Genetic structure of North African human populations

A complex history of admixture

Lara Rubio Araúna

TESI DOCTORAL UPF / 2017

DIRECTOR DE LA TESI

Dr. David Comas

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

upf. Universitat Pompeu Fabra *Barcelona*

"There is no barrier, lock or bolt that you can impose on the freedom of my mind "

Virginia Woolf

A Room of One's Own

Agradecimientos

"A woman must have money and a room of her own if she is to write fiction."

Virginia Woolf

Virginia Woolf expuso de manera excelente en "A Room of One's Own" porqué una mujer necesita dinero y una habitación propia para escribir ficción. Espero que el tema y las conclusiones de esta tesis no resulten ser ficción, pero nos servirá de metáfora.

Quiero agradecer a todas las personas que de una manera u otra me han aportado los recursos económicos y académicos necesarios, llamémosle dinero, para desarrollar y llegar a concluir esta etapa. Gracias a David por ayudarme con esta parte y por apoyarme con paciencia. Y gracias a mis compañeros por crecer juntos entre poblaciones.

Sobre todo, quiero agradecer a todos los que me habéis ayudado a crear un espacio de pensamiento dónde aprender y desarrollar conocimiento. Gracias por ayudarme a construir mis propias ideas, mi habitación propia.

En la misma obra Virginia Woolf afirma que uno no puede pensar bien, amar bien, dormir bien, si no ha comido bien. No podría estar más de acuerdo, y ahora no es metafórico, gracias a los que cada día me dais y me habéis dado de comer para que pueda seguir intentando hacer las cosas bien.

Abstract

North African populations show a complex genetic structure characterized by the admixture of at least four different components: Middle Eastern, sub-Saharan, European and autochthonous North African. However, there are few genome-wide studies focused on North Africa and only two Berber groups have been included in those analyses. In this thesis genotype array data that increases the number of available Berber samples is introduced. This data shows a heterogenous genetic structure of North African populations, including Berbers, and a lack of genetic differentiation between Berber and Arab groups. Admixture is the main process shaping North African diversity. In the results of this thesis different admixture events are described, mainly related to sub-Saharan gene-flow and Middle Eastern expansions. Furthermore, North African gene-flow into coastal surrounding populations is analyzed, showing a recent historical North African contribution from different geographical places in the European coast and the Canary Islands populations.

Resumen

Las poblaciones del norte de África presentan una estructura genética compleja caracterizada por la mezcla genética de al menos cuatro componentes: Oriente Medio, África subsahariana, Europa y autóctono del norte de África. Sin embargo, pocos estudios a escala genómica se centran en el norte de África y sólo dos grupos de bereberes han sido incluidos en los análisis. En esta tesis se presentan datos genotipados con chip que aumentan el número de muestras disponibles de bereberes. Estos datos genéticos muestran una estructura heterogénea de las poblaciones del norte de África, incluyendo los bereberes, y una falta de diferenciación genética entre grupos bereberes y árabes. La mezcla genética es el principal proceso que moldea la diversidad norte africana. En los resultados de esta tesis se describen diferentes procesos de mezcla, principalmente relacionados con flujo genético de África subsahariana y expansiones de Oriente Medio. Además, también se analiza el flujo genético desde el norte de África a poblaciones costeras circundantes, mostrando una contribución durante la historia reciente de zonas geográficas diferentes en la costa europea y en las Islas Canarias.

Preface

Different types of data contribute to an accurate and exhaustive reconstruction of human history. Although historical written records are many times biased and incomplete, they are a good guide and a valuable source of hypothesis. Archaeology and anthropology have produced a high amount of data that has solved plenty of questions regarding human origins and evolution. Linguistic evidence has also helped in the understanding of humans, particularly in the movement of people and the relation between the populations. In the last decades, genetic studies have joint efforts with these fields and have contributed extensively to the human history knowledge.

The large number of samples for which genetic data is currently available has allow the study of many different human populations. Genetically closeness of populations is explained by a similar history. In this sense, genetically close populations may be related because of a common ancestral origin or because of extensive gene-flow between them. Taken into account this principle and the information available, a detailed and rigorous reconstruction of human history and evolution has been possible at a worldwide scale. Genetic studies have given support and extended the knowledge about the African human origin, the out-of-Africa dispersal, and the colonization of new territories. The human population genetics field is growing rapidly and shedding light to human history. Particularly, ancient DNA studies are increasing insight into the first stages of human evolution and dispersal.

The focus of this thesis is the North African human populations. In this thesis, the genetic structure of those populations has been analyzed using genome-wide data and novel genomic methods. A special emphasis has been put in the admixture of the North African individuals with other populations. In this regard, the main objectives have been the identification of external genetic contribution to North Africa and the dating of the gene-flow arrival. Finally, the role of North Africa in the gene-pool of surrounding populations has been tested, showing an important contribution.

INDEX

"Tenere ti nin

Tenere nekkim

Taqqim tafuk tedu

Teyyed fel shinawan ashni

Itissed ehad hi tismesti"

"The desert is mine

Tenere, my homeland

We come to you when the sun goes down

Leaving a trail of blood across the sky

Which the black night wipes out"


+IO:I (Tinariwen), *Tenere Taqqim Tossam*

1 INTRODUCTION

1.1 THE PEOPLING OF NORTH AFRICA: FROM THE HUMAN ORIGINS TO THE PRESENT

1.1.1 The African cradle

All fields of study agree in an origin of modern humans in the East of sub-Saharan Africa around 200,000 years ago. Although it is controversial to determine morphologically whether the oldest fossils are anatomically modern humans (AMH) (Schwartz and Tattersall, 2010; Stringer, 2003), the fossil record places the earliest well-dated anatomically modern humans in southern Ethiopia (Omo I and Omo II) (McDougall et al., 2008) around 195,000 years ago (ya). However, more recent remains are also found in Ethiopia (Herto, 160,000 ya) supporting the East African origin of our species (White et al., 2003). Other ancient remains come from South Africa around 100,000 ya (Klasies River) (Rightmire et al., 2006) or Levant (Qafaeh and Skhul, 120,000 ya) (Vallois and Vandermeersch, 1975). The African recent origin of modern humans has been demonstrated using a large amount of genetic markers that have also signaled the larger effective population size of African groups (Armour et al., 1996; Cann et al., 1987; Ingman et al., 2000; Jorde et al., 1997; Li et al., 2008; Stoneking et al., 1997; Tishkoff et al., 1998, 2009, 1996; Vigilant et al., 1991). These two facts have been proposed to explain the high genetic diversity observed in African populations, especially in hunter-gatherer groups (Henn et al., 2011; Pickrell et al., 2012; Schlebusch et al., 2012; Tishkoff et al., 2009).However, ancient lineages of uniparental markers are present in both hunter-gatherers and non-hunter-gatherer groups (Batini et al., 2011; Mendez et al., 2013) contributing to the controversy of the geographic origin of modern humans and first population splits within Africa. A recent paper analyzing hunter-gatherer ancient samples (2,000 ya) from South Africa has pushed back the estimation of the divergence of modern humans at least 260,000 ya (Schlebusch et al., 2017). Furthermore, linguistics have also supported evidence from an African origin (Cysouw et al., 2012).

1.1.2 The Green Sahara

The Sahara Desert traces the southern limit of North Africa and has played a major role for the region. Although nowadays the Sahara represents rather a barrier, palaeoclimatology and archaeology have evidenced humid periods when the Sahara was a biological corridor. During the last Interglacial period, between 130,000 and 117,000 years ago, the Sahara was dotted of watercourses that flew towards the Mediterranean (Coulthard et al., 2013; Drake et al., 2011). At least three major river systems were present: Irharhar, crossing the territory currently covered by Algeria; and Sahabi and Kufrah, in Eastern Maghreb. The Irharhar watercourse has been proposed as the most probable route for hominins dispersion and Middle Stone Age artifacts have been found in its basin (Coulthard et al., 2013). As a result of this wide range of corridors, the industries that appeared in the Nile Valley are different from other North African assemblies (Osborne et al., 2008). The last wet period of the Sahara was during the Early Holocene, ~11,000 to 8,000 ya, when the Sahara was covered by a dense paleoriver network and even large paleolakes. Despite the important Nile corridor, those trans-Saharan corridors were more effective, which allowed the expansion of many
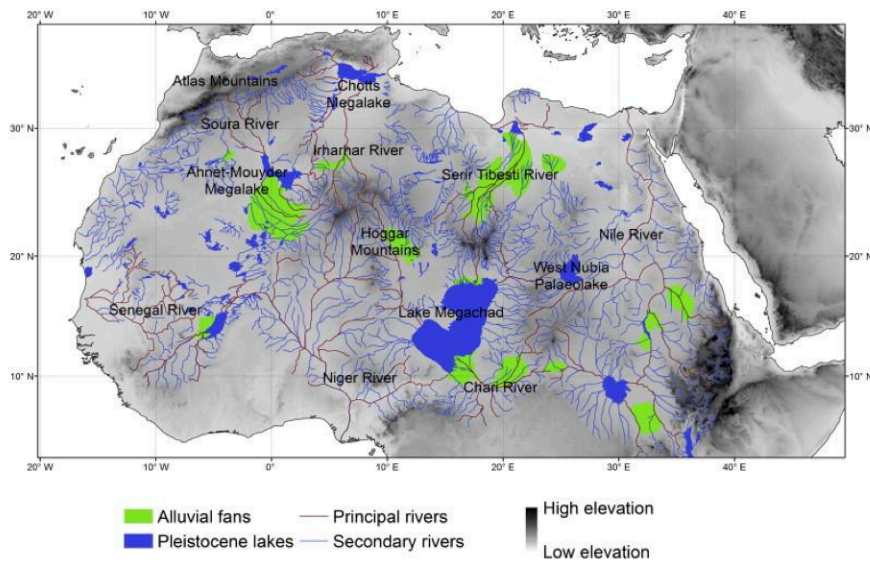
Fig 1. Topography of North Africa with the paleohydrology of MIS 5 (~130,000 to 75,000 ya) overlain. The hydrological systems that form connections throughout the Sahara are shown in brown, in order to highlight them. The alluvial fans that in some cases allow interconnections between river systems are also displayed. (Scerri et al., 2014)

4

different species (Drake et al., 2011).

In accordance to palaeoclimatic evidences, *Homo* fossils have been found in North Africa, for example *Homo erectus* remains are found in Sidi Abderrahman (Moroccan Eastern coast) and Ternifine (Algerian coast) archaeological sites (Newman, 1995). Very recently, the cataloging as AMH of Jebel Irhoud fossils (in the East of Morocco) and their dating 314,000±34,000 ya, has challenged the knowledge about *Homo sapiens* origins (Hublin et al., 2017). During some time, studies have pointed to a higher implication of North Africa in early human prehistory than previously thought (Balter, 2011; Stringer and Barton, 2008). This latest study highlights the need of including North Africa in the early steps of *Homo sapiens* and building a broader African perspective that takes into account the geoclimatic knowledge of the region.

The dating of Middle Stone Age sites in North Africa has been controversial and has changed with the irruption of new methodologies (Balter, 2006; Jacobs and Roberts, 2007). Before the last results of Hublin and colleagues (2017), the oldest *Homo sapiens* fossils dated were also from Jebel Irhoud. This site has classically produced the oldest dates of North African *Homo sapiens*. However, the first estimation dates of the fossil remains were around 40,000 ya and the remains were related to Neanderthals (Ennouchi, 1962). Nevertheless, in 2007 the juvenile Irhoud 3 sample was dated 160,000 ya and the Neanderthal association was discarded, placing the remains closer to *Homo sapiens* than to other *Homo* (Smith et al., 2007).

During the Middle Stone Age the Aterian culture stands out in North Africa. Previous to the Aterian, the Mousterian was extended in other parts of the globe; nevertheless, this culture is not well described in North Africa. The Mousterian was first described in Eurasia and then it was extrapolated to North Africa; however, the Mousterian in North Africa has been recently proposed to be closer to Middle Stone Age in Africa than to European Middle Paleolithic. Moreover, the differences between the Aterian and the Mousterian cultures have been questioned due to the technological similarities, and a recent study has considered them as variants of the same entity (Dibble et al., 2013). As a whole, there is a poor definition and scanty studies of the pre-Aterian culture and Early Stone Age in North Africa (Bouzouggar and Barton, 2012; Garcea and Giraudi, 2006). Due to dating problems, the Aterian was for a long time

dated between 40,000-50,000 ya; however, many studies in the last two decades have shifted backwards those dates until reaching an oldest date for the Aterian in the Atlantic Morocco 110,000 ya. The timing, the spread, and the characteristics of the Aterian culture have linked it to the dispersal of AMH (Barton et al., 2009). Although the Aterian is a very flexible technology that presents variations, some characteristics are remarkable including technological features (pedunculates, tanged pieces, bifacial foliates, small Llevallois, discoidal cores, blades and bladalets) and more interestingly symbolic artifacts, as stone-limed hearths, shell beads for ornamentation and red ochre (Barton et al., 2009; Bouzouggar et al., 2007).

The symbolic features represent abstract thinking, which is an essential characteristic of modern humans. Some of these cultural novelties seem to have developed earlier in Northwest Africa than in East or South Africa, reinforcing the need of a whole African perspective for understanding the human origins. For example, shell ornaments have been dated 82,000 ya (Bouzouggar et al., 2007) and even as far as 150,000 ya (d'Errico et al., 2009); however, there is a gap of production between 70,000 to 40,000 ya, showing that discontinuities in cultural transmission occurred 70,000 ya probably as a result of climatic and environmental changes after the last Interglacial period (Bouzouggar and Barton, 2012; Bouzouggar et al., 2007; d'Errico et al., 2009). This evidence could be important for the understanding of human behavior and evolution, and therefeore, among
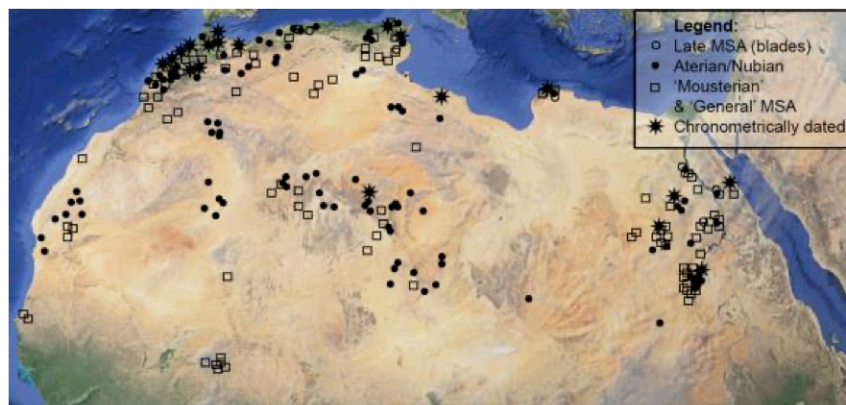
Fig 2. Map of North Africa Middle Stone Age sites, marking key chronometrically dated sites and sites assigned to particular industries in the literature. Dense concentrations can be observed in the Maghreb, and to a lesser degree, the Nile Valley, illustrating where research has been most intensive. (Scerri, 2017)

other North African evidences, it should be taken into account for the study of early humans.

The transition between the Middle Stone Age to the Later Stone Age (between 40,000 to 20,000 ya) has been poorly studied in North Africa. The Late Stone Age is characterized by the use of microlithic technology and the appearance of the Iberomaurusian culture. The beginning of this technology has been dated around 25,000 ya in Eastern Algeria, which places the disappearance of the Middle Stone Age during the Heninrich Event 2, a climatological event that produced the cooling of the Atlantic and the Western Mediterranean waters and produced an increase in continental aridity (Barton et al., 2013; Hogue and Barton, 2016). The continuity of the human population since the Middle Stone Age has been under debate and remains not completely unresolved (Bouzouggar et al., 2008; Irish, 2000). However, recent studies support a population and cultural continuity in the region (Barton et al., 2008, 2016). A Maghrebian origin of this new technology has been suggested (the oldest remains have been found in Eastern Algeria), followed by a subsequent expansion to the east (towards Tunisia and Lybia) and to the west (towards the Moroccan coast) (Hogue and Barton, 2016). In any case, it seems that the climate changes had an important effect in the populations settlement and adaptation, which produced technological and cultural innovations (Bouzouggar et al., 2008). Moreover, a lot of local variation has been found in the Iberomaurusian assemblies. Even a concrete pattern of differences between coastal rockshelters and hinterlands open-air-sites has been detected. These differences could suggest adaptations of the same population to different environments or extensive movement of people and ideas (Barton et al., 2016; Sari, 2014). A complex and heterogeneous funerary behavior has been described in Iberomaurusian assemblies (Humphrey and Bocaege, 2008; Mariotti et al., 2014), which is an interesting feature of humans from a population and anthropological perspective. It is important to stress that the preliminary link between the Iberomaurusian and Iberian cultures was discarded, and this culture has shown stronger African affinities (Barton et al., 2013).

The arid period that may lead to the Iberomaurusian emergence, was overcame by a wet period in the Sahara between 11,000 to 8,000 ya. Again, coincident with a climatic change, a new culture rose around 10,000 ya (Newman, 1995): the Capsian. However, the extension of this culture is not homogeneous in all North Africa, it is prevalent in Tunisia

and Eastern Algeria, while for example in Cyrenaica, a coastal region in Eastern Libya, the Dabban, Oranian and Lybicocapsian cultures succeeded (Newman, 1995). The Capsian sites are known as *escargotières* because snail shells have been found. In these sites there is also an accumulation of ashes, burned stones, and knapped flint. (Camps, 1995; Rahmani, 2004). The Capsian communities were larger and more sedentary than previous cultures in North Africa (Newman, 1995). Three periods are usually described: the Upper Capsian, the Typical Capsian, and the Neolithic Capsian (Irish, 2000; Rahmani, 2004). The environmental changes around 8,000 ya promoted the changes in subsistence and technology between the Upper and Typical Capsian cultures (Jackes and Lubell, 2008). As for every transition in North Africa, a debate arose around the population continuity with the emergence of the Capsian. Some authors argue a population replacement and different origins of the newcomers have been proposed, being probably an eastern origin the most accepted (Camps, 1995). However, other authors claim population continuity, arguing gradual and continuous changes and a shift due to a different subsistence economy, among other reasons (Barich and Garcea, 2008; Irish, 2000; Peter Sheppard., 1990).

Upper Capsian groups changed from a hunter-gatherer way of life to the Neolithic later than the populations of the nearby areas. The reason could be the high adaptation and flexibility of those groups to changing conditions (Rahmani, 2004). In western Maghreb and Libya, the Iberomaurisian finished with the transition to the Neolithic (Rahmani, 2004). The beginning of the Neolithic was a gradual process that started around 7,600 ya in the North African coast (Linstädter et al., 2012). The oldest date of a domesticated plant is one lentil dated around 7,611 ya in Morocco (Morales et al., 2013). Moreover, around 7,000 ya Neolithic pottery, probably adopted from neighboring populations, is found in hunter-gatherer sites. By 6,000 ya the Neolithic was extended in all the region (Linstädter et al., 2012).

A demic diffusion model for the Neolithic arrival and/or putative interactions with local groups is widely accepted (Mulazzani et al., 2016). A Middle Eastern origin of the newcomers is the most accepted hypothesis (Morales et al., 2013). In Libya, agriculture might have been introduced from Egypt and the Nile Valley (Garcea, 2004; Morales et al., 2013). The similarities between Southern Iberian and Northwest African Neolithic have been explained by three hypotheses. Two of these

hypotheses imply direct contact between Iberia and North Africa: an Iberian diffusion to North Africa or a diffusion of some specific traits through the creation of networks that transported technology, ideas, and maybe people (Linstädter et al., 2012; Mulazzani et al., 2016). However, the most accepted hypothesis nowadays is a synchronous arrival of Neolithic to both places through an east-to-west maritime spread along both shores of the Mediterranean, followed by acculturation and integration of the Neolithic elements (Morales et al., 2013; Mulazzani et al., 2016).

In conclusion, North Africa has been populated since the first stages of *Homo sapiens*. Although the North African prehistoric studies are underrepresented compared to other regions, a large amount of data is available. The main debates around North African prehistory are about the population continuity and replacement in different periods of time, and then the origin of the subsequent cultures.

1.1.3 The Mediterranean: a breeding ground

The Nile river has been the germ of a series of great civilizations. Egypt has played a major role in the history of North Africa, connecting Africa to Asia. In the Early Dinastic Period (3,000-2,686 BCE) a trade network connected Egypt with the Eastern Mediterranean, and probably even before this period of time, Egypt played an important role in the dissemination of Neolithic ideas. During the Old Kingdom (2,686-3,260 BCE) rulers linked spiritual and political royal authority and built the pyramids. After that, a period of uncertainty and instability arrived, known as the First Intermediate Period (2,160-2,055 BCE). The Middle Kingdom (2,055-1,650 BCE) restored the order and settled a golden age for Egypt literature and emphasized social ethics and moral responsibility. However, the Hyksos (coming from Western Asia) gradually entered into Egypt provoking a Second Intermediate Period (1650-1550 BCE). The Hyksos and the Kushites were expelled and the Pharaohs became the rulers of the New Kingdom (1550-1069 BCE). Once again, Egypt's pivotal location drew different peoples (Libyans, Hyksos and Kushites) which put pressure in the territory and finally the Persian Empire overwhelmed Egypt (525 BCE). From a population perspective, different peoples (Canaanites, Hittites, people arriving from the sea, etc.) converged in Egypt. (Naylor, 2009; Newman, 1995)

The Egyptian declined in the 12ᵗʰ century BCE favored the Phoenician expansion. The Phoenicians originated in the Levant coast, in Middle East, and developed a flourishing civilization. The Phoenicians had an economy based on trade which pushed them to sail through the rest of the Mediterranean. Their settlements were based on city-sates along the coast, the most important ones in the Levant were Sidon and Tyre, but they created many settlements along the Mediterranean coast that catalyzed the commercial activities. The North African coast was colonized in the 8ᵗʰ century BCE by Phoenicians. Among the most important city-states in North Africa were Utica, which is the oldest city and it is dated 1,100 BCE; Lixus, in west Morocco; and Carthage, which became later independent and absorbed(Kenrick, 1855; Phoenician commercial colonies under new Carthaginian rule. (Kenrick, 1855; Naylor, 2009; Newman, 1995)

The Carthaginians remained in the North African coastline and did not intended to conquer the interior, which was ruled by Berbers, and both groups kept good relations. The Carthaginians traded with the Berbers and acted as middlemen between them and Mediterranean areas. This commerce facilitated a trans-Saharan trade carried by the Berbers, which even included slaves. In this commerce, the Carthaginians traded with the Garamantes (a group of Berbers). The transcultural transmission between Carthage and Berbers was fluid and influenced the development of proximate groups, such as the Masaesyli and the Mauri. In the 5ᵗʰ century BCE Carthage was defeated by the Greeks from Syracuse in the Battle of Himera, in Sicily. After that, Carthage changed its policy and extended hinterland and by the end of the 5ᵗʰ century Carthage dominated the Western Mediterranean. (Camps and Vela i Aulesa, 1998; Mokhtar, 1981; Naylor, 2009)

The Greeks colonized some territories of North Africa including the Cyrene city and cities nearby (631 BCE). They had contact with the Libyans (the designation they used for North African people) and a cordial relation with the Carthaginians. In the 332 BCE Alexander the Great occupied Egypt. This period, known as the Hellenistic North Africa, had a magnificence development characterized by the splendor of Alexandria. (Naylor, 2009; Newman, 1995)

In 509 BCE Carthage recognized the Roman independence from the Etruscans and agreed in free Mediterranean trade for Carthage. However,

10

Rome started to gain power and conquered Sicily in the First Punic War (246-241 BCE) against Carthage and the Greeks. This fact meant the beginning of the decline of Carthage. Nevertheless, the Carthaginians still tried to expand and extended their power to the Iberian Peninsula. That expansion produced the Second Punic War (218-201 BCE). Despite the Hannibal (a Carthaginian general) effort to cross the Alps with elephants (Mahaney et al., 2017) aiming to defeat Rome, Rome succeeded and subdued the North African colonies. Hannibal army was composed of men from different origins, such as Gauls, Ligurians, Mauri or Numidians, which illustrates the Carthaginian interaction with close populations. In the Third Punic war (149-146 BCE) Carthage lost its last and more important bastion: Carthage. Numidia acted as ally of Rome in this last war, but after their victory Rome defeated Numidia and other Berber kingdoms in the Jugurthine war. Therefore, the Roman domination of the North African coast and some hinterland regions finished the Hellenistic Egypt and the Berber Kingdoms rule (1st century BCE). (Naylor, 2009; Newman, 1995)

The Romans remained mainly in the coast of North Africa, while in the hinterland different Berber Kingdoms took power. The Berbers provided the Romans with different goods, as wool and leather, and also slaves, some of them from sub-Saharan Africa. However, the relationships were not always cordial, especially with Garamantes. Berber Revolts occurred since the 1st century CE. Around the 3rd and 4th century CE a series of revolts weakened the Roman authority. In the 420 CE the Vandals (Germanic invaders of the Western Roman Empire) crossed from the Iberian Peninsula to North Africa. They made good relationships with the Berbers and promoted transculturalism. The Vandals dominated the Western Mediterranean and most of its islands with an extraordinary fleet mainly composed by Berbers. In the 533 CE the Byzantines defeated the Vandals, but once again the Berbers rebellions challenged the Byzantine domination. (Brett and Fentress, 1997; Naylor, 2009)

In the 610 CE Muhammad had the revelation that began the Islam. In the 639 CE Egypt, under Byzantine rule, was invaded by the Arabs. Egypt became the starting point for the North African expansion. Three years later the Arabs conquered Cyrenaica (West Libya) and soon after Tripolitania (Northeast Libya). In their expansion wave, the Arabs built new cities in North Africa. The Umayyad Dynasty (661-750 CE) continued the expansion thorough the Maghreb and reached the Atlantic.

Although the Arabs were tolerant toward other religions, there was discrimination against non-Arabs (even the converted: Mawali). As a result, the Berbers, allied with the Byzantines, tried to resist the Arab domination. Dihya (or Al-Kahina for the Arabs) symbolizes the last Berber resistance to the Ummayad. She was defeated in 698 CE in the final battle for the domination. In 711 CE Tariq crossed the Mediterranean and entered Iberia, an expansion that would culminate with the al-Andalus brilliance. (Camps and Vela i Aulesa, 1998; Naylor, 2009; Newman, 1995)

However, the Berbers were not satisfied with the Arab rule and in 741 CE a Berber revolt spread into Algeria and al-Andalus. At the same time, the Abbasids rose up in the east and, supported by the Shi'a, defeated the Umayyads in the 750 CE. The Abbasid had to assume the descentrality of their Caliphate. Egypt was ruled by the Tulumids and the Ikshids. Today's territoy of Morocco was ruled by the Idrisids, who were not Shi'i and then were persecuted by the Abbasids. The Idrisids were close to the Berbers, specially the Awraba. In 808, Idriss II founded the city of Fez. Eastern to the Idrididis territory another persecuted group, the Kharijites, found shelter among the Berbers. Their good relationships allowed them to rule the Rustamid dynasty. Their main city was Tahart, which had a valuable location for the Saharan trade. Finally, the Aghlabids controlled the Western Mediterranean and some of its Islands in a territory between today's Northern Algeria and Tunisia. This period finished with the arrival of the Fatimids to the power. (Camps and Vela i Aulesa, 1998; Naylor, 2009)

The Fatimids were the unique Shi'i caliphate in history. They started to expand from Algeria to the whole North Africa. The Fatimid period was characterized by the mixture of cultures. Their army was composed of different peoples, especially Berber (Kutama and Masmuda) that were engaged by the Fatimids, but also Bedouins, Turks, Persians and sub-Saharan Africans. In the 11th century CE, the Fatimids experienced a socioeconomic crisis. At that moment, Arab Bedouins (led by Banu Hilal and Banu Sulaym) entered the Maghreb and defeated their rulers. This second wave of Arabs meant an important transformation of the Maghreb. The Arabic language usage increased and also the intermarriage between Arabs and Berbers. Theferore, this invasion was probably the most important event in the mixture of cultures and peoples in the Maghreb. (Camps and Vela i Aulesa, 1998; Naylor, 2009)

Soon after the invasion of the 11th century, the Almoravids started a new period. The Almoravids emerged from the Lamtuna Berbers, in today's Mauritania and Western Sahara. They expanded through the Maghreb, founded Marrakesh in 1062 CE and dominated the al-Andalus until the 1145 CE. However, the Almoravids were soon defeated by the Almohades. Ibn Tumart founded the Almohades, but the Masmuda and the Zanata Berbers were close followers. Marrakesh fell in 1147 CE to Almohade's power, and the Maghreb and al-Andalus were united in an apogee of power, culture and influence. (Naylor, 2009; Newman, 1995)

After the fall of the Almohades in the 13th century the Maghreb was reformed. The Hafsids ruled in the Northern Algeria, Tunisia and Libya. The Zayyanids, Zanata Berbers, ruled in Central Maghrib and had the capital in Tlemecen which was a strategic point for trans-Saharan and Mediterranean trade. At last the Marinids, also Zanata, ruled in Morocco. (Camps and Vela i Aulesa, 1998; Naylor, 2009)

In the 15th century the conflicts in the Iberian Peninsula extended to the Maghreb. The Spanish and the Portuguese won coastal strategic positions (Oran, Tripoli, Tangier, etc.). In Morocco, the Sharifian Dynasty expelled the Portuguese and the Wattasids, who had the power. In the 17th century the Alaouites came to power thanks to an army of sub-Saharan African slaves. (McKenna, 2010; Naylor, 2009)

Meanwhile, in the east, the Ottoman Turks began a new occupation. They occupied Egypt, Algeria, Tunisia and Libya in the 16th century. Algeria rebelled in 1689 and achieved some independence from the Ottoman government. In 1830 internal conflicts and the gain of influence of European powers favored the French invasion of Algeria. Tunisia also achieved some independence of government from the Ottoman rule, and afterwards the Husainid dynasty ruled Tunisia until 1957 as a French Protectorate. In Libya, the Karamanlis governed until 1835, when the Ottomans had again the control. During the 18th and 19th centuries, the European colonialism controlled in different ways each North African region. (McKenna, 2010; Naylor, 2009)

In conclusion, the history of North Africa is extremely complex and the Mediterranean has been a breeding ground for many civilizations. As a result, North Africa encompass an amalgam of cultures and peoples. The

detailed study of history highlights the importance of transculturalism in the understanding of the North African peoples.

1.1.4 The Berber peoples

The North African population is composed by a high diversity of peoples with different ancestral origins, languages and cultures. It is an oversimplification to classify North African populations into only Arab and Berber peoples; however, many questions arise when these two groups are compared. The Berbers are usually considered the ancestral North African population, while the Arab arrived during the last two millennials.

The origin of the term "Berber" is controversial. It probably comes from the Latin "barbarus" and was later spread by the Arabs. However, they call themselves Amazigh (or Imazighen in plural) (Camps, memorie et identite). The main characteristic that all Berbers share is the Berber language (Tamazigh). The writing of Berber languages is called Tafineq and it is thought to have developed during the Carthaginian period (General history of Africa I). These languages are classified within the Afroasiatic linguistic family (Camps, 1995; Newman, 1995). The origins of the Berber peoples are still under debate. The most claimed hypothesis is that they come from the Capsian culture that emerged around 10,000 ya (Camps, 1995). Linguistic and genetic data support this hypothesis (Fire and Disturbance, 2004; Maca-Meyer et al., 2003). The place of origin of Afroasiatic languages is uncertain, but the Northern Sudan Nile valley is the most plausible (Newman, 1995). The time of origin is more certain, dated between 10,000 and 15,000 ya, clearly before the Neolithic (Fire and Disturbance, 2004; Maca-Meyer et al., 2003). The expansion of the Afroasiatic languages has been correlated with a demographic expansion linked to the U6a1 mitochondrial haplogroup (Maca-Meyer et al., 2003). Moreover, genome-wide data (Henn et al., 2012) has pointed to a population replacement in North Africa before 10,000 ya. Taken all together, a probable origin of proto-Berber people coming from the east and dragging Afroasiatic language is well supported. However, the population continuity during pre-history in North Africa has been questioned at many temporal stages, as stated previously. Demographic or cultural replacements before or after 10,000 ya could place the Berber origin in a different stage.

The proto-Berber, or Libyans as were called by the Romans, were not a unique group but rather a composite of different populations that inhabited different regions and had different lifestyles (Camps, 1995; Camps and Vela i Aulesa, 1998; Naylor, 2009; Newman, 1995). Agriculture was well spread long before the Phoenician arrival and was practiced by many groups, while others were nomadic (Ki-Zerbo, 1981). Among the different Berber groups that were present before the Phoenician arrival were the Garamantes, Mauri, Gaetuli, Masaesyli, Numidians and many others (Camps, 1995; Camps and Vela i Aulesa, 1998; Newman, 1995). The history of Berbers intertwined with that of the Phoenicians and the subsequent peoples and civilizations that arrived in North Africa, as has been explained above (Brett and Fentress, 1997; Naylor, 2009). After the Arab expansion, groups of neoberbers arrived to North Africa coming from the east and the south. They were the Butr, Zanata, Levatha, Laguatan and Luwata. Those Berber groups brought nomadism to areas that were previously cultivated (Camps and Vela i Aulesa, 1998).

Lifestyles, contacts and influences from external populations are extremely variable depending on the Berber group considered. For instance, the Tuareg are nomadic Berbers (Camps, 1995), but most of the Berber groups are not. The Garamantes had continuous contacts with sub-Saharan Africans through trans-Saharan routes, while other groups may had other influences (Mokhtar, 1981). Regarding the Arabization and the adoption of Islam, Berbers also present differences. Although nowadays most Berbers groups adopt the Sunnism, other Islam branches have been relevant in different historical periods. For instance, the Kutama played an important role in the spreading of Islam and adopted the Shiism, and the Barghwata developed a branch of Khawarij. Overall, during the Middle Ages the Berbers adopted easily any sociocultural movement that questioned the established authority (Camps, 1995).
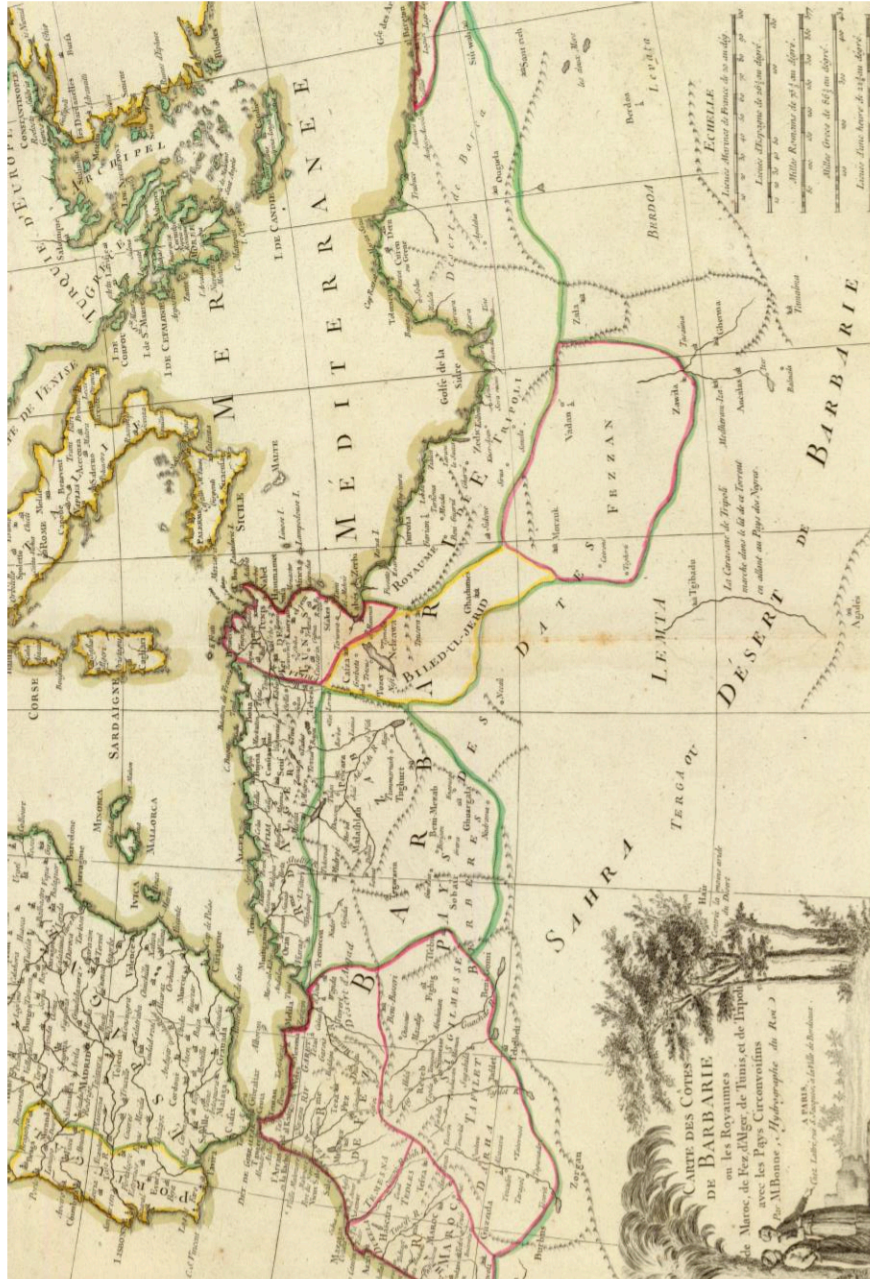
Fig 3. Bonne, Rigobert (1791). Cartes des cotes de Barbarie, ou les Royaumes de Maroc, de Fez, d'Alger, de Tunis, et de Tripoli avec les pays circonvoisins (Lattre & Delalain).

1.2 THE GENETIC PERSPECTIVE OF THE HUMAN SETTLEMENT IN NORTH AFRICA

The analyses of genetic and genomic data allow the reconstruction of the populations history. The genetic markers analyzed are the mitochondrial DNA (mtDNA), the Y chromosome and autosomal markers. MtDNA and Y chromosome are uniparental markers that are inherited without recombination from the mother and from the father, respectively. Autosomal data include a wide variety of markers, from Short Tandem Repeats (STRs) to whole genome sequence data.

1.2.1 A paleolithic mother

In North Africa mtDNA is the oldest genetic evidence, as some of the haplogroups present are related to the Middle Stone Age. The analysis of U6 and M1 mtDNA haplogroups has been interpreted as a back-flow to Africa during that period (Secher et al., 2014). The mitochondrial U6 haplogroup shows higher frequency (around 10%) and diversity in North Africa than everywhere else (Plaza et al., 2003), which has been explained by a North African origin and expansion of this haplogroup (Plaza et al., 2003). The divergence time of U6 was previously estimated around 35,000 ya (Maca-Meyer et al., 2003; Secher et al., 2014); however, the analysis of an ancient sample from Romania carrying a basal U6 lineage pushed back that estimation to 49,600 ya. Therefore, the reviewed history of U6 haplogroup suggests that it probably emerged in Western Asia, and afterwards expanded to North Africa (probably around 22,000 ya (Pereira et al., 2010a)), where it constituted a local background, particularly in Northwest Africa (Plaza et al., 2003). During this expansion, the U6 haplogroup diversified until the emergence of all U6 lineages found nowadays (Hervella et al., 2016). The effective population size estimated for the U6 reflects the population growth changes of North African groups: from 17,000 to 13,000 ya there was a population stasis, and then 13,000 ya it started to grow again (Secher et al., 2014). In parallel to the U6 expansion, the U5 haplogroup emerged around 29,900 ya and expanded from Central Asia to Europe (Secher et al., 2014). In Europe, U5 haplogroup is nowadays frequent in the Baltic and Basque Country and was common in Mesolithic Europeans (Matisoo-Smith et al., 2016).

A recent study analyzed ancient mitochondrial DNA from 28 Iberomaurusian samples from Algeria and Morocco (7 and 21 samples

respectively) (Kefi et al., 2016). The Algerian site analyzed (Afalou) is dated between 23,000 and 10,800 ya, and the Moroccan site (Taforalt) was dated between 15,000 and 11,000 ya. All the mtDNA haplogroups found in those samples are classified into Euroasiatic lineages: H, T, JT, J, V and U. Two samples were identified as U6, which indicate that the U6 lineage was already present in the Iberomaurusian individuals. Moreover, the lack of sub-Saharan mtDNA haplogroups suggests that the Iberomaurusian do not have a sub-Saharan origin and that the sub-Saharan lineages found in current populations might have arrived after 10,000 ya (Kefi et al., 2016), as previously suggested (El Moncer et al., 2010). The finding of the U6 lineage in Iberomaurusian samples is in accordance with previous hypotheses that linked the U6 and the M1 haplogroups expansions from Southwest Asia to the Iberomaurusian culture (Pennarun et al., 2012). Moreover, the expansion of mtDNA haplogroups during the Late Stone Age seems to be correlated with the expansion of Afro-Asiatic languages in North Africa (Pennarun et al., 2012).

The arrival of other Euroasiatic mtDNA lineages (H, U5, V, etc) has been hypothesized to be more recent, in Holocene times (Pereira et al., 2010a). However, the presence of some of those lineages in the ancient Iberomaurusian samples challenges this previous hypothesis. The study of Kefi and colleagues (2016) also showed a genetic affinity between the ancient samples and some modern North African (Berbers from North Morocco and Jerba Island in Tunisia) and some South European populations (Valencia, Baleares, Spain, Sardinia an Italy), which the authors claim to evidence the existence of a common Mediterranean mitochondrial gene pool. Probably one of the main conclusions of the study is the support for some genetic continuity in North Africa at least after the Late Stone Age.

The presence of U6 mtDNA haplogroups in some European Mediterranean populations has been suggested as evidence for gene-flow between both shores in different times including prehistoric periods (Maca-Meyer et al., 2003; Plaza et al., 2003). For instance, the mtDNA U6a haplogroup radiated from the Maghreb to the east, and might have reached the Iberian Peninsula (Secher et al., 2014). Moreover, the mitochondrial DNA of a Phoenician sample from Carthage has been recently analyzed and assigned to mtDNA haplogroup U5b2c1, which probes that this haplogroup might be introduced in North Africa from the European shores at least in the 6th century BCE, and probably earlier

(Matisoo-Smith et al., 2016). The mtDNA haplogroup U5b2c1 has not been found in modern Lebanese or ancient samples from the Levant, which could suggest another Mediterranean origin of that sample, in accordance with a wide movement of people in the Mediterranean during the Phoenician period (Matisoo-Smith et al., 2016).
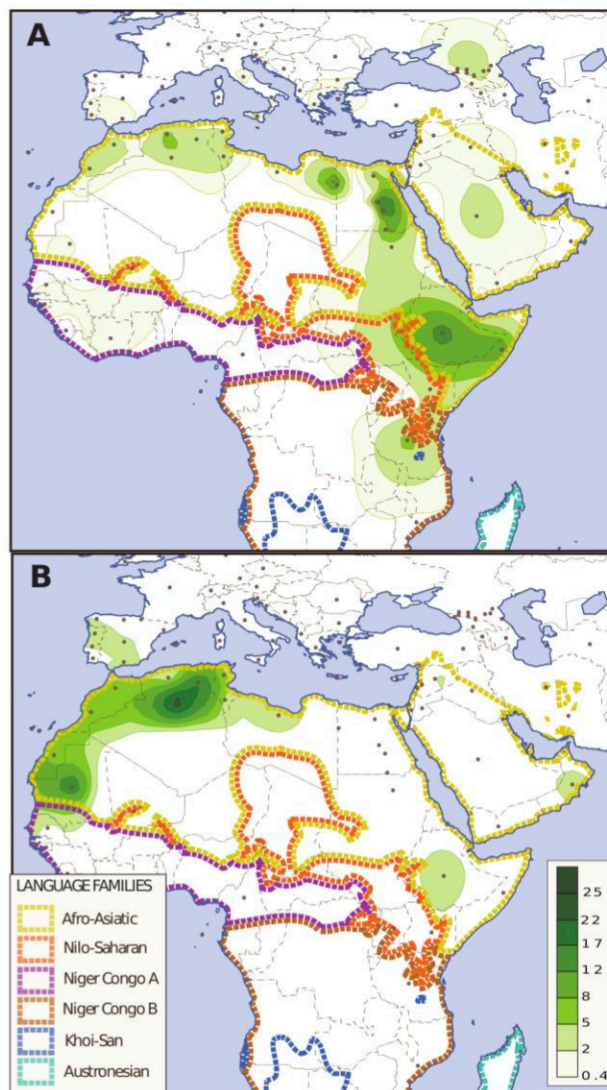
Fig 4. Spatial distribution of haplogroup a) M1 and b) U6, with languages' phyla. The dots indicate the populations geographic locations. (Pennarun et al., 2012)

Despite the importance of U6 haplogroup for the understanding of the North African populations origin, most of the mitochondrial gene pool in North Africa is represented by Eurasiatic lineages. One of the most frequent mtDNA lineages is haplogroup H, which has been linked to European migrations since the Last Glacial Maximum (around 15,000 ya) mainly from the Iberian Peninsula. The mtDNA subhaplogroups H1 and H3 are the most frequent within the H clade in the Iberian Peninsula with frequencies of 45 and 16%, respectively, and those frequencies decrease towards Eastern Europe. In North Africa, the frequencies of those mtDNA subhaplogroups are similar to the ones found in Europe (42 and 13%, respectively), and these frequencies also decrease towards the east. The highest frequencies of mtDNA haplogroup H in North Africa are found in the west, in Morocco, were they reach up to 37%, while in Egypt it ranges between 21 and 14%. (Ennafaa et al., 2009; Hernández et al., 2017)

Haplogroups H1 and H3 have been associated to hunter-gatherer expansions after the Last Glacial Maximum from the Franco-Cantabrian refuge. The coalescence time in European samples for H1 has been estimated around 14,000 ya, and for H3 around 10,000 ya. Therefore, the presence of those haplogroups in North Africa has been hypothesized to come from paleolithic expansions from the Iberian Peninsula, which could explain its higher frequency in Western North Africa. The coalescence for H1 haplogroup in North Africa has been estimated 8,000 to 9,000 ya (Ottoni et al., 2010), and the H1 phylogeny in North Africa is very diverse and show an early radiation. This fact supports the Iberia to North Africa direction of this gene-flow. There are historical evidences of maritime transport (Broodbank, 2006; Hernández et al., 2017), which makes possible the contact between the Iberian Peninsula and the Maghreb by sea by that time. Some other mitochondrial haplogroups, such as haplogroup V and U5b, have been suggested as gene-flow from the Iberian Peninsula during the postglacial expansion (Cherni et al., 2009; Ennafaa et al., 2009; Hernández et al., 2017; Ottoni et al., 2010). In accordance to this finding, archaeological similarities have been described between the Iberomaurusian and European Magdalenian, which support the population contact during the Late Stone Age (Hernández et al., 2017).

In summary, the North African mitochondrial gene pool is predominantly composed by a back-to-Africa gene-flow in the Paleolithic represented by haplogroups U6 and M1, and a high frequency of Eurasian haplogroups that arrived mainly in postglacial times. Historical gene-flow has been

suggested to have a lower impact in the mitochondrial gene pool in the region (Hernández et al., 2017; Ottoni et al., 2010). However, some lineages are related to historical gene-flow from Middle East and sub-Saharan Africa.

Sub-Saharan gene-flow into North Africa has been suggested to occur in more recent times. In Northwest Africa, the sub-Saharan mitochondrial haplogroups show frequencies of 25.9 ± 2.1% (Plaza et al., 2003). Mitochondrial L haplogroups are traces of that gene-flow. These L lineages found in North Africa originated in sub-Saharan Africa during the Holocene (around 10,000 ya) and their coalescence corroborates their presence in North Africa after the Euroasiatic expansions of the Stone Age (Harich et al., 2010). Moreover, the geographical distribution of the mtDNA L subhaplogroups correlates with slave-trade routes from sub-Saharan Africa which, together with a gradual cline towards the north, suggests an origin in historical times (Bekada et al., 2013; Brakez et al., 2001; Harich et al., 2010; Kujanová et al., 2009). The study of ancient mtDNA from Iberomaurusian samples (Kefi et al., 2016) corroborates the lack of sub-Saharan haplogroups until 10,000 ya. However, other studies have suggested a more ancient gene-flow, in pre-Holocene times (Ottoni et al., 2009). Finally, some mitochondrial haplogroups, such as R0a and J1b could have arrived to North Africa from Middle East in historical times, for example with the Arab expansion, although some Middle Eastern haplogroups may predate that period and could have arrived during the Neolithic (haplogroups J and T) (Coudray et al., 2009; Ennafaa et al., 2011; Kujanová et al., 2009).

1.2.2 A late fatherhood

Most Y-chromosome lineages found in North African individuals have been classified as descendants of the Y-chromosome haplogroup E-M35, which has been suggested to have an Eastern African origin, and its derivative lineages might have spread through the rest of the African continent (Luis et al., 2004; Trombetta et al., 2015). However, the identification of the geographical origins of some of the derived E-M35 lineages is not straightforward. The derived lineages of E-M35 include E-M81 and E-M78, among others. Those two lineages are the main substrate of the Y chromosome gene-pool in North Africa.

The most frequent Y chromosome haplogroup in North Africa is E-M81 (Solé-Morata et al., in press). Its distribution is not uniform: high frequencies are found in the west reaching 56.1% in Morocco (Fadhlaoui-Zid et al., 2013), and its frequency decreases towards the East of Maghreb (Algeria around 45% (Bekada et al., 2013), Libya 35.88% (Fadhlaoui-Zid et al., 2013)), and Egypt where the lowest frequencies are found (11.9% (Bekada et al., 2013)). However, this gradient is not homogeneous and differences can be found within restricted geographical areas. For instance, in Tunisia, E-M81 reaches its highest frequency in North Africa (71-79%) when considering a pool of populations together (Fadhlaoui-Zid et al., 2011, 2013), but focusing in specific populations the results are highly variable. In the Berber samples from Chenini and Jradou, this is the only haplogroup found (100% of frequency); while in a cosmopolitan Tunisian sample the frequency is not higher than 55% (Fadhlaoui-Zid et al., 2011). This heterogeneous pattern is shown in other populations, such as the Tuaregs from Libya where E-M81 shows a frequency up to 81% (Pereira et al., 2010b), and in general Berbers show the highest frequencies of E-M81(Cruciani et al., 2004).

The high frequency of E-M81 in North Africa has been the focus of most Y chromosome studies in the region. Its origin, both temporally and geographically, has been one of the main debates around the Y chromosome in North Africa. In a first study by Bosch et al. in 2001, a Paleolithic origin for E-M81 was suggested (Bosch et al., 2001), but soon after, in 2004, Arredi and colleagues argued for a more recent origin, during the Neolithic (Arredi et al., 2004). A recent study by Sole-Morata and colleagues (Solé-Morata et al., in press) addresses this question using massive parallel sequencing, which has produced the most extensive and updated study of E-M81. This study points to even a more recent origin around 2000 ya. Regarding the geographical origin of E-M81, a Middle Eastern origin has been proposed as the most probable in different studies (Arredi et al., 2004; Solé-Morata et al., in press). The east-west gradient of variation and the higher diversity of this haplogroup in the Middle East are evidences for this hypothesis. However, though unlikely, a Northwest African origin cannot be discarded due to the high frequencies of E-M81 in the west and its frequency gradation towards east (Arredi et al., 2004; Solé-Morata et al., in press).

E-M78 (another descendant haplogroup of E-M35, as stated above) is also frequent in North Africa. It is found mostly in the Middle East and Northeast African populations (in Libya its frequency is around 11% and in Egypt 36%, while in Morocco is decreases until 6% and in Algeria estimates are between 6% and 2%) (Bekada et al., 2013; Fadhlaoui-Zid et al., 2013). Initially, an East African origin was suggested for the E-M78 (Semino et al., 2004); however, a Northeastern African origin is the most probable hypothesis according to the diversity found for this haplogroup in that region (Cruciani et al., 2004; Fadhlaoui-Zid et al., 2004). The coalescence time of E-M215 haplogroup (the ancestral lineage of E-M78) was estimated around 22,400 ya (Cruciani et al., 2004), and placed the introduction of the E lineage in Northeast Africa during the Late Stone Age. The presence of E-M78 lineages in East Africa has been explained as later back migrations. Altogether highlights the relevance of the Nile Valley as a pivotal genetic corridor (Cruciani et al., 2007).

Besides the E lineages, J haplogroups are also present in North Africa. J haplogroup has a Middle Eastern origin (Arredi et al., 2004; Semino et al., 2004), where it reaches its higher frequencies (Fadhlaoui-Zid et al., 2013). More specifically, J haplogroup has two main derived lineages: J1-M267 and J2-M172. J1-M267 is more prevalent in the Arabic Peninsula and has been associated to the Arab expansion (Zalloua et al., 2008), while J2-M172 is more frequent in the Levant and Anatolia and it is associated to
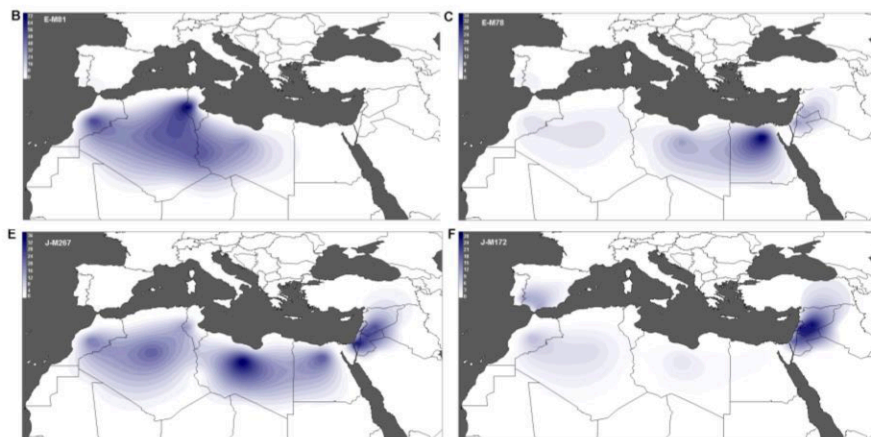
Fig 5. Frequency of the major Y-chromosome haplogroups in North Africa and surrounding regions, b)E-M81, c)E-M78, e)J-M267 and f)J-M172. The intensity of the colors reflects the frequency of each haplogroup in the studied populations. (Fadhlaoui-Zid et al., 2013)

the Phoenician maritime routes (Haber et al., 2011). The TMRCA for the J chromosomes in North Africa has been estimated between 4,400 and 11,100 ya, around the Neolithic (Arredi et al., 2004). These dates support a recent demographic event, in accordance with an expansion of haplogroup J1-M267 during the Arab expansion, but some studies have also suggested an earlier origin in prehistoric times (Arredi et al., 2004; Bekada et al., 2013).

Finally, some lineages found in North Africa are related to sub-Saharan gene flow, in more recent times, probably as a result of slave trade, such as haplogroups A, B, E-M96, E-M2, and E-M35, which in total represent the 16.3% of the Tunisian lineages (Ennafaa et al., 2011; Luis et al., 2004).

1.2.3 The autosomal inherited diversity

a) *Classical markers and STRs*

The first studies of genetic autosomal data analyzed classical markers (blood groups, red cell enzymes, serum proteins and HLA antigens), Alu insertions, and STRs. A compilation of studies focused on classical markers showed an east-west gradient of genetic differentiation (Bosch et al., 1997). In summary, these studies showed that Libya and Egypt differentiate from the rest of the Maghrebi countries and are closer to Europeans. In order to explain this genetic gradient, a demic diffusion from the Levant in Neolithic times and subsequent differentiation of Western populations has been suggested (Bosch et al., 1997). According to this hypothesis, the closeness of North Africans to Southern Europeans has been proposed as an indicative of a common origin in the Upper Paleolithic or the Neolithic (Coudray et al., 2006). Moreover, a north-south differentiation has also been described when studying specific Moroccan populations (Harich et al., 2002).

STR and Alu analyses have shown a substantial genetic differentiation between Iberian and North African populations, which points to the Gibraltar strait acting as a genetic barrier. However, this genetic barrier has not been total and limited gene-flow between both shores has been described. Moreover, continuous sub-Saharan gene-flow has been detected in North Africa, which might have increased the differentiation between both Mediterranean shores. The sub-Saharan gene-flow has also

created a south-north genetic gradient in North Africa. (Bosch et al., 2000; Comas et al., 2000; González-Pérez et al., 2010; El Moncer et al., 2010)

Regarding Berbers, a high heterogeneity and a lack of differentiation between Arab and Berbers was found when analyzing all types of markers (Amir et al., 2015; Bosch et al., 1997, 2000; Coudray et al., 2006; Harich et al., 2002; El Moncer et al., 2010). Moreover, low levels of genetic sharing with Middle Eastern and North African populations has been also described studying classical markers, Alu insertions, and STRs in different geographical populations (Bentayebi et al., 2014; Hajjej et al., 2015; Halima et al., 2014).

b) Genome-wide data

Genome-wide analyses have contributed to valuable knowledge in human population genetics in the last years (Hellenthal et al., 2014; Lao et al., 2008; Leslie et al., 2015; Li et al., 2008; Novembre et al., 2008). Unfortunately, the number of genome-wide studies focused on North Africa is scanty and have focused to a limited number of populations. However, some pioneer studies before the ones presented in this thesis described some general aspects of the genome of North Africans.

The genetic continuity of North African populations has been under a strong debate from archaeology, anthropology and genetic studies based on uniparental markers. The first study that included a wide dataset of North African populations (Henn et al., 2012) showed that those populations were an amalgam of four different components: European, Middle Eastern, sub-Saharan and autochthonous North African. The North African component showed a gradient decreasing from west to east, in contrast with the opposite pattern for the Middle Eastern component, from east to west (Henn et al., 2012). The North African component was proved to come from a back-to-Africa migration in pre-Holocean times (more than 12,000 ya). Therefore, in this study at least one genetic replacement from an out-of-Africa population was confirmed (Henn et al., 2012). Accordingly, Neanderthal admixture has been described in North African samples independently of their European and Middle Eastern components. This indicates that the ancestral population from which the autochthonous North African component derived was already admixed with Neanderthals (Sánchez-Quinto et al., 2012).

A genome-wide study of admixture in different populations hypothesized that African gene-flow in Europe and the Middle East might have come from North Africa in historical times, maybe during the Roman Empire (Moorjani et al., 2011). Although this study was not conclusive due to the lack of North African samples, it predicted the results found in forthcoming analyses. The study of Henn and colleagues (2012) showed a sub-Saharan contribution to the North African gene-pool. This admixture process was continuous and around 1200 ya in South Morocco, but as a result of a pulse and a more recent (around 700 ya) in Egypt. A recent analysis of ancient DNA from Egyptian mummies has shown an increase of sub-Saharan ancestry in modern Egyptian samples (14-21% of the total ancestry) compared to the ancient samples (6-15%) and, therefore, confirms sub-Saharan gene-flow in post-Roman times (Schuenemann et al., 2017). Further studies have shown an extensive admixture with sub-Saharan Africa along the whole North Africa (Fadhlaoui-Zid et al., 2013). The North African gene-flow has been also studied in the opposite direction: from North Africa to external populations. It has been shown that recent migrations from North Africa have increased the genetic diversity in Southwestern Europe (Botigue et al., 2013). The studies
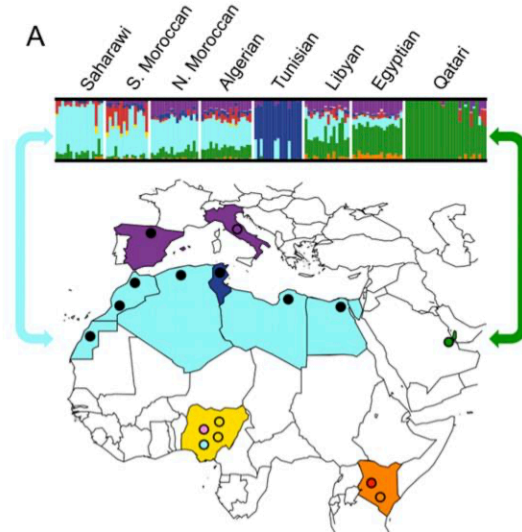
Fig 6. Map of samples and population structure of North Africa and neighboring populations. The barplot shows an ADMIXTURE analysis in which each color represents one ancestral component most frequent in the population that match the color in the map. The decreasing proportions of North African ancestry is indicated in a west-to-east gradient of blue across North Africa. (Henn et al., 2012)

presented in this thesis go in depth into this admixture processes with sub-Saharan populations and expand it to other populations such as the Middle East.

The North African samples analyzed in genome-wide studies are mainly from Arab-speaking groups, and a large group of North African Jews that showed high genetic affinity with other non-African Jewish groups (Campbell et al., 2012). Before the data presented in this thesis, only genome-wide data from two Berber-speaking samples were available, the Mozabites from Algeria (Cann, 2002) and a Tunisian group from Chenini, both described as isolated and outlier populations (Henn et al., 2012), which challenged the complete knowledge of the population history of the region.

1.3 STATE OF THE ART OF POPULATION GENETIC ADMIXTURE

The completion of the Human Genome Project (Lander et al., 2001; Venter, 2001) remodeled the functional and association genetic studies, but also the human population genetics field. Soon after the publication of the human genome sequences, different projects aiming to study the human genetic variation from a genome-wide perspective emerged. In 2002 began the implementation of the Human Genome Diversity Project (Cann, 2002), and in 2003 the HapMap Project (The International HapMap Consortium, 2005; The International HapMap Project, 2003). Only seven years later, the 1000 Genomes Phase I was released (Durbin et al., 2010). The availability of this high amount of genomic data was accompanied by the developing of computational methods for studying the genetic variation among populations. In 2008, Li and colleagues published a pioneer study of human genetic variation applying a set of those novel methods to the HGDP panel that includes a wide number of populations. The results showed a genetic structure highly explained by geographic variation and confirmed the sub-Saharan origin of all the studied populations followed by a serial founder effect (Li et al., 2008). Several studies followed this line of research applied to different sets of human populations (Bryc et al., 2010; Lao et al., 2008; Novembre et al., 2008).

The analytical methods for studying population genetics are rapidly becoming more complex and diverse. Principal Component Analyses (PCA) have been widely used in population genetics since the classical studies of Cavalli-Sforza (Cavalli-Sforza et al., 1995). PCA is a descriptive method that allows the study of the population genetic structure. Some conclusions regarding closeness of the populations can be deduced with PCA. However, although it could suggest specific demographic events, such as drift and admixture, other methods are required to probe them. Model-based clustering methods or global ancestry methods, such as ADMIXTURE, STRUCTURE and FRAPPE, are used as a complement of PCA to study the grouping of individuals that share common underlying allele frequencies (Alexander et al., 2009; Pritchard et al., 2000; Tang et al., 2005). However, to statistically test for admixture between populations, several methods have been developed. In

this sense, F statistics' (f3, D statistic, F4 ratio and qpGraph) are a series of statistics that allow testing the occurrence or not of admixture based on the sharing of allele frequencies between populations (Patterson et al., 2012; Reich et al., 2009).

All the methods described above are based on independent markers, usually SNPs (Single Nucleotide Polymorphisms). However, the linkage disequilibrium (LD) between markers can be particularly informative for admixture. The LD between SNPs along the genome generates an exponential distribution, whose decay is correlated with the time since admixture. Methods as Alder and ROLLOFF use this approach (Loh et al., 2013; Patterson et al., 2012). The combination of LD and phasing information increases the power of the methods. Local ancestry methods take profit of these two characteristics and allow us the assignation of each genomic region of the studied population to one ancestral population. There are several softwares that use different algorithms in order to infer the local ancestry: HAPMIX, PCAdmix, LAMP, RFMix, among others (Brisbin et al., 2012; Maples et al., 2013; Price et al., 2009; Sankararaman et al., 2008). However, in order to perform a local ancestry analysis, non-admixed populations that serve as reference of the ancestral populations that experienced the admixture event should be provided. Finally, other set of methods benefit from the haplotype information that LD and phasing produce: ChromoPainter, fineSTRUCTURE and GLOBETROTTER (Hellenthal et al., 2014; Lawson et al., 2012).

ChromoPainter reconstructs the chromosomes of a recipient individual as a series of chunks in LD (or haplotypes within a chromosome) from the other donor individuals in the sample. This information is gathered in a coancestry matrix that is built under a Hiden Markov Model (HMM). The values within this matrix count the number of recombination events leading to one individual being most closely related to another. ChromoPainter analyses can be performed in two different approaches. In a first approach, all the individuals can be considered as both recipient and donors, without providing any population classification. In this analysis we obtain information of the chunks received by each individual from the rest of the individuals in the dataset, which can be summarized in a squared coancestry matrix. In a second approach, individuals are grouped in clusters or populations and each of the populations can be classified as recipient, donor or both. In this analysis the coancestry matrix contains

only the information of the individuals that were included in a recipient population. (Lawson et al., 2012)

FineSTRUCTURE is a clustering analysis that uses a model-based bayesian approach applied to the coancestry matrix from ChromoPainter. A Markov chain Monte Carlo (MCMC) algorithm is used to assign the individual to populations or clusters of individuals. Finally, the fineSTRUCTURE clustering is recommended to be summarized in a tree, which is calculated performing hill-climbing from the MCMC iteration. (Lawson et al., 2012)

GLOBETROTTER models admixture events based on the estimation of coancestry curves. In order to build the coancestry curves, the information of the ChromoPainter coancestry matrix is used. The curves represent a distribution of the relative probability of finding a pair of chunks assigned to two donors separated by a given distance. The ancestry curves discern whether there has been or not admixture between the tested donors. Therefore, when applied to a set of different donor candidates, the most likely sources of admixture can be inferred. Moreover, recombination reduces the length of the admixed chunks in each generation and then the rate of decay of the curves is correlated to the time since admixture. The original or true admixing groups (sources) sometimes cannot be sampled because they no longer exist, have admixed or experienced high drift, or are unknown. GLOBETROTTER allows to infer admixture even if the original source is not sampled. The admixture events that gave place to the target or studied population are analyzed though a set of surrogates (donors), which might be admixed. The surrogates that most probably contributed to the admixture are classified in the major or the minor source of admixture. However, the process is sometimes more complex and more than two sources could contribute to the admixture events. Although it is difficult to interpret the relation among surrogates in these cases, the model is able to discriminate when more than two sources contribute to the admixture. (Hellenthal et al., 2014)

Is the last three year many population genetic studies have applied the haplotype-based methodology of ChromoPainter, fineSTRUCTURE and GLOBETROTTER. Montinaro and colleagues (2015) studied admixed American populations and reported a fine-scale characterization of the source populations, providing a precise description of the admixture events. (Montinaro et al., 2015). In another study a broad overview of

gene-flow and admixture events in Eurasian populations was described (Busby et al., 2015), showing an accurate performance in the detection of admixture and specially in the dating. Thanks to the high resolution that haplotypes allow, the structure of less differentiated samples can be analyzed. A study focused on British populations showed this ability and reveal a fine structure of ancestry (Leslie et al., 2015). Moreover, Van Dorp and colleages (2015) applied studied of Ethiopian populations and showed how this method can improve the study of populations that have experienced high genetic drift (van Dorp et al., 2015). Some other studies have applied this methodology and new studies are implementing widely (Busby et al., 2016; Patin et al., 2017; Sarno et al., 2017). Ancient DNA studies are also improving thanks to the application of ChoromoPainter to datasets (Broushaki et al., 2016; Cassidy et al., 2016).

In conclusion, several methods have been developed for the studying of human population admixture and each analysis require the application of
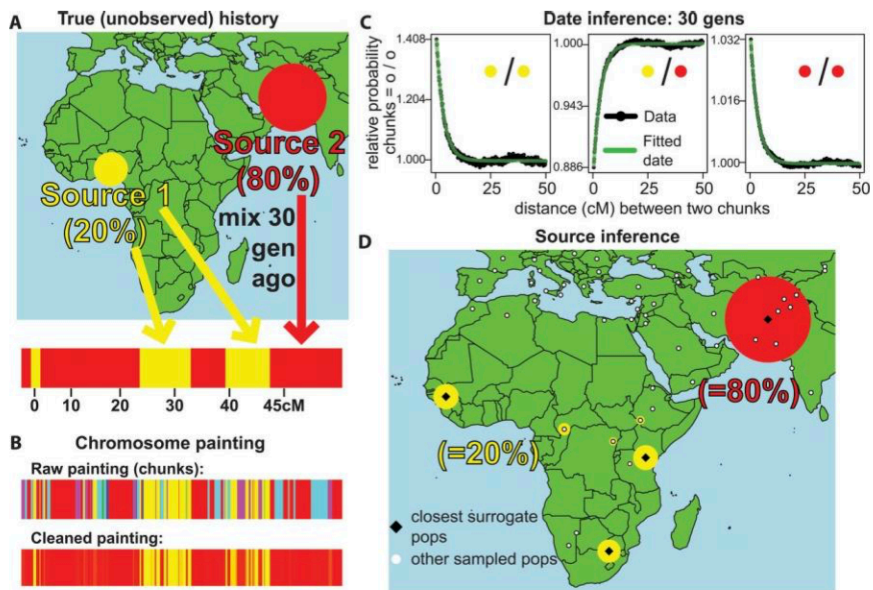
Fig 7. Scheme of an ancestry painting and admixture analysis of simulated admixture using ChromoPainter and GLOBETROTTER. A) Simulated event. B) ChromoPainter's painting showing haplotypic segments shared between the groups. C) The coancestry curves closely fit an exponential decay (green line) with a rate of 30 generations. The positive slope for the Balochi-Mandenka curve (middle) implies that these donors represent different admixing sources. D) GLOBETROTTER's source inference, black diamonds indicate sampled populations with greatest similarity to true sources and circles other sampled populations. (Hellenthal et al., 2014)

those that better fit to the data and the asked questions. Most of the methods cited above study the admixture between populations and an accurate combination of them allows a detailed and precise inference of admixture. Among them, haplotype-based methods are one of the most valuable tools for studying demographic processes due to the high resolution that is achieved, and these have been used in the present thesis to study the admixture processes that have taken place in North African human groups.

2 RESULTS

2.1 GENETIC HETEROGENEITY IN ALGERIAN HUMAN POPULATIONS

Asmahan Bekada*, Lara R. Arauna*, Tahria Deba, Francesc Calafell, Soraya Benhamamouch, David Comas

*equal contribution

2.2 RECENT HISTORICAL MIGRATIONS HAVE SHAPED THE GENE POOL OF ARABS AND BERBERS IN NORTH AFRICA

Arauna, L.R., Mendoza-Revilla, J., Mas-Sandoval, A., Izaabel, H., Bekada, A., Benhamamouch, S., Fadhlaoui-Zid, K., Zalloua, P., Hellenthal, G., and Comas, D.

Arauna LR, Mendoza-Revilla J, Mas-Sandoval A, Izaabel H, Bekada A, Benhamamouch S, et al. Recent Historical Migrations Have Shaped the Gene Pool of Arabs and Berbers in North Africa. Mol Biol Evol. 2017;34(2):318–29. DOI: 10.1093/molbev/msw218

2.3 GENETIC HETEROGENEITY BETWEEN BERBERS AND ARABS

Arauna, L.R., and Comas, D

eLS. 2017.

Arauna LR, Comas D. Genetic Heterogeneity between Berbers and Arabs. In: eLS. Chichester, UK: John Wiley & Sons, Ltd; 2017. p. 1–7. DOI: 10.1002/9780470015902.a0027485

2.4 DISSECTING HUMAN NORTH AFRICAN GENE-FLOW INTO ITS COASTAL SURROUNDINGS

Arauna, L.R., and Comas, D

In preparation

3 DISCUSSION

3.1 NORTH AFRICAN POPULATIONS

3.1.1 Prehistorical evidences

The human settlement in North Africa has been poorly studied and many questions remain unsolved. As evidenced from the recent study of Hublin and collaborators (Hublin et al., 2017), and has been pointed before (Balter, 2011), North Africa may have a crucial role in the emergence and dispersal of early humans. In this context an extension of the prehistoric Africa that encompass the Sahara and Northern Sahara regions could help in the understanding of human evolution.

This perspective is not only supported by archaeological and anthropological evidences, but also from paleoclimatic studies. According to this idea, a Eurocentrism perspective has dominated the archaeological and genetic studies in North Africa. In this regard, the extrapolation of European paleolithic cultures and tempos to North Africa and the intend to find similarities and connections with the European Neolithic have challenged the advance of knowledge (Garcea, 2006).

Moreover, North Africa is sometimes understood as a single unit, while its geographical extension and diversity suggest the need of understanding each specific region in order to build the complete view. Nevertheless, the knowledge about North African prehistory questions the population continuity in the region. Although there is a lack of consensus, it is established that at least one population replacement took place in North Africa before the Holocene (10,000 ya), probably before or during the Iberomaurusian period. There is also agreement in a demic diffusion during the Neolithic with admixture with pre-Neolithic populations and without a complete replacement. However, if more replacements occurred or whether those were complete or partial it is a question that remains unsolved. It is interesting to note how climatological changes, such as the degree of permeability of the Saharan barrier, have played a major role in population changes. This data should be taken in consideration in further archaeological and genetic studies.

3.1.2 Genetic structure

Heterogeneity, diversity, and complexity are probably the most frequent words in North African population genetic studies (Bosch et al., 2001;

Henn et al., 2012). Indeed, those are the main characteristics of the North African genetic structure. The complexity of this region is evinced by the lack of a complete correlation between the genetic structure of the populations and geography or linguistic and ethnic affiliations, in contrast to European populations, where a good correlation with geography has been described (Lao et al., 2008; Novembre et al., 2008). In this sense, there is not a genetic differentiation between Berber and Arab groups and a high heterogeneity among Berber groups has been shown (Arauna and Comas, 2017; Arauna et al., 2016; Bosch et al., 2000).

Many demographic processes, such as genetic drift and inbreeding, have shaped the genetic structure of those populations and have increased the heterogeneity of the region. However, admixture is the most relevant of those processes (Arauna et al., 2016). North African populations are composed by an amalgam of at least four main components: sub-Saharan, Middle Eastern, European, and autochthonous North African (Henn et al., 2012). Those components are present in the individuals of North African populations in differential frequencies, meaning that the proportion of each of them varies between populations, but also within populations, for instance within a Berber group (Arauna et al., 2016; Arredi et al., 2004; Coudray et al., 2009). Then, the differential admixture is the best explanation of the genetic landscape observed.

The main sources of admixture occurred in recent historical times, around the last 2,000 years. The main historical events that might be related to this admixture process are the Arab expansion and the sub-Saharan influence probably due to slave trade routes. The Arabization produced a complex transformation of the genetic structure, not only because of gene-flow from the Middle East, but also because it dragged gene-flow from other populations, such as sub-Saharan Africa. Moreover, the arrival of the Arabs and the Islam produced a deep sociological change that affected and disrupted the genetic structure mainly through admixture. Therefore, beyond the incoming external gene-flow, an increase in the admixture between North African populations took place in this period. Similar genetic influences might have occurred in other historical events, such as the Roman control of North Africa, although the genetic impact of substantial and more recent events might mask the detection of older processes. It is important to highlight the contact with sub-Saharan Africa and specially the slave-trade as a source of continuous gene-flow into North Africa (Arauna et al., 2016; Henn et al., 2012). Admixture and

118

population replacement already played an important role in prehistoric times, as it is attested by mitochondrial and Y chromosome studies (Arredi et al., 2004; Kujanová et al., 2009; Maca-Meyer et al., 2003). Overall, North Africa has experienced continuous gene-flow since prehistoric times that has constantly modified its gene-pool.

This pattern is also observed on a smaller scale, as has been illustrated in this thesis with the Algerian populations mtDNA study (Bekada et al., 2015). Algerian populations are a composite of the four main components in North Africa and also there is no correlation between genetics and Berber and Arab samples, or a specific genetic structure within linguistic groups. However, the focus on a local geographical area allows us to explore some sociological questions. For instance, in our study of mtDNA Algerian groups, urban populations were highly admixed and this fact decreases the genetic structure among linguistic groups.

3.2 THE FUTURE OF POPULATION GENETIC STUDIES WITH A FOCUS ON NORTH AFRICA

3.2.1 Methodology

The human population genetics field has experienced a fast growing in the last years. The number of samples and markers available has increased significantly, and therefore a large amount of data is being analyzed. In parallel with the data generation, computational methods for the analyses have been developed. Many methods are based in the analyses of allele frequencies (such as the set of F statistics' (Patterson et al., 2012; Reich et al., 2009)). However, haplotype-based methods are growing and increasing the power of previous methods.

Haplotypes are composed by a set of SNPs in linkage disequilibrium. The combination of SNPs in haplotypes allows a higher resolution than SNPs considered as separate units. The comparison of samples through their haplotypes for instance in a PCA, allows a higher differentiation of samples than a PCA that captures the allele frequency variation of SNPs separately. Moreover, haplotypes are less affected by the fixation of derived alleles through drift. Then, those methods (Hellenthal et al., 2014; Lawson et al., 2012) are more suitable for studying populations that had undergone recent demographic events as bottlenecks or founder effects.

Most part of this thesis has been focused on the analysis of haplotypes in human populations of North Africa, which has allowed to refine the demographic history of these populations. However, the complex population history of North Africa, with multiple and continuous migrations from different sources with different genetic impact, represents also a challenge for these haplotype-based methodologies. The demographic processes where more than two sources with similar genetic profiles admixed in recent times are difficult to unravel with the present methodologies. In addition, present haplotype methodologies have been proved to be very precise to detect and estimate dates of recent admixture events, but ancient admixture processes are not estimated with the same level of precision. Therefore, future haplotype-based methods assessing for more ancient events and including multiple sources in the admixture events will allowing us to refine some of the challenges found in North African populations.

3.2.2 Population datasets

The study of North African populations requires a diverse and large dataset that captures most of the North African diversity. Having a diverse dataset allows to incorporate most of the genetic heterogeneity found in North Africa and, therefore, it is a pivotal fact for analyzing the population structure. Moreover, the North African autochthonous component might be higher than sometimes estimated due to the lack of diverse samples in many genetic studies. It is frequent to include only the Mozabites or one Moroccan population as reference for the whole North African region, which is not enough to capture all the genetic diversity due to the heterogeneity of the region (Hellenthal et al., 2014; Mallick et al., 2016; Moorjani et al., 2011; Rodriguez-Flores et al., 2016). Therefore, in order the study North African neighbor populations, a large North African dataset is also important. Furthermore, North Africa is a large area (its geographic extension is almost the same of Europe), which explains part of its diversity and justify the need of building larger datasets.

Admixture analyses usually need ancestral populations as references. A good ancestral population should show no (or extremely little) admixture, low levels of inbreeding, and capture most of the population diversity. Until now, none of the North African populations sampled fulfill these characteristics. The best solution for the lack of an ancestral population is

the inclusion of many populations with little admixture, which at least increase the diversity captured in the study.

North Africa is not the only region that is not well represented in genetic datasets. The Middle East and the Arabian Peninsula are extremely important for the study of North Africa and only data from few samples are available. Those regions are inhabited by many different populations that are also very heterogeneous. Sub-Saharan Africa is the most diverse region worldwide (Henn et al., 2011; Tishkoff et al., 2009) and many times its diversity is reduced to one or two populations. The studies presented here could be refined with wider Middle Eastern and sub-Saharan datasets, which should be considered for future studies.

The studies presented in this thesis include a wide North African dataset that has allowed the understanding of the complex genetic structure. Despite the lack of a strong correlation between genetic structure and geography, we have been able to correlate genetic clusters with geographic areas, such as the Mediterranean coast. The genetic clusters are partially explained by differential admixture, together with other demographic effects, such as inbreeding. This dataset has been generated with a genotyping array platform and it is a good resource for future studies. However, array data present some limitations. Firstly, the ascertainment bias due to a non-random and European-biased selection of SNPs may affect the results. Secondly, arrays only include known and common SNPs, which challenge the analysis of populations that are not extensively studied and whose private SNPs are not included in the arrays. The populations that are underrepresented in the global panels, as it is the case of North Africa, are the ones that can show more bias in array analyses. Finally, the merge between different array platforms is sometimes a problem because the number of overlapping SNPs between the platforms is small, which reduces the power of the analyses. The merge between platforms may not only reduce the number of SNPs but also introduce a bias because of the non-random sharing of SNPs between two platforms. For instance, one genomic region could be over-represented while information for another could be missing. Sequence data might solve most of these limitations and, therefore, whole-genome sequences of North African samples would improve the dataset and the analyses. The sequence unbiased data that includes private SNPs could allow a higher resolution in the differentiation of North African

populations, as probably most of the North African internal diversity is not represented in array data.

3.2.3 Future studies

Ancient DNA (aDNA) samples would improve the North African dataset and shed light in many different aspects of North African population genetics. There are many open questions about the prehistory of the region for which aDNA data may be the only source that can provide answers: whether the origin of the ancestral North African populations is more related to Europe, the Middle East or sub-Saharan Africa; if admixture has been constant in prehistoric times or if it has been more frequent in historical times; the role of the Sahara Desert and sub-Saharan populations in North Africa during prehistory; among many others.

However, the most valuable answers that aDNA could provide might be related to population continuity. The continuity of North African populations has been questioned many times. There are plenty of studies that hypothesize different populations replacements during prehistory (Barton et al., 2008, 2013; Hublin, 2012; Irish, 2000) and at least one replacement has been genetically proved (Henn et al., 2012). Ancient DNA from different periods of time and modern samples could be compared in order to address the continuity of populations but also to go in depth into some demographic parameters. For instance, aDNA might allow to unravel if a population replacement occurred in pre-Holocean times and if this replacement was complete or some traces of the upper Paleolithic North Africans are still present in current or historical samples.

Moreover, aDNA could also solve questions regarding the origin and role of North African populations in the human early development. The higher diversity and effective population size of sub-Saharan populations has supported the sub-Saharan, particularly East African, origin of modern humans (Tishkoff et al., 2009). The comparison of North African Paleolithic samples and sub-Saharan populations could help to understand the relationships between both populations and the role of North Africa in the early human development.

Ancient DNA could also be useful in the understanding of questions regarding current genetic structure. As has been stated above, there is a lack of good North African ancestral samples among modern populations to use as reference for the North African component as one of the four

ancestral components described. This is because of the high admixture that most of the North African populations have experienced in recent times and also because some populations with lower levels of admixture show high levels of inbreeding and genetic drift, which make them not good representatives of the whole diversity of the region. Ancient samples that predate the most recent admixture events could be good ancestral references. Therefore, not only prehistoric ancient samples could be interesting, since samples from different historical periods could be good resources. For instance, the comparison of samples dated soon before and after the Arab expansion would help to understand the demographic impact of the Arabization.

An additional challenge that the human population genetics field is facing is the agreement among different types of markers. This is particularly interesting regarding uniparental markers. Sometimes the lack of agreement between uniparental markers and genome-wide data is explained because of the intrinsic characteristics of each marker. For instance, the higher drift and lack of recombination of Y chromosome and mitochondrial DNA make them more useful for tracing ancient events or to test population continuity, and may seem discordant with genome-wide results. However, mitochondrial, Y chromosome, and genome-wide markers are also revealing different histories for females and males. The incorporation and interpretation of previous information from uniparental markers studies in the light of genome-wide studies is important for a good understanding of the population genetics. Moreover, the generation of new high quality data for Y chromosome, mitochondrial, and the understudied X chromosome, could allow comparative analyses that would explain the differences and similarities in the male and female separate histories. This perspective is sometimes understudied not only from genetics but from many other disciplines, and it could help in the building of an unbiased story, especially for females whose history has been many times biased and undervalued (Hedenstierna-Jonson et al., 2017).

Most studies regarding North African populations, including those presented here, focus on the Arab-Berber dichotomy. However, this is an oversimplification of the socio-cultural landscape. Dense information about demography, sociological, cultural, linguistic, and in general anthropological characteristics is pivotal to improve our knowledge of North African history. The population genetics evolution cannot be

summarized in a simple dichotomy or in one sociological aspect such as language. This idea is applicable to all populations, but North Africa is the paradigm due to the convergence of cultures and the complex sociological and historical features. The population genetics field needs to be multidisciplinary and then to integrate different kinds of data. The collection of all possible information regarding socio-cultural, demographic and also phenotypic characteristics is important and opens a wide field of study.

In conclusion, the consideration of North Africa as a relevant population in genetic studies would improve the understanding of the modern history and the genetics of the populations, such as recent admixture inside and outwards North Africa, but also it would improve the understanding of prehistory and the early stages of human evolution.

4 REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

Amir, N., Sahnoune, M., Chikhi, L., and Atmani, D. (2015). STR-based genetic structure of the Berber population of Bejaia (Northern Algeria) and its relationships to various ethnic groups. Gene *574*, 140–148.

Arauna, L.R., and Comas, D. (2017). Genetic Heterogeneity between Berbers and Arabs. In eLS, (Chichester, UK: John Wiley & Sons, Ltd), pp. 1–7.

Arauna, L.R., Mendoza-Revilla, J., Mas-Sandoval, A., Izaabel, H., Bekada, A., Benhamamouch, S., Fadhlaoui-Zid, K., Zalloua, P., Hellenthal, G., and Comas, D. (2016). Recent historical migrations have shaped the gene pool of Arabs and Berbers in North Africa. Mol. Biol. Evol. *34*, 318–329.

Armour, J.A.L., Anttinen, T., May, C.A., Vega, E.E., Sajantila, A., Kidd, J.R., Kidd, K.K., Bertranpetit, J., Pääbo, S., and Jeffreys, A.J. (1996). Minisatellite diversity supports a recent African origin for modern humans. Nat. Genet. *13*, 154–160.

Arredi, B., Poloni, E.S., Paracchini, S., Zerjal, T., Fathallah, D.M., Makrelouf, M., Pascali, V.L., Novelletto, A., and Tyler-Smith, C. (2004). A Predominantly Neolithic Origin for Y-Chromosomal DNA Variation in North Africa. Am. J. Hum. Genet. *75*, 338–345.

Balter, M. (2006). GEOCHEMISTRY: Radiocarbon Dating's Final Frontier. Science. *313*, 1560–1563.

Balter, M. (2011). Was North Africa the Launch Pad for Modern Human Migrations? Science. *331*, 20–23.

Barich, B.E., and Garcea, E.A.A. (2008). Ecological patterns in the upper pleistocene and holocene in the Jebel Gharbi, Northern Libya: Chronology, climate and human occupation. In African Archaeological Review, pp. 87–97.

Barton, N., Bouzouggar, A., Humphrey, L., Berridge, P., Collcutt, S., Gale, R., Parfitt, S., Parker, A., Rhodes, E., and Schwenninger, J.-L. (2008). Human Burial Evidence from Hattab II Cave and the Question of Continuity in Late Pleistocene–Holocene Mortuary Practices in Northwest Africa. Cambridge Archaeol. J. *18*, 195–214.

Barton, R.N.E., Bouzouggar, A., Collcutt, S.N., Schwenninger, J.-L., and Clark-Balzan, L. (2009). OSL dating of the Aterian levels at Dar es-Soltan I (Rabat, Morocco) and implications for the dispersal of modern Homo sapiens. Quat. Sci. Rev. *28*, 1914–1931.

Barton, R.N.E., Bouzouggar, A., Hogue, J.T., Lee, S., Collcutt, S.N., and Ditchfield, P. (2013). Origins of the Iberomaurusian in NW Africa: New AMS radiocarbon dating of the Middle and Later Stone Age deposits at Taforalt Cave, Morocco. J. Hum. Evol. *65*, 266–281.

Barton, R.N.E., Bouzouggar, A., Collcutt, S.N., Carrión Marco, Y., Clark-Balzan, L., Debenham, N.C., and Morales, J. (2016). Reconsidering the MSA to LSA transition at Taforalt Cave (Morocco) in the light of new multi-proxy dating evidence. Quat. Int. *413*, 36–49.

Batini, C., Lopes, J., Behar, D.M., Calafell, F., Jorde, L.B., Van Der Veen, L., Quintana-Murci, L., Spedini, G., Destro-Bisol, G., and Comas, D. (2011). Insights into the demographic history of African pygmies from complete mitochondrial genomes. Mol. Biol. Evol. *28*, 1099–1110.

Bekada, A., Fregel, R., Cabrera, V.M., Larruga, J.M., Pestano, J., Benhamamouch, S., and González, A.M. (2013). Introducing the Algerian Mitochondrial DNA and Y-Chromosome Profiles into the North African Landscape. PLoS One *8*, e56775.

Bekada, A., Arauna, L.R., Deba, T., Calafell, F., Benhamamouch, S., and Comas, D. (2015). Genetic Heterogeneity in Algerian Human Populations. PLoS One *10*, e0138453.

Bentayebi, K., Abada, F., Ihzmad, H., and Amzazi, S. (2014). Genetic ancestry of a Moroccan population as inferred from autosomal STRs. Meta Gene *2*, 427–438.

Bosch, E., Calafell F, Peérez-Lezaun A, Comas D, Mateu E, and Bertranpetit J (1997). Population History of North Africa: Evidence from Classical Genetic Markers. Hum. Biol. *69*, 295–311.

Bosch, E., Calafell, F., Pérez-Lezaun, A., Clarimón, J., Comas, D., Mateu, E., Martínez-Arias, R., Morera, B., Brakez, Z., Akhayat, O., et al. (2000). Genetic structure of north-west Africa revealed by STR analysis. Eur. J. Hum. Genet. *8*, 360–366.

Bosch, E., Calafell, F., Comas, D., Oefner, P.J., Underhill, P.A., and 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.

Botigue, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., et al. (2013). Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. Proc. Natl. Acad. Sci. *110*, 11791–11796.

Bouzouggar, A., and Barton, R.N.E. (2012). The Identity and Timing of the Aterian in Morocco. pp. 93–105.

Bouzouggar, A., Barton, N., Vanhaeren, M., D'Errico, F., Collcutt, S., Higham, T., Hodge, E., Parfitt, S., Rhodes, E., Schwenninger, J.-L., et al. (2007). 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. Proc. Natl. Acad. Sci. *104*, 9964–9969.

Bouzouggar, A., Barton, R.N.E., Blockley, S., Bronk-Ramsey, C., Collcutt, S.N., Gale, R., Higham, T.F.G., Humphrey, L.T., Parfitt, S., Turner, E., et al. (2008). Reevaluating the Age of the Iberomaurusian in Morocco. African Archaeol. Rev. *25*, 3–19.

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

Brett, M., and Fentress, E. (1997). The Empire and the Other: Romans and Berbers. In The Berbers, (Library of Congress Cataloging-in-Publication Data), pp. 50–80.

Brisbin, A., Bryc, K., Byrnes, J., Zakharia, F., Omberg, L., Degenhardt, J., Reynolds, A., Ostrer, H., Mezey, J.G., and Bustamante, C.D. (2012). PCAdmix: Principal Components-Based Assignment of Ancestry Along Each Chromosome in Individuals with Admixed Ancestry from Two or More Populations. Hum. Biol. *84*, 343–364.

Broodbank, C. (2006). The origins and early development of Mediterranean maritime activity. J. Mediterr. Archaeol. *19*, 199–230.

Broushaki, F., Thomas, M.G., Link, V., López, S., van Dorp, L., Kirsanow, K., Hofmanová, Z., Diekmann, Y., Cassidy, L.M., Díez-del-Molino, D., et al. (2016). Early Neolithic genomes from the eastern Fertile Crescent. Science. *353*, 499–503.

Bryc, K., Auton, A., Nelson, M.R., Oksenberg, J.R., Hauser, S.L., Williams, S., Froment, A., Bodo, J.-M., Wambebe, C., Tishkoff, S.A., et al. (2010). Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc. Natl. Acad. Sci. *107*, 786–791.

Busby, G.B., Band, G., Si Le, Q., Jallow, M., Bougama, E., Mangano, V.D., Amenga-Etego, L.N., Enimil, A., Apinjoh, T., Ndila, C.M., et al. (2016). Admixture into and within sub-Saharan Africa. Elife *5*.

Busby, G.B.J., Hellenthal, G., Montinaro, F., Tofanelli, S., Bulayeva, K., Rudan, I., Zemunik, T., Hayward, C., Toncheva, D., Karachanak-Yankova, S., et al. (2015). The Role of Recent Admixture in Forming the Contemporary West Eurasian Genomic Landscape. Curr. Biol. *25*, 2518–2526.

Campbell, C.L., Palamara, P.F., Dubrovsky, M., Botigué, L.R., Fellous, M., Atzmon, G., Oddoux, C., Pearlman, A., Hao, L., Henn, B.M., et al. (2012). North

African Jewish and non-Jewish populations form distinctive, orthogonal clusters. Proc. Natl. Acad. Sci. U. S. A. *109*, 13865–13870.

Camps, G. (1995). Les Berbères : mémoire et identité (Errance).

Camps, G., and Vela i Aulesa, C. (1998). Los bereberes: de la orilla del Mediterráneo al límite meridional del Sáhara (Icaria).

Cann, H.M. (2002). A Human Genome Diversity Cell Line Panel. Science. *296*, 261b–262.

Cann, R.L., Stoneking, M., and Wilson, A.C. (1987). Mitochondrial DNA and human evolution. Nature *325*, 31–36.

Cassidy, L.M., Martiniano, R., Murphy, E.M., Teasdale, M.D., Mallory, J., Hartwell, B., and Bradley, D.G. (2016). Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. Proc. Natl. Acad. Sci. *113*, 368–373.

Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1995). The History and Geography of Human Genes (Princeton University Press).

Cherni, L., Fernandes, V., Pereira, J.B., Costa, M.D., Goios, A., Frigi, S., Yacoubi-Loueslati, B., Amor, M. Ben, Slama, A., Amorim, A., et al. (2009). Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. Am. J. Phys. Anthropol. *139*, 253–260.

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

Coudray, C., Guitard, E., Kandil, M., Harich, N., Melhaoui, M., Baali, A., Sevin, A., Moral, P., and Dugoujon, J.M. (2006). Study of GM immunoglobulin allotypic system in Berbers and Arabs from Morocco. Am. J. Hum. Biol. *18*, 23–34.

Coudray, C., Olivieri, A., Achilli, A., Pala, M., Melhaoui, M., Cherkaoui, M., El-chennawi, F., Kossmann, M., Torroni, A., and Dugoujon, J.M. (2009). The complex and diversified mitochondrial gene pool of Berber populations. Ann. Hum. Genet. *73*, 196–214.

Coulthard, T.J., Ramirez, J.A., Barton, N., Rogerson, M., and Brücher, T. (2013). Were Rivers Flowing across the Sahara During the Last Interglacial? Implications for Human Migration through Africa. PLoS One *8*, e74834.

Cruciani, F., La Fratta, R., Santolamazza, P., Sellitto, D., Pascone, R., Moral, P., Watson, E., Guida, V., Colomb, E.B., Zaharova, B., 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., La Fratta, R., Trombetta, B., Santolamazza, P., Sellitto, D., Colomb, E.B., Dugoujon, J.M., Crivellaro, F., Benincasa, T., Pascone, R., 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.

Cysouw, M., Dediu, D., and Moran, S. (2012). Comment on "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa." Science. *335*, 657–657.

d'Errico, F., Vanhaeren, M., Barton, N., Bouzouggar, A., Mienis, H., Richter, D., Hublin, J.-J., McPherron, S.P., and Lozouet, P. (2009). Out of Africa: modern human origins special feature: additional evidence on the use of personal ornaments in the Middle Paleolithic of North Africa. Proc. Natl. Acad. Sci. U. S. A. *106*, 16051–16056.

Dibble, H.L., Aldeias, V., Jacobs, Z., Olszewski, D.I., Rezek, Z., Lin, S.C., Alvarez-Fernández, E., Barshay-Szmidt, C.C., Hallett-Desguez, E., Reed, D., et al. (2013). On the industrial attributions of the Aterian and Mousterian of the Maghreb. J. Hum. Evol. *64*, 194–210.

van Dorp, L., Balding, D., Myers, S., Pagani, L., Tyler-Smith, C., Bekele, E., Tarekegn, A., Thomas, M.G., Bradman, N., and Hellenthal, G. (2015). Evidence for a Common Origin of Blacksmiths and Cultivators in the Ethiopian Ari within the Last 4500 Years: Lessons for Clustering-Based Inference. PLOS Genet. *11*, e1005397.

Drake, N.A., Blench, R.M., Armitage, S.J., Bristow, C.S., and White, K.H. (2011). Ancient watercourses and biogeography of the Sahara explain the peopling of the desert. Proc. Natl. Acad. Sci. *108*, 458–462.

Durbin, R.M., Altshuler, D.L., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Collins, F.S., De La Vega, F.M., Donnelly, P., et al. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.

Ennafaa, H., Cabrera, V.M., Abu-Amero, K.K., González, A.M., Amor, M.B., Bouhaha, R., Dzimiri, N., Elgaaïed, A.B., and Larruga, J.M. (2009). Mitochondrial DNA haplogroup H structure in North Africa. BMC Genet. *10*, 8.

Ennafaa, H., Fregel, R., Khodjet-el-khil, H., González, A.M., Mahmoudi, H.A. El, Cabrera, V.M., Larruga, J.M., and Benammar-Elgaaïed, A. (2011). Mitochondrial DNA and Y-chromosome microstructure in Tunisia. J. Hum.

Genet. *56*, 734–741.

Ennouchi, E. (1962). Un Néandertalien: L'homme du Jebel Irhoud (Maroc). Anthropologie *66*, 279–299.

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

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

Fadhlaoui-Zid, K., Haber, M., Martínez-Cruz, B., Zalloua, P., Benammar Elgaaied, A., and Comas, D. (2013). Genome-Wide and Paternal Diversity Reveal a Recent Origin of Human Populations in North Africa. PLoS One *8*, e80293.

Fire, P., and Disturbance, N. (2004). Prescribed Fire and Natural Disturbance. Science. *306*, 1680–1682.

Garcea, E.A.A. (2004). An Alternative Way Towards Food Production: The Perspective from the Libyan Sahara. J. World Prehistory *18*, 107–154.

Garcea, E.A.A. (2006). Semi-permanent foragers in semi-arid environments of North Africa. World Archaeol. *38*, 197–219.

Garcea, E.A.A., and Giraudi, C. (2006). Late Quaternary human settlement patterning in the Jebel Gharbi. J. Hum. Evol. *51*, 411–421.

González-Pérez, E., Esteban, E., Via, M., Gayà-Vidal, M., Athanasiadis, G., Dugoujon, J.M., Luna, F., Mesa, M.S., Fuster, V., Kandil, M., et al. (2010). Population relationships in the mediterranean revealed by autosomal genetic data (Alu and Alu/STR compound systems). Am. J. Phys. Anthropol. *141*, 430–439.

Haber, M., Platt, D.E., Badro, D.A., Xue, Y., El-Sibai, M., Bonab, M.A., Youhanna, S.C., Saade, S., Soria-Hernanz, D.F., Royyuru, A., et al. (2011). Influences of history, geography, and religion on genetic structure: the Maronites in Lebanon. Eur. J. Hum. Genet. *19*, 334–340.

Hajjej, A., Almawi, W.Y., Hattab, L., El-Gaaied, A., and Hmida, S. (2015). HLA class I and class II alleles and haplotypes confirm the Berber origin of the present day tunisian population. PLoS One *10*, e0136909.

Halima, A. Ben, Bahri, R., Esteban, E., Aribia, M.H. Ben, Moral, P., and Chaabani, H. (2014). Ethnic composition and genetic differentiation of the Libyan population: insights on Alu polymorphisms. Ann. Hum. Biol. *41*, 229–237.

Harich, N., Esteban, E., Chafik, A., Lopez-Alomar, A., Vona, G., and Moral, P. (2002). Classical polymorphisms in Berbers from Moyen Atlas (Morocco): Genetics, geography, and historical evidence in the Mediterranean peoples. Ann. Hum. Biol. *29*, 473–487.

Harich, N., Costa, M.D., Fernandes, V., Kandil, M., Pereira, J.B., Silva, N.M., and Pereira, L. (2010). The trans-Saharan slave trade - clues from interpolation analyses and high-resolution characterization of mitochondrial DNA lineages. BMC Evol. Biol. *10*, 138.

Hedenstierna-Jonson, C., Kjellström, A., Zachrisson, T., Krzewińska, M., Sobrado, V., Price, N., Günther, T., Jakobsson, M., Götherström, A., and Storå, J. (2017). A female Viking warrior confirmed by genomics. Am. J. Phys. Anthropol. *0*, 1–8.

Hellenthal, G., Busby, G.B.J., Band, G., Wilson, J.F., Capelli, C., Falush, D., and Myers, S. (2014). A Genetic Atlas of Human Admixture History. Science. *343*, 747–751.

Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodriguez-Botigue, L., Ramachandran, S., Hon, L., Brisbin, A., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. Proc. Natl. Acad. Sci. *108*, 5154–5162.

Henn, B.M., Botigué, L.R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J.K., Fadhlaoui-Zid, K., Zalloua, P. a, Moreno-Estrada, A., Bertranpetit, J., et al. (2012). Genomic Ancestry of North Africans Supports Back-to-Africa Migrations. PLoS Genet. *8*, e1002397.

Hernández, C.L., Dugoujon, J.M., Novelletto, A., Rodríguez, J.N., Cuesta, P., and Calderón, R. (2017). The distribution of mitochondrial DNA haplogroup H in southern Iberia indicates ancient human genetic exchanges along the western edge of the Mediterranean. BMC Genet. *18*, 46.

Hervella, M., Svensson, E.M., Alberdi, A., Günther, T., Izagirre, N., Munters, A.R., Alonso, S., Ioana, M., Ridiche, F., Soficaru, A., et al. (2016). The mitogenome of a 35,000-year-old Homo sapiens from Europe supports a Palaeolithic back-migration to Africa. Sci. Rep. *6*, 25501.

Hogue, J.T., and Barton, R.N.E. (2016). New radiocarbon dates for the earliest Later Stone Age microlithic technology in Northwest Africa. Quat. Int. *413*, 62–75.

Hublin, J.-J. (2012). Modern Origins: A North African Perspective (Springer).

Hublin, J.-J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., et al. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens.

Nature *546*, 289–292.

Humphrey, L.T., and Bocaege, E. (2008). Tooth evulsion in the Maghreb: Chronological and geographical patterns. In African Archaeological Review, pp. 109–123.

Ingman, M., Kaessmann, H., Paabo, S., and Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. Nature *408*, 708–713.

Irish, J.D. (2000). The Iberomaurusian enigma: North African progenitor or dead end? J. Hum. Evol. *39*, 393–410.

Jackes, M., and Lubell, D. (2008). Early and Middle Holocene Environments and Capsian Cultural Change: Evidence from the Télidjène Basin, Eastern Algeria. African Archaeol. Rev. *25*, 41–55.

Jacobs, Z., and Roberts, R.G. (2007). Advances in optically stimulated luminescence dating of individual grains of quartz from archeological deposits. Evol. Anthropol. *16*, 210–223.

Jorde, L.B., Rogers, A.R., Bamshad, M., Watkins, W.S., Krakowiak, P., Sung, S., Kere, J., and Harpending, H.C. (1997). Microsatellite diversity and the demographic history of modern humans. Proc. Natl. Acad. Sci. *94*, 3100–3103.

Kefi, R., Hechmi, M., Naouali, C., Jmel, H., Hsouna, S., Bouzaid, E., Abdelhak, S., Beraud-Colomb, E., and Stevanovitch, A. (2016). On the origin of Iberomaurusians: new data based on ancient mitochondrial DNA and phylogenetic analysis of Afalou and Taforalt populations. Mitochondrial DNA Part A DNA Mapping, Seq. Anal. 1–11.

Kenrick, J. (1855). Phoenicia (B. Fellowes).

Ki-Zerbo, J. (1981). General History of Africa I: Methodology and African Prehistory (Heinemann).

Kujanová, M., Pereira, L., Fernandes, V., Pereira, J.B., and Černý, V. (2009). Near eastern neolithic genetic input in a small oasis of the Egyptian Western Desert. Am. J. Phys. Anthropol. *140*, 336–346.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lao, O., Lu, T.T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A., Balascakova, M., Bertranpetit, J., Bindoff, L.A., Comas, D., et al. (2008). Correlation between Genetic and Geographic Structure in Europe. Curr. Biol. *18*, 1241–1248.

Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of Population Structure using Dense Haplotype Data. PLoS Genet. *8*, e1002453.

Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., Hutnik, K., Royrvik, E.C., Cunliffe, B., Lawson, D.J., et al. (2015). The fine-scale genetic structure of the British population. Nature *519*, 309–314.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., et al. (2008). Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. Science. *319*, 1100–1104.

Linstädter, J., Medved, I., Solich, M., and Weniger, G.C. (2012). Neolithisation process within the Alboran territory: Models and possible African impact. Quat. Int. *274*, 219–232.

Loh, P.R., Lipson, M., Patterson, N., Moorjani, P., Pickrell, J.K., Reich, D., and Berger, B. (2013). Inferring admixture histories of human populations using linkage disequilibrium. Genetics *193*, 1233–1254.

Luis, J.R., Rowold, D.J., Regueiro, M., Caeiro, B., Cinnioğlu, C., Roseman, C., Underhill, P.A., Cavalli-Sforza, L.L., and Herrera, R.J. (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., González, A.M., Pestano, J., Flores, C., Larruga, J.M., and Cabrera, V.M. (2003). Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. BMC Genet. *4*, 15.

Mahaney, W.C., Allen, C.C.R., Pentlavalli, P., Kulakova, A., Young, J.M., Dirszowsky, R.W., West, A., Kelleher, B., Jordan, S., Pulleyblank, C., et al. (2017). Biostratigraphic Evidence Relating to the Age-Old Question of Hannibal's Invasion of Italy, I: History and Geological Reconstruction. Archaeometry *59*, 164–178.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., et al. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature *538*, 201–206.

Maples, B.K., Gravel, S., Kenny, E.E., and Bustamante, C.D. (2013). RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. *93*, 278–288.

Mariotti, V., Condemi, S., and Belcastro, M.G. (2014). Iberomaurusian funerary customs: New evidence from unpublished records of the 1950s excavations of the Taforalt necropolis (Morocco). J. Archaeol. Sci. *49*, 488–499.

Matisoo-Smith, E.A., Gosling, A.L., Boocock, J., Kardailsky, O., Kurumilian, Y., Roudesli-Chebbi, S., Badre, L., Morel, J.-P., Sebaï, L.L., and Zalloua, P.A. (2016). A European Mitochondrial Haplotype Identified in Ancient Phoenician

Remains from Carthage, North Africa. PLoS One *11*, e0155046.

McDougall, I., Brown, F.H., and Fleagle, J.G. (2008). Sapropels and the age of hominins Omo I and II, Kibish, Ethiopia. J. Hum. Evol. *55*, 409–420.

McKenna, A. (2010). The history of Northern Africa (Britannica Educational Publishing).

Mendez, F.L., Krahn, T., Schrack, B., Krahn, A.-M., Veeramah, K.R., Woerner, A.E., Fomine, F.L.M., Bradman, N., Thomas, M.G., Karafet, T.M., et al. (2013). An African American Paternal Lineage Adds an Extremely Ancient Root to the Human Y Chromosome Phylogenetic Tree. Am. J. Hum. Genet. *92*, 454–459.

Mokhtar, G. (1981). General History of Africa II: ancient civilizations of Africa (United Nations Educational).

El Moncer, W., Esteban, E., Bahri, R., Gayà-Vidal, M., Carreras-Torres, R., Athanasiadis, G., Moral, P., and Chaabani, H. (2010). Mixed origin of the current Tunisian population from the analysis of Alu and Alu/STR compound systems. J. Hum. Genet. *55*, 827–833.

Montinaro, F., Busby, G.B.J., Pascali, V.L., Myers, S., Hellenthal, G., and Capelli, C. (2015). Unravelling the hidden ancestry of American admixed populations. Nat. Commun. *6*, 6596.

Moorjani, P., Patterson, N., Hirschhorn, J.N., Keinan, A., Hao, L., Atzmon, G., Burns, E., Ostrer, H., Price, A.L., and Reich, D. (2011). The history of african gene flow into Southern Europeans, Levantines, and Jews. PLoS Genet. *7*, e1001373.

Morales, J., Pérez-Jordà, G., Peña-Chocarro, L., Zapata, L., Ruíz-Alonso, M., López-Sáez, J.A., and Linstädter, J. (2013). The origins of agriculture in North-West Africa: Macro-botanical remains from Epipalaeolithic and Early Neolithic levels of Ifri Oudadane (Morocco). J. Archaeol. Sci. *40*, 2659–2669.

Mulazzani, S., Belhouchet, L., Salanova, L., Aouadi, N., Dridi, Y., Eddargach, W., Morales, J., Tombret, O., Zazzo, A., and Zoughlami, J. (2016). The emergence of the Neolithic in North Africa: A new model for the Eastern Maghreb. Quat. Int. *410*, 123–143.

Naylor, P.C. (2009). North Africa: a history from antiquity to the present (University of Texas Press).

Newman, J.L. (1995). The peopling of Africa : a geographic interpretation (Yale Univ. Press).

Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A.R., Auton, A., Indap, A., King, K.S., Bergmann, S., Nelson, M.R., et al. (2008). Genes mirror geography within Europe. Nature *456*, 98–101.

Osborne, A.H., Vance, D., Rohling, E.J., Barton, N., Rogerson, M., and Fello, N. (2008). A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. Proc. Natl. Acad. Sci. *105*, 16444–16447.

Ottoni, C., Martínez-Labarga, C., Loogväli, E.L., Pennarun, E., Achilli, A., De Angelis, F., Trucchi, E., Contini, I., Biondi, G., and Rickards, O. (2009). First genetic insight into Libyan Tuaregs: A maternal perspective. Ann. Hum. Genet. *73*, 438–448.

Ottoni, C., Primativo, G., Hooshiar Kashani, B., Achilli, A., Martínez-Labarga, C., Biondi, G., Torroni, A., and Rickards, O. (2010). Mitochondrial Haplogroup H1 in North Africa: An Early Holocene Arrival from Iberia. PLoS One *5*, e13378.

Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., Laval, G., Perry, G.H., Barreiro, L.B., Froment, A., et al. (2017). Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. Science. *356*, 543–546.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

Pennarun, E., Kivisild, T., Metspalu, E., Metspalu, M., Reisberg, T., Moisan, J.-P., Behar, D.M., Jones, S.C., and Villems, R. (2012). Divorcing the Late Upper Palaeolithic demographic histories of mtDNA haplogroups M1 and U6 in Africa. BMC Evol. Biol. *12*, 234.

Pereira, L., Silva, N.M., Franco-Duarte, R., Fernandes, V., Pereira, J.B., Costa, M.D., Martins, H., Soares, P., Behar, D.M., Richards, M.B., et al. (2010a). Population expansion in the North African Late Pleistocene signalled by mitochondrial DNA haplogroup U6. BMC Evol. Biol. *10*, 390.

Pereira, L., Cerný, V., Cerezo, M., Silva, N.M., Hájek, M., Vasíková, A., Kujanová, M., Brdicka, R., Salas, A., Černý, V., et al. (2010b). Linking the sub-Saharan and West Eurasian gene pools: maternal and paternal heritage of the Tuareg nomads from the {Africa}n Sahel. Eur J Hum Genet *18*, 915–923.

Peter Sheppard., D.L. (1990). Early Holocene Maghreb prehistory : an evolutionary approch. Sahara *3*, 63–69.

Pickrell, J.K., Patterson, N., Barbieri, C., Berthold, F., Gerlach, L., Güldemann, T., Kure, B., Mpoloka, S.W., Nakagawa, H., Naumann, C., et al. (2012). The genetic prehistory of southern Africa. Nat. Commun. *3*, 1143.

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

312–328.

Price, A.L., Tandon, A., Patterson, N., Barnes, K.C., Rafaels, N., Ruczinski, I., Beaty, T.H., Mathias, R., Reich, D., and Myers, S. (2009). Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. PLoS Genet. *5*, e1000519.

Pritchard, J., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics *155*, 945–959.

Rahmani, N. (2004). Technological and Cultural Change Among the Last Hunter-Gatherers of the Maghreb: The Capsian (10,000–6000 B.P.). J. World Prehistory *18*, 57–105.

Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009). Reconstructing Indian population history. Nature *461*, 489–494.

Rightmire, G.P., Deacon, H.J., Schwartz, J.H., and Tattersall, I. (2006). Human foot bones from Klasies River main site, South Africa. J. Hum. Evol. *50*, 96–103.

Rodriguez-Flores, J.L., Fakhro, K., Agosto-Perez, F., Ramstetter, M.D., Arbiza, L., Vincent, T.L., Robay, A., Malek, J.A., Suhre, K., Chouchane, L., et al. (2016). Indigenous Arabs are descendants of the earliest split from ancient Eurasian populations. Genome Res. *26*, 151–162.

Sánchez-Quinto, F., Botigué, L.R., Civit, S., Arenas, C., Ávila-Arcos, M.C., Bustamante, C.D., Comas, D., and Lalueza-Fox, C. (2012). North African Populations Carry the Signature of Admixture with Neandertals. PLoS One *7*, e47765.

Sankararaman, S., Sridhar, S., Kimmel, G., and Halperin, E. (2008). Estimating Local Ancestry in Admixed Populations. Am. J. Hum. Genet. *82*, 290–303.

Sari, L. (2014). Technological change in Iberomaurusian culture: The case of Tamar Hat, Rassel and Columnata lithic assemblages (Algeria). Quat. Int. *320*, 131–142.

Sarno, S., Boattini, A., Pagani, L., Sazzini, M., De Fanti, S., Quagliariello, A., Gnecchi Ruscone, G.A., Guichard, E., Ciani, G., Bortolini, E., et al. (2017). Ancient and recent admixture layers in Sicily and Southern Italy trace multiple migration routes along the Mediterranean. Sci. Rep. *7*, 1984.

Scerri, E.M.L. (2017). The North African Middle Stone Age and its place in recent human evolution. Evol. Anthropol. Issues, News, Rev. *26*, 119–135.

Scerri, E.M.L., Drake, N.A., Jennings, R., and Groucutt, H.S. (2014). Earliest evidence for the structure of Homo sapiens populations in Africa. Quat. Sci. Rev. *101*.

Schlebusch, C.M., Skoglund, P., Sjodin, P., Gattepaille, L.M., Hernandez, D.,

Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M.G.B., et al. (2012). Genomic Variation in Seven Khoe-San Groups Reveals Adaptation and Complex African History. Science. *338*, 374–379.

Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R., Steyn, M., Soodyall, H., Lombard, M., et al. (2017). Ancient genomes from southern Africa pushes modern human divergence beyond 260,000 years ago. bioRxiv 145409.

Schuenemann, V.J., Peltzer, A., Welte, B., van Pelt, W.P., Molak, M., Wang, C.-C., Furtwängler, A., Urban, C., Reiter, E., Nieselt, K., et al. (2017). Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat. Commun. *8*, 15694.

Schwartz, J.H., and Tattersall, I. (2010). Fossil evidence for the origin of Homo sapiens. Am. J. Phys. Anthropol. *143*, 94–121.

Secher, B., Fregel, R., Larruga, J.M., Cabrera, V.M., Endicott, P., Pestano, J.J., and González, A.M. (2014). The history of the North African mitochondrial DNA haplogroup U6 gene flow into the African, Eurasian and American continents. BMC Evol. Biol. *14*, 109.

Semino, O., Magri, C., Benuzzi, G., Lin, A. a, Al-Zahery, N., Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P.J., 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.

Smith, T.M., Tafforeau, P., Reid, D.J., Grun, R., Eggins, S., Boutakiout, M., and Hublin, J.-J. (2007). Earliest evidence of modern human life history in North African early Homo sapiens. Proc. Natl. Acad. Sci. *104*, 6128–6133.

Solé-Morata, N., García-Fernandez, C., Urasin, V., Bekada, A., Fadhalaoui-Zid, K., Zalloua, P., Comas, D., and Calafell, F. (in press). Whole Y-chromosome sequences reveal an extremelu recent origin of the most common North African paternal lineage E-M183 (M81).

Stoneking, M., Fontius, J.J., Clifford, S.L., Soodyall, H., Arcot, S.S., Saha, N., Jenkins, T., Tahir, M.A., Deininger, P.L., and Batzer, M.A. (1997). Alu Insertion Polymorphisms and Human Evolution : Evidence for a Larger Population Size in Africa Alu Insertion Polymorphisms and Human Evolution : Evidence for a Larger Population Size in Africa. Genome Res. *7*, 1061–1071.

Stringer, C.B. (2003). Out of Ethiopia. Nature *423*, 692–694.

Stringer, C., and Barton, N. (2008). Putting North Africa on the map of modern human origins. Evol. Anthropol. Issues, News, Rev. *17*, 5–7.

Tang, H., Peng, J., Wang, P., and Risch, N.J. (2005). Estimation of individual

admixture: Analytical and study design considerations. Genet. Epidemiol. *28*, 289–301.

The International HapMap Consortium (2005). A haplotype map of the human genome. Nature *437*, 1299–1320.

The International HapMap Project (2003). The International HapMap Project. Nature *426*, 789–796.

Tishkoff, S.A., Goldman, A., Calafell, F., Speed, W.C., Deinard, A.S., Bonne-Tamir, B., Kidd, J.R., Pakstis, A.J., Jenkins, T., and Kidd, K.K. (1998). A Global Haplotype Analysis of the Myotonic Dystrophy Locus: Implications for the Evolution of Modern Humans and for the Origin of Myotonic Dystrophy Mutations. Am. J. Hum. Genet *62*, 1389–1402.

Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.-M., Doumbo, O., et al. (2009). The Genetic Structure and History of Africans and African Americans. Science. *324*, 1035–1044.

Tishkoff, S. a, Dietzsch, E., Speed, W., Pakstis, a J., Kidd, J.R., Cheung, K., Bonné-Tamir, B., Santachiara-Benerecetti, a S., Moral, P., and Krings, M. (1996). Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science. *271,* 1380-1387.

Trombetta, B., D'Atanasio, E., Massaia, A., Ippoliti, M., Coppa, A., Candilio, F., Coia, V., Russo, G., Dugoujon, J.M., Moral, P., et al. (2015). Phylogeographic Refinement and Large Scale Genotyping of Human Y Chromosome Haplogroup E Provide New Insights into the Dispersal of Early Pastoralists in the African Continent. Genome Biol. Evol. *7*, 1940–1950.

Vallois, H.V., and Vandermeersch, B. (1975). The Mousterian skull of Qafzeh (Homo VI). An anthropological study. J. Hum. Evol. *4*, 445–455.

Venter, J.C. (2001). The Sequence of the Human Genome. Science. *291*, 1304–1351.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., and Wilson, A.C. (1991). African Populations and the Evolution of Human Mitochondrial DNA. Science. *253*, 1503–1507.

White, T.D., Asfaw, B., DeGusta, D., Gilbert, H., Richards, G.D., Suwa, G., and Clark Howell, F. (2003). Pleistocene Homo sapiens from Middle Awash, Ethiopia. Nature *423*, 742–747.

Zalloua, P.A., Xue, Y., Khalife, J., Makhoul, N., Debiane, L., Platt, D.E., Royyuru, A.K., Herrera, R.J., Hernanz, D.F.S., Blue-Smith, J., et al. (2008). Y-Chromosomal Diversity in Lebanon Is Structured by Recent Historical Events. Am. J. Hum. Genet. *82*, 873–882.