International Journal of Modern Anthropology
Int. J. Mod. Anthrop. (2014) 7: 100 - 111
Available online at: www.ata.org.tn ; doi: http://dx.doi.org/10.4314/ijma.v1i7.5

Review synthesis Report

Synthetic review on the genetic relatedness between North Africa and Arabia deduced from paternal lineage distributions

Soumaya Triki-Fendri, Ahmed Rebai

Dr Ahmed Rebai is professor and researcher (group leader) at the Centre of Biotechnology of Sfax. He completed his PhD in Quantitative genetics from the Institut National Agronomique of Paris, France in 1995. He has a long research and teaching experience in genetics, biostatistics and bioinformatics. His expertise includes genetics, genetic anthropology, population genetics and genetic diversity of human populations.

Centre of Biotechnology of Sfax, 3018 Sfax Tunisia; email: ahmed.rebai@cbs.rnrt.tn

Abstract -Y-chromosome single-nucleotide polymorphisms are highly used for phylogenetic construction and in the study of human migration patterns and evolution. Knowing that these genetic markers are associated with certain aspects of human culture like languages, it has been reported that some specific haplogroups characterize the Arab world. In this review, we draw the main conclusions referring to these polymorphisms in the Arab world, in order to provide an anthropological approach to the analysis of the genetic landscape of these populations. In the Middle East, the predominant categories of Y chromosomes are varieties associated with haplogroup J-M304. It has been hypothesized that the center of origin of sub-haplogroup J1-M267 would be the southern Arabian Peninsula whereas J2-M172 seems to be originating from the Fertile Crescent region. In North African populations, the distribution of E-M81, the most common haplogroup there, closely matches the present area of Berber-speaking population's allocation on the continent, suggesting a close haplogroup-ethnic group parallelism. Remarkably, J1-M267 and J2-M172 were also encountered in North African populations but with less frequency than the Middle Eastern ones, showing an important paternal gene flow from the Middle East towards North Africa. This is in agreement with historical data such as the Phoenician migration flows, from the Fertile Crescent, and mainly to the Arab expansion during the spread of Islam and the important migration of Arabic tribes like the Hilalians that led to a large scale Arabization of North Africa.

***Key words**:* Y-SNP, haplogroup, Arab-speaking populations, North Africa, Middle East

Introduction

Arabic ranks fifth in the world’s league table of languages, with an estimated 300 million native Arabic speakers, comprising 3.8 % of the world’s population, and it is the official lingua franca of 27 countries from the Maghreb (North Africa) to the Middle East (Procházka, 2006; Versteegh, 1997). The standard definition of the Arab world comprises the 22 countries and territories stretching from the Atlantic Ocean in the west to the Arabian Sea in the East, and from the Mediterranean Sea in the North to the Horn of Africa and the Indian Ocean in the southeast. Historically, many important events have occurred in the Arab word leading to the shaping of the modern populations inhabiting the region. For instance, The Middle East, essentially inhabited by Arabs, is well known for its important role in human history, particularly as a theatre for great historical events that changed the face of the world during the Neolithic period (Al-Zahery et al., 2011). Moreover, North Africa which was originally inhabited mainly by Berbers is currently mixed with some peoples originating from different civilizations, such as Phoenicians and Romans. However, the Berber language was not seriously threatened until the Islamic Arabs expanded their religion and culture to the Maghreb, since the end of the 7th Century onwards. Widely superseded by Arabic, Berber dialects are today confined to the more mountainous and desert rural areas of the region.

Given the geographic scope of the Arab world added to the important historical events occurring there, several population genetic studies using many types of genetic markers were undertaken in different Arab populations providing insight into the structure and relationships among Arab-Speaking populations. In this review, we aimed to gather the main conclusions referring to the non-recombining uniparental Y-chromosome DNA polymorphisms in this region, in order to provide an anthropological approach in the analysis of the genetic landscape there.

Y-SNPs: useful tools in population genetics

Y-chromosome single-nucleotide polymorphisms (Y-SNPs) are the most used markers for phylogenetic construction because they are biallelic, relatively stable (Xue et al., 2009), and easy to genotype (van Oven et al., 2011; 2013). The Y-SNPs have been widely used in the study of human migration patterns and evolution. Moreover, a wealth of studies addressing Y-SNP variation conducted in the last two decades enabled the construction of a robust and detailed phylogenetic tree whose current version is based on a core set of 417 branch-defining Y-SNPs (van Oven et al., 2014). Interestingly, some bi-allelic markers of the Y chromosome not only have geographically defined distributions, but are also associated with certain facets of human culture like languages (Cavalli-Sforza et al., 1997; Cruciani et al., 20010) and practice of pastoralism (Henn et al., 2008), all of which contributing to the phenomenon of genetic drift, probably the most single key element in shaping population genetic structures (Chiaroni et al., 2009; Gebremeskel et al., 2014).

Major haplogroups observed in the Arab-Speaking populations

Haplogroup J

Haplogoup J has been considered to represent the signature of the Neolithic demic diffusion associated with the spread of agriculture (Semino et al., 1996). It can be subdivided into two major clades, J1 and J2, characterized by the markers M267 and M172, respectively, plus the paragroup J*(xJ1, J2). The predominant categories of Y chromosomes in the Middle East are varieties associated with haplogroup J-M304 (Cinnioğlu et al., 2004). The frequency distribution of haplogroup J has shown a radial decline from the Middle East toward Central Asia, Caucasus, North Africa and Europe (Quintana-Murci et al., 2001; Cadenas et al., 2008; Al-Zahery et al., 2011).

Although both sub-haplogroups J1-M267 and J2-M172 have evolved *in situ* and have contributed to the Neolithic revolution, their opposite latitudinal gradients in the Middle East suggest two separate histories. While J2-M172 has been linked to the development and the expansion of agriculture in the wetter northern zone and is also considered the Y-chromosome marker for the spread of farming into South East Europe, J1-M267 has been associated with pastoralism in the semi-arid area of the Arabian Peninsula (Chiaroni et al., 2008; 2010; Al-Zahery et al., 2011). In fact, an inverse correlation between J1-M267 frequency and mean annual rainfall in the Middle East populations has been demonstrated (Chiaroni et al., 2008). This finding was interpreted as a founder effect associated with small groups of Neolithic herder–hunters moving into the arid regions of the Arabian Peninsula with a pastoral economy, whereas another ancestral population with a closely associated sister clade, J2a-M410, remained mainly in the regions of the Fertile Crescent that had sufficient rainfall to support a Neolithic farming economy. In fact, a decreasing frequency of haplogroup J1-M267 moving from Southern Arabia northwards has been revealed: Yemen (72.6 %), Qatar (58.3 %), Iraq (56.4 %), Kuwait (45 %), Oman (38 %), Egypt (20 %), Lebanon (12.5 %) and Turkey (9 %) (Cadenas et al., 2008; Luis et al., 2004; Al-Zahery et al., 2011; Triki-Fendri et al., submitted manuscript; Semino et al., 2000; Cinnioğlu et al., 2004).

Thus, it has been hypothesized that the place of origin of this haplogroup J1-M267 would be the southern Arabian Peninsula. In fact, the distribution of this haplogroup agrees with the first major expansions proposed from GM marker analysis (Chaabani, 2002) and that of mtDNA haplogroup R0a (Cerny et al., 2011) providing evidence for a deep genetic root in southern Arabia (Chaabani, 2014). Moreover, J1-M267 is also remarkably widespread in the Arab-speaking populations of North Africa such as in Libya (27.4 %), Tunisia (20.7 %), Algeria (22.5 %) and Morocco (7.5 %) (Triki-Fendri et al., submitted; Fadhlaoui-Zid et al., 2011; Robino et al., 2008, Arredi et al., 2004).

Hence, the high frequencies of J1-M267 among the Arabic-speaking populations of the Middle East added to its low variance in North Africa compared to that in the Middle East (Semino et al., 2004) has been related to an important paternal gene flow from the Middle East, and particularly from the Arabian Peninsula towards North Africa (Semino et al., 2004; Cadenas et al., 2008; Zalloua et al., 2008; Tofanelli et al., 2009). This result is in agreement with historical data such as the Arab expansion during the spread of Islam in the 7th Century CE (Semino et al., 2004; Luis et al., 2004; Di Giacomo et al., 2004; Arredi et al., 2004; Cadenas et al., 2008; Zalloua et al., 2008) as well as the important migration flows, during the 11th Century of Arabic tribes like the Hilalians, estimated as high as 200.000 families, leading to the Arabization of North Africa by imposing their social organization, values, and language (Metz, 1987; Camps, 1996).

Similarly, the second sub-haplogroup J2-M172 is widespread in the Middle East particularly in the Levant [Iraq (43.6 %); Turkey (41 %) Lebanon (26 %), Iran (16 %); Saudi Arabia (16 %) and Kuwait (10 %) (Al-zahery et al., 2011; Cinnioglu et al., 2004; Zalloua et al., 2008; Regueiro et al., 2006; Abu-Amero et al., 2009; Triki-Fendri et al., submitted manuscript)]. Actually, the Fertile Crescent region has been considered as the most probable origin of the earliest dispersions of this sub-haplogroup. Besides, a considerable decrease in the frequencies of J2-M172 was detected in Pakistan between the South (18.7 %) and the North (4.7 %) (Sengupta et al., 2006). In North Africa, J2 was also detected in smaller frequencies compared to that in the Middle East (Libya 7 % (Triki-Fendri et al., submitted manuscript). The presence of this haplogroup in North Africa could be probably due to migration flow of the Phoenicians, from the Fertile Crescent, who settled in North Africa by the 5th Century BCE.

Haplogroup E

Haplogroup E is the most frequent haplogroup in Africa, but is also found in the Middle East, southern Europe and Asia (Cruciani et al., 2002; Semino et al., 2004; Karafet et al., 2008). Among its sub-clades, E-M81 and E-M78 seem to be of North African origin with Paleolithic and Neolithic expansions that reached surrounding areas (Arredi et al., 2004; Cruciani et al., 2007).

Firstly, E-M81 is the most common haplogroup in North Africa showing its highest concentrations in Northwestern Africa (76 % in Saharawis in Morocco (Arredi et al., 2004)) with cline frequencies decreasing eastward: Algeria (45 %), Libya (34 %) and Egypt (10 %) (Robino et al., 2008; Triki-Fendri et al., submitted; Arredi et al., 2004). Besides, Ottoni et al., (2011) have reported that E-M81 appear to constitute a common paternal genetic matrix in the Tuareg populations where it was encountered at high frequency (89 %). Hence, the distribution of this haplogroup in Africa closely matches the present area of Berber-speaking population’s allocation on the continent, suggesting a close haplogroup-ethnic group parallelism (Bosch et al., 2001; Cruciani et al., 2002; 2004; Arredi et al., 2004; Fadhlaoui-Zid et al., 2011; Bekada et al., 2013). However, knowing that the Berber dialects have been replaced by Arabic in North African populations, carriers of E-M81 haplogroup are currently Arab-speaking peoples whose ancestors were Berber-speaking.

Outside of Africa, E-M81 is almost absent in the Middle East and in Europe (with the exception of Iberia and Sicily). The presence of E-M81 in the Iberian Peninsula (12 % in southern Portugal) (Cruciani et al., 2004) has been attributed to trans-Mediterranean contacts linked to the Islamic influence, since it is typically Berber (Bosch et al., 2001; Semino et al., 2004; Beleza et al., 2006; Alvarez et al., 2009; Cruciani et al., 2007; Trombetta et al., 2011).

Secondly, E-M78 is widespread in both eastern and northeastern Africa like in Somalia (77.6 %), Sudan (25.6 %), Ethiopia (22.7 %), Libya (17.7 %) and Egypt (18 %) (Underhill et al., 2000; Sanchez et al., 2005; Hassan et al., 2008; Luis et al., 2004; Triki-Fendri et al., submitted). However, this haplogroup is rare in northwestern Africa and in the Middle East (2 % in Algeria; 3.4 % in Kuwait (Arredi et al., 2004; Robino et al., 2008; Triki-Fendri et al., submitted)). According to Cruciani et al., (2007), the peripheral geographic distribution of the most derived sub-haplogroups with respect to northeastern Africa, as well as the results of quantitative analysis of UEP and microsatellite diversity are strongly suggestive of a northeastern rather than an eastern African origin of E-M78. Northeastern Africa thus, seems to be the place from where E-M78 chromosomes started to disperse to other African regions and outside Africa.

Haplogroup R

More than 50 % of the human males in Europe are affiliated to haplogroup R, mainly to the sub-haplogroup R1-M173 (Jobling and Tyler-Smith, 2003; Chiaroni et al., 2009). It is likely that M173 arose initially in Central Asia, and those subpopulations carrying M173 migrated westward into Europe soon thereafter. Haplogroup R was also observed in the Arab-speaking populations, particularly in the Middle East: 19.4 % in Iraq, 9.4 % in Kuwait; and 11.2 % in the Southern Arabian Peninsula as well as in North Africa: 7 % in Libya; 12.7 % in Algeria (Triki-Fendri et al., submitted; Al-Zahery et al. 2011; Cadenas et al., 2008; Robino et al., 2008).

In fact, Between the Umayyad and the Abbasid Empires, many slaves were brought from Iberia, especially Cordoba and Andalusia (Bassaam, 1977). Furthermore, envoys of the British Empire to the Persian Gulf and the Middle East between the early 19th and early 20th Centuries CE, mention in their memoirs the enslaving of Georgians, Armenians and Circassians (Philby, 1923). These slaves were naturalized and given tribal affiliations under slavery abolishment treaties signed between the tribal chieftains—or emirs—and the envoys of the British Empire. This could be a reason for the presence of haplogroup R in the Middle East (Mohammad et al., 2009).

The presence of the European male lineages R in North Africa could be the male counterpart of the maternal gene flow signaled by the mtDNA haplogroups H1, H3 and HV0. In fact, there are several haplogroups with clear geographical origins from

European or North African sides of the Mediterranean, but also present on the opposite side highlighting the presence of gene flow between both areas (Bekada et al., 2013). Moreover, it is well known that North Africa was colonized during the 19th and 20th Centuries by France, the United Kingdom, Spain and Italy. These European envoys could have brought the typically European Lineage R-M173 to North Africa. Furthermore, phylogenetic evidence and coalescence time estimates suggest that R-P25 chromosomes may have been carried to Africa by an Asia-to-Africa mid-Holocene back migrations (Cruciani et al., 2010). It has been proposed that the origin of the V88-derived allele (belonging to haplogroup R1b1-P25) has occurred in Central-West Africa and then dispersed, across the Sahara to North Africa (Gonzalez et al., 2013).

Global overview of the genetic landscape in the Arab world

The gene pool of each geographic region is, obviously, influenced by the historical events occurring there. On the one hand, the populations of the Middle East, particularly those of Arabia, are highly characterized by the typically Arabic haplogroup J1-M267 whose highest frequency was observed in Yemen, in southern Arabia. This haplogroup was also encountered in North Africa but with lower variance, reinforcing the hypothesis of gene flow from Arabia towards North Africa which is supported by historical important migration flows like the expansion of Islam and the migration of Hilalians. Furthermore, the Fertile Crescent is known to be the most probable origin of J2-M172 which was then spread into different regions like North Africa, probably, during the migration flow of the Phoenicians, from the Fertile Crescent towards North Africa by the 5th Century BCE. On the other hand, it is well established currently, that the most frequent haplogroup in North African populations is E-M81, which is typically Berber. And this is not strange as the Berbers are the original population inhabiting the region. However, the genetic pool of the modern populations living there is highly modeled by that of populations issued from different civilizations such as the Phoenicians (J2-M172), the Arabs (J1-M267) and different regions such as Europe (R-M173) and Africa (R-v88).

Figure 1 summarizes the information using a Principal Component Analysis performed on four major and discriminative haplogroups: E-M78, E-M81, J-M304 (xM172) and J-M172. We see that Arab-speaking populations are discriminated in three major groups; populations of the Gulf region and Middle East (Lebanon, Iraq and Palestine) are located on the right of component 1 (highly correlated to J-

M304(xM172), *r*=0.82) indicating that they have high frequency of J-M304(xM172); while component 2 (highly correlated positively to E-M78 and negatively to E-M81) separates western-north African populations (Libya, Tunisia, Algeria and Morocco) characterized by high frequency of E-M81 from geographically intermediate populations between North Africa and Arabia (Egypt, Sudan, Ethiopia and Somalia) having high frequency of E-M78.

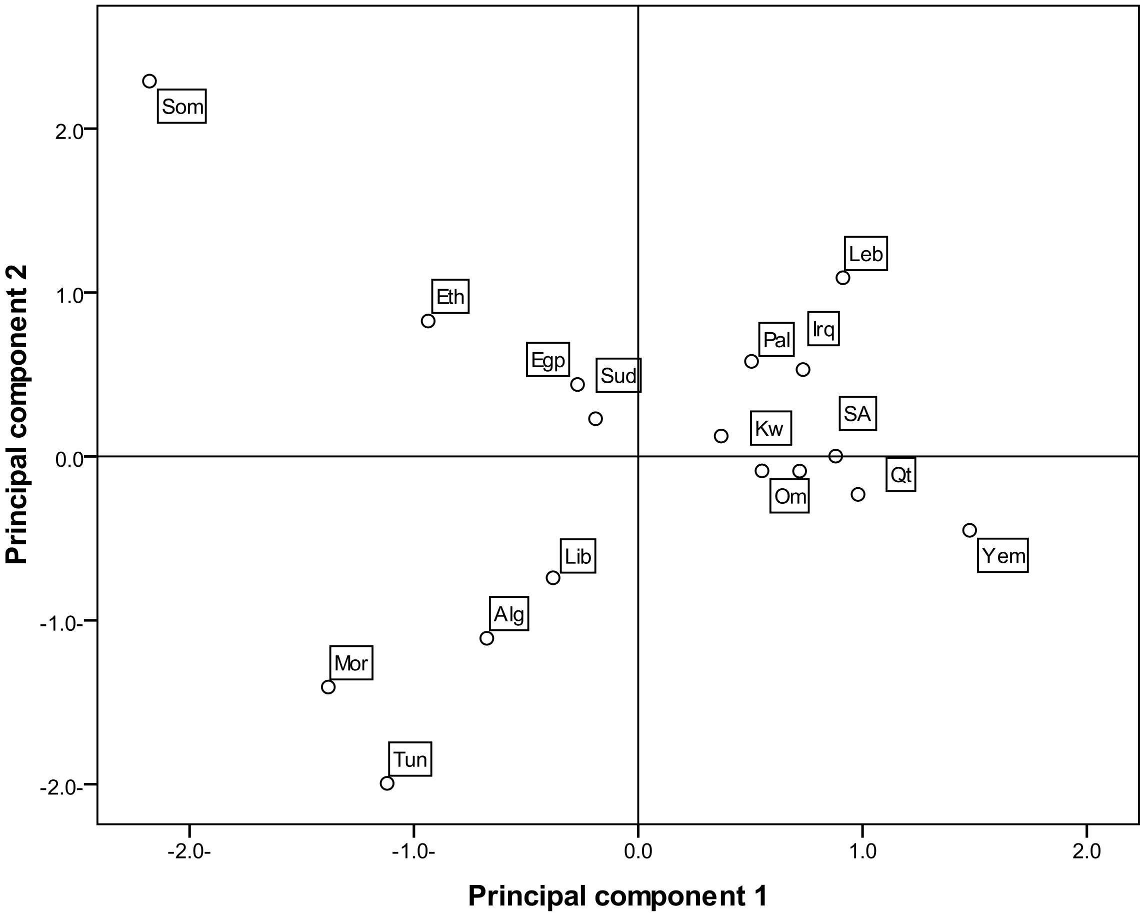

Figure 1. Mapping of populations on the plane of the two first axes from Principal Component Analysis (PCA) using frequencies of four major haplogroups (E-M78, E-M81, J-M304 (xM172), J-M172).

(Tun: Tunisia, Lib: Libya, Alg: Algeria, Mor: Morocco, Egp: Egypt, Sud: Sudan, Eth: Ethiopia, Som: Somalia, Om: Oman, Yem: Yemen, Qt: Qatar, Kw: Kuwait, SA: Saudi Arabia, Pal: Palestine, Irq: Iraq, Leb: Lebanon).

References

Abu-Amero KK., Hellani A., Gonzalez AM., Larruga JM., Cabrera VM., Underhill PA., 2009. Saudi Arabian Y-Chromosome diversity and its relationship with nearby regions, *BMC Genet* 10:59-68.

Alvarez L., Santos C., Montiel R., Caeiro B., Baali A., *et al.,* 2009. Y chromosome variation in South Iberia: insights into the North African contribution, *Am J Hum Biol* 21:407-409.

Al-Zahery N., Pala M., Battaglia V., Grugni V., Hamod MA., *et al.*, 2011. In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. *BMC Evol Biol* 11:288.

Arredi B., Poloni ES., Paracchini S., Zerjal T., Fathallah DM., *et al.* 2004. A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet* 75:338–345.

Bassaam AA., 1977. Ulamaa Najd Khilaal Sitat Quroon. Muktabat Al-Nahdha Al-Hadeethah; Damascus.

Bekada A., Fregel R., Cabrera VM., Larruga JM., Pestano J., *et al.,* 2013. Introducing the Algerian mitochondrial DNA and Y-chromosome profiles into the North African landscape. *PLoS One*. 8:e56775.

Beleza S., Gusmao L., Lopes A., Alves C., Gomes I., *et al.,* 2006 Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann Hum Genet* 70:181-194.

Bosch E., Calafell F., Comas D., Oefner PJ., Underhill PA., Bertranpetit J. 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.

Cadenas AM., Zhivotovsky LA., Cavalli-Sforza LL., Underhill PA., Herrera RJ., 2008. Y-chromosome diversity characterizes the Gulf of Oman. *Eur J Hum Genet* 16:374–386.

Camps G., 1996. The Berbers, Edisud, France.

Cavalli-Sforza LL., 1997. Genes, peoples, and languages. *Proc Natl Acad Sci USA*; 94: 7719–7724.

Cerny V., Mulligan CJ., Fernandes V., Silva NM., Alshamali F., *et al.,* 2011. Internal diversification of mitochondrial haplogroup R0a reveals post-last glacial maximum demographic expansions in South Arabia. *Mol Biol Evol* 28:71-78.

Chaabani H., 2002. GM polymorphism and the evolutionary history of modern humans. *Annales de Génétique* (current *European Journal of Medical Genetics*) 45: 197-206.

Chaabani H., 2014. Recent out of Yemen: new version of the theory of unique and recent origin of modern man. *Int. J. Mod. Anthrop* 7:13–41.

Chiaroni J., King RJ., Myres NM., Henn BM., Ducourneau A. *et al.,* 2010. The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. *Eur J Hum Genet* 18:348-353.

Chiaroni J., King RJ., and Underhill PA., 2008. Correlation of annual precipitation with human Y-chromosome diversity and the emergence of Neolithic agricultural and pastoral economies in the Fertile Crescent. *Antiquity* 82:281–289.

Chiaroni J., Underhill PA., Cavalli-Sforza LL., 2009. Y chromosome diversity, human expansion, drift, and cultural evolution. *Proc Natl Acad Sci USA* 106: 20174–20179.

Cinnioğlu C., King R., Kivisild T., Kalfoğlu E., Atasoy S., *et al.,* 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:127-148.

Cruciani F., La Fratta R., Santolamazza P., Sellitto D., Pascone R., 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., La Fratta R., Trombetta B., Santolamazza P., Sellitto D., *et al.* 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. *Mol Biol Evol* 24:1300–1311.

Cruciani F., Santolamazza P., Shen P., Macaulay V., Moral P., *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.

Cruciani F., Trombetta B., Sellitto D., Massaia A., Destro-Bisol G., *et al.,* 2010. Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. *Eur J Hum Genet* 18: 800–807.

Di Giacomo F., Luca F., Popa LO., *et al.,* 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum Genet* 115: 357–371.

Fadhlaoui-Zid K., Martinez-Cruz B., Khodjet-el-khil H., Mendizabal I., Benammar-Elgaaied A., Comas D., 2011. Genetic structure of Tunisian ethnic groups revealed by paternal lineages. *Am J Phys Anthropol* 146:271-280.

Gebremeskel EI., Ibrahim ME., 2014. Y-chromosome E haplogroups: their distribution and implication to the origin of Afro-Asiatic languages and pastoralism. *Eur J Hum Genet.* 22:1387-1392.

González M., Gomes V., López-Parra AM., Amorim A., Carracedo A., *et al.,* The genetic landscape of Equatorial Guinea and the origin and migration routes of the Y chromosome haplogroup R-V88. *Eur J Hum Genet.* 21:324-331.

Hassan HY., Underhill PA., Cavalli-Sforza LL., Ibrahim ME., 2008. Y-chromosome variation among Sudanese: restricted gene flow, concordance with language, geography, and history. *Am J Phys Anthropol* 137:316-323.

Henn BM., Gignoux C., Lin AA., Oefner PJ., Shen P., *et al.*, 2008. Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. *Proc Natl Acad Sci USA* 105: 10693–10698.

Jobling MA., Tyler-Smith C., 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 4:598–612.

Karafet TM., Mendez FL., Meilerman MB., Underhill PA., Zegura SL., Hammer MF., 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830–838.

Luis JR., Rowold DJ., Regueiro M., Caeiro B., Cinnioğlu C., *et al.*, 2004. The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am J Hum Genet.* 74:532-44.

Metz HC., 1987. Libya: A Country Study. Washington: GPO for the Library of Congress.

Mohammad T., Xue Y., Evison M., Tyler-Smith C., 2009. Genetic structure of nomadic Bedouin from Kuwait, *Heredity* (Edinb) 103:425-433.

Ottoni C., Larmuseau MH., Vanderheyden N., Martinez-Labarga C., Primativo G., *et al.*, 2011. Deep into the roots of the Libyan Tuareg: a genetic survey of their paternal heritage. *Am J Phys Anthropol* 145:118-124.

Philby HSJB., 1923. The heart of Arabia: a record of travel & exploration. New York: Putnam GP and Son.

Procházka S., 2006. Arabic. Encyclopedia of language and linguistics.

Quintana-Murci L., Krausz C., McElreavey K., 2001. The human Y chromosome: function, evolution and disease. *Forensic Sci Int* 118:169-181.

Regueiro M., Cadenas AM., Gayden T., Underhill PA., Herrera RJ., 2006. Iran: Tricontinental nexus for Y-driven migration. *Hum Hered* 61:132–143.

Robino C., Crobu F., Di Gaetano C., Bekada A., Benhamamouch S., 2008. Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample. *Int J Legal Med* 122:251-255.

Sanchez J., Hallenberg C., Borsting C., Hernandez A., Morling N., 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., Magri C., Benuzzi G., Lin AA., Al-Zahery N., *et al.*, 2004. Origin, diffusion,and differentiation of Y-chromosome haplogroups E and J: inferences on theneolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023-1034.

Semino O., Passarino G., Brega A., Fellous M., Santachiara-Benerecetti AS., 1996. A view of the neolithicdemic diffusion in Europe through two Y chromosome-specific markers. *Am J Hum Genet* 59:964–968.

Semino O., Passarino G., Oefner PJ., Lin AA., Arbuzova S. *et al.*, 2000. The genetic legacy of Paleolithic Homo sapiens in extant Europeans: a Y chromosome perspective. *Science* 290: 1155-1159.

Sengupta S., Zhivotovsky LA., King R., Mehdi SQ., Edmonds CA., 2006. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am J Hum Genet* 78:202-221.

Tofanelli S., Ferri G., Bulayeva K., Caciagli L., Onofri V., *et al.*, 2009. J1-M267 Y lineage marks climate-driven pre-historical human displacements. *Eur J Hum Genet* 17:1520–1524.

Triki-Fendri S., Sánchez-Diz P., Rey-González D., Alfadhli S., Ayadi I., *et al.*, Genetic structure of Kuwait revealed by paternal lineages. *Submitted manuscript.*

Triki-Fendri S., Sánchez-Diz P., Rey-González D., Ayadi I, Carracedo A, Rebai A., Paternal lineages in Libya inferred from Y-chromosome haplogroups, *submitted manuscript.*

Trombetta B., Cruciani F., Sellitto D., Scozzari R., 2011. A new topology of the human Y chromosome haplogroup E1b1 (E-P2) revealed through the use of newly characterized binary polymorphisms. *PLoS One* 6:e16073.

Underhill PA., Shen P., Lin AA., Jin L., Passarino G., 2000. Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358-361.

Van Oven M., Ralf A., Kayser M., 2011. An efficient multiplex genotyping approach for detecting the major worldwide human Y-chromosome haplogroups. *Int J Legal Med* 125:879–885.

Van Oven M., Toscani K., van den Tempel N., Ralf A., Kayser M., 2013. Multiplex genotyping assays for fine-resolution subtyping of the major human Y-chromosome haplogroups E, G, I, J and R in anthropological, genealogical, and forensic investigations. *Electrophoresis* 34:3029–3038.

Van Oven M., Van Geystelen A., Kayser M., Decorte R., Larmuseau MH., 2014. Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. *Hum Mutat* 35:187-191.

Versteegh, K., 1997. The Arabic language. Edinburgh: Edinburgh University Press.

Xue Y., Wang Q., Long Q., Ng BL., Swerdlow H., *et al.*, 2009. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol* 19:1453–1457.

Zalloua PA., Xue Y., Khalife J., *et al.*, Y-chromosomal diversity in Lebanon is structured by recent historical events. *Am J Hum Genet* 82: 873–882.