



BRILL

Iran and the Caucasus 15 (2011) 00-00



The Location of Azaris on the Patrilineal Genetic Landscape of the Middle East (A Preliminary Report)

**Levon Yepiskoposian, Ashot Margarian,
Laris Andonian, Vahid Rashidvash**

*Institute of Molecular Biology, National Academy of Sciences, Yerevan
Tehran University of Medical Sciences
Yerevan State University*

Abstract

The origin of the Turkic-speaking population of the north-western provinces of Iran, the so-called Azaris, is the subject of long-year debate. Here, we present preliminary results on testing of several hypotheses concerning their origin: 1) the Azaris are the descendants of the Turkic ethnic groups migrated from Central Asia; 2) they have an autochthonous origins; 3) they are of Iranian origin; and 4) they have mixed ethnic origin with unknown proportions of source populations' contribution. The results show that Azaris have much weaker genetic affinity with the populations from Central Asia and the Caucasus than with their immediate geographic neighbours. Relying on these outcomes one can suggest that language replacement (change) with regard to Azaris occurred through "elite dominance" mechanism rather than "demic diffusion" model.

Keywords

Azaris, Genetic Origin, Language Change (Replacement)

The Iranians display considerable genetic diversity consistent with patterns observed in populations of the Middle East overall, reinforcing the notion of Persia as a vital crossroad for human disseminations (Regueiro et al. 2006; Rafiee et al. 2009). The geographic area is remarkable for its high level of ethnic and linguistic diversity, comprising the major language families (Indo-European, Altaic, and Afro-Asiatic) currently spoken by more than seventy ethnically different populations. This demonstrates the role Iran played in population dispersal across the latitudinal belt spanning from Western Anatolia to the Indus Valley. However,

there have been gaps in high-resolution genetic analyses for this region to uncover population history at a fine scale, for example, for particular ethnic and linguistic groups. Among them, the Turkic-speaking Iranian Azari population constitutes the largest ethno-linguistic group after Persian-speakers, accounting for 24 percent of Iran's population (CIA 2010).

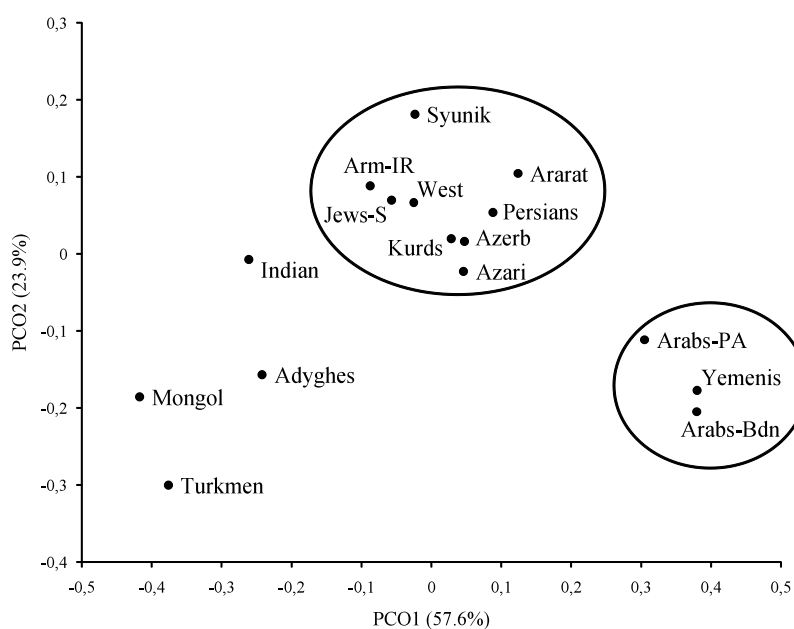
Some sources tend to designate them as the descendants of the Turkic ethnic groups who invaded the area from 11th century onward during several military campaigns (Golden 1992). Still, the majority of sources view the Azaris as having mixed ethnic origin going back to indigenous Caucasian populations and Iranians, who were influenced by Turkic languages throughout centuries of occupation under Turkic dynasties (Nasidze *et al.* 2004; Garthwaite 2005). The Iranian roots of Azaris, according to protagonists of this view, can likely be traced back to ancient Iranian tribes, such as the Medes, who lived on the territory of Azerbaijan (ancient *Āturpatakān*, roughly covering the modern Iranian provinces of Eastern Azerbaijan, Western Azerbaijan, and Ardabil), and Scythian invaders, who arrived during the 8th century B.C. (Zadok 2002). Some information is also available about the Caucasian Albanians as the main ethnic constituent of the inhabitants of the modern-day Republic of Azerbaijan in the South Caucasus, called Azerbaijanis by the name of the state (Walker 1991).

The principal aim of this paper is to test the existing principal hypotheses on the Azaris' origin through identification of their place on the genetic landscape of the Middle East. We used paternally transmitted genetic markers located on Y-chromosome, which enable the reconstruction of patrilineal genetic history of human populations.

As comparative datasets we used several ethnically distinct groups currently living in the Middle East, the Caucasus, Central Asia, and the Indian subcontinent. They consist of 16 populations representing Indo-European, Afro-Asiatic, North Caucasian, and Altaic language families (the list of ethnic groups is presented in the legends of the Figure below). The Turkmens (representing a Turkic-speaking group from Central Asia) and Adyghes (being Caucasians by origin) are considered here as possible source populations for Azaris genetic origin.

DNA samples were collected in 99 male Azaris living in Tabriz (north-west Iran) with strong local identity, whose paternal grandfathers lived in the same region at least three generations. All samples were screened for 18 molecular genetic markers broadly used in population genetics studies.

On the plot of genetic distances (see the Figure below) we can clearly identify several distinct clusters of the populations, which are in good agreement with geographical and, in some cases, linguistic classification of the groups. Armenians, Turks, Kurds, Jews, and Persians and form a distinct sub-cluster of the Middle East, while Arabs (Palestinian Arabs, Yemenis, and Bedouin Arabs) shape another separate entity. Other ethnic datasets, including hypothesised source populations of Turkmen and Adyghes, are located far beyond the limits of the two sub-clusters of the Middle Eastern origin.



SPATIAL RELATIONSHIPS BETWEEN THE GROUPS BASED ON GENETIC DISTANCES

Abbreviations used: *Arabs-Bdn*: Bedouin Arabs; *Arabs-PA*: Palestinian Arabs; *Ararat*: Armenians living in Ararat Valley; *Arm-IR*: Iranian Armenians; *Azerb*: inhabitants of the Republic of Azerbaijan; *Azari*: population of north-west Iran; *Jews-S*: Sephardic Jews; *West*: Western Armenians; *Persians*: Persian-speakers of Iran.

To further test the hypothesis of Central Asian and Caucasian origin of Azaris we used special software designed for admixture analysis, Admix v2.0, which enables the evaluation of the relative genetic contribution of any source population in the origin of the given group (http://web.unife.it/progetti/genetica/Isabelle/admix2_0.html). The results un-

ambiguously support the pattern presented on the Figure. The genetic contribution of Turkmens and Adyghes is much less (4-5 times) while compared with the contribution of Persian-speakers (living, for instance, in Fars).

Concluding, the results obtained indicate that genetic distance between the populations considered depends in the first instance on the geographic proximity than on the common linguistic background. In this context, the results were highly expected taking into account the actual geographic location of the populations. The Azaris, being situated between the Republic of Azerbaijan and Eastern Turkey, had more possibilities of genetic contacts with the closest neighbours, while gene flow between the populations of the same Republic and Eastern Turkey could have been rather limited.

Relying on the results, we can make rather strong inferences about the genetic relatedness between the populations under consideration. The principal one is that the Azaris have much weaker genetic affinity with Turkmens and Adyghes than with their immediate neighbours—Persians, Kurds, Azerbaijanis, and Armenians. We have all grounds to suggest that language replacement might have taken place through “elite dominance” phenomenon rather than “demic diffusion” model (Renfrew 2002). In addition, the modal genetic variant of Turkmens (most frequently encountered in the given population), which could be considered as a signature of this group is virtually absent in Azaris, thus confirming our findings about the mode of the language change. Here, as a statement, the following observation by Igor M. Diakonoff (based on the studies of L.V. Oshanin on the stability of certain genes in the different Turkic-speaking populations) is worth quoting. “What actually happens in history,—the author says,—can be illustrated by a linguistic ‘migration’, which occurred in historical times, namely, the spread of the Turkic languages... It is well known that in a biologically stable population the recessive and the dominant genes will be maintained in the same proportion. The first historically attested Turkic tribes were anthropologically Mongoloid and hence had the typical Mongoloid feature of the epicanthic fold of the eyelid, depending upon a dominant gene. If the movement of the Turkic languages from the borders of modern China to those of modern Turkey were actually a mass population movement, then the percentage of the epicanthus in the speakers of Turkish, Azerbaijani, Turkmen, Uzbek, Kazakh and Qyrghyz would be approximately identical. The actual measurements, however, show the retention percentage of epicanthic fold to be, in males, 51% for the Qyrghyz, 22% for the Kazakhs, 11% for the Uzbeks, 6% for the Turkmens,

and below 2% for the Azerbaijanis; and in females, 83% for the Qyrghyz, 53% for the Kazakhs, 18% for the Uzbeks, and 10% for the Turkmens. For Turks, the retention percentage for both sexes tends toward zero. This means, of course, that the Turkic language was adopted each time by a certain local population from the newcomers, and was then passed on to the next population; intermarriage of the local and immigrating population, and inclusion of local tribes in the general nomadic ethnic entity, also played their role. The farther West the wave of migration passed, the less it included individuals who were bearers of the original genes" (Diakonoff 1990: 55).

BIBLIOGRAPHY

- CIA, *Central Intelligence Agency*, (2010), *The World Factbook* (<https://www.cia.gov/library/publications/the-world-factbook>).
- Diakonoff, Igor M. (1990), "Language Contacts in the Caucasus and in the Near East", T. Markey; J. Greppin (eds.), *When Worlds Collide: Indo-Europeans and Pre-Indo-Europeans*, Karoma Publ.: 53-66.
- Garthwaite, G. R. (2005), *The Persians*, Oxford.
- Golden, Peter B. (1992), *An Introduction to the History of the Turkic Peoples: Ethnogenesis and State-formation in Medieval and Early Modern Eurasia and the Middle East*, Otto Harrassowitz.
- Nasidze, I.; Ling, E. Y. S.; Quinque, D.; Dupanloup, I.; Cordaux, R.; Rychkov, S.; Naumova, O.; Zhukova, O.; Sarraf-Zadegan, N.; Naderi, G. A.; Asgary, S.; Sardas, S.; Farhud, D. D.; Sarkisian, T.; Asadov, C.; Kerimov, A.; Stoneking, M. (2004), "Mitochondrial DNA and Y-Chromosome Variation in the Caucasus", *Ann Hum Genet.*, 68: 205-221.
- Rafiee, M.-R.; Sokhansanj, A.; Naghizadeh, M.-A.; Farazmand, A. (2009), "Analysis of Y-Chromosomal Short Tandem Repeat (STR) Polymorphism in an Iranian Sadat Population", *Russian Journal of Genetics*, 45: 969-973.
- Regueiro, M.; Cadenas, A. M.; Gayden, T.; Underhill, P. A.; Herrera, R. J. (2006), "Iran: Tricontinental Nexus for Y-Chromosome Driven Migration", *Human Heredity*, 61: 132-143.
- Renfrew, C. (2002), "The Emerging Synthesis: The Archaeogenetics of Farming/Language Dispersals and Other Spread Zones", P. Bell-

- wood; C. Renfrew (eds.), *Examining the Farming/Language Dispersal Hypothesis*, McDonald Institute for Archaeological Research: 3-16.
- Walker, Christopher J. (1991), *Armenia and Karabagh: The Struggle for Unity*, Minority Rights Group Publications.
- Zadok, R. (2002), *The Ethno-linguistic Character of Northwestern Iran and Kurdistan in the Neo-Assyrian Period*, Tel Aviv: Archaeological Center Publications.