An Exploratory Picture of the Iranian mtDNA Landscape

S. Farjadian[1,2], M. Sazzini[3], S. Tofanelli[4], L. Castrì[3], L. Taglioli[4], D. Pettener[3], A. Ghaderi[1,5], G. Romeo[6], D. Luiselli[3]

[1] Department of Immunology Shiraz University of Medical Sciences, Shiraz, Iran
[2] Allergy Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[3] Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Bologna, Italy.
E-mail: marco.sazzini2@unibo.it
[4] Dipartimento di Biologia, Università di Pisa, Pisa, Italy
[5] Institute for Cancer Research, Shiraz University of Medical Sciences, Shiraz, Iran
[6] Unità Operativa di Genetica Medica, Dipartimento di Scienze Ginecologiche, Ostetriche e Pediatriche, Policlinico Sant'Orsola Malpighi, Università di Bologna, Bologna, Italy

KEY WORDS: mtDNA, human population genetics, Iranian ethnic groups.

Introduction

Iran occupies a pivotal position in the South Western Asian corridor, having long represented a natural hub for the expansion of genes and cultures since *H. sapiens* early colonization of the Eurasian landmass (Kivisild *et al.*, 1999; Macaulay *et al.*, 2005), even though the Neolithic and Metal Ages seem to have left the deepest imprint on its genetic landscape. Proto-Elamo-Dravidian cultures, which spread eastwards from the Zagros Mountains with the movement of farmers to the Indus Valley, were almost completely replaced around 1,300 BC by Aryans, Indo-European nomadic tribes considered plausible ancestors of most Iranians. Medes, Persians, Bactrians, Parthians later populated Iran and subsequent invasions by Macedonians, Arabs, Turks and Mongols, while largely contributing to its ethno-linguistic heterogeneity, affected only marginally its pre-existent gene pool (Quintana-Murci *et al.*, 2004). The overlapping of ancient and more recent events interacting with geographical and cultural barriers has thus yielded a tangled patchwork of anthropological types within this region, although the genetic diversity of few Iranian groups has been investigated so far (Metspalu *et al.*, 2004; Shepard *et al.*, 2005; Regueiro *et al.*, 2006) and the depiction of a fine-grained picture of the overall Iranian mtDNA landscape is still to be drawn.

Materials and Methods

mtDNA variability at both HVS-I and coding regions was surveyed on around seven hundred unrelated individuals belonging to the following Iranian ethnic groups: Balochis (B), Gilakis (G), Jews (J), Kurds (K), Lurs from Khoramabad (L_K), Lurs from Yasouj (L_Y), Mazandaranis (M), Parsees (P), Qashqaees (Qa), people from Qeshm (Qe), Turkmens (T), Zoroastrians (Z). Apportionment of genetic variance among linguistically/geographically-based groups of populations was investigated with AMOVA and SAMOVA. Spatial PCA was carried out to explore whether geographical features structure the observed variability and correspondence analysis focused on the frequencies of the main haplogroups in each population was computed to identify lineages mainly contributing to ethnic groups' differentiation.

Results

Relatively low haplotype diversity (ranging from H=0.910 to H=0.963) and high mean number of pairwise differences (ranging from MPD=4.849 to MPD=7.043) due to deep haplotype genealogies evolved in small populations of constant size were observed for L_Y, B, J, Z, in accordance to smooth unimodal mismatch curves and significant negative Tajima's D values for all groups, except Z (D=-0.84036, p=0.209; r=0.034, p<0.001) and J (D=-0.70590, p=0.283; r=0.080, p<0.001). Most of the observed variation (97.27%, F_{ST}=0.027, p<0.001) was due to within-population differences and when samples were grouped according to language/geography the among-groups component of variance was low and not significant (F_{CT}=-0.0027/F_{CT}=0.001). Sequential increase of population clusters by SAMOVA pointed out a maximization of genetic differentiation among groups with 3 clusters (Z, J, others), with the greater among-groups variance (4.26%) and significant low F_{CT} (0.043; p<0.001), emphasizing a genetic discontinuity among Z, J and the bulk of Iranian samples. No significant spatial autocorrelations (Global_test=0.117; p=0.581 and Local_test=0.149; p=0.391) were found, suggesting that the examined groups are not more similar/dissimilar to neighbors than expected in a random spatial distribution. On a total of 71 haplogroups, 90.8% were West Eurasian lineages (*H, HV, I, J, K, N, T, U, V, W*) among which *H* and *U* were predominant (16.4%), with *H* being the most frequent in all groups, except B, J, Z, L_Y,

and *U7* reaching 34.5% in L_Y. Most groups tended to cluster due to this common West Eurasian background. East Eurasian lineages (*A, C, D, F, G, Z*) (2.8%) instead cumulated in T (10.9%) and Qe (10.3%), South Asians (*M*) (2.5%) in B (11.3%), Sub-Saharan Africans (*L1, L2, L3*) (2.2%) in Qe (12.7%) and Ab (11.1%). In particular, *A* and *L** clades differentiated Qe, *W* (24.2%) and *M* clades (11.3%) differentiated B, *R0* (17.6%) and *T* (35.3%) lineages characterized Z and *T* (32.7%) and *J2b* (12.73%) characterized J.

Discussion

High ethno-linguistic and low mtDNA heterogeneity are observed, reflecting the role of crossroad for human migrations played by Iran since prehistoric times, being geographic/linguistic barriers insufficient to generate a significant structure in the bulk of its mtDNA variability. Exceptions are due to limitations to exogamous marriages after the spread of Islam for members of religious minorities (Z, J), which experienced long-term isolation, endogamy, drift and/or founder effects leading to genetic discontinuity with respect to the other groups. Their autochthonous pre-Indo-Iranian genetic backgrounds survived relatively recent foreign contributions, likely representing the inheritance of early groups of farmers involved in the expansion of agriculture from Elam. Also mtDNAs of B and L_Y can be considered plausible relics of autochthonous prehistoric/Neolithic tribes inhabiting the area before the arrival of Aryans. Long-term cultural isolation exacerbated by geographical features, such as mountainous and desert districts, indeed limited introgression of Indo-European lineages in their mtDNA pools.

References

Kivisild T., Bamshad M.J., Kaldma K., Metspalu M., Metspalu E., Reidla M., Laos S., Parik J., Watkins W.S., Dixon M.E., Papiha S.S., Mastana S.S., Mir M.R., Ferak V., Villems R. 1999. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr. Biol.*, 9: 1331-1334.

Macaulay V., Hill C., Achilli A., Rengo C., Clarke D., Meehan W., Blackburn J., Semino O., Scozzari R., Cruciani F., Taha A., Shaari N.K., Raja J.M., Ismail P., Zainuddin Z., Goodwin W., Bulbeck D., Bandelt H.J., Oppenheimer S., Torroni A., Richards M. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*, 308: 1034-1036.

Metspalu M., Kivisild T., Metspalu E., Parik J., Hudjashov G., Kaldma K., Serk P., Karmin M., Behar D.M., Gilbert M.T., Endicott P., Mastana S., Papiha S.S., Skorecki K., Torroni A., Villems R. 2004. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet.*, 5: 26.

Quintana-Murci L., Chaix R., Wells R.S., Behar D.M., Sayar H., Scozzari R., Rengo C., Al-Zahery N., Semino O., Santachiara-Benerecetti A.S., Coppa A., Ayub Q., Mohyuddin A., Tyler-Smith C., Qasim Mehdi S., Torroni A., McElreavey K. 2004. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am. J. Hum. Genet.*, 74: 827-845.

Regueiro M., Cadenas A.M., Gayden T., Underhill P.A., Herrera R.J. 2006. Iran: tricontinental nexus for Y-chromosome driven migration. *Hum. Hered.*, 61: 132-143.

Shepard E.M., Herrera R.J. 2005. Iranian STR variation at the fringes of biogeographical demarcation. *For. Sci. Int.*, 158: 140-148.