Fruits 74(2), 73–81 | ISSN 0248-1294 print, 1625-967X online | https://doi.org/10.17660/th2019/74.2.3 | © ISHS 2019

Original article

Microsatellite (SSR) markers reveal genetic diversity and population structure in Tunisian pistachio

S. Choulak[1], K. Chatti[1,a] S. Rhouma-Chatti[1], K. Guenni[2], A. Salhi-Hannachi[2], K. Said[1] and N. Chatti[1]

[1] Laboratoire de recherche de Génétique, Biodiversité et Valorisation des Bioressources (LR11ES41), ISBM, Université de Monastir, Monastir 5000, Tunisie

[2] Laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie, Faculté des Sciences de Tunis, Université de Tunis El Manar, Campus Universitaire, Tunis, Tunisie

Summary

Introduction – **Little attention has been directed toward the conservation and evolution of Pistachio's genetic resources in Tunisia. Despite the increased importance of this crop, local pistachio germplasm is far from being adequately studied and used.** *Material and methods* – **The present research aims to study genetic diversity and relationships among three populations (El Guetar, Gafsa Sidi and Bouzid) of Tunisian pistachio (42 accessions) using a set of nine simple sequence repeats (SSR) markers.** *Results and discussion* – **The results revealed that 29 alleles were observed across all the studied genotypes with an average value of 2.9. The observed and expected levels of heterozygosity for the majority of locus indicated that the genotype frequencies in pistachio deviated from the Hardy-Weinberg equilibrium. In addition, *P. atlantica* trees (Battoum) present specific alleles amplified by two different primers. Factorial Correspondence Analysis (FCA) and Neighbor-Joining tree are structured independently from the plant's sex but structured according to geographical origin. Bayesian analysis of STRUCTURE showed that all pistachio individual could most likely be assigned to two genetics' pools named "South" and "Center".** *Conclusion* – **The exposed SSR alleles were successfully used to discriminate between different studied accessions. In addition, this study could be considered as an important step ensuring the conservation and the development of the national database.**

Keywords

Tunisia, *Pistacia vera*, genetic resources management, molecular polymorphism

Significance of this study

What is already known on this subject?

- In Tunisia, important morphological traits were identified according to the pistachio descriptor list (IPGRI, 1997) by Ghrab *et al.* (2012). In addition, ISSR, SRAP and chloroplastic markers were successfully used to study the genetic diversity of Tunisian pistachio (Farès *et al.*, 2009; Guenni *et al.*, 2016; Choulak *et al.*, 2015).

What are the new findings?

- The current study reported herein presented the first comprehensive analysis of diversity in the Tunisian pistachio germplasm using SSR markers. The exposed SSR alleles were successfully used to discriminate between different studied accessions. In addition, our results showed that *Pistacia atlantica* (Battoum) presents specific alleles amplified by two different primers.

What is the expected impact on horticulture?

- By growing mixtures of diversified local germplasm of crops, farmers have been able to select varieties adapted to local farming systems and environmental conditions. Additionally, our result helps to varietal identification as early as the juvenile stage of the plant.

Introduction

Pistacia vera L. (2*n*=30) is a dioecious and wind pollinated member of the *Anacardiaceae* family. It is the only cultivated and commercialized species in the genus *Pistacia*, which produces comestible nuts great enough to be commercially acceptable (Zohary, 1996; Pazouki *et al.*, 2010). Pistachio is cultivated over large areas around the world. Two centers of diversity have been defined, an area from Iran into eastern Turkey, and the second one in southern Turkmenistan and northern Afghanistan (Joley, 1969). *P. vera* has been long spread for nuts all over the Middle East and Mediterranean basin (Badenes and Byrne, 2012). Numerous reports suggest that the Romans were accountable for the propagation of this species within the Mediterranean regions (Hormaza *et al.*, 1994). In Tunisia, pistachio is cultivated since Carthaginians time (Jacquy, 1973).

Despite this long history, little is known today about its genetic variability and wealth (Ghrab *et al.*, 2012). Actually, there is 44,000 ha grown pistachio areas and the national production in 2014 reached 1,200 tonnes (Faostat, 2016). Pistachio trees represent about 11% of the Tunisian total area planted with drupe fruit trees (excluding olive trees and date palm). Several Iranian and Syrian varieties were introduced in Tunisia. The 'Matteur' variety, which resembles the Syrian variety 'Achoury', is the only extensively employed variety in Tunisia (Ghorbel *et al.*, 2001).

The main Tunisian regions producing pistachio are Kasserine, Sidi Bouzid, Gafsa, Sfax, and Kairouan. However, cultivation and productivity of pistachio are still narrow due to uncontrolled pollination and drought condition. In fact, pistachio trees are expensive to expand for the reason of their

[a] Corresponding author: ckhaled2000@yahoo.fr.

long juvenile period and to the need of artificial pollination (Ghorbel *et al*., 2001). On the other hand, a relatively small number of cultivars are currently used in the largest scaled pistachio production (Tous and Ferguson, 1996). This progress in production for few improved cultivars is unfortunately accompanied by a significant decline of wild material and land races (see Ghorbel et *al*., 2001). In addition, such cultivations are also risked to be based on a very strait genetic base (Maggs, 1973; Tous and Ferguson, 1996), making them susceptible to pest or disease attacks. As a result, pistachio breeding is now exposed to a severe loss of genetic diversity.

Researches on genetic diversity and relationships within cultivated species are needed to attract attention to the priorities for the conservation strategies of plant germplasm. Since the middle 1980's, genome identification and selection has surely improved with the advantage of polymerase chain reaction (PCR) skill (Mullis *et al*., 1986). In fact, many studies have described the use of RAPD, AFLP, ISSR methods and chloroplast DNA (Katsiotis *et al.*, 2003; Golan-Goldhirsh *et al*., 2004; Baghizadeh *et al*., 2010; Choulak *et al*., 2015) to identify *Pistacia vera* L. cultivars. These studies have confirmed the efficiency of the used molecular markers to evaluate genetic diversity within studied genotypes. Furthermore, because of their high degree of polymorphism, the possibility of automated scoring of genotypes and their random distribution throughout the genome, microsatellite loci have been proven to be one of the most potent tools for inferring with genetic variability (Burford and Wayne, 1993). In recent times, microsatellites were developed from *Pistacia lentiscus* (Albaladejo *et al*., 2008) and *P. vera* (Ahmad *et al*., 2003; Kolahi-Zonoozi *et al*., 2014; Ziya-Motalebipour *et al*., 2016; Topcu *et al*., 2016) which are applied to discriminate pistachio varieties collected from different regions (Baghizadeh *et al*., 2010; Pazouki *et al*., 2010). Thus, in order to improve the conservation of Tunisian pistachio germplasm, the identification and the preservation of the genetic variability of natural populations, molecular markers are used. The study reported herein presented the first comprehensive analysis of diversity in the Tunisian pistachio via simple sequence repeats (SSR) markers, which are the most consistent, and reproducible molecular markers for genetic diversity studies. This will be of great support for the management of the species.

Materials and methods

Plant material

Field visits were carried out from April 2012 to July 2013 in the central (Sidi Bouzid) and southern (Gafsa and El Guetar) traditional areas of pistachio culture in Tunisia. Forty-two Tunisian pistachio accessions, including 30 domestic female trees, nine males trees or 'Dhokkars' (*P. vera*) and three *P. atlantica* or 'Battoum', were used in the study (Table 1). Several information has been noted: the sex of trees, the name if possible, and grafted or wild plants. Plant material consists of young leaves of each adult tree. Samples were frozen at -80 °C awaiting their use for DNA purification.

DNA extraction and PCR amplification

Total genomic DNA was extracted from frozen leaves of single adult trees by means of a modified MATAB method (Risterucci *et al*., 2000). Before extraction, leaves were ground in liquid nitrogen using a ball mill (type MM2; Retsch, Haan, Germany). The DNA was re-suspended in molecular biology water after isopropanol evaporation. DNA concentrations were measured using the spectrophotometer and its integrity was checked by agarose-gel electrophoresis (Sambrook *et al*., 1989).

According to Ahmad *et al*. (2003), nine SSR loci were considered in this study. Amplifications were realized using a Bio-Rad thermocycler (Bio-Rad Laboratories Inc., Hercules, CA, USA). Each reaction was carried out in a 10-µL reaction mix containing 25 ng of genomic DNA, 0.5 mM $MgCl_2$, 0.1 µM of the reverse and forward primers, 0.1 µM of M13 primer, 0.2 mM of each dNTP, and 1 unit of Taq polymerase. Reverse primers were unlabeled while forward primers were 5'-labeled by one of the following fluorescent compounds: FAM, PET, VIC or NED, to facilitate analysis on automated sequencers. PCR products were observed using capillary electrophoresis on an ABI 3500xl DNA Analyzer. Samples were prepared by mixing 2 µL of diluted PCR products with 10 µL of the Mix (250 µL formamide and 3 µL GeneScan 600 LIZ (Applied Biosystems).

Genetic diversity analysis

The size of different SSR fragments for the 42 studied accessions was identified with the GENEMAPPER V4.1 software. The total number of alleles, the expected (H_e) heterozygosity values, and the observed (H_o) were calculated in GenAlEx (Peakall and Smouse, 2006) according to Nei (1978). Genetic diversity measures (H_t, total gene diversity; H_s, mean genetic diversity within the population; G_{st}, the coefficient of gene differentiation) were also tested using GenAlEx.

Genetic relationships

A dendrogram was constructed using the Darwin software (Perrier and Jacquemoud-Collet, 2006). This program deducing evolutionary and phylogeny trees from distance based methods of evolutionary similarity. Based on this dissimilarity, an unrooted neighbor-joining tree was constructed (Saitou and Nei, 1987). Additionally, Factorial Correspondence Analysis (FCA) was performed with GENETIX 4.05 program, to build a tri-dimensional graphic representation of the genetic variability of accessions according to their genetic origin.

Genetic differentiation and population structure

Wright (1965) developed two indices of fixation: the "fixation index" or F_{st} and F_{is} indices. The estimator (F_{st}) presents the proportion of genetic diversity due to allele frequency differences among populations. F_{is} indices are used to estimate the average departure of genotype frequencies from Hardy-Weinberg expectations within populations. These fixation indices were computed according to the formula of Weir and Cockerham (1984) using the program GENETIX 4.05 (Belkhir *et al*., 2004).

To infer population structure of studied pistachio cultivars, Bayesian Markov Chain Monte Carlo algorithm, STRUCTURE, was used (Pritchard *et al*., 2000). This program allows detecting the presence of distinct populations, assign individuals to populations and study hybrid zones. Moreover, this program presumes Hardy-Weinberg equilibrium (HWE) and linkage equilibrium within clusters. Without any information about the geographic origin of the cultivars, the STRUCTURE program was run for ten independent replicate runs for each k value (k ranging from 1 to 10). Therefore, a possible number of genetic clusters (k) in our sample set was investigated by a web-based program for collating results produced by the STRUCTURE program. In fact, Structure Harvester 0.694 (Dent, 2014; web version) provided a fast way to evaluate

and envisage likelihood values from multiple values of *k* and hundreds of iterations for the easier revelation of the number of genetic groups that best fit the data.

Results and discussion

SSR analysis

A total of 29 alleles were observed across all studied genotypes (Table 2). The number of alleles per locus ranged from one for primers Ptms-9, Ptms-40, and Ptms-47, to eight for Ptms-7, with an average value of 2.9. With the exception of Ptms-9, Ptms-40, and Ptms-47 for which the three populations were monomorphic at these loci, all the other used primers revealed different levels of polymorphisms. This result highlights a high degree of genetic diversity within the studied Tunisian pistachios. Additionally, average allele per locus was significantly higher than that observed for Syrian, Turkish, and American cultivated pistachios (Ahmad *et al.*, 2003), but less than that detected in Iranian pistachio, where different species were included in the study (Pazouki *et al.*, 2010).

Table 1. Label, sex and origin of Tunisian pistachio cultivars studied.

Geographic origin	Labels	Sex	Species	Plant nature
El Guetar (Tunisia)	GT1	♀	*P. vera*	Grafted plant
	GT2	♀	*P. vera*	Grafted plant
	GT3	♂	*P. vera*	Grafted plant
	GT4	♂	*P. vera*	Grafted plant
	GT5	♀	*P. vera*	Grafted plant
	GT6	♂	*P. vera*	Grafted plant
	GT7	♀	*P. vera*	Grafted plant
	GT8	♀	*P. vera*	Grafted plant
	GT9	♂	*P. vera*	Grafted plant
	GT10	♀	*P. vera*	Grafted plant
	GT11	♀	*P. vera*	Grafted plant
	GT12	♀	*P. vera*	Grafted plant
	GT13	♀	*P. vera*	Grafted plant
	GT14	♀	*P. vera*	Grafted plant
	GT15	♂	*P. vera*	Grafted plant
	GT16	♀	*P. vera*	Grafted plant
	GT17	♀	*P. vera*	Grafted plant
	GT18	♀	*P. vera*	Grafted plant
	GT19	♀	*P. vera*	Grafted plant
	GT20	♀	*P. vera*	Grafted plant
Gafsa (Tunisia)	GF1	♂	*P. vera*	Grafted plant
	GF2	♂	*P. vera*	Grafted plant
	GF8	♀	*P. vera*	Grafted plant
	GF9	♀	*P. vera*	Grafted plant
	GF10	♀	*P. vera*	Grafted plant
	Battoum1		*P. atlantica*	Wild seedling
Sidi Bouzid (Tunisia)	SB1	♀	*P. vera*	Grafted plant
	SB2	♀	*P. vera*	Grafted plant
	SB3	♀	*P. vera*	Grafted plant
	SB4	♀	*P. vera*	Grafted plant
	SB5	♂	*P. vera*	Grafted plant
	SB6	♂	*P. vera*	Grafted plant
	SB7	♀	*P. vera*	Grafted plant
	SB8	♀	*P. vera*	Grafted plant
	SB9	♀	*P. vera*	Grafted plant
	SB10	♀	*P. vera*	Grafted plant
	SB11	♀	*P. vera*	Grafted plant
	Matteur1	♀	*P. vera*	Grafted plant
	Matteur2	♀	*P. vera*	Grafted plant
	Irani	♀	*P. vera*	Grafted plant
	Battoum2		*P. atlantica*	Wild seedling
	Battoum3		*P. atlantica*	Wild seedling

Table 2. Genetic parameters of nine used loci. Na: number of alleles per locus, Ne: Mean effective number of alleles, H_e: expected heterozygosity, H_o: observed heterozygosity, F_{st} and F_{is}: Wright's analysis of hierarchical F-statistics, H_t: total genetic diversity and H_s: the mean genetic diversity within population.

Locus	Repeat motif	Allele size	Na	Ne	H_o	H_e	F_{st}	F_{is}	G_{st}	H_t	H_s
Ptms-3	(CA) 16	132-140	3	1.268	0.000	0.204	0.037	1.000	0.037	0.212	0.204
Ptms-7	(CA) 15	175-193	5	2.786	0.972	0.620	0.027	0.569	0.027	0.673	0.619
		197-201	3	2.057	0.833	0.495	0.135	0.682	0.048	0.569	0.541
Ptms-9	(CA) 7	128	1	1.000	0.000	0.000	-	-	-	0.000	0.000
Ptms-11	(CT) 13	162-202	4	2.613	1.000	0.604	0.108	0.656	0.108	0.677	0.603
Ptms-12	(CT) 21	137-203	4	2.256	1.000	0.546	0.011	0.831	0.001	0.552	0.546
Ptms-31	(CT) 20	135-203	5	2.961	1.000	0.639	0.158	0.565	0.158	0.759	0.639
Ptms-40	(CTTT) 4	199	1	1.000	0.000	0.000	-		-	0.000	0.000
Ptms-42	(CTT) 10	162-201	2	2.000	1.000	0.500	0.000	1.000	0.000	0.500	0.500
Ptms-47	(CTT) 13	146	1	1.000	0.000	0.000	-	-	-	0.000	0.000
All loci			2.9	1.894	0.581	0.360	0.077	0.609	0.064	0.558	0.522

For every locus, allele frequency was calculated using GenAlEx. Values ranged from 0.015 of Ptms-12 to 0.737 of Ptms-3 (Supplemental Information, Table S1). In accordance with Ahmad *et al*., (2003), in the present survey, no correlation was established between the repeated number of microsatellites and their polymorphism or their genetic parameters.

Furthermore, our results showed that *Pistacia atlantica* (Battoum) presents specific alleles amplified by two different primers. In fact, markers Ptms-3 and Ptms-7 were yielded to 132 bp, and 181 bp, 191 bp respectively (Supplemental Information, Table S1). On the other side, it is worth noting that the resultant genetic diversity was made independently from the sex of trees.

Genetic diversity

A heterozygote excess was recorded for the Ptms-7, Ptms-11, Ptms-12, Ptms-31, and Ptms-42 (Table 2). Only Ptms-3 showed a deficiency of heterozygosity. The Ptms-7 locus offered the highest number of alleles (8) and Ptms-31 gave the maximum expected heterozygosity value (Table 2).

The observed heterozygosity (H_o) ranged from 0 to 1 and the mean H_o was 0.581. The expected heterozygosity (H_e) for individual loci varied from 0 to 0.639 (Ptms-31), with an average value of 0.360 (Table 2). The quantities H_s and H_t are calculated for all the loci examined. Based on the studied SSR markers, the average within-population genetic diversity (H_s) was 0.522 and the total diversity (H_t) was 0.558. When monomorphic loci were included, H_s and H_t decreased to zero. The H_s and the H_t values were almost similar suggesting that maximum of local variability was preserved for all studied loci.

In addition, we tried to calculate the heterozygosis level in the geographical groups. The studied pistachio varieties were subdivided into three groups (El Guetar, Gafsa and Sidi Bouzid). Sampling areas showed an excess of heterozygosis and an important deviation from the Hardy-Weinberg equilibrium ($H_0 > H_e$) (Supplemental Information, Table S2). The privileged value of expected heterozygosity (H_e) was observed for Gafsa samples compared with the other regions. The population of Gafsa showed the highest genetic diversity, followed by Sidi Bouzid and El Guetar populations. This result indicated that the level of genetic diversity within the population of Gafsa was important and carried rich genetic variability. Consequently, Gafsa cultivars might be the results of few introduced genotypes, which have been, in the first place, propagated in El Guetar and Sidi Bouzid regions.

In general, Tunisian pistachio varieties displayed a moderate level of heterozygosity. This conclusion is similar to the obtained in several previous studies, on pistachio cultivars, focusing on a single country, like Iran (Salehi Shanjani *et al*., 2009; Pazouki *et al*., 2010) or numerous countries and species (Ahmad *et al*., 2003; Mirzaei *et al*., 2006; Arabnezhad *et al*., 2011). These studies supported low genetic variation in cultivated pistachio using several molecular markers including SSR markers. In fact, detecting a low level of polymorphism may suggest the presence of a narrow genetic variability among cultivated pistachios (Ahmad *et al*., 2003).

Furthermore, the levels of heterozygosity values, for different studied loci, showed that pistachio's genotype frequencies diverged from Hardy-Weinberg equilibrium. The observed deviation from the HWE in our study could be due to the limited sample size and the system of mating in pistachio (Templeton, 2006; Kolahi-Zonoozi *et al*., 2014). Otherwise, any deviation from HWE may indicate a biological process like the selection by farmers of a genotype combination for interesting agronomical traits (Zehdi-Azzouzi *et al*., 2015), or else it can be related to the monovarietal crops (or monoculture) of mains varieties cultivated in Tunisia ('Matteur' and 'El Guetar'). Alternatively, breeding events and the vegetative propagation system of *Pistacia* species can, also, explain this divergence. In fact, the long juvenility imposes severe limits on breeding efforts by farmers, who would have to wait several years before the fruits can be selected and grown. The first farmers avoided this problem by adopting vegetative reproduction. Vegetative propagation is the main form of reproduction in Tunisian pistachio. These results are in agreement with other work on dioecious species as the Fig tree (*Ficus carica* L.) (Chatti *et al*., 2010) and date palm (*Phoenix dactylifera* L.) (Zehdi *et al*., 2012).

Genetic differentiation and population structure

Genetic differentiation (F_{st}) values ranged from 0 to 0.158 at Ptms-42 and Ptms-31 respectively (Table 2). Furthermore, the coefficient (F_{is}) showed negative values; indicate heterozygote excess compared with HWE expectations. Otherwise, Nei's G_{st} was typically used for describing the average amount of differentiation observed over multiple loci, the multi-locus values of G_{st} were 0.064 (Table 2).

Wright showed that estimates of F_{st} are important in the measure of the genetic differentiation and in the descriptive statistics of evolutionary genetics, while F_{st} is more than just a genetic differentiation measurement (Holsinger and Weir,

2009). F_{st} is conversely related to the degree of resemblance among individuals within populations. If F_{st} is small, it means that allele frequencies within each population are very similar (Holsinger and Weir, 2009). In addition, F_{st} parameter is an indirect estimate of gene flow, since $F_{st} = 1 / (4Nm+1)$ (where *Nm* is the mean number of migrants per generation among populations.), *Nm* increases, F_{st} decreases, and vice versa (Whitlock and Mccauley, 1999). The low values of F_{ST} indicate high connectivity between populations, and this easy connection makes high breeding possible and gene exchanges among regions. This can be the result of the relative proximity of geographical sampling areas for the studied accessions (only 110 km apart). Same interpretations are cited by Guenni *et al*. (2016), where Tunisian pistachio cultivars from different geographic groups (Kasserine, El-Guetar, Gabes, and Sfax), showed a large genetic diversity and low level of genetic differentiation between them. These observations were justified by the high gene flow level (*Nm* = 1.127).

To illustrate a synthetic representation of the genetic variability distribution of the considered accessions, a multivariate analysis (FCA) was performed. The three first dimensions of factorial correspondence analysis represented, respectively, 29.95%, 16.37% and 14.86% of the total genetic variability. These observations gave an idea about a genetic diversity, which structured according to species (Figure 1). Moreover, the cultivars of *Pistacia atlantica* (Battoum), originating from Gafsa and Sidi Bouzid, were clearly different from *Pistacia vera* accessions. However, Gafsa cultivars represented a larger scale distribution compared to the other regions and showed an overlap between them (Figure 1). We noted also that El Guetar and Sidi Bouzid ecotypes were clearly separated from each other (Figure 1). This result supported the dispersion of cultivars exposed by the NJ tree (Figure 2).

In fact, the dendrogram classified the 42 studied accessions into two major clusters. The first one (I) contained all Sidi Bouzid genotypes, some varieties of Gafsa (GF8, GF9 and GF10) and Bottoum's cultivars. The second cluster (II) represented the totality of El Guetar genotypes with some varieties of Gafsa (GF1 and GF2). Topological NJ dendrogram showed that groupings of Tunisian varieties were made independently of the sex of the trees, but depending on their geographical origin. Indeed, El-Guetar varieties and Sidi-Bouzid ones are clearly separated from each other.

However, *Pistacia atlantica* (Battoum) cultivars belonged to the same sub-cluster and they were obviously detached from the *P. vera* ecotypes. Foreign cultivar *Irani* introduced in the Tunisian pistachio plantations was not significantly separated from the autochthonous ones (Figure 2). Therefore, considered together, the factorial correspondence analysis and the NJ tree showed, in the first part, the presence of a geographical distribution of *Pistacia vera* in Tunisia and the significant differentiation of *P. atlantica* (Battoum), in the second part (Figure 2).

The STRUCTURE algorithm, using multi-locus genotype data to investigate population structure, was run without precedent information about the geographic coordinates of the samples. This program was run 10 times for every k value from 1 to 6, with each run including a burn-in period of 250,000 iterations and a length of 2.10^6 iterations. According to Evanno *et al*. (2005), the log likelihood between following k values (ΔK) showed a higher level of clustering at k = 2 for the considered pistachio accessions (Figure 3). At k = 2, pistachio accessions were clearly differentiated into two geographic groups, the first one, named "South" cluster, consisted of accessions from El Guetar and some accessions originated from Gafsa. The second cluster, named "Center",

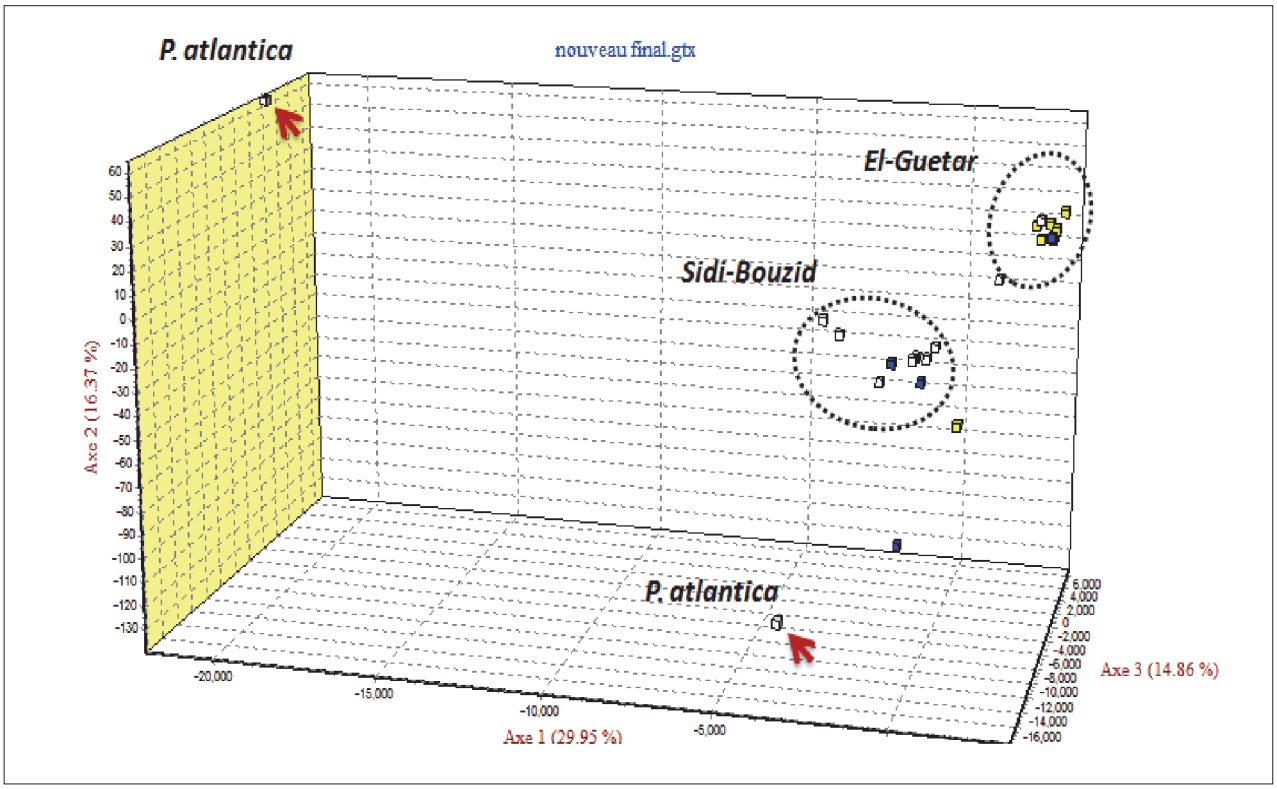

Figure 1. Factorial Correspondence Analysis scatter plot of 42 pistachio cultivars based on the first three principal coordinates. *Battoum* cultivars are represented by arrows. Yellow dots, 'El Guetar' varieties; blue dots, Gafsa accessions; white dots, Sidi Bouzid accessions.

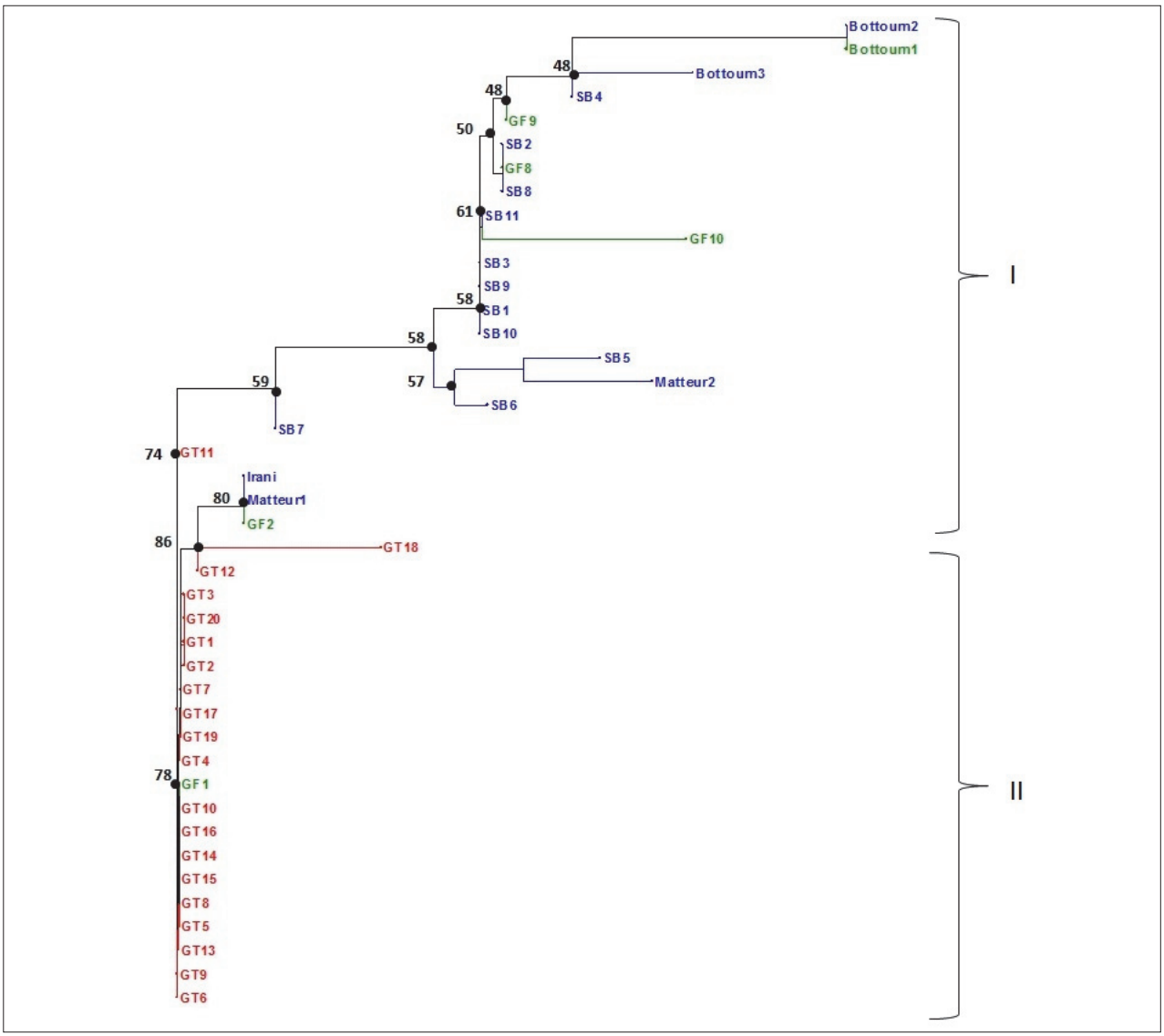

Figure 2. Dendrogram of different cultivar's groups produced by the Neighbour-joining clustering method and based on microsatellites markers. Colour of the branches represent accession/population assignment, red: El-Guetar (GT) accessions; green: Gafsa accessions (GF/ Battoum1); blue: accessions from Sidi Bouzid region (SB/ Matteur1 and Matteur2/ Battoum2 and Battoum3).

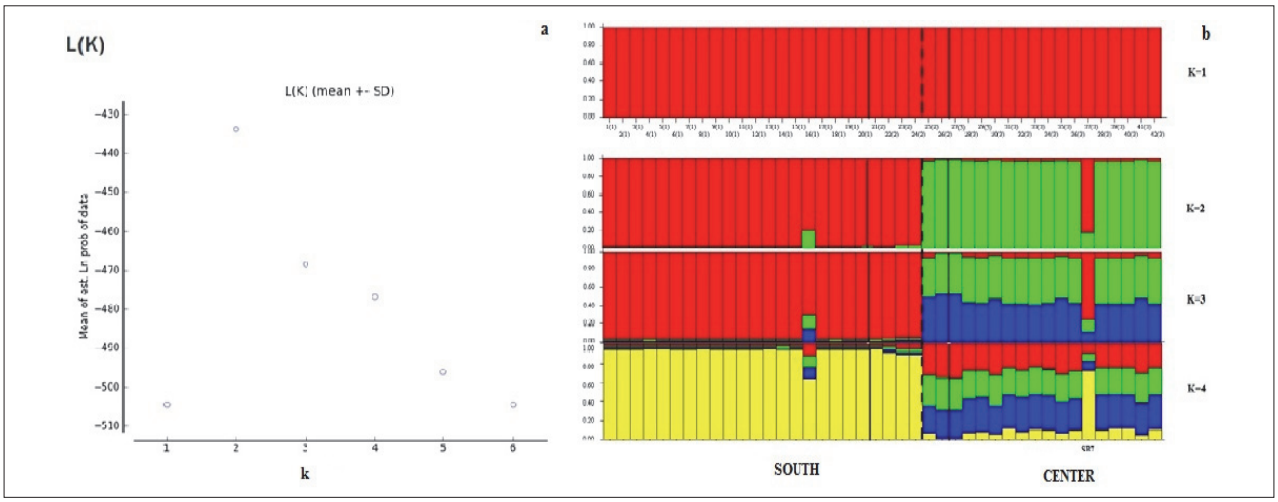

Figure 3. Estimated population structure based in microsatellites markers using STRUCTURE software. a) Mean log likelihood Ln P(X|K) as a function of the number of genetic clusters (K), b) Population structure, each color indicates a Group with the same genetic background. Bold lines represent the limit of each population; Population 1: El-Guetar (red); Population 2: Gafsa (green); Population 3: Sidi-Bouzid (blue). The interrupted line represents the limit of the two clusters: Center and South.

was composed of the accessions from Sidi Bouzid and a few samples from Gafsa. Thus, Gafsa cultivars could be the consequence of few introduced genotypes that have been initially propagated in El Guetar and Sidi Bouzid regions. Certainly, human impact on this region and geographic localization of Gafsa farms were the reasons for these observations. It is important to note that SB7 was the only admixture observed in the center cluster, this accession has been introduced from south to Sidi Bouzid farms. Sidi Bouzid accessions significantly deviated from the El Guetar ones, suggesting two different autochthonous origins. Thus, this discrimination between cultivars proved the presence of multiple events of domestication in the prospected areas. The admixtures observed by STRUCTURE analysis may be a result of cross-pollination and breeding processes. In agreement with Kolahi-Zonoozi *et al*. (2014), only moderate genetic structure was observed in 45 commercial cultivars of Iranian pistachio using microsatellites markers. Although pistachio is dioecious, out-breeding and clonally propagation as well as cultivator selection pressure for commercial traits, could have led to the medium level of genetic structure observed in this and in other studies (Ahmad *et al*., 2003; Arabnezhad *et al*., 2011; Kolahi-Zonoozi *et al*., 2014). Additionally, the STRUCTURE analysis reinforced the results obtained by the F_{st} estimator and phylogenetic reconstruction.

Conclusions

This study is an important step that ensures the conservation and the development of a national database. It could be used to plan sampling programs and assist in germplasm maintenance. Moreover, co-dominance and high polymorphism can determine variation in breeding strategies systems. SSR markers were also efficient in the establishment of genetic relationships between the investigated species. In fact, the revealed alleles were successfully used to discriminate *Pistacia atlantica* cultivars at the molecular level. In addition, the absence of SSR alleles in association with the tree's sex suggested that *Pistacia vera* has a similar ancestor for female and male ('Dhokkars') trees.

Nevertheless, the exploitation of additional SSR markers and the enlargement of the surveys areas clearly provide more precise results. Besides, more molecular markers are accessible to investigate pistachio systematic and cultivar relatedness. These tools will significantly promote breeding programs, genetic conservation, and management of the species.

Acknowledgments

This work was supported by grants from the Tunisian Ministry of Higher Education and Scientific Research. This study was conducted in the Laboratory of CIRAD-BIOS-UMR AGAP (France). We extend our thanks to Dr. Angelique D'hont and Céline Cardi for their assistance.

References

Ahmad, R., Ferguson, L., and Southwick, S.M. (2003). Identification of pistachio (*Pistacia vera* L.) nuts with microsatellite markers, J. Am. Soc. Hortic. Sci. *128*, 898–903. https://doi.org/10.21273/JASHS.128.6.0898.

Albaladejo, R.G., Sebastiani, F., Aparicio, A., Buonamici, A., Gonzalez-Martinez, S.C., and Vendramins, G.G. (2008). Development and characterization of eight polymorphic microsatellite loci from *Pistacia lentiscus* L. (*Anacardiaceae*). Mol. Ecol. Res. *8*, 904–906. https://doi.org/10.1111/j.1755-0998.2008.02110.x.

Arabnezhad, H., Bahar, M., and Tajabadi-Pour, A. (2011). Evaluation of genetic relationships among Iranian pistachios using microsatellite markers developed from Pistacia khinjuk Stocks. Sci. Hortic. *128*(3), 249–254. https://doi.org/10.1016/j.scienta.2011.01.028.

Badenes, M.L., and Byrne, D.H. (2012). Fruit Breeding. Handbook of Plant Breeding *8*, 208. https://doi.org/10.1007/978-1-4419-0763-9.

Baghizadeh, A., Noroozi, S.H., and Jalali-Javaran, M. (2010). Study on genetic diversity of some Iranian Pistachio (*Pistacia vera* L.) cultivars using random amplified polymorphic DNA (RAPD), inter sequence repeat (ISSR) and simple sequence repeat (SSR) markers: A comparative study. AJB *9*(45), 7632–7640.

Belaj, A., Satovic, Z., Cipriani, G., Baldoni, L., Testolin, R., Rallo, L., and Trujillo, I. (2003). Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. Theor. Appl. Genet. *107*(4), 736–744. https://doi.org/10.1007/s00122-003-1301-5.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations (Montpellier, France: Université de Montpellier II, Laboratoire Génome, Populations, Interactions CNRS UMR 5000).

Burford, M.W., and Wayne, R.K. (1993). Microsatellites and their application to population genetic studies. Curr. Opin. Genet. Dev. *3*, 939–943. https://doi.org/10.1016/0959-437X(93)90017-J.

Chatti, K., Baraket, G., Ben Abdelkrim, A., Saddoud, O., Mars, M., Trifi, M., and Salhi Hannachi, A. (2010). Development of molecular tools for characterization and genetic diversity analysis in Tunisian Fig (*Ficus carica* L.) cultivars. Biochem. Genet. *48*, 789–806. https://doi.org/10.1007/s10528-010-9360-1.

Choulak, S., Rhouma-Chatti, S., Marzouk, Z., Said, K., Chatti, N., and Chatti, K. (2015). Chloroplast DNA analysis of Tunisian pistachio (*Pistacia vera* L.): Sequence variations of the intron *trn*L (UAA). Sci. Hortic. *191*, 57–64. https://doi.org/10.1016/j.scienta.2015.04.037.

Dent, E.A. (2014). STRUCTURE HARVESTER version 0.694. http://taylor0.biology.ucla.edu/structureHarvester/.

Elshibli, S., and Korpelainen, H. (2008). Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan. Genetica *134*, 251–260. https://doi.org/10.1007/s10709-007-9232-8.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Faostat (2016). http://faostat3.fao.org.

Ghorbel, A., Ben Salem-Fnayou, A., Chatibi, A., and Twey, M. (2001). Genetic resources of *Pistacia* in Tunisia. In Towards a Comprehensive Documentation and Use of *Pistacia* Genetic Diversity in Central and West Asia, North Africa and Europe, S. Padulosi, and A. Hadj-Hassan, eds. (Rapport of the IPGRI Workshop in Irbid, Jordan), p. 62-71.

Ghrab, M., Zribi, F., Ben Mimoun, M., and Rhouma, A. (2012). Morphological investigation of genetic diversity of pistachio (*Pistacia vera* L.) germplasm in arid land of Tunisia. Plant. Ecol. Evol. *145*(3), 363–372. https://doi.org/10.5091/plecevo.2012.614.

Golan-Goldhirsh, A., Barazani, O., Wang, Z.S., Khadkal, D.K., Saunders, J.A., Kostiukovsky, V., and Rowland, L.J. (2004). Genetic relationships among Mediterranean *Pistacia* species evaluated by RAPD and AFLP markers. Plant. Syst. Evol. *246*, 9–18. https://doi.org/10.1007/s00606-004-0132-4.

Guenni, K., Aouadi, M., Chatti, K., and Salhi-Hannachi, A. (2016). Analysis of genetic diversity of Tunisian pistachio (*Pistacia vera* L.) using sequence-related amplified polymorphism (SRAP) markers. GMR *15*(4), 1–15. https://doi.org/10.4238/gmr15048760.

Hormaza, J.I., Dollo, L., and Polito, V.S. (1994). Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. Theor. Appl. Genet. *89*, 9–13. https://doi.org/10.1007/BF00226975.

IPGRI (1997). Descriptors for Pistachio (*Pistacia vera* L.) (Rome, Italy: International Plant Genetic Resources Institute).

Jacquy, P. (1973). La culture du pistachier en Tunisie. AGP. TUN/72/003, Tunis, 97 pp.

Joley, L.E. (1969). Pistachio. In Handbook of North American Nut Trees, R.A. Jaynes, ed. (Knoxville: Northern Nut Grower Assoc.), p. 348–361.

Katsiotis, A., Hagidimitriou, M., Drossoul, A., Pontikis, C., and Loukas, M. (2003). Genetic relationships among species and cultivars of *Pistacia* using RAPDs and AFLPs. Euphytica *132*, 279–286. https://doi.org/10.1023/A:1025027323184.

Kolahi-Zonoozi, S.H., Mardi, M., Zeinalabedini, M., Pirseyedi, S.M., Mahmoodi, P., Tabatabaei, I., Mosavi-Derazmahalleh, S.M., Farsi, M., Ebrahimi, M.A., Khayam-Nekoui, S.M., and Ahmadi, K. (2014). Development of 12 new SSR markers for genetic diversity and structure analysis in pistachio (*Pistacia vera* L.). JHSB *89*(6), 707–711.

Maggs, D.H. (1973). Genetic resources in Pistachio. Plant. Genet. Res. Newsl. *2*, 7–15.

Mirzaei, S., Bahar, M., and Sharifnabi, B. (2006). A phylogenetic study of Iranian wild pistachio species and some cultivars using RAPD markers. Acta Hortic. *726*, 30–43.

Mullis, K., Faloona, S., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harbor Symp. Quant. Biol. *51*, 263–273. https://doi.org/10.1101/SQB.1986.051.01.032.

Nei, M. (1978). Molecular Evolutionary Genetics (New York, USA: Columbia University Press).

Pazouki, L., Mardi, M., Salehi-Shanjani, P., Hagidimitriou, M., Pirseyedi, S.M., Naghavi, M.R., Avanzato, D., Vendramin, E., Kafkas, S., Ghareyazie, B., Ghaffari, M.R., and Khayam-Nekouei, S.M. (2010). Genetic diversity and relationships among Pistacia species and cultivars. Conserv. Genet. *11*, 311–318. https://doi.org/10.1007/s10592-009-9812-5.

Peakall, R., and Smouse, P.E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes *6*, 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x.

Perrier, X., and Jacquemoud-Collet, J.P. (2006). DARwin software. http://darwin.cirad.fr/darwin.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.

Risterucci, A.M., Grivet, L., N'goran, J., Pieretti, I., Flament, M.H., and Lanaud, C. (2000). A high-density linkage map of *Theobroma cacao* L. Theor. Appl. Genet. *101*(5), 948–955. https://doi.org/10.1007/s001220051566.

Saitou, N., and Nei, M. (1987). The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*(4), 406–425.

Salehi-Shanjani, P., Mardi, M., Pazouki, L., Hagidimitriou, M., Avanzato, D., Pirseyedi, S.M., Ghaffari, M.R., and Khayam-Nekoui, S.M. (2009). Analysis of the molecular variation between and within cultivated and wild *Pistacia* species using AFLPs. Tree Genet. Genomes *5*, 447–458. https://doi.org/10.1007/s11295-008-0198-1.

Sambrook, J., Frithsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (New York, USA: Cold Spring Harbor Laboratory).

Templeton, A.R. (2006). Population Genetics and Micro-evolutionary Theory. (Hoboken, NJ, USA: John Wiley and Sons Inc.), 681 pp. https://doi.org/10.1002/0470047356.index.

Topcu, H., Nergiz, C., and Kafka, S. (2016). Novel microsatellite markers in *Pistacia vera* L. and their transferability across the genus *Pistacia*. Sci. Hortic. *198*, 91–97. https://doi.org/10.1016/j.scienta.2015.11.012.

Tous, J., and Ferguson, L. (1996). Mediterranean fruits. In Progress in New Crops, J. Janick, ed. (Arlington, VA: ASHS Press), p. 416–430.

Weir, B., and Cockerham, C. (1984). Estimating F statistics for the analysis of population structure. Evolution *38*(6), 1358–1370.

Whitlock, M.C., and Mccauley, D.E. (1999). Indirect measures of gene flow and migration: F_{ST} = 1/(4Nm+ 1). Heredity *82*(2), 117–125. https://doi.org/10.1038/sj.hdy.6884960.

Wright, S. (1965). The interpretation of population structure by *F* statistics with special regard to system of mating. Evolution *19*, 395–420. https://doi.org/10.1111/j.1558-5646.1965.tb01731.x.

Zehdi, S., Cherif, E., Rhouma, S., Santoni, S., Salhi-Hannachi, A., and Pintaud, J.C. (2012). Molecular polymorphism and genetic relationships in date palm (*Phoenix dactylifera* L.): The utility of nuclear microsatellite markers. Sci. Hortic. *148*, 255–263. https://doi.org/10.1016/j.scienta.2012.10.011.

Zehdi-Azouzi, S., Cherif, E., Moussouni, S., Gros-Balthazard, M., Abbas-Naqvi, S., Ludeña, B., Castillo, K., Chabrillange, N., Bouguedoura, N., Bennaceur, M., Si-Dehbi, F., Abdoulkader, S., Daher, A., Terral, J.F., Santoni, S., Ballardini, M., Mercuri, A., Ben Salah, M., Kadri, K., Othmani, A., Littardi, C., Salhi-Hannachi, A., Pintaud, J.C., and Aberlenc-Bertossi, F. (2015). Genetic structure of the date palm (*Phoenix dactylifera*) in the Old World reveals a strong differentiation between eastern and western populations. Ann. Bot. *116*(1), 101–112. https://doi.org/10.1093/aob/mcv068.

Zhebentyayeva, T.N., Reighard, G.L., Gorina, V.M., and Abbott, A.G. (2003). Simple Sequence Repeat (SSR) for assessment of genetic variability in apricot germplasm. Theor. Appl. Genet. *106*(3), 435–444. https://doi.org/10.1007/s00122-002-1069-z.

Ziya-Motalebipour, E., Kafkas, S., Khodaeiaminjan, M., Çoban, N., and Gözel, H. (2016). Genome survey of pistachio (*Pistacia vera* L.) by next generation sequencing: Development of novel SSR markers and genetic diversity in *Pistacia* species. BMC Genomics *17*(1), 998. https://doi.org/10.1186/s12864-016-3359-x.

Zohary, D. (1996). The genus *Pistacia* L. In Taxonomy, Distribution, Conservation and Uses of *Pistacia* Genetic Resources, S. Padulosi, T. Caruso, and E. Barone, eds. (Palermo: IPGRI), p. 1–11.

Received: Jan. 00, 2018
Accepted: Jan. 00, 2018

Table S1. Allele size and frequency estimated for each SSR locus from three origins of pistachio accessions.

Locus	Allele	Size	Allele frequency				
			P. vera			*P. atlantica*	Total
			El Guetar	Gafsa	Sidi Bouzid		
Ptms-3	A1	132	0.000	0.000	0.000	0.666	0.166
	A2	136	0.050	0.000	0.000	0.333	0.095
	A3	140	0.950	1.000	1.000	0.000	0.737
Ptms-7a	B1	175	0.454	0.500	0.500	0.000	0.363
	B2	181	0.000	0.000	0.000	0.500	0.125
	B3	189	0.000	0.125	0.115	0.000	0.060
	B4	191	0.000	0.000	0.000	0.500	0.125
	B5	193	0.541	0.375	0.384	0.000	0.325
Ptms-7b	B6	197	0.500	0.250	0.500	–	0.312
	B7	199	0.000	0.250	0.000	–	0.062
	B8	217	0.500	0.500	0.500	–	0.375
Ptms-9	C1	128	1.000	1.000	1.000	1.000	1.000
Ptms-11	D1	162	0.000	0.250	0.500	0.500	0.312
	D2	164	0.500	0.250	0.000	0.000	0.187
	D3	200	0.000	0.062	0.045	0.333	0.110
	D4	202	0.500	0.437	0.454	0.166	0.389
Ptms-12	E1	137	0.500	0.437	0.500	0.500	0.484
	E2	139	0.000	0.062	0.000	0.000	0.015
	E3	201	0.500	0.437	0.500	0.500	0.483
	E4	203	0.000	0.062	0.000	0.000	0.015
Ptms-31	F1	135	0.000	0.000	0.090	0.000	0.022
	F2	137	0.500	0.250	0.045	0.000	0.198
	F3	139	0.500	0.250	0.045	0.000	0.198
	F4	201	0.000	0.250	0.454	0.500	0.301
	F5	203	0.500	0.250	0.045	0.000	0.198
Ptms-40	G1	199	1.000	1.000	1.000	1.000	1.000
Ptms-42	H1	162	0.530	0.530	0.530	0.530	0.530
	H2	201	0.460	0.480	0.450	0.490	0.470
Ptms-47	I1	146	1.000	1.000	1.000	1.000	1.000

Table S2. Genetic diversity parameters within each group of pistachio identified on the basis of their geographic origin in Tunisia. H_e: expected heterozygosity, H_o: observed heterozygosity, F_{is}: the inbreeding coefficient of an individual

Populations	H_e	H_o	F_{is}	Mean number of allele/locus
El Guetar	0.3092	0.5917	-0.9086	1.7
Gafsa	0.4164	0.6000	-0.3308	2.7
Sidi Bouzid	0.3709	0.6000	-0.5962	2.6