Evidence for prehistoric origins of the G2019S mutation in the North African Berber population

Rafiqua Ben El Haj[1], Ayyoub Salmi[2], Wafa Regragui[1,3], Ahmed Moussa[2], Naima Bouslam[3], Houyam Tibar[1,3], Ali Benomar[1,3], Mohamed Yahyaoui[1,3], Ahmed Bouhouche[1,3]*

1 Research Team in Neurology and Neurogenetics, Genomics Center of Human Pathologies, Medical School and Pharmacy, Mohammed V University, Rabat, Morocco, **2** Laboratory of Information and Communication Technologies, National School of Applied Sciences, Abdelmalek Essaadi University, Tanger, Morocco, **3** Department of Neurology and Neurogenetics, Specialties Hospital, IBN Sina University Hospital Center, Rabat, Morocco

* a.bouhouche@um5s.net.ma

Abstract

The most common cause of the monogenic form of Parkinson's disease known so far is the G2019S mutation of the *leucine-rich repeat kinase 2* (*LRRK2*) gene. Its frequency varies greatly among ethnic groups and geographic regions ranging from less than 0.1% in Asia to 40% in North Africa. This mutation has three distinct haplotypes; haplotype 1 being the oldest and most common. Recent studies have dated haplotype 1 of the G2019S mutation to about 4000 years ago, but it remains controversial whether the mutation has a Near-Eastern or Moroccan-Berber ancestral origin. To decipher this evolutionary history, we genotyped 10 microsatellite markers spanning a region of 11.27 Mb in a total of 57 unrelated Moroccan PD patients carrying the G2019S mutation for which the Berber or Arab origin was established over 3 generations based on spoken language. We estimated the age of the most recent common ancestor for the 36 Arab-speaking and the 15 Berber-speaking G2019S carriers using the likelihood-based method with a mutation rate of 10^{-4}. Data analysis suggests that the shortest haplotype originated in a patient of Berber ethnicity. The common founder was estimated to have lived 159 generations ago (95% CI 116–224) for Arab patients, and 200 generations ago (95% CI 123–348) for Berber patients. Then, 29 native North African males carrying the mutation were assessed for specific uniparental markers by sequencing the Y-chromosome (E-M81, E-M78, and M-267) and mitochondrial DNA (mtDNA) hypervariable regions (HV1 and HV2) to examine paternal and maternal contributions, respectively. Results showed that the autochthonous genetic component reached 76% for mtDNA (Eurasian and north African haplogroups) and 59% for the Y-chromosome (E-M81 and E-M78), suggesting that the G2019S mutation may have arisen in an autochthonous DNA pool. Therefore, we conclude that *LRRK2* G2019S mutation most likely originated in a Berber founder who lived at least 5000 years ago (95% CI 3075–8700).

Introduction

Parkinson's disease (PD) is the most common neurodegenerative disorder after Alzheimer's disease, affecting about 1–2% of people over 60 years of age [1]. PD is characterized by clinical features that comprise a large spectrum of motor symptoms including, resting tremor, rigidity and bradykinesia. These manifestations are associated with a severe and progressive degeneration of dopaminergic neurons within the substantia nigra *pars compacta* [2].

In the past two decades, 15 genes have been described for inherited forms of PD. Among them, there are at least 5 confirmed genes responsible for autosomal dominant forms of PD: *SNCA* (PARK1/4), *LRRK2* (PARK8), *VPS35* (PARK17), *DNAJC13* and *CHCHD2* [3,4]. Mutations in the *LRRK2* gene are the most frequently identified as monogenic causes of PD. Indeed, they have been found in about 1.8% of healthy controls, 3.6% of sporadic PD cases and 10% of autosomal dominant familial PD cases [1]. Approximately 80 variants have been identified in *LRRK2* gene so far [5]. Of these variants, 7 mutations are thought to be pathogenic but the most common one remains the G2019S mutation, which accounts for 5–6% of familial and 1–2% of sporadic PD cases [1,6]. Intriguingly, the frequency of this mutation varies widely depending on geographic origin and ethnicity. Nevertheless, the highest frequencies are those recorded in Ashkenazi Jews and North African populations reaching up to approximately20% and 40% of PD patients respectively [7].

Several studies on different populations around the world attempted to define the origin of the *LRRK2* G2019S mutation and date the age of emergence. Presently, the mutation is associated with three different haplotypes. While haplotype 2 has been found in individuals of European descent, haplotype 3 was first identified in Turkish and Japanese families. These two haplotypes have not been well studied and seem to be of more recent origin than haplotype 1[8–11]. Despite being widespread across the world, haplotype 1 follows from a common ancestral founder [12]. A study on a European cohort estimated the common founder occurrence to 725 years ago [13], though a subsequent study on Ashkenazi, European and European-American families has estimated it to 2250 years ago; a period that overlaps with the Jewish Diaspora from 586 BC to 70 AD [9]. In addition, Warren et al. [14] dated the most recent common ancestor of patients with haplotype 1 from Tunisia, America, European, and Middle-Eastern countries to 2600 years ago. Furthermore, Bar-Shira et al. [15] estimated that the common ancestor of Ashkenazi Jewish G2019S carriers lived about 1830 years ago. However, the most significant study carried out by Lesage et al. on a large sample of patients belonging to several ethnicities, which suggested that haplotype 1 originated in a Near-Eastern founder at least 4000 years ago [11]. Finally, a recent study on North African patients concluded that the G2019S mutation originated in a Moroccan Berber patient who lived 3840 years ago [16].

While the *LRRK2* G2019S mutation is largely widespread throughout the North African populations, which are composed of two predominant ethnic groups Berbers and Arabs. It should be noted that in all these previous studies, the ethnic origin of this population was considered as Arab or Arab-Berber without a clear distinction between the two ethnic groups, or not documented at all. In addition to having a small sample of Moroccan patients, this discrepancy places further limitations on both the power and clarity of experimental analyses in previous studies. To better distinguish between these ethnic backgrounds, we employ three generations of spoken language as a tool of differentiation. In the present study, we aim to date the age of the most recent common ancestor of the G2019S carriers within Moroccan PD patients and to determine whether the mutation is of Berber or Arab origin by studying maternal and paternal lineages using both mitochondrial and Y-chromosomal genetic markers.

Subjects and methods

Patients and controls

A total of 57 unrelated PD patients with the G2019S mutation were selected from the Department of Neurology and Neurogenetics (Specialties Hospital, Rabat, Morocco). All patients agreed to participate and were recruited from October 2013 to June 2016 according to criteria described earlier [17]. Age at examination ranged from 35 to 86 years old and the mean age was 59 ± 11.4 years. In addition, 77 control individuals with no family history of neurological disease were recruited at the Blood Transfusion Center (Rabat, Morocco). All individuals, patients and controls included in the study, are native Moroccans and their ethnical origin was determined on the basis of spoken language (Berber or Arab) through three generations. Participants were classified as Arab- or Berber-speaking if all of the individuals' grandparents speak Arab or Berber languages respectively, and mixed if at least one grandparent speaks a different language. This study was approved by the biomedical research ethics committee of the Medical School and Pharmacy of Rabat (CERB), and written informed consent was obtained from all subjects.

Genotyping and estimation of the age of the G2019S mutation

Genomic DNA was extracted from peripheral blood leukocytes using Isolate II Genomic DNA kit from Bioline. Haplotypes were constructed using the following microsatellite markers: D12S1648, D12S2080, D12S2194, D12S2514, D12S2516, D12S2518, D12S2519, D12S2520, D12S1048 and D12S1301, covering an interval of 11.27Mb. The genotyping of microsatellite markers was performed using fluorescently labeled primers, and the products were pooled and analyzed on a 310 Genetic Analyzer using Genotyper Software version 3.7 (Applied Biosystems). The CEPH DNA sample 1331–01 was used as external standard to control for consistency between runs. We used Phase v2.1.1 software [18] to infer the haplotypic phase for all subjects. The age of the most recent common ancestor of the 36 Arab-speaking and the 15 Berber-speaking G2019S carriers was estimated separately using a likelihood-based method implemented in the ESTIAGE program [19]. We used allele frequencies obtained from the control individuals and a stepwise mutation model with a mutation rate of 10^{-4} at each marker per generation was assumed [20].

Analysis of uniparental markers

Since the G2019S mutation can be transmitted by either parent, and the uniparental markers can only be studied simultaneously in males, the 29 males carrying the G2019S mutation (18 Arab-speakers, 8 Berber-speakers and 3 mixed) were studied for their Y-chromosome and mtDNA markers. These markers are known to be inherited unchanged from one male or female generation to the next unless mutations occur.

The biallelic markers E-M81 and E-M78, reported as being the specific male lineage of autochthonous Berbers of North Africa [21,22] and J1-M267, the most prominent genetic marker of males from the Levantine [23,24], were analyzed in the 29 male carriers. All markers were amplified by PCR using primers as described previously [25,26]. For the maternal lineages, we amplified the two mtDNA hypervariable regions, HV1 and HV2, using primers F15971/R16451 for HV1 and F15/R484 for HV2 as described by Levin BC et al. (1999) [27]. All PCR fragments were sequenced using Big Dye Terminator Cycle Ready Reaction 3.1 Kits and sequence analysis was done through an ABI 3130xl sequencer. Data were analyzed by SeqScape2.1 software (Applied Biosystems, Foster City, CA). The mtDNA haplogroup assignment was done using the HaploGrep2 software based on the underlying classification tree PhyloTree Build 17 [28].

Results

Genotyping and estimation of the age of the G2019S mutation

Among the 57 unrelated PD patients carrying the *LRRK2* G2019S mutation, 29 were males and 28 were females. Of these patients, 52 were heterozygous and 5 were homozygous for the mutation. Their ethnic origin was investigated over three generations on the basis of their spoken language (Berber or Arab). Surprisingly, we found that 63% (36/57) of patients with the G2019S mutation were from Arab-speaking families and only 26% (15/57) were from Berber-speaking families for at least three generations. In the six remaining families (11%), the patients' grandparents were found to speak both languages, and were therefore considered to be of mixed descent.

The 36 Arab-speaking and 15 Berber-speaking patients were genotyped for 10 microsatellites around the G2019S mutation; patients of mixed origin were excluded because it is not possible to determine precisely the parent from whom the mutation was inherited, since their DNA was not available for analysis. The shared haplotype was inferred from patients homozygous for the G2019S mutation (Table 1). This haplotype, covering a distance of 575 Kb and spanning microsatellite markers D12S2194 to D12S1048, corresponds with the most frequent haplotype obtained by PHASE software. This finding only differs from previous studies in that the inferred haplotype concerns the shared allele at marker D12S2194, which was 261bp in our study instead of 257bp. In addition to the fact that 80% (4/5) of patients homozygous for the G2019S mutation were also homozygous for the 261bp allele, this particular allele was also found in 48% (49/102) of chromosomes in PD patients and in only 8% (13/154) of chromosomes in control individuals; whereas the frequency of the 257bp allele was 18% (18/102) and 20% (31/154) for these groups, respectively (S1 and S2 Tables). The use of the 261bp allele at D12S2194 reduced the number of recombination events at this locus and thus enabled a more accurate estimation of its age.

Interestingly, 75% (38/51) of G2019S carriers showed the shared haplotype of 575 kb in length, which corresponds with the already reported MENA haplotype and is referred to as haplotype 1, indicating that they shared a common ancestor. The shortest haplotype was found in a patient of Berber origin (S1 Table). We estimated the age of the common ancestor of the 51 Moroccan G2019S carriers to 174 generation ago (95% CI 133–230). If a generation is defined as 25 years, then this corresponds to 4350 years (95% CI 3325–5750). The common founder was estimated to have lived 159 generations ago (95% CI 116–224) for the 36 G2019S Arab-speaking carriers and 200 generations ago (95% CI 123–348) for the 15 G2019S Berber-speaking patients, corresponding to 3975 (95% CI 2900–5600) and 5000 years (95% CI 3075–8700) respectively (Table 2).

Analysis of uniparental markers

In Table 3, results of the Y-chromosome marker analysis show that the E-M81 was the most frequent haplogroup with a prevalence of 48% (14/29), whereas E-M78 and J1-M267 were found in 10% (3/29) and 14% (4/29) of male carriers, respectively. The paternal genetic components of native Berbers (E-M81 and E-M78) represented 87.5% (7/8) of Berber-speakers, 50% (9/18) of Arab-speakers, and 33% (1/3) of mixed patients, whereas the most prominent male lineage of the LevantineJ1-M267, represented 12.5% (1/8), 16,6% (3/18) and 0% (0/3) respectively.

The maternal lineage was evaluated by sequencing the two mtDNA hypervariable regions, HV1 and HV2. Haplotypes and their haplogroup assignment for the 29 G2019S male carriers are presented in S3 Table and Table 3. The results show that haplogroups H, HV0, U (without

Table 1. Shared haplotype inferred from five G2019S homozygous patients.

Marker	Patients ID					Shared haplotype
	3332	3561	3601	3665	3811	
D12S1648	112/108	120/112	108/124	104/128	108/112	
D12S2080	188	180/188	180	180	180/188	
D12S2194	261	261	261	261	261/257	**261**
D12S2514	291	291	291	291	291	**291**
D12S2516	254	254	254	254	254	**254**
G2019S	A/A	A/A	A/A	A/A	A/A	**A/A**
D12S2518	154	154	154	154	154	**154**
D12S2519	132	132	132	132	132	**132**
D12S2520	260	260	260	260	260	**260**
D12S1048	214	214/223	214	214	214	**214**
D12S1301	110	110	110/114	102	94/98	

https://doi.org/10.1371/journal.pone.0181335.t001

U6), J, K, N1, and N2, belonging to the Eurasian lineages, were the most frequent haplogroups, overlapping with 66% (19/29) of male carriers. Among them, haplogroups belonging to the HV branch were the most common, comprising 28% (8/29). Haplogroups U6 and M1 of the North African lineages represented 10% (3/29), whereas sub-Saharan lineages (L2 and L3) were found in 24% (7/29) of our sample.

Furthermore, analysis of uniparental markers revealed that most males (13/29) carrying the G2019S mutation have both parents of Berber origin. The 4 male patients with the Arab J1-M267 marker have both of the native maternal lineage (mtDNA) haplogroups, H and K.

Discussion

In this study, to determine whether the G2019S mutation has a Berber or an Arab origin in Morocco, the ethnic background of 57 unrelated Moroccan patients carrying the mutation was investigated. Surprisingly, the results suggest that the majority of patients with the G2019S mutation were from Arab-speaking families 63% (36/57) and only 26% (15/57) were from Berber-speaking families for the last three generations. Based on these relative frequencies, the G2019S mutation seems to be of Arab origin. If this is the case, however, one would expect to find the mutation at high frequencies in other Middle-Eastern Arab countries from where the Islamic conquest descended to Morocco and North Africa. The only studies on the G2019S mutation in Arab countries have come recently from Egypt, which reported a frequency among sporadic PD patients of 9.7% (11/113) in the north [29] and only 1.45% (1/69) in the South of the country [30], and also from Saudi Arabia and all Gulf countries where the mutation was reported to be absent [31,32]. Together, these data suggest that it is highly unlikely that the mutation is of Arabic origin.

To address this issue, we estimated the age of the most recent common ancestor in the 36 G2019S Arab-speaking patients and the 15 G2019S Berber-speaking patients separately by genotyping 10 flanking microsatellites. Using best-call haplotypes, we estimated the age of the

Table 2. Age estimation of the G2019S mutation using an intergenerational time interval of 25 years.

Patients	Number of generations	Age estimation
Arab-speaking (N = 36)	159 (95% CI 116–224)	3975 (95% CI 2900–5600)
Berber-speaking (N = 15)	200 (95% CI 123–348)	5000 (95% CI 3075–8700)
All (N = 51)	174 (95% CI 133–230)	4350 (95% CI 3325–5750)

https://doi.org/10.1371/journal.pone.0181335.t002

Table 3. mtDNA and Y-chomosomeuniparental markers analysis for the 29 G2019S male carriers.

Patient	Langage	Y-chromosome markers			mtDNAhaplogroups
		E-M81	E-M78	J1-M267	
3076	A	-	-	-	U2e2
3336	A	+	-	-	I5a
3443	A	-	-	-	L2b
3450	A	-	-	-	N1a3a
3453	A	-	+	-	HV0f
3462	A	-	-	-	U6
3463	A	-	-	+	HV0
3540	A	+	-	-	L2c1a
3564	A	+	-	-	L3d1b1b
3589	A	-	-	-	L2a1b
3590	A	-	-	+	K2b1b
3639	A	+	-	-	U2
3665	A	+	-	-	HV0
3777	A	-	-	-	W5a
3783	A	+	-	-	U4a2a
3785	A	+	-	-	L3d
3787	A	-	-	+	H2a2a
3796	A	+	-	-	M1a5
3215	B	+	-	-	L2a1a
3280	B	-	+	-	H6
3331	B	+	-	-	B6
3393	B	-	-	+	H1e1a4
3434	B	+	-	-	U6
3445	B	+	-	-	K
3664	B	-	+	-	H6
3762	B	+	-	-	J1c3
3373	M	+	-	-	J2a2
3433	M	-	-	-	H13a2b3
3788	M	-	-	-	L2b1a

A: Arab-speaking;

B: Berber-speaking;

M: Mixed.

https://doi.org/10.1371/journal.pone.0181335.t003

common ancestor of the 36 Arab-speaking carriers with haplotype 1 to 159 generations ago (95% CI 97–161), and considering an intergenerational time interval of 25 years, this corresponds to about 3975 years. The common founder in the 15 G2019S Berber-speaking patients was estimated to have lived 200 generations ago (95% CI 123–348) corresponding to 5000 years (95% CI 3075–8700). These results, along with the fact that the shortest haplotype was observed in a Berber-speaking patient, indicate that the G2019S mutation might have arisen in Morocco for the first time in a Berber ancestor about 5000 years ago.

Our age estimation of the G2019S mutation in Berbers is older than what is reported in recent studies using the same study design. Also considering an intergenerational time interval of 25 years and using the ESTIAGE program, Lesage and collaborators reported an age estimate of about 3000 years using data collected from 67 North African patients and 8 microsatellites (D12S2080 to D12S1301) [11]. However, the age estimate provided by Lucotte et al. [16]

was about 3200 years using 7 microsatellites (D12S2194 to D12S1048). In the first study, only 10 patients were of Moroccan origin and were considered Arabs, while in the second study, the authors focused on only 5 Berber- and 2 Arab-speaking patients of Moroccan origin, neither of which is representative of the Moroccan population. Our study, however, focuses on a large sample of 51 Moroccan patients for whom the Arabic or Berber ethnicities were evaluated on the basis of three generations of spoken language.

The estimated age of the G2019S mutation for Arab-speaking patients (3975 years) is older than the date of the Islamic conquest, which occurred 1300 years ago. This suggests that current Arab-speaking populations in Morocco are the result of interbreeding between Berbers and Arabs over many generations, and that the majority of people who declare themselves as Arabs are in fact Arabized Berbers [12,21,33]. Indeed, the Arab influence in the Maghreb was more cultural than a demographic replacement of the autochthonous Berber population. This hypothesis follows from the fact that the Arab invasion of the Maghreb in the 7th century was made by only few thousand people compared to millions of indigenous Berbers of the time [34], strengthening the hypothesis that the G2019S mutation occurred in a Berber founder. As an observation that lends additional support to this hypothesis is that the frequency distribution of the G2019S mutation along the coastal regions of the Mediterranean Basin and in Europe is correlated strongly with that of the Y-chromosome E-M81 haplogroup, which would have appeared in North Africa during the Neolithic period about 14200 years ago (www.yfull.com/tree/E-M81). This haplogroup, referred to as a genetic Berber marker, reaches a mean frequency of 42% in North Africa, with a decreasing gradient ranging from 98.1% in the southern and Berber regions of Morocco to 11.7% in the north of Egypt [21,22,35]. In the Middle East, this Berber marker was reported in 1.3% of Lebanese [36], 5% among Sephardic Jews, and 3.7% of Turkish [37], but it was absent in Arab countries [23,38]. In Europe, E-M81 haplogroup, widespread but rare, was found relatively higher in southern countries particularly in Spain and Portugal with the highest frequency found in Cantabria at 18.6% [39]. In agreement with our hypothesis, the community of Cantabria has also been reported to have the highest frequency of G2019S carriers at 8.7% [40].

In our sample of G2019S male carriers, the specific autochthonous Y-chromosome haplogroups, E-M81 and E-M78, which coalescence time has been estimated in North Africa to 15000 years [41], are the most prevalent and represent 58% of patients. These markers are found in 87.5% of Berber-speaking G2009S male patients and in 50% of Arab-speaking males, supporting the hypothesis of the Arabic cultural influence on Berbers, since the Arab conquest occurred six centuries ago. Indeed, the most prominent Arab Y-chromosome marker in Arabs, J1-M267 [23,24], has been found only in 14% of our G2019S male carriers. On the other hand, analysis of the maternal phylogeny of the males carrying the G2019S mutation revealed that their mitochondrial pool was characterized by a high frequency of West-Eurasian haplogroups (66%), a lower frequency of sub-Saharan L lineages (24%), and the presence of autochthonous haplogroups U6 and M1 (10%). This haplotype diversity of the G2019S carriers has been reported in similar proportions in the modern day Moroccan population [42–44] and also in Iberomaurusian skeletons unearthed from the archaeological site of Taforalt in Northeastern Morocco dating back to 12000 years [45], suggesting that the Moroccan mitochondrial pool has been more typical of the Mediterranean for at least the last 12000 years. When both uniparental markers are taken into account, our results showed that the autochthonous genetic components in our sample reached 76% for mtDNA (Eurasian and north African haplogroups) and 58% for the Y-chromosome (E-M81 and E-M78) which suggests that the G2019S mutation may have arisen in an autochthonous DNA pool.

Since antiquity, the Berber people of the Maghreb have been invaded by many civilizations, all of which have been assimilated to various degrees: Phoenicians, Romans, Vandals,
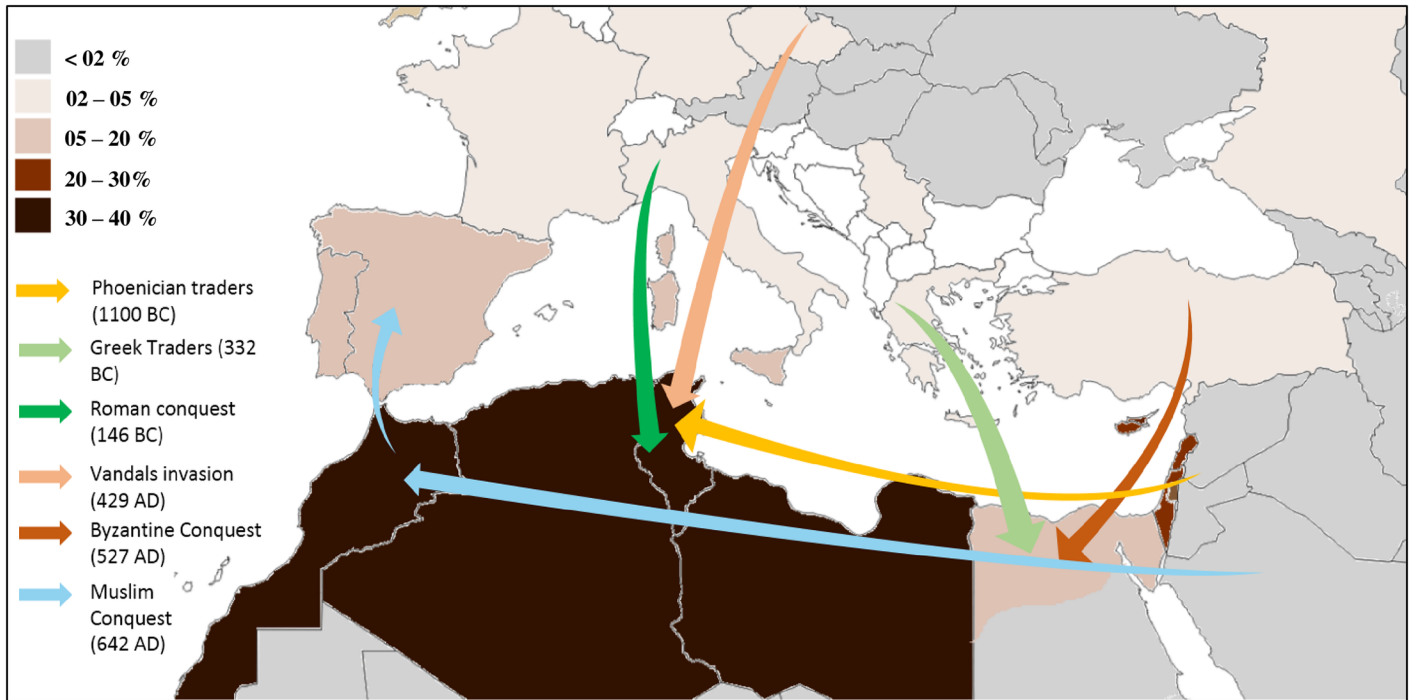
Fig 1. Map of the Mediterranean Basin illustrating the distribution of the G2019S mutation frequencies and the main conquests and invasions that the Maghreb has experienced since antiquity.

Byzantines and Arabs (Fig 1). Moreover, many Moors (Islamic Spaniards) and Jews arrived from Andalusia at the end of the 15th century. Despite this flux of humans into Maghreb area, the indigenous people have had genetic continuity since prehistoric time or even since the beginning of humanity, as the discovery in Morocco of the oldest *Homosapien* dating back to 300000 years supports [46].

The G2019S mutation is mainly found in Mediterranean populations with the highest frequencies seen in the Maghreb countries (Fig 1). Among them, Morocco presents with the highest prevalence and incidence of this mutation [17]. The general rule in population genetics stated that the geographic center of the origin of a mutation corresponds to the region in which the mutation is most frequent [47,48]. This conceptual framework, along with the results of uniparental markers, brings us to the most likely scenario for the origin of the G2019S mutation. This mutation would have appeared for the first time in a Berber founder living in the Maghreb at least 5000 years ago, with its center of dispersal located apparently further west, probably in Morocco. Over the course of several hundred years the mutation has increased its geographical range across the Northern coast of Africa as far as Egypt. The geographic distribution of G2019S frequencies in Europe (Fig 1), showing a north-south gradient with high values particularly in Spain and Portugal [11,12,40,49], could be explained by gene flow from Berber-Arab populations (mainly Berbers) into the Iberian Peninsula during the Muslim expansions from the 8th to 15th century[50]. Alternatively, it could also be caused by an earlier gene flow from Jewish populations, since their arrival was probably due to their migration from North Africa where they were already well-established. By genome-wide analysis, Campbell et al [51] showed that Jewish and non-Jewish North Africans formed distinct clusters, and that their mixture was not a recent event. Indeed, this event may go back to the Israelite's arrival to the Maghreb by the earliest Phoenician traders. Despite this ancient

connection, it is unlikely that the G2019S mutation is of Sephardic Jewish origin because the studies comparing Sephardic to Berber carriers all point to a Berber origin of the mutation. Berbers showed the highest prevalence and incidence of the mutation as well as the shorter haplotype compared to Sephardic Jews [16,50,52]. However, a problem arises when we attempt to explain the high frequency of this mutation in the Ashkenazim population. The G2019S mutation in Ashkenazim was reported to arise 4550 years (3250–6425) years ago [11] using a multi-ethnic ancestral haplotype. This age estimation, being slightly younger than that of our Berber ethnic group, is prior to the beginning of Jewish Diaspora and its establishment as an ethnic Jewish group. The presence of Jews in North Africa has been attested to as early as the Iron Age by the presence influence of memorial monuments allocated on architecture, practices and rituals to Jewish populations [53,54]. These Judeo-Berbers were reinforced by several waves of emigration, especially after the destruction of Jerusalem by Titus in 70 A.D.[55,56]. Judaism is thought to have spread among Berbers through proselytism, and its influence covered a significant portion of North Africa to the extent that it was even established as a state religion by the Berber Queen Kahina [57,58]. During the 7th century, some of these Judeo-Berbers would not have been converted to Islam by the Arab conquest and would have instead migrated to the Near-East as Jews, where some migrants might have served as Ashkenazi cofounders. Therefore, the high frequency of the G2019S mutation in the Ashkenazi population could be a result of isolation and genetic drift. Indeed, the origin of Ashkenazi Jews remains controversial, and their founders are thought to have lived in the Rhine Valley during the first millennium of the Common Era [59]. About 500 years ago, the size of the Ashkenazi population was estimated to consist of only several thousand people [60], but this period was followed by rapid population growth and a high rate of endogamy. Despite this consanguinity, it was reported that, due to Diasporas and migrations, the genome of the current Jewish people (including Ashkenazi) is very heterogeneous and has no specific genetic signature due to the influence of neighboring peoples [61]. This scenario gains support from the study at the genomic level by Henn BM and collaborators [62] who postulated that indigenous North African ancestry, spanning from the northwest to northeast, was most closely related to populations outside of Africa, and that their divergence from Near-Eastern and European groups preceded the Holocene for more than 12000 years. Therefore, it is important that further phylogenetic studies on G2019S carrier patients from other populations, Ashkenazi and Sephardic Jews from North Africa and Southern Europe in particular, are conducted to ascertain this scenario.

Supporting information

S1 Table. Haplotypes of 51 patients with Parkinson disease carrying the *LRRK2* G2019S mutation.
(XLSX)

S2 Table. Allele frequencies of 10 microsatellites in 39 Arab-speaking and 38 Berber-speaking control individuals.
(XLSX)

S3 Table. Mitochondrial DNA polymorphisms for HV1 and HV2 in 29 males carrying the G2019S mutation. A: Arab-speaking; B: Berber-speaking; M: Mixed.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Ahmed Bouhouche.

Formal analysis: Rafiqua Ben El Haj, Ayyoub Salmi, Ahmed Moussa.

Funding acquisition: Ahmed Bouhouche.

Investigation: Rafiqua Ben El Haj, Wafa Regragui, Naima Bouslam, Houyam Tibar, Ali Benomar, Mohamed Yahyaoui, Ahmed Bouhouche.

Methodology: Ahmed Bouhouche.

Project administration: Ahmed Bouhouche.

Resources: Ahmed Bouhouche.

Software: Ayyoub Salmi, Ahmed Moussa.

Supervision: Mohamed Yahyaoui, Ahmed Bouhouche.

Validation: Rafiqua Ben El Haj, Wafa Regragui, Ahmed Moussa, Houyam Tibar, Ali Benomar, Mohamed Yahyaoui, Ahmed Bouhouche.

Visualization: Rafiqua Ben El Haj, Wafa Regragui, Ahmed Moussa, Houyam Tibar, Ali Benomar, Mohamed Yahyaoui, Ahmed Bouhouche, Ahmed Bouhouche.

Writing – original draft: Rafiqua Ben El Haj, Ahmed Bouhouche.

Writing – review & editing: Ahmed Bouhouche.

References

1. Corti O, Lesage S, Brice A. What genetics tells us about the causes and mechanisms of Parkinson's disease.Physiol Rev. 2011; 91(4):1161–1218. https://doi.org/10.1152/physrev.00022.2010 PMID: 22013209

2. Lees AJ, Hardy J, Revesz T. Parkinson's disease. Lancet. 2009; 373(9680):2055–2066. https://doi.org/10.1016/S0140-6736(09)60492-X PMID: 19524782

3. Bonifati V. Genetics of Parkinson's disease—state of the art, 2013. Parkinsonism Relat Disord. 2014; 20 Suppl 1:S23–28. https://doi.org/10.1016/S1353-8020(13)70009-9

4. Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, et al. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. Lancet Neurol. 2015 Mar; 14(3):274–782. https://doi.org/10.1016/S1474-4422(14)70266-2 PMID: 25662902

5. Nuytemans K, Theuns J, Cruts M, Van Broeckhoven C. Genetic Etiology of Parkinson Disease Associated with Mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 Genes: A Mutation Update. Hum Mutat. 2010 Jul; 31(7):763–780. https://doi.org/10.1002/humu.21277 PMID: 20506312

6. Li JQ, Tan L, Yu JT. The role of the LRRK2 gene in Parkinsonism. Mol Neurodegener. 2014 Nov 12; 9:47. https://doi.org/10.1186/1750-1326-9-47 PMID: 25391693

7. Schulte C, Gasser T. Genetic basis of Parkinson's disease: inheritance, penetrance, and expression. Appl Clin Genet. 2011; 4:67–80. https://doi.org/10.2147/TACG.S11639 PMID: 23776368

8. Tomiyama H, Li Y, Funayama M, Hasegawa K, Yoshino H, Kubo S, et al. Clinicogenetic study of mutations in LRRK2 exon 41 in Parkinson's disease patients from 18 countries. Mov Disord. 2006 Aug; 21 (8):1102–1108. https://doi.org/10.1002/mds.20886 PMID: 16622854

9. Zabetian CP, Morino H, Ujike H, Yamamoto M, Oda M, Maruyama H, et al. Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. Neurology. 2006 Aug 22; 67 (4):697–9. https://doi.org/10.1212/01.wnl.0000227732.37801.d4 PMID: 16728648

10. Pirkevi C, Lesage S, Condroyer C, Tomiyama H, Hattori N, Ertan S, et al. A LRRK2 G2019S mutation carrier from Turkey shares the Japanese haplotype. Neurogenetics. 2009 Jul; 10(3):271–273. https://doi.org/10.1007/s10048-009-0173-5 PMID: 19172321

11. Lesage S, Patin E, Condroyer C, Leutenegger AL, Lohmann E, Giladi N, et al. Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans. Hum Mol Genet. 2010 May 15; 19(10):1998–2004. https://doi.org/10.1093/hmg/ddq081 PMID: 20197411

12. Benamer HT, de Silva R. LRRK2 G2019S in the North African population: a review. Eur Neurol. 2010; 63(6):321–325. https://doi.org/10.1159/000279653 PMID: 20413974

13. Lesage S, Leutenegger AL, Ibanez P, Janin S, Lohmann E, Dürr A, et al. LRRK2 Haplotype Analyses in European and North African Families with Parkinson Disease: A Common Founder for the G2019S Mutation Dating from the 13th Century. Am J Hum Genet. 2005 Aug; 77(2):330–332. https://doi.org/10.1086/432422 PMID: 16145815

14. Warren L, Gibson R, Ishihara L, Elango R, Xue Z, Akkari A, et al. A founding LRRK2 haplotype shared by Tunisian, US, European and Middle Eastern families with Parkinson's disease. Parkinsonism Relat Disord. 2008; 14(1):77–80. https://doi.org/10.1016/j.parkreldis.2007.02.001 PMID: 17433753

15. Bar-Shira A, Hutter CM, Giladi N, Zabetian CP, Orr-Utrreger A. Ashkenazi Parkinson's disease patients with the LRRK2 G2019S mutation share a common founder dating from the second to fifth centuries. Neurogenetics. 2009 Oct; 10(4):355–358. https://doi.org/10.1007/s10048-009-0186-0 PMID: 19283415

16. Lucotte G, David D, Change N. New contribution on the LRRK2 G2019S mutation associated to Parkinson's disease: age estimation of a common founder event of old age in Moroccan Berbers. Int J Mod Anthrop. 2012; 1(5):11–22. http://dx.doi.org/10.4314/ijma.v1i5.1

17. Bouhouche A, Tibar H, Ben El Haj R, El Bayad K, Razine R, Tazrout S, et al. LRRK2 G2019S Mutation: Prevalence and Clinical Features in Moroccans with Parkinson's Disease.Parkinsons Dis. 2017; 2017:2412486. https://doi.org/10.1155/2017/2412486 PMID: 28465860

18. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data.Am J Hum Genet.2003; 73:1162–1169. https://doi.org/10.1086/379378 PMID: 14574645

19. Génin E, Tullio-Pelet A, Begeot F, Lyonnet S, Abel L. Estimating the age or rare disease mutations: the example of triple-A syndrome. J Med Genet. 2004 Jun; 41(6):445–449. https://doi.org/10.1136/jmg.2003.017962 PMID: 15173230

20. Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, et al. Mutation patterns at dinucleotide microsatellite loci in Humans. Am J Hum Genet. 2002 Mar; 70(3):625–634. https://doi.org/10.1086/338997 PMID: 11793300

21. Bekada A, Fregel R, Cabrera VM, Larruga JM, Pestano J, Benhamamouch S, et al. Introducing the Algerian Mitochondrial DNA and Y-Chromosome Profiles into the North African Landscape. PLoS One. 2013; 8(2):e56775. https://doi.org/10.1371/journal.pone.0056775 PMID: 23431392

22. Reguig A, Harich N, Barakat A, Rouba H. Phylogeography of E1b1b-M81 haplogroup and analysis of its subclades in Morocco. Hum Biol. 2014; 86(2):105–112. https://doi.org/10.3378/027.086.0204 PMID: 25397701

23. Abu-Amero KK, Hellani A, González AM, Larruga JM, Cabrera VM, Underhill PA. Saudi Arabian Y-Chromosome diversity and its relationship with nearby regions. BMC Genet. 2009 Sep 22; 10:59. https://doi.org/10.1186/1471-2156-10-59 PMID: 19772609

24. Al-Zahery N, Pala M, Battaglia V, Grugni V, Hamod MA, Hooshiar Kashani B, et al. 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. 2011 Oct 4; 11:288. https://doi.org/10.1186/1471-2148-11-288 PMID: 21970613

25. Underhill PA, Passarino G, Lin AA, Shen P, Mirazón Lahr M, Foley RA, et al. The phylogeography of Y chromosome binary haplotypes the origins of modern human populations. Ann Hum Genet. 2001 Jan; 65(Pt1):43–62. PMID: 11415522

26. Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, et al. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. Mol Biol Evol. 2007 Jun; 24(6):1300–1311. https://doi.org/10.1093/molbev/msm049 PMID: 17351267

27. Levin BC, Cheng H, Reeder DJ. A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection.Genomics. 1999 Jan 15; 55(2):135–146. https://doi.org/10.1006/geno.1998.5513 PMID: 9933560

28. Van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat. 2009; 30(2):E386–E394. https://doi.org/10.1002/humu.20921 PMID: 18853457

29. Hashad DI, Abou-Zeid AA, Achmawy GA, Allah HM, Saad MA. G2019S mutation of the leucine-rich repeat kinase 2 gene in a cohort of Egyptian patients with Parkinson's disease. Genet Test Mol Biomarkers. 2011 Dec; 15(12):861–866. https://doi.org/10.1089/gtmb.2011.0016 PMID: 21699405

30. El Desoky ES, Khedr EM, Khalil MS, Gasser T. Genetic Analysis of Leucin-Rich Repeat Kinase 2 (LRRK2) G2019S Mutation in a Sample of Egyptian Patients with Parkinson's disease, a Pilot Study. Br J Med Med Res. 2015; 5(3):404–408. https://doi.org/10.9734/BJMMR/2015/10458

31. Al-Mubarak BR, Bohlega SA, Alkhairallah TS, Magrashi AI, AlTurki MI, Khalil DS, et al. Parkinson's disease in Saudi patients: A Genetic Study. PLoS One. 2015 Aug 14; 10(8):e0135950. https://doi.org/10.1371/journal.pone.0135950 PMID: 26274610

32. Alamri Y, MacAskill M, Anderson T, Benamer H. Parkinson's Disease in the Gulf Countries: An Updated Review. Eur Neurol. 2015 74(3–4):222–225. https://doi.org/10.1159/000442283 PMID: 26613525

33. Nebel A, Landau-Tasseron E, Filon D, Oppenheim A, Faerman M. Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. Am J Hum Genet. 2002 Jun; 70(6):1594–1596. https://doi.org/10.1086/340669 PMID: 11992266

34. Lugan B. Histoire des Berbères, un combat identitaire plurimillénaire. Bernard Luganed. 2012.

35. Triki-Fendri S, Sánchez-Diz P, Rey-González D, Ayadi I, Carracedo Á, Rebai A. Paternal lineages in Libya inferred from Y-chromosome haplogroups. Am J Phys Anthropol. 2015 Jun; 157(2):242–251. https://doi.org/10.1002/ajpa.22705 PMID: 25677690

36. Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, Haber M, et al. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. Am J Hum Genet. 2008 Nov; 83 (5):633–462. https://doi.org/10.1016/j.ajhg.2008.10.012 PMID: 18976729

37. Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, et al. 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. 2004 May; 74(5):1023–1034. https://doi.org/10.1086/386295 PMID: 15069642

38. Triki-Fendri S, Sánchez-Diz P, Rey-González D, Alfadhli S, Ayadi I, Ben Marzoug R, et al. Genetic structure of the Kuwaiti population revealed by paternal lineages. Am J Hum Biol. 2016 Mar 4; 28 (2):203–212. https://doi.org/10.1002/ajhb.22773 PMID: 26293354

39. Capelli C, Onofri V, Brisighelli F, Boschi I, Scarnicci F, Masullo M, et al. Moors and Saracens in Europe: estimating the medieval North African male legacy in southern Europe. Eur J Hum Genet. 2009 Jun; 17 (6):848–852. https://doi.org/10.1038/ejhg.2008.258 PMID: 19156170

40. Sierra M, González-Aramburu I, Sánchez-Juan P, Sánchez-Quintana C, Polo JM, Berciano J, et al. High frequency and reduced penetrance of LRRK2 G2019S mutation among Parkinson's disease patients in Cantabria (Spain). Mov Disord. 2011 Nov; 26(13):2343–2346. https://doi.org/10.1002/mds.23965 PMID: 21954089

41. Fadhlaoui-Zid K, Haber M, Martínez-Cruz B, Zalloua P, Benammar Elgaaied A, David Comas D. Genome-Wide and Paternal Diversity Reveal a Recent Origin of Human Populations in North Africa. PLoS One. 2013; 8(11):e80293. https://doi.org/10.1371/journal.pone.0080293 PMID: 24312208

42. Coudray C, Olivieri A, Achilli A, Pala M, Melhaoui M, Cherkaoui M, et al. The complex and diversified mitochondrial gene pool of Berber populations. Ann Hum Genet. 2009 Mar; 73(2):196–214. https://doi.org/10.1111/j.1469-1809.2008.00493.x PMID: 19053990

43. Turchi C, Buscemi L, Giacchino E, Onofri V, Fendt L, Parson W, et al. Polymorphisms of mtDNA control region in Tunisian and Moroccan populations: An enrichment of forensic mtDNA databases with Northern Africa data. Forensic Sci Int Genet. 2009 Jun; 3(3):166–172. https://doi.org/10.1016/j.fsigen.2009.01.014 PMID: 19414164

44. Aboukhalid R, Sturk-Andreaggi K, Bouabdellah M, Squalli D, Irwin JA, Amzazi S. Mitochondrial DNA control region variation from samples of the Moroccan population. Int J Legal Med. 2013 Jul; 127 (4):757–759. https://doi.org/10.1007/s00414-012-0813-4 PMID: 23283404

45. Kéfi R, Stevanovitch A, Bouzaid E, Béraud-Colombe. Diversité mitochondriale de la population de Taforalt (12.000 ans BP-MAROC): une approche génétique à l'étude du peuplement de l'Afrique du Nord. Anthropologie. 2005; XLIII (1):1–11.

46. Richter D, Grün R, Joannes-Boyau R, Steele TE, Amani F, Rué M, et al. The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. Nature. 2017 Jun 7; 546 (7657):293–296. https://doi.org/10.1038/nature22335 PMID: 28593967

47. Watterson GA, Guess HA. Is the most frequent allele the oldest?. Theor Popul Biol. 1977 Apr; 11 (2):141–160. PMID: 867285

48. Donnelly P,Tavaré S. The ages of alleles and a coalescent. Adv Appl Prob. 1986 March; 18(1):1–19. https://doi.org/10.2307/1427237

49. Healy DG, Wood NW, Schapira AH. Test for LRRK2 mutations in patients with Parkinson's disease. Pract Neurol. 2008 Dec; 8(6):381–385. https://doi.org/10.1136/jnnp.2008.162420 PMID: 19015299

50. Diéterlen F, Lucotte G. Origin of the G2019S mutation associated to Parkinson's disease in Europeans and in North Africans. Int J Mod Anthrop. 2010; 3:81–96

51. Campbell CL, Palamara PF, Dubrovsky M, Botigué LR, Fellous M, Atzmon G, et al. North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. Proc Natl Acad Sci USA. 2012 Aug 21; 109(34):13865–13870. https://doi.org/10.1073/pnas.1204840109 PMID: 22869716

52. Change N, Mercier G, Lucotte G. Genetic screening of the G2019S mutation of the LRRK2 gene in Southwest European, North African, and Sephardic Jewish subjects. Genet Test. 2008 Sep; 12(3):333–9. https://doi.org/10.1089/gte.2007.0098 PMID: 18666856

53. Nebot D. Les Tribus oubliées d'Israël. Romillat ed. 2004, 223 p., 1rst ed. 1994; Collection: Terra Hebraïca—ISBN: 2878940830

54. Bokbot Y, Rodriguez-Rodriguez A, Onrubia-Pintado J, Rodriguez-Santana CG, Velasco-Vazquez J, Amarir A. Le complexe funéraire et cultuel d'Adrar Zerzem (Anti-Atlas, Maroc) Résultats préliminaires. Études d'antiquités africaines. 2008; 1(1):21–29. oai: persee:issue/etaf_0768–2352_2008_act_1_1

55. Allouche-Benayoun J, Bensimon D. Les Juifs d'Algérie. Mémoires et identités plurielles. Privat ed. Toulouse. 1989.

56. Ayoun R, Cohen B, Les Juifs d'Algérie, deux mille ans d'histoire. Paris: J-C Lattès ed. 1982; p. 27.

57. Khaldoun Ibn. Histoire des Berbères et des dynasties musulmanes de l'Afrique septentrionale, trad. MacGuckin W.. Paris: Paul Geuthner ed. 1978.

58. Nebot D. La Kahéna, reine d'Ifrikia: Anne Carrière ed. 1998.

59. Ben-Sasson HH. A History of the Jewish People. Harvard University Press, Cambridge, Mass. 1976. ISBN 10: 0674397312

60. Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, et al. Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat Genet. 1995 Feb; 9(2):152–159. https://doi.org/10.1038/ng0295-152 PMID: 7719342

61. Falk R. Genetic markers cannot determine Jewish descent. Front Genet. 2015; 5:462. https://doi.org/10.3389/fgene.2014.00462 PMID: 25653666

62. Henn BM, Botigue LR, Gravel S, Wang W, Brisbin A, Byrnes JK, et al. Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS Genet. 2012 Jan 12; 8(1):e1002397. https://doi.org/10.1371/journal.pgen.1002397 PMID: 22253600