



Biolog. Journal of Armenia, 1 (67), 2015

Y CHROMOSOME DIVERSITY IN THE ARMENIAN POPULATION OF KARABAKH

A.S. POGHOSYAN, H.H. HOVHANNISYAN, A.A. HOVHANNISYAN,
Z.A. KHACHATRYAN, L.M. YEPISKOPOSYAN

*Laboratory of Ethnogenomics, Institute of Molecular Biology, NAS RA
grant.hovhannisyanyan@gmail.com*

Armenians from the Karabakh region are an indigenous ethnic group of the Armenian Highland with strong and distinct national identity. Today, there is only one study on the genetic structure of this geographic group of Armenians. Here, we have analyzed the high resolution Y-chromosomal DNA typing data of 105 unrelated males representing the region of Karabakh and compared it with those of other neighboring Armenian and non-Armenian populations. The results indicate that patrilineal gene pool of Karabakh population predominantly composed of haplogroups of Neolithic origin, while Central Asian lineages are almost absent from its genetic structure. Besides, population of Karabakh demonstrates close genetic affinity to other Eastern Armenian groups.

Karabakh – Y chromosome – population genetics

Ղարաբաղի հայերը Հայկական բարձրավանդակի բնիկ էթնոխումբ, որին բնորոշ է արտահայտված ազգային ինքնագիտակցությունը: Մինչ այժմ կատարվել է տվյալ աշխարհագրական խմբի պոպուլյացիոն գենետիկական կառուցվածքի ընդամենը մեկ ուսումնասիրություն: Ներկա աշխատանքում հետազոտվել են Ղարաբաղում բնակվող արական սեռի 105 ներկայացուցիչների Y քրոմոսոմի գենոտիպավորման տվյալները՝ համեմատելով դրանք հարակից այլ հայկական և ոչ հայկական պոպուլյացիաների համապատասխան տվյալների հետ: Բացահայտվել է, որ Ղարաբաղի հայրագծային գենոֆոնդում առավելապես գերակշռում են նեոլիթյան տոհմագծերը, մինչդեռ կենտրոնասիական գենետիկական հետքեր գրեթե չեն հայտնաբերվում: Բացի այդ, Ղարաբաղի հայերը սերտ գենետիկական ցեղակցություն են ցուցաբերում արևելյան Հայաստանի այլ խմբերի հետ:

Ղարաբաղ – Y քրոմոսոմ – պոպուլյացիոն գենետիկա

Армяне Карабаха являются коренной этногруппой Армянского Нагорья с выраженной национальной идентичностью. До настоящего времени было проведено лишь одно исследование популяционно-генетической структуры данной географической группы армян. Мы проанализировали результаты высокоразрешающего генотипирования Y хромосомы у 105 не родственных по отцовской линии мужчин, жителей Карабаха, и сравнили их с соответствующими данными других армянских и неармянских популяций. Полученные результаты выявили, что большая часть патрилинейного генофонда популяции Карабаха состоит из гаплогрупп неолитического происхождения, в то время как центральноазиатские родословные практически отсутствуют из генетической структуры данной группы. Кроме того, карабахская популяция имеет высокое генетическое сходство с другими группами Восточной Армении.

Карабах – Y хромосома – популяционная генетика

The origins of the population from the eastern regions of Historical Armenia, including Karabakh, is an actively debated issue which is difficult to unearth, relying heavily on broadly applied methods of history, linguistics, archaeology, etc. Unfortunately, due to numerous invasions, earthquakes, and other disturbances that took place on the territory of the Armenian Highland, the bulk of historically valuable evidence has been lost, destroyed, or expunged. Further, as a result of the geographic division of Historical Armenia between its neighboring countries, the archival and archeological studies have become problematic to implement. On the other hand, the evaluation of events that took place in prehistoric times must be done on the basis of information that virtually did not change during the millennia of upheaval. Only genetic data meets this stringent requirement, and the subsequent methods of genetic anthropology that have developed in the last few decades have been widely used to reconstruct the demographic history of geographically and ethnically diverse human populations [12].

Along with the historical destiny that Armenia acquired during long centuries, the Karabakh region also had its specific background and peculiarities. Since the beginning of the 5th century CE, due to the well fortified and mountainous geography, the area was periodically isolated from its historical motherland and environment, without losing its deep links with Armenian cultural heritage and strong ethnic identity [1, 10, 2]. In this regard, the genetic study of the Karabakh population might help elucidate the origins of this Armenian geographic group. So far, the genetic structure of this population is still poorly investigated. Moreover, the only genetic study of this group is based on a low resolution Y-chromosomal genotyping data [14]. To address the question of the origins of the population of Karabakh, we have analyzed data on high resolution Y-chromosomal SNP (single nucleotide polymorphisms) markers and compared the output with corresponding results of other Armenian regional groups and neighboring populations.

Materials and methods. Here, we have used Y-chromosomal data of nine geographic Armenian groups: Central Armenia (CA, n=200), West Armenia (WA, n=148), Van region (n=103), Sasun (n=104), Salmast (n=199), Ararat Valley (ARV, n=110), Gardman (GRD, n=96), Karabakh (n=105) and Syunik (n=105). All samples were typed for 32 Y-chromosomal SNP markers which defined 24 haplogroups (Table 1). The Armenian samples were compared with those of other ethnic groups – Lebanese (n=845) [16], Syrians (n=189) [17], Palestinians (n=281) [17], Turkish (n=467) [5], Ossetians (n=227) [10], Chechens (n=98) [4], Abkhazians (n=57) [10], Georgians (n=65) [15], Jews-Israelites (n=670) [9], Iranians (n=270) [8].

Fst genetic distance, gene diversity index (h) and exact test of population differentiation were computed by Arlequin v. 3.5 package [6]. Principal Coordinate Analysis (PCA) was performed using Genstat software [7].

Results and Discussion. The analysis of haplogroup distributions (Table 1) demonstrates that the patrilineal gene pool of the population from Karabakh is described by 15 SNP haplogroups, the majority of which (ca. 89%) are presented by the haplogroups of Neolithic origin. The modal haplogroup for this population is the lineage R1b1a2-M269 which came to the region of the Armenian Highland from the Near East [3, 11] and is encountered at frequency of 39.0%. This haplogroup is a modal one for all Armenian groups except Sasun. It is worth mentioning that haplogroups of Central Asian origin are virtually absent in the genetic pool of all Armenian populations (tab. 1).

Gene diversity index (h) calculated from haplogroup frequency data reveals that the populations of Karabakh and Syunik have the lowest rate of this parameter (0.8137 and 0.8139, respectively), which are significantly lower ($p < 0.05$) than those of Central and Western Armenian populations. The decreased genetic diversity can be explained by the geographic isolation of these mountainous groups.

Table 1. The frequency distribution of Y-chromosomal haplogroups in geographically different Armenian populations

Haplogroups	Populations								
	WA (n=148)	CA (n=200)	ARV (n=110)	GRD (n=96)	Karabakh (n=105)	Syunik (n=105)	Van (n=103)	Sasun (n=104)	Salmast (n=199)
E1b1b1a1-M78	0.027	0.020	-	-	0.029	0.010	0.039	-	0.070
E1b1b1b2a-M123	0.014	0.025	0.055	0.042	0.029	0.019	0.039	0.029	0.040
E1b1b1-M35	-	-	-	-	-	-	-	-	0.010
F-M89	-	0.005	-	-	-	-	-	-	-
G2a-P15	0.115	0.145	0.091	0.052	0.086	0.057	0.078	0.125	0.055
G-M201	0.020	0.015	0.018	0.010	0.019	0.010	-	-	0.020
I-M170	0.041	0.035	0.027	0.083	0.019	0.019	0.019	-	0.015
J-M304	-	-	-	-	-	-	-	-	0.010
J1-M267	0.061	0.080	0.082	0.021	0.114	0.057	0.078	0.067	0.070
J1a2-L136	0.061	0.090	0.064	0.021	0.038	0.124	0.058	0.029	0.126
J2a1b1-M92	-	-	0.091	0.073	-	-	0.010	-	0.025
J2a1b-M67	0.068	0.070	0.036	0.073	0.086	0.057	0.068	0.077	0.050
J2a1-L27	0.007	-	-	-	-	-	-	-	0.121
J2a-M410	0.196	0.145	0.100	0.146	0.086	0.181	0.194	0.096	0.005
J2b-M12	0.007	0.040	0.009	0.031	0.019	-	0.019	-	0.010
L-M20	0.020	0.005	0.009	0.010	0.019	-	0.019	0.038	0.020
N-M231	-	-	-	0.010	-	-	-	-	-
O-M175	0.007	-	-	-	-	-	-	-	-
Q-M242	-	-	-	-	0.019	-	0.010	-	0.015
R1-M173	-	0.005	-	-	0.010	-	-	-	0.005
R1a1a-M198	0.054	0.030	0.009	0.052	-	0.038	-	0.010	0.055
R1b1a2-M269	0.223	0.215	0.373	0.313	0.390	0.362	0.320	0.154	0.236
R2a-M124	0.034	0.015	-	-	-	0.019	0.010	0.173	0.015
T-M184	0.047	0.060	0.036	0.063	0.038	0.048	0.039	0.202	0.025

In order to depict the pattern of genetic affinity between Armenians and comparative datasets in a two-dimensional scale, we have conducted principal coordinates analysis based on F_{ST} genetic distance values. The PCA plot in fig. 1 places the Karabakh population in a distinct cluster mainly represented by the geographic groups of Eastern Armenia (Figure 1).

For deeper insight into the genetic relationship between the members of this cluster, we have performed Fisher's exact test for population differentiation. The results reveal that the population from Karabakh doesn't significantly differ ($p > 0.05$) from those of Syunik, ARV and Van, with only the GRD population being essentially different ($p < 0.05$). This pattern of differentiation points to the genetic homogeneity of Armenians living in the eastern region of Historical Armenia.

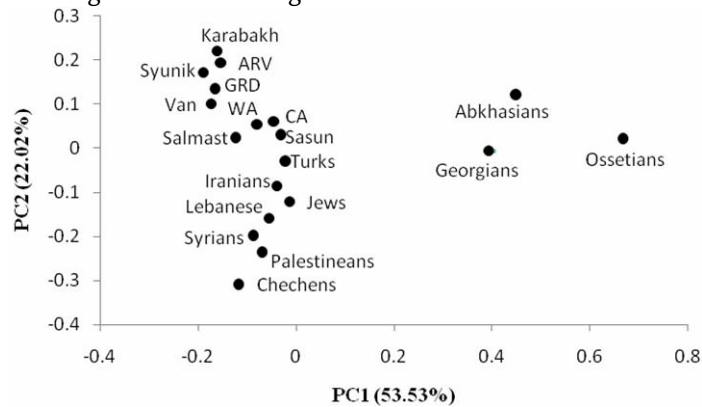


Figure 1. Principal coordinates plot (first two axes) based on F_{ST} values calculated from haplogroup frequency data. Numbers in brackets show the percentage of total variation explained by each axis.

In summary, we came to the following conclusions regarding the patrilineal genetic structure of the Karabakh population: it is mainly composed of Neolithic lineages, it demonstrates a high level of genetic affinity with other Eastern Armenian geographic groups, and no appreciable central Asian genetic influence is detected.

REFERENCES

1. Աճառյան Հր. Հայ բարբառագիտություն, Մոսկվա-Նոր Նախիջևան, էջ 61, 1911:
2. *Асрян Ю.* Урартские клинообразные надписи, М., с. 310, 446, 1960.
3. *Balaresque, P., et al.* A predominantly Neolithic origin for European paternal lineages. *PLoS biology*, *8*, 1, e1000285, 2010.
4. *Balanovsky, O., et al.* Parallel evolution of genes and languages in the Caucasus region. *Molecular biology and evolution*, *28*, 10, 2905-2920, 2011.
5. *Cinnioğlu, C., et al.* Excavating Y-chromosome haplotype strata in Anatolia. *Human genetics*, *114*, 2, 127-148, 2004.
6. *Excoffier L., et al.* Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online*, *1*, 47-50, 2005.
7. Genstat 5 Committee. (1993). Genstat 5 release 3 reference manual. Oxford University Press, Inc.
8. Haber, M., et al. Influences of history, geography, and religion on genetic structure: the Maronites in Lebanon. *European Journal of Human Genetics*, *19*, 3, 334-340, 2010.
9. *Hammer, M. F., et al.* Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood. *Human genetics*, *126*, 5, 707-717, 2009.
10. *Hewsen, R.H., & Salvatico, C.C.* Armenia: A historical atlas. University of Chicago Press, 2001.
11. *Herrera, K. J., et al.* Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. *European Journal of Human Genetics*, *20*, 3, 313-320, 2011.
12. *Jobling, M.A., & Tyler-Smith, C.* The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, *4*, 8, 598-612, 2003.
13. *Underhill, P. A., & Kivisild, T.* Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu. Rev. Genet.*, *41*, 539-564, 2007.
14. *Weale, M. E., et al.* Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group. *Human genetics*, *109*, 6, 659-674, 2001.
15. *Yunusbayev, B., et al.* The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Molecular biology and evolution*, *29*, 1, 359-365, 2012.
16. *Zalloua, P.A., et al.* Y-chromosomal diversity in Lebanon is structured by recent historical events. *The American Journal of Human Genetics*, *82*, 4, 873-882, 2008.
17. *Zalloua, P.A., et al.* Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *The American Journal of Human Genetics*, *83*, 5, 633-642, 2008.

Received on 08.10.2014