Analysis of genetic diversity of Tunisian pistachio (*Pistacia vera* L.) using sequence-related amplified polymorphism (SRAP) markers

K. Guenni[1], M. Aouadi[1], K. Chatti[2] and A. Salhi-Hannachi[1]

[1]Laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie,
Faculté des Sciences de Tunis, Université de Tunis El Manar,
Campus Universitaire, Tunis, Tunisie
[2]Laboratoire de Génétique, Biodiversité et Valorisation des Bio Ressources,
Institut Supérieur de Biotechnologie de Monastir, Université de Monastir,
Monastir, Tunisie

Corresponding author: A. Salhi-Hannachi
E-mail: Amel.SalhiHannachi@fsb.rnu.tn

ABSTRACT. Sequence-related amplified polymorphism (SRAP) markers preferentially amplify open reading frames and were used to study the genetic diversity of Tunisian pistachio. In the present study, 43 *Pistacia vera* accessions were screened using seven SRAP primer pairs. A total of 78 markers was revealed (95.12%) with an average polymorphic information content of 0.850. The results suggest that there is strong genetic differentiation, which characterizes the local resources (G_{ST} = 0.307). High gene flow (N_m = 1.127) among groups was explained by the exchange of plant material among regions.

Analysis of molecular variance revealed significant differences within groups and showed that 73.88% of the total genetic diversity occurred within groups, whereas the remaining 26.12% occurred among groups. Bayesian clustering and principal component analysis identified three pools, El Guettar, Pollenizers, and the rest of the pistachios belonging to the Gabès, Kasserine, and Sfax localities. Bayesian analysis revealed that El Guettar and male genotypes were assigned with more than 80% probability. The BayeScan method proposed that locus 59 (F13-R9) could be used in the development of sex-linked SCAR markers from SRAP since it is a commonly detected locus in comparisons involving the Pollenizers group. This is the first application of SRAP markers for the assessment of genetic diversity in Tunisian germplasm of *P. vera*. Such information will be useful to define conservation strategies and improvement programs for this species.

Key words: *Pistacia vera*; Genetic diversity; SRAP markers; Structure analysis

INTRODUCTION

Pistachio (*Pistacia vera* L.) is a diploid (2n = 30), dioecious, wind-pollinated tree (Zohary, 1952). It has the ability to adapt to arid conditions, representing a typical characteristic of species favored for cultivation in arid and semi-arid areas of Tunisia. Accordingly, pistachios contribute to the development of the economy in these regions and in the country. Although pistachio breeding in Tunisia has occurred for many years, productivity remains very low. Lack of varietal and rootstock diversity is among the factors contributing to low productivity (Chaabouni, 2013). Because of the dioecious nature of *P. vera*, selecting the most suitable male tree as a pollenizer can improve production and fruit quality in this species. A traditional crop method is practised, which leads to short-term success, thus it generates two deficiencies: i) it does not meet the increasingly demanding productivity requirements, ii) slow restoration of crop management. These drawbacks are attributed to a lack of necessary information and a lack of studies relating to genetic resources, an area of research that is rarely undertaken, especially in the case of the pistachio species. Molecular tools are now increasingly used to study diversity and for sex identification (Yakubov et al., 2005; Sun et al., 2014). Molecular markers based on DNA are stable, detectable in all tissues, and are independent of environmental or seasonal conditions. One important application of molecular markers is the assessment of genetic variation in the context of conservation (Karp et al., 1997). According to Robarts and Wolfe (2014), the application of sequence-related amplified polymorphism (SRAP) markers has increased since their introduction (Li and Quiros, 2001), especially in recent years, and the usefulness of this technique is justified in several studies for various purposes. Most SRAP investigations have focused on agronomic applications, and have demonstrated the capacity of this technique to uncover discrete levels of genetic diversity within advanced hybrid lineages. SRAP markers have proven their usefulness for testing various hypotheses of hybridization at different taxonomic levels (Robarts and Wolfe, 2014). Moreover, this technique is used to address hypotheses involving lower-order systematics, and has similarly demonstrated its value in the detection of intra- and interspecific diversity.

To our knowledge, this is the first report on the use of SRAP markers to examine genetic diversity in Tunisian pistachio germplasm, although this technique has previously been applied to the genus *Pistacia* (Talebi et al., 2012). Although, the genetic diversity of Tunisian pistachios has been studied using other molecular markers, no attempts have been made to use SRAP markers, which target open-reading frames as functional regions of the pistachio genome. The results provide relevant information for identifying units of conservation and highlight the genetic processes that take place in the studied populations, including patterns of genetic flux and genetic drift. Thus, the objectives of this study were, i) to explain molecular polymorphism, establish genetic relationships among 43 Tunisian pistachio using SRAP markers, and ii) to elucidate the genetic structure involving male specimens in the sample studied.
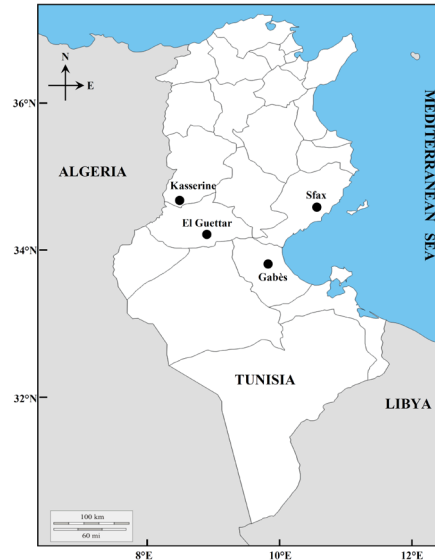
MATERIAL AND METHODS

Plant material

The plant material consisted of young leaves from adult trees. We used a set of 43 accessions (Table 1 and Figure 1), consisting of 8 male and 35 female cultivars. The male trees were sampled in three regions as follows: 'Pollenizer-1' to 'Pollenizer-4' from Gabès, 'Pollenizer-5' to 'Pollenizer-7' from Sfax, and 'Pollenizer-8' from El Guettar.

Cultivars were divided into four groups according to their geographic origin: El Guettar, Gabès, Kasserine, and Sfax. Male trees were gathered into one group named 'Pollenizers'.

Table 1. Localities of the female *Pistacia vera* trees used in this study.

Locality (code)	Characteristics		Accession name
El Guettar (GU)	Sample size	9	El Guettar-1
	Crop germplasm	Oases	El Guettar-2
	Latitude (North)	34°20'	El Guettar-3
	Longitude (East)	08°54'	El Guettar-4
	Altitude (m)	220-245	El Guettar-5
			El Guettar-6
			El Guettar-7
			El Guettar-8
			El Guettar-9
Gabès (GB)	Sample size	8	Mumtez
	Crop germplasm	Collection orchard	Mateur-C
	Latitude (North)	33°53'	Liby Blanc
	Longitude (East)	09°53'	Kermizi
	Altitude (m)	69	Lovy
			Red Aleppo
			Achoury
			Liby Rouge
Kasserine (KS)			
Fériana orchard	Sample size	4	Mateur F-3
	Crop germplasm	Orchard	Mateur F-4
	Latitude (North)	35°01'	Mateur F-5
	Longitude (East)	08°25'	Mateur F-6
	Altitude (m)	963	
Magel Bel Abbès (Orchard 1)	Sample size	4	Mateur MB-1
	Crop germplasm	Orchard	Mateur MB-2
	Latitude (North)	34°44'	Mateur MB-3
	Longitude (East)	08°32'	Mateur MB-4
	Altitude (m)	634	
Magel Bel Abbès (Orchard 2)	Sample size	6	Mateur MB-5
	Crop germplasm	Orchard	Mateur MB-6
	Latitude (North)	34°46'	Mateur MB-7
	Longitude (East)	08°44'	Mateur MB-8
	Altitude (m)	549	Mateur MB-9
			Mateur MB-10
Sfax (SX)	Sample size	4	Thyna
	Crop germplasm	Collection orchard	Safeed
	Latitude (North)	34°56'	Ohadi-1
	Longitude (East)	10°36'	Ohadi-2
	Altitude (m)	118	

Figure 1. Map of Tunisia showing the localities from which genotypes of pistachio were sampled.

DNA extraction

Young leaves were sampled, lyophilized, frozen in liquid nitrogen, and ground into powder. Genomic DNA was extracted and purified using the DNAeasy Plant Mini Kit (Qiagen, USA). The quality and quantity of the DNA were evaluated by analysis on 0.8% agarose gel and diluted appropriately to a working concentration of 30 ng/μL for PCR assays.

DNA amplification

Seven SRAP primer-pair combinations were selected from three forward and four reverse primers (Table 2) (Vandemark et al., 2006; Saha et al., 2010). These primer pairs were designated F8-R9, F8-R15, F9-R8, F9-R9, F13-R7, F13-R9, and F13-R15.

Amplifications were performed using a thermal gradient cycler TProfessional TRIO Thermocycler (Biometra, Germany) in a total volume of 25 μL containing 30 ng genomic DNA, 10X PCR buffer, 0.25 mM each primer (reverse and forward), 3 mM $MgCl_2$, 1 mM dNTPs, 1.5 unit Taq DNA Polymerase (Qbiogene, France), and double-distilled water. The thermal cycling profile for all reactions consisted of a single cycle of 95°C for 5 min followed by 5 cycles of 94°C for 1 min, 40°C for 1 min, 72°C for 1 min, 35 cycles of 94°C for 1 min, 47°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 10 min. PCR products were resolved on 2% agarose gel stained with 0.5 pg/mL ethidium bromide and were electrophoresed in 0.5X TBE buffer (pH 8.0) run at 100 V for 2 h.

The bands were visualized under UV light by means of a Gel-Doc 2000 image analysis system (Bio-Rad, USA). Amplified bands were scored for the presence (1) or absence (0) of bands of the same size for each primer combination to generate the 0/1-matrix.

The core sequences 'CCGG' in the forward primers and 'AATT' in the reverse primers are in bold.

Table 2. Forward and reverse sequence-related amplified polymorphism primers used in this study.	
Forward primer	Forward primer sequence
F8	5'- GTA GCA CAA G**CC GG**A AT -3'
F9	5'- GTA GCA CAA G**CC GG**A CC -3'
F13	5'- CGA ATC TTA G**CC GG**C AC -3'
Reverse primer	Reverse primer sequence
R7	5'- GAC ACC GTA CGA **ATT** TGC -3'
R8	5'- GAC ACC GTA CGA **ATT** GAC -3'
R9	5'- GAC ACC GTA CG**A ATT** TGA -3'
R15	5'- CGC ACG TCC GT**A ATT** CCA -3'

Data analysis

The ability of the primers to distinguish between genotypes was assessed by calculating their resolving power (Rp) (Prevost and Wilkinson, 1999) for SRAP primer combinations based on the formula:

$$Rp = \sum I_b \qquad \text{Equation 1}$$

where band informativeness is $I_b = 1 - (2x|0.5 - p|)$ and p is the proportion of genotypes containing band I. In addition, the polymorphism information content (PIC) values for all selected markers amplified by a particular primer pair were calculated to characterize the reliability of each primer combination to detect polymorphic loci among and within the populations. The PIC value was calculated by use of the formula:

$$PIC = 1 - \sum Pi^2 \qquad \text{Equation 2}$$

where Pi is the frequency of the ith allele (Smith et al., 1997). In addition, the percentages of polymorphic bands (PPB) of each primer combination were calculated.

Assuming Hardy-Weinberg equilibrium, the resulting binary data matrix of the SRAP phenotypes was analyzed using the PopGene software (Yeh et al., 1999) to compute the indices of population-level genetic diversity as follows: PPB, Nei's (1973) gene diversity (h), and Shannon's information index (I). The coefficient of gene differentiation (G_{ST}) was measured using Nei's (1977) gene diversity statistics. Gene flow estimates (N_m) were calculated as described by McDermott and McDonald (1993):

$$N_m = 0.5x(1 - G_{ST})/G_{ST} \qquad \text{Equation 3}$$

Analysis of molecular variance (AMOVA) was performed to estimate variance components for SRAP data, partitioning the variation within and among populations, by use of GenAlEx 6.5 program (Peakall and Smouse, 2012). The following groupings were made: group 1: El Guettar (GU), group 2: Gabès (GB), group 3: Kasserine (KS), group 4: Sfax (SX), and group 5: male tree samples. The significance of P values was tested non-parametrically, after 1000 random permutations. Based on the 0/1 matrix, a cluster analysis was conducted

based on Jaccard's dissimilarity coefficients. The tree topology was constructed based on the neighbor-joining (NJ) method and bootstrapping analysis with 1000 re-samplings in order to test the reliability of the clusters was performed using the software program DARwin version 5.0.158 (Perrier and Jacquemoud-Collet, 2006). Genetic relationships among individuals were assessed by a principal component analysis (PCA) performed with PAST version 2.17c (Hammer et al., 2001).

Population structure was inferred using Bayesian clustering methods implemented in STRUCTURE version 2.3.4 (Pritchard et al., 2000). In STRUCTURE, we assumed an admixture model with prior information of sampling location. Ten replicates were performed for each level of K (K = 1 to K = 9) with 50,000 iterations and 70,000 burn-in steps. To detect the number of genetically homogeneous groups (K) that best fits the data; we used the STRUCTURE Harvester website (Earl and vonHoldt, 2012), which implements the Evanno method (Evanno et al., 2005). To obtain optimal alignment of the independent runs, the CLUMPP version 1.1.2 software program (Jakobsson and Rosenberg, 2007) was used with greedy algorithm, 10,000 random input orders and 10,000 repeats, to calculate the average pairwise similarity (H') of runs. The output was used as input for the cluster Distruct version 1.1 visualization program (Rosenberg, 2004). To identify and justify the signatures of natural selection on individual genetic markers, a Bayesian genome scan approach generalized from the method described by Beaumont and Balding (2004) was implemented within the program BayeScan version 2.1 (Foll and Gaggiotti, 2008), which handles dominant marker data and permits direct estimation of the posterior probability that a given locus is under selection (Krauss et al., 2013). The rationale was to discriminate between the effects on population genetic variation portioning (F_{ST} values) that are specific to each population and to each locus. Here, the F_{ST} value for each locus is treated as an independent data point and is evaluated against a baseline distribution derived from the entire data set. Loci exhibiting extreme F_{ST} values relative to the baseline ('outlier loci') are then interpreted as being directly or indirectly influenced by divergent selection. This method uses a hierarchical bayesian genome scan approach to estimate the posterior probabilities of two alternative models, one including the effects of selection, and one excluding it (Krauss et al., 2013). In addition, we used the false discovery rate (FDR) to control for multiple testing. The FDR is defined as the expected proportion of false positives among outlier markers. In this context, BayeScan defines a q-value, which is the FDR analogue of the P value. The q-value of the given locus is the minimum FDR at which this locus may become significant. A locus-specific statistic, alpha, which is decomposed from the F_{ST} values that reflect the locus-wise difference of allele frequency, was used to infer the selection (Foll and Gaggiotti, 2008). Departure from neutrality is assumed for the loci when alpha differs significantly from zero, whereas positive and negative alpha values indicate balancing or diversifying selection, respectively. We performed pairwise population comparisons for selection signatures using the genome scan method in BayeScan. The Bayesian likelihood method in BayeScan, which compares a model of selection versus neutrality via a reversible-jump MCMC algorithm, was used to detect outliers and to estimate the posterior probability of each model (Foll and Gaggiotti, 2008). Posterior odds are the ratio of posterior probabilities, and indicate how more likely the model with selection compares to the neutral model. Log10 values of the posterior odds (PO) >0.5 and 2.0 were taken as 'substantial' and 'decisive' evidence for selection, respectively (Jeffreys, 1961). An advantage of the posterior probability approach is that it

directly allows for control of the FDR. Here, the FDR was set at 0.05 and 0.01, adjusting the Log10 (PO) significance thresholds corresponding to the 0.5 and 2.0 values considered before correction. Graphical illustrations of the results were obtained using R-PACKAGE version 3.0.1 (R Development Core Team, 2013).

RESULTS

SRAP polymorphism

A total of 12 SRAP primer pairs using three forward and four reverse primers were tested for their ability to amplify nine randomly selected genotypes by PCR. Only seven primer pairs showing reproducible patterns were selected and used to analyze all of the 43 pistachio genotypes. The features of the primers across the 43 tested genotypes are summarized in Table 3. From the seven primer pairs, 82 reliable bands ranging from 150 to 4072 bp were scored, of which 78 (95.12%) were polymorphic. The number of polymorphic fragments amplified using different SRAP primer pairs was 8 (F13-R15) to 14 (F13-R7). The highest level of polymorphism (100%) was shown by three primer pairs: F8-R15, F9-R8, and F13-R9 (Table 3). The value of the primer pairs ranged from 2.930 (F13-R9) to 6.279 (F8-R15). The PIC, measured as the percentage of polymorphic fragments for all primer pairs, was high, and varied within a relatively narrow range of 0.763 (for primer pair F13-R9) to 0.904 (for primer pair F13-R7), with an average of 0.850.

Table 3. Polymorphism and incidence of seven SRAP primer pairs used to amplify 43 genomic DNA templates of pistachio.

Primer pairs	Band size (bp)	Total bands	NPB	PPB(%)	Rp	PIC
F8-R9	400-3054	10	9	90	4.349	0.833
F8-R15	250-4072	13	13	100	6.279	0.903
F9-R8	350-4072	12	12	100	5.674	0.874
F9-R9	250-1800	14	13	92.86	3.628	0.836
F13-R7	150-3054	15	14	93.33	5.744	0.904
F13-R9	150-3054	9	9	100	2.930	0.763
F13-R15	150-4072	9	8	88.89	3.674	0.838
Total		82	78	-	32.279	-
Mean		11.714	11.143	95.12	4.611	0.850

NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; Rp: resolving power; PIC: polymorphic information content.

Cluster and principal component analysis

The dendrogram of genetic relationships among genotypes revealed by the NJ method and the two-dimensional PCA (component-1 and component-2) plots are presented in Figure 2. For the NJ method (Figure 2A), four main clusters are identified and labeled as A, B, C, and D. In cluster A, a grouping of eight genotypes from GU ('El Guettar-1' to 'El Guettar-8') was observed in a sub-cluster (gray branches). Cluster B includes all pistachio genotypes from KS except 'Mateur MB-4', which is found in cluster C. All male genotypes ('Pollenizer-1' to 'Pollenizer-8') were grouped in a sub-cluster (blue branches) of cluster C. Finally, cluster D consisted of four genotypes: 'Ohadi-1', 'Ohadi-2', 'Liby Rouge', and 'Thyna'. The two sub-clusters: 'Pollenizers' and 'El Guettar' are visible in the PCA plot (Figure 2B). This shows that both methods generate similar clusterings.
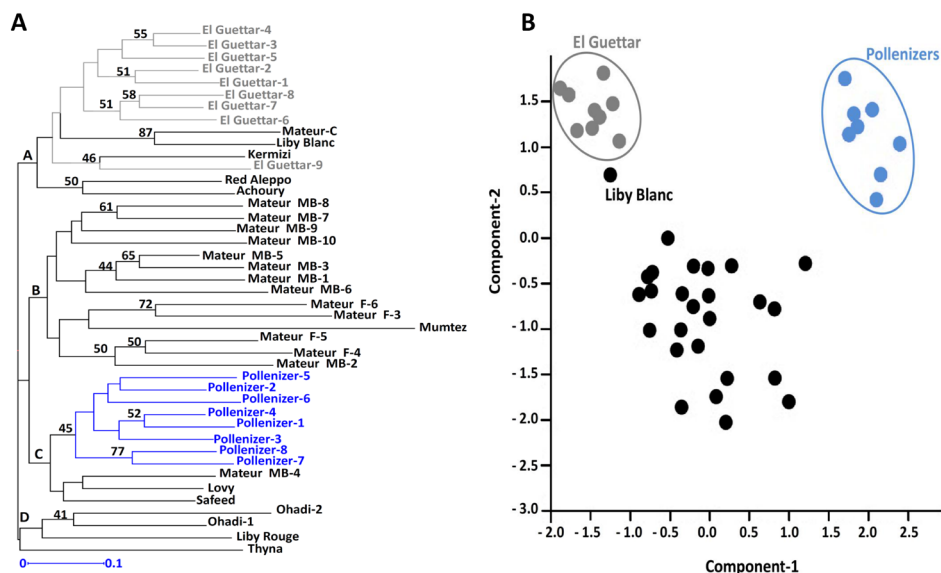
Figure 2. Neighbor-joining tree and principal component analysis (PCA) derived from sequence-related amplified polymorphism (SRAP) markers. **A.** Neighbor-joining tree with branch lengths reflecting phylogenetic distances based on Jaccard's genetic distances. Numbers associated with branches indicate the percent support based on 1000 bootstrap replicates. Branches with less than 40% support were collapsed. **B.** Biplot of PCA. Two groups are circled: El Guettar and Pollenizers groups.

Genetic diversity and genetic structure investigated by SRAP markers

The 35 female pistachio genotypes were grouped on the basis of their planting sites into four groups: 'El Guettar', 'Gabès', 'Kasserine', 'Sfax', with the addition of a 'Pollenizers' group (eight genotypes) (Table 4). Overall, among the five pistachio groups, the highest values for h, I, and PPB were obtained for the KS group (h = 0.255; I = 0.382; PPB = 71.95%), and the lowest were observed for the GU group (h = 0.129; I = 0.197; PPB = 39.02%).

The G_{ST} value was 0.307 and AMOVA revealed that 73.88% of the total genetic variation occurred within groups and 26.12% occurred among groups (Table 5). The low level of genetic differentiation between the five groups is justified by the high gene flow (N_m = 1.127).

Table 4. Genetic diversity and differentiation parameters for SRAP markers in five groups of *Pistacia vera*.

Groups (code)	Ss	NPB	PPB (%)	h ± SD	I ± SD
El Guettar (GU)	9	32	39.02	0.129 ± 0.181	0.197 ± 0.264
Gabès (GB)	8	58	70.73	0.243 ± 0.183	0.369 ± 0.261
Kasserine (KS)	14	59	71.95	0.255 ± 0.188	0.382 ± 0.266
Sfax (SX)	4	33	40.24	0.169 ± 0.211	0.245 ± 0.304
Pollenizers (POL)	8	41	50	0.181 ± 0.194	0.272 ± 0.284
Population level	-	44.6	54.39	0.195 ± 0.191	0.293 ± 0.276
Species level	43	78	95.12	0.289 ± 0.164	0.443 ± 0.212

Ss: sample size; NPB: number of polymorphic bands; PPB (%): percentage of polymorphic loci; h: Nei's genetic diversity; I: Shannon's information index; SD: standard deviation.

**Table 5.** Results of analysis of molecular variance (AMOVA) for five groups of *P. vera*.

Source of variation	d.f.	Sum of squares	Mean squares	Percentage of variation	P value
Among group	4	149.498	37.374	26.12%	<0.001
Within group	38	360.921	9.498	73.88%	<0.001
Total	42	510.419		$\Phi st = 0.261$	<0.001

d.f.: degrees of freedom; P value: significance after 1000 random permutations; Fst: differences among populations.

Population structure

The genetic structure of Tunisian pistachio was examined according to the model with two to seven clusters (K = 2 and 7, respectively). The ad-hoc quantity based on the second order rate of change of the likelihood function (ΔK) (Evanno et al., 2005) revealed a first level of clustering at K = 3 for the investigated Tunisian pistachios (ΔK = 25.116) and a sub-clustering at K = 2 (ΔK = 12.827). Based on the permuted average Q-matrix generated by Clumpp for the 10 Structure runs, the highest H' was observed for K = 3 (H' = 0.989) and K = 2 (H' = 0.983), indicating the stability of the results for these two models (Figure 3). According to the model at K = 3, pistachio genotypes were assigned to three genetically different clusters identified by STRUCTURE analysis: cluster 1 (gray) included accessions from GU, cluster 2 (blue) included the 'Pollenizers' group, and cluster 3 (yellow) included accessions from GB, KS, and SX. The corresponding genotypes of the GU 'Pollenizers' groups were assigned with over 80% probability to cluster 1 (gray) and cluster 2 (blue), respectively. At K = 2, pistachio genotypes from GU in cluster 1 (gray) were separated from those in cluster 2 (blue) with a probability greater than 80%. At K = 4 to K = 7, we found that cluster 1 (gray) and cluster 2 (blue) remained distinguished from the other groups (with an assigned probability greater than 80%).
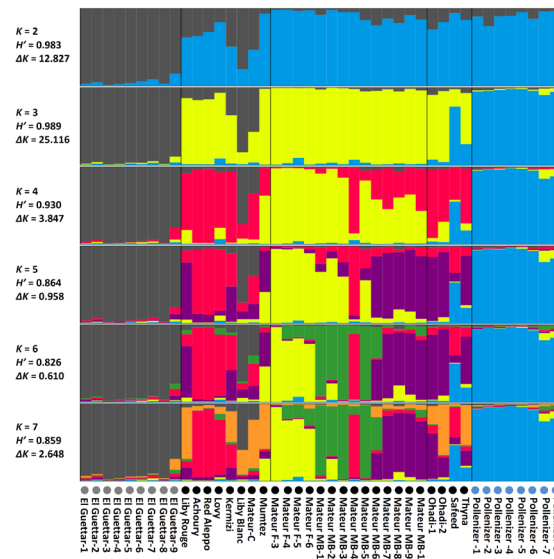

Figure 3. Genetic clustering obtained from the STRUCTURE analysis (N = 43). Each individual is represented by a single vertical column, divided into K colors. The colored segment shows the individual's estimated proportion of membership to that genetic cluster. Symbols (color-filled circle) correspond to those represented in the PCA plots (Figure 2B).

The three groups (GB, KS, and SX), show a high level of admixture in which it is difficult to identify sub-clusters for assignment above 70%. However, the KS group was divided into three sub-clusters (yellow, green, and purple) at K = 6 and K = 7, although, the genetic structure of these models is not convincing, since the ΔK and the highest H' are smaller than for K = 2 and K = 3. The genotype 'Liby Blanc' is assigned to cluster 1 (gray) with an assignment probability of 60 to 83% in all models. This is consistent with the close positioning of the 'Liby Blanc' genotype with the GU cluster in the PCA plot. Furthermore, the PCA plot showed two distinct groups (El Guettar and Pollenizers) within the pistachio genotypes, which was consistent with the structure-based assignment for the corresponding genotypes. However, the NJ tree revealed sub-clustering for these two groups, which was consistent with the structure-based assignment for these genotypes: a sub-cluster of the El Guettar genotypes in cluster A and a sub-cluster of the Pollenizers group in cluster C.

Outlier detection

We performed pairwise group comparisons to identify F_{ST} outliers using the BayeScan method. To minimize the presence of false positives, we considered a SRAP locus to have greater support as an outlier if it was identified by the BayeScan method.

BayeScan analysis produced low differentiation loci at a threshold between 0.5 and 1.5 (posterior probabilities higher than 0.75 and less than 0.94), corresponding to 11 identified loci (12, 15, 48, 50, 54, 59, 61, 68, 69, 73, and 74) of 82 investigated loci (13.41%) for 10 pairwise population comparisons. The 11 identified loci had significantly higher F_{ST} values than were found across the other loci [$F_{ST} > 0.23$; alpha > 0.75; log10(PO) > 0.5], when the FDR was set at 0.05 (Figure 4).
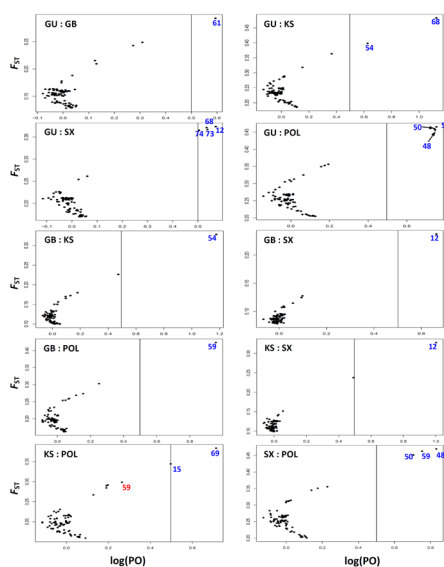
Figure 4. Genome scan to identify selective outlier loci with the BAYESCAN approach. Plots represent pairwise comparisons for the five *Pistacia vera* groups investigated. Each point corresponds to a SRAP locus. F_{ST} is plotted against the log10 of the posterior odds (PO), which provides evidence on whether the locus is subjected to selection. The vertical line shows the decisive threshold value [log10 (PO) = 0.5] used for identifying outlier loci.

According to the BayeScan approach, at least one outlier was identified for each peer comparison of pistachio groups. The number of identified outliers did not exceed four (case GU:SX) for peer comparisons. Nevertheless, the lower detection of outliers may be due to the low sample size. BayeScan demonstrated two loci of interest, which were commonly detected at loci 59 and 12. Locus 59 (F13-R9) was commonly detected during the comparison involving the Pollenizers group, except in pair KS:POL, in which locus 59 was determined at $F_{ST} = 0.29$ and $\log 10(PO) = 0.26 < 0.5$ (Figure 4). Similarly, locus 12 was commonly detected in the comparison involving the SX group, except in pair SX:POL, in which locus 12 was determined at $F_{ST} = 0.31$ and $\log 10(PO) = 0.017 < 0.5$.

DISCUSSION

Assessment of genetic variation within germplasm collections of pistachio is crucial for the conservation and efficient use of these resources in breeding programs, and could be significantly improved by using molecular fingerprinting tools. This study described the use of the SRAP technique to characterize and survey molecular polymorphism of Tunisian pistachio cultivars for the first time. SRAP molecular technology is very useful in plant breeding programs and has been commonly recognized as a useful technique for estimating plant genetic diversity (Li et al., 2013). Therefore, the SRAP technique was tested to characterize species and accessions of three species of *Pistacia*, which generated sufficient polymorphisms for assessment of genetic diversity among these species and accessions (Talebi et al., 2012). SRAP is a targeted fingerprinting marker (TFMs). A fundamental characteristic of this type of markers is the anchor elements (e.g., gene promoters), which are added to various parts of the primers to amplify sites flanking the target region. Another favorable feature of TFMs is the exploitation of common genomic features, which makes them easily transferable between organisms (Poczai et al., 2013).

In the present study, the average PPB was 95.12%, which was higher than that obtained previously for the *Pistacia* genus (Talebi et al., 2012). The high rate of polymorphism reflects the extent of genetic divergence revealed by SRAP combinations and the effectiveness of SRAP markers at detecting genetic diversity in pistachio and various plant species (Alghamdi et al., 2012). The PIC value, which depends on the number of detectable alleles and their distribution, indicates a marker's utility for detecting polymorphism within a population (Botstein et al., 1980). The average PIC value obtained in this study was 0.850, indicating that all of the SRAP markers showed high polymorphism and could be used to obtain basic information on the genetics of pistachios for breeding research. Assuming Hardy-Weinberg equilibrium, a strong genetic differentiation coefficient was calculated among groups (G_{ST} = 0.307), and AMOVA revealed that 73.88% of the total genetic diversity occurred within groups, whereas the remaining 26.12% of the total variation occurred among groups. Variation among the *P. vera* groups confirms there is a large genetic diversity base despite the level of N_m between the groups ($N_m = 1.127 > 1$), which indicates the presence of genetic differentiation among populations. Our results are consistent with those previously reported for pistachio, and a high level of variation is found within the groups and populations of *P. vera* or/and species of the *Pistacia* genus (Basha et al., 2007; Shanjani et al., 2009; Pazouki et al., 2010; Talebi et al., 2012). Cross-pollination reproduction due to the dioecy of species significantly influences the high variation within groups and explains the level of polymorphism detected. Despite the differences in crop management within the four areas [oases (GU), orchards (KS),

experimental sites (SX and GB)], more variability is observed within groups than between groups. This could be explained by the exchange of plant material (pollen, seeds, rootstock) together with the recurrent reciprocal transfer of specimens, especially the male trees, among the four regions studied.

The oasis of GU is very old, and is a rich gene pool for *P. vera* L. species. Furthermore, the oasis environment and the specific crop method practiced for pistachio production contribute to form an enriched gene pool. Furthermore, its vitality to the continued expansion of breeding areas has helped to perpetuate the traditional crops. Due to their favorable agronomic characteristics (resistance to pests and tolerance to abiotic stresses), the oases specimens are planted in areas of new crops. We noted some specimens originating from GU planted in orchards of KS, SX, and GB. Furthermore, the transfer of specimens, especially between GU and KS (facilitated by their geographical proximity), and the exchange of plant materials, may have decreased genetic variability between *P. vera* groups. Transfer and random exchanges between small populations inherently influence the effects of genetic drift. Gene flow plays a major role in the spatial organization of genetic diversity and represents high potential migration. The SRAP markers revealed the highest level of genetic diversity in the group from KS. Genotypes from KS in this study originated from three orchards (Magel Bel Abbès: orchard 1, orchard 2, and Fériana orchard). The high level of heterogeneity detected within the KS group could be explained by differences between the plantation sites despite their geographical proximity. In the case of the GB group, the variety has contributed to the high heterogeneity within the group; however, a small number of samples was studied, such as those in the SX group. Despite the low genetic diversity of the El Guettar samples, both the NJ and PCA methods discriminated the genotypes from GU from the other accessions. Conversely, male genotypes were congregated by both methods. Indeed, the male genotypes were not clustered in an outgroup but formed a subgroup of cluster C (NJ method) in association with three female genotypes 'Mateur B-4', 'Lovy', and 'Safeed'.

Structure analysis reinforced the results obtained by the PCA plot and NJ method, showing the identification of three pools: El Guettar, Pollenizers and the rest of pistachio belonging to Gabès, Kasserine and Sfax localities, at K = 2 and K = 3. This structure is consistent with the results of the PCA method. Even in the less informative clusters (K = 4 to K = 7), both groups showed no significant overlap with other clusters, and remained conserved groups. Incidentally, genotype 'Liby Blanc' is very close genetically to the GU group. In the current study, male genotypes formed a distinct subgroup from the female varieties. In addition, BayeScan revealed a low percentage of outliers (13.41%) for 10 pairwise comparison groups. Two loci, 59 and 12, were commonly detected in comparisons involving the Pollenizers and SX groups, respectively.

Regarding the Pollenizers group, it could be hypothesized that locus 59 has a role in sex-linked discrimination, despite the low numbers of samples. Zhou et al. (2011) implemented SRAP markers for sex determination in Buffalograss species. The ability to distinguish genders at an early vegetative stage would be useful, and is a major commercial interest in the case of *P. vera* (Yakubov et al., 2005). Hormaza et al. (1994) identified one specific band from *P. vera* using the random amplified polymorphic DNA (RAPD) technique. In a subsequent study, Yakubov et al. (2005) converted the RAPD primer to a sequence characterized amplified regions (SCAR) marker and used the touchdown-PCR technique to identify sexual genotypes in *P. vera*. Recently, Esfandiyari et al. (2012) confirmed those results in wild *Pistacia* species and Sun et al. (2014) used RAPD and inter-simple sequence repeat patterns to identify a SCAR marker for sex identification of *Pistacia chinensis*. However, in dioecious plants, even

if sex determination is regulated genetically, no heteromorphic sex chromosomes have been found. Therefore, sex-linked genetic markers seem to offer an appropriate technique for use in breeding programs. They are also useful for understanding the genetic map and dioecism of *Pistacia* species (Esfandiyari et al., 2012). Further studies should be conducted to develop sex-linked SCAR markers from SRAP markers, and specifically, locus 59 (F13-R9). The SCAR approach from SRAP markers has already been used by Mutlu et al. (2008) for the identification of markers linked to a gene involved in *Fusarium* wilt resistance in eggplant.

In the present study, the accessions were collected from key regions (e.g., El Guettar Province), although limited sample numbers were used. Therefore, future research on the genetic diversity of *P. vera* should involve more extensive germplasm resources. In conclusion, SRAP markers have proven their usefulness for estimating genetic diversity in *P. vera* and provide appropriate information for population genetic structure of the pistachio. The SRAP technique has advantages in terms of convenience, good reproducibility, and high polymorphism, and can therefore be used more extensively to study the genetic diversity of plants. Furthermore, the molecular relationships delineated in this study should be useful in breeding and conservation programs for *P. vera*.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Alghamdi SS, Al-Faifi SA, Migdadi HM, Khan MA, et al. (2012). Molecular diversity assessment using sequence related amplified polymorphism (SRAP) markers in *Vicia faba* L. *Int. J. Mol. Sci.* 13: 16457-16471. http://dx.doi.org/10.3390/ijms131216457

Basha AI, Padulosi S, Chabane K, Hadj-hassan A, et al. (2007). Genetic diversity of Syrian pistachio (*Pistacia vera* L.) varieties evaluated by AFLP markers. *Genet. Resour. Crop Evol.* 54: 1807-1816. http://dx.doi.org/10.1007/s10722-006-9202-5

Beaumont MA and Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13: 969-980. http://dx.doi.org/10.1111/j.1365-294X.2004.02125.x

Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Chaabouni AC (2013). Capacité germinative et composition minérale du pollen de quelques génotypes mâles de pistachier (*Pistacia vera* L.). *Revue des Régions Arides.* 35: 329-335.

Earl DA and vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361. http://dx.doi.org/10.1007/s12686-011-9548-7

Esfandiyari B, Davarynejad GH, Shahriari F, Kiani M, et al. (2012). Data to the sex determination in *Pistacia* species using molecular markers. *Euphytica* 185: 227-231. http://dx.doi.org/10.1007/s10681-011-0527-6

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. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x

Foll M and Gaggiotti O (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180: 977-993. http://dx.doi.org/10.1534/genetics.108.092221

Hammer Ø, Harper DAT and Ryan PD (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electronica* 4: 9.

Hormaza JI, Dollo L and Polito VS (1994). Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. *Theor. Appl. Genet.* 89: 9-13. http://dx.doi.org/10.1007/BF00226975

Jakobsson M and Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801-1806. http://dx.doi.org/10.1093/bioinformatics/btm233

Jeffreys H (1961). Theory of Probability. In: Oxford Classic Texts in the Physical Sciences. Oxford University Press, Oxford, United Kingdom.

Karp A, Kresovich S, Bhat KV, Ayad WG, et al. (1997). Molecular tools in plant genetic resources conservation: A guide to the technologies. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy.

Krauss SL, Sinclair EA, Bussell JD and Hobbs RJ (2013). An ecological genetic delineation of local seed-source provenance for ecological restoration. *Ecol. Evol.* 3: 2138-2149. http://dx.doi.org/10.1002/ece3.595

Li G and Quiros CF (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica. Theor. Appl. Genet.* 103: 455-461. http://dx.doi.org/10.1007/s001220100570

Li G, McVetty BE and Quiros CF (2013). SRAP Molecular Marker Technology in Plant Science. In: Plant Breeding from Laboratories to Fields (Andersen SV, eds.). Copenhagen University, Denmark.

McDermott JM and McDonald BA (1993). Gene flow in plant pathosystems. *Annu. Rev. Phytopathol.* 31: 353-373. http://dx.doi.org/10.1146/annurev.py.31.090193.002033

Mutlu N, Boyaci FH, Göçmen M and Abak K (2008). Development of SRAP, SRAP-RGA, RAPD and SCAR markers linked with a *Fusarium* wilt resistance gene in eggplant. *Theor. Appl. Genet.* 117: 1303-1312. http://dx.doi.org/10.1007/s00122-008-0864-6

Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70: 3321-3323. http://dx.doi.org/10.1073/pnas.70.12.3321

Nei M (1977). F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* 41: 225-233. http://dx.doi.org/10.1111/j.1469-1809.1977.tb01918.x

Pazouki L, Mardi M, Shanjani PS, Hagidimitriou M, et al. (2010). Genetic diversity and relationships among *Pistacia* species and cultivars. *Conserv. Genet.* 11: 311-318. http://dx.doi.org/10.1007/s10592-009-9812-5

Peakall R and Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics* 28: 2537-2539. http://dx.doi.org/10.1093/bioinformatics/bts460

Perrier X and Jacquemoud-Collet JP (2006). DARwin Software. Available at [http://darwin.cirad.fr/darwin].

Poczai P, Varga I, Laos M, Cseh A, et al. (2013). Advances in plant gene-targeted and functional markers: a review. *Plant Methods* 9: 6. http://dx.doi.org/10.1186/1746-4811-9-6

Prevost A and Wilkinson MJ (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor. Appl. Genet.* 98: 107-112. http://dx.doi.org/10.1007/s001220051046

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

R Development Core Team (2013). R: A Language and Environment for Statistical Computing. Version 3.0.1. R Foundation for Statistical Computing, Vienna, Austria.

Robarts DWH and Wolfe AD (2014). Sequence-related amplified polymorphism (SRAP) markers: A potential resource for studies in plant molecular biology(1.). *Appl. Plant Sci.* 2: 1400017. http://dx.doi.org/10.3732/apps.1400017

Rosenberg NA (2004). DISTRUCT: a program for the graphical display of population. *Mol. Ecol. Notes* 4: 137-138. http://dx.doi.org/10.1046/j.1471-8286.2003.00566.x

Saha GC, Sarker A, Chen W, Vandemark GJ, et al. (2010). Inheritance and Linkage Map Positions of Genes Conferring Resistance to Stemphylium Blight in Lentil. *Crop Sci.* 50: 1831-1839. http://dx.doi.org/10.2135/cropsci2009.12.0709

Sun Q, Yang X and Li R (2014). SCAR marker for sex identification of Pistacia chinensis Bunge (Anacardiaceae). *Genet. Mol. Res.* 13: 1395-1401. http://dx.doi.org/10.4238/2014.February.28.12

Shanjani PS, Mardi M, Pazouki L, Hagidimitriou M, et al. (2009). Analysis of the molecular variation between and within cultivated and wild *Pistacia* species using AFLPs. *Tree Genet. Genomes* 3: 447-458. http://dx.doi.org/10.1007/s11295-008-0198-1

Smith JSC, Chin ECL, Shu H, Smith OS, et al. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. *Theor. Appl. Genet.* 95: 163-173. http://dx.doi.org/10.1007/s001220050544

Talebi M, Kazemi M and Sayed-Tabatabaei BE (2012). Molecular diversity and phylogenetic relationships of *Pistacia vera, Pistacia atlantica* subsp. *mutica* and *Pistacia khinjuk* using SRAP markers. *Biochem. Syst. Ecol.* 44: 179-185. http://dx.doi.org/10.1016/j.bse.2012.05.013

Vandemark GJ, Ariss JJ, Bauchan GA, Larsen RC, et al. (2006). Estimating genetic relationships among historical sources of alfalfa germplasm and selected cultivars with sequence related amplified polymorphisms. *Euphytica* 152: 9-16. http://dx.doi.org/10.1007/s10681-006-9167-7

Yakubov B, Barazani O and Golan-Goldhirsh A (2005). Combination of SCAR primers and touch-down PCR for sex identification in *Pistacia vera* L. *Sci. Hortic. (Amsterdam)* 103: 473-478. http://dx.doi.org/10.1016/j.scienta.2004.06.008

Yeh F, Yang R and Boyle T (1999). POPGENE Version 1.32. Microsoft Window-based Freeware for Population Genetic Analysis. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton.

Zhou Y, Wang X and Zhang X (2011). Development and application of a SRAP marker for the identification of sex in *Buchloe dactyloides. Euphytica* 181: 261-266. http://dx.doi.org/10.1007/s10681-011-0419-9

Zohary M (1952). A monographical study of genus *Pistacia. Palestine J. Bot. Jerusalem Ser.* 5: 187-228.