

Haplogroup R1a-M198 in Turkic-speaking populations of the North Caucasus

Dzhaubermezov M.A.^{1,2},
Ekomasova N.V.^{1,2},
Litvinov S.S.¹,
Gabidullina L.R.¹,
Khusnutdinova E.K.^{1,2}

¹ Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences, Ufa

² Bashkir State University, Russia, Ufa



Objective: to study the genetic structure and diversity of gene pools of populations of the North Caucasus based on the variability of the haplogroup R1a-M198 of the Y chromosome.

Material and methods. The analysis of the genetic diversity of the Caucasian populations was carried out according to the results of the variability of the haplogroup R1a-M198 of the Y chromosome.

Results. Based on the results of genotyping, it was revealed that the haplogroup R1a-Z2123 is the major haplo group in the Turkic-speaking populations of the Central and Western regions of the North Caucasus, while it turned out to be uncharacteristic for the Turkic-speaking populations of Dagestan and the Kuban Nogai. We conducted an analysis of the frequency distribution of the branches of the haplogroup R1a-M198 of the Y chromosome in the Balkarian population, taking into account their belonging to different subethnic groups (Baksans, Chegemians, Kholamets, Bezengi, Malkars).

Conclusion. The subethnic groups of Bezengis and Kholamets are characterized by the lowest frequency of lines of the haplogroup R1a-M198, and in the subpopulation of cholama, only haplogroup R1a-Z2123 is detected.

Keywords: genetic structure, gene pool, haplogroup R1a-M198, turkic-speaking population, North Caucasus

For citation: Dzhaubermezov M.A., Ekomasova N.V., Litvinov S.S., Gabidullina L.R., Khusnutdinova E.K. Haplogroup R1a-M198 in Turkic-speaking populations of the North Caucasus. Ecological medicine 2019;2(1):33-38. DOI: 10.34662/2588-0489.2019.2.1.33-38.

For correspondence: Murat A. Dzhaubermezov - Institute of Biochemistry and Genetics, Ufa Scientific Center, Bashkir State University, Russia, e-mail: murat-kbr@mail.ru.

Translated from Ecologicheskaya medicina 2019. Vol. 2, N 1, pp. 35-41.

Translated by Kakvaeva S.B.

Active development of new lands and the aggressive campaigns of the Turkic-speaking tribes led to the settlement of the vast territories of the Eurasian continent by the peoples of this language group. Representatives of the Turkic group of languages in the Central and Western Caucasus are populations of Balkarians and Karachais living in the mountains and foothills of the region. In the course of many territorial transformations that took place in the Russian Empire and in the former Union of Soviet Socialist Re-

publics (USSR), the single Karachay-Balkarian people (endo-ethnonym - taulu) was administratively divided into two entities. Currently, most Karachais and Balkars live compactly in the Karachay-Cherkess and Kabardino-Balkarian republics, which are part of the Russian Federation. Among the Balkars, 5 sub-ethnic groups are distinguished: Baksans, Chegemians, Kholamts, Bezengis and Malkars. In some early works, the population of various subpopulations of Balkars is referred to as peoples at all [1]. At different times, linguists distin-

guished up to 4 different dialects in the Balkar language [5-10]. According to the 2010 census, the number of Balkars in Russia is 112.9 thousand people, Karachais - 218.4 thousand people [11]. For the period between the censuses of 1897 and 2010 the number of Balkars and Karachais (in general) increased from 50 thousand to 331,327, i.e. more than 6 times. In general, the number of Balkars and Karachais in Russia and foreign countries totals about 460 thousand people [12]. As a result of greater interest on the part of historians in the problems of ethnogenesis of the Balkarians and Karachais, 2 conferences were devoted to this subject. The first one took place in 1959 in the city of Nalchik, with the participation of leading academic institutions of the USSR, where, comprehensively, with the involvement of specialists in the field of historiography, archeology and anthropology, the whole spectrum of the historical process of formation and development of these peoples was considered [14]. The second conference was held in 2014 at the Institute of Ethnology and Anthropology named after N.N. Miklouho-Maclay of Russian Academy of Sciences with the involvement of specialists in the field of human population genetics [15]. The results of these conferences testify in favor of the fact that various North Caucasian, Iranian-speaking and Turkic-speaking tribes played the main role in the ethnogenesis of these peoples [13, 14].

To date, a number of studies have been carried out on the anthropological characteristics of the Balkars and Karachays [15-17]. They studied a wide range of anthropological disciplines, such as: craniology (studying the variation in the shape of the skull), somatology (studying the anthropometric characteristics), odontology (studying the dentoalveolar system), dermato-glyphics (studying the pattern of the skin). As a result of a comprehensive study, conclusions were drawn about: the Caucasian physical type of the Balkars and Karachays, the indigenous origin of the Balkars and Karachays and their kinship with representatives of neighboring peoples [14-16].

To assess the genetic relationship of the studied populations with the neighboring

peoples of the region, as well as to better understand the processes of ethnogenesis, we considered it necessary to conduct a detailed study of the Karachai population and sub-ethnic groups of the Balkars using the latest markers of the haplogroup R1a-M198 of the Y chromosome.

Objective: to study the genetic structure and diversity of the gene pools of the populations of the North Caucasus according to the results of the variability of the haplogroup R1a-M198 of the Y chromosome.

Materials and methods

The study used samples from 2002-2015 from 22 populations of the Caucasus from the "Collection of Human Biological Materials" Institute of Bioregulation and Gerontology Ufim Federal Research Center of the Russian Academy of Sciences, supported by the Program of Bioresource Collections Federal Agency of the Russian Federation, a total of 1942 samples belonging to populations of the North and South Caucasus. As a result of the questionnaire, the ethnicity of the subjects was established, indicating ancestors up to the third generation. All participants received informed consent to participate in the study.

DNA was isolated from peripheral blood by the standard phenol-chloroform extraction method [18]. To determine the haplogroups of the Y chromosome, 13 markers of the non-recombining region of the Y chromosome were analyzed: M9, M198, PAGE07, Z282, Z93, Z95, M458, M558, M582, Z2125, Z2123, Z2122, M207. These polymorphisms were analyzed for the most part using appropriate restriction enzymes. In their absence, the analysis was carried out by sequencing. Sequencing was performed on an automated sequencer Applied Biosystems (ABI) 3730 XL DNA Analyzer or Applied Biosystems (ABI) 3500. The sequencing results were analyzed using the ChromasPro 2.4.1 software and Sequencer 5.1 (GeneCode Corp).

Results and discussion

Genotyping data indicate that 7 out of 33 haplogroups found in the studied samples belong to the branch of one of the most widespread haplogroups in the world - R1a. More than 10% of men living in the territory from South Asia to Scandinavia belong to

this haplogroup. The distribution of R1a subhaplogroups in the world shows its clear division into two main branches: the European (R1a-Z282) and the Central and South Asian (R1a-Z93) [20]. We also studied additional markers within the haplogroup R1a, the distribution data of which were not previously published in the study of the genetic structure of the peoples of the world. In view of this, we activated samples from the genetic bank of the Institute of Bioregulation and Gerontology Ufim Federal Research Center of the Russian Academy of Sciences - 1942 sample from 22 populations of the North and South Caucasus. It was revealed that the dominant line of the haplogroup R1a-M198 of the Y chromosome in the populations of Karachais (30.1%) and Balkars (16.2%) is R1a-Z2123 (table 1). In other peoples of the region, this haplogroup is much less common, and maximum frequencies are observed among representatives of the Adyghe-Abkhaz language group: Circassians (7.9%), Abazin (7.9%), Q-Bardins (6.4%), Adyghe people (4.5%) and Abkhazians (1.8%). Speaking about other Turkic-speaking peoples of the North Caucasus, this haplogroup has the maximum value among the Karanogais (3.9%), while in the Kumyks its frequency is 1.4%, and it was not found at all in the Kuban Nogai population. In other populations, the haplogroup R1a-Z2123 is found in isolated cases. In the sub-populations of the Balkars, it is distributed almost evenly, demonstrating the maximum value among the Chegemians (20.3%) and the minimum among the Kholamets (10.7%).

Of particular interest is the high frequency of the haplo group R1a-Z2123 in the Bashkir population, which reaches 44% in some ethno-geographical groups (our unpublished data). But the data obtained by sequencing the complete sequence of the Y chromosome in the Bashkir and Balkars attributed them to the various branches of this haplo group, combining, in turn, the Bashkir with a representative of the Jewish community from India [20]. Despite the poor knowledge of this marker in world populations and the lack of large-scale population studies on it, relying on data on

the frequency distribution of the ancestral haplogroup R1a-Z2125, which was found with high frequencies in populations of Kyrgyz and Uzbeks, Turks, Afghan Pashtuns (40%), and among other peoples of Afghanistan, and in some Iranian populations [19], we can assume that the presence of the haplogroup R1a-Z2123 in the populations of Karachais and Balkars South and Central Asian influence. On the other hand, analysis of the complete sequence of the Y-chromosome with R1a-Z2123 in the Balkar showed the presence of a separate deep branch [20], which indicates the earlier presence of this haplogroup in the Caucasus. In this regard, the data obtained as a result of the study of the remains of representatives of the Srub-Scythian archaeological cultures, which are direct descendants of the so-called Kurgan historical and archaeological community and the identification of East Eurasian haplogroups in their environment, are of particular interest. R1a (Z93, Z94, as well as Z2123) [21-25]. Considering the fact that the major haplogroup in the populations of Karachais, Balkarians, and Bashkirs living in the ancient distribution area of the Srub and Scythian cultures is haplogroup R1a-Z2123, the probable connection of these peoples with Balkars and the Karachais look logical. The opinion existing in historical literature about the ubiquitous Iranian language of the Scythians does not seem so correct, especially in the light of recent works showing the exceptional connection of the Eastern Scythians with Turkic-speaking peoples, especially those whose languages belong to the Kypchak group of the Altai language family [26].

At the same time, groups of western Scythians have much in common with populations of the Caucasus and Central Asia [26–33]. In populations of North and South Ossetians, which, according to historiography, are direct descendants of Scythians and Alans, the haplo group R1a-Z2123 was found only in a single case. A similar picture is also observed in the work devoted to the Don Alans, where, in addition to the haplogroups of the Y chromosome characteristic for the Caucasus cascade, haplogroups R1a-Z94 and R1a-Z95 were also

found [34]. This hypothesis still needs to be taken with caution, since the haplogroup R1a-Z2123 with extremely low frequencies was found in the population of Turkic-speaking Volga Tatars and Chuvashs and, at the same time, the ancestral haplogroup for R1a-Z2123-R1a-Z2125, albeit with lower frequencies than the Kyrgyz, found among Iranian-speaking Tajiks and Pashuns. The presence and continuity of R1a-Z2123 from the designated ancient cultures requires a detailed study using new homogenous markers, as well as the results of the complete Y-chromosome sequencing of ancient samples of designated cultures.

With a much lower frequency, the haplogroup R1a-Z2122 is found in the studied populations. With a frequency of 2.4%, she was found in the Karachai population and 3.4% in the Balkarian population. Interestingly, in the Chegemian sub-population, the frequency R1a-Z2122 is 6.8%, which is the absolute maximum for all the studied peoples of the region. All samples were genotyped for the M582 marker, but the results were negative. Previously, the haplogroup R1a-Z2122 (xM582) was found in native speakers of the New Aramaic, while R1a-M582 was characteristic of the Jewish population with their bifurcation about 4000 years ago [20, 35]. For an in-depth study of this haplogroup, 23 Y-STR haplotyping was performed, followed by the construction of a median network. The extremely scarce Y-STR data with haplotypes from populations of other regions did not allow us to determine the genetic relationship with other regions, but nevertheless a rather significant variability was revealed within our sample. Separately, it is necessary to say about the frequency distribution of the haplogroups R1a-M458 and R1a-M558, which are mainly found in Central and Eastern Europe. In this case, the peak of occurrence of the haplogroup R1a-M458 falls on Central Europe, and its frequency decreases to the Urals, while R1a-M558 has several peaks, including in the Volga-Ural region [19]. In this regard, these haplogroups can be used

as markers for the distribution of the East European gene pool. In the populations studied by us, these haplogroups were detected with low frequencies or were not detected at all. This testifies in favor of a slight East European influence on the studied populations along the male line.

Conclusions

Thus, we can conclude that the major haplogroup in the Turkic-speaking populations of the Central and Western Caucasus (Balkars and Karachays) is the haplogroup R1a-Z2123, which is not typical for the Turkic-speaking populations of Dagestan, as well as the Kuban Nogais.

This work was carried out as part of the state task of the Ministry of Education and Science of the Russian Federation (No. AAAA-A16-116020350032-1) with the support of the Russian Foundation for Basic Research "Studying the genetic structure of the Turkic-speaking peoples of Northern Eurasia according to homogeneous markers" No. 17-44 -020748 p_a. DNA samples for research were taken from the Collective Use Center for the Human Biology Institute of Bioregulation and Gerontology, Federal Research Center of the Russian Academy of Sciences supported by the Biological Resource Collection Program of the Federal Agency for Scientific Organizations of Russia (agreement No. 007-030164/2).

Information on financing and conflict of interest

The authors declare no apparent and potential conflicts of interest related to the publication of this article.

Search and analytical work on the preparation of the manuscript was carried out at the personal expense of members of the team of authors.

The participation of the authors: the concept and design of the research - Dzhaubermmezov M.A., Khusnutdinova E.K.; collection and processing of materials - Ekomasova N.V., Litvinov S.S.; analysis of the received data, writing the text - Gabidullina L.R.

References

1. Berger AP. A brief overview of mountain tribes in the Caucasus. Nalchik: Polygraph-combi-nat them. Revolution of 1905, 1992. 48

pages. [Berzhe AP. Kratkii obzor gorskikh plemen na Kavkaze Nalchik Poligrafkombi-nat im Revoliutsii 1905 g 1992. 48 (In Russian)].

2. Tulchinsky NP. Five Mountain Societies of Kabarda. Vladikavkaz: Type. Tersky region Board, 1903. 86 [Tulchinskii NP. Piat gorskikh obshchestv Kabardy. Vladikavkaz: Tip. Terskogo obl. Pravleniia, 1903. 86 (In Russian)].
3. Filonenko VI. Grammar of the Balkar language. Phonetics and morphology. Nalchik: Kabardino-Balkarian State Publishing House, 1940. 88 [Filonenko VI. Grammatika balkarskogo iazyka Fonetika i morfologiya. Nalchik Kabardino-Balkarskoe gosudarstvennoe izdatelstvo, 1940. 88 (In Russian)].
4. Filonenko VI. Mysteries of the Highlanders of the North Caucasus. Scientists notes. Pyatigorsk, 1957. [Filonenko VI. Zagadki gortsev Severnogo Kavkaza. Uchenye zapiski. Piatigorsk, 1957. (In Russian)].
5. Appaev AM. Dialects of the Balkar language in their relation to the Balkar literary language. Nalchik: Kabard.-Balkar. Prince publishing house, 1960. 78 [Appaev AM. Dialekty balkarskogo iazyka v ikh otnoshenii k balkarskomu literaturnomu iazyku Nalchik Kabard-Balkar. Kn. izd-vo, 1960. 78 (In Russian)].
6. Akbaev ShKh. Phonetics of the dialects of the Kara-Chay-Balkar language. (Experience of comparative and comparative historical study). Cherkessk: Nalchik: Kabard.-Balkar. Prince publishing house, 1963. 166 [Akbaev ShKh. Fonetika dialektov Kara-chaevo-Balkarskogo iazyka Opyt sravnitelnogo i sravnitelno-istoricheskogo izucheniia Cherkessk Nalchik Kabard-Balkar. Kn. izd-vo, 1963. 166 (In Russian)].
7. Federal State Statistics Service <http://www.gks.ru>.
8. Khapaev SA. Karachais. Balkarians. Moscow, 2014. 815 pp. [Khapaev SA. Karachaevtsy. Balkartsy. Moskva, 2014. 815 s. (In Russian)].
9. Materials of the scientific session on the problem of the origin of the Balkarian and Karachai peoples. Nalchik: Kabard.-Balkar. Prince publishing house, 1960, 322 [Materialy nauchnoi sessii po probleme proiskhozhdeniia balkarskogo i karachaevskogo narodov Nalchik Kabard-Balkar. Kn. izd-vo, 1960, 322 (In Russian)].
10. Karachais. Balkarians. Ed. M. D. Karaketov, X.-M.A. Sabanchiev. Moscow: Science, 2014. 815 p. [Karachaevtsy Balkartsy. Otv. Red. M.D. Karaketov Kh.-M.A. Sabanchiev. M., Nauka, 2014, 815 s. (In Russian)].
11. Dzhanderidze GK. On the origin of the Balkars and Karachais. Materials of the scientific session on the problem of the origin of the Balkarian and Karachai peoples. Nalchik: Kabard.-Balkar. Prince publishing house, 1960. 335 [Dzhanderidze GK. O proiskhozhdenii balkartsev i karachaevtsev Materialy nauch-noi sessii po probleme proiskhozhdeniia balkarskogo i karachaevskogo narodov. Nalchik, Kabard-Balkar. Kn. izd-vo, 1960. 335 (In Russian)].
12. Alekseev VP. The origin of the peoples of the Caucasus. Craniological study. Moscow: Science, 1974. 320 [Alekseev VP. Proiskhozhdenie narodov Kavkaza Kraniologicheskoe issledovanie. M.: Nauka, 1974. 320 (In Russian)].
13. Khit GL. Dermatoglyphics and racogenesis of the Caucasian population. The Ancient Caucasus: Retro-Specification of Cultures. XXIV Krupnov readings on archeology of the North Caucasus. Moscow. 2004; 198-200 [Khit GL. Dermatoglifika i rasogenez naseleniia Kavkaza. Drevnii Kavkaz retrospektsiia kultur XXIV Krupnovskie chteniia po arkheologii Severnogo Kavkaza. M. 2004; 198-200 (In Russian)].
14. Mathew CG. The isolation of high molecular weight eukaryotic DNA. Methods Mol. Biol. 1984;(2):31-34
15. Grudianov AI. Means and methods of prophylaxis of inflammatory diseases of periodontitis. Moscow. 2012. S. 92 [Grudianov AI. Sredstva i metody profilaktiki vospalitelnykh zabolevanii parodonta. M. 2012. S. 92 (In Russian)].
16. Underhill P, Poznik D, Rootsi S, et al. The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. *European Journal of Human Genetics* 2015;(23)1;124-31.
17. Karmin M, Saag L, Vicente M, et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res* 2015;(25):4. P. 459-66.
18. Krivtsova-Grakova OA. Steppe Volga region and Black Sea coast in the Late Bronze Age. Moscow: MIA, 1955;(46). [Krivtsova-Grakova OA. Stepnoe Povolzhe i Prichernomore v epokhu pozdnei bronzy. M.: MIA, 1955;(46) (In Russian)].
19. Merpert IYa. The oldest herders of the Volga-Ural interfluvium Moscow: Nauka, 1974. 173 [Merpert IYa. Drevneishie skotovody Volzhsko-Uralskogo mezhdurechia M.: Nauka 1974. 173 (In Russian)].
20. Miziev IM. The history of Balkaria and Karatea from ancient times to the campaigns of Timura. Nalchik: Publishing Center, 1996. 372 [Miziev IM Istoriia Balkarii i Karachaia s drevneishikh vremen do pokhodov Timura. Nalchik: Izdatelskii tsentr, 1996. 372 (In Russian)].

21. Wells RS, Yuldasheva N, Ruzibakiev R, et al. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* 2001;102:44-9.
22. Mathieson I, Lazaridis I, Rohland N, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 2015;(528):7583:499-503.
23. Unterlander M, Palstra F, Lazaridis I, et al. Ancestry and demography of descendants of Iron Age nomads of the Eurasian Steppe. *Nature Communications* 2017;14:615.
24. Miller V. Epigraphic traces of Iran in southern Russia. *Journal of the Ministry of Education of the Ministry of Education* 1886.(9) [Miller V. Epigraficheskie sledy iranstva na iuge Rossii. *Zhurnal Mini-sterstva Narodnogo Prosveshcheniia* 1886.(9) (In Russian)].
25. Abaev VI. Fundamentals of Iranian language knowledge. Ancient Iranian languages. Moscow: Science, 1979. 279 c. [Abaev VI. Osnovy iranskogo iazykoznaniiia Drevneiranskie iazyki M.: Nauka 1979.279. (In Russian)].
26. Abaev VI. Fundamentals of Iranian language knowledge. Ancient Iranian languages. M.: Science, 1999.107 [Ishaev MI. Skifskii iazyk Iazyki mira Iranskie iazyki III Vostochnoiranskie iazyki M.: Indrik 1999.107 (In Russian)].
27. Tokhtasiev SR. The problem of the Scythian language in modern science Ethnic Contacts and Cultural Exchanges North and West of the Black Sea from the Greek Colonization to the Ottoman Conquest. *Edited by Victor Cojocaru Iași Trinitas* 2005;59-108 [Tokhtasiev SR. Problema skifskogo iazyka v sovremennoi nauke. Ethnic Contacts and Cultural Exchanges North and West of the Black Sea from the Greek Colonization to the Ottoman Conquest. *Edited by Victor Cojocaru Iași Trinitas* 2005;59-108 (In Russian)].
28. Edelman CI. Some problems of comparative historical Iranian linguistics. Questions of linguistic kinship. Moscow, 2009;(1):81-94 [Edelman DI. Nekotorye problemy sravnitelno-istoricheskogo iranskogo iazykoznaniiia. *Voprosy iazykovogo rodstva*. M. 2009;(1):81-94 (In Russian)].
29. Ivanchik AI. To the question of the Scythian language. *Bulletin of ancient history*. 2009;(2):62-88 [Ivanchik AI. K voprosu o skifskom iazyke. *Vestnik drevnei istorii* 2009;(2):62-88 (In Russian)].
30. Kullanda S. Scythians: language and ethnos. Bulletin of the Russian State Humanitarian University 2015;(2):63 [Kullanda S Skify iazyk i etnos *Vestnik RGGU* 2015;(2):63 (In Russian)].
31. Afanasyev GE, Wen W, Thun S, et al. Kharaz Confederates in the Don basin. Natural science research methods and the paradigm of modern archeology. 2015;146-153 [Afanasev GE, Ven SH, Tun S, i dr. Khazarskie konfederaty v basseine Dona Estestvennonauchnye metody issledovaniia i paradigma sovremennoi arkheologii 2015;146-153 (In Russian)].
32. Rootsi S, Behar DM, Järve M, et al. Phylogenetic applications of whole Y-chromosome sequences and the Near Eastern origin of Ashkenazi Levites. *Nat Commun* 2013;(4):2928.

Author credentials

Murat A. Jaubermezov, Junior Researcher, Laboratory of Human Molecular Genetics, Institute of Biochemistry and Genetics, Ufa Scientific Center, RAS, Assistant, Department of Genetics and Fundamental Medicine, Bashkir State University, e-mail: murat-kbr@mail.ru.

Natalia V. Ekomasova, Ph.D. in Biology, researcher at the Laboratory of Molecular Human Genetics, Institute of Biochemistry and Genetics, UFIC RAS, Associate Professor, Department of Genetics and Fundamental Medicine, Bashkir State University, e-mail: trofimova_nata_@mail.ru.

Sergey S. Litvinov, Candidate of Biological Sciences, Senior Researcher, Laboratory of Human Molecular Genetics, Institute of Biochemistry and Genetics, URC RAS, e-mail: seregtg@gmail.com.

Liliya R. Gabidullina, Graduate student of the Department of Genetics and Fundamental Medicine, Bashkir State University, e-mail: liliya.gab@gmail.com.

Elsa K. Khusnutdinova, Doctor of Biological Sciences, acting Director of the Institute of Biochemistry and Genetics, UFIC RAS, Head of the Department of Genetics and Fundamental Medicine, Bashkir State University, Corresponding Member of the Russian Academy of Education, e-mail: ekkh@anrd.ru.