Human mitochondrial DNA as a molecular tool for population studies; the case of North Morocco

Taha Rhouda[a,B], Yahya Dahmani[a], Noureddine Elmtili[b], Eduardo Ruiz-Pesini[a], Mouhamed Idaomar[b], Julio Montoya[a], Carmen Diez-Sanchez[a], Manuel J. Lopez-Perez[a]

[a] *Universidad de Zaragoza, Departamento de Bioquímica y Biología Molecular y Celular, Miguel Servet 177 Zaragoza, Spain*
[b] *Universidad Abdelmalek Essaadi, Facultad de ciencias, Departamento de Biología, BP 2121, 93002 Tetouan, Morocco*

Abstract

Mitochondrial DNA (mtDNA) variability is the most common genetic marker used for the study of population heterogeneity and human phylogenetic studies. The population of North Morocco has not been well studied. To check out this diversity and to compare it with the Spanish population, mtDNA hipervariable region (HVR-I) was sequenced and RFLP analyses were performed in a sample of 260 men coming from Oujda and Tetouan. This study is also interesting for experimental designs to study mitochondrial disorders linked to mtDNA polymorphisms. Moreover, our data can be used for forensic studies.

Keywords : mtDNA, variability, diseases, Morocco

Introduction

The human mtDNA is a circular double-stranded molecule of 16.569 bp, coding for 37 genes: 13 polypeptides of the complexes of the respiratory chain, 2 ribosomal RNAs (rRNA) and 22 transfer RNAs (tRNA). It also encompasses a non-coding region of 1122 bp, containing the majority of elements for the replication and transcription control [2,3].

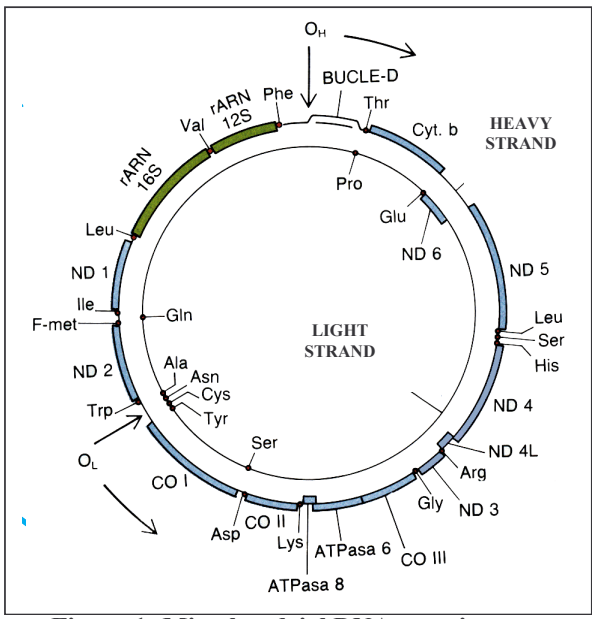

Figure 1: Mitochondrial DNA genetic map

Human mtDNA presents some particular characteristics, such as maternal inheritance and lack of recombination, that fix appearing non pathological mutations in human maternal phylogeny . Because of this, it has been used as an important tool for the study of evolution and genetics of human population. Since 1981 when human mitochondrial genome was sequenced, several studies have reported sequences of mtDNA that show individual differences, the history of women's evolution, and therefore that of the humanity.

The individual lineages of mtDNA have radiated along migrations from Africa towards the different continents during the 150,000 years of human prehistory. In this time, point mutations (single nucleotide polymorphisms or SNPs) accumulated in mtDNA appears with high frequency in the population and they are known like specific mtDNA polymorphisms of the different continents, or *mitochondrial haplogroups*.

There are four major African lineages L0, L1, L2 and L3 [10]. All non-African mtDNA sequences descended from mitochondrial cluster L3. The non-African L3 sequences belong to the M and N superclades. Caucasian population mostly harbors four clades; HV, JT, KU, and IWX. Each of these is defined by certain relatively ancient and stable polymorphic sites located in the coding region.

DNA sequence variations can be used to construct a phylogenetical tree, or several alternative trees arranged in a network, to display the evolutionary relationships between individual sequences. When the women migrated from Africa, additional mutations were accumulated becoming specific variations of discrete geographical regions, and then giving rise to the already mentioned mitochondrial haplogroups[9].

The Caucasian haplogroups derived from African superhaplogroup L3. Studies of Restriction Fragment Length Polymorphisms (RFLP), of samples of Caucasian individuals that lived in the United States

and Canada, revealed four specific haplogroups (H, I, J, K) that included 64% of the population [11]. Some early works in three Caucasian populations of Finland, Italy and Sweden, also showed other five new mitochondrial haplogroups (V, W, X, T, U). By this way, the nine Caucasian haplogroups were identified.

Africa presents the most complex genetic picture of any continent, due to the great diversity of African mtDNA sequences produced by time depth of their genetic lineages. Studies on the distribution and frequency of superhaplogroup M, concluded that the exit route of migration (Out-of-Africa) took place from East Africa towards Near East. From this point, lines of migrations to other continents arose [12].

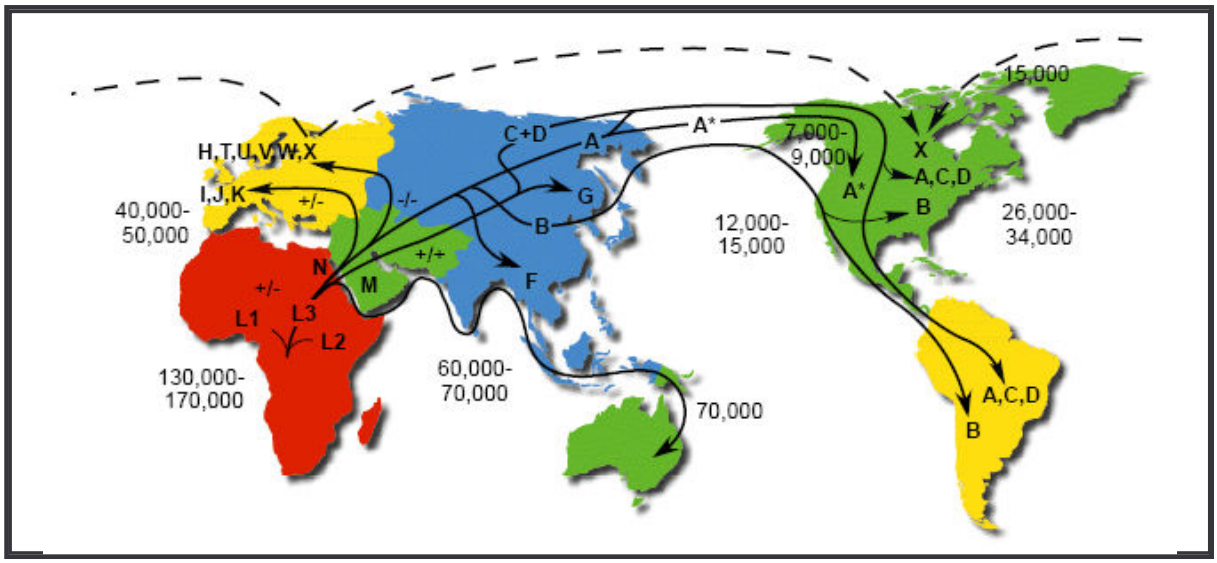

Figure 2: Mitochondrial DNA migrations from www.mitomap.com

	Brakez et al, 2001	Plaza et al, 2003	
Haplogroups	Berbers from Souss area N= 50	Arabs N= 50	Berbers N= 64
H	32	26	42
V	10	4	6,2
J	10	4	9,4
T	4	4	15,6
U6	6	8	7,8
Urest	10	12	6,2
K	2	4	7,8
X	0	4	0
L	26	32	3,2
M	0	2	0
Others	0	0	1,8

Table 1: mtDNA haplogroup distribution in Moroccan population according to [4,7].

The African mtDNAs were grouped into four haplogroups L0, L1, L2 and L3; L1 is found in West and Central sub-Saharan Africa. However, the earliest mitochondrial lineage was L0, which arose with Mitochondrial Eve. And all existing human haplogroups are descended from L1. Among these descendants are the African haplogroups L2 and L3, the latter giving rise to all non-African haplogroups.

Haplogroups A, B, C, D, E, F, G M* and R* embrace the majority of the lineages described for Asia, Oceania and Native Americans. The geographic distribution of derived branches of these haplogroups has shed light on crucial aspects of human history, such as the probable origin and approximate dating of migrations into the New World and Polynesia.

First studies on the variability of the Moroccan population, revealed a high heterogeneity of mtDNA sequences within Berber and Arab populations. Distribution of haplogroups found in the two previous studies on Moroccan population are described in table 1.

The most frequent haplogroup in both Berber and Arab populations was H. However, important differences in the contribution of haplogroup L were reported, being notably lesser in the case of Berber population studied by Plaza et al. In this last case, origin of the samples was not clearly stated. The rest of Caucasian haplogroups (V,J,T,U6, Urest, K and X), appear in different proportions. In this respect, differences between Arabs and Berbers are important for haplogroup X, which is distributed in Arabs but absent in Berbers in both studies. Interestingly, this absence was also noted in M haplogroup.

U6 haplogroup is almost absent in Caucasian population but was well represented in these reports. Then, these data confirmed that North Africa

population is enriched to U6 haplogroup. Moreover, other works reporting the Berber repartition in adjacent countries, as Algeria, Tunis, and Mauritani, show that U6 haplogroup is well represented there.
Considering that Morocco is an African peripheral region in close proximity to Europe, the African influence represented by L haplogroup is smaller than expected, and reflects an intracontinental heterogeneity. The geographical proximity to the Iberian Peninsula and possible human contacts or even movements between the two parts of the strait, makes North of Morocco an interesting area to study the relationship between South of Europe and North of Africa. Table 2, presents the distribution of mtDNA haplogroups in Spain and Andalusia, the nearest European land to North Morocco.
Haplogroup H originated in Near East though its frequency in Spanish and Caucasian populations (40-50%) is higher than in Near East. These frequencies are also higher than those found in Morocco. However, no differences were observed in its distribution between whole Spain and Andalusia.

Haplogroups	Spain N=686	Andalusia N= 158
H	47,1	46,2
V	5	5,7
J	8,9	7
T	6,9	4,4
U	22,6	12
K	—	6,3
X	1,2	3,2
L	—	1,9
M	—	3,8
Others	8,3	9,5

Table 2 : Mitochondrial haplogroups diversity in Spain and Andalusia [5,7].

To complete this puzzle and in order to reveal further intra-population differences, we decided to characterize the North Moroccan population. To collect samples, we have chosen Tetouan and Oujda, two North Moroccan cities situated respectively in the West and East part of the Rifain region, known to be enriched in Berber population.

Material and methods

Blood samples were collected, DNA was extracted and genetic analyses were performed. Blood was obtained from 260 unrelated healthy volunteers between 20 and 45 years old, including 80 from Tetouan and 180 from Oujda. Total DNA was extracted from blood (3-5 ml in EDTA tube) using Proteinase K, followed by standard phenol-chloroform method, and the DNA was precipitated using absolute ethanol, as was previously described [6,8].
For the genetic characterization of the individuals, all the samples were analyzed by RFLP and by sequencing the hipervariable region I (HVR-I). Amplifications were carried out in 50µl of reaction mixture and PCR products were loaded in 2% agarose gel. Positive amplifications were digested at 37˚C in 10µl reaction volumes. The polymorphisms used and PCR conditions are shown in table 3 as it has been previously described [8]. A fragment of 468 bp (HVS-I) was amplified using the primers L15977 (5'-CCACCATTAGCACCCAAAGC-3') and H16455 (5'-CGAGGAGAGTAGCACTCTTG-3'), where "L" and "H" refer respectively to "Light" and "Heavy" strands of mtDNA. The sequences were aligned to the revised Cambridge Reference Sequence (rCRS) using BLAST program. The results obtained are shown in table 4.

Results and discussion

In our study 58,1 % of the North Moroccan mtDNA haplogroups belong to Caucasian variants. As commented before, haplogroup H, is less represented than in Caucasian populations; though is the most frequent in both cases. On the other hand, L mtDNA genotypes found in our samples are less represented than in Souss area (13,5% in North Morocco and 26% in Souss area) and it is almost absent in the Spanish population. Nevertheless, it has been found at low frequencies in Andalusia, suggesting a slight but coherent South-North decline from rest of Spain, Andalusia, North Morocco and Souss.
Haplogroup X appears in our studies. Probably due to the higher size of our samples as compared to those described in table 1. Interestingly, haplogroup M shows a higher frequency in Oujda than in Tetouan. This result could suggest a different penetrance of this genetic character from East to West part of North Morocco. Whether this different penetrance is due to human migrations to North Africa from Near-East or Eastern Africa, where M haplogroup is somehow frequent, remains to be investigated.
It is well known that mtDNA haplogroups contain other genetic sub-lineages or sub-haplogroups. Characterization of these sub-lineages allows a deeper knowledge of the phylogenetical structure of a population, and then facilitates demographic studies. Since our results open new questions on the demographic structure of North Morocco, we decided to deep in the genetic analysis and to study some sub-haplogroups of our samples.

Table 3: Conditions used for identification of polymorphisms. mtDNA fragments amplified under the PCR conditions here defined, were digested for RFLP analysis. One of the primers used for the detection of 12308 polymorphism have a mismatch.

Amplicon	Rstriction enzyme	Haplogroup identified	Primers 5'- - - - - - - - - - - - - - - - 3'	Annealing temperature (C)	Digested PCR product -- (bp) +	
3350-3680	HpaI	L	TCGCAATGGCATTCCTAATG GAGTTTGATGCTCACCCTGA	62	331	243/88
4096-4407	Afl III	J-T	CTACTTCTAACCTCCCTGTT CTTACTTTAGGATGGGGTGT	60	312	121/191
4308-4739	Nla III	V	GGAGCTTAAACCCCCTTA GGAGCTTAAACCCCCTTA	60	432	273/159
6874-7134	Alu I	H	TCGCCACACTCCACGGAAG TGGCGTAGGTTTGGTCTAGG	65	183/78	152/31/78
10270-10579	Alu I	M	TCCTTTTACCCCTACCATGAG ATTATTCCTTCTAGGCATAGTAG	62	310	128/182
12101-12338	Hinf I	K-U	TCCCTCAACCCCGACATCATTACCG CTTTTATTTGGAGTTGCACCAAGATT	64	67/168	67/138/30
14430-14580	Acc I	X	ATGCCTCAGGATACTCCTCAATAGCCGTC TTGATTGTTAGCGGTGTGGT	60	151	36/115

Table 4: Comparison of the distribution (number of individuals between brackets)

Haplogroups	**Tetouan** n=80	**Oujda** n=180	**North of Morocco N=260**
H	31,3(25)	27,7(50)	28,8
V	6,3(5)	2,7(5)	3,8
J	1,3(1)	5,5(10)	4,2
T	7,5(6)	4,4(8)	5,3
U_{res}	5(4)	4,4(8)	4,6
U_5	3,7(3)	2,7(5)	3
U_6	11,3(9)	9,4(17)	10
K	3,7(3)	6,1(11)	5,4
X	3,7(3)	2,7(5)	3
L	11,3(9)	14,4(26)	13,5
M	3,7(3)	5(9)	4,6
Others	11,3(9)	14,4(26)	13,8

On the other hand, mtDNA variants-associated diseases are in some cases depending of these sub-haplogroups. Since haplogroup H was the most frequent in North Morocco and different sub-haplogroups H have been clearly defined, we have initiated a study to further characterize these haplogroups in our samples. Preliminary results obtained are shown in table 5. The criteria of sub-haplogrouping have been those described by [1].

Very interestingly, a different genetic structure appears in Tetouan and Oujda. H1 is the most frequent one in Spain and also appears to be the most penetrated in both areas. However, H3 that is more frequent in Italy than in Spain, shows a higher frequency in Oujda. On the other hand, H5 more frequent in Spain than in Italy, appears in Tetouan and it is absent in Oujda. These preliminary results point out to a different genetic structure of haplogroup H that remains to be confirmed and explained.

However, these differences in sub-haplogroups H, together with those of M and U6 frequencies, suggest that genetic structure of the maternally inherited mtDNA lineages show some differences between Tetouan and Oujda. These two cities, no more than 500 Km far are separated by Riff Mountains, a natural geographic barrier. We suggest that this

barrier could explain the differences found between both populations.

H sub-haplogroups	Tetouan N=25	Oujda N=50
H1	11 (44%)	24 (48%)
H2	1 (4%)	0
H3	1 (4%)	7 (14%)
H5	2 (8%)	0
H8	0	0
H*	10 (40%)	19 (38%)

Table 5: Distribution of H sub-haplogroups in Tetouan and Oujda

Acknowledgment

We express our appreciation to the blood donors who participated in the present study. We especially thank S. Morales for his technical assistance. This work was supported by grants from Agencia Española de Cooperación Internacional (AECI 145-03-P), the Diputación General de Aragón (Grupos Consolidados B33), Spanish Fondo de Investigación Sanitaria (FIS-04-0009 and 05/0647), and Networks for Mitochondrial Disorders and Ataxias (G03-011/G03-056).

References

[1]Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogvali EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, 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

[2]Anderson.S, Bankier A.T, Barrell B.G, De Bruijn M.H., Coulson A.R. and Drouin J, et al., (1981). Sequence and organization of the human mitochondrial genome, Nature ; **290**, 457–465.

[3]Andrews R.M, Kubacka I, Chinnery P.F, Lightowlers R.N, Turnbull D.M and Howell N. (1999). Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. Nat. Genet ; **23**, p. 147.

[4]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.

[5]Dahmani Y, Marcuello A, Montiel-sosa FJ, Montoya J, Diez-sanchez C, López-pérez MJ, Ruiz-pesini, E. (2006) Mitochondrial lineales distribution in the Spanish population: anticipating association studies. An. R. Acad. Nac. Farm ; **72**: 37-47

[6]Marcuello A, Gonzalez-Alonso J, Calbet JA, Damsgaard R, Lopez-Perez MJ, Diez-Sanchez C. (2005). Skeletal muscle mitochondrial DNA content in exercising humans. J Appl Physiol; **99**(4):1372-7.

[7]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**(Pt 4):312-28.

[8]Rhouda T, Dahmani Y, Elmtili N, Ruiz-pesini E, Idaomar M, Montoya J, Diez-sanchez C, López-pérez MJ. (2005). Mitochondrial genetic variability of North Morocco population. Moroccan Journal of Biology – In press

[9]Richards M, Rengo C, Cruciani F, Gratrix F, Wilson J.F, Scozzari R, Macaulay V, Torroni A. (2003). Extensive female-mediated gene flow from sub-Saharan Africa into near eastern Arab populations. Am. J. Hum. Genet ; **72**, pp. 1058–1064

[10]Salas, A., M. Richards, T. De la Fe, M.-V. Lareu, B. Sobrino, P. Sanchez-Diz, V. Macaulay, A. Carrac edo. 2002. The making of the African mtDNA landscape. Am. J. Hum. Genet ; **71**:1082–1111

[11]Torroni A, M. T. Lott, et al. (1994). "mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region." Am J Hum Genet ; **55**(4): 760-76.

[12]Wallace, D. C. (1995). "1994 William Allan Award Address. Mitochondrial DNA variation in human evolution, degenerative disease, and aging." Am J Hum Genet ; **57**(2): 201-23