Genet Resour Crop Evol (2019) 66:1189–1203
https://doi.org/10.1007/s10722-019-00784-8

RESEARCH ARTICLE

Assessment of genetic diversity among local pea (*Pisum sativum* L.) accessions cultivated in the arid regions of Southern Tunisia using agro-morphological and SSR molecular markers

Amina Mohamed · Santiago García-Martínez · Mohamed Loumerem · Pedro Carbonell · Juan José Ruiz · Mohsen Boubaker

Received: 19 January 2019 / Accepted: 3 May 2019 / Published online: 8 May 2019

Abstract The pea (*Pisum sativum* L.) is arguably an important winter pulse crop and a fundamental source of proteins for human and animal nutrition. In Tunisia, pea cultivation is characterized by its instability, especially in the arid region of Southern Tunisia. We carried out an investigation of the agro-morphological and genetic diversity of 12 local pea accessions cultivated in the arid region. We analyzed 21 qualitative and 26 quantitative traits following the UPOV descriptors. Furthermore, 8 SSRs were employed to examine genetic polymorphism, differentiation, and population structure. Molecular and agro-morphological data distinguished all the accessions under investigation. A considerable phenotypic diversity among accessions was observed for many characters, including some related to agronomical performance. At the molecular level, all the SSRs were polymorphic, with an average of 0.44 PIC value per locus. Our work indicated the presence of a wide-ranging variation among the local pea accessions evaluated. The overall results indicated that the agro-morphological traits and SSR markers were reliable and effective for assessing the genetic diversity of phenotypic pea accessions.

Keywords Genetic diversity · Agro-morphological traits · SSR · Population structure · Pea

Introduction

Nowadays, pea (*Pisum sativum* L. subsp. *sativum*) is one of the world's most important winter pulse crops (Mikic et al. 2011). As a member of the large legume family *Fabaceae*, it has both interesting biological features and attractive ecological services. Due to the symbiosis it establishes with atmospheric nitrogen fixing *Rhizobacteria*, this winter crops do not require nitrogen fertilizer inputs and provide nitrogen to the following crop. Furthermore, pea seeds, as well as leaves, are a good source of plant proteins for human and animal nutrition (Ubayasena et al. 2010).

In Tunisia, the area occupied by this crop is about 20,000 ha and its production is characterized by its instability (DGEDA 2011). Abiotic stresses and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10722-019-00784-8) contains supplementary material, which is available to authorized users.

A. Mohamed · M. Loumerem
Dryland and Oases Cropping Laboratory, Arid Land Institute, Street El Jorf, 4119 Medenine, Tunisia

A. Mohamed · M. Boubaker
Higher Agronomic Institute, Chott Mariem, IRESA-University of Sousse, B.P. 47, 4042 Chott Mariem, Sousse, Tunisia

S. García-Martínez (✉) · P. Carbonell · J. J. Ruiz
Department of Applied Biology, Miguel Hernandez University, Carretera de Beniel, km 3.2, 03312 Orihuela, Alicante, Spain
e-mail: sgarcia@umh.es

inefficient production techniques are the main reasons, especially in the southern region characterized by severe aridity (Kharrat et al. 2006). Therefore, a large collection of local pea populations from arid region of Southern Tunisia has been surveyed and evaluated, considering two major compelling factors, the erosion in the genetic diversity in the local breeding program and the need for better adapted genotypes in the context of an increased impact of climate change and environmental variability (De Giorgeo and Polignano 2001).

The study of genetic diversity of this winter crop is of interest for the conservation of genetic resources and practical applications in breeding programs. Several studies have been carried out to estimate the genetic variation within pea germplasm through different approaches such as morphological descriptors, biochemical and molecular markers (Hoey et al. 1996; Baranger et al. 2004; Kwon et al. 2012). In connection with this, the evaluation of morphological traits is a traditional and important method for the description and the determination of relationship among pea landraces (Smykal et al. 2008a). On the other hand, the expressions of most of these morphological traits are generally influenced by environmental factors and cultivation practices.

Molecular markers that reveal polymorphism at the DNA level have been considered as a powerful tool for the estimation of plant genetic diversity characterization and also to discriminate different morphological individuals from different sources (Keneni et al. 2005). Various molecular markers have been successfully used to characterize pea germplasm, such as ISSR (Lázaro and Aguinagalde 2006), SNP (Duarte et al. 2014), SRAP (Esposito et al. 2007), and SSR (Baranger et al. 2004). Among them, microsatellites (SSRs) have gained popularity because of cost effectiveness, speed, reproducibility and polymorphism (Snowdon and Friedt 2004).

The main objective of this study was to characterize and evaluate the genetic diversity of 12 Tunisian pea accessions cultivated in the arid region (Southern Tunisia) using agro-morphological traits and SSR markers, in order to develop strategies to preserve the endangered genetic resources of this specie.

Materials and methods

Plant material and crop conditions

The study took place at the experimental field of the Arid Land Institute of Medenine in Tunisia (33°29′57.80″N, 10°38′32.96″E, Altitude 16 m), during 2015–2016; 2016–2017 and 2017–2018 period. The investigation was carried out on eleven accessions of local peas (*Pisum sativum* L.), which are a peasant seeds that come directly from various sources of arid region farmers in Tunisia, characterized and evaluated for selection purposes and maintained by the Laboratory of Dryland and Oases Cropping (Table 1), as well as two commercial pea varieties 'Baddar' (Tunisia) and 'Lincoln' (France) were included as references. For the latter, we added only in the molecular analysis. For the seeds used in the 3 year evaluation, they are coming from the dynamic management and selection work based on the yield parameters. Undamaged clean seeds of each accession, selected to a reasonably uniform size by hand sorting, were planted on the seedbeds. They were arranged in random complete blocks design with three replications; with ten plants per replicate per population on each block. The seeds were sown in rows spaced 40 cm. Irrigation was distributed twice weekly with equal quantities during the trial period and supplied with a mineral fertilizer (NPK 20–20–20) as base dressing. During the growing season, weeds were hand-controlled, while pests were handled through chemical management.

Agro-morphological characterization

Agro-morphological characteristics of 12 pea accessions in the 3 years were determined on 10 randomly selected plants per population. A total of 48 morphological traits, including quantitative, qualitative and phenological characters were evaluated in plant, stem, leaf and leaflet, wing, stipule, flower and pod characters (supplementary Tables 1 and 2), according to the UPOV guidelines TG 7/9 *Pisum sativum* L. (UPOV 1990). Six traits showed to be monomorphic.

Agro-morphological data analysis

Mean, standard deviations and range values were calculated for quantitative data in the 3 years. Inter-accession variability was estimated by calculating the

Table 1 Locations and characteristics of seeds of pea accessions collected from the oases of southern Tunisia

Accession code	Accession name	Site of collection	Province	Latitude (N)	Longitude (E)	Altitude (m)
P1	Baddar	–	Baddar Agricole Tunisia (khatima)	36°49′	10°09′	23
				36°49′	10°09′	23
P2	P1001	Ksar Jawama (Beni khedach)	Medenine	33°15′	10°11′	506
				33°15′	10°11′	506
P3	P3001	Ksar hallouf (Beni khedach)	Medenine	33°29′	10°15′	386
P4	P5001			33°29′	10°15′	386
P5	P6001					
P6	P6003					
P 7	P7001	Mareth	Gabes	33°37′	10°16′	48
P8	P7002			33°37′	10°16′	48
P9	P9001					
P10	P9002					
P11	P9003					
P12	P1002					
P13	Lincoln	–	Semillas Fito (France)			

coefficient of variation (CV) of each quantitative trait in each accession. For qualitative traits, the frequency of each category was considered. Analysis of variance (ANOVA) was conducted along with a Duncan test of means comparison ($P < 0.05$). For each parameter, a Pearson correlation analysis was used to estimate the relationship between the studied variables. In order to classify the accessions and to identify the most important traits, multivariate analysis methods were used: principal component analysis (PCA) was performed for quantitative traits, multiple correspondence analysis (MCA) for qualitative data and multiple factor analysis (MFA), useful for the combination of quantitative and qualitative data. These analyses were conducted using the 'FactoMineR' v. 1.34 (Le et al. 2008) and 'factoextra' v. 1.0.5 (Kassambara and Mundt 2017) packages in R. In addition, accessions were clustered by the unweighted pair-group method with arithmetic averages (UPGMA) (Sneath and Sokal 1973). Finally, a cophenetic value matrix of the UPGMA clustering was also used to test for the goodness-of-fit of the clustering to the resemblance matrix. These analyses were conducted using the 'dendextend' v. 1.8.0 package in R (Galili 2015).

Molecular characterization

DNA extraction

DNA was extracted from young leaves of 1-week old seedlings following the CTAB method with slight modifications (Doyle and Doyle 1990). DNA was isolated from five individuals for each accession. The DNA thus obtained was quantified and diluted to 10 ng μl^{-1}. Concentration and purity of DNA samples were assessed using NanoDrop 2000 Spectrophotometer, while the integrity of the DNA was analyzed using a 2% agarose gel and stored at − 20 °C until use.

PCR optimization and microsatellite selection

To investigate the genetic diversity among the selected accessions, 22 SSR markers were chosen based on pea genetic map obtained from Loridon et al. (2005). Primer sequences and annealing temperature (Tm) are reported in Table 2. The PCR master mix was prepared as described by Smykal et al. (2008a). In brief, PCRs were carried out in a total reaction volume of 20 µl, containing 10 ng DNA, 25xPCR buffer, 10 µM of both forward and reverse primers, 10 mM dNTPs, and 5U NZYTaq-DNA polymerase (nZytech). PCR amplification was conducted on the Eppendorf Master Cycler Gradient (Germany) under the

Table 2 Primer sequence and annealing temperature for the markers selected to assess the genetic variability among pea accessions, taken from Loridon et al. (2005)

Primer name	Sequence	Tm (°C)
AA205	F:tacgcaatcatagagtttggaa	51
	R:aatcaagtcaatgaaacaagca	
AD270	F:ctcatctgatgcgttggattag	51
	R:aggttggatttgttgtttgttg	
D21	F:tcaaaatttccttattctcctc	51
	R:gtcaaaattagccaaattcctc	
A5	F:gtaaagcataaggggattctcat	51
	R:cagcttttaactcatctgacaca	
AA430902	F:ctggaattcttgcggtttaac	54
	R:cgttttggttacgatcgagcat	
A9	F:gtgcagaagcatttgttcagat	56
	R:cccacatatatttggttggtca	
AB53	F:cgtcgttgttgccggtag	51
	R:aaacacgtcatctcgacctgc	
AB25	F:ttttcactcaaaacactcggct	51
	R:gatgccattgctgaaggagatt	
PSGAPA1	F:gacattgttgccaataactgg	51
	R:ggttctgttctcaatacaag	
AF004843	F:ccatttctggttatgaaaccg	54
	R:ctgttcctcattttcagtggg	
AA72	F:taatcgattgggcataggtgtc	61
	R:ttgtcttgttgcgtctgagtga	
AA153	F:tttgatagtccgacttttccat	51
	R:gtgacaaaagaattcaaaacgc	
AA179	F:tggaccagatgaaaattttgtt	61
	R:gatgtatgtgaaggagagtagcg	
AA170	F:gtgcacaaggtgaaatgaatga	61
	R:ttgatgagttaaatcccgaagg	
AA238	F:tatcatcaaggtccaatttagt	51
	R:agctaaatcgtacctaatctgt	
AA278	F:ccaagaaaggcttatcaacagg	61
	R:tgcttgtgtcaagtgatcagtg	
AA335	F:acgcacacgcttagatagaaat	51
	R:atccaccataagttttggcata	
AB30	F:gattcttgaacatcgtgcagtg	61
	R:catttgagctttctggatgacg	
AB47	F:tccacaataccatctaaatgcca	51
	R:aatttgttcagttgaaatttcgtttc	
AC21	F:ttctgagttcaagccgaagttt	61
	R:tgctaatgatgcgtatgctgtt	
AD61	F:ctcattcaatgatgataatccta	51
	R:atgaggtacttgtgtgagataaa	
AD135	F:tggcattagattctccagcaca	61
	R:tgaggaggtgaacgtaaaagca	

conditions of an initial degradation at 95 °C (3 min), followed by 35 cycles of 94 °C (30 s), an annealing step with a range of melting temperatures (51 to 61 °C) for 1 min, and 72 °C (60 s), with a final extension at 72 °C (20 min), and there after maintaining at 4 °C. The amplified PCR products were separated on 2% agarose prepared in TAE buffer, together with a DNA molecular weight marker GeneRuler 100 bp Plus DNA Ladder (ThermoFisher Scientific, Waltham, USA) and visualized under UV after staining with ethidium bromide and documented using a gel documentation and image analysis system (VilberLourmat, Collégien, France). For each marker, amplification in part of the samples was repeated twice and only those markers that produced reproducible and consistent bands were selected for data generation.

Microsatellite data analysis

Amplicon profiles produced by microsatellites were compiled into a binary data matrix with each band scored "1" for presence and "0" for absence of each marker allele. For each SSR, we calculated the number of different alleles (Na), the Shannon's information index ($I = -1 \times \Sigma(pi \times \ln(pi))$, the observed heterozygosity (Ho; number of heterozygotes/N), the polymorphic information content ($PIC = 1 - \Sigma pi^2$, equivalent to the expected heterozygosity) (Anderson et al. 1993) and the fixation index (F). The detected intra-accessions variability (DIAV) per SSR represents the percentage of within-accessions polymorphic loci. Analysis of Molecular Variance (AMOVA) was carried out using a distance matrix. Permutations (999) were used to test for significance. All these calculations were performed using the Genalex 6.5 software (Peakall and Smouse 2012). For calculating D (Discrimination power) for each SSR marker the following formula was used $Dj = 1 - \Sigma[(pi \times (Npi - 1))/(N - 1)]$; where N is the total number of *Pisum* accessions and pi is the frequency of the ith allele at a given jth SSR locus. Multiplex ratio (MR = TB/TP where TB is total band and TP is total primers used), effective multiplex ratio (EMR = MR × FP where FP is fraction of polymorphism) and marker index (MI = EMR × Mean PIC) were calculated as proposed by Powell et al. (1996).

Phylogenetic relationships among accessions were estimated from the molecular data, using the package NTSYSpc 2.0 (Adams et al. 1998). Dendrograms were

constructed using the unweighted pair group method with arithmetic averaging (UPGMA) clustering analysis based on the genetic similarity coefficient matrices (Nei and Li 1979). Principal coordinate analysis (PCoA) was carried out to classify the local accessions of pea. Bootstraps analysis was performed with WinBoot (Yap and Nelson 1996) to evaluate the robustness of the nodes. Bootstrapping of the resulting dendrogram was conducted with 1000 permutations using Nei-Li similarity coefficients.

Population structure was estimated using a model-based Bayesian procedure implemented in the software Structure v. 2.3.4 (Pritchard et al. 2000). An admixture model with correlated allele frequencies without prior population information was used. The most informative number of subpopulations was identified using the K method (Evanno et al. 2005) with the aid of Structure Harvester (Earl and Vonholdt 2012). The estimated cluster membership coefficient matrices of the 20 runs were permuted so that all replicates have the closest match possible and then averaged across replicates, using the Greedy algorithm of the software CLUMMP (Jakobsson and Rosenberg 2007). To validate the predefined or the estimated population structure, we calculated pairwise Fst and Nei's standard genetic distance between populations (Nei and Li 1979). The reference distribution for P value calculation of the Fst analysis was calculated using 10,000 permutations. These analyses were performed with the Genalex 6.5 software (Peakall and Smouse 2012).

Results

Qualitative traits

Twenty-one qualitative characters were examined in 12 pea accessions. Six traits (stem fasciation, presence of leaflets, standard base shape, pod color, color of cotyledon and black hilum in seed) were not polymorphic, while the remaining showed a considerable level of variation among accessions, but not between the 3 years (supplementary Table 3). Those traits can be considered as suitable characters for classification and selection. Considering the polymorphic phenotypic characters that were covered, the traits that displayed a remarkable variability was the presence of anthocyanin, the development of the stipule and the leaflet size and dentation. Large differences were also present for the color of the wing and standard and the most frequent color was reddish purple and light purple, respectively. In this study, the correlations between the studied qualitative variables are shown in Fig. 1. To understand sources of variance among pea accessions, a multi-component analysis (MCA) was carried out. The first three axes described 55.37% of the total variation (Supplementary Fig. 1). Hence, three groups were obtained by multi-component analysis: the first group consisted of seven accessions (P2, P4, P6, P8, P9, P10, and P11) which were characterized by violet flower color, the second group with four accessions (P3, P5, P7, and P12) and characterized by white flower color, while the third group harbored just one accession, P1, with white flower color. Similarity coefficients were used to build a dendrogram using the Gower coefficient (Supplementary Fig. 2). This analysis grouped the 12 accessions into two groups white and violet flower color similar to the multi-component analysis.

Quantitative traits

The analysis of 27 quantitative traits indicated a widespread intraccession variability (the average coefficient of variation was 30.18%). Analysis of variance revealed that all sources of variation, including genotypes and years, were significant for all traits studied with probability $P < 0.05$, except for the first day of emergence of pea seeds for the factor year (Table 3), indicating that the performance of pea genotypes could be significantly affected by environmental conditions between the 3 years. High degree of variation was observed for all the traits with a coefficient of variation ranging between 15.25 and 75.5%, including variables that are considered the main components of legume productivity (pods per plant, seeds per pod, weight of 100 seeds), the phonological traits (emergence of seeds, days to the appearance of the first flower, days to first pod) and the physiological traits (plant height, number of stems per plant, the number of nodes per stem…). These characters were dependent on the environment. Flowering is considered very late when the number of days between sowing and appearance of the first flower exceeds 60 days (Solberg et al. 2015). Thus, all genotypes studied were divided into two groups; the earliest population with an average shorter than the

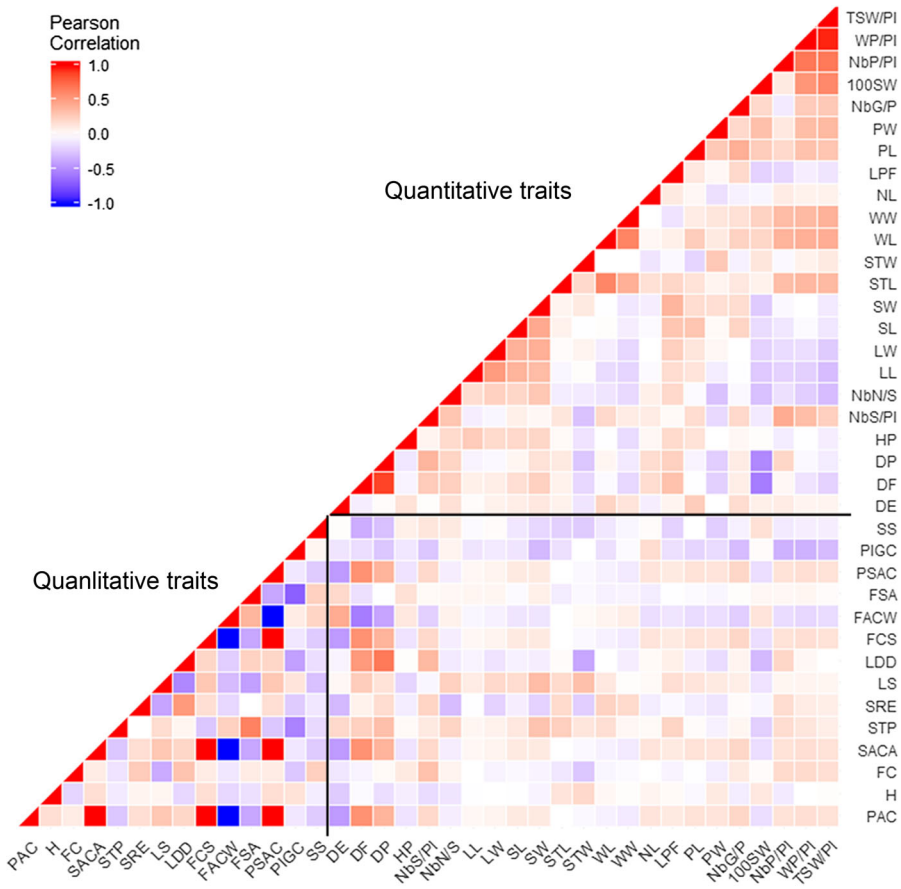

Fig. 1 Correlation matrix between the agro-morphological traits. Trait codes are in Supplementary Tables 1 and 2

overall genotype means 70.7 days and the late populations with the latest average of flowering. In this study, the correlations between the studied quantitative variables are shown in Fig. 1. The pod weight per plant was significantly and positively correlated with the number of pod per plant, total seed weight per plant and weight of 100 seeds with r = 0.68, 0.96 and 0.54 ($P < 0.01$) respectively. A negative phenotypic correlation was also found between the weight of 100 seeds and days to flowering and maturity with r = − 0.56 and − 0.51, respectively. To understand sources of variance among pea accessions, a principal component analysis (PCA) was carried out that grouped most traits into the first three axes describing 45.31% of the total variation (Supplementary Fig. 3). As was the case for the principal component analysis, hierarchical cluster analysis using 27 quantitative traits grouped 12 accessions of pea mainly into two clusters (Supplementary Fig. 4). Cluster I consisted of six accessions (P2, P3, P5, P6, P7, P8, and P9) which were characterized by high yielding, cluster II with three accessions (P10, P11, and P12) characterized by low yielding, and the two remaining accessions P4 (characterized by late flowering and maturity) and P1 (characterized by an early flowering and maturity), respectively, were grouped outside the two clusters and failed to form a cluster. Tunisian pea genotypes were clustered independently from their geographic origin.

Combined cluster and multiple factor analysis

In this study, the correlations between the studied quantitative and qualitative traits are shown in Fig. 1. The flowering days was significantly and positively correlated with the color of wing and standard in anthocyanin plant with r = 0.56. The multiple factor analysis (MFA) was performed on morphological quantitative and qualitative data (Supplementary Fig. 5). It is shown that 51.23% of the total variance

Table 3 Variation in quantitative traits among the 12 pea accessions

Traits	Mean			Mean	CV (%)	F	Sig
	Year 2016	Year 2017	Year 2018				
DE	5.59 ± 1.67	5.48 ± 2.03	5.67 ± 1.87	5.56 ± 1.86	33.45	0.793	Ns
D50E	8.9 ± 2.69	9.88 ± 3.17	10.5 ± 3.07	9.72 ± 3.06	31.48	29.84	< 0.0001***
DF	66.78 ± 12.45	71.33 ± 16.62	74 ± 16.65	70.7 ± 15.64	22.12	20.325	< 0.0001***
D50F	72.98 ± 13.97	77.09 ± 17.04	82.75 ± 17.13	77.6 ± 16.59	21.37	33.408	< 0.0001***
DP	103.15 ± 22.47	99.09 ± 20.93	104.09 ± 21.09	102.1 ± 21.6	21.15	5.5	< 0.05*
D50P	107.9 ± 21.41	102.92 ± 19.43	112.59 ± 19.67	107.8 ± 20.21	18.74	21.357	< 0.0001***
HP	91.97 ± 23.22	114.54 ± 15.11	108.65 ± 24.51	105.12 ± 23.38	22.25	108.258	< 0.0001***
NbS/Pl	11.48 ± 3.77	12.25 ± 4.22	8.9 ± 3.22	10.88 ± 4.01	36.85	78.124	< 0.0001***
NbN/S	15.13 ± 4.63	18.02 ± 4.56	18.92 ± 5.01	17.36 ± 5	28.80	62.615	< 0.0001***
LL	3.43 ± 1.26	4.97 ± 0.56	4.79 ± 0.78	4.4 ± 1.14	25.90	299.801	< 0.0001***
LW	2.19 ± 1.06	3.23 ± 0.54	3 ± 0.69	2.8 ± 0.91	32.5	171.646	< 0.0001***
SL	5.38 ± 1.66	6.75 ± 0.90	6.52 ± 0.75	6.21 ± 1.32	21.25	140.847	< 0.0001***
SW	4.67 ± 1.59	6.21 ± 1.52	6.45 ± 1.54	5.78 ± 1.74	30.1	138.653	< 0.0001***
STL	2.2 ± 0.49	2.1 ± 0.49	1.9 ± 0.44	2.07 ± 0.49	23.67	32.229	< 0.0001***
STW	3.38 ± 0.47	3.18 ± 0.47	3.26 ± 0.52	3.27 ± 0.5	15.29	14.872	< 0.0001***
WL	1.78 ± 0.43	1.58 ± 0.43	1.49 ± 0.30	1.62 ± 0.41	25.3	15.062	< 0.0001***
WW	1.75 ± 0.38	1.5 ± 0.38	1.53 ± 0.31	1.59 ± 0.38	23.89	51.87	< 0.0001***
NL	4.98 ± 0.91	4.89 ± 0.92	4.59 ± 0.85	4.82 ± 0.91	18.87	18.797	< 0.0001***
LPF	6.07 ± 2.7	7.4 ± 2.44	7.81 ± 2.21	7.09 ± 2.56	36.1	49.162	< 0.0001***
PL	5.85 ± 1.32	6.15 ± 1.47	5.58 ± 0.83	5.86 ± 1.26	21.5	18.426	< 0.0001***
PW	14.02 ± 2.58	13.32 ± 1.98	13.49 ± 1.92	13.61 ± 2.2	16.16	9.937	< 0.0001***
NbG/P	5.35 ± 1.8	4.63 ± 1.24	5.17 ± 1.39	5.05 ± 1.53	30.29	22.623	< 0.0001***
100SW	20.78 ± 7.63	16 ± 4.68	14.84 ± 4.03	17.21 ± 6.22	36.14	111.218	< 0.0001***
NbP/Pl	130.29 ± 68.15	105.93 ± 41.52	64.94 ± 20.09	100.38 ± 54.6	54.39	173.948	< 0.0001***
WP/Pl	130.58 ± 66.86	100.10 ± 45.5	53.73 ± 27.09	94.8 ± 58.48	61.68	222.284	< 0.0001***
TSW/Pl	130.54 ± 66.9	75.2 ± 47.62	39.82 ± 23.34	81.85 ± 61.8	75.50	309.785	< 0.0001***

For each trait, the table reports the coefficient of variation (CV %), average (mean) and significance value of the quantitative traits analyzed. Trait codes are in Table 3

***Very highly significant; **highly significant; *significant; *Ns* non-significant

was explained by the first five axes. Traits related to yield, color of flower and time of flowering and maturity can be considered powerful for studying the genetic diversity of pea accessions. Gower distance coefficients were calculated based on quantitative and qualitative traits (Fig. 2), and varied from 0.19 to 0.66, with a mean of 0.42 between all accessions. The coefficient of genetic similarity obtained in the present study ranged from 0 to 0.5, indicating that a high level of genetic diversity existed among pea accessions. Comparisons of data and cluster analysis (UPGMA) generate a dendrogram where 12 pea accessions were grouped into two main clusters. Cluster I contains (P1, P3, P5, P7, and P12) which are characterized by white flower color, while the Cluster II is composed by (P2, P4, P6, P8, P9, P10, and P11) characterized by violet flower. In the present study, the cluster results were similar to those of MFA analysis. The phylogenetic tree based on morphological data showed that the selected accessions did not correspond to patterns of geographical distribution.

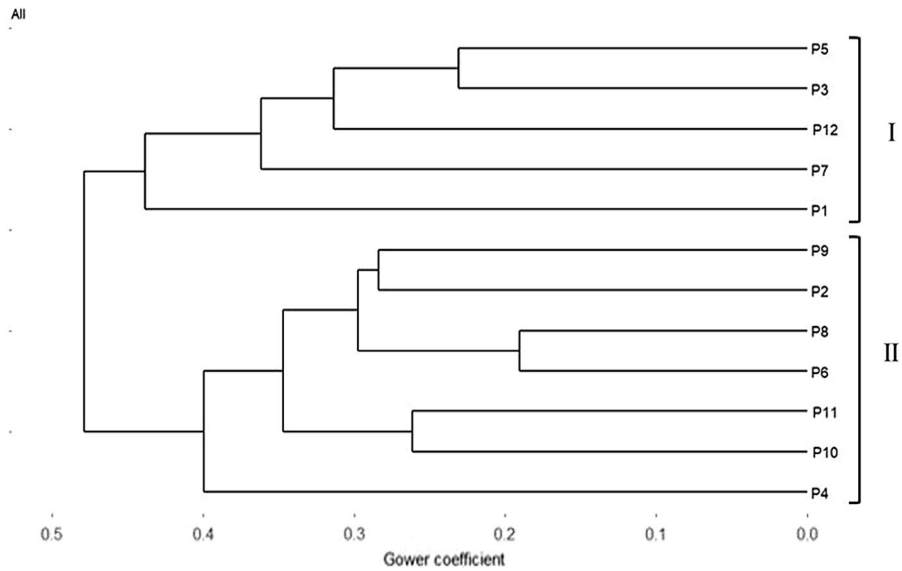

Fig. 2 Hierarchical cluster analysis of the 12 pea accessions based on quantitative and qualitative traits. UPGMA dendrogram based on similarities calculated with Gower distance coefficient

Molecular characterization

Genetic diversity and hierarchical classification

An analysis of the genetic diversity was carried out using 8 SSRs selected from the literature. All these loci were polymorphic in our germplasm collection. Main genetic parameters of the accession under investigation are presented in Table 4. We detected 23 alleles, whose length ranged from 120 to 420 bp. Differences were present for the number of alleles per locus, which ranged from 2 (PSGAP1, AA179, AD61) to 4 (AA72, A9) with an average of 2.88 per locus. Allele frequency spanned from a maximum value of 0.94 (PSGAP1) to a minimum of 0.27 of the AA72 locus. As anticipated for a selfing species, the observed heterozygosity was very low (on average, 5%) detected at AA72. A significant correlation between the number of alleles per locus and its observed heterozygosity was not present ($P > 0.05$; Spearman's Rho). PIC and D-values for each marker revealed the informativeness of marker in resolving the diversity among the accessions. A significant correlation between them was observed ($P < 0.01$; Spearman's Rho). The most informative locus was AA72, showing the higher diversity, number of alleles and PIC. Considerable differences among SSRs were present in the PIC, which ranged from 0.09 (PSGAP1) to 0.75 (AA72) with an average of 0.44. PIC significantly correlated with the number of alleles per SSR ($P < 0.05$; Spearman's Rho). Similarly, the highest D-value of 0.81 was for the marker AA72, followed by AA170 (0.70) and A9 (0.68). The PSGAP1 marker possessed the lowest number of diversity, number of alleles, PIC and D score. In addition, the mean values of Shannon information index per SSR marker ranged from 0.19 (PSGAP1) to 1.33 (AA72), with an average of 0.76. The Fixation index showed substantial positive values and limited differences among SSRs. All the employed SSRs were able to detect intra-accessions variability, although with a different efficiency. Intra-accessions variability was detected at D21, PSGAP1 and AD61 loci with a percentage of 8%. In order to perform a hierarchical partition of the genetic variation among accessions, we used the AMOVA statistical procedure. The results indicated that a considerable fraction of genetic variation (10%) was present within accessions (Table 5). Furthermore, Fst analysis was used to ascertain the degree of genetic differentiation among accessions. Out of the 78 possible pairwise combinations, 5 were not significant ($P > 0.05$). All of these indexes indicated that the polymorphic SSRs developed in the present study would be useful for the evaluation of the genetic variation of pea germplasm resources.

The principal coordinate analysis of the 13 pea accessions (Supplementary Fig. 6) accounted for a 68.33% of the total variation; the first and the second principal coordinates explained 41.52 and 26.81% of the molecular variance, respectively. The P4 accession

Table 4 Main genetic indices of the pea accessions obtained by SSR analysis

SSR	Size (pb)	Na	PIC	MAF	Ho	F	I	D	DIAV (%)
D21	280–380	3	0.45	0.71	0.00	1.00	0.78	0.48	8
PSGAP1	160–170	2	0.09	0.94	0.00	1.00	0.19	0.13	8
AA179	295–360	2	0.26	0.85	0.00	1.00	0.43	0.28	0
AA72	350–420	4	0.75	0.27	0.38	0.49	1.39	0.81	0
A9	360–390	4	0.63	0.46	0.00	1.00	1.19	0.68	0
AD61	120–150	2	0.46	0.65	0.00	1.00	0.65	0.50	8
AA170	290–350	3	0.64	0.44	0.00	1.00	1.06	0.70	0
AA278	150–190	3	0.27	0.85	0.00	1.00	0.53	0.29	0
Range/mean		2.875	0.44	0.65	0.05	0.94	0.78	0.5	3
Mean PIC	MR = TB/TP			EMR = MR*F			MI = EMR*Mean PIC		
0.44	2.89			2.67			1.15		

Na number of different alleles, *PIC* polymorphic information content, *MAF* major allele frequency, *Ho* observed heterozygosity, *F* fixation index, *I* Shannon's information index, *D* discrimination power, *DIAV* detected intra-accession variability, *TB* total number of bands scored, *TP* total primers, *MR* multiplex ratio, *EMR* effective multiplex ratio, *MI* marker index

Table 5 Analysis of molecular variance among and within accessions

Source	*df*	SS	Est. Var.	%	*P* value
Among accessions	12	217,200	1.795	90	0.001
Among individuals	52	7,600	0.000	0	0.997
Within individuals	65	13,500	0.208	10	0.001
Total	129	238,300	2.003	100	

df Degree of freedom, *SS* sum of squares calculated from a square genetic distance matrix, *Est. Var.* estimated variance

was separated from the rest and appears as composed of genetically distinct genotypes. The accessions (P1, P3, P5, P10, and P11) were discriminated along the coordinate 1 and closely placed in the plot area indicating a low variability among those accessions. The remaining accessions (P2, P6, P7, P8, P9, P12, and P13) were included in the same group in this analysis.

The results obtained by microsatellite DNA analysis revealed a clear separation of most pea accessions and showed a significant degree of inter-accession genetic diversity. A dendrogram was made using the Nei-Li similarity coefficient (Fig. 3). A high range of similarity was found among analyzed samples ranging from 0.24 to 1.00 (with an average of 0.62). The dendrogram depicts the pattern of relationships between the studied accessions. There is no clear clustering of accessions in relation with their growing area. Nevertheless, two major clusters can be defined by cutting the dendrogram at the lowest range of similarity value 0.24. The first cluster included one accession, P4, with intra-accession differences. The second group was formed by the remaining accessions. However, at around 0.43 similarity level, the second cluster subdivided into two subclusters. The first subcluster grouped the accessions P2, P6, P7, P8, P9, P12, and P13, while the second group was formed by accessions P1, P3, P5, P10 and P11.

Genetic structure analysis

The identification of genetically similar groups of plants was performed using an admixture model-based clustering analysis implemented in the software Structure. The Evanno's test indicated that the most

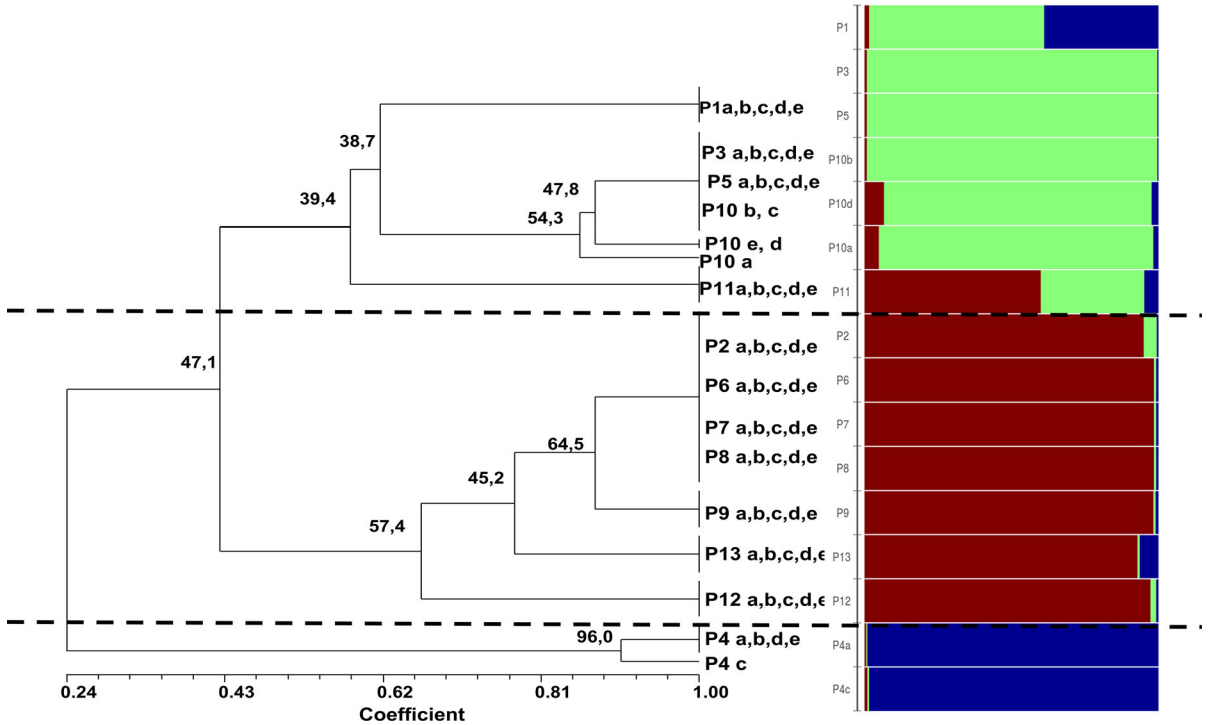

Fig. 3 Genetic relationship among accessions, **a** Dendrogram (UPGMA) of the pea accessions based on genetic distances calculated with the genetic similarity coefficient matrices The numbers near nodes represent the percentage of time when the node occurred among 10,000 bootstraps, **b** estimated population structure of the pea genotypes for K = 3. Each genotype is represented by a horizontal line, which is partitioned into colored segments that represent the estimated membership fractions in the K clusters

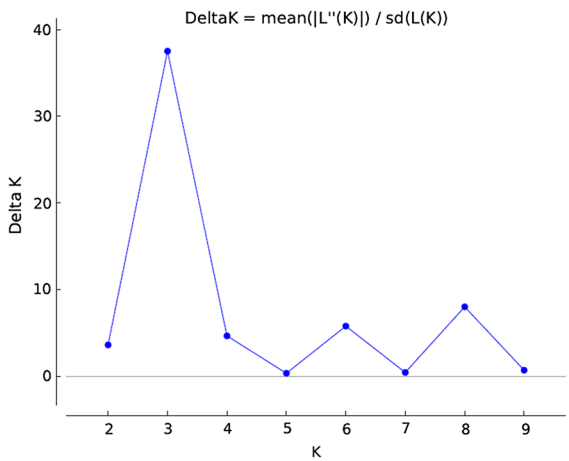

Fig. 4 Estimation of the optimum number of clusters for the pea genotypes according to the Evanno's method. The graph displays the Delta K [mean (lL″(K)l/SD(L(K))] for each K value

informative number of subpopulations (K) is 3, suggesting the existence of three major clusters in present pea accessions (Fig. 4). The inferred population structure is presented in Fig. 3. The groups defined by the Structure's analysis represent statistically different subpopulations, as indicated by the evaluation of genetic differentiation. Bayesian analysis performed in Structure confirmed the results obtained with PCA and the UPGMA dendrogram. All the accessions formed three subpopulations: subpopulation 1 consisting of accessions P1, P3, P5, P10 and P11, subpopulation 2 including accessions P2, P6, P7, P8, P9, P12 and subpopulation 3 containing just accession P4. Nonetheless, 90% of the samples had a membership coefficient higher than 0.8, indicating that the majority of the genotypes were strongly assigned to subpopulations. At K = 3, the five analyzed plants per accessions were attributed to the same subpopulation, suggesting that the genetic differences within accessions represent true accession variability. These results suggest that the country of origin does not represent a major factor of differentiation for the investigated accessions.

Comparison between agro-morphological and molecular cluster

The cophenetic coefficient (r = 0.13) estimated by molecular and agro-morphological data showed a lower correlation in comparison with the cophenetic coefficient with only the quantitative data (r = 0.54). For morphological traits, a clear separation of violet and white flower accessions was observed, where as using molecular markers no clear structure was detected. Nevertheless, some common groupings have been observed in both analyses based on UPGMA clustering. P4 was clustered alone in agro-morphological and molecular markers. The accessions (P2, P6, P8 and P9) and (P3 and P5) were grouped together in both analyses.

Discussion

Genetic diversity assessment of a species can facilitate the establishment of conservation strategies, the use in breeding programs, and the study of the crop evolution. To date, research on the genetic diversity of species grown in the arid region has been relatively limited. Previous studies have revealed that the different regions of Southern Tunisia are considered rich in traditional varieties of horticultural species, including pea (Loumerem et al. 2004). In this arid region, which are characterized by environmental stress such as drought, water salinity and soil fertility (Mezni et al. 2002), there are several traditional pea types well adapted to oasis conditions. However, until today, we cannot find in the market pea varieties improved for arid regions of Tunisia. This work represents the first investigation focused on the agro-morphological and molecular characterization of pea accessions collected from the arid region of Tunisia with the main goal to evaluate their genetic diversity and their population structure.

Morphological characterization is the first step in the description and classification of the germplasm, there have been a lot of successful examples in this area, such as maize (Couto et al. 2013), wheat (Li et al. 2012) or pea (Jha et al. 2013). These studies suggested that high diversity in morphological traits could be a useful tool for germplasm identification and selection. In this study, we selected 27 quantitative and 21 qualitative traits to characterize local pea accessions. Most of the characters analyzed here presented a considerable phenotypic diversity among the accessions. All the local accessions could be distinguished by at least one character, suggesting a significant diversification of the crop in the arid region of Tunisia. Similar results were also obtained on pea by Gixhari et al. (2014) and Wani et al. (2013). The parameters which showed the highest coefficient of variation and powerful discriminative potential were: 100 seeds weight, pod weight, plant height, flowering and maturity days, and standard and wing color. The results for these traits were consistent with previous studies (Ceyhan and Avci 2015; Georgieva et al. 2016). According to Bonny (2011), correlation data are very significant assets in varietal selection and breeding programs as they allow the improvement of several variables using just a few. Thus, the results showed that yield parameters were correlated with the variables related to the vegetative part of the plant. Similar results were obtained on pea by Martín-Sanz et al. (2011) in Spain, Ouafi et al. (2016) in Algeria and Gixhari et al. (2014) in Albania. PCA results provided information on the most appropriate morphological traits for varietal discrimination. In addition, the qualitative and quantitative phenotypic measurements revealed two accessions groups with white and violet color of the flower. The cophenetic correlation between qualitative dendrogram (Supplementary Fig. 2) and global dendrogram (Fig. 2) is 0.8, while between quantitative dendrogram (Supplementary Fig. 4) and global dendrogram (Fig. 2) is only 0.31. This result demonstrated the importance of the qualitative traits in this study. It is difficult or almost impossible to clarify the links between these accessions without making reference to other characterizations. Thus, molecular characterizations were needed to really classify local pea accessions collected from the arid region of Tunisia.

Twenty-three primers used in the present study were selected among a set of microsatellite markers developed for the Pea Microsatellite Consortium (Loridon et al. 2005), but just eight SSRs primers exhibited clear fragments and polymorphism on profiling. The number of alleles and a locus-specific allelic richness revealed the presence of a diffuse genetic variation. It was observed that the microsatellite markers used in our investigation were multi-allelic, detecting a number of alleles ranged from 2 to 4. This value is similar to that reported by Jain et al.

(2014) in ninety-six pea cultivars from wide geographical origins. However, Hagenblad et al. (2014) reported 5 to 10 alleles in the Swedish garden pea. The level of polymorphism was mainly assessed by PIC and D-values of markers (Shete et al. 2000). In our study, the average value of the PIC coefficient was 0.44. It was similar to the PIC value obtained previously for pea by Gong et al. (2010) and Ahmad et al. (2012) with an average of 0.41 and 0.46 respectively, but lower to other studies with an average of 0.52 and 0.62 reported by Loridon et al. (2005) and Smykal et al. (2008b), respectively. Marker AA72 was highly informative and had the maximum level of polymorphism with the highest PIC value of 0.75 and D-value of 0.81. A slightly lower level of polymorphism for AA72 (PIC 0.50) was detected by Loridon et al. (2005). The AMOVA analysis showed that a significant part of the genetic variation is present within accessions, indicating that, despite autogamy, the analysis of genetic variability on some plants per accessions might be useful in breeding programs. A similar result was found by Scarano et al. (2014) in common bean. Our study divided all 13 pea accessions into three main clusters that were further subdivided into subclusters, comparable to previous studies where cluster analysis grouped 35 pea accessions into two major clusters and eight subclusters (Xu et al. 2012; Zhuang et al. 2013). The information revealed by cluster analysis may be useful in designing a breeding program (Ahmad et al. 2012). Results of this study were congruent with results of Cupic et al. (2009), who suggested low to medium correlations among molecular and morphological traits. Clustering estimated both by SSR markers data and with distances estimated using agro-morphological traits showed no correlation between pea accessions and patterns of geographical origin. These results were in accordance with previous researches with several species, as pea (Solberg et al. 2015; Keneni et al. 2005), safflower (Khan et al. 2009) and mung beans (Sangiri et al. 2007). Baloch et al. (2015) observed a lack of association between genotypes and their origin of collections in cluster analysis among Turkish pea germplasm and suggested that it could be due to random genetic variation within Turkey. The random genetic variation could be due to gene flow arising from short distances between agroecologies which has led to routine exchange of planting materials among and between farmers who have been growing peas for decades. Migration of landraces among regions, followed by combining and introgression with pre-existent germplasm, may well be one more reason. The observed subpopulation structure in this study indicated that our germplasm collection could be clearly subdivided into three sub-populations, showing a convergence with the UPGMA cluster analysis. The membership coefficient of the genotypes to specific sub-populations was very high and possible admixture was detected in a reduced number of accessions. Pairwise analysis of Fst supported this hypothesis since lack of statistically significant genetic differentiation was present only for some comparisons between accessions.

Conclusions

The results showed the existence of wide spectrum diversity between the studied Tunisian pea accessions. For agro-morphological characterization, we found that qualitative traits were of greater importance than quantitative traits. SSRs were employed to study the genetic diversity and the population structure. Similar grouping was found with agro-morphological and SSR markers. Hence, we report that the agro-morphological and molecular data confirmed that the selected local pea accessions do not follow patterns of geographical origin. This genetic diversity can be used in future conservation and breeding programs in order to investigate the advantageous adaptation potentialities of this material.

Acknowledgements This work was partially supported by the Drylands and Oases Cropping Laboratory of the Arid Land Institute of Medenine (Tunisia). We gratefully acknowledge the Ministry of Higher Education and Scientific Research of Tunisia for the doctoral scholarship providing for a short stay (Amina Mohamed) at the Miguel Hernández University (Spain). The authors are thankful to the CEBAS-CSIC Fruit Breeding Group for the DNA quantification assistance. We also wish to thank all the members of the Department of Applied Biology of Miguel Hernandez University of Orihuela for their help in carrying out this work. We thank Ms. Yahia Hédi for his help in carrying out the field research work. The authors wish to thank the anonymous reviewers for helpful comments on the manuscript.

Compliance with ethical standards

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

References

Adams D, Kim J, Jensen R, Marcus L, Slice DE, Walker J (1998) NTSYSpc Version 2.02c. Applied Biostatistic Inc, St. Paul

Ahmad S, Singh M, Lamb-Palmer ND, Lefsrud M, Singh J (2012) Assessment of genetic diversity in 35 *Pisum sativum* accessions using microsatellite markers. Can J Plant Sci 92:1075–1081. https://doi.org/10.1007/s11105-016-1006-y

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

Baloch FS, Alsaleh A, Saenz de Miera LS, Hatipoglu R, Çiftçi V, KarakVoy T, Yıldız M, Ozkan H (2015) DNA based iPBS-retrotransposon markers for investigating the population structure of pea (*Pisum sativum*) germplasm from Turkey. Biochem Syst Ecol 61:244–252. https://doi.org/10.1016/j.bse.2015.06.017

Baranger A, Aubert G, Arnau G, Laine AL, Deniot G, Potier J, Weinachter C, Lejeune-Hénaut I, Lallemand J, Burstin J (2004) Genetic diversity within *Pisum sativum* using protein and PCR-based markers. Theor Appl Genet 108(7):1309–1321. https://doi.org/10.1007/s00122-003-1540-5

Bonny S (2011) L'agriculture écologiquement intensive: nature et défis. Cahagric 20(6):451–462. https://doi.org/10.1684/agr.2011.0526

Ceyhan E, Avci MA (2015) Determination of some agricultural characters of developed pea (*Pisum sativum* L.) lines. Int J Biol Biomol Agric Food Biotechnol Eng 9:12. https://doi.org/10.3923/javaa.2013.798.802

Couto EGdO, Davide LMC, Bustamante FdO, Pinho RGV, Silva TN (2013) Identification of haploid maize by flow cytometry, morphological and molecular markers. Ciência Agrotecnologia 37:25–31. https://doi.org/10.1590/S1413-70542013000100003

Cupic T, Tucak M, Popovic S, Bolaric S, Grljusic S, Kozumplik V (2009) Genetic diversity of pea (*Pisum sativum* L.) genotypes assessed by pedigree, morphological and molecular data. J Food Agric Environ 7:343–348. https://doi.org/10.1007/s11105-016-1006-y

De Giorgeo D, Polignano GB (2001) Evaluating the biodiversity of almond cultivars from germplasm collection field in Southern Italy. Sustain Global Farm 56:305–311

DGEDA (2011) Agrarian statistics of tunisia. General direction of studies and development agricultural of Tunisia. Ministry of Agriculture of Tunisia

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Duarte J, Rivière N, Baranger A, Aubert G, Burstin J, Cornet L, Lavaud C, Lejeune-Hénaut I, Martinant JP, Pichon JP, PiletNayel ML, Boutet G (2014) Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. BMC Genom 15:126. https://doi.org/10.1186/1471-2164-15-126

Earl DA, Vonholdt BM (2012) Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Esposito MA, Milanesi LA, Martin EA, Cravero VP, Lopez AFS, Cointry EL (2007) Principal component analysis based on morphological characters in pea (*Pisum sativum* L.). Int J Plant Breed 1(2):135–137

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620

Galili T (2015) dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. Bioinformatics 31:3718–3720

Georgieva N, NikolovaI I, Kosev V (2016) Evaluation of genetic divergence and heritability in pea (*Pisum sativum* L.). J Biol Sci Biotechnol 5(1):61–67

Gixhari B, Pavelkov ÁM, Ismaili H, Vrapi H, Jaupi A, Smýkal P (2014) Genetic diversity of Albanian pea (*Pisum sativum* L.) landraces assessed by morphological traits and molecular markers. Czech J Genet Plant 50:177–184. https://doi.org/10.1007/s11105-016-1006-y

Gong YM, Xu SC, Mao WH, Hu QZ, Zhang GW, Ding J, Li YD (2010) Developing new SSR markers from ESTs of pea (*Pisum sativum* L.). J Zhejiang Univ Sci B 11:702–707. https://doi.org/10.1631/jzus.B1000004

Hagenblad J, Bostrom E, Nygards L, Leino MW (2014) Genetic diversity in local cultivars of garden pea (*Pisum sativum* L) conserved on farm and in historical collections. Genet Resour Crop Evol 2(61):413–422. https://doi.org/10.1007/s10722-013-0046-5

Hoey BK, Crowe KR, Jones VM, Polans MO (1996) A phylogenetic analysis of *Pisum* based on morphological characters, allozyme and RAPD markers. Theor Appl Genet 92:92–100. https://doi.org/10.1007/BF00222957

Jain S, Kumar A, Mamidi S, McPhee K (2014) Genetic diversity and population structure among Pea (*Pisum sativum* L.) cultivars as revealed by simple sequence repeat and novel genic markers. Mol Biotechnol 56:925–938. https://doi.org/10.1007/s12033-014-9772-y

Jakobsson M, 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

Jha AB, Arganosa G, Tar'an B, Diederichsen A, Warkentin TD (2013) Characterization of 169 diverse pea germplasm accessions for agronomic performance, *Mycosphaerella* blight resistance and nutritional profile. Genet Resour Crop Evol 60:747–761. https://doi.org/10.1007/s10722-012-9871-1

Kassambara A, Mundt F (2017) factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.5. https://CRAN.R-project.org/package=factoextra. Accessed Dec 2017

Keneni G, Jarso M, Wolabu T, Dino G (2005) Extent and pattern of genetic diversity for morpho-agronomic traits in Ethiopian highland pulse landraces: 1. Field pea (*Pisum sativum* L.). Genet Resour Crop Evol 52:539–549. https://doi.org/10.1007/s10722-003-6016-6

Khan M, Witzke-Ehbrecht SV, Maass BL, Becker HC (2009) Relationships among different geographical groups, agromorphology, fatty acid composition and RAPD marker diversity in safflower (*Carthamus tinctorius*). Genet Resour Crop Evol 56:19–30. https://doi.org/10.1007/s10722-008-9338-6

Kharrat M, Le Guen J, Tivoli B (2006) Genetics of resistance to 3 isolates of *Ascochyta fabae* on faba bean (*Vicia faba* L.) in controlled conditions. Euphytica 151:49–61. https://doi.org/10.1007/s10681-006-9127-2

Kwon S, Brown AF, Hu J, McGee R, Watt C, Kisha T, Timmerman-Vaughan G, Grusak M, McPhee KE, Coyne CJ (2012) Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (*Pisum sativum* L.) core collection. Genes Genom 34:305–320. https://doi.org/10.1007/s13258-011-0213-z

Lázaro A, Aguinagalde I (2006) Genetic variation among Spanish pea landraces revealed by inter simple sequence repeat (ISSR) markers: its application to establish a core collection. J Agric Sci 144:53–61. https://doi.org/10.1017/S0021859605005848

Le S, Josse J, Husson F (2008) Facto MineR: an R package for multivariate analysis. J Stat Softw 25(1):1–18. https://doi.org/10.18637/jss.v025.i01

Li X, Xu X, Yang X, Li X, Liu W, Gao A, Li L (2012) Genetic diversity among a founder parent and widely grown wheat cultivars derived from the same origin based on morphological traits and microsatellite markers. Crop Pasture Sci 63:303–310. https://doi.org/10.1071/CP11302

Loridon K, McPhee K, Morin J, Dubreuil P, Pilet-Nayel ML, Aubert G, Rameau C, Baranger A, Coyne C, Lejeune-Hènaut I, Burstin J (2005) Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). Theor Appl Genet 111:1022–1031. https://doi.org/10.1007/s00122-005-0014-3

Loumerem M, Moussa M, Bellachheb C (2004) La collecte et l'étude de la diversité génétique des éspèces cultivées aux aménagements hydrauliques dans les régions arides tunisiennes. In: Revue des Régions Arides, Actes de Séminaire International, Aridoculture et Cultures Oasiennes

Martín-Sanz A, Palomo JL, de la Vega MP, Caminero C (2011) Identification of pathovars and races of *Pseudomonas syringae*, the main causal agent of bacterial disease in pea in North-Central Spain, and the search for disease resistance. Eur J Plant Pathol 29:57–69. https://doi.org/10.1007/s10658-010-9691-0

Mezni M, Albouchi A, Bizid E, Hamza M (2002) Effet de la salinité des eaux d'irrigation sur la nutrition minérale chez trois variétés de luzerne pérenne (*Medicago sativa*). Agronomie 22:283–291

Mikić A, Mihailović V, Ćupina B, Kosev V, Warkentin T, McPhee K, Ambrose M, Hofer J, Ellis N (2011) Genetic background and agronomic value of leaf types in pea (*Pisum sativum*). Rat povrt/Field Veg Crop Res 48:275–284

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Ouafi L, Alane F, Bouziane H, Abdelguerfi A (2016) Agro-morphological diversity within field pea (*Pisum sativum* L.) genotypes. Afr J Agric Res 11(40):4039–4047. https://doi.org/10.5897/ajar2016.11454

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for gemplasm analysis. Mol Breed 2:225–238. https://doi.org/10.1007/BF00564200

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi locus genotype data. Genetics 155:945–959

Sangiri C, Kaga A, Tomooka N, Vaughan DA, Srinives P (2007) Genetic diversity of the mung bean (*Vigna radiata*, Leguminosae) gene pool on the basis of microsatellite analysis. Aust J Bot 55:837–847. https://doi.org/10.1071/BT07105

Scarano D, Rubio F, Ruiz JJ, Rao R, Corrado G (2014) Morphological and genetic diversity among and within common bean (*Phaseolus vulgaris* L.) landraces from the Campania region (Southern Italy). Sci Hort 180:72–78. https://doi.org/10.1016/j.scienta.2014.10.013

Shete S, Tiwari H, Elston RC (2000) On estimating the heterozygosity and polymorphism information content value. Theor Popul Biol 57:265–271. https://doi.org/10.1006/tpbi.2000.1452

Smykal P, Hybl M, Corander J, Jarkovsky J, Flavell AJ, Griga M (2008a) Genetic diversity and population structure of pea (*Pisum sativum* L.) varieties derived from combined retrotransposon, microsatellite and morphological marker analysis. Theor Appl Genet 117:413–424. https://doi.org/10.1007/s00122-008-0785-4

Smykal P, Horacek J, Dostalova R, Hybl M (2008b) Variety discrimination in pea (*Pisum sativum* L.) by molecular, biochemical and morphological markers. J Appl Genet 49:155–166. https://doi.org/10.1007/BF03195609

Sneath P, Sokal R (1973) Numerical taxonomy: the principles and practice of numerical classification. S. W. Freeman, San Francisco

Snowdon RJ, Friedt W (2004) Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. Plant Breed 123:1–8. https://doi.org/10.1111/j.1439-0523.2003.00968.x

Solberg SO, Brantestam AK, Olsson K, Leino MW, Weibull J, Yndgaard F (2015) Diversity in local cultivars of *Pisum sativum* collected from home gardens in Sweden. Biochem Syst Ecol 62:194–203. https://doi.org/10.1016/j.bse.2015.09.004

Ubayasena L, Bett K, Tar'an B, Vijayan P, Warkentin T (2010) Genetic control and QTL analysis of cotyledon bleaching resistance in green field pea (*Pisum sativum* L.). Genome 53:346–359. https://doi.org/10.1139/g10-117

UPOV (1990) International Union for the protection of New Varieties of Plants (1981 and 1990). Guidelines for the conduct of tests for distinctness, uniformity and stability. UPOV TG/7/10. Geneva, Switzerland

Wani G, Mir B, Shah M (2013) Evaluation of diversity in pea (*Pisum sativum* L.) genotypes using agro-morphological characters and RAPD analysis. Int J Curr Res Rev 5(10):17–25. https://doi.org/10.1007/BF03195609

Xu S, Gong Y, Mao W, Hu Q, Zhang G, Fu W, Xian Q (2012) Development and characterization of 41 novel EST-SSR markers for *Pisum sativum* (Leguminosae). Am J Bot 1:1–4. https://doi.org/10.3732/ajb.1100445

Yap V, Nelson RJ (1996) Win Boot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. International Rice Research Institute, Manila, Philippines (online). http://irri.org/science/software/winboot.asp (cited 23 December 2005). Accessed Nov 2005

Zhuang X, McPhee KE, Coram TE, Peever TL, Chilvers MI (2013) Development and characterization of 37 novel EST-SSR markers in *Pisum sativum* (Fabaceae). Appl Plant Sci 1:1–4. https://doi.org/10.3732/apps.1200249