6-3-2014

Phylogeography of E1b1b1b-M81 Haplogroup and Analysis of its Subclades in Morocco

Ahmed Reguig
Laboratoire de Génétique Moléculaire Humaine, Département de Recherche Scientifique, Institut Pasteur du Maroc,
ahmed.reguig@yahoo.fr

Nourdin Harich
Equipe des sciences Anthropogénétiques et Biotechnologie, Département de Biologie, Faculté des Sciences El Jadida, Université Chouaib Eddoukkali

Abdelhamid Barakat
Laboratoire de Génétique Moléculaire Humaine, Département de Recherche Scientifique, Institut Pasteur du Maroc

Hassan Rouba
Laboratoire de Génétique Moléculaire Humaine, Département de Recherche Scientifique, Institut Pasteur du Maroc

Phylogeography of E1b1b1b-M81 Haplogroup and Analysis of its Subclades in Morocco

Ahmed Reguig[1*], Nourdin Harich[2], Abdelhamid Barakat[1], and Hassan Rouba[1]

1. Laboratoire de Génétique Moléculaire Humaine, Département de Recherche Scientifique, Institut Pasteur du Maroc.

2. Equipe des sciences Anthropogénétiques et Biotechnologie, Département de Biologie, Faculté des Sciences El Jadida, Université Chouaib Eddoukkali.

*Correspondence to: Ahmed Reguig, Laboratoire de Génétique Moléculaire Humaine, Département de Recherche Scientifique, Institut Pasteur du Maroc.
E-mail: ahmed.reguig@yahoo.fr

Keywords: Y chromosome, Haplogroup, Northern Africa, Biallelic markers, Berber-speaking.

Suggested running head: Phylogeography of E1b1b1b-M81 Haplogroup

Abstract. In this work, we have analyzed a total of 295 unrelated Berber-speaking men from the northern, center and southern of Morocco, in order to characterize frequency of E1b1b1b-M81 haplogroup and to refine the phylogeny of its subclades: E1b1b1b1-M107, E1b1b1b2-M183 and E1b1b1b2a-M165. For this

purpose, we have typed four biallelic polymorphisms: M81, M107, M183 and M165. As results, a large majority of the Berber-speaking male lineages belong to the Y chromosomal E1b1b1b-M81 haplogroup. The frequency ranged from 79.1 to 98.5% in all localities sampled. Then, the E1b1b1b2-M183 was the most dominant subclade in our samples, which ranged from 65.1% to 83.1%. In contrast, the E1b1b1b1-M107 and E1b1b1b2a-M165 subclades weren't found in our samples. Our results suggest a predominance of E1b1b1b-M81 haplogroup among Moroccan Berber-speaking male with a decreasing gradient from south to north. Then, the most prevalent subclade in this haplogroup was E1b1b1b2-M183 in which difference between these three groups was statistically significant between central and southern groups.

Introduction

The tree of binary polymorphisms in the non recombining portion of the human Y chromosome allowing important insight to trace the pattern of demographic history of human populations and helping to decipher the paternal genetic relationships among human populations. According to a hierarchical phylogenetic nomenclature of the Y chromosome tree reported by Underhill et al. (2001) and up dated by the *Y Chromosome Consortium* (*YCC*) (2002) and Karafet et al. (2008). The E major-haplogroup was observed in Africa, Europe and the Near

East (Semino et al. 2004), the most widespread and frequent cluster among its clades was E1b1b1-M35, at this level of molecular resolution, the main derivative haplogroups were E1b1b1a-M78 as a home land the Eastern Africa, E1b1b1c-M123 spread in the Near Eastern Asia and E1b1b1b-M81 was the most prevalent male-lineage in North Western Africa (Semino et al. 2004). The E1b1b1b-M81 Haplogroup can be further characterized at one additional level of Unique Event Polymorphisms markers (UEPs), E1b1b1b*, E1b1b1b1 given by M107 and E1b1b1b2 given by M183, at this last level we can find E1b1b1b2* and E1b1b1b2a given by M165 (Figure 1).

Berber-speaking as ethnic group represents the ancient inhabitants of the North Western Africa. In Morocco, Berber groups speak several dialects, the three main spoken dialects are Tarifit, Tamazight and Tachelhit, respectively distributed in the Northern, Center and Southern of the country.

The aim of the present study was to give a further description of the Y-chromosomal E1b1b1b-M81 haplogroup and its subclades among Moroccan Berber-speaking groups. For this purpose, we have typed four biallelic polymorphisms M81, M107, M183 and M165 in samples of the Berber-speaking groups collected from different regions of Morocco. The second aim was to give an overview about the distribution of this lineage in Mediterranean area confronted with historical events.

Subjects and Methods

Subjects. We had analyzed samples of 295 unrelated Moroccan men, speaking Berber as a first dialect, including 43 men from Northern (Rif) speaking Tarifit, 187 men from Center speaking Tamazight and 65 men from Southern speaking Tachlhit (Figure 2). These samples were collected from individuals who reported that their four grandparents belonged to the same geographic localization. Appropriate informed consent was obtained from all participants in this study. DNA was extracted from fresh peripheral blood by standard phenol-chloroform protocols.

Biallelic Polymorphisms Typing. All biallelic polymorphisms were amplified by PCR using the primers described by Underhill et al. (2001) (table 1). The standard protocol of PCR was modified, as final volume was 25µL, the composition of the reaction mix was the same for all reactions: 400 nM of each primer, 200 µM dNTP, 1 unit of Taq DNA polymerase, 100 ng of nuclear DNA, 1.5 mmol/L of MgCl2 and 2.5 µL of 10× PCR buffer. The M81, M107 and M165 markers were analyzed by RFLP using respectively restriction enzymes: HpyCH4IV (ACG^T), BspHI (TCATG^A) and BstZ17I (GTA^TAC), electrophoresis was performed on an agarose gel (5%). The M183 was typed by direct sequencing with the automated 4-capillary ABI PRISM 3130 Genetic

Analyzer (Applied Biosystems), using the Big Dye Terminator 1.1. Comparisons of biallelic markers frequencies between groups were performed by R software, using the Chi-square test, except comparisons of M81 frequencies between Northern-Southern and Center-Southern groups were calculated by the Fisher's test. P values considered statistically significant for values equal to or less than 0.05.

Results and Discussion

E1B1B1B-M81 Haplogroup and its Phylogeography. The E1b1b1b-M81 haplogroup was most frequently observed among 295 men Berber-speaking originating from Morocco, its frequency was 79.1% among Northern men, 89.8% among men from Center and 98.5% among Southern men (Table 2). We found that the statistically significance difference was between the Northern-Southern (P-value = 0.0009961) and between Center-Southern. (P-value = 0.03051).

It seems that the E1b1b1b-M81 haplogroup is characteristic of the Berber-speaking populations and make this male lineage of North-Western Africa had a largely distinct genetic history compared to Europe, Near East and Sub-Saharan Africa. This difference might suggest an Upper Paleolithic background (Bosch et al. 2001) or a Neolithic origin (Arredi et al. 2004) of this lineage in North-Western Africa. The E1b1b1-M35 clade that includes E1b1b1b-M81, E1b1b1a-

M78 and E1b1b1c-M123 haplogroups may have occurred in Eastern Africa (Semino et al. 2004), then E1b1b1b-M81 dispersed mainly toward North-Western Africa (Luis et al. 2004). Furthermore, some socio-cultural factors can also explain the present results, like the effect of the patrilocality that men live closer to their birth place than women and the high rate of consanguinity founded among Berber populations as a consequence of marriage traditions, still persisted in isolated tribe at nowadays.

In previous studies, the E1b1b1b-M81 haplogroup was a variable frequency in North-Western Africa, ranged from 23% to 76 % in some region of Morocco (Table 3). The frequency of this haplogroup among other groups of Berber-speaking has been reported nearly 80% among Algerian Mozabite-Berbers (Cruciani et al. 2004) and 49% among Tuaregs from Libya (Ottoni et al. 2011).

However, a relatively moderate frequency of E-M81 was found among the Arabic speaking from North Africa, among Moroccans the values were 33% (Cruciani et al. 2002) and 23% (Bosch et al. 2001), among Algerians the frequency was 45.1% (Robino et al. 2008), among Tunisians the frequency was 27.6% (Semino et al. 2004) and 9.3% among Egyptians (Figure 3). In contrast, E1b1b1b-M81 was absent in the deep sub-Saharan Africa, Middle East and Northern Europe, with the exception of Southern Europe (Semino et al. 2004; Cruciani et al. 2004 and Arredi et al. 2004). Figure 3 shows the geographical

distribution of E1b1b1b-M81 haplogroup in Mediterranean area and in the immediate South of North Africa.

In the immediate South of Northern Africa, Underhill et al. (2000) found E-M81 with a frequency of 29.5% among Malians. Cruciani et al. (2004) found this haplogroup at 9.1% among Tuareg-Berbers from Niger and in the Eastern Africa at 5% among Sudanese. Also, in Middle East, E-M81 was found among Lebanese (1.2–2.4%), Turkish (2.5–8.7%), Bedouins (3.6%) and Sephardic Jews (5%) (Table 3). On the other hand, this haplogroup has been found at low frequencies among South European populations (Table 3). In the Iberian Peninsula, it occurs at low frequencies in Spain (1.6–5.3%), with the exception of Pasiegos (18%) and Cantabrian non-Lebaniego (17%) (Maca-Meyer et al. 2003). E-M81 has been reported to occur with a frequency of 5–8% in Portugal and in the two North Atlantic Archipelagos of Madeira and Açores (Gonçalves et al. 2005). In Italy, Di Gaetano et al. (2009) estimated a contribution of 6% in the Sicilian gene pool (Table 3).

This contribution at different level of the E1b1b1b-M81 in the Southern Europe suggests a recent arrival from Northern Africa, probably during the Islamic influence in (A.D. 711–1492) at different parts of the Iberian Peninsula and Sicily (Hitti 1990). This recent gene flow was supported by the absence of differentiation between Iberian and North-West African populations in the network of the Y chromosome microsatellite of E1b1b1b-M81 haplogroup, as

reported by Cruciani et al. (2004) and Semino et al. (2004). In addition, estimation of the most recent common ancestor (TMRCA) based on Y chromosome microsatellite variation observed within E1b1b1b-M81 haplogroup in North-Western Africa and Portugal was respectively 8.6 ± 2.3 ky (Semino et al. 2004) and 8.1 ± 3.2 ky (Gonçalves et al. 2005).

Historical data indicate that the two main dynasties of Berbers, Almoravids and Almohades reigned in Morocco respectively in (A.D.1056–1147) and (A.D. 1145–1248), which extended their governed territories from the south of Morocco to the Niger river, the Senegal river and also to the Iberian peninsula (Hitti 1990), this expansion of Berbers dynasties could explain this wide geographic distribution of E1b1b1b-M81 haplogroup. In addition, the extent of the Ottoman Empire to North Africa except Morocco, which lasted from 1517 to 1924 (Hitti 1990), could explain the presence of this haplogroup in the Turkish populations and let us assume that probably there had a genetic exchange between populations in their conquered territories.

Subclades of E1b1b1b-M81. In the present study, no individual belonging to the E1b1b1b1-M107 subclade was found. In contrast, the other E1b1b1b2-M183 subclade has been dominated with high frequencies in our samples: 65.1%, 67.4% and 83.1%, respectively in Northern, Center and Southern (Table 2). The statistical significant difference was noted in the distribution of E-M183

frequencies between Center-Southern groups (P-value = 0.0242). The M165 marker that define E1b1b1b2a subclade hasn't been found in our samples, whereas only one individual carrying this marker was detected by Underhill et al. (2000) in the Middle Eastern populations.

The absence of E1b1b1b1-M107 and E1b1b1b2a-M165 in our 295 samples assumes that these tow subclades were private, which means that M107 and M165 were found in only one individual/family in the world and suggest a late occurrence of these tow variants. In addition, this absence might also due to the effect of random genetic drift.

Conclusion

In conclusion, our results suggest a higher frequency of E1b1b1b-M81 haplogroup among Moroccan Berber-speaking than reported by previous studies, with a decreasing gradient from south to north. Then, only the E1b1b1b2-M183 subclade was found with a high frequency, which can be used as a second specific marker in typing North-Western Africa men. Because the powerful phylo-geographical information provided by the M81 marker, we envisage using this marker in medical genetics and forensic studies especially in Mediterranean populations.

Literature Cited

Arredi, B., S. P. Estella, P. Silvia, et al. 2004. A Predominantly Neolithic Origin for Y-chromosomal DNA Variation in North Africa. *Am. J. Hum. Genet*. 75: 338–345.

Bekada, A., R. Fregel, VM. Cabrera, et al. 2013. Introducing the Algerian Mitochondrial DNA and Y-Chromosome Profiles into the North African Landscape. *PLoS ONE*. 8(2):e56775.

Bosch, E., F. Calafell, D. Comas, et al. 2001. High-resolution Analysis of Human Y-chromosome Variation Shows a Sharp Discontinuity and Limited Gene Flow Between Northwestern Africa and the Iberian Peninsula. *Am. J. Hum. Genet.* 68: 1019–1029.

Cruciani, F., R. La Fratta, P. Santolamazza, et al. 2004. Phylogeographic Analysis of Haplogroup E3b (E-M215) y Chromosomes Reveals Multiple Migratory Events Within and Out of Africa. *Am. J. Hum. Genet*. 74: 1014–1022.

Cruciani, F., P. Santolamazza, P. Shen, et al. 2002. A Back Migration from Asia to sub-Saharan Africa Is Supported by High-resolution Analysis of Human Y-chromosome Haplotypes. *Am. J. Hum. Genet*. 70: 1197–1214.

Flores, C., N. Maca-Meyer, J. A. Pérez, et al. 2003. A Predominant European Ancestry of Paternal Lineages from Canary Islanders. *Ann. Hum. Genet*. 67: 138–152.

Gonçalves, R., A. Rosa, A. Freitas, et al. 2003. Y-chromosome Lineages in Cabo Verde Islands Witness the Diverse Geographic Origin of Its First Male Settlers *Hum. Genet*. 113 : 467–472.

Hitti, PK. (1990) The Arabs: a short history. *Gate way Editions, Washington*, DC.

Karafet, T. M., F. L. Mendez, M. B. Meilerman, et al. 2008. New Binary Polymorphisms Reshape and Increase Resolution of the Human Y Chromosomal Haplogroup Tree. *Genome Res*. 18: 830–838.

Kujanovà, M., L. Pereira, V. Fernandes, et al. 2009. Near Eastern Neolithic Genetic Input in a Small Oasis of the Egyptian Western Desert. *Am. J. Phys. Anthropol*. 140: 336–346.

Luis, J. R., D. J. Rowold, M. Regueiro, et al. 2004. The Levant Versus the Horn of Africa: Evidence for Bidirectional Corridors of Human Migrations. *Am. J. Hum. Genet*. 74: 532–544.

Maca-Meyer, N., A. M. González, J. Pestano, et al. 2003. Mitochondrial DNA Transit Between West Asia and North Africa Inferred from U6 Phylogeography. *BMC Genet*. 4: 15.

Ottoni, C., M. H. Larmuseau, N. Vanderheyden, et al. 2011. Deep into the roots of the Libyan Tuareg: A genetic survey of their paternal heritage. *Am. J. Phys. Anthropol.* 145: 118–24.

Robino, C., F. Crobu, C. Di Gaetano, et al. 2008. Analysis of Y-chromosomal SNP Haplogroups and STR Haplotypes in an Algerian Population Sample. *Int. J. Legal Med*. 122: 251–255.

Sanchez, J. J., C. Hallenberg, C. Borsting, et al. 2005. High Frequencies of Y Chromosome Lineages Characterized by E3b1, DYS19-11, DYS392-12 in Somali Males. *Eur. J. Hum. Genet*. 13: 856–866.

Semino, O., A. S. Santachiara-Benerecetti, F. Falaschi, et al.2002. Ethiopians and Khoisan Share the Deepest Clades of the Human Y-Chromosome Phylogeny. *Am. J. Hum. Genet*. 70: 265–268.

Semino, O., C. Magri, G. Benuzzi, et al. 2004. Origin, Diffusion, and Differentiation of Y-chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area. *Am. J. Hum. Genet*. 74: 1023–1034.

Underhill, P. A., G. Passarino, A. A. Lin, et al. 2001. The Phylogeography of Y Chromosome Binary Haplotypes and the Origins of Modern Human Populations. *Ann. Hum. Genet*. 65: 43–62.

Underhill, P. A., P. Shen, A. A. Lin, et al. 2000. Y Chromosome Sequence Variation and the History of Human Populations. *Nature Genetics*. 26: 358–361.

YCC. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 12:339–348.

Zalloua, P. A., Y. Xue, J. Khalife, et al. 2008. Y-chromosomal Diversity in Lebanon Is Structured by Recent Historical Events. *Am. J. Hum. Genet*. 82: 873–882.

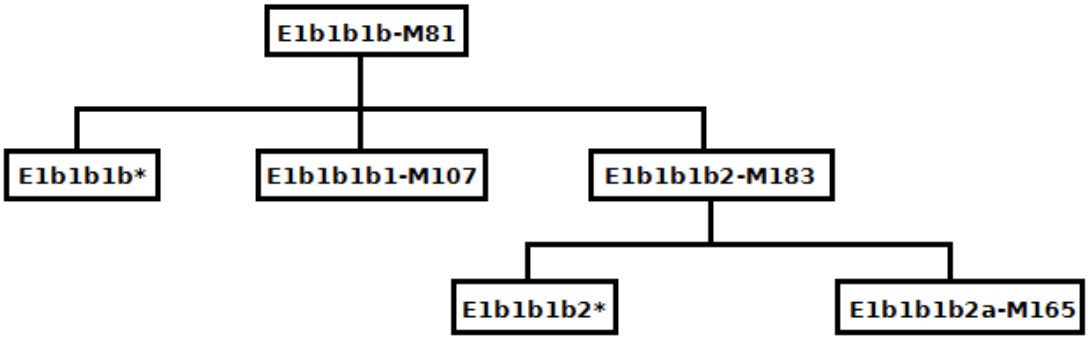

Figure 1: Phylogeny of E1B1B1B-M81 haplogroup

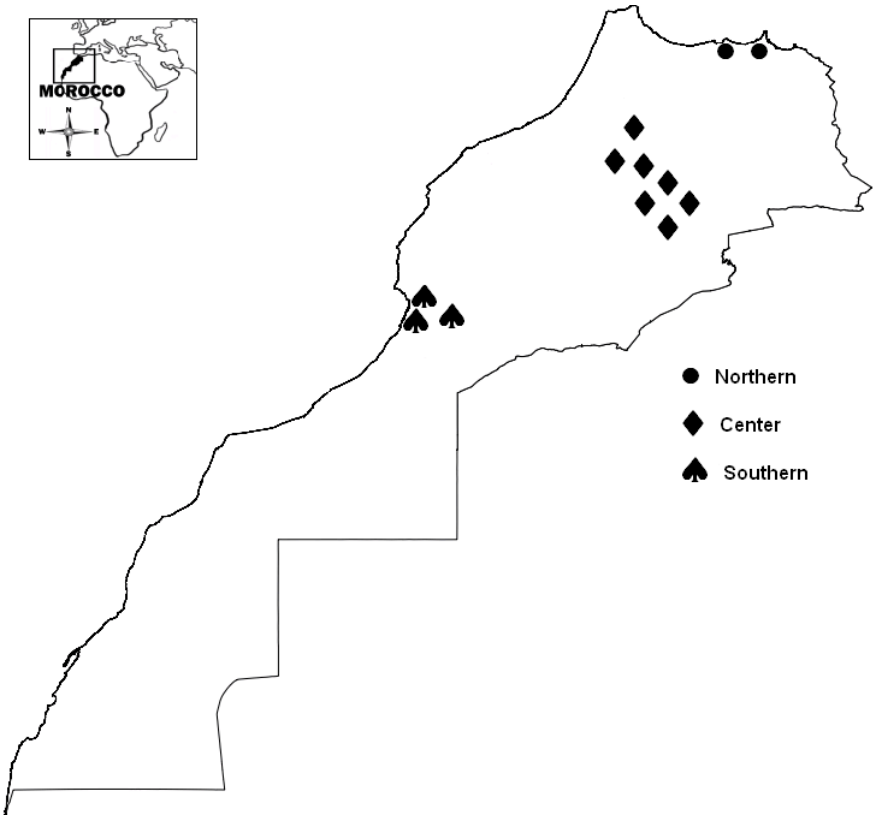

Figure 2: Geographic localization of the studied populations

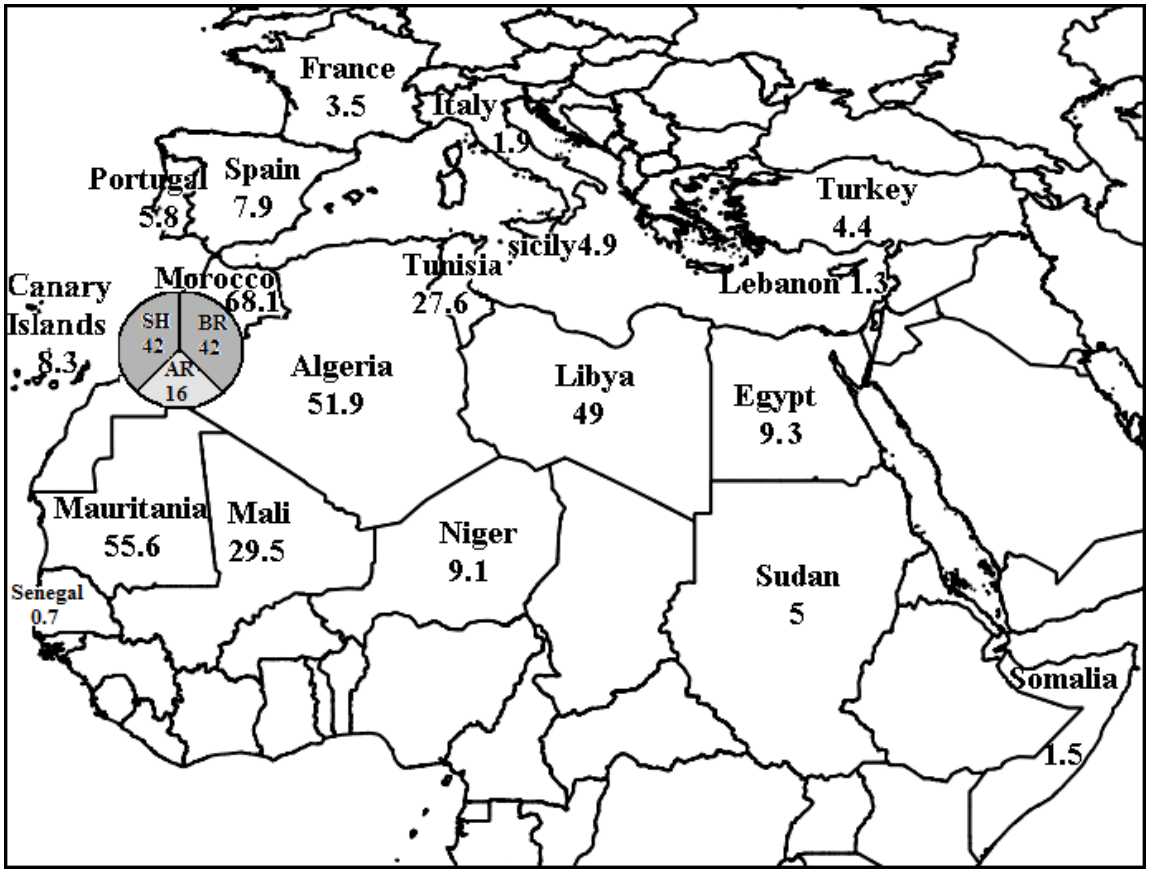

Figure 3: Geographical distribution of E1b1b1b-M81 haplogroup frequency (%) in the world and among Moroccan groups (pie chart represents: SH: Sahrouis, BR: Berber-speaking, AR: Arab-speaking)

Table 1: Description of used biallelic markers and their genotyping

Locus (Size)	Primer sequences	Discriminating Enzymes	Ancestral to derived	Position bp*
M81 (422 bp)	R : TTCATGGAGATGTCTGTATCTGG F : ACTTAATTTATAGTTTCAATCCCT	HpyCH4IV : ACG^**T**	C to T	147
M107 (376 bp)	R : CTTTACTCCCACTTATGCAACG F : CAAAAGCACTCGGGTTCCT	BspHI : TCATG^**A**	A to G	298
M183 (427bp)	R : TTCCTTTTAACCTATTATTACTTTCC F : ACTGGGTAAATATGACTATGATTGAG	Sequencing	A to C	324
M165 (340bp)	R : TTTTAGCAAGTTAAGTCACCAGC F : AAAGCGAGAGATTCAATCCAG	BstZ17I : GT**A**^TAC	A to G	132

*: Nucleotide position from 5' end of forward primer

Table 2: Frequency (%) of E1b1b1b-M81 haplogroup and its sub-clades among 295 Berber-speaking men

Studied regions	Dialect of population	E1b1b1b-M81	E1b1b1b1-M107	E1b1b1b2-M183	E1b1b1b2a-M165
Northern (N=43)	Tarifit	34(79.1)	0.0	28(65.1)	0.0
Center (N=187)	Tamazight	168(89.8)	0.0	126(67.4)	0.0
Southern (N=65)	Tachelhit	64(98.5)	0.0	54(83.1)	0.0

Table 3: Geographical distribution of E1b1b1b-M81 haplogroup in the world

Region and Population size	%	Region and Population size	%
Northern Africa		Southern Europe	
North Moroccan Berbers (N=43)*	79.1	Pasiegos spain (N=82) [I]	18
Central Moroccan Berbers (N=187)*	89.8	Cantabrians NP and NL (N=88) [I]	17
South Moroccan Berbers (N=65)*	98.1	Spanish Basques (N=103) $[D]^{+}[B]$	2.9
Moroccan Sahraouis (N=29) [A]	75.9	Southern Spaniards(N= 62) [B]	1.6
Marrakesh Berbers (N=29) [B]	72.4	Andalusians (N=113) $[A]^{+}[D]$	5.3
Moyen Atlas Berbers (N=69) [B]	71	Catalans (N=33) [D]	3
South Moroccan Berbers (N=40) [A]	26	Basques (N=44) [H]	2.3
NCMB (N=63) [A]	41	Northern Portugal (N=151) $[B]^{+}[J]$	5.3
Berbers (N=64) [C]	69	Southern Portugal (N=149) $[B]^{+}[J]$	8.1
Moroccan Arabs (N=44) [A]	23	Center of Portugal (N=102) $[B]^{+}[J]$	4.9
Moroccan Arabs (N=49) [C]	33	Madeira (N=129) $[B]^{+}[J]$	5.4
Canary Islanders (N=652) [Q]	8.3	Açores (N=121) $[B]^{+}[J]$	5
Mozabite Algerians (N=20) [B]	80	Piazza Armerina Sicily (N=28) [K]	7.1
Algerians (N=32) [D]	53.1	Sciacca Sicily (N=28) [K]	3.6
Northwest Algerians (N=102) [E]	45.1	Caccamo Sicily (N=16) [K]	6.3
Tunisians (N=58) [D]	27.6	Santa Ninfa Sicily (N=31) [K]	3.2
North Egyptians (N=21) [B]	4.8	French (N=85) [B]	3.5
Arab Egyptians (N=147) [G]	5.4	North Italians (N=67) [B]	1.5
Egyptian Western Desert (N=35) [N]	28.6	Central Italians(N=89) [B]	2.2
Mauritanians + Sahraouis (N=189) [F]	55.6	Middle East	
Tuaregs from Libya (N=47) [O]	49	Sephardim Jewishs (N=40) [D]	5
Immediate South of North Western Africa		Sephardim Turkishs (N=19) [B]	5.3
Sudaneses (N=40) [H]	5	Istanbul Turkishs (N=81) $[B]^{+}[D]$	3.7

Malians (N=44) [H]	29.5	Southwestern Turkishs (N=40) [B]	2.5
Tuaregs from Niger (N=22) [B]	9.1	Northeastern Turkishs (N=41) [B]	2.4
Somalis (N=201) [M]	1.5	Turkish Cypriots (N=46) [B]	8.7
Senegaleses (N=139) [P]	0.7	Lebaneses (N=956) $[D]^{+}[L]$	1.3
		Bedouins (N=28) [B]	3.6

NCMB: North Central Moroccan Berbers. NP: Non-Pasiego, NL: Non-Lebaniego.*: Present study. [A] Bosch et al. (2001) - [B] Cruciani et al. (2004) - [C] Cruciani et al. (2002) - [D] Semino et al. (2004) - [E] Robino et al. (2008) - [F] Bekada et al. (2013) - [G] Luis et al. (2004) - [H] Underhill et al. (2000) - [I] Maca-Meyer et al. (2003) - [J] Gonçalves et al. (2005) - [K] Di Gaetano et al. (2009) - [L] Zalloua et al. (2008) - [M] Sanchez et al. (2005) - [N] Kujanovà et al. (2009) - [O] Ottoni et al. (2011) - [P] Semino et al. (2002) - [Q] Flores et al. (2003).