TITLE: Mitochondrial DNA structure in North Africa reveals a genetic discontinuity in the Nile Valley.

AUTHORS: Karima Fadhlaoui-Zid [1,*], Laura Rodríguez-Botigué [2,*], Nejib Naoui [3], Amel Benammar-Elgaaied [1], Francesc Calafell [2,4], David Comas [2,4]

*Both authors contributed equally to this work

[1]Laboratory of Genetics, Immunology and Human Pathology, Faculty of Science of Tunis, University Tunis El Manar, 2092 Tunis, Tunisia.

[2] Institut de Biologia Evolutiva (CSIC-UPF), Department de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08003 Barcelona, Spain.

[3] Laboratory of Medical Biology. Tunis, Tunisia.

[4] CIBERESP, 08003 Barcelona, Spain

26 text pages, 3 figures, 3 tables (plus 2 tables and 1 figure as Supplementary material).

RUNNING TITLE: mtDNA structure in North Africa

KEYWORDS: uniparental marker, haplogroup, gene flow

Corresponding author: DAVID COMAS

Institut de Biologia Evolutiva (CSIC-UPF), Department de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08003 Barcelona, Spain.

Tel : +34 93 3160843, Fax: +34 93 3160901

E-mail: david.comas@upf.edu

Abstract

Human population history in North Africa has been constrained in an east-west direction due to the geographical barriers imposed by the Sahara Desert and the Mediterranean Sea. Although these barriers have not completely impeded human migrations, genetic studies have shown that an east-west genetic gradient exists. However, the lack of genetic information of certain geographical areas and the focus of some studies in parts of the North African landscape have limited the global view of the genetic pool of North African populations. In order to provide a global view of the North African genetic landscape and population structure, we have analyzed ~2,300 North African mitochondrial DNA lineages (including 269 new sequences from Libya, in the first mtDNA study of the general Libyan population). Our results show a clinal distribution of certain haplogroups, some of them more frequent in Western (H, HV0, L1b, L3b, U6) or Eastern populations (L0a, R0a, Nib, I, J, M1) that might be the result of human migrations from the Middle East, sub-Saharan Africa, and Europe. Despite this clinal pattern, a genetic discontinuity is found in the Libyan/Egyptian border, suggesting a differential gene flow in the Nile River Valley. Finally, frequency of the post-LGM subclades H1 and H3 is predominant in Libya within the H sequences, highlighting the magnitude of the LGM expansion in North Africa.

North Africa is a region characterized by a complex history of demographic events and the extent of its genetic effect on extant human populations is still far from being known. Despite being part of the African continent, its demographic history, conditioned in an east-west direction by the barriers imposed by the Mediterranean Sea and the Sahara Desert, has been completely different from the rest of the continent. According to archaeological records, the first modern humans in North Africa produced the Aterian stone industry during the Early Upper Paleolithic, around 45,000 years ago (ya) (Garcea and Giraudi 2006). No clear connections have been established between this first human industry and subsequent cultures in the region, such as the Ibero-Maurusian industry (22,000 - 9,500 ya) (Newman 1995). The Ibero-Maurusian culture was followed by the Capsian industry (10,000 – 4,700 ya) (Desanges 1990) that persisted well after the adoption of farming and agriculture, which began around 5,500 years ago in the region. The persistence of a pre-Neolithic culture in Neolithic times might indicate cultural replacement with admixture, rather than a population replacement of the autochthonous pre-Neolithic people by Neolithic farmers originated in the Middle East. In general terms, the prehistoric cultural changes in North Africa were quite independent of the dynamics on the European shores of the Mediterranean. Historical records document trade routes across the Sahara Desert and contacts between both Mediterranean shores and the Middle East, such as Phoenicians, Romans, Vandals, and Byzantines. The first Arab invasion, initially confined to Egypt, started in A.D. 643 and may have involved only a few thousand individuals (McEvedy and Jones 1980). The Arabs began to impose their religion and language over the Berber autochthonous population, a process that culminated with the second and more numerous Arab wave in which the Bedouin reached the Maghreb (northwest Africa) in the 11th century. The later arrivals to northern Africa in colonial times include Europeans and Ottoman Turks, mainly in Egypt.

The genetic data available for North Africa is scarce, which limits the power to test population history hypotheses. Most of the African genetic diversity studies have been focused on the origin of our species and the first dispersals out of Africa (see for instance Tishkoff et al. (2009)), processes in which North Africa had a marginal role, which made the region less attractive to human population geneticists. The analyses based on frequencies of classical genetic polymorphisms (blood groups, red cell enzymes and serum proteins) have shown that the genetic landscape in northern Africa presents an east-west pattern of variation without differences between Arabs and Berbers, pointing to a sizeable Upper Paleolithic component in current northern African populations, whereas the Neolithic diffusion in the region was more a cultural than a demic process (Barbujani et al. 1994; Bosch et al. 1997). As for autosomal markers, only some STRs (Bosch et al. 2000; Khodjet-El-Khil et al. 2008) and *Alu* polymorphisms (Comas et al. 2000; Flores et al. 2000; Frigi et al. 2010; González-Pérez et al. 2003) have been analyzed in a few northern African samples. Concerning massive analysis of genome-wide markers, only 30 Mozabite individuals (a Berber isolate from Algeria) have been analyzed for 650,000 SNPs (Li et al. 2008), showing various degrees of admixture between sub-Saharans, Middle Easterners and Europeans. The analysis of Y chromosome lineages has shown a high frequency of two specific North African haplogroups (E1b1b1a and E1b1b1b), although their origins have been controversial since some analyses have suggested a Paleolithic component (Bosch et al. 2001), whereas others have pointed to a Neolithic origin (Arredi et al. 2004; Cruciani et al. 2004; Cruciani et al. 2007; Semino et al. 2004). The analysis of mitochondrial (mtDNA) lineages has shown that, in spite of an important sub-Saharan contribution, most haplogroups in North Africa are of Eurasian origin (Fadhlaoui-Zid et al. 2004; Harich et al. 2010; Krings et al. 1999; Plaza et al. 2003; Rando et al. 1998). Some can be traced to ancient Paleolithic times (such as haplogroups U6, M1, which are almost specific of northern African populations); however, some maternal lineages

have been recently acquired from Europe or the Middle East (such as haplogroups U5, V, R0a, J1b, U3) (Gonzalez et al. 2007; Maca-Meyer et al. 2003; Olivieri et al. 2006). Several studies suggest that at the end of the Last Glacial Maximum (LGM), around 13,000 ya, humans expanded from the Franco-Cantabrian refuge towards Europe and North Africa, spreading mtDNA haplogroups H1, H3 and V (Achilli et al. 2004; Cherni et al. 2009; Ennafaa et al. 2009; Loogvali et al. 2004; Pereira et al. 2005; Rhouda et al. 2009; Torroni et al. 1998; Torroni et al. 2001). However, a recent analysis of mtDNA diversity on Iberian populations points to the opposite conclusion: it suggests the absence of such an expansion (Garcia et al. 2010). In addition, a large degree of genetic heterogeneity has been shown in North African maternal lineages compared to other geographical regions such as Europe (Fadhlaoui-Zid et al. 2004; Plaza et al. 2003).

One of the main limitations of the genetic analyses of North African populations is the lack of representative and homogeneously distributed samples. For instance, most of the studies have focused in the north western samples and Egypt, being Libya a region with almost no genetic data with the exception of a Tuareg sample (Ottoni et al. 2009). The presence of gaps in the coverage of genetic studies across North Africa creates an artificial division between Eastern (Egypt) and Western populations (Morocco, Algeria and Tunisia) and prevents explicitly geography-based analyses. Due to this lack of data, most of population genetic research in the area has a local scope rather than being comprehensive and covering the whole region.

Within the present analysis we aim to address several questions concerning the population history of North Africa. Is there any genetic structure that differentiates North African populations? What is the influence of the trade routes and natural corridors such as the Nile? Conversely, have the particularly inhospitable conditions of the Western Egyptian and Libyan deserts created a genetic barrier? These questions have not been successfully

answered yet, partly because of the ~2,000 Km gap in sampling between Egypt and Tunisia. In the present work we analyze for the first time the mtDNA sequences of a set of 269 Libyans representing the general population of the country that will allow us filling this gap and study the whole region.

MATERIALS AND METHODS

Mitochondrial DNA sequencing and SNP genotyping in Libyan individuals

DNA was extracted from fresh blood from a total of 269 individuals from Libya using standard phenol-chloroform methods. The mtDNA control region was PCR amplified using primer pairs L15996 and H408, purified using the GFX PCR DNA and Gel Band purification Kit (GE Healthcare), and sequenced for both mtDNA hypervariable segments (HVSI and HVSII) as described previously (Plaza et al. 2004) using primer pairs L15996, H16401, L29, and H408 (Vigilant et al. 1989). Positions 16024 to 16391 for HVSI and positions 63 to 323 for HVSII (Anderson et al. 1981; Andrews et al. 1999) were considered for further analysis.

Four TaqMan® probes (Applied Biosystems) were used, following supplier's recommendations, to genotype positions 3594, 10873, 12705, and 14783, diagnostic for major lineages L0-L2/L5-L6, L/M, R, and M, respectively. After this first classification, 105 samples were subsequently refined by genotyping eight SNPs in the mtDNA coding region (7028, 10400, 10873, 11251 11719, 12308, and 12705 diagnostic for haplogroups H, M, L/M, J/T, R, U, and N and respectively) by means of a SNaPshot™ Multiplex kit (Applied Biosystems), as described previously (Bosch et al. 2006). Two of them (10873 and 12705) were typed with both methods and were used as controls. Finally, a dissection of haplogroup H was carried out using an additional SNaPshot™ multiplex reaction: positions 3010, 4793, 4336, 6776, and 14872 were typed in order to classify individuals into subhaplogroups H1, H7, H5a, H3, and H13, respectively. Sequences of primers used for PCR amplification and primers used for genotyping can be found in Supplementary Table 1.

Samples were assigned to haplogroups with the joint information of the control region sequence and the SNPs in the coding region following the nomenclature previously described (Behar et al. 2008; Finnilä et al. 2001; Kivisild et al. 2004; Loogvali et al. 2004; Maca-Meyer et al. 2003; Maca-Meyer et al. 2001; Metspalu et al. 2004; Olivieri et al. 2006; Palanichamy et al. 2004; Richards et al. 2000).

Statistical and phylogenetic analyses

Previously published data of HVSI sequences ranging from positions 16024 to 16383 of 28 North African populations was used for population comparisons. Their names and references can be found in Table 1. For the purposes of the present analysis, and given our focus on Libya, we define eastern North Africa as Egypt and Northern Sudan (Nubia), and western North Africa as Morocco, Western Sahara, Mauritania, Algeria, and Tunisia. Tunisian samples were 13 out of the 21 original samples in the database (62%); since Tunisia was clearly overrepresented (it contains ~6.5% of the population of the region), we sought to pool some of the Tunisian samples. The Tunisian samples that were not significantly different from each other and could be merged were identified with an AMOVA (Analysis of the Molecular Variance). In all subsequent analyses the so-called Tunisian "Andalousian" populations were pooled together in a single group and the Tunisian from Plaza et al. (2003) and the Urban Tunisian from Ottoni et al. (2009) were also merged into a single population. Sample locations are shown in Figure 1A. Intrapopulation diversity parameters (number of unique sequences, sequence diversity, nucleotide diversity and mean pairwise differences) of samples in the database were calculated using Arlequin 3.0 software (Excoffier et al. 2005). Analyses of the molecular variance (AMOVAs) were also performed with the Arlequin software.

A Correspondence Analysis (CA) was built using absolute haplogroup frequencies with the SPSS 15.0 software (SPSS Inc., Chicago, Illinois). Haplogroups with relative

frequencies below 1% were grouped with their superior clade. After using this criterion, a total of 23 haplogroups were included in the analysis. In addition, Andalusians, Macedonians, and Saudi Arabians were included as southern European and Middle Eastern representatives (Abu-Amero et al. 2008; Larruga et al. 2001; López-Soto and Sanz 2000; Plaza et al. 2003; Zimmermann et al. 2007), and South Sudanese and Mandenka as sub-Saharan representatives (Graven et al. 1995; Krings et al. 1999).

The geographical pattern of haplogroup distribution was investigated by computing Pearson's correlation coefficients (using SPSS 15.0 software, SPSS Inc., Chicago, Illinois) between the frequencies of each haplogroup and the geographical longitude of each population sample. Given the linear, East-West disposition of human settlement in N. Africa, longitude captures most of the geographical distance between populations.

Haplogroup specific median networks for haplogroups U6, M1 and H1 present in the data set and on Saudi Arabian populations were generated with the median joining algorithm (Bandelt et al. 1999) using the Network 4.5.1.0 program (http://www.fluxus-engineering.com). Networks were weighed taking into account the mutation rate of each position (Allard et al. 2002). Positions 16182 and 16183 were not considered because they mutate recurrently and therefore they are not phylogenetically informative. Time estimates were calculated using the rho statistic (Saillard et al. 2000) with one nucleotide substitution every 19,171 years for the HVSI sequences according to Soares et al. (2009).

RESULTS

Mitochondrial DNA sequence diversity and haplogroup composition in Libya

A total of 164 different sequences were found in the analysis of both hypervariable segments (HVSI and HVSII) in 269 Libyan individuals. Sequence diversity in Libyans is similar to other North African samples (Table 1), and when comparing only HVSI regions, 78 (29%) Libyan sequences were not found in the dataset used. Sequences and haplogroup

frequencies of the Libyan population are shown in Supplementary Table 2. Lineages of west Eurasian origin are the most frequent in Libyans (65%), followed by sub-Saharan L lineages (28%), and North African specific haplogroups U6 and M1, which have an overall frequency of 7%.

Haplogroups L2a1, L3b, L3f1b, and L1b are the most frequent (over 3%) sub-Saharan haplogroups in Libyans. Haplogroup L2a1 is common and apparently scattered throughout Africa (Salas et al. 2002), and therefore its geographical origin is difficult to assess. However, L1b, L3b, and L3f1b have more restricted locations in Africa. These haplogroups, together with other minor lineages in Libya such as L2b, L2c, L3d, and L3e have a typical Western Africa distribution (Harich et al. 2010; Salas et al. 2002). Nonetheless, other minor lineages present in Libya such as L0a, L3f, L3h, and L3x are more frequent in Eastern Africa. Such haplogroup frequency distribution suggests a predominantly Western origin of L lineages in Libya with some minor admixture of Eastern lineages.

The back to Africa M1 and U6 lineages are mainly present in North Africa and show opposite frequency gradients, being U6 significantly more frequent in the West, whereas M1 is more frequent in the East. Interestingly, these haplogroups display similar frequencies in the Libyan mtDNA pool (4.1% for U6 and 3.3% for M1).

Eurasian haplogroups HV0, H1, and K are the most frequent in Libyans (7.4%, 6.3% and 5.2% respectively). In order to trace back the geographical origin of the H lineages in Libya, we dissected haplogroup H in several subclades (see Material and Methods). Compared with previous published data (Ennafaa et al. 2009), Libyan individuals exhibit an admixture of western and eastern H subclades (Table 2). As in North West African populations, H1 and H3 are the most frequent subclades, and account for 48% of the H lineages. However, these frequencies are lower than those found in Maghreb populations

because of the relative high proportion of H5, H7, and H13 subgroups in Libya, which are more frequent in the Near East (Roostalu et al. 2006).

The age estimation of haplogroup H1 based on the HVSI region in Europe is 16.0 kya (Pereira et al. 2005) and 11.7 kya in Tunisia (Cherni et al. 2009). When H1 Libyan sequences are taken into account, coalescence age estimates in Libya (14.7 ± 4.4 kya) are compatible with those found in Tunisia. The haplogroup H1 network can be found in supplementary Figure 1.

North African maternal lineage landscape

In order to have a general view of the maternal genetic landscape in North Africa, a Correspondence Analysis (CA) based on haplogroup frequencies was built (Fig. 2). The first dimension separates sub-Saharan and southern samples on one edge, characterized by L haplogroups (except for some L3e subgroups), and Saudi Arabians on the opposite edge, characterized by R0a and J lineages. The second dimension follows a longitudinal pattern, grouping the Saudi Arabian, the Egyptian and the Sudanese samples on one edge and the Moroccan and European ones at the opposite side. North African populations form a large cluster in the center of the chart, without a clear structure. Nevertheless, it is noticeable that Egyptian populations and Nubians are placed in one edge of the second dimension, whereas Maghreb and European populations are grouped in the opposite edge.

A series of Analyses of Molecular Variance (AMOVA) were performed in order to test the proportion of the genetic variance within and among samples in North Africa. When all North African populations were considered as a single group, 3.88% ($p<0.01$) of the genetic variance was attributed to differences among samples. Then, we aimed to test how the apportionment of the maternal genetic variance was distributed across North Africa when two groups were considered in a west-east axis. In order to perform this test, we divided the whole region along four sections that were roughly limited by the actual geopolitical boundaries in

the region (Fig. 1B). Next, we performed a series of AMOVA analyses: in each new analysis the border between the two groups was moved progressively eastwards. Results showed that the amount of genetic variation was maximal (1.09%, p-value < 0.01) when the Eastern group was defined only by Egyptian and Sudanese populations.

In order to assess which haplogroups might be responsible of the differences found in the AMOVAs between eastern North Africa (Egypt and Sudan) and western North Africa together with Libya, we performed a correlation analysis between haplogroup frequencies and the longitude coordinates of the populations in our dataset (Table 3). Some lineages have higher frequencies in the West and decrease significantly towards the East, such as Eurasian H and HV0 haplogroups, sub-Saharan L1b and L3b haplogroups, and the North African U6 haplogroup. On the contrary, some lineages are more frequent in eastern samples, such as L0a and Eurasian haplogroups R0a, N1b, I, and J lineages. Interestingly, M1 does not reach statistical significance (p=0.055).

Phylogeographic analysis of North African U6 and M1 lineages

In order to deeply analyze the distribution and relationships of U6 and M1 lineages, a phylogeographic analysis was performed. Figure 3A shows the Median network of U6 haplotypes. As expected, the most represented groups are U6a* and U6a1, both of which show starlike phylogenies. Interestingly, both subclades bear similar diversity and many of the derived nodes are unique lineages (found only once in the database). The Maghreb is largely represented in the U6a clade. Of note is that most of the non-Maghreb U6a sequences are indeed from Libya. Moreover, Libya bears many unique sequences placed in basal and intermediate nodes spread all over the network, showing a high level of variability, in contrast with more eastern samples that show little diversity since all their sequences but one bear the root motifs of haplogroups U6a and U6a1. This is consistent with the mtDNA pool from Libyans being genetically closer to the Maghreb than to the northeast. Previous studies have

shown that minor subclades U6b and U6c are restricted to local areas (Maca-Meyer et al. 2003). The distribution of U6b was restricted to Morocco, Algeria and Eastern Bedouins; however, it has been found in Libya and Saudi Arabia (Abu-Amero et al. 2008) as well, extending its presence to nearly the entire North African area.

The coalescence time estimate for the U6 network (except for the U6c branch) is 44.0 ± 21.6 kya. Our coalescence age estimation based on the HVSI region for the haplogroup U6a1 is 13.0 ± 5.7 kya, whereas for U6a* is 13.5 ± 3.7 kya.

In a similar way to U6, M1 network shows that the basal lineages of M1 and M1a1 are the most common and have a starlike phylogeny (Fig. 3B) as well. Unlike U6, this clade is mostly represented in Egypt, Sudan and Saudi Arabia, where most of the unique sequences are found. Moreover, it should be noted that all Maghreb sequences, with only one exception, belong to the M1 haplogroup. Haplogroup distribution, haplogroup diversity and the high number of terminal nodes suggest that M1a1 arose in North East Africa and subsequently spread westwards. Coalescence age estimation for M1a1 is 13.1 ± 7.0 kya. When the entire M1 clade is considered, the time estimate increases to 23.1 ± 9.2. Note that coalescence ages for all the nodes that show starlike shape -M1a1, U6a*, and U6a1- are similar, which may be concordant with a common expansion.

DISCUSSION

The maternal lineage background in North Africa shows a moderate degree of East-West differentiation, with a genetic discontinuity between Libya and Egypt. This difference is summarized in the AMOVA that attributes 1.09% of the North African genetic variance to differences between Eastern and Western groups. Despite that other groupings within North Africa also yield significant differences in the AMOVAs, the differences found between Eastern and Western groups defined in the Libyan-Egyptian border are more than double

compared to the rest. Overall, the genetic structure within North Africa is the result of different haplogroup frequency distribution of L, U6, M1, and probably H lineages.

Besides L2a1, which is widespread in Africa, most sub-Saharan mtDNA haplogroups found in North Africa exhibit a slight east-west cline. L1b, L3b, and L3f1b lineages, which have a mainly western African distribution (Harich et al. 2010; Salas et al. 2002) are more frequent in NW African samples and rare in NE African populations. Harich and collaborators (2010) proposed that the origin of most of the sub-Saharan sequences found in North Africa can be found in the impact of the trans-Saharan slave trade routes that were established during recent times. This hypothesis could well explain the results found in Libya. According to trans-Saharan slave trade routes from Segal (2002), northern Libya was directly connected to western Africa with the Chad basin and was also interconnected with Tunisia, Algeria and Morocco, which were in turn connected with other western Africa locations. This would explain as well differences found in L haplogroups between our Libyan results and those found in Libyan Tuareg populations, where 18% of the L sequences are L0a1, a typical eastern African haplogroup (Ottoni et al. 2009). Libyan Tuaregs live in south-western Libya, along a trade route that interconnects this region with Egypt. Therefore, differences in sub-Saharan haplogroup distribution between these two Libyan samples could be due to gene flow either across the trade routes connecting North Africa and sub-Saharan Africa, or across North Africa itself. Indeed, the significant gradient of frequencies of haplogroups L1b and L3b shown with the correlation analysis agrees with this sub-Saharan genetic exchange within North Africa.

The distribution of haplogroups U6 and M1 also suggests the presence of a discontinuity between Libya and Egypt, separating western North Africa from eastern North Africa. Even if both haplogroups are thought to have been carried by a back-to-Africa migration from the Near East, significant increased U6 frequencies have been detected in the

West compared to the East. The network of all U6 sequences found in the database presents two nodes with star-like shape, U6a* and U6a1. In a similar way, M1a1 is the node with star-like topology in haplogroup M1. Time estimates of these nodes are 13.5 ± 3.7, 13.0 ± 5.7 and 13.1 ± 7.0 kya for haplogroups U6a*, U6a1, and M1a1 respectively. The most plausible explanation of the frequency distribution of M1 and U6 lineages, the coalescence age estimates, and the starlike shape would be an early split in the back to Africa migration followed by a period of stability and a period of expansion. The split would have produced two different migration waves, one westward, represented by U6, and the other southwards, represented by M1. Each haplogroup would have increased its frequency by drift and subsequently accumulated diversity over time. Coalescent time estimates point to an expansion of these haplogroups at the end of the LGM, simultaneously with some Eurasian haplogroups, as suggested by Olivieri et al. (2006). Moreover, all but one M1a1 sequence are found in eastern North Africa, which suggests that this subclade might have appeared in the East, and only after that have migrated westwards.

A similar East-West structure has been found with haplogroups related to the post-LGM expansion in the European Franco-Cantabrian area. A declining gradient of frequencies from west to east is detected for haplogroups H1 and H3. Moreover, the estimate age of haplogroup H1 agrees with previous estimates in North Africa, being 14.7 ± 4.4, 11.7 ± 3.6, and 11.3 ± 2.3 ya for Libya (present study), Tunisia (Cherni et al. 2009), and North Africa (Ennafaa et al. 2009) respectively. These dates set an upper limit for the presence of H1 in North Africa, which in any case is unlikely to have entered the region before the LGM. They are compatible with the posited post-LGM expansion from the Franco-Cantabrian glacial refuge area, although subsequent introductions cannot be ruled out. Unfortunately, no data is available for haplogroup H subclades in Egypt. The dissection of Egyptian H lineages would help to discern whether H1 is ubiquitous along North Africa or if a clear genetic barrier exists

between Libya and Egypt. Moreover, it would also be possible to discern whether a similar pattern has taken place from a post-LGM expansion in the Near Eastern refuge, considering that Libya has an increased frequency of typically Near Eastern haplogroups as H5, H7 and H13 compared to western North Africa.

The most plausible explanation for the differences found between NW and NE Africa is the presence of a demographic corridor along the Nile Valley. This corridor might have allowed the contact between Egypt, East Africa, and the Near East; leaving the rest of NW Africa apart from this Eastern contacts. Later, the Arab movements tied to the expansion of Islam did not apparently bridge the gap, at least for the female-transmitted mtDNA. The arid conditions west of the Nile may have conditioned population movement throughout much of human (pre)history, to the point of partially isolating the genetic pool of Egypt from those of countries to its west, including Libya.

ACKNOWLEDGEMENTS

We would like to thank Roger Anglada, Stephanie Plaza, and Mònica Vallés (UPF, Barcelona, Spain) for technical support. We also would like to thank all the volunteers that donated their DNA, making this study possible. The present research was supported by the Dirección General de Investigación, Ministerio de Ciencia e Innovación, Spain (CGL2010-14944/BOS); Tunisian Ministry of High Education and Scientific Research ; Programa de Cooperación Interuniversitaria e Investigación Científica (AECI), Ministerio de Asuntos Exteriores y Cooperación, Spain (A/5180/06, A/8394/07, B/018514/08); and Direcció General de Recerca, Generalitat de Catalunya (2009SGR1101).

LITERATURE CITED

Abu-Amero KK, Larruga JM, Cabrera VM, and Gonzalez AM. 2008. Mitochondrial DNA structure in the Arabian Peninsula. BMC Evolutionary Biology 8(1):45.

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V and others. 2004. The Molecular Dissection of mtDNA Haplogroup H Confirms That the Franco-Cantabrian Glacial Refuge Was a Major Source for the European Gene Pool. The American Journal of Human Genetics 75(5):910-918.

Allard M, Miller K, Wilson M, Monson K, and Budowle B. 2002. Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA dataset for 1771 human control region sequences. Scientific Working group on DNA analysis methods. Journal of Forensic Sciences 47:1215-1223.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F and others. 1981. Sequence and organisation of the human mitochondrial genome. Nature 290:457 - 465.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, and Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nature Genetics 23:147.

Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, and Tyler-Smith C. 2004. A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. American Journal of Human Genetics 75(2):338-345.

Bandelt HJ, Forster P, and Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16:37 - 48.

Barbujani G, Pilastro A, De Domenico S, and Renfrew C. 1994. Genetic variation in North Africa and Eurasia: neolithic demic diffusion vs. Paleolthic colonisation. American Journal of Physical Anthropology 95(2):137-154.

Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D and others. 2008. The Dawn of Human Matrilineal Diversity. The American Journal of Human Genetics 82(5):1130-1140.

Bosch E, Calafell F, González-Neira A, Flaiz C, Mateu E, Scheil HG, Huckenbeck W, Efremovska L, Mikerezi I, Xirotiris N and others. 2006. Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Annals of Human Genetics 70(4):459-487.

Bosch E, Calafell F, Pérez-Lezaun A, Clarimón J, Comas D, Mateu E, Martínez-Arias R, Morera B, Brakez Z, Akhayat O and others. 2000. Genetic structure of north-west Africa revealed by STR analysis. European Journal of Human Genetics 8(5):360-366.

Bosch E, Calafell F, Pérez-Lezaun A, Comas D, Mateu E, and Bertranpetit J. 1997. Population history of north Africa: evidence from classical genetic markers. Human Biology 69(3):295-311.

Bosch E, Clarimón J, Pérez-Lezaun A, and Calafell F. 2001. STR data for 21 loci in northwestern Africa. Forensic Science International 116(1):41-51.

Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, and Sajantila A. 2000. Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. Human Genetics 107(4):312-319.

Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B and others. 2004. Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. American Journal of Human Genetics 74(5):1014-1022.

Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, Dugoujon J, Crivellaro F, Benincasa T, Pascone R and others. 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia:new clues from Y-chromosomal haplogroups E-M78 and J-M12. Molecular Biology and Evolution 24(6):1300-1311.

Cherni L, Fernandes V, Pereira JB, Costa MD, Goios A, Frigi S, Yacoubi-Loueslati B, Amor MB, Slama A, Amorim A and others. 2009. Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. The American Journal of Human Anthropology 139(2):253-260.

Desanges J. 1990. The proto-Berbers. Mokhtar G, editor. Paris: Unesco.

Ennafaa H, Cabrera V, Abu-Amero K, Gonzalez A, Amor M, Bouhaha R, Dzimiri N, Elgaaied A, and Larruga J. 2009. Mitochondrial DNA haplogroup H structure in North Africa. BMC Genetics 10(1):8.

Excoffier L, Laval G, and Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fadhlaoui-Zid K, Plaza S, Calafell F, Amor MB, Comas D, Bennamar A, and El g. 2004. Mitochondrial DNA Heterogeneity in Tunisian Berbers. Annals of Human Genetics 68(3):222-233.

Finnilä S, Lehtonen MS, and Majamaa K. 2001. Phylogenetic network for European mtDNA. The American Journal of Human Genetics 68:1475 - 1484.

Flores C, Maca-Meyer N, González AM, and Cabrera VM. 2000. Northwest African distribution of the CD4/Alu microsatellite haplotypes. Annals of Human Genetics 64(Pt 4):321-327.

Frigi S, Ennafaa H, Amor MB, Cherni L, and Ammar-Elgaaied BA. 2010. Assessing human genetic diversity in Tunisian Berber populations by Alu insertion polymorphisms. Annalls of Human Biology(epub).

Garcea EA, and Giraudi C. 2006. Late Quaternary human settlement patterning in the Jebel Gharbi. Journal of Human Evolution 51:411-421.

Garcia O, Fregel R, Larruga JM, Alvarez V, Yurrebaso I, Cabrera VM, and Gonzalez AM. 2010. Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. Heredity(epub).

González-Pérez E, Via M, Esteban E, López-Alomar A, Mazieres S, Harich N, Kandil M, Dugoujon JM, and Moral P. 2003. Alu insertions in the Iberian Peninsula and north west Africa -- genetic boundaries or melting pot? Collegium Antropologicum 27(2):491-500.

Gonzalez AM, Larruga JM, Abu-Amero KK, Shi Y, Pestano J, and Cabrera VM. 2007. Mitochondrial lineage M1 traces an early human backflow to Africa. BMC Genomics 8:223.

Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti AS, Langaney A, and Excoffier L. 1995. Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of large Senegalese Mandenka sample. Molecular Biology and Evolution 12(2):334-345.

Harich N, Costa MD, Fernandes V, Kandil M, Pereira JB, Silva NM, and Pereira L. 2010. The trans-Saharan slave trade - clues from interpolation analyses and high resolution characterization of mitochondrial DNA lineages. BMC Evolutionary Biology 10:138-156.

Khodjet-El-Khil H, Fadhlaoui-Zid K, Gusmão L, Alves C, Benammar-Elgaaied A, and Amorim A. 2008. Substructure of a Tunisian Berber population as inferred from 15 autosomal short tandem repeat loci. Human Biology 80(4):435-448.

Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, Pennarun E, Parik J, Geberhiwot T, Usanga E, and Villems R. 2004. Ethiopian Mitochondrial DNA Heritage: Tracking Gene Flow Across and Around the Gate of Tears. The American Journal of Human Genetics 75:752 - 770.

Krings M, Salem A-eH, Bauer K, Geisert H, Malek AK, Chaix L, Simon C, Welsby D, Di Rienzo A, Utermann G and others. 1999. mtDNA Analysis of Nile River Valley Populations: A Genetic Corridor or a Barrier to Migration? The American Journal of Human Genetics 64(4):1166-1176.

Larruga JM, Diez F, Pinto F, Flores C, and Gonzalez AM. 2001. Mitochondrial DNA characterization of European isolates: the Maragatos from Spain. European Journal of Human Genetics 9(9):708-716.

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL and others. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. Science 319(5866):1100-1104.

Loogvali EL, Roostalu U, Malyarchuk BA, Derenko MV, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk HV, Parik J and others. 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. Molecular Biology and Evolution 21:2012 - 2021.

López-Soto M, and Sanz P. 2000. Polimorfismos de ADN Mitocondrial en individuos residentes en Andalucía y Extremadura. Cuadernos de Medicina Forense 20:17-29.

Maca-Meyer N, Gonzalez A, Pestano J, Flores C, Larruga J, and Cabrera V. 2003. Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. BMC Genetics 4(1):15.

Maca-Meyer N, Gonzalez AM, Larruga JM, Flores C, and Cabrera VM. 2001. Major genomic mitochondrial lineages delineate early human expansions. BMC Genetics 2:13.

McEvedy C, and Jones R. 1980. Atlas of the world population history. Hardmonswoth, UK.

Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, Serk P, Karmin M, Behar D, Gilbert MT and others. 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 Genetics 5(1):26.

Newman JL. 1995. The peopling of Africa: a geogrpahic interpretation. Press YU, editor: New Haven and London.

Olivieri A, Achilli A, Pala M, Battaglia V, Fornarino S Al-Zahery N, Scozzari R, Cruciani F, Behar DM, Dugoujon JM, Coudray C and others. 2006. The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. Science 314:1767 - 1770.

Ottoni C, Martínez-Labarga C, Loogvali EL, Pennarun E, Achilli A, De Angelis F, Trucchi E, Contini I, Biondi G, and Rickards O. 2009. First genetic insight into Libyan Tuaregs: a maternal perspective. Annals of Human Genetics 73(4):438-448.

Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V, and Zhang YP. 2004. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. The American Journal of Human Genetics 75:966 - 978.

Pereira L, Richards M, Gozos A, Alonso A, Albarran C, Garcia O, Behar DM, Golge M, Hatina J, Al-Gazali L and others. 2005. High-resolution mtDNA evidence for the late-glacial resettlementof Europe from an Iberian refugium. Genome Research 15:19 - 24.

Plaza S, Calafell F, Helal A, Bouzerna N, Lefranc G, Bertranpetit J, and Comas D. 2003. Joining the Pillars of Hercules: mtDNA Sequences Show Multidirectional Gene Flow in the Western Mediterranean. Annals of Human Genetics 67(4):312-328.

Plaza S, Salas A, Calafell F, Corte-Real F, Bertranpetit J, Carracedo Á, and Comas D. 2004. Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. Human Genetics 115(5):439-447.

Rando JC, Pinto F, Gonzalez AM, Hernandez M, Larruga JM, Cabrera VM, and Bandelt HJ. 1998. Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, Near-Eastern, and sub-Saharan populations. Annals of Human Genetics 62:531 - 550.

Rhouda T, Martinez-Redondo D, Gomez-Duran A, Noureddine E, Idaomar M, Diez-Sanchez C, Montoya J, Lopez-Perez MJ, and Ruiz-Pesini E. 2009. Moroccan mitochondrial genetic

background suggests prehistoric human migrations across the Gibraltar Strait. Mitochondrion 9:402-407.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T and others. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. The American Journal of Human Genetics 67:1251 - 1276.

Roostalu U, Kutuev I, Loogvali EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, and Villems R. 2006. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Molecular Biology and Evolution 24:436 - 448.

Saillard J, Forster P, Lynnerup N, Bandelt HJ, and Norby S. 2000. mtDNA variation among Greenland eskimos: the edge of the Beringian expansion. The American Journal of Human Genetics 67:718 - 726.

Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sanchez-Diz P, Macaulay V, and Carracedo A. 2002. The making of the African mtDNA landscape. The American Journal of Human Genetics 71:1082 - 1111.

Segal R. 2002. Islam's black slaves: a history of Afrca's other black diaspora. Books A, editor. London.

Semino O, Magri C, Benuzzi G, Lin AA, Al-Zaheri N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner P and others. 2004. Origin, diffusion, and differentiation

of Y-chromosome haplogroups E and J: inferences on the neoltihization of Europe and later migratory events in the Mediterranean area. American Journal of Human Genetics 74(5):1023-1034.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, and Richards MB. 2009. Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. The American Journal of Human Genetics 84(6):740-759.

Tishkoff SA, Reed FA, Friedlaender F, Ehret C, Ranciaro A, Froment A, Hirbo JB, Awomoyi AA, Bodo JM, Doumbo O and others. 2009. The genetic structure and history of Africans and African Americans. Science 324(5930):1035-1044.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B and others. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. The American Journal of Human Genetics 62:1137 - 1152.

Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E and others. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. The American Journal of Human Genetics 69:844 - 852.

Vigilant L, Pennington R, Harpending H, Kocher T, and Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proceedings of the National Academy of Sciences of the United States of America 86(23):9350-9354.

Zimmermann B, Brandstätter A, Duftner N, Niederwieser D, Spiroski M, Arsov T, and Parson W. 2007. Mitochondrial DNA control region population data from Macedonia. Forensic Science International: Genetics 1(3-4):e4-e

FIGURE LEGENDS

Fig.1. A. Map showing the location of the populations used in the present study. Boxes with numbers show the limits between sections used to divide the region. Populations are: **1** Moroccan Arab (MAR); **2** Moroccan Berber (MBE); **3** Figuig Berber (FIG); **4** Asni Berber (ASN); **5** Bouhria Berber (BOU); **6** Souss (SOU); **7** West Saharan (WSH); **8** Saharawi (SAH); **9** Mauritanian (MAU); **10** Algerian (ALG); **11** Mozabites (MZA); **12** Western Tuareg (WTUA); **13** Tunisian Urban (TUN_URB); **14** Matmata Berber (TMA); **15** Sened Berber (TSE); **16** Chenini-Douriet Berber; **17** Kesra Berber (KES); **18** Zriba Arab (ZRI); **19** Skira Berber (SKI); **20** Tunisian Andalusian (TUN_AND); **21** (DJE) Djerba; **22** Eastern Tuareg (ETUA); **23** Egyptian (EGY); **24** Upper Egypt (UPE); **25** Gurna (GUR); **26** Siwa (SIW); **27** Northern Nubian (NNUB); **28** Southern Nubian (SNUB); **33** Libya (LIB). B. Series of AMOVA results between and within groups including North African populations. Sample locations are represented in the map by dots. Five transects have been defined by the numbered white lines. Each analysis is represented by a raw in the bottom of the Figure. When two groups are defined, the split is located in one of the barriers limiting two transects, and populations laying on the left represent the Western group and populations on the right represented the Eastern group.

Fig.2. Results of the Correspondence Analysis performed with the dataset used and three additional populations, the Sub-Saharan Mandenka and the Europeans Macedonian and Andalusians. Haplogroups are represented by dots and they have been colored according to its prevalence. Populations are represented by squares and have been colored according to the Transect where they belong. Population codes are the same as in Figure 1.

Fig.3. Median joining network of sequences present in the dataset that belong to A) haplogroup U6, B) haplogroup M1. In both images, each haplotype is represented by a circle and its dimension is proportional to the number of individuals that bear that haplotype. Haplogroups are located beside their most probable "root" haplotype and numbers separating haplotypes correspond to the positions of the HVSII region that change from one haplotype to the other (positions are under the form "position – 16000"). Small red dots represent reticulation.

TABLE 1. Diversity measures within mtDNA HVSI in North African samples.

Population	n	k	Seq. diversity	π	PD	Reference
Moroccan Arabs	50	44	0.993 ± 0.007	0.0195 ± 0.0103	7.037 ± 3.356	Plaza et al. 2003
Moroccan Berbers	64	42	0.968 ± 0.013	0.0126 ± 0.0069	4.521 ± 2.251	Plaza et al. 2003
Figuig	94	29	0.937 ± 0.014	0.0171 ± 0.0091	6.173 ± 2.958	Coudray et al. 2009
Asni	53	36	0.963 ± 0.016	0.0151 ± 0.0082	5.424 ± 2.650	Coudray et al. 2009
Bouhria	70	38	0.964 ± 0.011	0.0157 ± 0.0084	5.661 ± 2.744	Coudray et al. 2009
Souss	50	34	0.961 ± 0.018	0.0128 ± 0.0071	4.604 ± 2.295	Brakez et al. 2001
West Saharans	25	20	0.973 ± 0.022	0.0148 ± 0.0082	5.340 ± 2.658	Rando et al. 1998
Saharaui	56	41	0.976 ± 0.012	0.0151 ± 0.0082	5.448 ± 2.659	Plaza et al. 2003
Mauritanian	64	43	0.979 ± 0.008	0.0178 ± 0.0095	6.407 ± 3.071	Rando et al. 1998
Algerian	47	29	0.965 ± 0.012	0.0164 ± 0.0088	5.894 ± 2.861	Plaza et al. 2003
Mozabites	85	30	0.943 ± 0.010	0.0134 ± 0.0073	4.822 ± 2.375	Macaulay et al. 1999
Western tuareg	23	21	0.992 ± 0.015	0.0190 ± 0.0103	6.838 ± 3.330	Watson et al. 1997
Tunisian Urban	98	83	0.992 ± 0.004	0.0172 ± 0.0094	6.433 ± 3.070	Plaza et al. 2003, Cherni et al. 2009

Tunisian Berber Matmata	49	29	0.946 ± 0.021	0.0140 ± 0.0077	5.050 ± 2.490	Fadahlaoui-Zid et al. 2004
Tunisian Berber Sened	53	37	0.975 ± 0.011	0.0209 ± 0.0110	7.527 ± 3.565	Fadahlaoui-Zid et al. 2004
Tunisian Berber Chenini-Douiret	53	23	0.939 ± 0.017	0.0189 ± 0.0100	6.823 ± 3.259	Fadahlaoui-Zid et al. 2004
Zriba Arab	50	16	0.904 ± 0.022	0.0110 ± 0.0062	3.948 ± 2.008	Cherni et al. 2005
Kesra Berbers	47	20	0.931 ± 0.021	0.0174 ± 0.0093	6.264 ± 3.022	Cherni et al. 2005
Tunisian Andalusian	155	84	0.965 ± 0.010	0.0155 ± 0.0083	5.581 ± 2.693	Cherni et al. 2009
Skira Berbers	20	14	0.937 ± 0.043	0.0118 ± 0.0068	4.237 ± 2.185	Cherni et al. 2009
Djerba	59	43	0.977 ± 0.011	0.0153 ± 0.0083	5.517 ± 2.687	Loueslati et al. 2006
Eastern tuareg	129	20	0.677 ± 0.046	0.0115 ± 0.0064	4.131 ± 2.068	Ottoni et al. 2009
Libyan	269	163	0.988 ± 0.003	0.0189 ± 0.0099	6.746 ± 3.189	Present study
Egyptian	344	232	0.993 ± 0.002	0.0190 ± 0.0099	6.832 ± 3.224	Krings et al. 1999, Saunier et al. 2009
Upper Egypt	24	24	1.000 ± 0.012	0.0234 ± 0.0125	8.427 ± 4.028	Stevanovitch et al. 2004
Gurna, Egypt	34	29	0.989 ± 0.010	0.0231 ± 0.0122	8.331 ± 3.947	Stevanovitch et al. 2004
Siwa	78	22	0.914 ± 0.014	0.0151 ± 0.0081	5.436 ± 2.644	Coudray et al. 2009
Egyptian Nubian	80	53	0.977 ± 0.008	0.0228 ± 0.0118	8.203 ± 3.840	Krings et al. 1999

Sudanese Nubian	76	66	0.995 ± 0.003	0.0236 ± 0.0122	8.482 ± 3.963	Krings et al. 1999

k stands for number of different sequences;

π stands for nucleotide diversity;

PD stands for mean Number of pairwise differences.

TABLE 2. Frequencies (%) of haplogroup H subclades within haplogroup H observed along the southern shore of the Mediterranean.

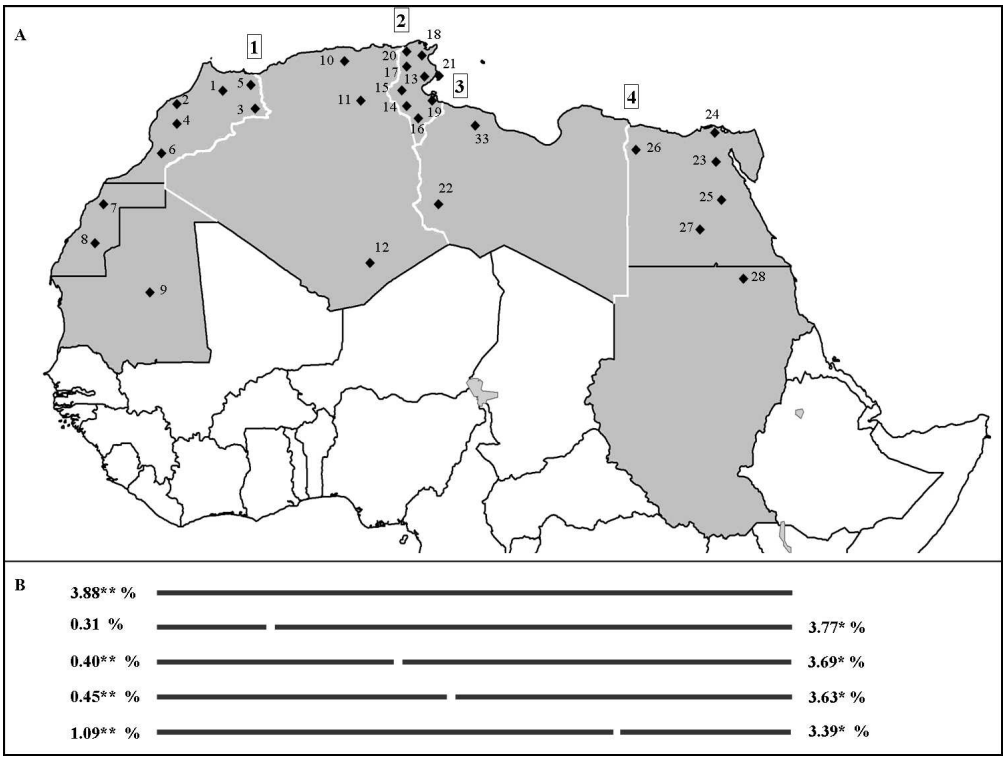
	Haplogroup H subclades								
Mauritania	H*	H1	H2a	H3	H5	H6	H7	H8	H13
Saharawi	18	64	-	9	-	-	9	-	-
Moroccan Arabs	14	68	-	7	.	11	-	-	-
Moroccan Berbers	14	63	-	17	6	-	-	-	-
Tunisian Berbers	47	34	4	4	2	4	2	-	-
Tunisian	28	24	-	21	3	-	17	3	-
Libyan	50	24	-	15	4	-	7	-	-
North Africa	24	37	2	11	9	2	7	-	9
Arabian Peninsula	32	43	1	13	2	2	5	<1	-
Near East	49	7	20	-	3	17	3	-	1
	61	14	1	-	10	4	4	1	4

[1] Data from Ennafaa et al. 2009

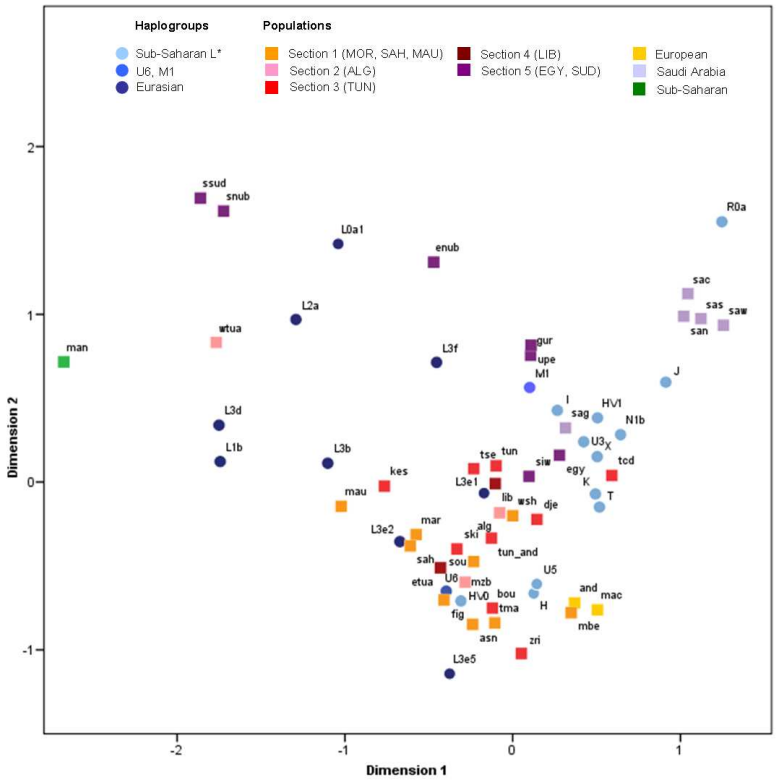
TABLE 3. Pearson correlation indexes and significance observed for the correlation between the longitudinal coordinate and the haplogroup frequencies for North African samples. Negative values represent higher frequencies in western samples, whereas positive values represent higher frequencies in Eastern samples.

	Pearson correlation		Pearson correlation
H	-0.508 (0.001)	L0a1	0.323 (0.048)
HV1	0.075 (0.656)	L1b	-0.569 (<0.001)
HV0	-0.420 (0.009)	L2a	-0.059 (0.727)
R0a	0.614 (<0.001)	L3b	-0.364 (0.025)
J	0.527 (0.001)	L3d	-0.097 (0.561)
T	0.123 (0.464)	L3e1	0.068 (0.687)
U3	0.181 (0.276)	L3e2	-0.193 (0.247)
U5	-0.296 (0.071)	L3e5	-0.271 (0.100)
K	-0.040 (0.814)	L3f	0.038 (0.820)
N1b	0.345 (0.034)	U6	-0.515 (0.001)
I	0.328 (0.044)	M1	0.314 (0.055)
X	-0.098 (0.557)		

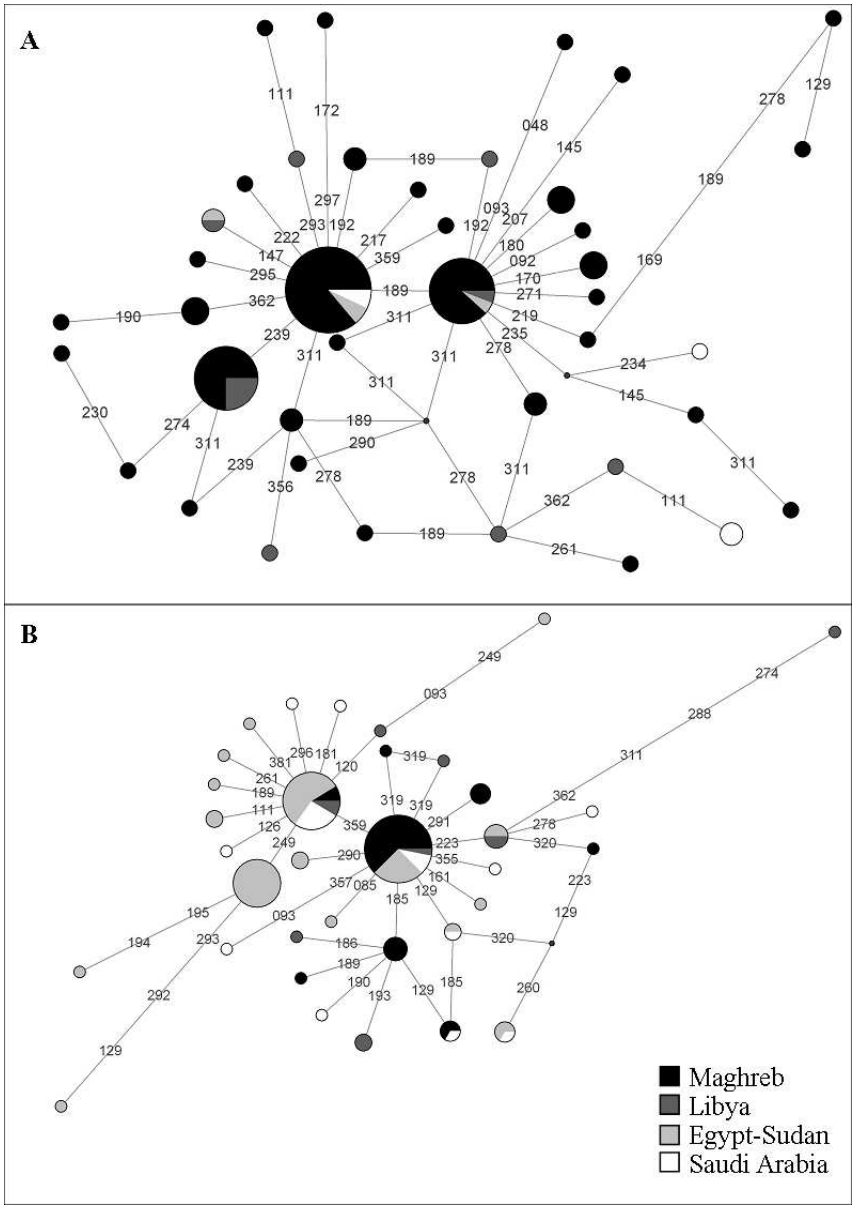
Significance (2 tailed) is shown in brackets



254x191mm (150 x 150 DPI)



234x187mm (150 x 150 DPI)



134x189mm (150 x 150 DPI)