Annals of human genetics

doi: 10.1111/j.1469-1809.2008.00493.x

The Complex and Diversified Mitochondrial Gene Pool of Berber Populations

C. Coudray[1*], A. Olivieri[2], A. Achilli[2,3], M. Pala[2], M. Melhaoui[4], M. Cherkaoui[5], F. El-Chennawi[6], M. Kossmann[7], A. Torroni[2] and J. M. Dugoujon[1]

[1]*Laboratoire d'Anthropobiologie, CNRS FRE2960, Université Paul Sabatier, Toulouse, France*
[2]*Dipartimento di Genetica e Microbiologia, Università di Pavia, Pavia, Italy*
[3]*Dipartimento di Biologia Cellulare e Ambientale, Università di Perugia, Perugia, Italy*
[4]*Université Mohamed 1er, Faculté des Sciences d'Oujda, Oujda, Morocco*
[5]*Laboratoire d'Ecologie Humaine, Université Cadi Ayyad, Marrakech, Morocco*
[6]*Faculté de Médecine, Université de Mansoura, Mansoura, Egypt*
[7]*Department of Languages and Cultures of Africa, Leiden University, Leiden, The Netherlands*

Summary

The mitochondrial DNA variation of 295 Berber-speakers from Morocco (Asni, Bouhria and Figuig) and the Egyptian oasis of Siwa was evaluated by sequencing a portion of the control region (including HVS-I and part of HVS-II) and surveying haplogroup-specific coding region markers. Our findings show that the Berber mitochondrial pool is characterized by an overall high frequency of Western Eurasian haplogroups, a somehow lower frequency of sub-Saharan L lineages, and a significant (but differential) presence of North African haplogroups U6 and M1, thus occupying an intermediate position between European and sub-Saharan populations in PCA analysis. A clear and significant genetic differentiation between the Berbers from Maghreb and Egyptian Berbers was also observed. The first are related to European populations as shown by haplogroup H1 and V frequencies, whereas the latter share more affinities with East African and Nile Valley populations as indicated by the high frequency of M1 and the presence of L0a1, L3i, L4*, and L4b2 lineages. Moreover, haplogroup U6 was not observed in Siwa. We conclude that the origins and maternal diversity of Berber populations are old and complex, and these communities bear genetic characteristics resulting from various events of gene flow with surrounding and migrating populations.

Keywords: Berbers, mtDNA, north african settlement, morocco, Egypt

Introduction

Even if archaeological and paleoanthropological records testify to the ancient (Paleolithic) human occupation of North Africa, the evolution of human groups living in that area is still unclear. This is mainly due to the large number of successive prehistoric and historic events that occurred in that area after the arrival of the first modern humans. In North Africa, the presence of Berbers – a term to denote those populations which speak a Berber language (Camps 1980) – is well described since the Capsian (10,000–4700 years ago), although this industry derived from oldest cultures. Nevertheless, it was only during the Neolithic transition (around 6000 years ago in the Saharan areas and 5000 years ago in the Maghreb) that North Africa was incontestably marked by various cultural events. Then, Berbers experienced a long and complicated history with many invasions, conquests and migrations by Phoenicians, Romans, Vandals and Byzantines (Brett & Fentress 1996). The most significant event was the Arab conquest, begun during the 7th century, when North Africans were converted to Islam, and Arabic became the official unique language employed. In spite of strong resistance, Berbers acquiesced to Arab authority. Refractory groups were driven out and constrained to more isolated areas. This troubled past directly influenced the geographical distribution of Berber communities which are nowadays scattered in a vast region extending from Mauritania to Egypt (Siwa oasis) and from the Sahara desert to the Moroccan Atlas mountainous

*Corresponding author: Clotilde Coudray, Laboratoire d'Anthropobiologie, CNRS FRE 2960, 37 allées Jules Guesde, 31073 Toulouse, France. Tel: 0033.5.61.14.59.85 or 87; Fax: 0033.5.61.14.59.79; E-mail: clotilde.coudray@wanadoo.fr

areas. Over the course of time, the various populations that migrated to North Africa have probably left a footprint in the gene pool of modern Berbers.

A detailed analysis of human mtDNA could be a powerful tool in reconstructing the population history of a region, such as Northern Africa, which is at the crossroads of Europe, the Middle East, and sub-Saharan Africa, regions for which mtDNA phylogeography is known in detail. Many studies have attempted to describe the genetic diversity of Berber populations, evaluating mitochondrial DNA (mtDNA) sequence variation (Achilli et al. 2005; Brakez et al. 2001; Cherni et al. 2005; Fadhlaoui-Zid et al. 2004; Krings et al. 1999; Loueslati et al. 2006; Macaulay et al. 1999; Olivieri et al. 2006; Plaza et al. 2003; Rando et al. 1998; Stevanovitch et al. 2004). In this study, we present sequence data from mtDNA, detected by surveying a portion of the control region, including hypervariable segment I (HVS-I) and most of hypervariable segment II (HVS-II) and haplogroup diagnostic RFLP markers. We studied four Berber populations: three from Morocco (Asni, Bouhria and Figuig) and one from Egypt (Siwa oasis). Our main objectives were to describe the maternal lineages of the Berbers, to determine the genetic contributions to their mitochondrial pool from surrounding populations (particularly from the sub-Saharan area, southern Europe, and the Near East), and to link their present-day genetic diversity to the numerous human migrations which affected North Africa.

Materials and Methods

Populations

We studied 295 unrelated and healthy individuals from four different Berber populations. Three groups were from Morocco, Asni (N = 53), Bouhria (N = 70) and Figuig (N = 94), and one was from the Egyptian oasis of Siwa (N = 78). The geographical location of the sampled sites is presented in Figure 1. Asni is located in the Tacheddirt valley in the High Atlas Mountains (47 km from Marrakech). The Berbers from Asni speak Chleuh and belong to the Rhiraya tribe. Sidi Bouhria is located in north-eastern Morocco (Oujda wilaya). Berbers from Bouhria belong to the Beni Moussi Roua fraction of the Beni Iznasen tribe. Figuig is located in eastern Morocco, at the Algerian border, at the juncture of the High Plateaus and the north-western edge of the Sahara. It is an oasis encompassing seven villages, and the Berber population that we studied lives in the largest of these, Zenaga. The Siwa oasis is located in the western Egyptian desert, at 300 km from the Mediterranean coast and 25 km from the Libyan border. This population speaks Siwi, a Berber language. Before blood collection, people were interviewed in order to ascertain their ethnic origins and to obtain informed consent. All individuals and their families included in this study have been living in the area of interest for at least three generations. In addition, mtDNA haplogroup information from 58 European, Near Eastern, and African populations was used for comparative analyses (see Table 1).

DNA Extraction

For the Moroccan Berbers, genomic DNA was isolated from blood samples using a standard proteinase-K digestion followed by phenol-chloroform extraction and ethanol precipitation, while for the sample from Siwa, genomic DNA was isolated from blood using the NucleoSpin Blood Mini Kit (BD Biosciences Clontech, San Jose, USA).

MtDNA Amplification, Sequencing and RFLP Screening

A portion of the mtDNA control region encompassing the entire HVS-I and part of HVS-II was PCR amplified using primer pairs F15877 (5′-CAAATGGGCCTGTCCTTGTA-3′) and R468 (5′-GGAGTGGGAGGGGAAAATAA-3′) for samples from Bouhria, Asni and Siwa, and primers F15973 (5′-AACTCCACCATTAGCACCCA-3′) and R296 (5′-GGAAAT TTTTTGTTATGATGTCT-3′) for the Figuig sample. DNA purification was undertaken using the QIAquick PCR Purification Kit (QIAgen, Courtaboeuf, France). Reactions were carried out using the BigDye Terminator (version 3.1) Ready Reaction Cycle Sequencing kit, with AmpliTaq DNA polymerase (PE, Applied Biosystems, Foster City, USA). Sequences were run in

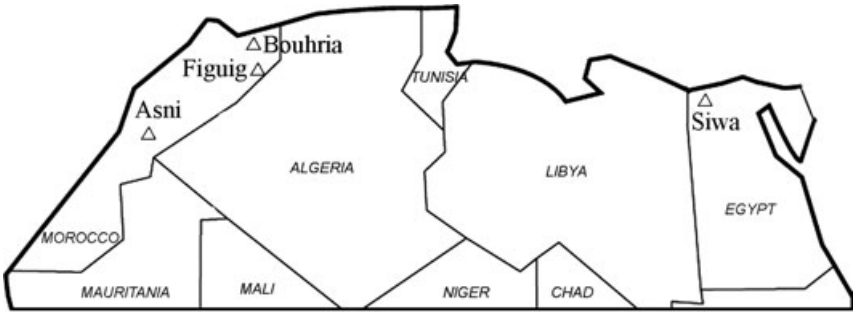

Figure 1 Geographical location of samples.

Table 1 Analyzed populations

	Country	Sample size	(References)	Identification code
	AFRICA			
Morocco	Asni	53	(present study)	1 or Asn
	Bouhria	70	(present study)	2 or Bou
	Figuig	94	(present study)	3 or Fig
	Souss	50	(Brakez et al. 2001)	4 or Sou
	Arabs	50	(Plaza et al. 2003; Rando et al. 1998)	5
	Saharawis	56	(Plaza et al. 2003)	6
Algeria	Mozabites	85	(Corte-Real et al. 1996; Macaulay et al. 1999)	7 or Moz
	Arabs	47	(Plaza et al. 2003)	8
Tunisia	Chenini Douiret	53	(Fadhlaoui-Zid et al. 2004)	9 or Che
	Sened	53	(Fadhlaoui-Zid et al. 2004)	10 or Snd
	Matmata	49	(Fadhlaoui-Zid et al. 2004)	11 or Mat
	Jerba Berbers	29	(Loueslati et al. 2006)	12 or Bjr
	Jerba Arabs	30	(Loueslati et al. 2006)	13
	Arabs	47	(Plaza et al. 2003)	14
Egypt	Siwa	78	(present study)	15 or Siw
	Gurna Egypt	34	(Stevanovitch et al. 2004)	16
	Adaima Muslims	100	(Dugoujon JM, unpublished data)	17
	Adaima Copts	100	(Dugoujon JM, unpublished data)	18
	Egypt	118	(Rowold et al. 2007)	19
	Niger-Nigeria	132	(Watson et al. 1997)	20
	Mauritania	64	(Gonzalez et al. 2006; Rando et al. 1998)	21
	Senegal	121	(Rando et al. 1998)	22
	Sierra Leone	270	(Jackson et al. 2005)	23
	Guinea Bissau	361	(Rosa et al. 2004)	24
	Benin	87	(Rowold et al. 2007)	25
	Central African Republic	30	(Batini et al. 2007)	26
	Popular Republic of Congo	94	(Batini et al. 2007)	27
	Sao Tomé e Principe	103	(Trovoada et al. 2004)	28
	Chad	448	(Cerny et al. 2007)	29
	Cameroon	441	(Coia et al. 2005)	30
	Mali	124	(Gonzalez et al. 2006)	31
	Angola	44	(Plaza et al. 2004)	32
	Rwanda	63	(Rowold et al. 2007)	33
	Kenya	100	(Brandstatter et al. 2004)	34
	Mozambique	416	(Pereira et al. 2001; Salas et al. 2002)	35
	Ethiopia	270	(Kivisild et al. 2004)	36
	Sudan	154	(Pereira et al. 2001; Underhill et al. 2000)	37
	NEAR EAST			
	Yemen	115	(Kivisild et al. 2004)	38
	Saudi Arabia	553	(Abu-Amero et al. 2008)	39
	Oman	105	(Rowold et al. 2007)	40
	Qatar	90	(Rowold et al. 2007)	41
	UAE	131	(Rowold et al. 2007)	42
	Iraq	116	(Richards et al. 2000)	43
	Israel	117	(Richards et al. 2000)	44
	Jordan	146	(Richards et al. 2000)	45
	Syria	118	(Richards et al. 2000; Vernesi et al. 2001)	46
	Turkey	340	(Di Benedetto et al. 2001; Quintana-Murci et al. 2004; Richards et al. 2000)	47

Table 1 Continued

	Country	Sample size	(References)	Identification code
	EUROPE			
	France	210	(Dubut et al. 2004)	48
Spain	Basque	173	(Bertranpetit et al. 1995; Corte-Real et al. 1996; Richards et al. 2000)	49
	Spain	305	(Alvarez et al. 2007)	50
	Cordoba	108	(Casas et al. 2006)	51
	Spain Center	148	(Larruga et al. 2001)	52
	Spain North East	179	(Crespillo et al. 2000; Plaza et al. 2003; Richards et al. 2000)	53
	Spain North West	216	(Gonzalez et al. 2003; Richards et al. 2000)	54
	Andalusia	114	(Larruga et al. 2001; Plaza et al. 2003; Richards et al. 2000)	55
	Portugal	542	(Gonzalez et al. 2003; Pereira, Prata, & Amorim 2000)	56
Italy	Italy Center	1273	(Torroni A, unpublished data)	57
	Italy North	346	(Torroni A, unpublished data)	58
	Italy South	539	(Torroni A, unpublished data)	59
	Sardinia	370	(Richards et al. 2000); (Torroni A, unpublished data)	60
	Sicily	105	(Richards et al. 2000); (Torroni A, unpublished data)	61
	Greece	155	(Richards et al. 2000; Vernesi et al. 2001)	62

an automatic Sequencer ABI PRISM 3730 (PE, Applied Biosystems). Both strands were sequenced using primers F15973 and R263 (5′-TGGCTGTGCAGACATTCAAT-3′), and only confirmed deviations from the reference sequence were considered. In addition to the control region sequencing, and in order to classify mtDNAs unambiguously within haplogroups, coding region diagnostic markers were surveyed (see supplementary material table).

Statistical and Phylogenetic Analyses

The mtDNA sequences were aligned using the BIOEDIT software (Tom Hall, Carlsbad, USA). Only segments from positions 16033 to 220 according to the revised Cambridge Reference Sequence rCRS (Anderson et al. 1981; Andrews et al. 1999) were considered for analysis. Haplogroup assignment followed the basic classification scheme of Torroni et al. (1996; 2006). Additional reports (Achilli et al. 2004; Achilli et al. 2005; Behar et al. 2008; Kivisild et al. 2004; Macaulay et al. 1999; Olivieri et al. 2006; Quintana-Murci et al. 2004; Quintana-Murci et al. 1999; Richards et al. 2000; Richards et al. 1998; Salas et al. 2004; Salas et al. 2002; Shen et al. 2004) were consulted to further delineate the haplogroups and to assign geographical labels (West Eurasian, East Eurasian, North African, and sub-Saharan African). Population internal genetic diversity parameters (number of different sequences, haplotype diversity, nucleotide diversity, and mean number of pairwise differences) were calculated for each sample using Arlequin 3.0 software (Excoffier et al. 2005). Based on the available haplotypes, population-specific median networks were generated with the median joining algorithm (Bandelt et al. 1995; 1999) using the NETWORK 4.5 software program developed by Fluxus Technology Limited (Sudbury, UK http://www.fluxus-engineering.com). The weighting scheme for the nucleotide positions used in this analysis (nps 16033–220) followed Richards et al. (1998). Subsequently, the most parsimonious tree was inferred and drawn manually and confirmed by using the Network MP-calculation option. For summarizing population relationships, Principal Component Analysis (PCA) of haplogroup frequencies was performed using the Excel Stat package (Microsoft). PCA provides a method for representing frequency data in a Euclidian space so that the results can be visually examined for structure (Greenacre 1992). For the PC analyses, we used the frequencies of the following haplogroups: H (including H∗ and all H subhaplogroups), HV0 (including HV0* and V), HV (including HV* and HV1), R0 (represented mainly by R0a), J, T, U (including all U lineages except U6 and K), U6, K, N1 (including N1 and I lineages), N2 (represented mainly by W), X, M (including M* and all M subhaplogroups except M1), M1, L0, L1, L2, L3, and L4-L5. Population genetic structure was tested by analysis of variance (AMOVA) and F statistics (Excoffier et al. 1992; Weir & Hill 2002) using the Arlequin program (Excoffier et al. 2005). The significance of the covariance components associated with the different levels of genetic structure was tested using a non-parametric permutation procedure (Excoffier, Smouse & Quattro 1992). For this analysis, population samples were grouped into major geographic

areas: North Africa, East Africa, sub-Sahara, the Near East and Europe. We estimated the genetic variation among populations within the same geographical group (FSC), among geographical groups (FCT), and among populations across the entire study area (FST). Molecular pairwise F_{ST} indexes between populations were calculated from haplogroup frequencies, and their significance was evaluated with a nonparametric permutation test using the Arlequin package (Excoffier, Laval & Schneider 2005).

Results

MtDNA Sequence Diversity in the Four Berber Populations

The total number of different control region haplotypes in our sample of 295 Berbers was 127. Among these, only 9 haplotypes were shared between two or three populations (4 between Asni and Bouhria, 3 between Asni and Figuig, 1 between Bouhria and Figuig and 1 among Asni, Bouhria and Figuig). Among the 53 Berbers from Asni, 42 different sequences belonging to 19 distinct haplogroups were found. The 70 Berbers from Bouhria were characterized by 38 different sequences and 26 haplogroups/sub-haplogroups. In the 94 Berbers from Figuig, 32 different sequences and 20 haplogroups were detected. Finally, among the 78 Egyptian Berbers, the 25 different observed sequences were subdivided into 19 distinct haplogroups.

Haplogroup frequencies in the four Berber populations are reported in Table 2. The four mtDNA pools show different compositions but overall harbor a high frequency of Eurasian haplogroups (H, HV0, HV, R0, J, T, U w/o U6, K, N1, N2, and X), and a relatively lower frequency of sub-Saharan L lineages (L0, L1, L2, L3, and L4-L5). Both North African-specific clades U6 and M1 are represented but with extremely different distributions. In detail, Asni and Bouhria samples showed slightly similar genetic profiles whereas those from Figuig and Siwa were quite different, in particular because of their higher frequency of L and M1 haplogroups, respectively. When lineages are considered one by one, several differences appear within Moroccan populations, and between Moroccans and Egyptians. For example, among West Eurasian lineages, haplogroup H was the most common in Morocco whereas haplogroups U (without U6 and K) and K were the most represented in Siwa. As for the sub-Saharan lineages in Morocco, the Berbers from Figuig were found to harbor the highest proportion of L mtDNAs. M1 mtDNAs were instead much more frequent in Siwa than in Morocco, whereas U6 was only observed in Morocco with the highest frequency in Asni.

Table 3 shows the descriptive parameters of the analyzed populations. Gene diversity values in Asni and Bouhria samples are very similar whereas Figuig and Siwa samples show a reduced diversity level, Siwa showing the lowest values. For the nucleotide diversity and mean number of pairwise differences, the sample from Figuig tends to have slightly higher values. The observed reduced diversity in the maternal gene pool of the Berbers from Figuig and Siwa could be attributable to genetic drift because of the relative isolation of the oases where they live.

Genetic diversity values were compared with those of a set of North African Berber populations. The diversity indices obtained here for Asni (0.983), Bouhria (0.966) and Figuig (0.956) samples are in full accordance with previous data. The observed high level of mtDNA sequence diversity is in the range of values generally found in North Africa: 0.941 for the Souss Valley (Brakez et al. 2001), 0.942 for Mozabites (Corte-Real et al. 1996; Macaulay et al. 1999), 0.939, 0.975 and 0.964 for Chenini-Douiret, Sened and Matmata (Fadhlaoui-Zid et al. 2004), respectively. Within all Berber populations, the Siwi show the lowest value of sequence diversity (0.921).

The Phylogeography of North African Berbers

To visualize the mtDNA diversity of the four Berber samples, a tree was constructed and presented in Figure 2 (for mitochondrial sequences, see additional file 1).

Eurasian Lineages in Moroccan and Egyptian Berbers

Thirty-three, 57, 47 and 45 individuals out of 53, 70, 94 and 78 samples from Asni, Bouhria, Figuig and Siwa (respectively) carried mtDNAs that are typical of western Eurasian populations, particularly from Europe.

Haplogroup H is by far the most frequent in western European populations. It accounts for 40%-50% of the mtDNA pool in most of Europe, and 20%–30% in the Near East and the Caucasus region (Richards et al. 2000; Roostalu et al. 2007). It is thought to have evolved in the Near East and Middle East ~23,000–28,000 years ago and to be involved in the late glacial expansions that took place into Europe ~13,000–15,000 years ago, mainly from the Iberian refuge (Achilli et al. 2004; Loogvali et al. 2004; Pereira et al. 2005; Torroni et al. 1998). H is observed at high frequencies in Moroccan samples (37.7%, 37.1% and 24.4% in Asni, Bouhria and Figuig, respectively) whereas it accounts for only 1.3% of the Egyptian Berbers. Different subclades of H lineages are present in Asni (H*, H1, H3, and H7), Bouhria (H*, H1, H2, H3, H5, H6, and H11), Figuig (H1 and H8), and Siwa samples (H1). Among these, H1, which is the most common sub-clade of H in Europe (Achilli et al. 2004; Finnila et al. 2001; Herrnstadt et al. 2002; Loogvali et al. 2004), is also the most represented in the Berber samples.

HV0 probably originated in Europe before the Last Glacial Maximum. MtDNAs belonging to this haplogroup are generally rare and apparently scattered throughout Europe,

Table 2 Frequencies of mtDNA haplogroups in the four Berber populations (Except for H*, the use of an asterisk for haplogroup labels (i.e. HV0*, K*, X*...) means that the samples belong to these haplogroups, since the sub-clades diagnostic mutations have not been checked. H* indicates H samples non belonging to tested sub-clades (H1–3, H5–8, H11).)

	Haplogroup frequencies (number of subjects) by population			
Haplogroup	Asni (N = 53)	Bouhria (N = 70)	Figuig (N = 94)	Siwa (N = 78)
H:	*0.377 (20)*	*0.371 (26)*	*0.244 (23)*	*0.013 (1)*
H*	0.132 (7)	0.071 (5)	...	...
H1	0.207 (11)	0.186 (13)	0.212 (20)	0.013 (1)
H2	...	0.014 (1)	...	...
H3	0.019 (1)	0.014 (1)	...	...
H5	...	0.043 (3)	...	...
H6	...	0.014 (1)	...	...
H7	0.019 (1)	...	...	...
H8	...	...	0.032 (3)	...
H11	...	0.029 (2)	...	...
HV0:	*0.113 (6)*	*0.143 (10)*	*0.085 (8)*	*0.127 (10)*
HV0*	0.038 (2)	0.129 (9)	0.053 (5)	0.127 (10)
V	0.075 (4)	0.014 (1)	0.032 (3)	...
HV (without H/HV0):	...	*0.014 (1)*	...	*0.013 (1)*
HV*	...	...	...	0.013 (1)
HV1	...	0.014 (1)	...	...
R0:	*0.038 (2)*	...	...	*0.026 (2)*
R0a	0.038 (2)	...	...	0.026 (2)
J:	...	...	*0.117 (11)*	*0.051 (4)*
J1c	...	...	0.053 (5)	...
J2	...	...	0.032 (3)	0.051 (4)
J2a	...	...	0.032 (3)	...
T:	...	*0.100 (7)*	*0.032 (3)*	*0.013 (1)*
T1a	...	0.086 (6)	...	0.013 (1)
T2	...	...	0.032 (3)	...
T2b	...	0.014 (1)	...	...
U (without U6):	*0.019 (1)*	*0.057 (4)*	*0.011 (1)*	*0.180 (14)*
U2b	...	0.014 (1)	...	...
U2e	...	0.029 (2)	...	...
U3	...	...	...	0.013 (1)
U3a	...	...	0.011 (1)	...
U4	...	0.014 (1)	...	...
U5b	0.019 (1)	...	...	0.167 (13)
U6:	*0.113 (6)*	*0.014 (1)*	*0.032 (3)*	...
U6a	0.094 (5)	0.014 (1)	0.032 (3)	...
U6d	0.019 (1)	...	...	...
K:	*0.038 (2)*	*0.057 (4)*	...	*0.115 (9)*
K*	0.038 (2)	0.057 (4)	...	0.115 (9)
N1:	*0.019 (1)*	*0.014 (1)*	...	...
N1b1	0.019 (1)	0.014 (1)	...	...
N2:	...	*0.029 (2)*	...	...
W*	...	0.029 (2)	...	...

Table 2 Continued

Haplogroup	Haplogroup frequencies (number of subjects) by population			
	Asni (N = 53)	Bouhria (N = 70)	Figuig (N = 94)	Siwa (N = 78)
X:	*0.019 (1)*	*0.029 (2)*	*0.011 (1)*	*0.038 (3)*
X*	0.019 (1)	0.029 (2)	0.011 (1)	0.038 (3)
M:	...	...	...	*0.013 (1)*
M33	...	...	...	0.013 (1)
M1:	*0.038 (2)*	*0.043 (3)*	*0.021 (2)*	*0.167 (13)*
M1a	0.038 (2)	...	...	...
M1a1	...	...	0.021 (2)	0.167 (13)
M1b	...	0.043 (3)	...	...
L0:	...	...	*0.043 (4)*	*0.013 (1)*
L0a1	...	...	...	0.013 (1)
L0f	...	...	0.043 (4)	...
L1:	*0.075 (4)*	*0.072 (5)*	*0.064 (6)*	*0.013 (1)*
L1b	...	0.029 (2)	...	...
L1b1	0.075 (4)	0.043 (3)	0.064 (6)	0.013 (1)
L2:	*0.057 (3)*	*0.028 (2)*	*0.043 (4)*	*0.051 (4)*
L2a	...	0.014 (1)	...	0.051 (4)
L2a1	0.019 (1)	0.014 (1)	0.032 (3)	...
L2b1	0.038 (2)	...	...	...
L2d2	...	...	0.011 (1)	...
L3:	*0.075 (4)*	*0.029 (2)*	*0.297 (28)*	*0.128 (10)*
L3b	...	...	0.064 (6)	...
L3b1	...	...	0.021 (2)	...
L3e1	...	...	...	0.115 (9)
L3e2	...	...	0.043 (4)	...
L3e5	0.075 (4)	0.029 (2)	0.169 (16)	...
L3i2	...	...	...	0.013 (1)
L4-L5:	*0.019 (1)*	...	...	*0.039 (3)*
L4*	...	...	...	0.026 (2)
L4b2	...	...	...	0.013 (1)
L5	0.019 (1)	...	...	...
Total Eurasian lineages (H, HV0, HV, R0, J, T, U (without U6), K, N1, N2, X)	0.623 (33)	0.814 (57)	0.500 (47)	0.576 (45)
Total Asian lineages (M)	...	...	...	0.013 (1)
Total Sub-Saharan lineages (L0, L1, L2, L3, L4-L5)	0.226 (12)	0.129 (9)	0.447 (42)	0.244 (19)
Total North African lineages (U6, M1)	0.151 (8)	0.057 (4)	0.053 (5)	0.167 (13)

although they are less uncommon in the Mediterranean area (including northwest Africa) (Torroni et al. 2001). Indeed, HV0 mtDNAs were observed in 6 subjects from Asni, 10 from Bouhria, 8 from Figuig and 10 from Siwa. HV0 lineages were represented by both HV0* (2, 9, 5 and 10 sequences in Asni, Bouhria, Figuig, and Siwa, respectively) and V sub-clades. V sequences generally represent 5% of European mtDNAs and tend to be restricted to western, central, and northern Europe (Torroni et al. 1998; Torroni et al. 2001). Four, one and three individuals belonging to this haplogroup were detected in the Asni, Bouhria and Figuig samples, respectively, whereas V was not observed in Siwa.

Table 3 Descriptive parameters in four Berber populations from Morocco and Egypt (N = sample size; K = number of different control-region sequences and percentage of sample size in brackets; H = haplotype diversity and standard error (SE) in parentheses; D = nucleotide diversity; M = mean number of pairwise differences)

Population	N	K (K/N)	H (SE)	D	M
Asni	53	42 (79.2)	0.983(0.0099)	0.011990	9.100145
Bouhria	70	38(54.3)	0.966(0.0103)	0.011833	8.981366
Figuig	94	32(34.0)	0.956(0.0091)	0.013886	10.511325
Siwa	78	25(32.1)	0.921(0.0141)	0.011798	8.955045

Only one individual from Bouhria showed a HV1 sequence (a sister clade of HV0 and H) and one Egyptian Berber bore a HV* sequence. Haplogroup R0a, a sister clade of HV, occurs in populations from the Near East, the Caucasus, and Mediterranean Europe. It was found in two Berbers from Asni and two Siwi.

J and T sister clades might have originated in the Near East ~50,000 years ago, but they could have been recently introduced into Europe during the Neolithic ~10,000 years ago (Finnila et al. 2001; Palanichamy et al. 2004; Richards et al. 2000). Haplogroup J is observed in populations from the southern Caucasus, the Near East, and North

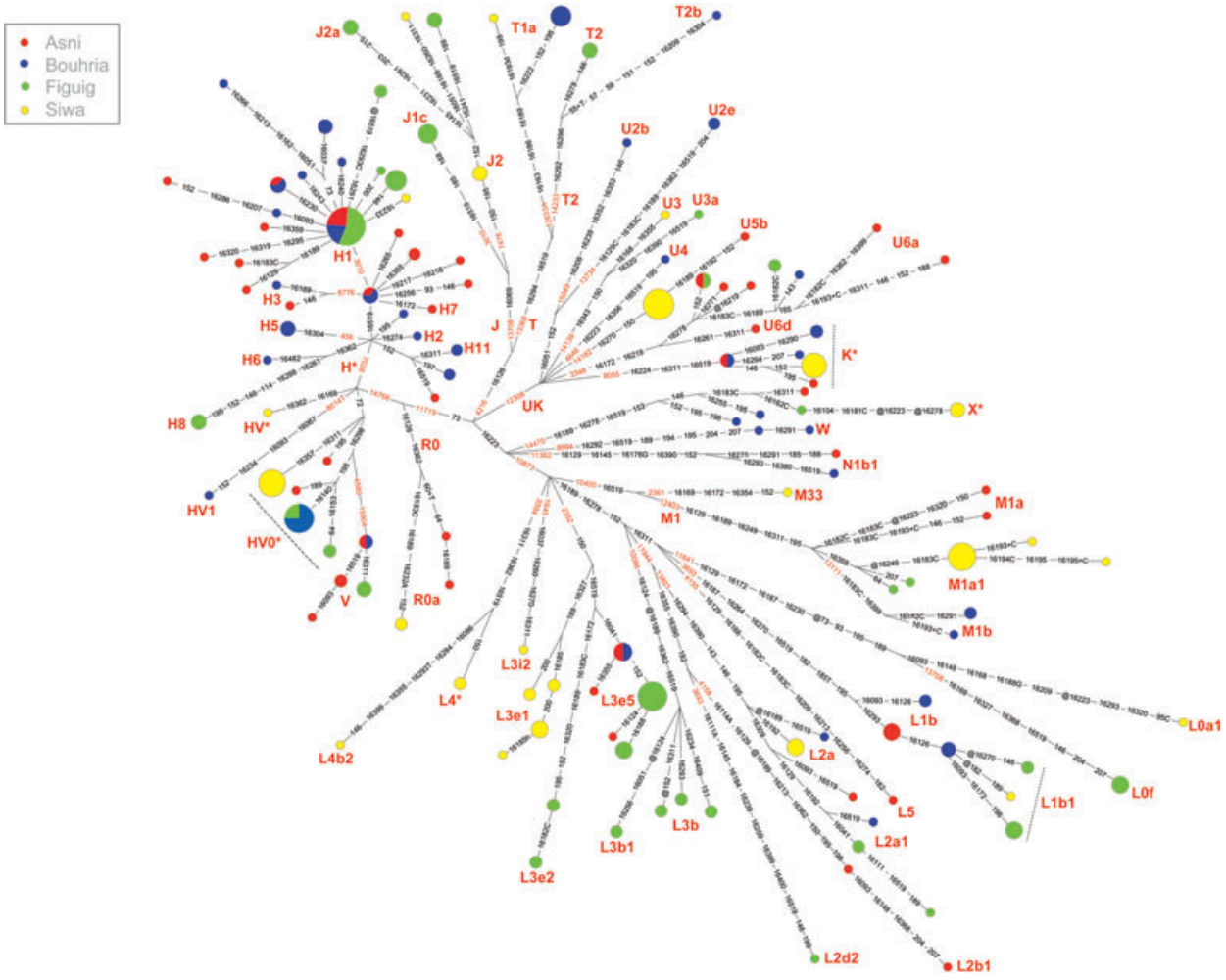

Figure 2 Tree of the 127 mtDNA haplotypes observed among the 295 Berbers
Haplogroups and subhaplogroups are reported in red. The numbers on the connecting branches refer to the revised reference sequence rCRS (Andrews et al. 1999), and indicate either a control region mutation or a haplogroup diagnostic site (in red). Mutations are transitions unless the base change is explicitly indicated by the suffix "A", "G", "T" or "C" for a transversion, "+ and the inserted nucleotide" for an insertion, "d" for a deletion, "h" for heteroplasmy, and @ for a reversion. The size of each circle is proportional to the haplotype frequency.

Africa (Brakez et al. 2001; Maca-Meyer et al. 2001; Plaza et al. 2003; Richards et al. 2000). It is not found in the Asni and Bouhria samples, whereas the Berbers from Figuig harbor five J1c, three J2, and three J2a mtDNAs, and the Berbers from Siwa show four J2 mtDNAs. The most numerous T sequences were observed in Bouhria (six T1a and one T2b) while the Figuig samples exhibited three T2 and the Siwa sample only one T1a mtDNA. Haplogroup T was not detected in Asni.

Eurasian U macrohaplogroup is subdivided into U1 to U9 (with the exception of U6 restricted to North Africa) and K lineages. U has an extremely broad geographical distribution and accounts for about 20% of European mtDNA sequences (Herrnstadt et al. 2002). In this study, it is represented by four haplogroups: U2, U3, U4 and U5. U2 can be found at low frequencies in populations of western Asia and the Caucasus. Here, it was observed only in the Berbers from Bouhria, by two subclades, U2b (1 sequence) and U2e (2 sequences). U3 is observed at 1.1% (U3a) and 1.3% (U3) in samples from Figuig and Siwa, respectively. This haplogroup has a frequency peak in the Near East (Achilli et al. 2007). U4 haplogroup is spread at moderate frequencies all over Europe, western Siberia, and southwestern Asia (Richards et al. 2000). Only one sequence from Bouhria belongs to U4. U5, one of the most ancient subhaplogroup of U, occurs in most cases as occasional haplotypes that are derived from European lineages (Achilli et al. 2005; Richards et al. 2000). In 16.7% of Siwi samples, it appears as U5b, a lineage suggesting back-migration of people from Europe to the South (Torroni et al. 2006). One Berber from Asni also bears this sequence. K lineages are observed at 3.8%, 5.7% and 11.5% in samples from Asni, Bouhria and Siwa, respectively.

N1b1 is a minor mtDNA haplogroup that has been observed at marginal frequencies in European and Near Eastern populations (Palanichamy et al. 2004; Richards et al. 2000). It occurs at a very low frequency in both Asni (1.9%) and Bouhria (1.4%) samples.

N2 sequences, represented by W* haplogroup, is only present in two Berbers from Bouhria.

X* mtDNAs occur at low frequency in Asni (1.9%), Bouhria (2.9%), Figuig (1.1%) and Siwa (3.8%) samples.

All major European mtDNA haplogroups are most likely of Middle Eastern origin (Torroni et al. 1996, 2001). Their presence at high frequencies in Berber samples indicates either gene flow from Europe to North Africa and/or a probable common ancestry of populations of both shores of the Mediterranean Sea. Despite total West Eurasian lineages occur at equally high frequency in the four Berber samples, they exhibit a different distribution between Morocco and Egypt, with the Siwi showing the lowest frequency of the West European H but the highest frequencies of HV0* and K haplogroups which are more prevalent in East Eurasia and the Near East.

Sub-Saharan Llineages in Moroccan and Egyptian Berbers

L African lineages which account for 22.6%, 12.9%, 44.7% and 24.4% of Asni, Bouhria, Figuig and Siwa mitochondrial pool are represented by six haplogroups: L0, L1, L2, L3, L4 and L5 (Gonder et al. 2007; Torroni et al. 2006).

Haplogroup L0 is the earliest offshoot of the mtDNA tree in Africa that appears as a sister group to the branch that holds all other haplogroups (Salas et al. 2002). L0a subclade probably originated in Eastern Africa and is common in Eastern, Central, and South Eastern Africa. One L0a1 sequence was found in an Egyptian Berber. Although L0f subclade is rare and geographically confined to eastern Africa, it was observed in four Berbers from Figuig.

Haplogroup L1 appears in our dataset as L1b and L1b1 sub-clusters. L1b and L1b1 are particularly concentrated in Western Africa, although with some overflow into Central and North Africa (Salas et al. 2002). L1b is observed in 2 sequences from Bouhria and L1b1 in 4, 3, 6, and 1 individuals from Asni, Bouhria, Figuig, and Siwa, respectively.

Three sequences from Asni, two from Bouhria, four from Figuig and four from Siwa, belong to L2 lineages. They appear as four clusters, L2a (in Bouhria and Siwa), L2a1 (only in Moroccan populations), L2b1 (in Asni), and L2d2 (in Figuig). While L2a is the most common and widespread L2 subclade all over Africa, L2b is largely restricted to western Africa and L2d is rather rare but also appears to be restricted to Western Africa (Salas et al. 2002; Torroni et al. 2001).

L3 lineages appear in our dataset as three clades, L3b, L3e and L3i (Salas et al. 2002). L3b is widely spread in Western African populations and Bantu-speaking South Eastern Africans and it is also observed in Ethiopia (Kivisild et al. 2004; Salas et al. 2002). L3b accounts for 8.5% of Figuig mtDNA sequences while it is absent in the three other populations.

L3e is prevalent in central (∼20%) and western (∼11%) Africa but it is less abundant (∼3%) in eastern Africa. It originated from the Sudan (Bandelt et al. 2001). It is observed as 9 L3e1 in Siwa, as 4 L3e2 in Figuig, and as 4, 2 and 16 L3e5 in Asni, Bouhria and Figuig samples, respectively. L3e1 is the oldest and most diverse clade among L3e types. Albeit practically omnipresent in sub-Saharan and South Eastern Africa, it is supposed to have had a western/central African origin. L3e2 is omnipresent in Africa (Bandelt et al. 2001). L3e5 is considered to be restricted to North Africans (Fadhlaoui-Zid et al. 2004; Torroni et al. 2006). L3i haplogroup seems to be restricted to Ethiopia and Yemen (Kivisild et al. 2004). One individual from the Siwa oasis exhibits an L3i2 sequence (Behar et al. 2008).

L4 haplogroup, which is an early branch of L3 lineages, is only represented in Berbers from Siwa by 2 L4* (L4 subclade not defined) and 1 L4b2 sequences (Behar et al. 2008). L4 is

an east African lineage and reveals high sequence diversity in Ethiopians (Kivisild et al. 2004).

L5 has been observed at low frequency only in eastern Africa (Kivisild et al. 2004). It is present in one individual from Asni.

L lineages are specific of the mitochondrial pool of sub-Saharan populations. Their presence at relatively high frequencies in North African Berbers could be attributed to a gene flow from the Saharan areas. However, these exchanges do not seem to have the same geographical origin: in Morocco, frequencies of L1b, L2b, L2d, L3b and L3e5 haplogroups indicate a predominantly west sub-Saharan influence, whereas in Egypt, L0a, L3e1, L3i and L4 lineages underline migration routes from eastern Africa.

M Asian, M1 and U6 North African Lineages in Moroccan and Egyptian Berbers

M1 is a subclade of southwest Asian ancestry that moved to Africa about 40,000 to 45,000 years ago (Olivieri et al. 2006). Its highest frequencies are found in Egypt (∼18%) and in eastern Africa (∼11% in Ethiopia and Somalia). It is not uncommon in the Mediterranean basin showing a peak in the Iberian Peninsula and it is also observed in the Middle East (Gonzalez et al. 2007). In our dataset, M1 is represented by two subclades, M1a, particularly abundant in Ethiopia, and M1b detected only in the Mediterranean area (Olivieri et al. 2006). We observed two M1a, three M1b, two M1a1, and 13 M1a1 sequences from Asni, Bouhria, Figuig, and Siwa samples, respectively. Finally, one mtDNA from Siwa was found to belong to the South Western Asian M33a clade (Abu-Amero et al. 2008).

U6 is a North African clade, that similar to M1, is of southwest Asian ancestry and most likely entered Africa ∼40,000–45,000 years ago (Olivieri et al. 2006). U6 is distributed from the Near East to Northwest Africa where it is found at its highest frequencies (Maca-Meyer et al. 2003; Olivieri et al. 2006). It was also observed at a sparse distribution in populations from the southwestern part of the Iberian Peninsula (Plaza et al. 2003). In our dataset, U6 as a whole is only observed in the Moroccan samples, where it is represented by its most significant subgroup, U6a (five mtDNAs in Asni, one in Bouhria, and three in Figuig). In Asni, one U6d mtDNA was also observed.

Although U6 is considered specific of Berber populations, its absence in the Siwa sample suggests a differentiation between Berber-speakers living at the extremes of their geographical distribution range. An opposite trend is that observed for M1 that is much more common in the Berbers from Egypt than in the Berbers from Morocco. This reveals a more significant east African influence in the North Eastern African genetic pool than in the North Western one.

Mitochondrial Diversity in a Broad Geographical Context

To observe the mtDNA patterns of Berber populations within a broad geographical context, we performed a first Principal Component Analysis (PCA) including data from 62 European, Near Eastern and African populations (Fig. 3a). Principal components 1 and 2 (axes 1 and 2 in Fig. 3) account for 46% of the total variation (32% and 14%, respectively). The two-dimensional pattern displays distinct population clusters: the sub-Saharan populations located in the most positive portion of axis 1; the East Africans and the Near Eastern groups in the upper-right and upper-left quadrant, respectively; the Europeans and some East Mediterranean populations located in the negative portion of axis 1; and North Africans sandwiched between the sub-Saharan cluster and European populations, but closer to the latter. If we detail the position of our Berber samples in the overall mitochondrial diversity, the Moroccan Berbers from Bouhria are closest related to Europeans whereas Figuig are furthest away, with Asni being in an intermediate position. Siwi Berbers are far from the other three Moroccan samples and have closest relationships with East Mediterranean populations of Jordania, Syria, Israel and Iraq.

The contribution of each haplogroup to the first two Principal Components is illustrated in Figure 3b. The population positions on the PCA are explained by the distinction between African and Eurasian haplogroups. As expected, sub-Saharan samples are characterized by high frequencies of L lineages (L0, L1, L2 and L3) whereas Europeans are mainly associated with lineages H, HV0, U, K, T, N2 and X. Then, East Africans show high frequencies of M1 and L4-L5 lineages, and Near Easterners are mainly associated with lineages HV, R0, J, N2, M, and N1. Finally, the centered position of North Africans is explained by higher frequencies of West Eurasian lineages with respect to the African haplogroups, by the small frequencies of East Eurasian and M1 lineages, and by the presence of U6 haplogroup.

Overall, the PCA plot based on mitochondrial haplogroup frequencies (Fig. 3a) reflects a clear differentiation between Berbers and sub-Saharans but a close relationship between Berbers and South Western European populations. This fits within the phylogeographical patterns established by previous genetic research (Plaza et al. 2003; Rando et al. 1998; Rosa et al. 2004).

We carried out analyses of variance (AMOVA) on the populations used for the PC analysis (Table 4). For this, population samples were grouped into major geographical areas: North Africa, sub-Sahara, East Africa, the Near East and Europe. Results show that most of the genetic variation (83.38%) occurs within the populations whereas there is a relatively little (2.61%) but significant differentiation among populations within the same geographical region ($F_{SC} = 0.030$). A

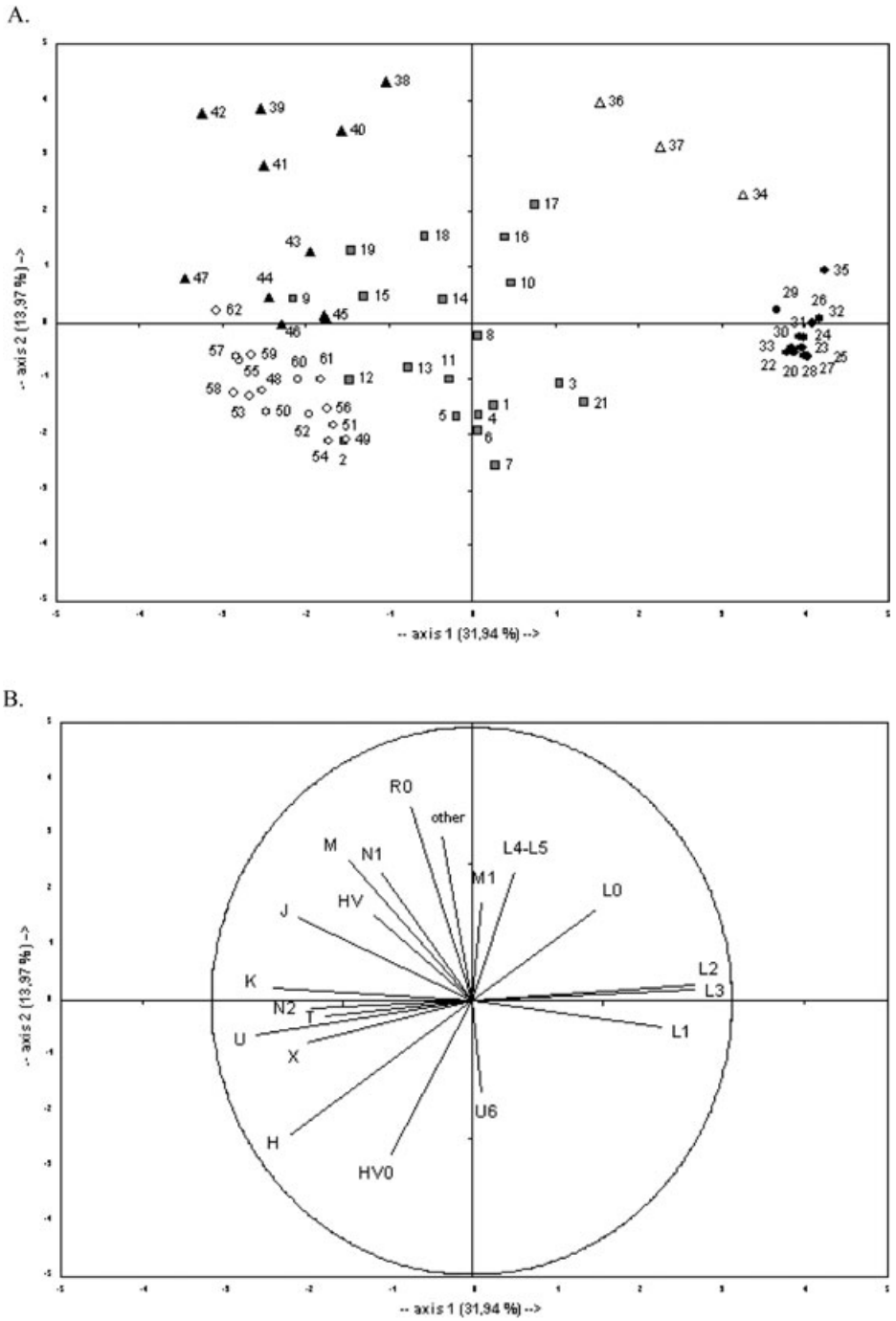

Figure 3 a) Principal Component Analysis (PCA) of mtDNA haplogroup profiles of 62 populations from Africa, Europe, and the Near East; b) Plot of the contribution of each haplogroup to the first two axes of the PCA
Gray squares: North Africans; White circles: Europeans; Black circles: sub-Saharans; White triangles: East Africans; Black triangles: Near Easterners. See Table 1 for population abbreviation code. Haplogroups and subhaplogroups included in the PCA are as follow: H (including H* and all H subhaplogroups), HV0 (including HV0* and V), HV (including HV* and HV1), R0 (represented mainly by R0a), J, T, U (including all U lineages except U6 and K), U6, K, N1 (including N1 and I lineages), N2 (represented mainly by W), X, M (including M* and all M subhaplogroups except M1), M1, L0, L1, L2, L3, and L4-L5.

Table 4 Analyses of Variance (AMOVA) in 62 population samples (Percentile distribution of the variance components and fixation indices (in parenthesis) – FCT, FSC and FST, respectively – are given. All values are significant ($p<0.05$) after 1000 permutations)

Groups	Among groups	Among populations within groups	Within populations
All populations (North Africa, sub-Sahara, East Africa, the Near East, Europe)	14.01	2.61	83.38
	(Fixation indices: FCT = 0.14011	FSC = 0.03031	FST = 0.16617)
North Africa vs. sub-Sahara	12.53	4.65	82.82
	(Fixation indices: FCT = 0.12526	FSC = 0.05316	FST = 0.17176)
North Africa vs. East Africa	4.08	4.29	91.63
	(Fixation indices: FCT = 0.04075	FSC = 0.04475	FST = 0.08368)
North Africa vs. the Near East	1.39	4.15	94.46
	(Fixation indices: FCT = 0.01390	FSC = 0.04211	FST = 0.05542)
North Africa vs. Europe	5.01	1.63	93.36
	(Fixation indices: FCT = 0.05010	FSC = 0.01714	FST = 0.06638)
sub-Sahara vs. East Africa	5.53	4.95	89.52
	(Fixation indices: FCT = 0.05530	FSC = 0.05235	FST = 0.10475)
sub-Sahara vs. the Near East	16.94	3.97	79.08
	(Fixation indices: FCT = 0.16942	FSC = 0.04785	FST = 0.20916)
sub-Sahara vs. Europe	25.72	1.76	72.52
	(Fixation indices: FCT = 0.25716	FSC = 0.02376	FST = 0.27481)
East Africa vs. the Near East	6.67	3.56	89.77
	(Fixation indices: FCT = 0.06672	FSC = 0.03813	FST = 0.10231)
East Africa vs. Europe	17.31	0.75	81.93
	(Fixation indices: FCT = 0.17312	FSC = 0.00911	FST = 0.18065)
The Near East vs. Europe	4.75	1.56	93.69
	(Fixation indices: FCT = 0.04753	FSC = 0.01638	FST = 0.06313)

14.01% of the variance was attributed to differences among geographical groups ($F_{CT} = 0.140$). This means that populations within each group are relatively homogeneous, that groups are significantly differentiated and that the mitochondrial diversity is structured according to a geographical classification. Moreover, F_{ST} index is definitely higher than 0.005 ($F_{ST} = 0.166$) showing a strong diversification among populations. By performing AMOVA between pairs of geographical groups, we measured the population structuring observed in the PC analysis (Table 4). We found that 12.53% of the variance was attributed to differences among North African and sub-Saharan groups, 4.08% among North Africans and East Africans, 1.39% among North Africans and Near Eastern groups, and 5.01% among North Africans and Europeans. However, the highest value is observed between Europeans and sub-Saharans (25.72% of the total variance).

Berber Mitochondrial Diversity

An additional PC analysis was performed only on the ten Berber populations of our database (Fig. 4a) by considering their haplogroup frequencies (Table 5). The first two principal components account for 41% of the total variation (22% and 19%, respectively). The first axis reveals a clear differentiation between Siwa samples and all other Berber populations from Morocco, Tunisia and Algeria, while the second axis shows the Berbers from Chenini-Douiret (Tunisia) as the following most differentiated group. Our Moroccan samples are located at the centre of the graph. Locations of the population on the PCA plot are explained by the distinction between M, L4-L5, M1, X, K, U, R0, HV0, and L0 haplogroups on positive values (by descending order of contribution on axis 1) and H, T, HV, L1, L2, U6, N1, J, L3, and N2 lineages for negative values of axis 1 (Fig. 4b). In particular, the Siwa mitochondrial pool is characterized by the highest frequencies of M1 lineages (16.7%, from 0% to 11.3% in the Maghreb, see Table 5), by the lowest frequencies of H haplogroups (1.3%, from 13.1% to 37.7% in the Maghreb) and by the absence of the U6 North African lineage (from 0% in Chenini-Douiret and Jerba to 28.3% in Mozabites) (Corte-Real et al. 1996; Loueslati et al. 2006; Macaulay et al. 1999).

Comparison of F_{ST} values from pairs of populations reveals that, with the exception of some pair-groups (Asni vs. Bouhria, Asni vs. Souss, Asni vs. Sened, Asni vs. Matmata, Bouhria vs. Souss, Bouhria vs. Matmata, Figuig vs. Matmata, Souss vs. Sened, Souss vs. Matmata, Souss vs. Jerba, Sened vs. Jerba), Berber samples displayed significant differences in

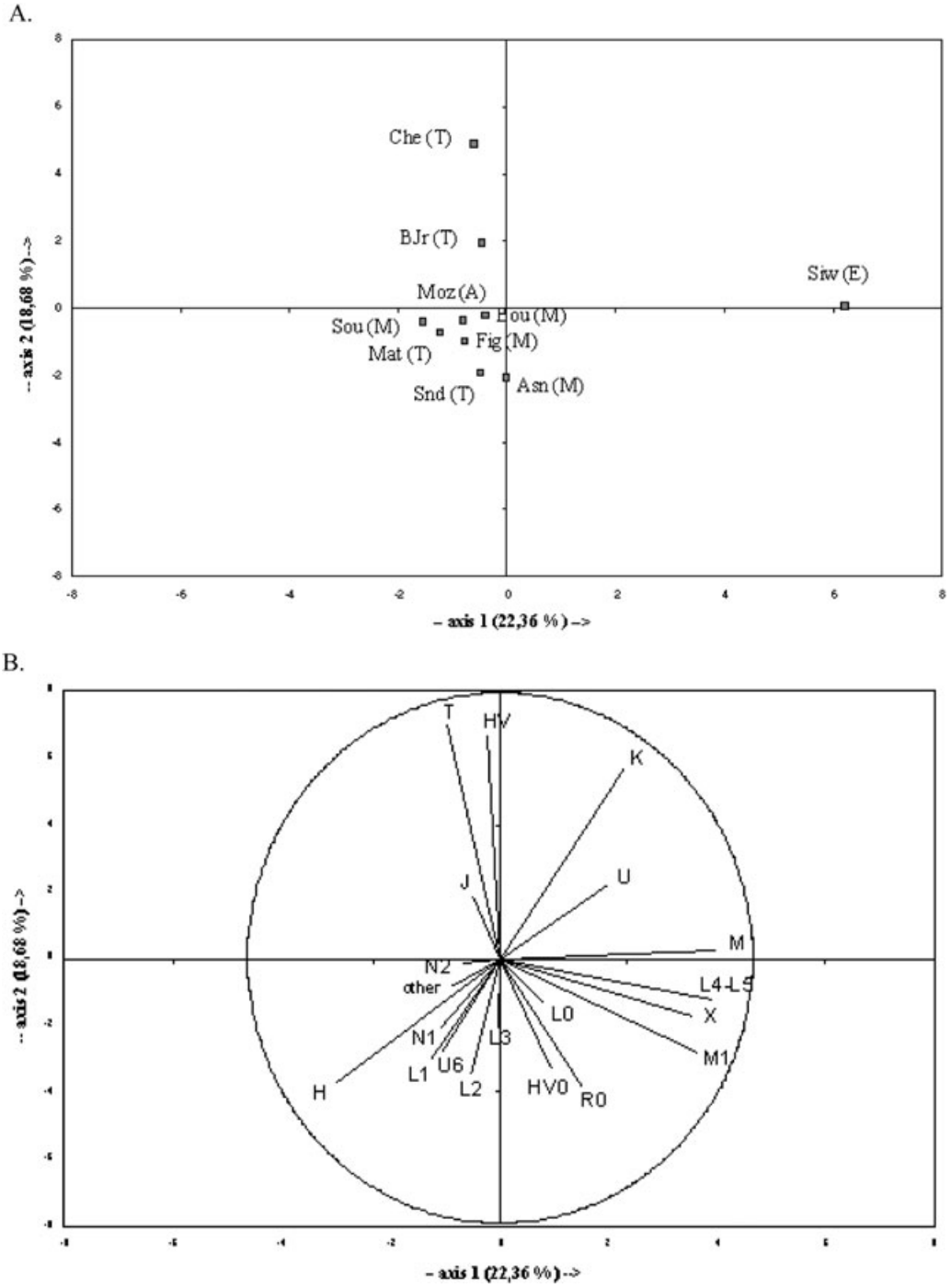

Figure 4 a) Principal Component Analysis (PCA) of mtDNA haplogroup profiles of ten Berber populations; b) Plot of the contribution of each haplogroup to the first two axes of the PCA
For population abbreviation codes, see Table 1. The "A", "E", "M" and "T" letters refer to the country where each population lives: Algeria, Egypt, Morocco and Tunisia, respectively.
Haplogroups and subhaplogroups included in the PCA are as follow: H (including H* and all H subhaplogroups), HV0 (including HV0* and V), HV (including HV* and HV1), R0 (represented mainly by R0a), J, T, U (including all U lineages except U6 and K), U6, K, N1 (including N1 and I lineages), N2 (represented mainly by W), X, M (including M* and all M subhaplogroups except M1), M1, L0, L1, L2, L3, and L4-L5.

Table 5 Frequencies of mtDNA haplogroups in the ten Berber populations included in the PCA

Population	Asni	Bouhria	Figuig	Souss	Mozabites	Chenini-Douiret	Sened	Matmata	Jerba	Siwa
Identification code	Asn	Bou	Fig	Sou	Moz	Che	Snd	Mat	BJr	Siw
Sample Size	53	70	94	50	85	53	53	49	29	78
H	0.377	0.371	0.244	0.320	0.247	0.131	0.245	0.265	0.233	0.013
HV0	0.113	0.143	0.085	0.100	0.083	–	–	0.164	0.067	0.127
HV	–	0.014	–	0.020	–	0.151	–	–	–	0.013
R0	0.038	–	–	–	–	–	0.056	–	–	0.026
J	–	–	0.117	0.100	0.035	0.038	0.038	0.040	0.167	0.051
T	–	0.100	0.032	0.040	0.047	0.321	0.038	0.041	0.134	0.013
U (without U6 and K)	0.019	0.057	0.011	0.080	0.129	0.057	0.095	0.061	0.234	0.180
U6	0.113	0.014	0.032	0.060	0.283	–	0.075	0.020	–	–
K	0.038	0.057	–	0.020	–	0.151	–	0.042	0.033	0.115
N1	0.019	0.014	–	–	–	0.019	0.075	0.083	–	–
N2	–	0.029	–	–	–	–	–	–	–	–
X	0.019	0.029	0.011	–	–	–	–	–	–	0.038
M	–	–	–	–	–	–	–	–	–	0.013
M1	0.038	0.043	0.021	–	0.047	–	0.113	0.020	–	0.167
L0	–	–	0.043	–	–	–	–	–	–	0.013
L1	0.075	0.072	0.064	0.060	–	0.037	0.056	0.020	–	0.013
L2	0.057	0.028	0.043	0.100	0.059	0.038	0.114	0.040	0.033	0.051
L3	0.075	0.029	0.297	0.100	0.070	0.057	0.095	0.184	0.099	0.128
L4-L5	0.019	–	–	–	–	–	–	–	–	0.039
other	–	–	–	–	–	–	–	0.020	–	–
Total Eurasian lineages (H, HV0, HV, R0, J, T, U (without U6), K, N1, N2, X)	0.623	0.814	0.500	0.680	0.541	0.868	0.547	0.696	0.868	0.576
Total Asian lineages (M)	–	–	–	–	–	–	–	–	–	0.013
Total sub-Saharan lineages (L0, L1, L2, L3, L4-L5)	0.226	0.129	0.447	0.260	0.129	0.132	0.265	0.244	0.132	0.244
Total North African lineages (U6, M1)	0.151	0.057	0.053	0.060	0.330	–	0.188	0.040	–	0.167

[a]For population references, see Table 1.

Table 6 Pairwise F_{ST} values between ten Berber populations
(In bold: significant F_{ST} values ($p < 0.05$). For population references, see Table 1.)

	Asni	Bouhria	Figuig	Souss	Mozabites	Chenini-Douiret	Sened	Matmata	Jerba	Siwa
Asni	–									
Bouhria	0.00065	–								
Figuig	**0.04101**	**0.05667**	–							
Souss	−0.00224	0.00434	**0.01858**	–						
Mozabites	**0.02727**	**0.05245**	**0.06937**	**0.02559**	–					
Chenini-Douiret	**0.10783**	**0.07428**	**0.10505**	**0.07625**	**0.10923**	–				
Sened	0.01324	**0.02683**	**0.03571**	0.00488	**0.02549**	**0.07463**	–			
Matmata	0.01013	0.01288	0.00857	-0.00120	**0.04306**	**0.07877**	**0.01403**	–		
Jerba	**0.05326**	**0.03461**	**0.04234**	0.00679	**0.04702**	**0.05257**	0.02324	**0.01894**	–	
Siwa	**0.09308**	**0.08826**	**0.07521**	**0.06826**	**0.08235**	**0.08946**	**0.04743**	**0.04954**	**0.04438**	–

mtDNA variance (Table 6). It also shows that the Egyptian Berbers and Tunisians from Chenini-Douiret exhibit highest F_{ST} values, confirming the clear-cut pattern of the PC plot (Fig. 4a).

Discussion

The purpose of this study was to analyze for the first time the mitochondrial gene pool of North African populations in a

broad phylogeographic context, in particular through a comparison with population groups from southern Europe, the Saharan areas and the Near East. Analysis of the 62 populations included in our haplogroup frequency database revealed that Berber populations are genetically close to southern Europeans, but significantly differentiated from sub-Saharan groups. This peculiarity is explained by their mitochondrial genetic structure, characterized by an overall high frequency of Western Eurasian haplogroups, a somehow lower frequency of sub-Saharan L lineages, and a significant (but differential) presence of North African haplogroups U6 and M1.

The genetic proximity observed between the Berbers and southern Europeans reveals that these groups shared a common ancestor. Two hypotheses are discussed: one would date these common origins in the Upper Paleolithic with the expansion of anatomically modern humans, from the Near East to both shores of the Mediterranean Sea; the other supports the Near Eastern origin, but would rather date it from the Neolithic, around 10,000 years ago (Ammerman & Cavalli-Sforza 1973; Barbujani et al. 1994; Myles et al. 2005; Rando et al. 1998). Common polymorphisms (i.e. those defining H and V lineages) between Berbers and south Europeans also could have been introduced or supported by genetic flows through the Straits of Gibraltar. For example, genetic exchanges could have taken place during prehistory, while European populations retreated from ice sheets and expanded from refuge, around 15,000 years ago (as evidenced by the H and U5b mitochondrial lineages). Alternatively, these exchanges could have occurred during history, with the invasion and the occupation during nearly seven centuries (from the 8^{th} to the 15^{th} century) of the Iberian Peninsula by Almoravide then Almohade Muslim Berber troops.

The differentiation observed between North Africans and sub-Saharan populations shows, first, that settlement of these areas was achieved by different migration waves and, then, that a genetic diversity was already observable in Africa since very old times. However, the Berber genetic heritage consists of a relatively high frequency of L lineages from various parts of Africa (i.e. L0a, L3i, L4, and L5 clades are from East Africa, L1b, L2b, and L3b are from West Africa, and L3e originated in the Sudan). It poses a question about the Sahara desert role in population movements and exchanges. It should be specified that the Sahara was not always a desert, because it also underwent enormous variation between wet and dry, offering green spaces favorable for human occupation and animal domestication (Aumassip et al. 1994; Said & Faure 1990). Thus, this is plausible that exchanges between African prehistoric populations took place; exchanges during which markers typical of sub-Saharan groups would have been introduced into the Berber gene pool. Contacts between North Africa and great sub-Saharan empires (such as those of Ghana, of Mali, or the Songhai Empire) are also reported by history during trans-Saharan trade of gold, salt and slaves.

The second analysis we conducted aimed to measure the relationship between the Berber communities on their whole geographical habitat. Although the Berber populations have the same overall mitochondrial genetic composition, the distribution of some markers is not the same along the Berber-speaker habitats. Our results highlighted a clear genetic differentiation between Berbers from the Maghreb and Egyptian Berbers. The first seems to be more related to European populations as shown by haplogroup H1 and V frequencies, whereas the latter share more affinities with East African and Nile Valley populations as indicated by the high frequency of M1 and the presence of L0a1, L3i, L4*, and L4b2 lineages. Moreover, haplogroup U6 was not observed in Siwa. Probably, such a maternal diversity between North African Berbers would have been the result of a conjunction of several geographical, prehistoric, and historic factors which guided contacts (and thus exchanges) between local populations and migrating groups. First, in addition to the geographical distance, which certainly increases the genetic distance, the geographical location of Berber populations is very peculiar: the Berbers from the Maghreb are at the end of a long migration route, whereas Berbers from Siwa are rather in a crossroads between the Middle East, East Africa, sub-Saharan areas and the North African corridor. Therefore, meetings and exchanges between local and migrating populations were not identical in North West and North East Africa. In addition, prehistory and history of the populations from Maghreb are different from those of the Egyptian group. For the Siwa oasis, there is very little information on which exact prehistoric period was the starting point of Berber culture. However we know that throughout history, the oasis was crossed by successive human groups, like pilgrims traveling to Mecca, Mediterranean tradesmen, or Sahelian slave merchants. Siwa was also repopulated by Libyan Berber-speakers driven from their land by Arab conquerors. Lastly, it experienced a period of decline and has faced, between the ninth and twelfth centuries, a drastic demographic reduction of its population (Fakhry 1973). The current gene pool of the Siwa people would therefore be the result of these various genetic exchanges which occurred in the past.

Overall, it is clear that neither the Strait of Gibraltar nor the Sahara desert appears to have forced the movements and exchanges between South Europeans, North Africans, and sub-Saharans. This interaction took place since prehistory in a human and genetic framework already diversified, and the maternal ancestors of the Berbers were able to exchange some markers with surrounding populations with different cultures and genetic pools. Influences from the Middle East and East Africa are marked in Siwa, while southwestern European influences are observed in the Maghreb. Although the origin of

these Eurasian and sub-Saharan lineages in the mitochondrial pool of Berbers is still questionable (whether due to common ancestry or past and/or current gene flow), certainly they were not diluted by the many historical invasions and migrations, leaving a clear maternal footprint in the contemporary populations.

We conclude that the origins and diversity of Berber populations are old and complex, and these communities bear genetic characteristics resulting from various events of gene flow with surrounding and migrating populations.

Acknowledgments

We are grateful to all the people and infrastructures who contributed to the collection of blood samples and to the genetic analyses. We particularly thank Evelyne Guitard, André Sevin, and the team of the Genopole Toulouse platform from Auzeville (http://genopole-toulouse.prd.fr/) for their technical help. We also make a particular homage to the linguist Naima Louali (deceased). This work was carried out within the framework of the E.C. Sixth Framework Programme under Contract no. ERAS-CT-2003-980409 (EUROCORES project of the European Science Foundation). It benefited from funding from Progetti Ricerca Interesse Nazionale 2007 (Italian Ministry of the University) (to A.T.), Compagnia di San Paolo (to A.T), CNRS (Centre National de la Recherche Scientifique) and the Conseil Régional de Midi-Pyrénées (Toulouse, France).

References

Abu-Amero, K. K., Larruga, J. M., Cabrera, V. M. & Gonzalez, A. M. (2008) Mitochondrial DNA structure in the Arabian Peninsula. *BMC Evol Biol* **8**, 45.

Achilli, A., Olivieri, A., Pala, M., Metspalu, E., Fornarino, S., Battaglia, V., Accetturo, M., Kutuev, I., Khusnutdinova, E., Pennarun, E., Cerutti, N., Di Gaetano, C., Crobu, F., Palli, D., Matullo, G., Santachiara-Benerecetti, A. S., Cavalli-Sforza, L. L., Semino, O., Villems, R., Bandelt, H. J., Piazza, A. & Torroni, A. (2007) Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. *Am J Hum Genet* **80**, 759–768.

Achilli, A., Rengo, C., Battaglia, V., Pala, M., Olivieri, A., Fornarino, S., Magri, C., Scozzari, R., Babudri, N., Santachiara-Benerecetti, A. S., Bandelt, H. J., Semino, O. & Torroni, A. (2005) Saami and Berbers–an unexpected mitochondrial DNA link. *Am J Hum Genet* **76**, 883–886.

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., Moral, P., Dugoujon, J. M., Roostalu, U., Loogvali, E. L., Kivisild, T., Bandelt, H. J., Richards, M., Villems, R., Santachiara-Benerecetti, A. S., Semino, O. & Torroni, A. (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* **75**, 910–918.

Alvarez, J. C., Johnson, D. L., Lorente, J. A., Martinez-Espin, E., Martinez-Gonzalez, L. J., Allard, M., Wilson, M. R. & Budowle, B. (2007) Characterization of human control region sequences for Spanish individuals in a forensic mtDNA data set. *Leg Med (Tokyo)* **9**, 293–304.

Ammerman, A. & Cavalli-Sforza, L. L. (1973) A population model for the diffusion of early farming in Europe. In: *The explanation of culture change: models in prehistory* (ed. C. Renfrew), pp. 348–358. London.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R. & Young, I. G. (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. & Howell, N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**, 147.

Aumassip, G., Ferhat, N., Heddouche, A. & Vernet, R. (1994) Le milieu saharien aux temps préhistoriques. In: *Milieux, hommes et techniques du Sahara préhistorique. Problèmes actuels* (eds. G. Aumassip, N. Ferhat & A. Heddouche), L'Harmattan, pp. 9–30. Paris.

Bandelt, H. J., Alves-Silva, J., Guimaraes, P. E., Santos, M. S., Brehm, A., Pereira, L., Coppa, A., Larruga, J. M., Rengo, C., Scozzari, R., Torroni, A., Prata, M. J., Amorim, A., Prado, V. F. & Pena, S. D. (2001) Phylogeography of the human mitochondrial haplogroup L3e: a snapshot of African prehistory and Atlantic slave trade. *Ann Hum Genet* **65**, 549–563.

Bandelt, H. J., Forster, P. & Rohl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**, 37–48.

Bandelt, H. J., Forster, P., Sykes, B. C. & Richards, M. B. (1995) Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753.

Barbujani, G., Pilastro, A., De Domenico, S. & Renfrew, C. (1994) Genetic variation in North Africa and Eurasia: neolithic demic diffusion vs. Paleolithic colonisation. *Am J Phys Anthropol* **95**, 137–154.

Batini, C., Coia, V., Battaggia, C., Rocha, J., Pilkington, M. M., Spedini, G., Comas, D., Destro-Bisol, G. & Calafell, F. (2007) Phylogeography of the human mitochondrial L1c haplogroup: genetic signatures of the prehistory of Central Africa. *Mol Phylogenet Evol* **43**, 635–644.

Behar, D. M., Villems, R., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., Scozzari, R., Makkan, H., Tzur, S., Comas, D., Bertranpetit, J., Quintana-Murci, L., Tyler-Smith, C., Wells, R. S. & Rosset, S. (2008) The dawn of human matrilineal diversity. *Am J Hum Genet* **82**, 1130–1140.

Bertranpetit, J., Sala, J., Calafell, F., Underhill, P. A., Moral, P. & Comas, D. (1995) Human mitochondrial DNA variation and the origin of Basques. *Ann Hum Genet* **59**, 63–81.

Brakez, Z., Bosch, E., Izaabel, H., Akhayat, O., Comas, D., Bertranpetit, J. & Calafell, F. (2001) Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area. *Ann Hum Biol* **28**, 295–307.

Brandstatter, A., Peterson, C. T., Irwin, J. A., Mpoke, S., Koech, D. K., Parson, W. & Parsons, T. J. (2004) Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database. *Int J Legal Med* **118**, 294–306.

Brett, M. & Fentress, E. (1996) *The Berbers*. Oxford: Blackwell.

Camps, G. (1980) *Les Berbères. Mémoire et identité*. Paris: Errance.

Casas, M. J., Hagelberg, E., Fregel, R., Larruga, J. M. & Gonzalez, A. M. (2006) Human mitochondrial DNA diversity in an archaeological site in al-Andalus: genetic impact of migrations

from North Africa in medieval Spain. *Am J Phys Anthropol* **131**, 539–551.

Cerny, V., Salas, A., Hajek, M., Zaloudkova, M. & Brdicka, R. (2007) A bidirectional corridor in the Sahel-Sudan belt and the distinctive features of the Chad Basin populations: a history revealed by the mitochondrial DNA genome. *Ann Hum Genet* **71**, 433–452.

Cherni, L., Loueslati, B. Y., Pereira, L., Ennafaa, H., Amorim, A. & El Gaaied, A. B. (2005) Female gene pools of Berber and Arab neighboring communities in central Tunisia: microstructure of mtDNA variation in North Africa. *Hum Biol* **77**, 61–70.

Coia, V., Destro-Bisol, G., Verginelli, F., Battaggia, C., Boschi, I., Cruciani, F., Spedini, G., Comas, D. & Calafell, F. (2005) Brief communication: mtDNA variation in North Cameroon: lack of Asian lineages and implications for back migration from Asia to sub-Saharan Africa. *Am J Phys Anthropol* **128**, 678–681.

Corte-Real, H. B., Macaulay, V. A., Richards, M. B., Hariti, G., Issad, M. S., Cambon-Thomsen, A., Papiha, S., Bertranpetit, J. & Sykes, B. C. (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann Hum Genet* **60**, 331–350.

Crespillo, M., Luque, J. A., Paredes, M., Fernandez, R., Ramirez, E. & Valverde, J. L. (2000) Mitochondrial DNA sequences for 118 individuals from northeastern Spain. *Int J Legal Med* **114**, 130–132.

Di Benedetto, G., Erguven, A., Stenico, M., Castri, L., Bertorelle, G., Togan, I. & Barbujani, G. (2001) DNA diversity and population admixture in Anatolia. *Am J Phys Anthropol* **115**, 144–156.

Dubut, V., Chollet, L., Murail, P., Cartault, F., Beraud-Colomb, E., Serre, M. & Mogentale-Profizi, N. (2004) mtDNA polymorphisms in five French groups: importance of regional sampling. *Eur J Hum Genet* **12**, 293–300.

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

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Fadhlaoui-Zid, K., Plaza, S., Calafell, F., Ben Amor, M., Comas, D. & Bennamar El Gaaied, A. (2004) Mitochondrial DNA heterogeneity in Tunisian Berbers. *Ann Hum Genet* **68**, 222–233.

Fakhry, A. (1973) *Siwa Oasis*. Cairo: The American University in Cairo Press.

Finnila, S., Lehtonen, M. S. & Majamaa, K. (2001) Phylogenetic network for European mtDNA. *Am J Hum Genet* **68**, 1475–1484.

Gonder, M. K., Mortensen, H. M., Reed, F. A., de Sousa, A. & Tishkoff, S. A. (2007) Whole-mtDNA genome sequence analysis of ancient African lineages. *Mol Biol Evol* **24**, 757–768.

Gonzalez, A. M., Brehm, A., Perez, J. A., Maca-Meyer, N., Flores, C. & Cabrera, V. M. (2003) Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am J Phys Anthropol* **120**, 391–404.

Gonzalez, A. M., Cabrera, V. M., Larruga, J. M., Tounkara, A., Noumsi, G., Thomas, B. N. & Moulds, J. M. (2006) Mitochondrial DNA variation in Mauritania and Mali and their genetic relationship to other Western Africa populations. *Ann Hum Genet* **70**, 631–657.

Gonzalez, A. M., Larruga, J. M., Abu-Amero, K. K., Shi, Y., Pestano, J. & Cabrera, V. M. (2007) Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics* **8**, 223.

Greenacre, M. (1992) Correspondence analysis in medical research. *Stat Methods Med Res* **1**, 97–117.

Herrnstadt, C., Elson, J. L., Fahy, E., Preston, G., Turnbull, D. M., Anderson, C., Ghosh, S. S., Olefsky, J. M., Beal, M. F., Davis, R. E. & Howell, N. (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* **70**, 1152–1171.

Jackson, B. A., Wilson, J. L., Kirbah, S., Sidney, S. S., Rosenberger, J., Bassie, L., Alie, J. A., McLean, D. C., Garvey, W. T. & Ely, B. (2005) Mitochondrial DNA genetic diversity among four ethnic groups in Sierra Leone. *Am J Phys Anthropol* **128**, 156–163.

Kivisild, T., Reidla, M., Metspalu, E., Rosa, A., Brehm, A., Pennarun, E., Parik, J., Geberhiwot, T., Usanga, E. & Villems, R. (2004) Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. *Am J Hum Genet* **75**, 752–770.

Krings, M., Salem, A. E., Bauer, K., Geisert, H., Malek, A. K., Chaix, L., Simon, C., Welsby, D., Di Rienzo, A., Utermann, G., Sajantila, A., Paabo, S. & Stoneking, M. (1999) mtDNA analysis of Nile River Valley populations: A genetic corridor or a barrier to migration? *Am J Hum Genet* **64**, 1166–1176.

Larruga, J. M., Diez, F., Pinto, F. M., Flores, C. & Gonzalez, A. M. (2001) Mitochondrial DNA characterisation of European isolates: the Maragatos from Spain. *Eur J Hum Genet* **9**, 708–716.

Loogvali, E. L., Roostalu, U., Malyarchuk, B. A., Derenko, M. V., Kivisild, T., Metspalu, E., Tambets, K., Reidla, M., Tolk, H. V., Parik, J., Pennarun, E., Laos, S., Lunkina, A., Golubenko, M., Barac, L., Pericic, M., Balanovsky, O. P., Gusar, V., Khusnutdinova, E. K., Stepanov, V., Puzyrev, V., Rudan, P., Balanovska, E. V., Grechanina, E., Richard, C., Moisan, J. P., Chaventre, A., Anagnou, N. P., Pappa, K. I., Michalodimitrakis, E. N., Claustres, M., Golge, M., Mikerezi, I., Usanga, E. & Villems, R. (2004) Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol Biol Evol* **21**, 2012–2021.

Loueslati, B. Y., Cherni, L., Khodjet El Khil, H., Ennafaa, H., Pereira, L., Amorim, A., Ben Ayed, F. & Ben Ammar Elgaaied, A. (2006) Islands inside an island: reproductive isolates on Jerba island. *Am J Hum Biol* **18**, 149–153.

Maca-Meyer, N., Gonzalez, A. M., Larruga, J. M., Flores, C. & Cabrera, V. M. (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet* **2**, 13.

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

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., Scozzari, R., Bonne-Tamir, B., Sykes, B. & Torroni, A. (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* **64**, 232–249.

Myles, S., Bouzekri, N., Haverfield, E., Cherkaoui, M., Dugoujon, J. M. & Ward, R. (2005) Genetic evidence in support of a shared Eurasian-North African dairying origin. *Hum Genet* **117**, 34–42.

Olivieri, A., Achilli, A., Pala, M., Battaglia, V., Fornarino, S., Al-Zahery, N., Scozzari, R., Cruciani, F., Behar, D. M., Dugoujon, J. M., Coudray, C., Santachiara-Benerecetti, A. S., Semino, O., Bandelt, H. J. & Torroni, A. (2006) The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science* **314**, 1767–1770.

Palanichamy, M. G., Sun, C., Agrawal, S., Bandelt, H. J., Kong, Q. P., Khan, F., Wang, C. Y., Chaudhuri, T. K., Palla, V. & Zhang, Y. P. (2004) Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am J Hum Genet* **75**, 966–978.

Pereira, L., Macaulay, V., Torroni, A., Scozzari, R., Prata, M. J. & Amorim, A. (2001) Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. *Ann Hum Genet* **65**, 439–458.

Pereira, L., Prata, M. J. & Amorim, A. (2000) Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Ann Hum Genet* **64**, 491–506.

Pereira, L., Richards, M., Goios, A., Alonso, A., Albarran, C., Garcia, O., Behar, D. M., Golge, M., Hatina, J., Al-Gazali, L., Bradley, D. G., Macaulay, V. & Amorim, A. (2005) High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res* **15**, 19–24.

Plaza, S., Calafell, F., Helal, A., Bouzerna, N., Lefranc, G., Bertranpetit, J. & Comas, D. (2003) Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. *Ann Hum Genet* **67**, 312–328.

Plaza, S., Salas, A., Calafell, F., Corte-Real, F., Bertranpetit, J., Carracedo, A. & Comas, D. (2004) Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. *Hum Genet* **115**, 439–447.

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.

Quintana-Murci, L., Semino, O., Bandelt, H. J., Passarino, G., McElreavey, K. & Santachiara-Benerecetti, A. S. (1999) Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet* **23**, 437–441.

Rando, J. C., Pinto, F., Gonzalez, A. M., Hernandez, M., Larruga, J. M., Cabrera, V. M. & Bandelt, H. J. (1998) Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, near-eastern, and sub-Saharan populations. *Ann Hum Genet* **62**, 531–550.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Golge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Cali, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di, R. A., Novelletto, A., Oppenheim, A., Norby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozari, R., Torroni, A. & Bandelt, H. J. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251–1276.

Richards, M. B., Macaulay, V. A., Bandelt, H. J. & Sykes, B. C. (1998) Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet* **62**, 241–260.

Roostalu, U., Kutuev, I., Loogvali, E. L., Metspalu, E., Tambets, K., Reidla, M., Khusnutdinova, E. K., Usanga, E., Kivisild, T. & Villems, R. (2007) Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Mol Biol Evol* **24**, 436–448.

Rosa, A., Brehm, A., Kivisild, T., Metspalu, E. & Villems, R. (2004) MtDNA profile of West Africa Guineans: towards a better understanding of the Senegambia region. *Ann Hum Genet* **68**, 340–352.

Rowold, D. J., Luis, J. R., Terreros, M. C. & Herrera, R. J. (2007) Mitochondrial DNA geneflow indicates preferred usage of the Levant Corridor over the Horn of Africa passageway. *J Hum Genet* **52**, 436–447.

Said, R. & Faure, H. (1990) Chronological framework: African pluvial and glacial epochs. In: *General history of Africa*. (ed. J. Ki-Zerbo), pp. 146–166. Paris: UNESCO.

Salas, A., Richards, M., De la Fe, T., Lareu, M. V., Sobrino, B., Sanchez-Diz, P., Macaulay, V. & Carracedo, A. (2002) The making of the African mtDNA landscape. *Am J Hum Genet* **71**, 1082–1111.

Salas, A., Richards, M., Lareu, M. V., Scozzari, R., Coppa, A., Torroni, A., Macaulay, V. & Carracedo, A. (2004) The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet* **74**, 454–465.

Shen, P., Lavi, T., Kivisild, T., Chou, V., Sengun, D., Gefel, D., Shpirer, I., Woolf, E., Hillel, J., Feldman, M. W. & Oefner, P. J. (2004) Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. *Hum Mutat* **24**, 248–260.

Stevanovitch, A., Gilles, A., Bouzaid, E., Kefi, R., Paris, F., Gayraud, R. P., Spadoni, J. L., El-Chenawi, F. & Beraud-Colomb, E. (2004) Mitochondrial DNA sequence diversity in a sedentary population from Egypt. *Ann Hum Genet* **68**, 23–39.

Torroni, A., Achilli, A., Macaulay, V., Richards, M. & Bandelt, H. J. (2006) Harvesting the fruit of the human mtDNA tree. *Trends Genet* **22**, 339–345.

Torroni, A., Bandelt, H. J., D'Urbano, L., Lahermo, P., Moral, P., Sellitto, D., Rengo, C., Forster, P., Savontaus, M. L., Bonne-Tamir, B. & Scozzari, R. (1998) mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet* **62**, 1137–1152.

Torroni, A., Bandelt, H. J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H. V., Tambets, K., Forster, P., Karger, B., Francalacci, P., Rudan, P., Janicijevic, B., Rickards, O., Savontaus, M. L., Huoponen, K., Laitinen, V., Koivumaki, S., Sykes, B., Hickey, E., Novelletto, A., Moral, P., Sellitto, D., Coppa, A., Al-Zaheri, N., Santachiara-Benerecetti, A. S., Semino, O. & Scozzari, R. (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet* **69**, 844–852.

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M. L. & Wallace, D. C. (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.

Trovoada, M. J., Pereira, L., Gusmao, L., Abade, A., Amorim, A. & Prata, M. J. (2004) Pattern of mtDNA variation in three populations from Sao Tome e Principe. *Ann Hum Genet* **68**, 40–54.

Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J. R., Mehdi, S. Q., Seielstad, M. T., Wells, R. S., Piazza, A., Davis, R. W., Feldman, M. W., Cavalli-Sforza, L. L. & Oefner, P. J. (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* **26**, 358–361.

Vernesi, C., Di Benedetto, G., Caramelli, D., Secchieri, E., Simoni, L., Katti, E., Malaspina, P., Novelletto, A., Marin, V. T. & Barbujani, G. (2001) Genetic characterization of the body attributed to the evangelist Luke. *Proc Natl Acad Sci U S A* **98**, 13460–13463.

Watson, E., Forster, P., Richards, M. & Bandelt, H. J. (1997) Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet* **61**, 691–704.
Weir, B. S. & Hill, W. G. (2002) Estimating F-statistics. *Annu Rev Genet* **36**, 721–750.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 MtDNA haplotypes of the four studied Berber populations (from Asni, Bouhria, Figuig, and Siwa). Variant positions from the rCRS are shown for each individual. Substitutions are transitions unless the base change or a deletion (d) is explicitly indicated. Insertions of a base and heteroplasmy are indicated by "i" and "h", respectively.

Please note: Wiley-Blackwell Publishing are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received: 31 July 2008
Accepted: 1 October 2008