



Contents lists available at ScienceDirect

Meta Gene



Genetic ancestry of a Moroccan population as inferred from autosomal STRs



K. Bentayebi^{a,*}, F. Abada^a, H. Ihzmad^b, S. Amzazi^a

^a Laboratoire de Biochimie Immunologie, Faculté des Sciences, Université Mohammed V, Morocco

^b Laboratoire d'épidémiologie, Institut National d'Hygiène, Rabat, Morocco

ARTICLE INFO

Article history:

Received 3 March 2014

Accepted 19 April 2014

Available online 21 June 2014

Keywords:

North Africa

Morocco

Population genetics

Iberia

STR

Ancestry

Gene flow

ABSTRACT

Detecting population substructure and ancestry is a critical issue for both association studies of health behaviors and forensic genetics. Determining aspects of a population's genetic history as potential sources of substructure can aid in design of future genetic studies. Within this context, fifteen autosomal short tandem repeat (STR), were used to examine population genetic structure and hypotheses of the origin of the modern Moroccan population from individuals belonging to three different ethnical groups from Morocco (Arab, Berber and Sahrawi), by comparing their autosomal STR variation with that of neighboring and non-neighboring populations in North Africa, Europe and Middle East as well as proposed ancestral populations in Morocco (Berber). We report on the results that the gradient of North African ancestry accounts for previous observations of low levels of sharing with Near East and a substantially increased gene flow especially from Morocco and Spain.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Introduction

The question of Moroccan ancestry has been addressed for ages to answer questions of human evolutionary, historical and medical significance. Human evolutionary or anthropological studies have focused on vast array of DNA-based genetic markers, each of which has different attributes but complementary roles (Garrigan and Hammer, 2006). The nuclear genome contains the higher rate of polymorphisms. These include microsatellites that consist of tandemly repeated DNA sequences with a

* Corresponding author at: 4 Avenu Ibn Battouta, B.P 1014 RP, Rabat, Morocco. Tel.: +212 624689630; fax: +212 537774261.
E-mail address: kaoutar.bentayebi@gmail.com (K. Bentayebi).

variable number of repeats from one individual to another. Thousands of microsatellite polymorphisms have now been identified in human. Their mutation rate is much higher than that of single nucleotides, approaching 10^{-3} per generation (White et al., 2007; Zhang and Hewitt, 2003). Because of this high mutation rate, microsatellites have the potential to provide information about recent evolutionary events which made them the markers of choice in population genetics and forensic applications (El Amri et al., 2011, Chbel et al., 2003). By comparing the autosomal and X chromosome, it has been clear that the lower mutation rate, smaller population size and higher linkage disequilibrium make the X chromosome less informative than autosomes but record substantially older histories than the Y chromosome (Bentayebi et al., 2012a, 2012b). In the other hand, the mitochondrial genome as well as the Y chromosome offer a very different perspective on human evolution. The absence of recombination in these regions of the genome and their uniparental characteristics (maternal for the mtDNA and paternal for the Y-chromosomal DNA) allows researchers to infer past human behaviors and evolutionary events such as migrations, founder events, population bottlenecks or expansions, relative male and female contributions to an admixed population, marriage practices, and mode of transmission of languages (Underhill and Kivisild, 2007; Xue et al., 2008). Interestingly, a joint analysis has the potential of providing more robust information (Jorde et al., 1998), as these markers give an independent (for autosomal and X chromosome) and dependent (for mitochondrial DNA and Y chromosome) picture of sex-specific demography (Ségurel et al., 2008; Ouborg et al., 2008, White et al., 1998, Schaffner, 2004; Bentayebi et al., 2012a, 2012b). Mitochondrial DNA studies of Moroccan population, at times, provided interesting results that together with autosomal and other genetic markers will help to understand the genetic landscape of the Moroccan population. In molecular evolution, a haplogroup is a group of similar haplotypes that share a common ancestor having the same polymorphism in all haplotypes. Some haplogroups exhibit specific geographic homelands (Aboukhalid et al., 2013). Haplogroup U6 in the mtDNA phylogenetic tree descends from the haplogroup U, who lived around 55,000 years ago. Recent studies based on the distribution of mtDNA of haplogroup U6 raised two theories about the origins of modern human populations in North Africa. The first put forward that groups of the proto-U6 lineage spread from the Near East to North Africa around 40–45 ka (thousands of years ago) (Olivieri et al., 2006), followed by some degree of regional continuity, while the second proposes a westward human migration from the Near East, followed by further demographic expansion at 22 ka centered on the Maghreb and associated with a microlithic bladelet culture known as the Iberomaurusian (Maca-Meyer et al., 2003; Pereira et al., 2003). In the other hand, a recent leading evolutionary theory envisages that the first modern humans in Europe, called the Cro-Magnon (Iberomaurusian), arrived from North-west Africa and are believed to have completely replaced the previous inhabitants “the Neanderthals” (Caramelli et al., 2008). Present day genetic diversity of North African populations has been revised through a wide variety of uniparental and autosomal genetic markers.

Morocco is an appropriate region for population genetic studies and a good example for answering and completing the previous questions and theories. Located in northwest Africa with a total of 13 living languages listed by Ethnologue (Gordon, 2005), Morocco is important for studies concerning human migration, because it is bounded by the Mediterranean Sea to the North with just 14 km away from Spain across the Gibraltar strait and the Atlantic Ocean to the west, which is part of the traditionally favored model of the migratory route out of Africa for anatomically modern humans (Bilgili and Weyel, 2009). Berber-speaking populations, as representative of the autochthonous inhabitants of North Africa, inhabit a wide ranging area extended from the Saharan desert to the Atlas Mountains and nowadays combining six different countries from Egypt to Mauritania, speaking 26 sorts of Berber dialects coming all from a unique and standard Berber language (Gordon, 2005; Hayward, 2000; Louali and Philippson, 2003). In spite of their self-identification by their common language, Berbers have a complex history of invasions, conquests and migrations, especially the Arabic conquest in the seventh century AD. The Arab-Muslim conquerors adopted a language policy that enabled them to spread Arabic and Islamic cultural values (Brunschvig R, 1975). However, today in Morocco, beside Standard Arabic, Berber language remains one of the two national languages of Morocco (Gordon, 2005).

The considerable ethnic and cultural diversity within Morocco reflects differential historic population influences and make the study of existing genetic diversity of the kingdom an attractive effort. Morocco has the most complex linguistic situation compared to the three other North African countries with a remarkably unequalled repartition of population: 90% of residents settle in the north part of Morocco. Although around half of the total population speaks one of the different Berber dialects, it is quite difficult

to trace a clear line among Berber, Arabized Berber and Arabic people in Morocco (Abderrahman Z, 2013). All these historical events together with recent demographical movements have left an important imprinting in their current genetic background.

Previous genetic studies in Morocco have mainly focused on classical markers, Alu polymorphisms (Aboukhalid et al., 2013; Bentayebi et al., 2011), mitochondrial DNA (Bentayebi et al., 2012a, 2012b), X chromosome STRs and SNPs (Picornell et al., 2011; Tomas et al., 2008) and Y-chromosome STRs and SNP (Aboukhalid et al., 2009, 2010) variation in Berber and Arabic speaking samples from small and limited parts of Morocco. The present work completes the previous studies investigating a wide range of samples from different part of Morocco (Arab, Berber and Sahrawi).

Analyses of autosomal STRs have become very useful tools, both in evolutionary studies and forensic casework (Gaibar et al., 2010). In the same way, AmpFISTR Identifiler PCR amplification kit provides an easily reproducible and highly reliable method for typing 15 highly polymorphic STR loci for forensic and evolutionary purposes (Anslinger et al., 2001; Applied Biosystems, 2001). The aim of this study is to analyze STR genotypes in a Moroccan population that has special interest because of its particular geographical, cultural and historical characteristics (Fernández-Santander et al., 1999).

The present study characterizes autosomal STR genotypes from Morocco to examine population substructure and genetic relationships with other groups, including testing of the proposed genetic affinities between Morocco and populations in Southern Europe, North Africa and Middle East (Aboukhalid et al., 2009, 2010, 2013; Bentayebi et al., 2011, 2012a, 2012b; Picornell et al., 2011). We predict that if the actual Moroccan population share a common ancestor (or have experienced more recent migration and resultant gene flow) with either populations in Southern Europe or Middle East (Arab), allele frequencies of autosomal STR loci will be similar among Morocco and these proposed related populations and genetic distances between these populations will be low. Alternatively, if Morocco is an autochthonous North African population (with no recent gene flow from groups in Middle East or Europe), autosomal STR frequencies will be within the range of other populations in North Africa, and Moroccan will be more genetically similar to other North African groups with a Berber ancestry. Previous studies presenting STR data from the Moroccan population either used small samples that were often collected in urban areas of a single province to study relationships between Morocco and other populations (Chbel et al., 2003; El Ossmani et al., 2009, 2010) (Table 1). This study represents one of the most comprehensive samples of Morocco yet analyzed for autosomal STR variation, with 377 male and female throughout the country.

Material and methods

DNA samples

Blood samples and buccal swabs were collected from 320 healthy unrelated men and women (men $n = 170$, women $n = 150$) belonging to the three ethnic groups (Arab $n = 116$, Berber $n = 104$, Sahrawi $n = 100$) living in Morocco and including 145 samples already typed for 12 X-STRs (Bentayebi et al., 2012a, 2012b), were analyzed. Genomic DNA was extracted using a DNA IQ™ System (Promega, Madison, USA) according to the manufacturer's instructions, and the DNA was quantitated using a NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA), following the manufacturer's

Table 1

Size and ethnical distribution of the Moroccan population in previous studies based on autosomal STRs.

Size	Ethnicity	Location
138	Unknown ethnicity	Moroccan living in Brussel Abdin et al. (2003)
180	Arab, Berber and Saharawi	Various parts of Morocco El Amri et al. (2011)
201	Berber	Azrou (Atlas Montains) El Ossmani et al. (2010)
204	Arab	Southern Morocco El Ossmani et al. (2009)
234	Arab	Casablanca (Central Morocco) Chbel et al. (2003)
387	Arab	Rabat, Sale, Zemmour, Zaer (Central Morocco) El Ossmani et al. (2008)

recommendation. Samples were obtained, with informed consent. All participants had all four of their grandparents born in the same region. The work was approved by the Ethical Committee of the University of Mohammed V-Agdal. Information about the linguistic origins was also recorded.

Polymerase chain reaction (PCR) amplification and genotyping

PCR amplification was performed according to the manufacturer's instruction (AmpFISTR Identifier™, Applied Biosystems, Foster City, CA, USA), to amplify the 15 autosomal short tandem repeats (CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX, vWA) included in the AmpFISTR Identifier PCR Amplification Kit™. The samples were amplified in single reactions for all loci using fluorescently labeled primers. Electrophoresis was performed on an ABI Prism 3130 XL Genetic Analyser (Applied Biosystems, Foster City, CA).

Comparative population data

To test hypotheses of population origins, autosomal STR data from our sample were compared to geographically assumed-related populations. In this way, data on 13 STRs (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, VWA, TPOX, D18S51, FGA, D5S818, D16S539) in a total of 21 populations, were selected from the available literature and compared to our sample, including populations from North Africa (Moroccan Berber, Tunisian Berber, Tunisian Arabic and Libyan) (Cherni et al., 2005; El Ossmani et al., 2010; Khodjet-el-Kkil et al., 2012); Southern Europe (Italy, Spain and Greece) (Brisighelli et al., 2009; Camacho et al., 2007; Kovatsi et al., 2006); and Middle East (Palestine, Syria, Lebanon, UAE, Saudi Arabia, Oman, Yemen, Iraq, Kuwait, Egypt and Sudan) (Abdin et al., 2003; Abu Halima et al., 2009; Alenizi et al., 2008; Alshamali et al., 2005; Babiker et al., 2011; Barni et al., 2007; Chouery et al., 2010; Omran et al., 2009) and Western Asia (Pakistan, Iran, Fars) (Alshamali et al., 2005; Hedjazi et al., 2013; Rakha et al., 2009).

Statistical analysis

Arlequin software ver 3.11 (Excoffier et al., 2005) was used to calculate allele frequencies, population pairwise genetic distances (FST), AMOVA test, and also to assess departures from Hardy–Weinberg equilibrium. Pairwise FST values for each STR were calculated between 22 populations, based on 13 of the 15 studied STRs (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338). They were averaged over loci and represented in a multidimensional scaling (MDS) plot using the statistical package SPSS v.17.0 (Levesque, 2007).

Statistical parameters of forensic interest (PI, paternity index; MP, matching probability; PD, power of discrimination; PIC, polymorphic information content and PE, power of exclusion) were calculated using PowerStats v1.2 (Promega Corporation, Madison, WI) software package (Tereba, 1999). They are defined as follows (Houston, 1998):

Matching probability (MP), also known as probability of match (pM), is the number of individuals that may be surveyed before finding the same DNA pattern in a randomly selected individual. This is represented as:

$$MP = \sum_{i=a}^n \sum_{j \geq 1}^n p_{ij}^2$$

where i and j represent the frequencies of all possible alleles a through n , p_{ij} represents the frequencies of all possible genotypes.

The paternity index (PI) reflects how many more times likely it is that the person being tested is the biological father, rather than a randomly selected individual. The typical paternity index is assigned to a locus rather than an individual case. Generally, a PI of less than one is indicative of non-relatedness. The PI is represented by the following equation:

$$PI = 1/2H$$

where H is the frequency of homozygotes.

The power of discrimination is defined as the probability of discrimination between two unrelated individuals. Thus, correlated to MP discussed above:

$$PD = 1 - MP.$$

The power of exclusion (PE) is defined as the fraction of individuals having a DNA profile that is different from that of a randomly selected individual in a typical paternity case. The value for each individual case will vary. The average for a given locus is represented by the following equation:

$$PE = h^2(1 - 2hH^2)$$

where h is the heterozygosity and H is the frequency of homozygotes.

The polymorphism information content (PIC) is the probability that one could identify which homologue of a given parent was transmitted to a given offspring, the other parent being genotyped as well.

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^n \sum_{j=i+1}^n 2p_i^2 p_j^2.$$

Results

Allele frequencies of 15 autosomal STR loci (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539) as well as the values for forensic parameters are shown in Table 2. D21S11 was the most polymorphic autosomal STR, with 18 alleles, while TH01, TPOX, CSF1PO, D5S818, D7S820 and D16S539 were the least polymorphic with 6 alleles each. In particular, the most frequently observed alleles of the TH01, TPOX, CSF1PO, D5S818, D7S820 and D16S539 loci had 23.0%, 41.0%, 34.4%, 39.2%, 34.0% and 30.5% occurrences, respectively. Informativeness can be quantitatively measured by the polymorphism information content or PIC. Theoretically, PIC values can range from 0 to 1. At a PIC of 0, the marker has only one allele. At a PIC of 1, the marker would have an infinite number of alleles. A PIC value of greater than 0.7 is considered to be highly informative. Clearly markers with greater numbers of alleles tend to have higher PIC values and thus are more informative (Hildebrand et al., 1992). Using the equation for PIC, eleven autosomal markers included in the Identifier are considered to be highly polymorphic: 73% of the offspring should be informative for D3S1358, 79% for vWA, 82.2% for FGA, 81.9% for TH01, 72.5% for D7S820, 80% for D8S1179, 80.1% for D21S11, 85.5% for D18S51, 72.7% for D16S539, 81.8% for D2S1338 and 80.1% for D19S433. While only four markers are considered to be moderately polymorphic: 68% of the offspring should be informative for TPOX, 66% for CSF1PO, 69.9% for D5S818 and 68.9% for D13S317. No significant deviations from the Hardy–Weinberg equilibrium were observed in these 15 markers (Table 2).

The observed heterozygosity (H_o) ranges from 0.737 (CSF1PO) to 0.860 (FGA). The power of discrimination varies between 0.855 (CSF1PO) and 0.970 (D18S51), and the combined power of discrimination for the 15 STR loci is 0.99999999. The probability of excluding paternity (PE) varies between 0.400 (TPOX) and 0.675 (TH01), and the combined probability of excluding paternity for the 15 loci is 0.9999. The present work of the Moroccan population was conducted in order to complete the STR profiling of the different ethnic groups of the country. Previous studies were published on the Arab and Berber speaking groups (Chbel et al., 2003; El Ossmani et al., 2010) but this is the first large work that encloses both groups plus a small Sahraoui pool. All statistical values reported in this work showed no significant difference from those reported previously (Table 2).

A first approach to population diversity by locus through the F_{st} statistic values range from – 0.56% for vWA between Morocco and Libya to an impressive 11.8% for TPOX between Morocco and Pakistan. A hierarchical AMOVA, assuming four geographical groups ('Maghreb', 'South Europe', 'Middle East' and 'Western Asia'), does not reveal any significant distances between the Mediterranean populations. The frequency variance between the four groups ($F_{ct} = 0.3\%$, $P = 0.154$) being clearly lower than the diversity among populations within groups ($F_{sc} = 0.44\%$, $P = 0.068$) (Table 3).

Table 2
Allele frequencies and statistical parameters of AmpFISTR Identifier PCR Amplification Kit loci in Moroccan population.

	D3S1358	vWA	FGA	TH01	TPOX	CSF1PO	D5S818	D13S317	D7S820	D8S1179	D21S11	D18S51	D16S539	D2S1338	D19S433
5								0.002							
5.3				0.006											
6				0.183	0.020										
7		0.002		0.230	0.027	0.017		0.002	0.005						
8		0.019		0.195	0.410	0.027	0.050	0.074	0.154	0.001		0.001	0.005		
9		0.003		0.252	0.193	0.025	0.035	0.051	0.089	0.001			0.103		
9.3		0.003		0.134	0.080										
10		0.003			0.271	0.306	0.065	0.047	0.340	0.100	0.001	0.005	0.054		0.001
11		0.009				0.344	0.273	0.325	0.266	0.138		0.022	0.305	0.004	0.008
12	0.001	0.001				0.280	0.392	0.364	0.146	0.125	0.005	0.150	0.300		0.140
13	0.001	0.010					0.184	0.135		0.203	0.001	0.124	0.203	0.004	0.248
13.2											0.003				0.040
14	0.072	0.129								0.234		0.140			0.278
14.2															0.070
15	0.280	0.197								0.197	0.001	0.147		0.003	0.150
16	0.261	0.226										0.160		0.048	0.046
17	0.243	0.232	0.003									0.145		0.282	0.017
18	0.142	0.148	0.003									0.070		0.099	0.002
18.2			0.001												
19		0.020	0.044								0.001			0.138	
20			0.140								0.003	0.036		0.170	0.002
21			0.159								0.001			0.060	
22			0.210												
23			0.201											0.090	
23.2											0.001				

Table 3

Locus-by-locus analysis of molecular variance of 13 autosomal short tandem repeat loci.

Locus	Among groups			Among populations			Within populations		
	Percent variance	F _{CT}	P	Percent variance	F _{SC}	P	Percent variance	F _{ST}	P
1	0.29	0.0029	0.0010	0.0000	0.0001	0.4448	99.71	0.0029	0.0176
2	0.27	0.0027	0.00000	0.08	0.00082	0.19844	99.65	0.00348	0.00000
3	0.05	0.0005	0.52688	1.57	0.01574	0.00000	98.38	0.01619	0.00000
4	0.71	0.0071	0.04301	1.02	0.01032	0.00000	98.27	0.01732	0.00000
5	1.03	0.0103	0.04692	1.38	0.01397	0.00000	97.59	0.02415	0.00000
6	0.17	0.0017	0.03617	99.54	0.00168	0.01564	0.17	0.00457	0.00000
7	0.07	0.0007	0.20821	0.35	0.00346	0.00098	99.58	0.00415	0.00000
8	0.13	0.0013	0.08309	0.22	0.00216	0.02444	99.65	0.00348	0.00000
9	0.31	0.0031	0.00098	0.17	0.00175	0.04790	99.51	0.00485	0.00000
10	0.53	0.0053	0.00098	0.16	0.00165	0.02053	99.31	0.00691	0.00000
11	0.03	0.0003	0.45846	0.14	0.00144	0.03617	99.82	0.00177	0.01564
12	-0.02	-0.0002	0.59922	0.30	0.00298	0.00000	99.73	0.00274	0.00000
13	0.37	0.0037	0.00098	0.11	0.00111	0.09580	99.52	0.00482	0.00000
Global estimates		0.0030	0.1540		0.0044	0.0680		0.0075	0.0030
Covariance estimates			V _a = 0.0012			V _b = 0.0018			V _c = 0.3944

*Abbreviations: FCT – fixation index among groups; FSC – fixation index among populations within groups; and FST – fixation index within populations.

The allele frequencies of the 15 autosomal STR loci were compared with the exact test of population differentiation. We found no significant differences between Moroccan and Spanish populations in 10 loci ($P < 0.05$); P -values of population comparison were 0.09910 (D3S1358), 0.15315 (vWA), 0.62162 (FGA), 0.09910 (TH01), 0.00000 (TPOX), 0.37838 (CSF1PO), 0.38739 (D5S818), 0.00901 (D13S317), 0.00000 (D7S820), 0.29730 (D8S1179), 0.90090 (D21S11), 0.88288 (D18S51) and 0.62162 (D16S539). While the

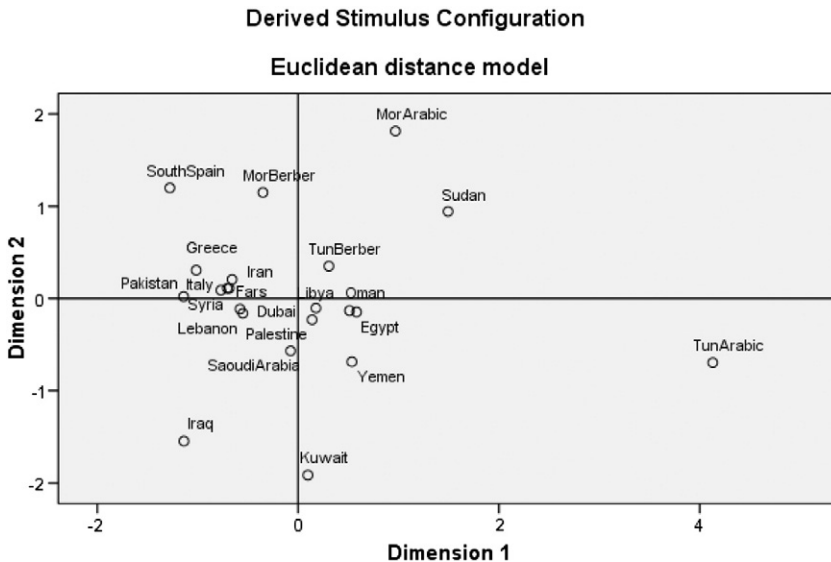


Fig. 1. Multidimensional scaling analysis (MDS) calculated between 22 population for 13 autosomal STRs. Stress: 0.14391 RSQ = 0.93846.

Iraqi population showed no significant differences with both Moroccan and Spanish populations in 9 loci. However, the Pakistani population showed significant differences from the Moroccan population ($P < 0.05$) in most loci. The MDS analysis showed that the Moroccan population form a separated plot with the Iberian Peninsula, while populations from Near East and south east Europe were clustered together. Tunisian Arabic remains isolated at the opposite side of the MDS (Fig. 1).

Discussion

Our study of the autosomal STR variation supports the hypotheses of a recent common ancestor between North Africa and Southern Europe (Morjani et al., 2011). The deviation of Arabic Tunisian was of particular concern.

This paper describes the pattern of the frequency distribution of 15 polymorphic STR of the autosomal chromosomes in the Moroccan population. In general, the allele frequencies found range within the general patterns described previously, but with a remarkable between population variations. In terms of variation, the samples were adjusted to a decreasing pattern of diversity from East to West (mean heterozygosity for Kuwait: 0.875, Syria: 0.822, Italy: 0.796 and Morocco: 0.785).

The hierarchical AMOVA analysis showed that only a small and non-significant part of the genetic variance could be attributed to the variation between East and West groups ($F_{CT} = 0.30\%$), indicating no particular genetic differentiation between both sides and arguing for a shared middle eastern ancestry. Botigué et al. have previously studied population relationships in the western and eastern Mediterranean basin using genome-wide SNP data from more than 2000 individuals. This survey indicated that Southwestern European population average between 4% and 20% of their genomes assigned to a North African ancestral whereas this value does not exceed 2% in southeastern European populations. They also found that migration event from North Africa to Europe would have occurred at least 6–10 generations ago (~240–300 ya) in Spain, and at least 5–7 generations ago in France and Italy. In the other hand, SNP data showed that the haplotype sharing between Europe and the Near East follows a southeast to southwest gradient, whereas sharing between Europe and the Maghreb follows the opposite pattern (Botigué et al., 2013; Henn et al., 2012); this suggests that gene flow from the Near East cannot account for the sharing with North Africa. It is possible that these patterns reflect more ancient migrations, perhaps dating back to the Neolithic, which resulted in a low level of short Near Eastern haplotypes across much of Europe. A model of gene flow from the Near East into both Europe and North Africa, such as a strong demic wave during the Neolithic, could result in shared haplotypes between Europe and North Africa. In other ways, studies focusing on mitochondrial DNA concluded that groups of the proto U6 lineage spread from the Near East to North Africa around 40–45 ka, followed by some degree of regional continuity. In Europe, U6 lineage was only found in the Iberian Peninsula, pointing likely to a northward expansion from Africa (Maca-Meyer et al., 2003; González et al., 2003). This expansion could be attributed to either the Arab-Berber occupation that lasted seven centuries or to prehistoric immigrations of North Africans to Iberia (Ibn Abd-el-Hakem, 1858) probably during the Capsian diffusion in North Africa and suggesting important demographic movements in the western area of this continent. For instance and in line with our results, it was found using classical genetic markers, that the Near East showed smaller genetic distances with East Africa than with West Africa (especially Arab groups) and Iberian Peninsula (Bentayebi et al., 2011; Botigué et al., 2013; Henn et al., 2012). Moreover, Cultural and religious changes within the last two millennia appear to have facilitated and maintained admixture between culturally similar populations from the Near East and North Africa. The deviation of Arabic Tunisian was of particular concern. This could be explained by the history of the Tunisian population, reflecting the influence of the ancient Phoenician settlers of Carthage followed, among others, by Roman, Byzantine, Arab and French occupations, according to historical records. Notwithstanding, other explanations cannot be discarded, such as the relative heterogeneity within current Tunisian populations, and/or the limited sub-Saharan genetic influence in this region as compared with other North African areas, without excluding the possibility of the genetic drift and the small size of the sampled population.

The total number of markers used in this study is quite small in comparison to many other available studies, but due to higher mutation rates and number of alleles per locus STRs provide much more information, on average, than SNPs for population assignment and population stratification (Listman et al., 2010). The fifteen used autosomal markers reveal a high population differentiation; this marker set may

be more useful for detecting recent admixture or founding events, such as those which formed the North African populations.

To sum up, our data on autosomal markers support, in general, the decreasing ancestry of the East-west populations described by other investigators (Henn et al., 2012; Maca-Meyer et al., 2003), providing, at the same time, detailed data of the frequency distribution of autosomal elements. These markers seem to perform well in fine-scale population differentiation studies. Also according to the statistical parameters the combined analysis of these 15 STR loci is a powerful tool for forensic identification and paternity testing in the Moroccan population.

References

- Abdin, L., Shimada, I., Brinkmann, B., Hohoff, C., 2003. Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria. *Legal Med.* 5, S150–S155 (Suppl.).
- Aboukhalid, R., Bouabdellah, M., Abbassi, M., Bentayebi, K., Elmzibri, M., Squalli, D., Amzazi, S., 2009. Haplotype frequencies for 17 Y-STR loci (AmpFISTR Y-filer) in a Moroccan population sample. *Forensic Sci. Int. Genet.* 4, e73–e74.
- Aboukhalid, R., Bouabdellah, M., Lai, R., Bradford, N., El Ossmani, H., Talbi, J., Bentayebi, K., Squalli, D., Podini, D., Amzazi, S., 2010. Y chromosomal SNP analysis using the minisequencing strategy in a Moroccan population samples. *J. Forensic Res.* 1 (2), 1000113.
- Aboukhalid, R., Sturk-Andreaggi, K., Bouabdellah, M., Squalli, D., Irwin, J.A., Amzazi, S., 2013. Mitochondrial DNA control region variation from samples of the Moroccan population. *Int. J. Legal Med.* 127 (4), 757–759 (July).
- Abu Halima, M.S., Bernal, L.P., Sharif, F.A., 2009. Genetic variation of 15 autosomal short tandem repeat (STR) loci in the Palestinian population of Gaza Strip. *Legal Med.* 11 (4), 203–204 (July).
- Alenzi, M., Goodwin, W., Ismael, S., Hadi, S., 2008. STR data for the AmpFISTR Identifier loci in Kuwaiti population. *Legal Med.* 10 (6), 321–325 (November).
- Alshamali, F., Alkhayat, A.Q., Budowle, B., Watson, N.D., 2005. STR population diversity in nine ethnic populations living in Dubai. *Forensic Sci. Int.* 152 (2–3), 267–279 (September 10).
- Anslinger, K., Rolf, B., Keil, W., 2001. Evaluation and application of the AmpF/STR profiler plus PCR amplification kit in a Bavarian population sample. *Int. J. Legal Med.* 114 (4–5), 278–280.
- Applied Biosystems, 2001. AmpFISTR Identifier™ User's Manual. Applied Biosystems, USA.
- Babiker, H.M., Schlebusch, C.M., Hassan, H.Y., Jakobsson, M., 2011. Genetic variation and population structure of Sudanese populations as indicated by 15 Identifier sequence-tagged repeat (STR) loci. *Investig. Genet.* 2 (1), 12.
- Barni, F., Berti, A., Pianese, A., Boccellino, A., Miller, M.P., Caperna, A., Lago, G., 2007. Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the middle-eastern region. *Forensic Sci. Int.* 167 (1), 87–92 (March 22).
- Bentayebi, K., Ramon, M.M., Castro, J.A., Barbaro, A., Aboukhalid, R., Amzazi, S., Picornell, A., 2011. Inferring ethnicity from the X-chromosome ALU insertions: data from Western Mediterranean human groups. *Forensic Sci. Int. Genet. Suppl. Ser.* e27–e28.
- Bentayebi, K., Picornell, A., Bouabdellah, M., Castro, J.A., Aboukhalid, R., Squalli, D., Misericordia, M.M., Amzazi, S., 2012a. Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population. *Forensic Sci. Int. Genet.* 6, e48–e49.
- Bentayebi, K., Picornell, A., Amzazi, S., 2012b. Genetic Diversity: New research: Genetic History of Morocco. Novapublishers, USA pp. 109–132 (Chapter 5).
- Bilgili, O., Weyel, S., 2009. Migration in Morocco: history, current trends and future prospects. Paper Series: Migration and Development Country Profiles. Maastricht Graduate School of Governance, p. 6.
- Botigué, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., Bertranpetit, J., Comas, D., Bustamante, C.D., 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl. Acad. Sci.* 110 (29), 11791–11796.
- Brisighelli, F., Capelli, C., Boschi, I., Garagnani, P., Lareu, M.P., Pascali, V.L., Carracedo, A., 2009. Allele frequencies of fifteen STRs in a representative sample of the Italian population. *Forensic Sci. Int. Genet.* 3 (2), e29–e30 (March).
- Brunschvig, R., 1975. Ibn Abd al-Hakam et la conquête de l'Afrique du Nord par les arabes. , 40. Túnex: J.Alloccio, Al-Andalus pp. 129–179.
- Camacho, M.V., Benito, C., Figueiras, A.M., 2007. Allelic frequencies of the 15 STR loci included in the AmpFISTR Identifier PCR Amplification Kit in an autochthonous sample from Spain. *Forensic Sci. Int.* 173 (2–3), 241–245 (December 20).
- Caramelli, D., Milani, L., Vai, S., Modi, A., Pecchioli, E., Girardi, M., Pilli, E., Lari, M., Lippi, B., Ronchitelli, A., Mallegni, F., Casoli, A., Bertorelle, G., Barbujani, G., 2008. A 28,000 years old Cro-Magnon mtDNA sequence differs from all potentially contaminating modern sequences. *PLoS ONE* 3 (7), e27000.
- Carl E., Hildebrand, Torney, D.C., Wagner, R.P., 1992. Informativeness of Polymorphic DNA Markers. , 20. Los Alamos Science, Los Alamos pp. 100–102.
- Chbel, F., Nadjifi, S., Martinez-Bouzcas, C., Louahlia, S., Azeddoug, H., Martinez De Pancorbo, M., 2003. Population genetic data of eight tetrameric short tandem repeats (STRs) in Casablanca resident population to use in forensic casework. *Forensic Sci. Int.* 132 (1), 82–83 (March 12).
- Cherni, L., Yaâcoubi, B.L., Pereira, L., Alves, C., Khodjet-El-Khil, H., El Gaaied, A.A.B., Amorim, A., 2005. Data for 15 autosomal STR markers (Powerplex 16 System) from two Tunisian populations: Kesra (Berber) and Zriba (Arab). *Forensic Sci. Int.* 147 (1), 101–106 (January).
- Chouery, E., Coble, M.D., Strouss, K.M., Saunier, J.L., Jalkh, N., Medlej-Hashim, M., Ayoub, F., Mégarbané, A., 2010. Population genetic data for 17 STR markers from Lebanon. *Legal Med.* 12 (6), 32432–32436 (November).

- El Amri, H., Squalli, D., Gazzaz, B., El Harrak, H.H., El Ossmani, H., 2011. Genetic structure of Moroccan population using 15 STRs of NGM kit. *Forensic Sci. Int. Genet. Suppl. Ser. 3* (1), e534–e535.
- El Ossmani, H., Talbi, J., Bouchrif, B., Chafik, A., 2009. Allele frequencies of 15 autosomal STR loci in the southern Morocco population with phylogenetic structure among worldwide populations. *Legal Med.* 11 (3), 155–158 (May).
- El Ossmani, H., Talbi, J., Bouchrif, B., Chafik, A., 2008. Exploitation de 15 STR autosomaux pour l'étude phylogénétique de la population Arabophone de Rabat-Salé-Zemmour-Zaer (Maroc). *Antropo.* 17, 15–23.
- El Ossmani, H., Bouchrif, B., Aboukhalid, R., Bouabdillah, M., Gazzaz, B., Zaoui, D., Chafik, A., Talbi, J., 2010. Assessment of phylogenetic structure of Berber-speaking population of Azrou using 15 STRs of Identifier kit. *Legal Med.* 12 (1), 52–56 (January).
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver.3.0: an integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* 1, 47–50.
- Fernández-Santander, A., Kandil, M., Luna, F., Esteban, E., Giménez, F., Zaoui, D., Moral, P., 1999. Genetic relationships between southeastern Spain and Morocco: new data on ABO, RH, MNSS, and DUFFY polymorphisms. *Am. J. Hum. Biol.* 11 (6), 745–752 (November).
- Gaibar, M., Esteban, E., Moral, P., Gómez-Gallego, F., Santiago, C., Bandrés, F., Luna, F., Fernández-Santander, A., 2010. STR genetic diversity in a Mediterranean population from the south of the Iberian Peninsula. *Ann. Hum. Biol.* 37 (2), 253–266 (April).
- Garrigan, D., Hammer, M.F., 2006. Reconstructing human origins in the genomic era. *Nat. Rev. Genet.* 7, 669–680.
- González, A.M., Brehm, A., Pérez, J.A., Maca-Meyer, N., Flores, C., Cabrera, V.M., 2003. Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am. J. Phys. Anthropol.* 120, 391–404.
- Gordon, R.G., 2005. *Ethnologue: Languages of the World*, Fifteenth edition. SIL International, Dallas, Texas.
- Hayward, R.J., 2000. African Languages: An Introduction. In: Heine, Bernd, Nurse, Derek (Eds.), *Afroasiatic*. Cambridge University Press, Cambridge, pp. 74–98.
- Hedjazi, A., Nikbakht, A., Hosseini, M., Hoseinzadeh, A., Hosseini, S.M., 2013. Allele frequencies for 15 autosomal STR loci in Fars province population, southwest of Iran. *Legal Med.* 15 (4), 226–228.
- Henn, B.M., Botigué, L.R., Gravel, S., Wang, W., Brisbin, A., 2012. Genomic ancestry of North Africans supports back-to-Africa migration. *PLoS Genet.* 8 (1), e1002397.
- Houston, K.A., 1998. Statistical Analysis of STR Data. Profiles in DNA, Gene Print™ pp. 14–15.
- Ibn Abd-el-Hakem, 1858. *The Islamic Conquest of Spain*. Medieval Sourcebook: Ibn Abd-el-Hakem (870 CE).
- Jorde, L.B., Bamshad, M., Rogers, A.R., 1998. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *Bioessays* 20 (2), 126–136.
- Khodjet-el-Kkil, H., Fadhlaoui-Zid, K., Gusmão, L., Alves, C., Benammar-Elgaied, A., Amorim, A., 2012. Allele frequencies for 15 autosomal STR markers in the Libyan population. *Ann. Hum. Biol.* 39 (1), 80–83 (January).
- Kovatsi, L., Parsons, T.J., Just, R.S., Irwin, J.A., 2006. Genetic variation for 15 autosomal STR loci (PowerPlex 16) in a population sample from northern Greece. *Forensic Sci. Int.* 159 (1), 61–63 (May).
- Levesque, R., 2007. *SPSS Programming and Data Management: A Guide for SPSS and SAS Users*, Fourth edition. SPSS Inc., Chicago Ill.
- Listman, J.B., Hasin, D., Kranzler, H.R., Malison, R.T., Mutirangura, A., Sughondhabirom, A., Aharonovich, E., Spivak, B., Gelernter, J., 2010. Identification of population substructure among Jews using STR markers and dependence on reference populations included. *BMC Genet.* 11 (48).
- Louali, N., Philippon, G., 2003. Protoméditerranéens Capsiens sont-ils des protoberbères? Interrogations de linguiste. *GALF (Groupement des Anthropologues de Langue Française)*, Marrakech pp. 22–25 (September).
- Maca-Meyer, N., González, A.M., Pestano, J., Flores, C., Larruga, J.M., Cabrera, V.M., 2003. Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. *BMC Genet.* 4 (15) (October 16).
- Morjani, P., Patterson, N., Hirschhorn, J.N., Keinan, A., Hao, L., Atzmon, G., Burns, E., Ostler, H., Price, A.L., Reich, D., 2011. The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* 7 (4), e1001373 (April).
- Olivieri, A., Achilli, A., Pala, M., Battaglia, V., Fornarino, S., Al-Zahery, N., Scozzari, R., Cruciani, F., Behar, D.M., Dugoujon, J.M., Coudray, C., Santachiara, A.S., Semino, O., Bandelt, H.J., Torroni, A., 2006. The mtDNA legacy of the Levantine early Upper Paleolithic in Africa. *Science* 314, 1767–1770.
- Omran, G.A., Ruttly, G.N., Jobling, M.A., 2009. Genetic variation of 15 autosomal STR loci in Upper (Southern) Egyptians. *Forensic Sci. Int. Genet.* 3 (2), e39–e44 (March).
- Ouborg, N.J., Piquot, Y., Van Groenendael, J.M., 2008. Population genetics, molecular markers and the study of dispersal in plants. *J. Ecol.* 87 (4), 551–568 (September 18).
- Pereira, L., Silva, N.M., Franco-Duarte, R., Fernandes, V., Preira, J.B., Costa, M.B., Martins, H., Soares, P., Behar, D.M., Richards, M.B., Macaulay, V., 2003. New chronometric data from Ifri n Ammar (Morocco) and the chronostratigraphy of the Middle Paleolithic in the Western Maghreb. *J. Hum. Evol.* 59, 672–679.
- Picornell, A., Bentayebi, K., Bouabdeallah, M., Castro, J.A., Bouabdillah, M., Squalli, D., Misericordia, M., Amzazi, S., 2011. X-chromosomal haplotype frequencies of four linkage groups in a North African population. *Forensic Sci. Int. Genet. Suppl. Ser.* e19–e20.
- Rakha, A., Yu, B., Hadi, S., Sheng-Bin, L., 2009. Population genetic data on 15 autosomal STRs in a Pakistani population sample. *Legal Med.* 11 (6), 305–307 (November).
- Schaffner, S.F., 2004. The X chromosome in population genetics. *Nat. Rev. Genet.* 5, 43–51 (January).
- Ségurel, L., Martínez-Cruz, B., Quintana-Murci, L., Balaresque, P., Georges, M., Hegay, T., Aldashev, A., Nasyrova, F., Jobling, M.A., Heyer, E., Vitalis, E.R., 2008. Sex-specific genetic structure and social organization in Central Asia: insights from a multi-locus study. *PLoS Genet.* 4 (9), e1000200.
- Tereba, A., 1999. Tools for analysis of population statistics. Profiles DNA, 2 pp. 14–16.
- Tomas, C., Sanchez, J.J., Barbaro, A., Brandt-Casadevall, C., Hernandez, A., Ben Dhiab, M., Ramon, M., Morling, N., 2008. X-chromosome SNP analyses in 11 human Mediterranean populations show a high overall genetic homogeneity except in North-west Africans (Moroccans). *BMC Evol. Biol.* 8 (75) (pp. 8:75 (February 29)).
- Underhill, P.A., Kivisild, T., 2007. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu. Rev. Genet.* 41, 539–564.

- White, T.L., Adams, W.T., Neale, D.B., 2007. Genetic markers – morphological, biochemical and molecular markers. Book: *Forest Genetics* pp. 53–74 (Chapter 4).
- Xue, F., Wang, Y., Xu, S., Zhang, F., Wen, B., Wu, X., Lu, M., Deka, R., Qian, J., Jin, L., 2008. A spatial analysis of genetic structure of human populations in China reveals distinct difference between maternal and paternal lineages. *Eur. J. Hum. Genet.* 16, 705–717.
- Zhang, D.X., Hewitt, G.F., 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Mol. Ecol.* 12, 563–584.