IRANIAN ARMENIANS IN THE ARMENIAN GENETIC LANDSCAPE

L. Andonian1, A. Harutyunyan2,3, A. Margaryan2, Z. Khachatryan3, L. Yepiskoposyan2,3,*

1 Dept. of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Iran,
2 Human Genetics Group, Institute of Molecular Biology NAS RA,
3 Institute of Man, Yerevan, Armenia

* Corresponding author: Levon Yepiskoposyan, Sc.D.; Institute of Molecular Biology NAS RA, 7 Hasratyan Str., 0014, Yerevan, Armenia; Tel.: (+374 93) 352 421; e-mail: lepiskop@yahoo.com

ABSTRACT

The place of modern Armenians living in Iran in the context of general Armenian population was clarified based on Y chromosomal single nucleotide (SNP) and microsatellite (MS) polymorphisms. 89 DNA samples in form of buccal swabs have been collected in ethnic Armenian males who are the direct descendants of Armenians forcibly moved to Iran in the beginning of 17th century AD. Twelve SNP and six MS markers have been typed to describe patrilineal genetic structure of the group studied and to compare with six geographic groups of Armenians representing different areas of Historical Armenia. The results indicate that modern Iranian Armenians are genetically closer to Western Armenians which is in agreement with historical records. Political and geographic isolation during the last four centuries had moderate influence on the genetic structure of the community that is expressed in lower diversity of their patrilineal genetic legacy.

Key words: Iranian Armenians, Y chromosome, single nucleotide polymorphism, microsatellite, genetic isolation

INTRODUCTION

Armenians have a long history of interaction with Persia/Iran since ancient times. The two peoples shared many ethnic, linguistic and cultural features, and even intermarriage among Iranian and Armenian nobility was common [3, 11]. However, Armenia’s adoption of Christianity as a state religion in the beginning of fourth century AD alienated them from the Zoroastrian Persians; and the Persian conversion to Islam after the Arab-Muslim conquests in the seventh century further deepened this alienation.

The period from the eleventh to sixteenth centuries has been crucial for the history of the Armenians, especially in terms of their relocation and settlement in Iran [11]. However, the greatest impact on modern Iran and the Iranian Armenian community was left by Shah Abbas the First in the late sixteenth and early seventeenth centuries [1-3, 7, 19, 20]. Persian leader followed the policy of moving large groups of Armenians to Iran to bring skilled merchants and artisans to capital Isfahan [3, 8, 12].

Between 1604 and 1605 some 250,000 to 300,000 Armenians were removed from their homeland. Thousands died on the roads of migration, while most of the Armenians eventually settled in Iranian Azerbaijan, where other Armenians had established earlier. Some ended up in the Mazandaran region and in the cities of Sultanieh, Qazvin, Mashhad, Hamadan, Arak, and Shiraz [11].

New Julfa (Nor Jugha) became the most important settlement of Armenians in Iran. They lived across the banks of the Zayandeh Rud in a new part of the city, which was then given the same name as their old hometown. Thus, in 1605 New Julfa was established as the Armenian quarter [1-3, 7, 8].

The Armenian community of New Julfa, a suburb of Isfahan, was highly valued by Shah Abbas I and became one of the economic bases of the Safavid state [7, 13, 19, 20]. Armenians were granted trading privileges and a monopoly on the silk trade, which changed the community into a prosperous and influential one [2, 3, 20]. In time, the Armenian population of New Julfa and surrounding Armenian villages grew to around 150,000, thus becoming the majority of the Christian population of Safavid Iran [2, 3]. Presently, the Armenian community in Iran comprises the largest Christian religious group among ethnic minorities in the country.

The aim of this paper is twofold: first, we intended to clarify the place of modern Armenians living in Iran in the frame of different regional groups of their compatriots studied so far; second, to check if virtual isolation during at least four centuries had any notable impact on their population genetic structure.
MATERIAL AND METHODS

Buccal swabs were collected from 90 ethnically Armenian men currently living in Isfahan, Iran. One sample was later discarded as the Y-chromosome typing was unsuccessful.

All donors were selected only if their paternal grandfathers were from the same region and they were unrelated to other donors at the grandfather level. Swabs were stored in a DNA preservative solution consisting of 0.5% sodium dodecyl sulphate and 0.05 M ethylenediaminetetraacetic acid for transport purposes. Samples were collected anonymously and informed consent was obtained from all individuals before samples were taken.

Our sample set was compared with those from several geographic groups of Armenians representing different areas of Historical Armenia [25, 27, 28]. All samples were typed for the same Y chromosomal markers. Standard phenol-chloroform DNA extractions were performed. The strategy adopted for typing samples was designed to ensure informative comparison with existing published data. NRY were characterised by 12 binary Y chromosome polymorphisms: 9R7, M9, M13, M20, S7, SY+465, SY4064, SY10831, Tat, M17, Alu insert – YAP, and p12f2, as described by Weale et al. [25] and Rosser et al. [17] and screened for six microsatellite (MS) markers: DYS189, DYS390, DYS391, DYS392, and DYS393, as described by Thomas et al. [22]. Haplogroups (hg) were defined by single nucleotide polymorphism (SNP) markers according to the Y Chromosome Consortium nomenclature [9]. Microsatellite repeat numbers were assigned according to Kayser et al. [10].

Unbiased genetic diversity index, \( h \), and its standard error were calculated using the formulae of Nei [14]. Nei’s Genetic Identity, \( I \), was calculated in accordance with Nei [14]. Pairwise genetic distances (\( F_{ST} \)) were estimated from analysis of molecular variance (AMOVA) \( \Phi_{ST} \) values with the aid of Arlequin software [5]. Tests for significant population differentiation were carried out using the exact test for population differentiation [14]. Testing for differences in \( h \) between two populations was performed by bootstrapping method [24].

Principal Coordinates Analysis was conducted on similarity matrices calculated as one minus Genetic distance (\( R_{ST} \)) or based on Nei’s Genetic Identity values. Figures along the main diagonal, representing the similarity of each population sample to itself, were calculated from the estimated genetic distance between two copies of the same sample. Signature haplotype analysis (high frequency modal haplotypes and modal clusters [21, 23, 26] was performed by hand. Phylogenetic tree (dendrogram) was constructed from the SNP+MS haplotype frequencies by using the Neighbor-Joining (NJ) method [18] with the standard genetic distances (1-\( I \)) by using PHYLIP software [6].

RESULTS AND DISCUSSION

We compared Armenians from Iran with the following geographic groups of Armenians: Ararat (representing in general Ararat valley in modern Armenia), North (predominantly northern regions of the Republic of Armenia), Syunik (southern, predominantly mountainous, region of Armenia), Artsakh (representing the population of Nagorno-Karabakh Republic), West (the descendants of population from historical Western Armenia, currently in Turkey), Hamshen (the descendants of Hamsheni population from the southern coast of Black Sea, currently living in Russia).

The SNP markers typed on the Y chromosome allowed distinguishing nine haplogroups (Tab. 1). The table shows that Iranian Armenians bear seven haplogroups and all of these are present in other Armenian territorial groups. In other words, Armenians from Iran do not have any unique SNP mutations which might differentiate them from global Armenian population and might witness about possible gene flow from neighboring ethnic groups.

![Table 1. SNP haplogroup frequencies in different Armenian geographic groups.](image)

The most common haplogroup in Iranian Armenians is hg J with 31.46% of frequency. This feature is encountered also in Armenians from Ararat, North and West, where this haplogroup is the most common as well (40.43%, 38.95% and 37.25%, respectively). Despite the fact that hg J has comparable rates in the remaining three geographic groups of Armenians, nevertheless this haplogroup is not the modal one in Hamshen (where the modal haplogroup is BR*(xDE,JR), present in 43.90%), Karabakh or Syunik (having modal haplogroup P*(xR1a), with 43.06% and 40.20% of frequency, respectively).
The second frequently encountered haplogroup is P*(xR1a), being present in 25.84% of Iranian Armenians. This haplogroup has the same position in Ararat (23.40%), North (22.11%) and West (24.25%), while it is much less frequent in Hamshen (8.54%).

The haplogroup BR*(xDE,JR) is the third frequent one (20.22%) in Armenians from Iran and is present almost at the same rate in Ararat, North and West (17.02%, 20.00% and 21.00%, respectively). This haplogroup is markedly less frequent in Karabakh and Syunik (11.57% and 9.12%, respectively).

Other four haplogroups detected in Iranian Armenians are R1a1, E*(xE3a), L and K*(xL, N3, O2b, P), which are encountered at less than 10% rate.

In general, this comparison allows assuming that Iranian Armenians are expected to differ significantly from Hamshen, Syunik and Karabakh according to haplogroup structure and be very close in this respect to Ararat, North and West. To check this hypothesis we applied an exact test for population differentiation [15] that proved the initial suggestion. The same result is obtained while performing principal coordinate analysis (PCO) based on genetic distance values (FST, Tab. 2). The resulting plot is shown on Fig. 1.

**Table 2.** FST values based on SNP haplogroups (lower left) and SNP+MS haplotypes (upper right triangle).

<table>
<thead>
<tr>
<th></th>
<th>Ararat (n=47)</th>
<th>Arm IR (n=89)</th>
<th>Hamshen (n=82)</th>
<th>Karabakh (n=216)</th>
<th>North (n=190)</th>
<th>Syunik (n=296)</th>
<th>West (n=400)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ararat</td>
<td>-</td>
<td>0.0055*</td>
<td>0.0204*</td>
<td>0.0012*</td>
<td>-</td>
<td>0.0058*</td>
<td>0.0006*</td>
</tr>
<tr>
<td>Arm IR</td>
<td>0.0080</td>
<td>-</td>
<td>0.0293*</td>
<td>0.0154*</td>
<td>0.0075*</td>
<td>0.0164*</td>
<td>0.0085*</td>
</tr>
<tr>
<td>Hamshen</td>
<td>0.0550*</td>
<td>0.0450*</td>
<td>0.0293*</td>
<td>-</td>
<td>0.0232*</td>
<td>0.0322*</td>
<td>0.0226*</td>
</tr>
<tr>
<td>Karabakh</td>
<td>0.0293*</td>
<td>0.0223*</td>
<td>0.1334*</td>
<td>-</td>
<td>0.0036*</td>
<td>0.0008*</td>
<td>0.0071*</td>
</tr>
<tr>
<td>North</td>
<td>-0.0100</td>
<td>-0.0011</td>
<td>0.0463*</td>
<td>0.0350*</td>
<td>-</td>
<td>0.0050*</td>
<td>0.0024*</td>
</tr>
<tr>
<td>Syunik</td>
<td>0.0184*</td>
<td>0.0224*</td>
<td>0.1359*</td>
<td>0.0004</td>
<td>0.0272*</td>
<td>-</td>
<td>0.0095*</td>
</tr>
<tr>
<td>West</td>
<td>-0.0048</td>
<td>-0.0003</td>
<td>0.0476*</td>
<td>0.0294*</td>
<td>-0.0022</td>
<td>0.0248*</td>
<td>-</td>
</tr>
</tbody>
</table>

*P<0.05 according to exact test for population differentiation.

The PCO plot visualizes the actual position of Iranian Armenians in the context of various Armenian territorial groups according to patrilineal genetic history. Indeed, Iranian Armenians, Ararat, North and West form rather distinct cluster, apparently different from another compact group including Syunik and Karabakh. Hamshen, is distant from both clusters which is in accordance with previously reported results [27].

Microsatellite markers allow analyzing the population genetic structure of different Armenian geographic groups at much higher level of resolution which provides more valid inferences on their genetic relationships. The six microsatellite markers in addition to SNP haplogroups define 51 compound SNP+microsatellite haplotypes. Homoplasy of microsatellite haplotypes across haplogroups was low (3/51=5.88% of haplotypes, single chromosome in both of the haplogroups involved in each case), consistent with previous reports [4, 23, 25]. 16 SNP+MS haplotypes are unique for Iranian Armenians (or 31.4%).

Genetic diversity values, h, based on SNP+MS haplotypes frequencies are at intermediate rate for Iranian Armenians (actual level is 0.9793, bootstrapped – 0.9684). The lowest level of genetic diversity is encountered in Hamshen which reflects their long centuries of genetic isolation [27]; the highest values are detected in North (Fig. 2). In case of using bootstrap technique, h value in Iranian Armenians significantly differs (P<0.05) only from Hamshen; while applying Bayesian approach, the significant level of difference (P<0.001) exists between Iranian Armenians from one side and Ararat, North, West and Hamshen from the other side (Tab. 3). In other words, Armenians from Iran have intermediate rate of genetic diversity, close to those of Syunik and Karabakh (P<0.05). These results might indicate some degree of reproductive isolation during four centuries of the recent history of Iranian Armenians’ community, although not to the same extent as in Hamshen.

The results of Analysis of Molecular Variation (AMOVA) show that most of the Y chromosome diversity is attributed to within-population differentiation (99.08%), while only 0.92% accounts for inter-population variability. This is in full accordance with previously reported results on Y chromosome diversity in different territorial groups of the same ethnicity [25].
Fig. 2. Genetic diversity, $h$, with bootstrap 95% confidence intervals across Armenian geographic groups.

Table 3. Pairwise differences in $h$ values. P values (two-tailed) for difference in $h$ based on bootstrap standard errors (lower left) and Bayesian approach (upper right triangle). Significant p values are in bold.

<table>
<thead>
<tr>
<th></th>
<th>Ararat (n=47)</th>
<th>Hamshen (n=82)</th>
<th>Artsakh (n=216)</th>
<th>Syunik (n=296)</th>
<th>West (n=400)</th>
<th>Arm IR (n=89)</th>
<th>North (n=190)</th>
<th>$h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ararat</td>
<td></td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0062</td>
<td>0.0000</td>
<td>0.6518</td>
</tr>
<tr>
<td>Hamshen</td>
<td>0.0310</td>
<td></td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.9898</td>
</tr>
<tr>
<td>Artsakh</td>
<td>0.8350</td>
<td>0.0549</td>
<td></td>
<td>0.5560</td>
<td>0.0000</td>
<td>0.2728</td>
<td>0.0000</td>
<td>0.9512</td>
</tr>
<tr>
<td>Syunik</td>
<td>0.9570</td>
<td>0.0818</td>
<td>0.8238</td>
<td></td>
<td>0.0000</td>
<td>0.0740</td>
<td>0.0000</td>
<td>0.9732</td>
</tr>
<tr>
<td>West</td>
<td>0.6706</td>
<td>0.0030</td>
<td>0.0300</td>
<td>0.0041</td>
<td></td>
<td>0.0000</td>
<td>0.0018</td>
<td>0.9866</td>
</tr>
<tr>
<td>Arm IR</td>
<td>0.9148</td>
<td>0.0312</td>
<td>0.6858</td>
<td>0.8332</td>
<td>0.2073</td>
<td></td>
<td>0.0000</td>
<td>0.9793</td>
</tr>
<tr>
<td>North</td>
<td>0.9559</td>
<td>0.0011</td>
<td>0.0044</td>
<td>0.0003</td>
<td>0.6758</td>
<td>0.0605</td>
<td></td>
<td>0.9903</td>
</tr>
</tbody>
</table>

The modal haplotype in Iranian Armenians, ht 15-13-21-10-11-14 (belonging to haplogroup BR*(xDE,JR)), is encountered at 7.87% and is at very low level in other Armenian geographic groups with total absence in Ararat, North and Hamshen (Tab. 4). The most representative in all Armenian groups haplotypes, almost evenly distributed, is ht 14-15-23-10-11-12 (haplogroup J) and is encountered at 3.37% in Iranian Armenians. It is necessary to underline that both Hamsheni and Karabakh and Syunik modal haplotypes are totally absent in Armenians from Iran.

Table 4. Frequently Encountered Haplotypes for seven Armenian regional groups.

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>MS</th>
<th>Karabakh (n=216)</th>
<th>North (n=190)</th>
<th>Syunik (n=296)</th>
<th>Arm IR (n=89)</th>
<th>Hamshen (n=82)</th>
<th>West (n=400)</th>
<th>Ararat (n=47)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P*(XR1a)</td>
<td>14 12 24 11 13 12</td>
<td>0.1111</td>
<td>0.0474</td>
<td>0.1216</td>
<td>-</td>
<td>-</td>
<td>0.0325</td>
<td>0.0426</td>
</tr>
<tr>
<td>BR*(xDE,JR)</td>
<td>15 13 21 10 11 14</td>
<td>0.0093</td>
<td>-</td>
<td>0.0135</td>
<td>0.0787</td>
<td>-</td>
<td>0.0025</td>
<td>-</td>
</tr>
<tr>
<td>BR*(xDE,JR)</td>
<td>16 12 23 10 12 13</td>
<td>-</td>
<td>0.0053</td>
<td>-</td>
<td>0.1585</td>
<td>0.0125</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>J</td>
<td>14 15 22 10 11 12</td>
<td>-</td>
<td>0.0105</td>
<td>0.0101</td>
<td>0.0225</td>
<td>-</td>
<td>0.0525</td>
<td>-</td>
</tr>
<tr>
<td>J</td>
<td>14 15 23 10 11 12</td>
<td>0.0741</td>
<td>0.0421</td>
<td>0.0541</td>
<td>0.0337</td>
<td>0.0732</td>
<td>0.0450</td>
<td>0.0851</td>
</tr>
</tbody>
</table>

It is worthwhile to add that Karabakh and Syunik modal haplotype is one-step neighbor of Atlantic Modal Haplotype [26], which is a signature haplotype of the Palaeolithic peopling of Europe. The Atlantic Modal Cluster is thought to represent a remnant paternal signal of an ancient, possibly pre-Neolithic population that spread from Southeast Asia into Europe. Its absence in Iranian Armenians might indicate that Armenians from Syunik and Karabakh did not leave any detectable signal in their genetic legacy.

The same pattern is observed while analyzing modal clusters, i.e. when adding one-step neighbors to modal haplotype (Tab. 5). The modal cluster in Iranian Armenians (7.87%) is present at low frequencies in all Armenian geographic groups but Hamshen. Modal clusters of Ararat and West are rather frequent (19.10% and 11.24%, respectively), while modal clusters of Hamshen and Syunik/Karabakh are absent or at low rate (0% and 5.62%, respectively). Thus, these results support abovementioned outcomes on predominantly central-western (related to Historical Armenia) genetic roots of Iranian Armenians in contrast with Eastern Armenia, i.e. Syunik and Karabakh. Microsatellite analysis conducted at higher resolution provides similar pattern of genetic relationships between different Armenian populations with some peculiarities. These relate to more detailed position of Iranian Armenians while
regarding its genetic distance from two distinct clusters: Ararat-North-West and Karabakh-Syunik. Genetic distance values ($F_{ST}$ based on SNP+MS) are shown in Table 2. Exact test of population differentiation (based on SNP+MS) reveals significant differences between Iranian Armenians and all other comparative data sets.

**Table 5.** Frequently Encountered Haplotype Clusters for seven Armenian regional groups.

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>MS</th>
<th>Karabakh (n=216)</th>
<th>Syunik (n=296)</th>
<th>Hamshen (n=82)</th>
<th>Ararat (n=47)</th>
<th>North (n=190)</th>
<th>West (n=400)</th>
<th>Arm_IR (n=89)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P^*(xR1a)$</td>
<td>14 12 24 11 13 12</td>
<td><strong>0.2639</strong></td>
<td><strong>0.2568</strong></td>
<td>0.0366</td>
<td>0.1064</td>
<td>0.0842</td>
<td>0.1125</td>
<td>0.0562</td>
</tr>
<tr>
<td>$BR^*(xDE,JR)$</td>
<td>16 12 23 10 12 13</td>
<td>-</td>
<td><strong>0.0034</strong></td>
<td><strong>0.3049</strong></td>
<td>-</td>
<td>0.0053</td>
<td>0.0175</td>
<td>-</td>
</tr>
<tr>
<td>J</td>
<td>14 16 23 10 11 12</td>
<td>0.1065</td>
<td>0.0912</td>
<td>0.0732</td>
<td><strong>0.1277</strong></td>
<td>0.1105</td>
<td>0.1275</td>
<td>0.0449</td>
</tr>
<tr>
<td>J</td>
<td>14 15 23 10 11 12</td>
<td>0.1204</td>
<td>0.1419</td>
<td>0.1098</td>
<td><strong>0.1277</strong></td>
<td><strong>0.1526</strong></td>
<td><strong>0.1800</strong></td>
<td><strong>0.1910</strong></td>
</tr>
</tbody>
</table>

To visualize this pattern we applied two types of data (Fst and Rst) for principal coordinate analysis (Fig. 3). Fig. 3 shows some interesting peculiarities supported by tests of significance. In the first instance, the level of genetic divergence among the Armenian geographic groups is rather high, almost at the level detected across ethno-national boundaries. In general it is possible to visually distinguish two dense clusters: Ararat-North-West and Karabakh-Syunik. Armenians from Iran display closer genetic affinity to the first cluster and are in rather distant position from the second one, thus confirming the results based on Fst, Rst and exact tests.

Another method (neighbor-joining tree) for visualizing the genetic affinity between different geographic groups of Armenians was used to construct the classification tree based on Nei’s Genetic Identity [14] values (Fig. 4). The obtained pattern of classification is in full accordance with the previous results of inter-population comparison conducted. The tree evidently separates two mountainous regions of Eastern Armenia, Syunik and Karabakh, as well as extremely isolated population of Hamshen, along with moderate level of separation of Iranian Armenians. The other geographic groups comprising West, North and Ararat actually reflect Western Armenian genetic legacy.

**CONCLUSION**

The first inference based on the results obtained is that Armenians living in Isfahan are genetically closer to Western Armenians than to the compatriots living in Eastern regions of Historical Armenia (Syunik and Karabakh). It might reflect the actual case that in the beginning of 17th century forcibly moved Armenians were predominantly from the regions which can be attributed to the western genetic province of Historical Armenia. It is worth mentioning that this outcome is supported by historical records. In the same time, we can also suggest that the mountainous and thus difficult to access regions of Syunik and Karabakh could mainly escape the forced resettlement four centuries ago. The results obtained give us an additional hint to suggest a moderate degree of genetic isolation of Iranian Armenians during last four centuries. The plausible reasons of it might be high level of reproductive isolation of Christian
Armenians surrounded by predominantly Muslim neighbors which significantly restricted Armenians’ gene exchange with other populations. It might be of great interest to check if the same pattern would be found while studying the population genetic structure of Iranian Armenians based on mitochondrial DNA and autosomal markers.

REFERENCES


Reviewer: Cor. Member of NAS RA, Prof. R. Aroutiounian