AMERICAN JOURNAL OF HUMAN BIOLOGY 28:203–212 (2016)

Original Research Article

Genetic Structure of the Kuwaiti Population Revealed by Paternal Lineages

SOUMAYA TRIKI-FENDRI,[1] PAULA SÁNCHEZ-DIZ,[2] DANEL REY-GONZÁLEZ,[2] SUAD ALFADHLI,[3] IMEN AYADI,[1] RIADH BEN MARZOUG,[1] ÁNGEL CARRACEDO,[2,4] AND AHMED REBAI[1]*
[1]*Laboratory of Molecular and Cellular Screening Processes, Centre of Biotechnology of Sfax, BP1177 Route Sidi Mansour Km 6, Sfax, Tunisia*
[2]*Forensic Genetics Unit, Institute of Forensic Science, University of Santiago De Compostela, Santiago De Compostela, Galicia, Spain*
[3]*Department of Medical Laboratory Sciences, Faculty of Allied Health Sciences, Kuwait University, Kuwait*
[4]*Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia*

Objective: We analyzed the Y-chromosome haplogroup diversity in the Kuwaiti population to gain a more complete overview of its genetic landscape.

Method: A sample of 117 males from the Kuwaiti population was studied through the analysis of 22 Y-SNPs. The results were then interpreted in conjunction with those of other populations from the Middle East, South Asia, North and East Africa, and East Europe.

Results: The analyzed markers allowed the discrimination of 19 different haplogroups with a diversity of 0.7713. J-M304 was the most frequent haplogroup in the Kuwaiti population (55.5%) followed by E-M96 (18%). They revealed a genetic homogeneity between the Kuwaiti population and those of the Middle East ($F_{ST} = 6.1\%$, P-value < 0.0001), although a significant correlation between genetic and geographic distances was found ($r = 0.41$, P-value $= 0.009$). Moreover, the nonsignificant pairwise F_{ST} genetic distances between the Kuwait population on the one hand and the Arabs of Iran and those of Sudan on the other, corroborate the hypothesis of bidirectional gene flow between Arabia and both Iran and Sudan.

Conclusion: Overall, we have revealed that the Kuwaiti population has experienced significant gene flow from neighboring populations like Saudi Arabia, Iran, and East Africa. Therefore, we have confirmed that the population of Kuwait is genetically coextensive with those of the Middle East. Am. J. Hum. Biol. 28:203–212, 2016.

Kuwait is an Arab country situated in the northeastern edge of the Arabian Peninsula. It is bordered by Saudi Arabia to the south, Iraq to the north, and the end of the Persian Gulf to the east. In 2012, the population of Kuwait was estimated to be 3,268,431 persons according to the Central Statistical office. In this census, the Kuwaitis comprised about 34.5% of the population while the rest were non-Kuwaitis including other Arabs, South Asians, and Iranians.

According to historical data, the Arabian Peninsula is considered a strategic junction linking Africa to Eurasia and thus may have been used by expanding modern human populations out of Africa (Cadenas et al., 2008; Lahr, 1994; Oppenheimer, 2003; Stringer, 2000). The original population of Kuwait is comprised of early settlers originating from the tribes of neighboring Arabian and Persian countries (particularly from Saudi Arabia and Iran) and from the nomadic Arabs of the desert, called Bedouins (Casey, 2007; Teebi, 1994). The first known settlement dates to the 3rd Century BCE, during the Greek colonization of Falaika Island, whose decline 200 years later, was, eventually, followed by the Romans' entry into the region. During the mid-7th Century CE, Kuwait came under Muslim rule. It is noteworthy that Kuwait was an important trading site for Muslim travelers, who paused there to rest from their travels.

In 1722, a severe drought struck the Arabian Peninsula leading to a famine in Saudi Arabia. Thus, many Arab groups, inter alia the Bani Utub group, including 10 extended families, were obliged to migrate northeastward to the Persian Gulf looking for better pastures for their herds. At Kuwait Bay, these immigrants settled at a small town that they named Kuwait. The name is a form of the Arabic word *Kut*, which means "a fortress near water." The Bani Utub were mixed with the autochthonous small population already present there. In 1899, Kuwait signed an agreement with the British, accepting their protection and giving them control over Kuwait's foreign affairs until 1961 (Di Piazza, 2008). Thus, we can consider that the modern Kuwaiti population began to take form at this stage.

In agreement with archeological and historical data indicating the important role of the region as a point of contact between distant populations, the Middle East displays a high degree of genetic diversity (Cavalli-Sforza et al., 1994; Cinnioğlu et al., 2004; Nasidze et al., 2004; Regueiro et al., 2006). According to previous studies of Y-haplogroup diversity, J-M304 and, particularly, its sub-haplogroups J-M267 and J-M172, are the most frequent haplogroups in the Middle East. The different geographic distribution of both sub-haplogroups J1-M267 and J2-M172 suggests two separate histories, even though they have evolved *in situ* leading to the Neolithic revolution. It is well known that J1-M267, whose maximum frequency

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare that they have no conflict of interest.

Contract grant sponsors: Ministry of Higher Education and Scientific Research, Tunisia and the Institute of Forensic Sciences "Luis Concheiro", Genomics Medicine Group of the University of Santiago de Compostela, Spain; Contract grant sponsor: Ministry of Education, Government of Spain (F.P.U. grant [Formación de Profesorado Universitario]; to D.R.G.); Contract grant number: AP2010-2616.

*Correspondence to: Ahmed Rebaï, Laboratory of Molecular and Cellular Screening Processes, Centre of Biotechnology of Sfax, BP1177 route Sidi Mansour km 6, 3018 Sfax, Tunisia. E-mail: ahmed.rebai@cbs.rnrt.tn

Received 13 May 2014; Revision received 18 June 2015; Accepted 25 July 2015

DOI: 10.1002/ajhb.22773
Published online 21 August 2015 in Wiley Online Library (wileyonlinelibrary.com).

is observed in the Southern Arabia, has been related to pastoralism in the semiarid region of Arabia, while J2-M172, mainly observed in the Fertile Crescent, has been the Y-chromosome marker for the spread of farming into Southeast Europe (Al-Zahery et al., 2011; Chiaroni et al., 2008, 2010; Semino et al., 1996, 2000, 2004).

The genetic structure of the general population of Kuwait has been described through the study of different genetic markers, particularly Y-chromosome Short Tandem Repeats (Y-STRs) (Triki-Fendri et al., 2010), autosomal STRs (Al-Enizi et al., 2013), autosomal SNPs (Alsmadi et al., 2013), and mtDNAs (Theyab et al., 2012). These studies suggest that the genetic structure of Kuwait resembles that of neighboring populations. Furthermore, the genetic diversity of autosomal STRs, Y-STRs, and Y-SNPs was investigated in six Bedouin tribes from Kuwait with these data indicating that these groups clustered with their geographical neighbors (Mohammad et al., 2009), thus adding to the strong evidence of genetic similarity of Kuwait with neighboring populations. In these Kuwaiti Bedouin populations, haplogroup J1 was observed at very high frequency reaching 84% of all the analyzed samples. To a lesser extent, other haplogroups such as R-M198 (6.75%) and E-M123 (6%) were detected in these particular Bedouin tribes.

Although many types of genetic markers have been studied in Kuwait during the last few years, Y-SNP diversity remains poorly investigated. In fact, El-Sibai et al. (2009) have analyzed the Y-SNPs in 885 samples from five different Middle Eastern populations, including 40 Kuwaitis. To obtain a more complete overview of the genetic landscape in this population, particularly in light of previous studies based on other types of markers, the present study focuses on the Y-haplogroup distribution in a sample of 117 males from the general population of Kuwait, through the analysis of 22 Y-SNPs. In addition, results from these Y-SNP analyses were interpreted in conjunction with those of other populations from the Middle East, South Asia, North and East Africa, and East Europe.

TABLE 1. Populations included in the comparative study

Group	Abreviation	Population	*N*	References
Arabian Peninsula	SA	Saudi Arabia	157	Abu-Amero et al. (2009)
	Yem	Yemen	62	Cadenas et al. (2008)
	Qt	Qatar	72	
	UAE	United Arab Emirates	164	
	Om	Oman	121	Luis et al. (2004)
Fertile Crescent	Irq	Iraq	154	Al-Zahery et al. (2011)
	Jds	Jordan (Dead Sea)	45	Flores et al. (2005)
	JA	Jordan (Amman)	97	
	Pal	Palestine	287	Zalloua et al. (2008a)
South Asia	Ind	India	728	Sengupta et al. (2006)
	Pak	Pakistan	175	
	IrA	Iranian Arabs	57	Grugni et al. (2012)
	IrP	Iranian Persians	160	
	Ir	Iran	715	
North Africa	Mar	Morocco	146	Arredi et al. (2004)
	Alg	Algeria	102	Robino et al. (2008)
	Tun	Tunisia	159	Fadhlaoui-Zid et al. (2011)
	Lib	Libya	175	Triki-Fendri et al. (2015)
	Egp	Egypt	147	Luis et al. (2004)
East Africa	SdN	Sudanese Nilo-Saharan-speaking group	53	Hassan et al. (2008)
	SdA	Arabs of Sudan	89	
	Som	Somalia	200	Sanchez et al. (2005)
	Rw	Rwanda	153	Luis et al. (2004)
	Eth	Ethiopia	124	Sengupta et al. (2006)
East Europe and Caucasus	Trk	Turkey	520	Cinnioğlu et al. (2004)
	Gce	Greece	148	Battaglia et al. (2009)
	Grg	Georgia	66	
	Crt	Croatia	118	
	Bsn	Bosnia	255	
	Alb	Albania	119	Sengupta et al. (2006)
Total			5568	

N, sample size.

MATERIALS AND METHODS

Population samples

We analyzed a total of 117 unrelated males who are natives of Kuwait (Supporting Information Fig. 1). All participants were verbally informed of this Research Project, following a research protocol authorized by the Bioethical Committee of the Galician Foundation of Genomic Medicine (GFMX-SERGAS). These samples were previously analyzed using Y-STR markers (Triki-Fendri et al., 2010). The DNA used for analyses was extracted from blood samples, using a standard phenol-chloroform procedure (Kawazaki, 1990).

Y-SNP genotyping

Y-haplogroups were determined through the analysis of 22 Y-SNPs (SRY10831, M213, M9, M22, Tat, 92R7, M173, P25, M269, M70, M96, M35, M78, M81, M123, M34, M201, M170, M26, M304, M62, and M172). The selection of these Y-SNPs was performed based on both the Y-SNP tree topology (Cruciani et al., 2010; Karafet et al., 2008; Trombetta et al., 2011) and information in the literature concerning the Y-haplogroups that have been found in the populations of the region (Abu-Amero et al., 2009; Al-Zahery et al., 2011; Cadenas et al., 2008; Mohammad et al., 2009). The selected method for allele discrimination was a single base extension reaction with the use of the SNaPshot multiplex kit (Applied Biosystems, Foster city) according to Blanco-Verea et al. (2008). To determine the haplogroup distribution in Kuwait, the Y-SNPs were hierarchically typed in three multiplex reactions as previously reported in the literature (Brión et al., 2005b,a; Brisighelli et al., 2012). These multiplexes were implemented in such a way to avoid the genotyping of unnecessary Y-SNPs to define the final haplogroup, saving effort and cost, allowing us to genotype one single multiplex in the best case and two in the worst case.

Statistical analysis

The haplogroup frequencies were estimated by direct counting. Arlequin software version 3.5 (Excoffier et al., 2007) was used to perform the Analysis of Molecular Variance (AMOVA) and to calculate the pairwise F_{ST} genetic distances between Kuwaitis and 30 populations from the Middle East, South Asia, North and East Africa, East Europe, and Caucasus region (Table 1). The Y-SNPs used in this work were the same as those used for comparisons with other populations. Genetic distances were visualized in two-dimensional space using the multidimensional

scaling (MDS) method included in the Statistical Package for the Social Sciences (SPSS) software 17.0 (SPSS Inc.).

The R environment for statistical computing and graphics (v2.15.2) was used to perform the Principal Component Analysis (PCA) based on haplogroup frequencies as well as the Mantel tests, to check for correlation between geographical and genetic distances. A Bonferroni correction for all multiple testing was applied (Hochberg, 1988). Data from 17 Y-STR haplotypes, typed in a previous study with the same samples (Triki-Fendri et al., 2013), were used to understand the genetic relationships between Y-STR haplotypes within specific SNP-haplogroups using Network 4.6.1.0 software available at (http://www.fluxusengineering.com) (Bandelt et al., 1999). The mutation-based weighting scheme proposed by Muzzio et al. (2010) was applied. The network with high-dimensional cubes was resolved using the reduced median algorithm to generate a file (*.rmf) and then applying the median joining network method to this file. For network construction, the number of repeats at DYS389II was calculated after subtracting the number of repeats at DYS389I. DYS385 was not considered for statistical analysis.

Demographic inferences

Bayesian Analysis of Trees With Internal Node generation (BATWING) (Wilson and Balding, 1998) was used to estimate the posterior distributions of (1) the time to the most recent common ancestor (TMRCA), (2) the effective population size (N), (3) the expansion time (β), and (4) the growth rate per generation (α) in the Kuwaiti population. BATWING generates TMRCAs with median values and 95% confidence intervals around those estimates of the major haplogroup in Kuwait as well as in some nearby populations. We used a population model of exponential growth from an initially constant-sized population of size N with growth starting at time *beta* before present. In this analysis, 10 Y-STR loci were used (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, and DYS439), and samples containing duplicated and missing alleles were avoided. In fact, it has been proposed that different sets of STRs behave differently and that any dating might be affected by which STRs are used (Busby et al., 2012). Populations were analyzed individually using weakly informative prior distributions for N, the effective population size before expansion [gamma (1, 0.0001): mean = 10.000, SD = 10.000]; α, the rate of growth per generation [gamma (2, 400): mean = 0.005, SD = 0.0035]; and β, the time in coalescent units when exponential growth began [gamma (2, 1): mean = 2, SD = 1.41]. However, we constructed a per-locus mutation rate prior distribution μ as [gamma (4, 1667): mean = 0.0024, SD = 0.0016] (Contu et al., 2008; Goedbloed et al., 2009; Shi et al., 2010; Wilson et al., 2003; Xue et al., 2006). Thus, the given 95% confidence intervals take into account uncertainty about the mutation rates, effective population size, expansion time, and population growth but not generation time which was fixed at 25 years. A total of 10^5 MCMC cycles were taken after discarding the first 3×10^3 samples as "burn-in." Convergence was confirmed by examining longer runs of 10^6 and 10^7 MCMC cycles for some haplogroups and finding the same posterior distribution. Out-files were postprocessed with the R environment.

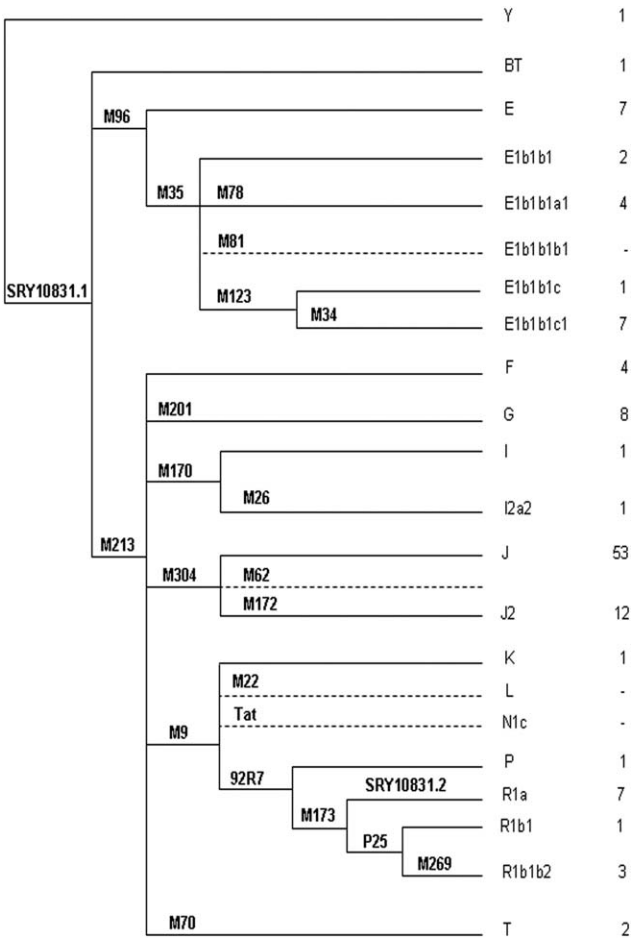

Fig. 1. Phylogenetic tree of Y-chromosome haplogroups searched in this study. The analyzed Y-SNPs are shown in each branch and the corresponding haplogroup and observed absolute frequencies are shown at the end of each branch according to Karafet et al. (2008), Cruciani et al. (2010), and Trombetta et al. (2011). *Broken line branches* correspond to haplogroups not found in our sample and *solid line branches* to those found here.

RESULTS

A sample of 117 unrelated males from the general population of Kuwait (Supporting Information Fig. 1), previously studied for 12 Y-STRs (Triki-Fendri et al., 2010), was genotyped for a set of 22 Y-SNPs. Both the SNP and STR data are shown in Supporting Information Table 1. These data were submitted to YHRD (http://www.yhrd.org/) and received the following accession number: YA003483.

Y-SNP haplogroup variation

The 22 Y-SNPs analyzed in this work allowed the discrimination of 19 different haplogroups and subhaplogroups represented in Figure 1, along with their respective frequencies in the Kuwaiti samples. The most common haplogroup in Kuwait was J-M304 (55.6%) followed by E-M96 (18%), R-M173 (9.5%), and finally G-M201 (7%). The other haplogroups occurred with lower frequencies, such as F-M213 (3.5%) and T-M70 (1.7%).

A haplogroup diversity of 0.7713 ± 0.026 was found in Kuwait based on the Y-SNP genotyping, which is higher

TABLE 2. AMOVA results

Source of variation	1	2	3	4	5
Among groups	–	0.4	–	24	13.6
Among populations within group	1.3	3.9	6.1	4.5	8.3
Within populations	98.7	95.8	93.9	71	78.1
F_{ST}	1.3*	4.2*	6.1*	29*	21.9*

1, Arabian Peninsula; 2, Arabian Peninsula—Fertile Crescent; 3, Middle East: Arabian Peninsula—Fertile Crescent—Iran–Turkey–Egypt; 4, Middle East and North Africa (without Egypt); 5, Middle East, South Asia, East Europe, North Africa and East Africa.
*P-value < 0.0001.

than those previously observed in five different Kuwaiti Bedouin tribes ranging from 0 in Ajman to 0.6 in Aniza, with a global haplogroup diversity of 0.3569 (Mohammad et al., 2009). However, the haplogroup diversity in Iran and Saudi Arabia were higher than that of Kuwait: 0.9419 and 0.8165, respectively (Abu-Amero et al., 2009; Regueiro et al., 2006). This indicates a higher diversity in heterogeneous populations.

Phylogeographic analysis

Haplogroup J. In agreement with the other populations of the Arabian Peninsula, haplogroup J-M304 is the most common lineage in Kuwait (55.6%), the subhaplogroup (J-M304(xM62,M172) (45.3%), followed by J2-M172 (10.2%).

A network of haplotype data was constructed to assess the relationship among J-M304(xM62,M172) Y-STR lineages between Kuwaitis, Yemenis (Cadenas et al., 2008), and Iranians (Grugni et al., 2012). The resulting network (Supporting Information Fig. 2A) is star-like, with a center occupied by the most frequent haplotype in the dataset of this analysis and from which the remaining haplotypes were derived by additional mutations. This central haplotype (DYS19*14; DYS389I*13; DYS389II*30; DYS390*23; DYS391*11; DYS392*11; DYS439*11) was shared by 15 Kuwaitis and three Yemenis having the same haplogroup J-M304(xM62,M172). It is noteworthy that this haplotype was observed in 12.8% of all of the analyzed Kuwaiti samples. In addition, according to the YHRD database, it has been commonly observed in the Afro-Asiatic speaking populations, such as those from Saudi Arabia (19%), the UAE (8%), Tunisia (10.5%), and Libya (7%).

Supporting Information Figure 2B illustrates the haplotype network of haplogroup J2-M172 in four populations: Kuwaitis (present work), Palestinians (Fertile Crescent) (Zalloua et al., 2008a), Libyans (North Africa) (Triki-Fendri et al., 2015), and Ukrainians (East Europe) (Mielnik-Sikorska et al., 2013). The center of the network is made up, mainly, of Palestinian and Kuwaiti haplotypes, from which derive the remaining haplotypes by additional mutations, whereas the Libyan and Ukrainian nodes are situated at the end of the branches.

Haplogroup E. In the Kuwaiti population, 18% of the males belong to haplogroup E-M96, among which two-thirds have the M35 mutation (12% of all of the samples), which defines E-M35 or E1b1b1*.

To illustrate the genetic relationships between haplotypes inside haplogroup E-M35, we constructed a network with Y-STR haplotype data from Kuwait (present work), Ethiopia (de Filippo et al., 2011) and Iran (Regueiro et al., 2006). The obtained network (Supporting Information Fig. 2C) indicated that the most common haplotype, shared by nine Ethiopian individuals, was represented by the central node of the network and from which derived all of the Kuwaiti and Iranian nodes.

Furthermore, among the E-M35 Kuwaiti Y-chromosomes, 57% belonged to the subhaplogroup E-M123 (particularly E-M34) (6.8% of all of the samples), whereas 14% belonged to E-M78 (3.4% of all of the samples). E-M123was observed in Ethiopia (11.2%), in the Middle East (3.7%), in Europe (1.7%), and in North Africa (0.9%) (Cruciani et al., 2004). In the analyzed Kuwaiti samples, we found seven chromosomes carrying the M34 mutation among the eight samples belonging to the E-M123 branch. Among the seven Kuwaiti samples belonging to haplogroup E-M78, only one individual carried both of the rare STR alleles DYS19*11 and DYS392*12, typical of cluster ɤ (Supporting Information Fig. 2D).

Other haplogroups. In the Kuwaiti population, 9.4% of the samples belonged to Haplogroup R-M173, among which 63% were from the R1a1-SRY10831.2 subhaplogroup (6% of all of the population), with the remaining samples belonging to R1b1-P25. Indeed, approximately 7% of Kuwaiti males were assigned to haplogroup G-M201. This frequency is higher than that found in the neighboring populations for the same haplogroup: 4.54% in Iraq, 4.2% in the UAE, 3.2% in Saudi Arabia, 2.7% in Qatar, and 1.6% in Yemen (Abu-Amero et al., 2009; Al-Zahery et al., 2011; Cadenas et al., 2008).

Population relationships and genetic structure

We conducted a comparative study of haplogroup frequency distribution based on the Y-SNP data of Kuwaitis and those of 30 other populations geographically organized into six groups: Arabian Peninsula, Fertile Crescent, South Asia, North Africa, East Africa, and finally East Europe and the Caucasus (Table 1).

Analysis of molecular variance

Table 2 displays the AMOVA results (Arlequin v 3.5) based on haplogroup frequencies in all comparative populations. When we considered the 31 populations organized in six groups, highly significant variance was observed (F_{ST} = 21.9%, P-value < 0.0001), indicating the genetic divergence of the selected populations. The highest fraction of the variability, as expected, resided within populations but there was a substantial percentage due to differences among groups (13.6%) and among populations within groups (8.3%).

However, a small value of F_{ST} was obtained when the AMOVA was applied to one group made up of the populations of Arabia (F_{ST} = 1.3%, P-value < 0.0001) (Table 2). In addition, the F_{ST} obtained when considering either the populations of Arabia and the Fertile Crescent (F_{ST} = 4.2%, P-value < 0.0001) or those of the Middle East (Arabian Peninsula, Fertile Crescent, Iran, Turkey, and Egypt) (F_{ST} = 6.1%, P-value < 0.0001) reflected the fact that all of the populations of the Middle East share the same genetic pool.

Moreover, 73% of the pairwise F_{ST} genetic distances between the six considered populations of the Arabian Peninsula were not significant (Supporting Information Table 2). In particular, no significant differences were observed between Kuwaitis and the remaining populations of this region, including Saudi Arabia (F_{ST} = −0.002,

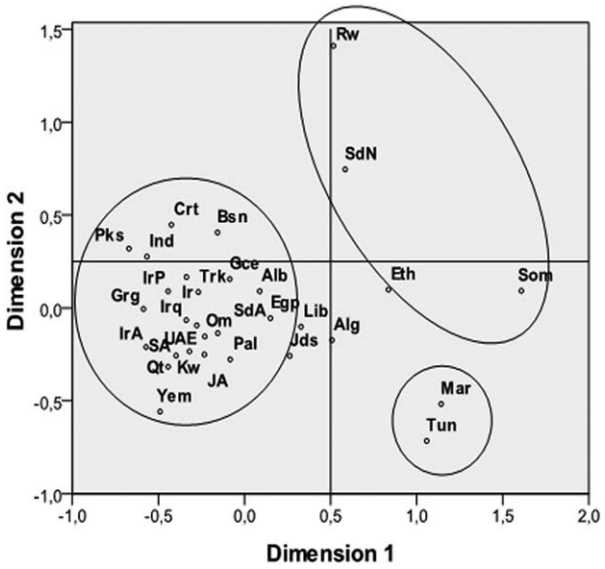

Fig. 2. Multidimensional Scaling (MDS). MDS was constructed using pairwise F_{ST} genetic distances based on Y-haplogroup frequency data from the populations described in Table 1 (the abbreviations are mentioned in Table 1) (Stress value = 0.051 and DAF = 0.979)

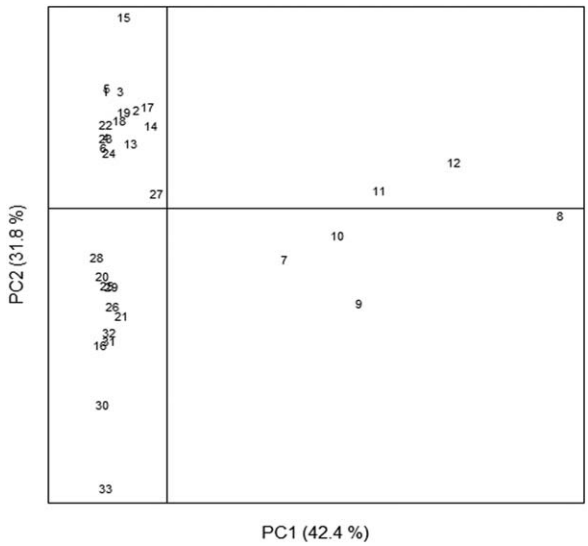

Fig. 3. Principal Component Analysis (PCA). PCA was constructed based on the Y-chromosome haplogroup frequency data of all the populations of Table 1. The populations were coded as follows: 1, Albania; 2, Bosnia; 3, Croatia; 4, Georgia; 5, Greece; 6, Turkey; 7, Libya; 8, Tunisian Berbers; 9, Tunisian Arabs; 10, Algeria; 11, Moroccan Arabs; 12, Moroccan Berbers; 13, Ethiopia; 14, Rwanda; 15, Somalia; 16, Sudanese Arabs; 17, Sudanese Nilo-Saharan-speaking group; 18, Pakistan; 19, India; 20, Palestine; 21, Jordan (Amman); 22, Jordan (Dead Sea); 23, other Iranians; 24, Persians; 25, Iranian Arabs; 26, Oman; 27, Egypt; 28, Iraq; 29, UAE; 30, Qatar; 31, Kuwait; 32, Saudi Arabia; 33, Yemen.

P-value = 0.513), Yemen (F_{ST} = 0.034, P-value = 0.009), Oman (F_{ST} = 0.0018, P-value = 0.306), and the UAE (F_{ST} = 0.002, P-value = 0.22). In addition, no significant differences were obtained between Kuwaitis and the populations of the Fertile Crescent (Iraq, Palestine, and Jordan [Amman]), excluding the population of the Dead Sea from Jordan (F_{ST} = 0.188, P-value < 0.00001).

We have to note that the Iranian Arabs were genetically similar to the Kuwaitis (F_{ST} = 0.022, P-value = 0.018) and to four other Arab-speaking populations, Saudi Arabia, Qatar, Iraq, and Jordan. In the same way, we have found that the Arabs of Sudan, unlike the Sudanese Nilo-Saharan-speaking group from the same population, are not significantly different from Kuwaitis (F_{ST} = 0.02, P-value = 0.018). Nevertheless, a high genetic variance was found between the populations of the Middle East and those of North Africa (F_{ST} = 29%, P-value < 0.00001) with 24% of variation occurring between groups.

Multidimensional scaling. Pairwise F_{ST} genetic distances (Supporting Information Table 2) were used to generate an MDS plot to assess the relationships between the populations of Table 1. Three clusters were obtained in the MDS plot (Fig. 2). First, Kuwaitis were found to be close to the populations of the Arabian Peninsula, and more generally, to all of the populations of the Middle East. In addition, East European populations were associated with this cluster. Second, North African populations clustered together. Finally, the third cluster in the MDS plot included the East African populations like Somalia, Ethiopia, Rwanda, and the Sudanese Nilo-Saharan-speaking group. It is noteworthy that the Arabs of Sudan belong to the Middle Eastern cluster instead of that of East Africa.

Principal component analysis. PCA was used to compare the haplogroup distribution between all the populations included in the comparative study (Table1). A considerable geographic structuring was obtained with the first two components accounting together for 74.2% of all of the genetic variance (Fig. 3).

Mantel test. We performed the Mantel test to assess the correlation between genetic and geographic distances. No correlation was observed when we considered the populations of the Arabian Peninsula (r = 0.46, P-value = 0.16, with 1,000 iterations). Nevertheless, we found a significant correlation between the genetic and geographic distances with all of the populations of the Middle East, including Iran, Turkey, and Egypt (r = 0.41, P-value = 0.009, with 1,000 iterations).

BATWING analysis results

BATWING was applied to all of the Y-STR haplotypes of the analyzed Kuwaiti population. We found a median TMRCA of 23.5 kya (95% CI 9.7–81.6) with an effective population size at the beginning of the expansion of 445 (95% CI 141–1,531), an expansion time of 14 kya (95% CI 5.8–49.4), and a growth rate per generation of 0.0051 (95% CI 0.0015–0.0122).

In addition, we have estimated the TMRCAs of the main haplogroups, based on haplotype data (Table 3). To evaluate these values, we calculated the TMRCA of R1-M173 and G-M201 in Iran (Haber et al., 2011) and found a median value of 53.3 kya [95% CI 18.7–184.7] and 15.9 kya [95% CI 5.5–60.2], respectively, and the TMRCA of haplogroup J1-M267 in Saudi Arabia (Abu-Amero et al., 2009) was about 19.5 kya [95% CI 6.5–77]. All of these haplogroup time estimates are higher than those observed

TABLE 3. TMRCA values of the most important haplogroups in Kuwait provided by BATWING analysis

Haplogroup	TMRCA[a]		
	Mean	Median	95% CI
E-M35	16.5	12.8	4.1–51.3
J(xJ1a,J2)-M304	19	14.3	4.5–60
J2-M172	11.8	9.3	3–35.3
R-M173	23.7	18.9	6.6–70
G-M201	16.4	12.7	4.2–50

[a]Estimates based on 10 Y-STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, and DYS437) inside of specific haplogroups.
CI, confidence interval.
Time is expressed in kya (thousands of years ago).
Generation time = 25 years.

in the Kuwaiti population. To the contrary, TMRCA of haplogroup E-M35 in Ethiopians (Sengupta et al., 2006) (12 kya [95% CI 4.5–50.5]) was almost similar to that of Kuwaitis (12.8 kya [95% CI 4.1–51.3]).

DISCUSSION

A sample of 117 unrelated males from the general population of Kuwait, was genotyped for a set of 22 Y-SNPs. The most common haplogroup in Kuwait is J-M304 (55.6%) followed by E-M96 (18%), R-M173 (9.5%), and finally G-M201 (7%).

Haplogroup J has been postulated to have a Middle Eastern origin (Semino et al., 2004). It has been considered to represent the signature of the Neolithic demic diffusion associated with the spread of agriculture and domestication of animals (Semino et al., 1996, 2004; Soares et al., 2010). The frequency distribution of haplogroup J exhibits a radial decline from the Middle East toward Central Asia, Caucasus, North Africa, and Europe (Al-Zahery et al., 2011; Cadenas et al., 2008; Quintana-Murci et al., 2001). The Southern Arabian Peninsula has been proposed to be the place of origin of haplogroup J. In fact, the distribution of this haplogroup agrees with the first major expansions proposed from GM marker analysis (Chaabani, 2002) and that of mtDNA haplogroup R0a (Cerny et al., 2011) providing evidence for its deep genetic root in Southern Arabia (Chaabani et al., 2000).

Based on the frequencies of both subhaplogroups J1 and J2, the mutation leading to the subclade J1 has, very likely, happened in Yemen (J1 frequency is 72.% vs. 9.6% for J2). By contrast, J2 would have arisen outside of Yemen, particularly in the northern region such as Syria (33.6% for J1 vs. 20% for J2), Turkey (9.1% vs. 24.2%), or Greece (1.9% vs. 18.1%) (Al-Zahery et al., 2011; Cadenas et al., 2008; Chaabani, 2000; Cinnioğlu et al., 2004; Luis et al., 2004; Semino et al., 2000, 2004).

On the one hand, previous studies on haplogroup J1-M267, which is the most frequent haplogroup in Kuwait (45.3%), have revealed a decreasing frequency moving from Southern Arabia northward: Yemen (72.6%), Qatar (58.3%), Iraq (56.4%), Oman (38%), Egypt (20%), Lebanon (12.5%), and Turkey (9%) (Al-Zahery et al., 2011; Cadenas et al., 2008; Cinnioğlu et al., 2004; Luis et al., 2004; Semino et al., 2000). Moreover, the Network constructed on the basis of haplotypic data belonging to haplogroup J-M304(xM62,M172) (Supporting Information Fig. 2A) showed a central node (DYS19*14; DYS389I*13; DYS389II*30; DYS390*23; DYS391*11; DYS392*11; DYS439*11) shared by 15 Kuwaitis and 3 Yemenis. This common haplotype, associated with haplogroup J-M304(xM62,M172), could be considered an ancestral haplotype originating in Arabic populations, particularly those of Southern Arabia. Indeed, the higher TMRCA estimate of haplogroup J1-M267 in Saudi Arabians compared with that observed in the Kuwaitis is concordant with historical data of gene flow from Saudi Arabia toward Kuwait.

Moreover, the Iranian nodes of the network (Supporting Information Fig. 2A) were located at the end of the branches, suggesting that the haplotypes belonging to haplogroup J-M304(xM62,M172) in Iran was derived from the typically Arabic one. This hypothesis is, generally, concordant with historical patterns of gene flow, particularly involving haplogroup J1-M267, of Arabic populations toward Iran (Grugni et al., 2012).

On the other hand, haplogroup J2-M172 was observed in 10.2% of the Kuwaiti samples. This haplogroup is highly represented in the Middle East (16% in the population of Saudi Arabia (Abu-Amero et al., 2009), and 23% in the population of Iran (Regueiro et al., 2006), particularly in the Fertile Crescent [25% in the Lebanese population (Zalloua et al., 2008b) and 24% in that of Turkey (Cinnioğlu et al., 2004)]. In fact, this region has been proposed as the most likely origin of haplogroup J2-M172, which was then spread to Europe during the Neolithic (Cinnioğlu et al., 2004; Semino et al., 2000, 2004). In addition, the network representing the relationship between haplotypes carrying haplogroup J2-M172 in four populations (Kuwaitis, Palestinians, Libyans, and Ukrainians) (Supporting Information Fig. 2B) illustrated that the central nodes are mainly, haplotypes from Palestinians and to a lesser extent from Kuwaitis. This finding is concordant with the hypothesis of a recent migration of this haplogroup from the Middle East and particularly from the Fertile Crescent toward North Africa and Europe.

In addition, 12% of all of the analyzed Kuwaiti samples carry the haplogroup E-M35 (E1b1a1*). It has been proposed that the origin of haplogroup E-M35 could most likely be Eastern Sub-Saharan Africa, where the highest frequencies of this subhaplogroup are observed: 19.2% in the Ethiopian population and 16.7% among the Khoisans of South Africa (Cruciani et al., 2002, 2004; Underhill et al., 2000). From there, this lineage could have been spread to the Middle East (5% in Iran) (Regueiro et al., 2006) and North Africa at the end of the Pleistocene (Underhill et al., 2001). During the Neolithic expansion, E-M35 would have been introduced in South Europe through the Middle East (Hammer et al., 1998; Semino et al., 2000; Underhill et al., 2001). The obtained network in Supporting Information Figure 2C reinforces the hypothesized East African origin of the E-M35 lineage, as the central node was shared by nine Ethiopian samples and from which derive all of the other Kuwaiti and Iranian haplotypes belonging to haplogroup E-M35.

Furthermore, we have reported that most of the Kuwaiti chromosomes sharing the haplogroup E-M123 are also carriers of the E-M34 mutation. This finding is in agreement with previous work on 3,401 samples (Cruciani et al., 2004). The Middle East has been proposed as the most likely origin for this lineage rather than East Africa. In fact, E-M34 seems to be restricted to Ethiopia, as it has not been detected in other populations in the region such as Somalia, Kenya (Cruciani et al., 2004), or Sudan (Underhill

et al., 2000). By contrast, E-M34 chromosomes have been found in a large majority of the populations from the Middle East, including Jordan (13%) (El-Sibai et al., 2009), Oman (12%) (Luis et al., 2004), Yemen (8%) (Cadenas et al., 2008), Turkey (5%) (Cinnioğlu et al., 2004), and Lebanon (4.2%) (Zalloua et al., 2008b). Thus, E-M34 chromosomes would have been introduced into Ethiopia from the Middle East (Cruciani et al., 2004).

Moreover, E-M78 carried by 3.4% of the analyzed Kuwaiti individuals, was observed over a wide area, including East Africa (21.5%), North Africa (18.5%), Middle East (5.8%), and Europe (7.2%). The Southern Mediterranean has been strongly imprinted by this haplogroup (Hammer et al., 1998; Semino et al., 2000, 2004). And the high frequency of E-M78, added to its high microsatellite diversity in Eastern Africa, has led to the hypothesis that it originated from this region, 23.2 ky ago (95% CI 21.1–25.4 ky) (Cruciani et al., 2004, 2007). The cluster ɤ, a particular branch of E-M78, defined by the presence of the rare Y-STR allele 11 in DYS19, is observed in Eastern Africa at an average frequency of 17.7% and is almost absent in the populations outside the Horn of Africa (Cruciani et al., 2004; Sanchez et al., 2005). In Kuwait, it was observed in 0.8% of the samples. This cluster has been reported to be associated with the unusual DYS392*12 allele (Semino et al., 2004).

Based on time estimation with BATWING, the median TMRCA obtained for haplogroup E-M35 in Ethiopia (12 kya [95% CI 4.5–50.5]) was almost similar to that of Kuwait (12.8 kya [95% CI 4.1–51.3]). In fact, it has already been proposed that the likely origin of the subhaplogroup E-M123 could be the Middle East (Cruciani et al., 2004), whereas the second subhaplogroup E-M78 was proposed to be issued from North-East Africa (Cruciani et al., 2007). Thus, the similarity between TMRCAs of E-M35 in Kuwait and Ethiopia suggests a bidirectional gene flow between both regions, concordant with historical data on people movements via Bab-el-Mandeb Strait (Chaabani, 2014; Lahr, 1994; Oppenheimer, 2003; Stringer, 2000).

Haplogroup R, observed in about 10% of the Kuwaiti samples, might reflect the impact of expansion and migration of Indo-Europeans from Central Asia (Sengupta et al., 2006). Based on the high frequency of R1-M173* and R1a-SRY1532 lineages in Iran compared with their corresponding frequencies in Turkey, Pakistan, or India, it has been proposed that the geographic origin of haplogroup R would, probably, be nearer to Persia (Regueiro et al., 2006). R-M173 would, probably, have arisen originally in this region, and subsequently, individuals bearing this lineage migrated toward Europe (Wells et al., 2001). Overall, the majority of human Y chromosomes in Europe belong to haplogroup R, and most of them carry the M173 mutation (Chiaroni et al., 2009; Jobling and Tyler-Smith, 2003). In the Middle East, R1-M173 was observed in Iraq (19.4%) (Al-Zahery et al., 2011), Southern Arabian Peninsula (11.2%) (Cadenas et al., 2008), Iran (25.6%) (Regueiro et al., 2006), and Pakistan (46.2%) (Sengupta et al., 2006). It is noteworthy that this haplogroup is represented in the Bedouin tribes of Kuwait with a frequency of 8.7% (Mohammad et al., 2009).

About 7% of the Kuwaiti males were assigned to haplogroup G-M201. The Middle East, Caucasus, and Southern Europe delineate the phylogeographic region of haplogroup G (Rootsi et al., 2012). Most of the chromosomes in the Caucasus carry haplogroup G with a frequency ranging between 13% in Iran (Regueiro et al., 2006) and 70% in the population of North Ossetia (Balanovsky et al., 2011; Yunusbayev et al., 2012).

Rootsi et al. (2012) have proposed that, despite the total absence of chromosomes carrying the paragroup G-M201* in their dataset, the place of origin of G-M201 would be near Eastern Anatolia, Armenia or Western Iran, the only regions marked by the presence of deep basal branches added to the occurrence of high subhaplogroup diversity.

The presence of both haplogroups R and G in Kuwait could be explained by historical data. In fact, many slaves were brought from Iberia, between the Umayyad and the Abbasid Empires (Bassaam, 1977). In addition, during the colonization of the Middle East, by the British Empire, between the 19th and 20th Centuries CE, many slaves of Georgian, Armenian, or Circassians origins were imported to this region (Philby, 1923). Afterward, these slaves were naturalized and given tribal affiliations (Mohammad et al., 2009).

Furthermore, the more ancient TMRCA estimates of haplogroups R1-M173 and G-M201 in the Iranian population compared with those found in Kuwait, corroborates historical data of ancient gene flow from Iran into Kuwait. Consequently, the presence of haplogroup R in Kuwait could also be explained by migration flow from Iran, knowing that the current population of Kuwait consists, inter alia, of Persians coming from Iran such as the Ajam tribe, which is an ethnic group of Iranian origin representing approximately one-third of the Kuwaiti population (Alsmadi et al., 2013).

The comparative study carried out between the Y-haplogroup frequencies in the population of Kuwait and those of 30 other populations has helped us to corroborate the genetic similarity of the population of Kuwait with those of Arabia ($F_{ST} = 1.3\%$, P-value < 0.0001) (Table 2). Similar results were obtained with haplotypic data from the populations of the peninsula [Kuwait, Saudi Arabia, Qatar, UAE, Yemen, Oman, and Dubai ($F_{ST} = 6.6\%$, P-value $= 0.042$) (Triki-Fendri et al., 2010)]. These results indicate that the paternal lineages, based on both Y-STR and Y-SNP data, support the close genetic relationship between the populations of Arabia. Besides, the lack of correlation between genetic and geographic distances, when considering the populations of the Arabian Peninsula ($r = 0.46$, P-value $= 0.16$, with 1,000 iterations), is in agreement with the genetic homogeneity of the populations in Arabia.

In the same way, the low genetic variance found when comparing the populations of Arabia with those of the Fertile Crescent, from the one hand ($F_{ST} = 4.2\%$, P-value < 0.0001), or those of the Middle East in general, on the other hand ($F_{ST} = 6.1\%$, P-value < 0.0001), indicates the genetic homogeneity of the populations in the region. In fact, the Middle East is defined by the region between the eastern shore of the Mediterranean Sea and the line drawn by the border between Iran on the one hand, Pakistan and Afghanistan on the other hand.

This genetic homogeneity is reinforced by the high proportion of nonsignificant pairwise F_{ST} genetic distances between Kuwait and other populations from either Arabia or the Fertile Crescent. However, the correlation between the genetic and geographic distances with all of the populations of the Middle East, including Iran, Turkey, and Egypt ($r = 0.41$, P-value $= 0.009$, with 1,000 iterations)

indicates that these three populations are more heterogeneous than those of the Arabian Peninsula.

Moreover, the MDS plot, constructed on the basis of these pairwise F_{ST} genetic distances (Fig. 2), indicated that Kuwaitis are close to the populations of the Arabian Peninsula, and more generally, to all of the populations of the Middle East. In addition, East European populations were associated with this cluster. This is not strange, as the R1-M173 haplogroup, particularly its subhaplogroup R1a1-SRY10831.2 which is common in East Europe, is also highly represented in the Middle East: 9.4% in Kuwait (present work), 19.4% in Iraq (Al-Zahery et al., 2011), and 25% in Iran (Regueiro et al., 2006). Indeed, the J2-M172 haplogroup, which is typical of Middle Eastern populations, was relatively common in the populations of Eastern Europe, including Croatia (29%); Bosnia (13%) (Battaglia et al., 2009); and Albania (18%) (Sengupta et al., 2006).

It is noteworthy that a significant difference was observed between the Kuwaiti population and the Jordanians from the Dead Sea (Jordan Valley). This finding could be attributed to the anomalous Y-chromosome pool of the population of the Dead Sea due to isolation. In fact, the Jordan valley is located 390 m below sea level and consequently under mild chronic hyperoxia. This peculiarity has reduced the immigration toward the Jordan Valley from other regions. Thus, the inhabitants might have kept the genetic features of the original population of this region (Flores et al., 2005).

The similarity observed between the population of Kuwait and the Arabs of Iran ($F_{ST} = 0.022$, P-value = 0.018) supports previous findings. In fact, it has already been reported that the subhaplogroups J1-Page08 and J2-M92, observed in the Iranian Y-chromosomal gene pool, demonstrates the presence of migration flow from Arab countries and Anatolia toward Iran (Grugni et al., 2012). Similarly, the nonsignificant difference between the Kuwaiti population and the Arabs of Sudan ($F_{ST} = 0.02$, P-value = 0.018) is concordant with the hypothesis of gene flow from Arabia toward Africa, through the Strait of Bab El-Mandeb (Cruciani et al., 2002; Hassan et al., 2008).

Nevertheless, the significant difference between Middle Eastern populations and those of North Africa ($F_{ST} = 29\%$, P-value < 0.00001) with 24% of variation occurring between groups, reflects a substantial difference between both regions. Based on haplogroup distribution, the genetic pool of the North African populations is made up of two components: the Berber genetic impact of the autochthonous inhabitants of the region (haplogroup E-M81) added to the Arabic one (haplogroup J1-M267), brought to North Africa by gene flow from the Middle East (Fadhlaoui-Zid et al., 2011; Robino et al., 2008; Triki-Fendri et al., 2015). In fact, it is well known that the distribution of E-M81 in Africa is highly related to the regions inhabited by Berber-speaking populations, suggesting a close haplogroup-ethnic group parallelism (Cruciani et al., 2004). Thus, the total absence of this Berber genetic component (haplogroup E-M81) in the Middle East could be one reason for the significant difference between both regions.

In addition, in the MDS plot (Fig. 2), Tunisian and Moroccan populations clustered together, representing North Africa. However, the populations of Libya and Algeria were found closer to the populations of the Middle East rather than those of North Africa. This peculiarity could be due to the bias of sampling as the Tunisian and Moroccan samples were made up essentially of Berbers (60 and 70%, respectively). Besides, and according to the PCA (Fig. 3), the first component separated the populations of North Africa from the remaining populations. It is obvious, according to the PCA, that haplogroup E-M81 typifies the North African region, particularly the Berber populations. Conversely, the second component differentiates the populations of the Arabian Peninsula [haplogroup J-M304(xM62,M172)] from those of East Europe and East Africa following the directions of R-M173 and E-M78, respectively. Note that the haplogroup J-M172 characterizes the populations of the Fertile Crescent.

BATWING posterior estimates based on the Y-STR haplotypes of the Kuwaiti samples suggest a largely pre-Neolithic settlement in Kuwait with gene flow from outside populations occurring later. However, archaeological data point to an Ubaid-related settlement in Kuwait, dating back 8,000 years, the first culture preceding urbanization while other studies indicate a more ancient settlement of small populations of fisherman (Fe Santiago, 2012).

In conclusion, the comparative study of Kuwait with other populations has helped to elucidate the genetic landscape of the region. First, Kuwait shares the same pattern of genetic diversity of the Y-chromosome as the populations of Arabia and those of the Fertile Crescent. Moreover, we have demonstrated that the populations of the Middle East are genetically homogeneous ($F_{ST} = 6.1\%$, P-value < 0.0001). This homogeneity was also supported by TMRCA estimates, suggesting a largely pre-Neolithic settlement in Kuwait with gene flow, likely due to historical immigrant of labor, from outside populations, particularly from Iran, based on the TMRCAs of haplogroups R-M173 and G-M201, and from Saudi Arabia, according to those of J1-M267. However, a higher genetic variance was found between the populations of the Middle East and those of North Africa ($F_{ST} = 29\%$, P-value < 0.00001), reflecting different settlement histories between both regions, most likely caused by the total absence of the Berber component (haplogroup E-M81) in the Middle East.

Moreover, the similarity between Kuwait and the Arabs of Sudan, based on the pairwise F_{ST} genetic distance, added to the TMRCAs estimates of haplogroup E-M35 in both populations, corroborates the hypothesis of bidirectional gene flow between Arabia and East Africa.

ACKNOWLEDGMENT

The authors are grateful to Dr. Ian Wilson for his valuable help with the expert use of BATWING program.

LITERATURE CITED

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

Al-Enizi M, Ge J, Ismael S, Al-Enezi H, Al-Awadhi A, Al-Duaij W, Al-Saleh B, Ghulloom Z, Budowle B. 2013. Population genetic analyses of 15 STR loci from seven forensically-relevant populations residing in the state of Kuwait. Forensic Sci Int Genet 7:e106–e107.

Alsmadi O, Thareja G, Alkayal F, Rajagopalan R, John SE, Alsmadi O, Thareja G, Alkayal F, Rajagopalan R, John SE, Hebbar P, Behbehani K, Thanaraj TA. 2013. Genetic substructure of Kuwaiti population reveals migration history. PLoS One 8:e74913.

Al-Zahery N, Pala M, Battaglia V, Grugni V, Hamod MA, Hooshiar Kashani B, Olivieri A, Torroni A, Santachiara-Benerecetti AS, Semino O. 2011. In search of the genetic footprints of Sumerians: a survey of Y-

chromosome and mtDNA variation in the Marsh Arabs of Iraq. BMC Evol Biol 11:288.
Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C. 2004. A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. Am J Hum Genet 75:338–345.
Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, Haber M, Platt D, Schurr T, Haak W, Kuznetsova M, Radzhabov M, Balaganskaya O, Romanov A, Zakharova T, Soria Hernanz DF, Zalloua P, Koshel S, Ruhlen M, Renfrew C, Wells RS, Tyler-Smith C, Balanovska E; Genographic Consortium. 2011. Genographic Consortium. Parallel evolution of genes and languages in the Caucasus region. Mol Biol E 28:2905–2920.
Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol E 16:37–48.
Bassaam AA. 1977. Ulama Najd Khilaal Sitat Quroon. Damascus: Muktabat Al-Nahdha Al-Hadeeth.
Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S, Marjanovic D, Primorac D, Hadziselimovic R, Vidovic S, Drobnic K, Durmishi N, Torroni A, Santachiara-Benerecetti AS, Underhill PA, Semino O. 2009. Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe. Eur J Hum Genet 17:820–830.
Blanco-Verea A, Brion M, Ramos-Luis E, Lareu MV, Carracedo A. 2008. Forensic validation and implementation of Y-chromosome SNP multiplexes. Forens Sci Int Genet Suppl Ser 1:181–183.
Brión M, Sanchez JJ, Balogh K, Thacker C, Blanco-Verea A, Børsting C, Stradmann-Bellinghausen B, Bogus M, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N. 2005a. Introduction of an single nucleodite polymorphism-based "Major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. Electrophoresis 26:4411–4420.
Brión M, Sobrino B, Blanco-Verea A, Lareu MV, Carracedo A. 2005b. Hierarchical analysis of 30 Y-chromosome SNPs in European populations. Int J Legal Med 119:10–15.
Brisighelli F, Álvarez-Iglesias V, Fondevila M, Blanco-Verea A, Carracedo A. 2012. Uni-parental markers of contemporary Italian population reveals details on its pre-Roman heritage. PLoS One 7:e50794.
Busby GB, Brisighelli F, Sánchez-Diz P, Ramos-Luis E, Martinez-Cadenas C, Thomas MG, Bradley DG, Gusmão L, Winney B, Bodmer W, Vennemann M, Coia V, Scarnicci F, Tofanelli S, Vona G, Ploski R, Vecchiotti C, Zemunik T, Rudan I, Karachanak S, Toncheva D, Anagnostou P, Ferri G, Rapone C, Hervig T, Moen T, Wilson JF, Capelli C. 2012. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. Proc Biol Sci 279:884–892.
Cadenas AM, Zhivotovsky LA, Cavalli-Sforza LL, Underhill PA, Herrera RJ. 2008. Y-chromosome diversity characterizes the Gulf of Oman. Eur J Hum Genet 16:374–386.
Casey MS. 2007. Desert origins and settlement (ca. 3000 BCE-1756 CE). In: Casey MS, editors. History of Kuwait. Connecticut: Greenwood Press. p 15–28.
Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton, NJ: Princeton University Press.
Cerny V, Mulligan CJ, Fernandes V, Silva NM, Alshamali F, Non A, Harich N, Cherni L, El Gaaied ABA, Al-Meeri A, Pereira L. 2011. Internal diversification of mitochondrial haplogroup R0a reveals post-last glacial maximum demographic expansions in south Arabia. Mol Biol Evol 28:71–78.
Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA. 2004. Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 114:127–148.
Chaabani H. 2002. GM polymorphism and the evolutionary history of modern humans. Ann Genet 45:197–206.
Chaabani H. 2014. Recent out of Yemen: new version of the theory of unique and recent origin of modern man. Int J Mod Anthrop 7:13–41.
Chaabani H, Sanchez-Mazas A, Sallami SF. 2000. Genetic differentiation of Yemeni people according to Rhessus and Gm polymorphisms. Ann Génét 43:155–162.
Chiaroni J, King RJ, Myres NM, Henn BM, Ducourneau A, Mitchell MJ, Boetsch G, Sheikha I, Lin AA, Nik-Ahd M, Ahmad J, Lattanzi F, Herrera RJ, Ibrahim ME, Brody A, Semino O, Kivisild T, Underhill PA. 2010. The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. Eur J Hum Genet 18:348–353.
Chiaroni J, King R, Underhill P. 2008. Correlation of annual precipitation with human Y-chromosome diversity and the emergence of Neolithic agricultural and pastoral economies in the Fertile Crescent. Antiquity 82:281–289.
Chiaroni J, Underhill PA, Cavalli-Sforza LL. 2009. Y chromosome diversity, human expansion, drift, and cultural evolution. Proc Natl Acad Sci USA 106:20174–20179.
Contu D, Morelli L, Santoni F, Foster JW, Francalacci P, Cucca F. 2008. Y-chromosome based evidence for pre-neolithic origin of the genetically homogeneous but diverse Sardinian population: inference for association scans. PLoS One 3:1430.
Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cali F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R. 2004. Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. Am J Hum Genet 74:1014–1022.
Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, Dugoujon JM, Crivellaro F, Benincasa T, Pascone R, Moral P, Watson E, Melegh B, Barbujani G, Fuselli S, Vona G, Zagradisnik B, Assum G, Brdicka R, Kozlov AI, Efremov GD, Coppa A, Novelletto A, Scozzari R. 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomalhaplogroups E-M78 and J-M12. Mol Biol E 24:1300–1311.
Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA. 2002. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am J Hum Genet 70:1197–1214.
Cruciani F, Trombetta B, Sellitto D, Massaia A, Destro-Bisol G, Watson E, Beraud Colomb E, Dugoujon JM, Moral P, Scozzari R. 2010. Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. Eur J Hum Genet 18:800–807.
de Filippo C, Barbieri C, Whitten M, Mpoloka SW, Gunnarsdóttir ED, Bostoen K, Nyambe T, Beyer K, Schreiber H, de Knijff P, Luiselli D, Stoneking M, Pakendorf B. 2011. Y-chromosomal variation in sub-Saharan Africa: insights into the history of Niger-Congo groups. Mol Biol E 28:1255–1269.
Di Piazza FD. 2008. Kuwait in pictures. Minneapolis, MN: Lerner Publishing Group. p 20–21.
El-Sibai M, Platt DE, Haber M, Xue Y, Youhanna SC, Wells RS, Izaabel H, Sanyoura MF, Harmanani H, Bonab MA, Behbehani J, Hashwa F, Tyler-Smith C, Zalloua PA; Genographic Consortium. 2009. Geographical structure of the Y-chromosomal genetic landscape of the Levant: a coastal-inland contrast. Ann Hum Genet 73:568–581.
Excoffier L, Laval G, Schneider S. 2007. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.
Fadhlaoui-Zid K, Martinez-Cruz B, Khodjet-el-khil H, Mendizabal I, Benammar-Elgaaied A, Comas D. 2011. Genetic structure of Tunisian ethnic groups revealed by paternal lineages. Am J Phys Anthropol 146: 271–280.
Fe Santiago M. 2012. Kuwait-Polish team finds Ubaid era village. Arab Times, Kuwait Editor's, Airport Road, Shuwaikh P.O.Box: 2270 13023 Safat. available: http://www.arabtimesonline.com/wp-content/uploads/pdf/2012/nov/30/04.pdf.
Flores C, Maca-Meyer N, Larruga JM, Cabrera VM, Karadsheh N, Gonzalez AM. 2005. Isolates in a corridor of migrations: a high-resolution analysis of Y-chromosome variation in Jordan. J Hum Genet 50:435–441.
Goedbloed M, Vermeulen M, Fang RN, Lembring M, Wollstein A, Ballantyne K, Lao O, Brauer S, Krüger C, Roewer L, Lessig R, Ploski R, Dobosz T, Henke L, Henke J, Furtado MR, Kayser M. 2009. Comprehensive mutation analysis of 17 Y-chromosomal shorttandem repeat polymorphisms included in the AmpFlSTRYfiler PCR amplification kit. Int J Legal Med 123:471–482.
Grugni V, Battaglia V, HooshiarKashani B, Parolo S, Al-Zahery N. 2012. Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. PLoS One 7:e41252.
Haber M, Platt DE, Badro DA, Xue Y, El-Sibai M, Bonab MA, Youhanna SC, Saade S, Soria-Hernanz DF, Royyuru A, Wells RS, Tyler-Smith C, Zalloua PA. 2011. Genographic Consortium. Influences of history, geography, and religion on genetic structure: the Maronites in Lebanon. Eur J Hum Genet 19:334–340.
Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. 1998. Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol E 15:427–441.
Hassan HY, Underhill PA, Cavalli-Sforza LL, Ibrahim ME. 2008. Y-chromosome variation among Sudanese: restricted gene flow, concordance with language, geography, and history. Am J Phys Anthropol 137:316–323.
Hochberg Y. 1988. A charper Bonferroni procedure for multiple test of significance. Biometrika 75:800–802.
Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4:598–612.

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

Kawazaki E. 1990. Sample preparation from blood, cells and others fluids. In: Innis M, Gelffand D, Snisky G, White T, editors. PCR protocols. A guide to methods and application. San Diago, CA: Academic press. p 146–152.

Lahr MM, Foley R. 1994. Multiple dispersals and modern human origins. Evol Anthropol 3:48–60.

Luis JR, Rowold DJ, Regueiro M, Caeiro B, Cinnioğlu C. 2004. The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. Am J Hum Genet 74:532–544. Erratum in: Am J Hum Genet 74:788.

Mielnik-Sikorska M, Daca P, Woźniak M, Malyarchuk BA, Bednarek J, Dobosz T, Grzybowski T. 2013. Genetic data from Y chromosome STR and SNP loci in Ukrainian population. Forensic Sci Int Genet 7:200–203.

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

Muzzio M, Muzzio JC, Bravi CM, Bailliet G. 2010. Technical note: a method for assignment of the weight of characters. Am J Phys Anthropol 143:488–492.

Nasidze I, Ling EYS, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA, Asgary S, Sardas S, Farhud DD, Sarkisian T, Asadov C, Kerimov A, Stoneking M. 2004. Mitochondrial DNA and Y-chromosome variation in the Caucasus. Ann Hum Genet 68:205–221.

Oppenheimer S. 2003. Out of Eden: the peopling of the world. London: Constable.

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

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

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

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

Rootsi S, Myres NM, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, Karachanak S, Hooshiar Kashani B, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. 2012. Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. Eur J Hum Genet 20:1275–1282.

Sanchez J, Hallenberg C, Borsting C, Hernandez A, Morling N. 2005. High frequencies of Y chromosome lineages characterized by E3b1, DYS19-11, DYS392-12 in Somali males. Eur J Hum Genet 13:856–866.

Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill PA, Santachiara-Benerecetti AS. 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.

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

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

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

Shi W, Ayub Q, Vermeulen M, Shao RG, Zuniga S, et al. 2010. A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations. Mol Biol E 27:385–393.

Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt HJ, Torroni A, Richards MB. 2010. The archaeogenetics of Europe. Curr Biol 20: 174–183.

SPSS Inc. Released 2008. SPSS Statistics for Windows, Version 17.0. Chicago: SPSS Inc.

Stringer C. 2000. Coasting out of Africa. Nature 405:24–27.

Teebi AS. 1994. Autosomal recessive disorders among Arabs: an overview from Kuwait. J Med Genet 31:224–233.

Theyab JB, Al-Bustan S, Crawford MH. 2012. The genetic structure of the Kuwaiti population: mtDNA inter- and intra-population variation. Hum Biol 84:379–403.

Triki-Fendri S, Alfadhli S, Ayadi I, Kharrat N, Ayadi H, Rebai A. 2010. Genetic structure of Kuwaiti population revealed by Y-STR diversity. Ann Hum Biol 37:827–835.

Triki-Fendri S, Sánchez-Diz P, Rey-González D, Ayadi I, Alfadhli S, Rebai A, Carracedo Á. 2013. Population genetics of 17 Y-STR markers in West Libya (Tripoli region). Forensic Sci Int Genet 7:e59–e61.

Triki-Fendri S, Sánchez-Diz P, Rey-González D, Ayadi I, Carracedo A, Rebai A. 2015. Paternal lineages in Libya inferred from Y-chromosome haplogroups. Am J Phys Anthropol 157:242–251.

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

Underhill PA, Passarino G, Lin AA, Shen P, Mirazón Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62.

Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonné-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ. 2000. Y chromosome sequence variation and the history of human populations. Nat Genet 26: 358–361.

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill, PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF. 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci USA 98:10244–10249.

Wilson IJ, Balding DJ. 1998. Genealogical inference from microsatellite data. Genetics 150:499–510.

Wilson J, Weale ME, Balding DJ. 2003. Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. J R Stat Soc A 166:155–201.

Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J. 2006. Male demography in East Asia: a north-south contrast in human population expansion times. Genetics 172:2431–2439.

Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, Yepiskoposyan L, Khusnutdinova EK, Underhill PA, Kivisild T, Villems R. 2012. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. Mol Biol E 29:359–365.

Zalloua PA, Platt DE, El-Sibai M, Khalife J, Makhoul N, Haber M, Xue Y, Izaabel H, Bosch E, Adams SM, Arroyo E, López-Parra AM, Aler M, Picornell A, Ramon M, Jobling MA, Comas D, Bertranpetit J, Wells RS, Tyler-Smith C; Genographic Consortium. 2008a. Genographic Consortium. Identifying genetic traces of historical expansions: phoenician footprints in the Mediterranean. Am J Hum Genet 83:633–642.

Zalloua PA, Xue Y, Khalife J, Makhoul, N, Debiane L, Platt DE, Royyuru AK, Herrera RJ, Hernanz DF, Blue-Smith J, Wells RS, Comas D, Bertranpetit J, Tyler-Smith C; Genographic Consortium. 2008b. Genographic Consortium. Y-chromosomal diversity in Lebanon is structured by recent historical events. Am J Hum Genet 82:873–882.