Genetic Geography of the Historical Ancestors of the Chuvash

Anton K. Salmin

Department of Eastern Slavs and Peoples of European Russia, Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint-Petersburg, Russia
Email: antsalmin@mail.ru

Abstract

Discourses concerning the ethnic history of the Chuvash, their origins, and their distant ancestors, have emerged over the course of three centuries. In the 21st century, the science of genetics began to play a major role in this discussion. In the present article, the author, relying on the achievements of genetic geographers, attempts to provide insights into a profoundly difficult nexus of problems and issues. Today it is accepted as a given that the historical ancestors of the Chuvash people—the Savars/Savirs/Suvars, had lived in the Caucasus Mountain region as far back as the 2nd–8th centuries CE. For this reason, research on the ethnogenesis of the Chuvash in the context of the genome history of the Caucasian peoples is more than warranted. There is also no doubt that the genetic lines of this region came directly out of the Near East. This is especially true of males, as it was the male population that first embarked on the migration process. Research has shown that the haplotype frequency of the Chuvash includes components from the Caucasus, Central Europe, Eurasia, and the Mediterranean. Among those, however, haplotypes of Turkish peoples have also been discovered. During the centuries-long contact with peoples of the Middle Volga region, neighboring peoples also had a genetic influence on them. As a result, the original gene pool of the Chuvash people carries traces that are characteristic of the ancient Caucasus, Near East, and Northern Mesopotamia.

Keywords

Chuvash, History, Genetic Geography, the Caucasus, the Near East, the Volga Region

1. Introduction

Today, fortunately, there is no barrier that might sequester or obscure the know-
Haplotype of the historical ancestry of the Chuvash from the basic theory of their origins in the Caucasus region. This is certainly the case as regards the 2nd to the 8th centuries. Preliminary studies have already been completed on this subject. (See, for example (Arnaiz-Villena et al., 2003: pp. 375-392; Salmin, 2013: pp. 197-201, 2014, 2016: pp. 17-34))

The genetic information witnesses to the direct descendence of the Caucasian lines from the Near East, in particular, the male lines. This is understandable: it was the male population that first took part in migratory flows. It is possible that colonization originated in the area between the Tigris River in the East and the North of Syria, and the adjacent East Anatolia in the West. On the whole, the most common haplogroups in the Caucasus were J1*-M267 (34%), J2a4b*-M67 (21%), G2a3b1-P303 (12%), and G2a1a-P18 (18%), which together accounted for 73% of the Y-chromosomes. All four basic Caucasian haplogroups clearly suggest their Near Eastern, rather than European, origins. They are either widespread in the population of the Near East (G2a3b1-P303 and J2a4b*-M67) or have ancestral lines present (G2a1a*-P16 and J1*-M267). Typical European haplogroups are very rare (I2a-P37.2) or limited to specific populations (R1a1a-M198) in the Caucasus. This regularity presupposes a unidirectional genetic flow from the Near East to the Caucasus, which might have occurred during the time of primitive Paleolithic settlement or the Neolithic proliferation of farming that followed (Balanovsky et al., 2011).

As we can see, geneticists write about the proximity of Caucasian populations to the population of the Asian Near East. Only the Kuban Nogays and the Karanogays, who tend toward the population of Europe and Central Asia, fall outside this purview. The formation of the autochthonous population of the Asian Near East is dated to the period following the end of the last ice age (or LGP, Last Glacial Period, approx. 24th c. BCE), in the southeast Circum-Pontic region. Afterward, the tribes migrated from that area to the Caucasus region. The most ancient lines of haplogroup H (standard human mitochondrial sequence, to which all others are compared) are found with the highest frequency on the territory of the Southern Caucasus and the northern part of the Asian Near East. This allows us to assume that the origin and initial evolution of haplogroup H took place precisely in this region. The conclusions of geneticists witness unequivocally to the autochthonous people of the Caucasus and point to the Transcaucasian origin of their genetic substratum. The greatest frequency of haplogroup H in the populations of the Caucasus and the Asian Near East occurs among Armenians, followed by Karachaevs, Balkars, and the peoples of Dagestan (Khusnutdinova et al., 2012: p. 752).

2. The Historical Forebears of the Chuvash According to Genogeographic Data

Haplogroup U forms a significant portion of the gene pool of the Caucasian population. The majority of these subclades originated and evolved in the region...
of the Asian Near East. This allows us to assume that the U subgroups, dispersed throughout the region of the Caucasus, originally came from the Asian Near East. For example, U1 is almost never encountered in Europe and Central Asia, but is commonly found among the populations of the Asian Near East, which together with its relatively low frequency may witness to the penetration of U1 in this region precisely out of the Asian Near East. The study of the main components of the Caucasian population demonstrates that this region falls into two separate clusters: one representing the Northwestern and Southern Caucasus, and the other representing the Northeast (Kutuev, 2010: pp. 15, 33, 36).

Regarding the Y-chromosome, here, a frequency distribution pattern of haplogroups specific to the autochthonous population of the Caucasus, can be clearly discerned: a high frequency of occurrence of G-M201, J-12f2 and R1-M173. A similar frequency distribution is characteristic only for the autochthonous population of the Caucasus, and is not seen anywhere else in the world. Haplogroup G-M201 is primarily encountered in the Caucasus, the Asian Near East, and the Mediterranean. The ethnic groups of the Northwest and Southern Caucasus are characterized by a higher proportion of haplogroups J2-M172 and G-M201, whereas their frequency drops significantly in the Northeast Caucasus, and the majority is represented by haplogroup J1-M267. For example, geneticists point out the connection of a sub-variant of haplogroup J2-M172 of the Abkhazians with the peoples of Transcaucasia (Teuchezh, 2013: p. 22). In addition, examining the frequency of the aforementioned major haplogroup, two separate regions within the Caucasus may be delineated. To wit, the populations of the Northwestern and Southern parts of the region are characterized by a higher frequency of haplogroups G2-P15, J2-M267, and R1a1a-M198. This region includes populations of Abkhazians, Adygeis, Balkars, Cherkessians, Kabaradians, Karachais, Ossetians, Megrelians, Armenians, and Abkhazians. For the Northeast Caucasus, represented by populations of the Andi, Avars, Bagualins, Chalmalins, Chechens, Dargins, Ingushetians, Kumyks, Lezgins, and Tatars, a high rate of haplogroups J1-M267 and R1b1b2-M269 may be observed (Khusnutdinova et al., 2012: pp. 754-755). There is a particularly high rate of occurrence of J1 among the minor native populations of Dagestan. For example, among the Hunzibs, it reaches 100%; the Tsezians, more than 99%; the Bezhtins, around 98% (Radzhabov et al., 2014: p. 219). Thus, the population of the Northeast Caucasus, on the one hand, and the Northwest and Southern Caucasus, on the other, are divided to a significant degree by genetic markers. At the same time, a general Asian Near-Eastern origin of the genetic substratum is evident in them. This picture, to a large extent, corresponds to the notion of two directions of settlement in the Caucasus connected with the Western Caspian Sea region, and the Eastern Black Sea region, well known since ancient times.

The study of the ethnic groupings of the Caucasus using Y-chromosome markers, mitochondrial DNA, as well as genome-wide association studies demonstrate the presence primarily (except in the Nogays) of an Asian Near-Eastern
The Asian Near-Eastern ancestral component characteristic of the peoples of the Caucasus is almost absent in their immediate northern neighbors—the peoples of the Eastern European plains (Khusnutdinova et al., 2012: p. 750).

Variations in the autosomal (not sex-chromosome) genes in the Caucasus show significant uniformity among the ethnic and linguistic diversity of the population, and correspond primarily to the version of the Near Eastern origin of the Caucasians, with insignificant external influences. In contrast to the autosomal and mtDNA variations, the indicators of the regional Y-chromosomes allow us to distinguish the East and West Northern Caucasians. In addition, the tribes of the Caucasus and the Eastern Europe in the period after the LGP, possibly served as a refugium for the peoples of Eastern Europe. This occurred 26,000 years ago. At that time, the first settlers of the Caucasus reached the Northern regions (Yunusbayev et al., 2012).

On the whole, the distribution of the mtDNA haplogroups of and the Y-chromosomes, as well as the analysis of lines in the populations of the Caucasus, witness to the fact that the genetic population of the Caucasus forms an autochthonous substratum.

An analysis of the autosomal data shows that the cluster (the unification of several homogeneous elements) \( K = 7 \) is the basic genealogical component of the Caucasian peoples. This cluster is present in the Near East, but is virtually absent among the immediate northern neighbors—the populations of the Eastern European plains. According to autosomal data, however, there is some evidence of an opposing stream of genes from Eastern Europe in the Northern Caucasus, which is supported by the presence of Y-chromosomes of haplogroup R1a1. Hence, the presence of genetic gaps between the Caucasus and Eastern European plains. This witnesses to the fact that the nucleus of the autosomal-genetic structure of the populations of the Caucasus was formed before its contemporary linguistic diversity.

The map of the gene pool of the Caucasus reveals the region to be an arena of penetration of the Northern Steppe and Asia Minor-Near Eastern impulses. At the same time, paths of penetration, such as through the Daryal and Caspian Shore, played an important role. Overall, the map demonstrates a Near-Eastern path of settlement of the Caucasus (Kovalevskaya, 2016: pp. 49-50).

At the same time, haplogroup H is among the groups with the highest frequency of distribution among the majority of peoples of the Volga-Ural region (the Bashkirs, Tatars, Chuvash, Mordvinians, and Mari). By distribution frequency in the populations of the Volga-Urals region, the U cluster occupies an important place. This is considered to be one of the most ancient, its evolutionary age being 50,000 years. This cluster has the most significance among the Chuvash (43.6%), the Komi-Zyryan (37.1%), the Tatars (29.4%), and the Bashkirs (28.5%). In the Volga-Urals region, haplogroups U5 and U4 were most widely dispersed. Along with the Finno-Ugrian populations, haplogroup U5 appears
with a great frequency among the Bashkirs, Tatars, and Chuvash. Sub-cluster U3 in Europe is rarely encountered, and has not been identified in Central Asia. A high level of genetic diversity of haplogroup U3 appears in the populations of the Caucasus (Ossetia, Georgia, Armenia) and Turkey. Among the populations of Tatars and the Chuvash, this sub-cluster is represented by two lines, found earlier among the Turks, Armenians, and Nogay (Bermisheva et al., 2002: pp. 992-994, 996).

Geneticists see the uniqueness of the gene pool of the Chuvash, in comparison with their neighbors, first and foremost in the sharply increased frequency of haplogroups E and J. But these haplogroups are typical for the Asian Near East. The frequency of haplogroup J is very characteristic of the peoples of the Caucasus, as well. For example, among the Kubachins, J1* makes up 99%; the Kajtags, 85%, and J2a4b* among the Ingush makes up 87% (Balanovsky, 2015: pp. 79, 181). Research has been done on the structure of the gene pool of the indigenous population of Dagestan, who belong to the Tsezian group of peoples through markers of Y-chromosomes. Populations representing the Tsezians, Hunzibs, Bezhtins, and Hinukhs. A fundamental characteristic of their gene pool is a very low genetic diversity, due to the predominance of haplogroup J1. Strong founder effects may be observed. It is well known that the Tzezian groups live in proximity to Georgian regions. Here, the frequency of haplogroup J1 reaches almost 100%. “The revealed structure of the gene pool witnesses to the strong and lasting genetic isolation of the Tzezian ethnic groups from other Caucasian peoples, and the preservation among them of the original haploid composition, characteristic of the first indigenous inhabitants of West Dagestan” (Glazunova et al., 2016: p. 30). As we can see, through the E and J indicators, the modern Chuvash are strongly connected to the peoples of the Caucasus and the Asian Near East.

Among the Uralo-Volga Turkish peoples, the Chuvash evidence the highest frequency of haplogroups H, K, U2 and V, which, doubtless, witness to the high percentage of mtDNA lines in their gene pool (Bulygina & Soloviev, 2013: pp. 73-74). To be more specific, one must acknowledge the Western Eurasian origins of the historical ancestors of the Chuvash. More precisely, they have historical and genetic ties with the autochthonous population of the Caucasus, including Transcaucasia.

The frequency of distribution of Y-chromosome haplogroups among the peoples of the Volga-Ural region suggests a predominance here of haplogroup R1a1a-M198 (Table 1 in (Trofimova et al., 2015)). The highest rate among the Mordvinians is 42.4%. Among the Chuvash, this rate reaches 29.5%. The Bessermians are close to the Chuvash in frequency of this group (32.1%). In this regard, the peoples of the Volga-Urals region are close to the autochthonous peoples of the Caucasus: among them, R1a1a-M198 also counts among the number of predominant major haplogroups. At the same time, the Chuvash have more similarities to the peoples of the Northwest and Southern parts of the Caucasus.
One must, however, keep in mind that haplogroup R1a1a-M198 was originally absent from modern Dagestan (Khusnutdinova et al., 2012: p. 755). The study of the populations in the genetic space of the Eurasian Steppe and adjacent regions by a set of 24 haplogroups of the Y-chromosome demonstrates the proximity of the Chuvash, first and foremost, to the Karachaevs and the Balkars (Figure 3 in Skhalyakho et al., 2013).

Among the peoples of the Volga-Urals region, however, is a fairly high indication of haplogroup N-M231, which has parallels with South and West Asia, as well as Inner Asia and Southern Siberia (Trofimova et al., 2015: p. 123). Among the Chuvash, 27.3% have N-M231. The Kazan Tatars show indicators close to this, at 28.3%. The highest N-M231 indicators are seen among the Mari, at 79.5%. By distribution frequency haplogroup N in the mtDNA among the populations of the Volga-Urals region, the Chuvash occupy the first place. This indicator reaches 69% among them. Among the Burzyan Bashkirs and Kazan Tatars, respectively, it numbers 57% and 56%. A nearly identical distribution frequency of Y-chromosome haplogroups is seen among the populations of the Volga-Ural region. The Chuvash, by these indicators, are close to the Mari, Komi, and Tuimazin Tatars. According to mtDNA analysis, on the whole, the Chuvash in the Volga-Ural region are closest of all to the Kazan Tatars and furthest of all from the Perm Bashkirs (Trofimova, 2015: pp. 8, 16, 17). From the perspective of haplogroup N, one must acknowledge the significant influence on the Chuvash of the tribes of Inner Asia and Southern Siberia, even more than on the Uralo-Volga Bashkirs and Tatars. At the same time, according to the DNA-loci of the nuclear genome, overall the Chuvash in the Volga-Ural region are closer to the Mari and the Udmurts than to the Tatars and Bashkirs. The genetic similarity of the population of Chuvash to Udmurts and Mari may be explained by their territorial proximity, as well as the historical mixing of the Chuvash with the Mari (Vasilieva et al., 2013: pp. 3, 18).

Contemporary studies of Europoid components in the mitochondrial (extranuclear) gene pool of the peoples of the Vola-Urals region reveal a large degree of Europoidism in the populations of Komi (83%) and Morvinians (76%), and a lesser amount among the Mari (53%) and Chuvash (63%). At the same time, analysis of the DNA-loci suggests that the Chuvash in the region demonstrate the largest proportion of Europoidism (98.3%, and less among the Bashkirs, at 66.3%).

This whole genetic picture witnesses, at least, to the combination in populations of Chuvash of the three basic genomes having Caucasian, European, and Southern Siberian convergence.

The samples from ninety-six unrelated individuals from Chuvashia were sequenced according to indicators of mitochondrial DNA molecules. The majority of the people studied, like all representatives of Western and Northern Europeans, demonstrate the presence of haplogroups H (31%), U (22%), and K (11%), but these indicators are absent among the Altai or Mongol populations.
Thus, studies of mtDNA markers reveal that the indigenous Chuvash are not related to the Altai and Mongol populations through the maternal line (Graf et al., 2012: pp. 17-67). At the same time, analyses of separate genes in the population of the Chuvash witnesses in favor of the Mongolidness (the 240 allele, for example), and some (the allele 244, for example) in favor of Europoidness (Khusnutdinova et al., 2003: p. 1557). While the Chuvash exchanged their autochthonous language for Turkish, the gene pool remained fundamentally unchanged.

The study of alleles of leukocyte antigens (Human Leukocyte Antigens, HLA) by means of DNA analysis is extraordinarily significant. The analysis of HLA alleleism, together with other genetic markers, represents a valuable instrument for tracing the migration of ancient peoples, as well as for determining the ethnic makeup of modern populations. This system is a useful instrument for studying the origin of particular groups of people, among them the Chuvash.

3. Conclusion

The haplotype frequency of the Chuvash attests to various contributions of Central Europeans and Caucasians, Eurasians and Mediterranean peoples. On the other hand, particular Turkish alleles or haplotypes (by comparison of the Chuvash with the modern Turkish population) were not found among the Turkish speaking Chuvash population. Several haplotypes of the Urals peoples found among the Chuvash population reflect the definite influence of populations living in areas adjacent to the Chuvash. The kinship of the Chuvash with Caucasians and the people of the Near East shows that the gene pool of the Chuvash is characteristic for the populations of ancient Caucasus, the Near East and Northern Mesopotamia.

Acknowledgements

The study was performed under the R&D plan of the Peter the Great Museum of Anthropology and Ethnography (the Kunstkamera) of the Russian Academy of Sciences “Factors of ethnocultural identity”.

Conflicts of Interest

The author declares no conflicts of interest.

References


Balanovsky, O., Dibirova, K., Dybo, A., Mudrak, O., Frolova, S., Pocheshkhova, E., Haber, M., Platt, D., Schurr, T., Haak, W., Kuznetsova, M., Radzhabov, M., Balaganskaya, O.,


Kutuev, I. A. (2010). *Genetičeskaia struktura i molekuliarnaia filogeografija narodov Kav-kaza [Genetic Structure and Molecular Phylogeography of the People of the Caucasus].* Avtoreferat doktorskoj dissertacii. In-t bioximii i genetiki UNC RAN.


A. K. Salmin


