	12	16	33	37	20	0	4	4	0	50	22	24	18	8	3
Center Portugal ^h	78	47	0	6	5	0	1	1	0	1	2	3	0	0	0	7	0	1	0	0	4
Croatia ^f	569	255	43	58	41	6	19	9	13	33	18	1	2	1	5	26	17	10	9	0	1
Czechs ⁱ	179	79	7	21	22	0	1	4	5	9	6	0	0	1	0	7	8	1	3	3	0
Finland ^j	200	80	9	9	12	0	1	0	4	13	36	0	1	0	0	13	7	5	5	0	0
France ^k	173	74	5	11	22	2	5	0	2	9	4	2	0	0	3	22	5	3	3	0	0
Galicia (Northwestern Spain) ^h	43	22	4	1	3	0	2	0	0	0	0	1	0	0	0	4	0	1	1	0	4
Germany ^l	213	90	0	20	26	1	1	1	12	10	6	0	0	0	3	14	4	7	2	1	0
Herzegovina ^f	130	55	14	11	8	2	3	0	6	3	3	0	0	0	0	12	3	3	1	2	0
Kaluga ^m	71	27	4	3	10	1	2	1	6	6	2	0	0	0	0	3	1	0	2	2	1
La Spezia and Tuscany ^d	52	29	2	3	4	0	0	0	1	1	0	2	0	0	0	7	2	0	0	0	0
North Portugal ^h	84	39	4	2	15	0	1	2	2	1	3	1	0	0	0	4	5	4	0	0	1
Northeast Italy ^d	108	42	8	6	16	1	1	1	2	6	5	0	0	0	0	9	5	4	0	2	0
Northern Greeks ⁿ	319	133	13	33	27	6	1	12	14	12	3	2	2	0	0	16	11	8	12	4	0
Northwest Italy ^d	162	73	14	19	12	1	5	4	3	7	1	0	0	0	0	10	5	2	3	1	0
Poland North ^o	253	109	18	20	25	0	3	1	15	14	12	0	0	3	0	8	11	8	4	1	0
Poland Northeast ^o	73	27	4	6	3	0	0	1	1	12	3	0	0	0	0	7	3	2	2	0	0
Poland South ^o	87	38	3	6	3	2	4	2	5	5	5	0	0	2	0	3	1	5	2	1	0
Pskov ^m	68	29	4	8	4	0	2	0	2	12	2	0	0	1	0	0	0	2	0	1	0
Romania ^p	128	67	9	8	8	0	1	0	7	6	2	0	0	1	0	7	0	7	1	1	0
Sardinia ^d	70	32	3	9	11	0	0	0	0	2	5	0	0	1	0	5	2	0	0	0	0
Serbians ^q	139	66	9	9	7	1	1	2	12	5	4	0	1	0	0	11	3	3	2	1	1
Slovaks ^h	207	94	14	19	22	1	2	0	11	4	12	0	1	0	0	3	11	6	3	2	2
Slovenia ^{e,f}	201	86	10	24	21	1	5	3	7	14	10	0	0	1	0	9	2	6	2	0	0
South Portugal ^h	137	66	4	11	9	0	1	1	3	3	2	0	0	0	0	8	1	4	5	2	17

Continued

Sweden (Saami excluded) ^f	284	134	13	27	23	1	4	0	8	14	11	0	2	2	0	18	8	4	7	4	4
Swedish Saami ^f	37	0	25	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0
Tula ^m	73	43	3	4	7	0	0	2	3	2	2	0	0	0	0	2	4	0	0	0	1
Ukrainians ^s	607	233	49	50	72	0	9	6	35	36	24	0	0	1	0	30	25	16	8	7	1
VelikiiNovgorod ^o	79	33	6	4	5	0	2	1	0	12	3	0	0	3	0	0	3	1	1	2	0
Vladimir ^m	72	25	1	6	14	0	1	1	2	5	0	0	1	0	0	5	6	3	1	1	0
Volot ^o	78	32	10	7	7	0	0	0	1	3	1	0	0	2	0	3	2	2	0	5	0
Yaroslavl ^m	41	20	3	2	1	0	0	0	2	4	0	0	0	0	0	3	3	0	2	1	0

^aHernandez et al., 2014; ^bBrandstätter et al., 2007; ^cKushniarevich et al., 2013; ^dBoattini et al., 2013; ^eMalyarchuk et al., 2003; ^fSarac et al., 2014; ^gKarachanak et al., 2012; ^hGonzález et al., 2003; ⁱMalyarchuk et al., 2008; ^jHedman et al., 2007; ^kRichard et al., 2007; ^lTetzlaff et al., 2007; ^mMalyarchuk et al., 2004; ⁿIrwin et al., 2008; ^oGrzybowski et al., 2007; ^pHervella et al., 2014; ^qDavidovic et al., 2015; ^rTillmar et al., 2010; ^sPshenichnov et al., 2013.

References

- Boattini, A., Martinez-Cruz, B., Sarno, S., Harmant, C., Useli, A., Sanz, P. et al. (2013). Uniparental Markers in Italy Reveal a Sex-Biased Genetic Structure and Different Historical Strata. *PLoS ONE*, 8, e65441. <https://doi.org/10.1371/journal.pone.0065441>
- Brandstätter, A., Niederstätter, H., Pavlic, M., Grubwieser, P., & Parson, W. (2007). Generating Population Data for the EMPOP Database—An Overview of the mtDNA Sequencing and Data Evaluation Processes Considering 273 Austrian Control Region Sequences as Example. *Forensic Science International*, 166, 164-175. <https://doi.org/10.1016/j.forsciint.2006.05.006>
- González, A. M., Brehm, A., Pérez, J. A., Maca-Meyer, N., Flores, C., & Cabrera, V. M. (2003). Mitochondrial DNA Affinities at the Atlantic Fringe of Europe. *American Journal of Physical Anthropology*, 120, 391-404. <https://doi.org/10.1002/ajpa.10168>
- Hedman, M., Brandstätter, A., Pimenoff, V., Sistonen, P., Palo, J., Parson, W., & Sajantila, A. (2007). Finnish Mitochondrial DNA HVS-I and HVS-II Population Data. *Forensic Science International*, 172, 171-178. <https://doi.org/10.1016/j.forsciint.2006.09.012>
- Hernandez, C. L., Reales, G., Dugoujon, J. M., Novelletto, A., Rodriguez, J. N., Cuesta, P., & Calderon, R. (2014). Human Maternal Heritage in Andalusia (Spain): Its Composition Reveals High Internal Complexity and Distinctive Influences of mtDNA Haplogroups U6 and L in the Western and Eastern Side of Region. *BMC Genetics*, 15, 11. <https://doi.org/10.1186/1471-2156-15-11>
- Hervella, M., Izagirre, N., Alonso, S., Ioana, M., Netea, M. G., & de-la-Rua, C. (2014). The Carpathian Range Represents a Weak Genetic Barrier in South-East Europe. *BMC Genetics*, 15, 56. <https://doi.org/10.1186/1471-2156-15-56>
- Irwin, J., Saunier, J., Strauss, K., Paintner, C., Diegoli, T., Sturk, K. et al. (2008). Mitochondrial Control Region Sequences from Northern Greece and Greek Cypriots. *International Journal of Legal Medicine*, 122, 87-89. <https://doi.org/10.1007/s00414-007-0173-7>
- Richard, C., Pennarun, E., Kivisild, T., Tambets, K., Tolk, H.-V., Metspalu, E. et al. (2007). An mtDNA Perspective of French Genetic Variation. *Annals of Human Biology*, 34, 68-79. <https://doi.org/10.1080/03014460601076098>
- Tetzlaff, S., Brandstätter, A., Wegener, R., Parson, W., & Weirich, V. (2007). Mitochondrial DNA Population Data of HVS-I and HVS-II Sequences from a Northeast German Sample. *Forensic Science International*, 172, 218-224. <https://doi.org/10.1016/j.forsciint.2006.12.016>
- Tillmar, A. O., Coble, M. D., Wallerstrom, T., & Holmlund, G. (2010). Homogeneity in Mitochondrial DNA Control Region Sequences in Swedish Subpopulations. *International Journal of Legal Medicine*, 124, 91-98. <https://doi.org/10.1007/s00414-009-0354-7>

Table S4. Absolute frequencies of haplogroups of the Y-chromosome in European populations.

Population	Total	E-M35	G-M201	I-M253	I-P37	I-M223	I-M170 (x M253, P37, M223)	J-M304	N-M231	T-M70	Q-M242	R-SRY1532	R-M343
Arkhangel ^{la,b}	287	1	2	16	14	0	13	6	118	0	1	92	22
Belarus ^{c,d}	761	24	9	30	129	9	1	25	73	3	1	406	48
Belgorod ^a	143	1	1	5	18	1	0	6	18	2	0	85	4
Bologna ^e	29	3	3	3	0	0	0	3	0	0	0	0	17
Bosnia-Croats ^f	90	8	1	0	66	0	0	2	0	0	0	11	2
Bosnians ^f	84	11	3	4	38	0	0	12	0	0	0	13	3
Bosnia-Serbs ^f	81	18	1	2	28	1	2	8	5	0	0	11	5
Bulgarians ^g	808	175	40	35	163	14	3	112	4	13	4	138	88
Catalonia ^h	2309	191	100	44	67	37	11	266	0	28	0	32	1525
Crete ⁱ	193	17	21	8	3	3	11	75	0	4	0	17	33
Croats ^f	118	11	5	8	38	1	0	7	2	0	0	35	11
Czechs ^f	75	0	3	7	0	1	0	4	2	1	0	31	26
Finland-Osterbotten ^j	40	0	2	8	0	0	0	0	26	0	0	3	1
Germany-Bavaria ^k	218	18	7	37	6	6	0	16	2	2	1	27	95
Germany-Mecklenburg ^k	131	7	2	19	3	7	1	4	3	1	0	18	66
Greeks ^{fi}	149	43	6	4	10	2	1	33	0	5	0	21	21
Hungarians ^{fi}	245	18	11	16	35	5	2	21	1	1	1	81	42
Italy Central ^e	77	8	12	1	0	1	1	21	0	0	0	3	29
Italy Northeast ^e	73	11	5	8	0	0	1	11	0	2	0	0	28
Italy Northwest ^e	161	19	13	5	1	1	6	17	0	2	0	5	89
Italy South ^e	198	35	34	7	4	2	1	42	0	9	0	8	51
Kursk ^{a,b}	85	5	0	4	11	1	0	2	11	1	0	46	4
La Spezia and Tuscany ^e	123	14	8	5	1	1	0	14	0	4	0	6	67
Livni ^a	110	1	0	9	15	1	0	1	6	3	1	69	4
Macedonian Greeks ^{fi}	114	24	3	6	15	1	2	21	0	2	0	19	19
Moldavians ^m	125	16	1	6	26	4	0	10	2	1	1	38	20
Ostrov ^a	75	3	0	5	7	0	1	2	21	0	0	34	2
Poles ^{fk}	619	19	8	56	28	9	14	14	17	0	3	371	80
Porhov ^a	57	2	0	2	6	0	1	0	11	0	0	30	3
Repievka ^a	96	1	2	3	16	0	1	3	6	0	0	57	5
Romanians ^{l,m}	232	39	7	12	53	5	5	36	1	1	1	46	25
Roslavl ^a	107	8	0	2	11	0	0	3	15	3	0	49	12
Sardinia ^e	82	8	11	0	32	2	0	8	0	1	0	0	18

Continued

Serbs ^a	103	19	6	8	30	1	0	8	2	0	0	21	8
Sicily ^e	141	25	17	2	2	2	1	36	0	3	0	8	43
Slovakia Western ^k	164	7	8	9	25	3	0	6	5	1	0	74	25
Slovenians ^f	75	2	2	5	15	2	1	3	0	0	0	29	16
Swedish Saami ^j	38	0	0	12	0	0	0	0	17	0	0	6	3
Tver ^{a,b}	111	4	0	5	10	3	0	5	13	0	1	63	6
Ukrainians ^{d,f}	748	54	22	38	145	2	17	27	43	9	3	338	50
Unzha ^a	52	3	0	6	6	2	0	4	7	0	0	17	6
Vologda ^a	121	1	3	14	10	2	0	1	47	0	2	40	1

^aBalanovsky et al., 2008; ^bMirabal et al., 2009; ^cKushniarevich et al., 2013; ^dKushniarevich et al., 2015; ^eBoattini et al., 2013; ^fBattaglia et al., 2009; ^gKarachanak et al., 2013; ^hSolé-Morata et al., 2015; ⁱKing et al., 2008; ^jKarlsson et al., 2006; ^kRebala et al., 2013; ^lMartinez-Cruz et al., 2012; ^mVarzari et al., 2013; ⁿReguiro et al., 2012.

References

- King, R. J., Özcan, S. S., Carter, T., Kalfoglu, E., Atasoy, S., Triantaphyllidis, C. et al. (2008). Differential Y-Chromosome Anatolian Influences on the Greek and Cretan Neolithic. *Annals of Human Genetics*, 72, 205-214. <https://doi.org/10.1111/j.1469-1809.2007.00414.x>
- Martinez-Cruz, B., Ioana, M., Calafell, F., Arauna, L. R., Sanz, P., Ionescu, R. et al. (2012). Y-Chromosome Analysis in Individuals Bearing the Basarab Name of the First Dynasty of Wallachian Kings. *PLoS ONE*, 7, e41803. <https://doi.org/10.1371/journal.pone.0041803>
- Solé-Morata, N., Bertranpetit, J., Comas, D., & Calafell, F. (2015). Y-Chromosome Diversity in Catalan Surname Samples: Insights into Surname Origin and Frequency. *European Journal of Human Genetics*, 23, 1549-1557. <https://doi.org/10.1038/ejhg.2015.14>
- Varzari, A., Kharkov, V., Nikitin, A. G., Raicu, F., Simonova, K., Stephan, W. et al. (2013). Paleo-Balkan and Slavic Contributions to the Genetic Pool of Moldavians: Insights from the Y Chromosome. *PLoS ONE*, 8, e53731. <https://doi.org/10.1371/journal.pone.0053731>