

Molecular Anthropology

Comparative Y-Chromosome Research in East Georgia Population

Ramaz Shengelia*, **Giorgi Andriadze****, **Liana Bitadze§**,
David Chitanava§, **Nino Chikovani****, **Eka Khmaladze§§**,
Merab Kekelidze§§, **Shorena Laliashvili§**

* *Department for the History of Medicine and Bioethics, Tbilisi State Medical University, Tbilisi, Georgia*

** *St. Andrew The First-Called Georgian University of the Patriarchate of Georgia, Tbilisi, Georgia*

§ *Department of Anthropology, Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia*

§§ *National Center for Disease Control and Public Health, Tbilisi, Georgia*

(Presented by Academy Member Giorgi Kvesitadze)

ABSTRACT. The Georgians occupy the territory in the center of the Caucasus and adjoin many Caucasian nations and ethnic groups. In the north Georgia borders on Karachai, Circassians, Kabardians, Balkars, Ossetians, Chechens, Ingushs and Dagestani, in the East and South – Armenians, Azerbaijani and Turks. Besides, the Georgian population is unique regarding its ethnographical diversity, anthropogenic composition and multi-ethnicity. Within the framework of the project we studied 136 males in different ethnographic groups of the Georgian population: Tushs, the Fereydanian Georgians, other Georgians without taking into consideration their ethnographic groups, also Azerbaijani, Armenians and the so called Ottoman Greeks living in Georgia. Each patrimonial haplotype was distinguished by its 17-STR profile. Y-chromosome lines were determined as unique combinations of SNP and STR data presented in samples. The paper considers the density of the Y-chromosome haplo-groups according to the data in historical sources, their settlement time and areas in Georgia. Despite the fewness of the studied individuals there are some noteworthy results: genetic proximity of the five populations is caused by different factors – existence of a single genetic group with a possible common language and culture on the territory of the Caucasus-Anatolia in the historical past; genetic role of the Albanian population in the formation of the gene pool of the population in the East Caucasus; common migration and assimilation processes; high percentage of L haplo-group in the Black Sea coast Greeks strengthens the hypothesis on genetic proximity of the Laz ethnoses living on the southern Black Sea coastline and the Greeks deported to Georgia, i.e. the possibility of Hellenization of some part of the Laz population at certain historical stages (or permanently). The Caucasus is one of the most significant regions regarding the reconstruction of the ethnic history and ethnogeny of ancient and modern nations of Europe and Southwest Asia. Therefore, purposeful, scrupulous, complex anthro-genetic and phylo-genetic researches containing much innovation are the matter of the near future. © 2017 Bull. Georg. Natl. Acad. Sci.

Key words: population genetics, Georgia, Caucasian populations

Comprehensive and unbiased genetic studies of Georgian people specifically and entire population of the Caucasus, generally including individual tribes and population of various regions associated with them, is necessary for correct understanding of the world history and accurate reconstruction of its ethnogenesis.

Currently analysis and comparison of various genetic data obtained in Georgia (Caucasus) provides the following picture: According to **mtDNA** the picture is quite diverse and, even though a moderate domination of U, HV, K and some of the other haplogroups sub-clades is clearly observable [1-3]. It also is obvious, that according to **Y-DNA** analysis there are basically two haplogroups dominating in Georgia – G2a and J2. There also are relatively rare though rather interesting E1b1b (or E3b) clade and almost all those haplogroups, which are characteristic of Mediterranean world (such as L, I, T and others). Relatively scarce is the R group [3]. G2a group is concentrated on the certain territories of the Caucasus, Anatolia and South Europe, and this group is recognized to be a genetic marker of people with agricultural culture [4-6]. This genetic flow is believed to travel from Levant, Near East, achieving Europe through passing Caucasus in Neolithic period [7-9]. The specialists [10] believe, that haplogroup G was entirely formed admittedly on the united territory of East Anatolia, Lesser Caucasus, and West Iran. The haplogroup J2 was also found within this region. This is the indicator of genographic development in this region [11,12]. It should be noticed that once very popular theory of *in situ* formation of this haplogroup in the Caucasus, North Anatolia, and in general the assumption, that the peoples of Levant-Anatolia-Caucasus (entire Mediterranean basin) are of the united genetic origin, nowadays is not in the focus of the scientists. We should remember, that from anthropological aspects the entire Caucasus has more common and uniting features than differences [13-15]. There are also interesting studies of genetic markers of the Caucasus, as well as of the language and

environmental links [16]. The projection of alleged glotto-chronological paradigms to phylogenetic trees and readily understandable graphical likeness is another obvious tendency, which are gaining the ground in research literature and is mainly pushed by the former Soviet Union scientific groups, who often overestimate its importance [17,18].

Materials and Methods. The first fully independent population and phylogenetic study in Georgia was launched last year and completed this year. Small groups of the following representatives of Georgian population were studied: Georgians (including Tushs), Armenians, Azerbaijanis the so-called Urum Greeks, Fereydan Georgians, total of 136 individuals. Since the samples of each group under study were small, we decided to call this study “mosaic”. However the results obtained seemed to be rather interesting from certain standpoint.

Extraction of genomic DNA was performed from oral cavity mucosa smear, which was studied using Short Tandem Repeat (STR) analysis.

Isolation of every paternal haplotype was performed using 17-STR profile. Y-chromosome lines were defined as unique combinations of SNP and STR data from the obtained samples. DYS389b was calculated by subtraction of DYS389I from DYS389II, which is routinely used in all statistical and nets analyses.

To identify the similarities and differences among various populations, the Y-chromosome data were analyzed using statistical and phylogenetic methods. Calculation of genetic diversity, haplotype distribution, population structure (AMOVA) and balance indices were performed using ARLEQUIN 3.5.1.3 software package [19]. and existing methods

Research of the small groups of Georgian population on fortified some conceptions about genetic spectrum and polymorphism of the population of the Caucasus, and at the same time yielded new interesting data. While acknowledging the significance of population genetics studies within geographical, cultural and confessional environment of the Caucasus (Georgia), a special attention was paid to the

Table 1. Groups studied and the number of individuals under research

Azerbaijanis	Pontiac Greeks	Armenians	Tushs	Fereydan Georgians	Georgians (total)
8	12	10	13	9	78

ethical aspects of the study considering a particularly sensitive attitude towards the results obtained.

Results of the Genetic Study. Total of 4 ethnic groups (Tushs and Fereydan Georgians are both Georgians) was studied using 130 samples obtained from Azerbaijanis, Pontiac Greeks (migrated from the south-east coast of the Black Sea), Armenians, Tushs, living in the east highland of Georgia and Fereydan Georgians, descendants of the migrants from east Georgia to Iran in the 17th century [20, 21], as well as other Georgians irrespective of the places of their origin (Table 1). Genetic material obtained from each person under study was examined for 17 chromosomal STR (short tandem repeat, micro-satellites) matched to relevant marker. Haplogroups of every individual were calculated based on the profiles of micro-satellites. Fig. 1 shows distribution of haplogroups.

Even though the number of the individuals studied in every group is not large, distribution of haplogroups provided several interesting results:

Pontiac Geeks had high incidence of haplogroup L, which previously was observed among the Lazs, this fact being probably indicative of some genetic peculiarities of Pontiac mountains population.

Similar distribution of the haplogroups was also observed in Tushs and Fereydan Georgians, which can be indicative of similarities-and-kinship between the Tushs and External Kakheti population of 17th century. In both groups the haplogroup J2 was pronouncedly dominating.

Two haplogroups – G2a and J2 are dominating in the studied population of Georgians. These groups are characteristic for West and East Caucasus population respectively [17, 16].

Armenians and particularly Azerbaijanis as indicated by the haplogroups of Y chromosome are characterized with significant diversity; both ethnic groups had haplogroups J2, J1, R1b, though Armenians also had haplogroups E1b1b and I2 (which are allegedly of Mediterranean shore origin), whereas Azerbaijanis had haplogroups R1a, T and C indicative of their Iranian

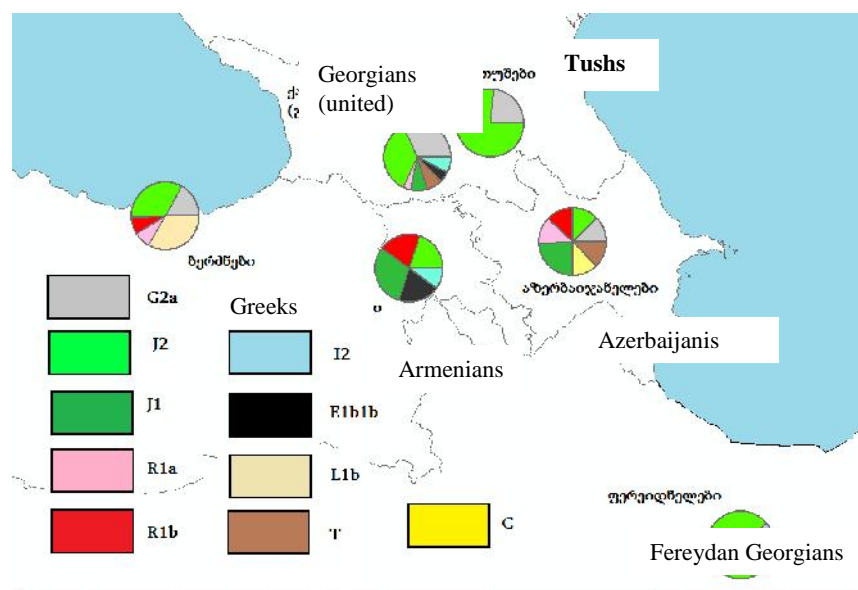


Fig. 1. Haplogroups distribution in the studied groups.

Table 2. Intra- and inter-population changes

The source of variability	Degree of freedom	Squares sum	Variance component	c
Inter-population	4	124.350	1.34457 Va	7.24
Intra-population	47	809.112	17.21514 Vb	92.76
Total	51	933.462	18.55971	
Fixation index FST :	0.07245			

Table 3. FST for every pair of populations

	1	2	3	4	5
1.Azerbaijanis	0.00000				
2.Greeks	0.06135	0.00000			
3.Armenians	-0.02498	0.02156	0.00000		
4Tushs	0.14995	0.10787	0.08832	0.00000	
5.Fereidan Georgians	0.03911	0.11010	0.07893	0.09296	0.00000

and Middle East origin. Below the results of population-and-genetic study based on the analysis of 5 populations are provided. "Unspecified" Georgians are not included into this analysis.

The AMOVA algorithm (software Arlequin 3.5) recommended for research of non-recombinant genes changes were used. Table 2 shows the results obtained.

Table 2 demonstrates that intra-population variability is rather high and exceeds inter-population variability, while validity test (1023 permutation) confirms, that there are statistically significant differences between the population, with $p=0.00391$ (in other words, probability that those differences are coincidental is less than 5%).

Table 3 shows the results of paired F-statistics. Negative values actually are indicative of null statistics.

Table 4 demonstrates significance of those values, or in other words, to what degree of statistical significance they differ from null. From this table it is evident that statistically significant difference between the studied population was observed only in Tushs ($P<0.05$) versus the remaining three populations (Armenians, Greeks, Azerbaijanis), though this is not true for Fereydan Georgians. Difference between Fereydan Georgians and the three other populations is not statistically significant. Only the difference between them and Greeks is statistically significant. This may be probably the result of the small number of individuals studied. However it can be said that there are relatively significant differ-

ences between the Tushs and other four populations; between Fereydan Georgians and other populations (though similarity is relatively high between Fereydan Georgians and Azerbaijanis); between Greeks and Azerbaijanis. In other words similarity more or less reflects geographical location.

Table 5 below shows the indices of genetic variability in all five groups. This Table shows the absence of statistically significant differences between the groups.

The same is true for Theta indices, which represent the ratio of inverted inbreeding. Low Theta (or high inbreeding ratio) was observed in Tushs and Fereydan Georgians, while the lowest values of Theta were observed in Azerbaijanis, whose level of mixture between closely related members is higher compared to other groups. The number of alleles is low in Tushs and Fereydan Georgians.

Of interest are Garza-Williamson index values (Table 6). This index shows the distance between the alleles and alleles numbers ratio. Its low values are indicative of the fact that historically the population passed through "the bottle neck", or in other words, in the history of this population there was a moment when the number of males in it dramatically decreased and subsequently the male part of population increased again.

It is evident, that in Tushs and Fereydan Georgians this index is relatively low, while the highest value of this index was observed in Pontic Greeks. Interrelationship between the groups under study is shown in Fig. 2.

Table 4. Significance of Fst values

	1	2	3	4	5
1.Azerbaijanis	*				
2.Greeks	0.08108	*			
3.Armenians	0.51351	0.28829	*		
4Tushs	0.00901	0.02703	0.02703	*	
5.Fereidan Georgians	0.25225	0.01802	0.13514	0.07207	*

Table 5. Basic characteristics

statistic s.d.	Azerbaijanis	Pontic Greeks	Armenians	Tushsi	Fereydan Georgians	Total
No. of gene copies-2.074	8	12	10	13	9	10.4
No. of loci - 0.0	17	17	17	17	17	17
No. of usable loci- 0.894	15	17	17	17	17	16.6
No. of polym. loci -0.837	15	17	16	16	15	15.8

Table 6. Garza-Williamson's index

Locus	Azerbaijanis	Pontic Greeks	Armenians	Tushs	Fereydan Georgians	Mean	s.d.
1	0.60000	0.40000	0.60000	1.00000	0.60000	0.64000	0.21909
2	1.00000	1.00000	1.00000	0.66667	0.33333	0.80000	0.29814
3	0.75000	0.75000	1.00000	0.75000	1.00000	0.85000	0.13693
4	1.00000	1.00000	1.00000	1.00000	0.75000	0.95000	0.11180
5	0.66667	0.66667	0.66667	0.66667	0.66667	0.66667	0.00000
6	1.00000	1.00000	0.33333	0.66667	1.00000	0.80000	0.29814
7	0.37500	0.50000	0.62500	0.25000	0.25000	0.40000	0.16298
8	0.50000	0.60000	0.70000	0.30000	0.30000	0.48000	0.17889
9	0.60000	1.00000	0.80000	0.40000	0.40000	0.64000	0.26077
10	1.00000	0.75000	0.50000	0.25000	0.50000	0.60000	0.28504
11	0.75000	1.00000	0.50000	0.50000	0.75000	0.70000	0.20917
12	0.50000	0.83333	0.83333	0.50000	0.50000	0.63333	0.18257
13	0.50000	0.66667	0.66667	0.33333	0.16667	0.46667	0.21731
14	0.75000	1.00000	0.75000	0.75000	0.75000	0.80000	0.11180
15	1.00000	1.00000	0.66667	0.66667	1.00000	0.86667	0.18257
16	0.66667	0.66667	0.66667	0.50000	0.50000	0.60000	0.09129
17	0.80000	1.00000	0.60000	0.40000	0.60000	0.68000	0.22804
Mean	0.73284	0.81373	0.70049	0.56471	0.59216	0.68078	0.10259
s.d	0.20872	0.20406	0.18452	0.2422	0.26128	0.21856	0.02973

Conclusions. Study, which was conducted on significantly limited cohort of population, provided the results, which do not contradict the literature data; however it provides a serious basis to critically review some theses and at the same time justifies the necessity to expand the areal of study as well as increase the number of studied individuals.

1. The populations of Tushs and Fereydan Georgians are more homogenous compared to other studied groups. The haplogroup J2 dominates in both

groups, though G2 haplogroup is also encountered. These two haplogroups are dominating among the Georgian population in general. However Georgian population is evidently more diverse genetically. Belonging of the both ethnographic groups to Georgian genetic universe is proved, even though at the same time it is possible to see the trends of their isolated, patrilineal development. These populations also are characterized by low level of heterozygosity, relatively low numbers of alleles and low Garza-

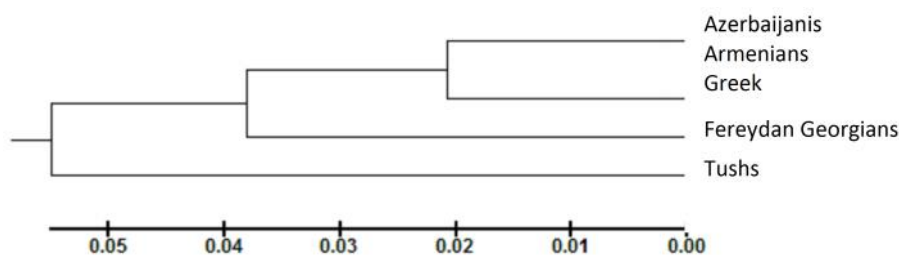


Fig 2. Graphic picture of conducted study.

Williamson's index values, which is indicative of the fact that somewhere in their history they might have passed through "the bottle neck", or in other words both populations suffered dramatic (critical) decrease in the number of males at least once in their history.

2. Armenian, Azerbaijani and Greek populations are more diverse genetically. This fact confirms the above thesis about multiple genetic mixtures occurring in those populations. Viz., Armenians have expressed traces of ancient population of Caucasus-Anatolia (probably even broader geography can be implied) as well as of Indo-European languages speaking ethnic groups, which appeared later (Iranians, Greeks) and to some extent there are traces of semitization (Assyrians, Arabs, Hebrews). Complete assimilation of Albanians should have played its role, which resulted in appearance of relevant haplogroups in general genetic spectrum. Specifically which clade or subclade is that can be determined only through studying Udieli population.

3. Nearly the same applies to Azerbaijani population as well, where the existing "diversity" is created by Turkman, Iranian and Middle Asian influx alongside with Caucasus-Anatolia-Mediterranean shoreline substrate, which possibly might be the result of the of previous assimilation of Albanians.

4. Thus, to our opinion, genetic closeness of all the above five populations should be determined by: a) the existence of population with the united genetic spectrum in the historical past of Caucasus-Anatolia, who possibly had the same language and culture; b) Special genetic contribution of Albanian population to particularly the formation of genofond in East-

ern Caucasus population; c) Existence of common migration processes; d) Intermixing process.

5. High incidence of L haplogroup in Pontic Greeks strengthened the theory about close genetic affinity between the Lazs residing along the Black Sea shore and the Greeks who migrated to Georgia; in other words at a certain historical stage (or even permanently) a certain part of Lazs probably underwent helenisation, which should not be something unheard of considering the well known imperial policy of Byzantine was not something unheard of. Garz-Williamson's index is particularly high in Pontic Greeks, meaning, that in not so far historical past the quantity of Greeks should have been rather high.

6. Tush's population stands somewhat apart from other populations under this study. Here, it seems, migration processes were less intensive, hence the development (through male lines) was somewhat of isolative nature. Compared to them, the remaining four populations more or less differ from each other. Of those four populations one can distinguish Fereydan Georgians who manifested likeness to Azerbaijani, which can be explained by their mutual Albanian roots and intermixing.

7. Male line of Armenians and Azerbaijani actually are indistinguishable, which can be explained by common Iranian, Albanian and Semitic genetic substrate, many years of residing in each others neighborhood and intermixing. Armenians have a strong affinity to Greeks which can be related to the process of total Indo-Europeanization (Hellenisation) of Anatolia occurring during the last 3-4 thousand years, as well as to the existence of common pre-hellenistic roots in Armenian and Greek populations.

The future purpose-oriented and conscientious population and phylogenetic study of the Caucasus (Georgia) seems to bring a lot of novelties, which could answer many questions on the regional on much broader level.

Aknowledgements. The study concerned the materials covering several Georgian ethnographic

groups and national minorities collected by the Center at the Patriarchate of Georgia. Genetic studies were performed at the National Center for Disease Control of Georgia. We acknowledge the contribution of prof. D. Tarkhishvili (Ilia University, Georgia) in the statistical and phylogenetic segments of our study.

მოლეკულური ანთროპოლოგია

აღმოსავლეთ საქართველოს მოსახლეობის შედარებითი Y-ქრომოსომული კვლევა

რ. შენგელია*, გ. ანდრიაძე**, ლ. ბითაძე§, დ. ჭითანაგა§, ნ. ჩიქოვანი**,
ე. ხმალაძე§§, მ. კეკელიძე§§, შ. ლალიაშვილი§§

*თბილისის სახელმწიფო სამედიცინო უნივერსიტეტი, მედიცინის ისტორიისა და ბიოეთიკის განყოფილება, თბილისი, საქართველო

**საქართველოს საპატრიარქოს წმიდა ანდრია პირველწოდებულის ქართული უნივერსიტეტი, თბილისი, საქართველო

§ ივანე ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი, ანთროპოლოგიის განყოფილება, თბილისი, საქართველო

§§ დაავადებათა კონტროლისა და საზოგადოებრივი ჯანმრთელობის ეროვნული ცენტრი, თბილისი, საქართველო

(წარმოდგენილია აკადემიის წევრის გ. კვეციანის მიერ)

საქართველოსა და მთლიანად კავკასიის მოსახლეობის გენეტიკური შესწავლისადმი დაინტერესება მეტად დიდია. საქართველოს მოსახლეობა უნიკალურია თავისი ეთნოგრაფიული მრავალფეროვნებით და გეოგრაფიული განაწილებით. ამ თვალსაზრისით, ის საუკეთესო მოდელი უნდა იყოს ახლო აღმოსავლეთისა და განსაკუთრებით ანატოლიის ეთნიკური სტრუქტურირების გასარკვევად.

წინამდებარე კვლევა ჩატარდა როგორც ზოგადად ქართულ პოპულაციაში, ასევე იზოლირებულად მის გარკვეულ ჯგუფებშიც: თუშებსა და ფერეიდნელებში. აგრეთვე, სომხებში, აზერბაიჯანელებსა და ე.წ. ურმელ ბერძნებში.

სომხების, აზერბაიჯანელების და ბერძნების პოპულაციური ჯგუფები, თუშებსა და ფერეიდნელებთან შედარებით, უფრო მრავალფეროვანია გენეტიკურად, რაც ადასტურებს ისტორიულ მონაცემებს შესაბამის ტერიტორიებზე შესაძლო “ეთნიკური შეხვედრების” შესახებ. თუშებსა და ფერეიდნელებში ჭარბობს ზოგადად ქართული პოპულაციებისთვის დამახასიათებელი J2 და G2a ჰაპლოჯგუფები.

აღნიშნული ხუთივე პოპულაციის მეტ-ნაკლები გენეტიკური სიახლოვე განპირობებული უნდა იყოს: ა) კავკასია-ანატოლიაში, ისტორიულ წარსულში, ერთიანი გენეტიკური სპექტრის მქონე მოსახლეობის ფენის არსებობით, შესაძლო ენობრივი და კულტურული ერთობით; ბ) ალბანური მოსახლეობის გენეტიკური წვლილით განსაკუთრებით აღმოსავლეთ კავკასიაში მცხოვრები მოსახლეობის გენოფონდის ფორმირებაში; გ) საერთო მიგრაციული პროცესების არსებობით; დ) პერმანენტული ურთიერთშერევის პროცესით.

REFERENCES

- Schurr TG., Yardumian A., Shengelia R., Bitadze L., Chitanava D., Laliashvili S., Laliashvili I., Sanders A. et al. (2015) Genetic diversity and population history in Svaneti, Northwest Georgia. *Caucasiological Papers*, 7: 306-328.
- Nasidze I., Ling EY., Quinque D., Dupanloup I., Cordaux R., Rychkov S. et al. (2004) Mitochondrial DNA and Y-chromosome variation in the Caucasus. *Annals of Human Genetics*, 68: 205-21.
- Nasidze I., Salamatin N.M. (1996) Genetic characteristics of the Georgian population. *Gene Geography*, 10: 105-112.
- Bar-Yosef O., Belfer-Cohen A., Adler DS. (2006) The implications of the Middle-Upper Paleolithic chronological boundary in the Caucasus to Eurasian prehistory. *Anthropology*, XLIV: 49-60.
- Bramanti B., Thomas MG., Haak W., Unterlaender M., Jores P. et al. (2009) Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, 326 (5949): 137-40.
- Zvelebil M. & KV. (1988) Agricultural transition and Indo-European dispersals. *Antiquity*, 62: 574-583.
- Algaze G. (1993) The Uruk expansion: The dynamics of expansion of early Mesopotamian civilization. Chicago: University of Chicago Press.
- Adler DS., Bar Oz G., Belfer Cohen A., Bar Yosef O. (2006) Ahead of the game: Middle and Upper Paleolithic hunting behaviors in the Southern Caucasus. *Current Anthropology*, 47:1: 89-118.
- Bruch AA. (2012) Temporal corridors of migration in the Southern Caucasus during [the] Late Early Pleistocene. *Quaternary International*, 279-280: 72-3.
- Rootsi S., Myres NM., Lin AA., Järve M., King RJ., Kutuev I., Cabrera VM., Khusnutdinova EK. et al. (2012) Distinguishing the co-Ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. *European Journal of Human Genetics*, 20, (12): 1275-82.
- Barbujani G., Nasidze I., Whitehead GN. (1994) Genetic diversity in the Caucasus. *Human Biology*, 66 (4): 639-668.
- Houwink Ten Cate PHL. (1995) Ethnic diversity and population movement in Anatolia. In: *Civilizations of the Ancient Near East*. Vol. I. Ed. by Sasson M, Baines J, Beckman G, Rubinson KS. New York: Scribner, pp. 259-270.
- Nasidze I. (2008) What can Genetics tell us about the origins of South and North Ossetians? *Bull. Georg. Acad. Sci.*, 2, 1: 102-108.
- Abdushelishvili MG. (1964), *Antropologia drevnego i sovremennogo naselenii Gruzii*. Tbilisi, Mecniereba (in Russian).
- Bitadze I., Chitanava D., laliashvili Sh., Kvavadze E., Zubiashvili T. (2011) Questions of ethnic entity of the population of Kartli Region and changes of anthropological type from III millenium, B.C. through XX c A.D. pp.850 (in Georgian).
- Tarkhishvili D., Gavashelishvili A., Murtskhvaladze M., Gabelaia M., Tevzadze G. (2014) Human paternal lineages, languages, and environment in the Caucasus. *Human Biology*, 86 (2): 113-30.
- Balanovsky O., Dibirova K., Dybo A., Mudrak O., Frolova S. et al (2011) Parallel evolution of genes and languages in the Caucasus region. *Molecular Biological Evolution*, 28 (10): 2905-2920.
- Bulayeva K.B., Jorde LB., Ostler C., Watkins S. (2003) Genetics and population history of Caucasus populations. *Human Biology*, 75(6): 837-853.
- Excoffier L., Lisher Y.E.L (2010) Arlequin Suitever 3.5: A new series of programs to perform population genetics analysis Linux and under Windows. *Molecular Ecology Resources*, 10: 564-567.
- Chelidze A. (2011) Georgians from Fereydani, pp.99,138. Tbilisi (in Georgian).
- Ethnicities in Georgia (2008) 129,132. Tbilisi (in Georgian).

Received September, 2017