Genetic heterogeneity of Berber peoples as a result of differential migration and admixture patterns

Lara R. Arauna[1], Javier Mendoza-Revilla[1,2], Hassan Izaabel[3], Asmahan Bekada[4], Soraya Benhamamouch[4], Karima Fadhlaoui-Zid[5], Pierre Zalloua[6], Garrett Hellenthal[2], David Comas[1]

[1]Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Barcelona, Spain; [2]Genetics Institute, University College London, London, United Kingdom; [3]Laboratoire de Biologie Cellulaire & Génétique Moléculaire (LBCGM), Université IBNZOHR, Agadir, Morocco; [4]Department of Biotechnology, Faculty of Sciences, University of Oran, Oran, Algeria; [5]University El Manar, Tunis, Tunisia; [6]The Lebanese American University, Chouran, Beirut, Lebanon

Introduction

- North Africa has a strategic place in the Mediterranean that has made it a crossroad for many civilizations. Even during prehistory, North Africa exhibited a complex pattern of cultures that is still not well understood. Genetic studies reveal an amalgam of North African, Middle Eastern, European and Sub-Saharan ancestries.

- A "back to Africa" migration has been proposed based on genetic data. North African autochthonous component has diverged from Middle Eastern and European ancestral populations around 12,000 years ago (ya) (Henn et al., 2012). The Berbers are thought to be the descendants of that back migration.

- After the Arabic expansion, around 1,400 ya, in accordance with historical data, the population of North Africa become a mixture of Arabic and Berber groups, nowadays differentiated mainly by their language.

Conclusions

- Population structure analyses show a high genetic heterogeneity in North African populations and a lack of correlation between geography or ethnic and linguistic affiliations with population structure.

- North Africa has experienced complex gene flow since old times from Sub-Saharan Africa, Middle East and Europe (admixture can be detected at least from the first century).

- Differential admixture patterns, specially with Sub-Saharan Africa, produces genetic substructure within geographical populations probably caused by social structure.
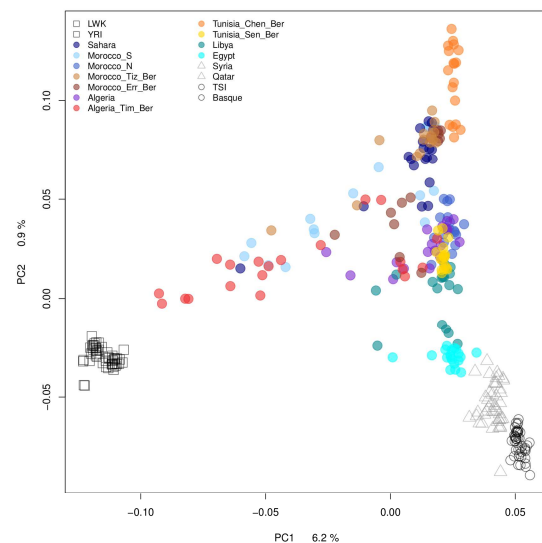

Figure 1. Principal Component Analysis.

ChromoPainter and Finestructure analysis

ChromoPainter is a method based on haplotype information that reconstructs the ancestral relationships of the individuals in a dataset. Each fragment of a haplotype (a chunk) is given from a donor to a recipient individual. The complete information of chunks copied from each donor allows to obtain a coancestry matrix.

Finestructure takes profit of the ancestral information reconstructed by ChromoPainter and identify population structure using a model-based Bayesian clustering.

The conjunction of both algorithms allows a fine analysis of the population by taking advantage of the genetic linkage information contained in the haplotypes.
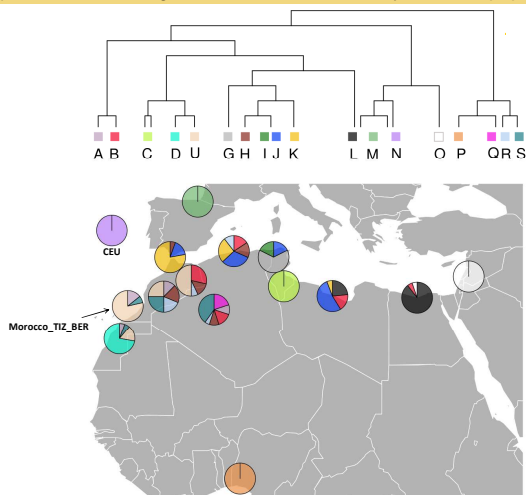

Figure 2. Finestructure tree based on the number of chunks that the individuals copy from each other. The map shows the distribution of the clusters established by the tree in the geographical populations.
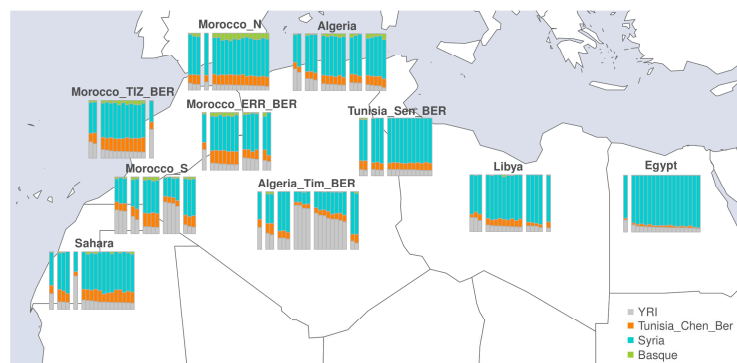

Figure 3. ChromoPainter was run setting four ancestral population donors. The plot shows the chunks length that each individual copies from each donor. The individuals within geographical populations are grouped by the Finestructure clusters (figure 2), for example, Morocco_TIZ_BER shows three groups that correspond to A, U and S clusters.

Globetrotter analysis

Globetrotter is a method developed to estimate the time of admixture between populations. It considers the information about ancestry sharing given by ChromoPainter. As a consequence of recombination, the distribution of the chunks copied from each donor depends on the time since admixture. This method constructs coancestry curves based on genetic distances and an estimation of the separation among chunks from the same donor. The date of admixture between the populations is then estimated from the decay of the curves.

For this analysis the populations were considered based on the clusters established by Finestructure. Four donor populations were considered taking into account the origin of gene flow observed in the region: Sub-Saharan, Middle Eastern, European and ancestral North African.
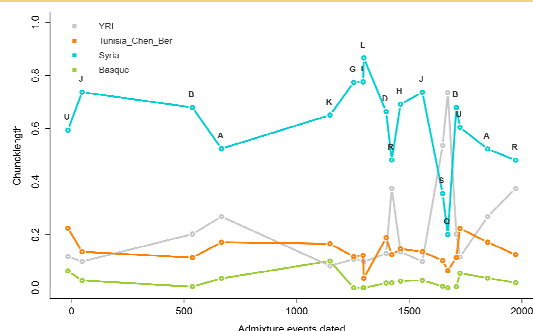

Figure 4. Graphical representation of the admixture dates estimated by Globetrotter. Each point corresponds to one cluster and the position in the x-axis indicates the admixture time, while the y-axis indicates the total chunklength that receives from each donor population.
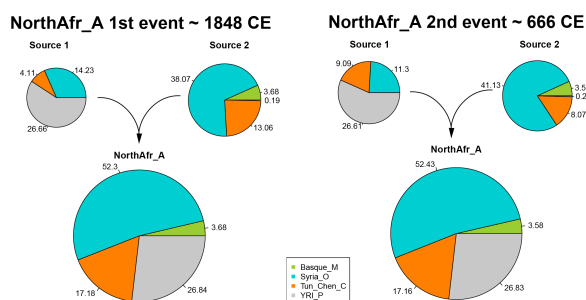

Figure 5. Example of a result of Globetrotter. NorthAfr_A cluster shows two admixture events, one around the 19th century and another in the 7th century. In each admixture event two sources give place to the final population.

References

- Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S. (2014). A genetic atlas of human admixture history. *Science (New York, N.Y.)*, *343*(6172), 747–51.
- Henn, B. M., Botigué, L. R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J. K., Comas, D. (2012). Genomic ancestry of North Africans supports back-to-Africa migrations. *PLoS Genetics*, *8*(1), e1002397.
- Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, *8*.
- The International HapMap Consortium. (2003). The International HapMap Project. *Nature*.