JJBS

Volume 11, Number 2,June 2018
ISSN 1995-6673
Pages 201 - 207

Jordan Journal of Biological Sciences

Genetic Diversity and Structure Analysis of Pea Grown in Iraq Using Microsatellite Markers

Nawroz A. R. Tahir[1,*], Djshwar D. Lateef[2], Dlshad A. Omer[3], Shadia H. S. Kareem[2], Dastan A. Ahmad[2] and Lanja H. Khal[1]

[1]*Horticulture Department;*[2]*Crop Science Department, College of Agricultural Sciences, University of Sulaimani ;* [3] *Ministry of Agriculture, General Directorate of Agriculture, Sulaimani, Iraq*

Received October 12, 2017; Revised November 15, 2017; Accepted November 20, 2017

Abstract

The diversity and population structure of twenty-two genotypes of peas (*Pisum sativum L.*) grown in Iraq were estimated using microsatellite [simple sequence repeats (SSRs)]. Nineteen SSR primers generated sixty-eight polymorphic bands with an average of 3.789 polymorphic bands per primer. The highest number of polymorphic bands stated for the locus AA5 (eight alleles). Means of gene diversity, PIC, marker index, resolving power and Shannon index were 0.562, 0.513, 2.090, 2.703 and 0.833, respectively. Principal component analysis and hierarchical cluster analysis clustered the pea genotypes into three major clades. Genotypes G1 (ORP-2), G5 (ORP-11) and G6 (ORP-12) and G7 (ORP-13) were genetically the most distant from the other genotypes. Analysis of molecular variance (AMOVA) revealed that differences among the populations accounted for 14 % of the total variation, whereas difference within the population was 86 %. The population structure assay conceded that the genotypes were grouped into two evident subpopulations. Only three genotypes: G4 (ORP-10), G11 (Ns minima) and G17 (Karina) were considered to be admixture. The multi-locus F_{ST} analysis revealed strong differentiation within the populations and significant isolation by distance. The results of this study revealed that different origins of genotypes had played a remarkable role in shaping the current patterns of genetic variation among these populations, many of which serve as good candidates for conservation and breeding programs.

Keywords: Clustering, Genetic diversity, *Pisum sativum*, SSRs.

1. Introduction

Pea (*Pisum sativum L.*) was the basic model organism used in Mendel's discovery of the laws of inheritance, making it the basis of modern plant genetics. Nevertheless, subsequent progress in pea genomics has become last among other crop species. Pea is a diploid plant with chromosome number 2n = 14 (Smýkal *et al*., 2012), it represents a major pulse crop grown for its protein content. In different zones of the world, it is an essential component of agro-ecological cropping systems. Genotypes identification and the evaluation of genetic variation of populations are important in genotypes protection and breeding program (Smýkal *et al*., 2012). Global climate change and the new technologies make pea breeders conduct more effective methods of selection and take benefit of the large genetic diversity present in the *Pisum sativum* gene pool (Smýkal *et al*., 2012).

Geneticists and plant breeders have affirmed the need for additional development in capturing and harnessing genetic variability. Several approaches are accessible to assess the diversity of genotypes. Traditional morphological or biochemical markers are restricted and are not fully-reliable because of the influence of the environment. The markers have been superseded by DNA-based methods that generate fingerprints (Hollingsworth, 2006). On this basis, various studies have been so far conducted on peas employing Random Amplification of Polymorphic DNA (RAPD), Simple Sequence Repeats (SSRs), *Retrotransposon-Based Insertion* Polymorphism (RBIP) markers, Single Nucleotide Polymorphism (SNP) marker, and Targeted Induced Local Lesions In Genomes (TILLING). These have been established as useful methods for the genetic diversity assessment (Deulvot *et al.*, 2010; Ahmad *et al*., 2015; Tahir *et al*., 2016). Among the most widely-used markers in crop species are the microsatellite [Simple Sequence Repeats (SSRs)] (Sarikamiş *et al*., 2010). SSR technique amplifies repetitive or motifs region (ranging in length from 1-6 or more base pairs) of DNA. The best benefits of microsatellite analysis include the accuracy, high polymorphism, co-dominance and genome coverage (Tahir, 2010; Lateef, 2015). These types of markers are frequently used for genome mapping, estimation of genetic diversity, gene tagging and marker-assisted selection (Smýkal *et al*., 2008). In the majority of the cultivated crops, very limited genome sequence information is

* Corresponding author. e-mail: nawroz.tahir@univsul.edu.iq.

available especially for pea compared to other legume crops (Deulvot *et al*., 2010). The present study was carried out to test the suitability of applying SSR techniques on pea genotypes, and to asses the efficency of the defined markers set for diversity studies in a collection of *Pisum sativum.*

2. Materials and Methods

2.1. Genotypes Collection

A total of twenty-two pea genotypes were investigated in the present work (Table 1); these genotypes were gathered from the Agriculture Research Station, Ministry of Agriculture in Sulaimani, Kurdistan region, Iraq.

Table 1. Name and origins of pea (*Pisum sativum L*.) genotypes examined in this study.

No.	Code	Genotype Name	Origin
1	G1	ORP-2	Iraq
2	G2	ORP-3	Iraq
3	G3	ORP-8	Iraq
4	G4	ORP-10	Iraq
5	G5	ORP-11	Iraq
6	G6	ORP-12	Iraq
7	G7	ORP-13	Iraq
8	G8	ORP-15	Iraq
9	G9	ORP-58	Iraq
10	G10	ORP-D1	Iraq
11	G11	Ns minima	Australia
12	G12	Green sage	USA
13	G13	Oregon	USA
14	G14	Provence	Italy
15	G15	Javor	Australia
16	G16	Legacy	Turkey
17	G17	Karina	Turkey
18	G18	Topaz	Turkey
19	G19	Vada nunheins	Turkey
20	G20	Bolero	Turkey
21	G21	Lancet	Germany
22	G22	Rainer	Turkey

2.2. DNA Extraction

By performing cetyltrimethylammonium bromide (CTAB), Genomic DNA was isolated from fresh pea leaves (Doyle, 1991). Then, the concentration and quality of the extracted DNA were determined by 1% (w/v) agarose gels using a Bio-Rad gel imaging system. Finally, the extracted genomic DNA was diluted to 40 ng/µL and stored at (-20 °C).

2.3. Molecular Marker (SSR) Assay

Out of the twenty-four markers, nineteen primers (Reis and Diogo, 2012; Bouhadida *et al*., 2013; Ahmad *et al*., 2015) were found to be polymorphic, and those were utilized for diversity and structure studies (Table 2). PCR reaction was conducted in a 20 µL reaction containing 1X PCR buffer, 200 mM dNTPs, 0.40 µL of primer, 4mM $MgCl_2$, 1µL Taq polymerase and 80ng templates DNA. PCR amplification was conducted using Applied Biosystems Thermocycler following the PCR protocol: Initial denaturation at 94ºC for seven minutes, thirty-seven cycles at 94ºC for one minute, 50, 54, 55 and 60ºC for one minute, 72ºC for two minutes and a final extension step at 72ºC for seven minutes. Amplified products were resolved on 2.4% agarose gels at 87 V in 1X TBE buffer, and fragment sizes were determined by 100-bp DNA ladder (Invitrogen, USA).

2.4. Statistical Data Analysis

The scorable bands were coded manually as either present (1) and absent (0). Scored data were applied for the calculation of Jaccard's similarity coefficient using XLSTAT 2017 software (XLSTAT, 2017). The Jaccard's coefficient was converted to dissimilarity. The Jaccard's coefficient was converted to dissimilarity. The dissimilarity matrix was used for unweighted pair-group method with arithmetic averages (UPGMA) dendrogram constructed by using XLSTAT 2017 software. The Polymorphism information content (PIC) for SSR markers was calculated using the following formula: $PIC_i = 1-\Sigma Pij^2$, where PICi is the PIC of marker i; Pij is the frequency of the j^{th} pattern for marker i (Kumari *et al*., 2013). Gene diversity and major of allele frequency were computed by using Power Marker version 3.25 software (Liu and Muse, 2005). To determine the relationship among different genotypes, the principal component analysis (PCA) was conducted by XLSTAT 2017 software. GenAlEx (version 6.5) software also used to estimate the molecular variance among and within the populations (Peakall and Smouse, 2012). Marker index (MI) and resolving power (RP) are measured according to (Powell *et al*., 1996; Prevost and Wilkinson, 1999). For the population structure, a model analysis was fulfilled to infer the genetic structure and to clarify the number of sub-populations using the software STRUCTURE (version 2.3.4) (Pritchard *et al*., 2000). The number of supposed populations (K) was set from one to ten, and the analysis was repeated two times. For each run, the burn-in and MCMC were fixed to 50,000 each, and iterations were deposited to 5. The run with the maximum likelihood was employed to set genotypes into sub-populations.

Table 2. Name, sequence, and annealing temperature of SSRs primers used in this research.

No	SSR markers	F 5'- 3'	R 5'- 3'	AT
1	A5	GTAAAGCATAAGGGGTTCTCAT	CAGCTTTTAACTCATCTGACA	50
2	A6	CTTAAGAGAGATTAAATGGACAA	CCAACTCATAATAAAGATTCAAA	50
3	AA205	TACGCAATCATAGAGTTTGGAA	AATCAAGTCAATGAAACAAGCA	50
4	AA355	AGAAAAATTCTAGCATGATACTG	GGAAATATAACCTCAATAACACA	50
5	AB53	CGTCGTTGTTGCCGGTAG	AAACACGTCATCTCGACCTGC	50
6	AD61	CTCATTCAATGATGATAATCCTA	ATGAGGTACTTGTGTGAGATAAA	50
7	AA92	AAGGTCTGAAGCTGAACCTGAAGG	GCAGCCCACAGAAGTGCTTCAA	50
8	AD73	CAGCTGGATTCAATCATTGGTG	ATGAGTAATCCGACGATGCCTT	50
9	AA372.1	GAGTGACCAAAGTTTTGTGAA	CCTTGAACCCATTTTTAAGAGT	50
10	D23	ATGGTTGTCCCAGGATAGATAA	GAAAACATTGGAGAGTGGAGTA	50
11	AA5	TGCCAATCCTGAGGTATTAACACC	CATTTTTGCAGTTGCAATTTCGT	50
12	AD59	TTGGAGAATGTCTTCTCTTTAG	GTATATTTTCACTCAGAGGCAC	50
13	A9	GTGCAGAAGCATTTGTTCAGAT	CCCACATATATTTGGTTGGTCA	50
14	X13.1a	GAACTAGAGCTGATAGCATGT	GCATGCAAAAGAACGAAACAGG	50
15	PSGAPA1a	GACATTGCCAATAACTGG	GGTTCTGTTCTCAATACAAG	54
16	PEACP LHPPSa	GTGGCTGATCCTGTCAACAA	CAACAACCAAGAGCAAAGAAAA	54
17	PSBLOX13.2a	CTGCTATGCTATGTTTCACATC	CTTTGCTTGCAACTTAGTAACAG	54
18	PSCAB66a	CACACGATAAGAGCATCTGC	GCTTGAGTTGCTTGCCAGCC	54
19	PSMPSAA278b	CCAAGAAAGGCTTATCAACAGG	TGCTTGTGTCAAGTGATCAGTG	60

3. Results and Discussion

3.1. Diversity Parameters

The SSRs are suitable markers of choice employed by different researchers for genetic variation analysis in different crops (Sarikamiş *et al*., 2010). In the present study, twenty-four SSRs primers were employed for the studying of twenty-two pea genotypes. Nineteen SSRs primers exhibited clear fragments and polymorphism on profiling (Tables 3). Nineteen SSRs makers produced sixty-eight polymorphic alleles with an average of 3.789, ranging from two to eight per primer. The level of polymorphism reported here is consistent with the data obtained in Burstin *et al.*(2001) where 3.6 alleles per polymorphic marker were observed for thirty-one markers derived from gene sequences, even though the panel included twelve genotypes. Cupic *et al.* (2009) reported an average of 4.5 alleles per locus using thirty SSRs markers in a population of eighteen pea accessions. However, a higher mean of 5.9 alleles per locus was detected in twenty pea varieties and fifty-seven wild pea accessions using ten SSRs markers (Nasiri *et al.*, 2010). This greater mean of alleles per locus may be due to the wild pea genotypes in their study thus making the average of total detected alleles higher. In this study, the highest number of polymorphic bands were stated for the locus AA5 (8 alleles), and the minimum was detected for A6, AD73, AA205 and PSMPSAA278b markers (2 alleles). Similarly, Nasiri *et al.* (2010) observed 8 alleles for marker AF004843. The number of unique positive bands was 7 across all genotypes. The maximum number of unique bands was recorded by PEACPLHPPSa (2 bands). PIC, marker index and resolving power values indicate the relative informativeness of each marker and the average PIC, marker index and resolving power values, in the present research, were found to be 0.513, 2.090 and 2.703, respectively. PIC, marker index and resolving power values ranged between 0.228 for PSCAB66a to 0.935 for AA5, 0.507 for A6 to 7.476 for AA5 and 0.636 for AA205 to 7.636 for AA5, respectively. This result suggests that alleles of marker AA5 were uniformly distributed among pea genotypes. In the present study, the high value of polymorphism parameters is owing to the efficiency of the selected SSRs primers. Heterozygosity was detected to be very low which may be due to autogamous nature of pea. Gene diversity calculated according to Nei (1973) varied from 0.298 (A6) to 0.938 (AA5) with the average of 0.562 (Table 3). Shannon index was determined, and was stated between 0.276 and 2.416 with a mean of 0.833. Loridon *et al.* (2005) obtained a PIC value of 0.73 for locus AA206. The PIC in this research is higher than that reported by Burstin *et al.* (2001) (four alleles). Ahmad *et al*. (2015) stated that marker AA121 was highly informative and had the maximum level of polymorphism with the highest PIC value of 0.887 and a resolving power value of 0.901.

Table 3. Details of 19 SSRs primers and their genetic diversity parameters used for genotyping in the 22 pea genotypes.

Primers	TAF	TPB	TPB%	Size	NUB	Major Allele Frequency	Gene Diversity	PIC	Marker index	Resolving power	H
A5	4	4	100.000	200-1000	0	0.591	0.603	0.569	2.276	3.909	0.931
A6	2	2	100.000	150-180	0	0.818	0.298	0.253	0.507	2.000	0.474
A9	6	5	83.333	340-1200	0	0.455	0.736	0.709	3.546	2.727	1.264
AA205	3	2	66.667	260-660	0	0.682	0.479	0.427	0.854	0.636	0.276
AA355	4	3	75.000	190-720	1	0.636	0.541	0.496	1.489	2.818	0.528
AA372.1	5	4	80.000	225-1150	0	0.318	0.777	0.744	2.974	3.273	1.037
AA430902a	2	2	100.000	420-450	0	0.545	0.496	0.373	0.746	2.000	0.689
AA5	8	8	100.000	180-840	0	0.091	0.938	0.935	7.476	7.636	2.416
AA92	8	5	62.500	150-950	1	0.364	0.744	0.706	3.530	2.636	1.119
AB53	5	4	80.000	250-1200	1	0.591	0.545	0.471	1.884	2.818	0.517
AD59	5	4	80.000	530-1240	0	0.818	0.322	0.311	1.245	2.273	0.316
AD61	7	7	100.000	120-500	1	0.364	0.806	0.786	5.505	6.364	1.493
AD73	3	2	66.667	220-275	0	0.545	0.591	0.522	1.045	1.909	0.653
D23	3	3	100.000	150-200	0	0.409	0.649	0.574	1.721	2.000	1.070
PEACP LHPPSa	5	3	60.000	150-650	2	0.818	0.318	0.302	0.905	2.000	0.368
PSCAB66a	4	3	75.000	200-700	1	0.864	0.244	0.228	0.685	0.455	0.576
PSGAPA1a	4	3	75.000	150-1200	0	0.409	0.702	0.653	1.958	1.909	0.842
PSMPSAA278b	2	2	100.000	180-200	0	0.727	0.397	0.318	0.636	2.000	0.586
X13.1a	3	2	66.667	250-750	0	0.591	0.483	0.367	0.733	2.000	0.677
Total	83	68	81.928		7						
Mean	4.368	3.789				0.560	0.562	0.513	2.090	2.703	0.833

TAF: Total number of fragments, TPB: Total number of polymorphic bands, NUB: Number of unique bands, PIC: Polymorphism information content, H: Shannon Index.

3.2. Clustering and Genetic Relationship

Clustering and principal component analysis (PCA) avail as a platform to supply a spatial clarification of the comparative genetic dissimilarity between the genotypes. It also estimates the robustness of the differentiation among the clusters assorted by the dendrogram (Liu *et al.*, 2013). A fan dendrogram of the SSRs data demonstrated clear groupings of pea genotypes on the basis of origin (Figure 1). The UPGMA depended on cluster analysis for SSRs alleles which exhibited that twenty-two genotypes were clustered into three major clades at the dissimilarity coefficient of 0.520. Clade I had four genotypes: G3 (ORP-8), G8 (ORP-15), G7 (ORP-13) and G11 (Ns minima). Clade II is the large cluster including thirteen genotypes which are represented by five sub-groups at a dissimilarity coefficient of 0.45. The first sub-group consisted of G17 (Karina), G18 (Topaz), G19 (Vada nunheins), G20 (Bolero) and G22 (Rainer). G10 (ORP-D1), G4 (ORP-10), G14 (Provence) and G13 (Oregon) belonging to sub-group two. The third sub-group composed of G12 (Green sage) and G15 (Javor) while the two genotypes G16 (Legacy) and G21 (Lancet) created the fourth and fifth sub-groups, respectively. Group III could be divided into two sub-clades at a dissimilarity coefficient of 0.48. The subgroup I consisted of two genotypes, *viz.*, G2 (ORP-3) and G9 (ORP-58). The subgroup II comprised G1 (ORP-2), G5 (ORP-11), and G6 (ORP-12). The minimum genetic dissimilarity of 0.179 was found between G12 (Green sage) and G15 (Javor) followed by 0.216 between G18 (Topaz) and G19 (Vada nunheins). The maximum genetic dissimilarity of 0.707 was found between G5 (ORP-11) and G12 (Green sage) followed by 0.705 between G5 (ORP-11) and G21 (Lancet). Ahmad *et al.* (2015) scored a range of 0.075-0.875 of genetic similarity among thirty-five accessions of pea using SSRs markers.

In our PCA scatter plots, the first two principal components stated 9.80 and 16.70% of the total variation, respectively. In accordance with the dendrogram result, pea genotypes were clearly separated into three major clusters (Figure 2). As shown in Figures 1 and 2, the genotype G1 (ORP-2), G5 (ORP-11) and G6 (ORP-12), and G7 (ORP-13) were far distant genotypes from the others. These results reveal that the genotypes taken in the study are genetically diverse. The genetic distance identified in this research can be employed in the crossing programs. The number of clusters in the current study is higher than that obtained by Bouhadida *et al.* (2013) (2 clusters) and Reis and Diogo (2012) (2 clusters), and is lower than that stated by Ahmad *et al.* (2015) (4 clusters).

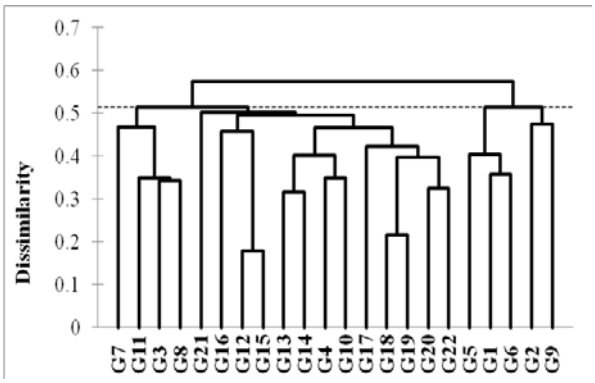

Figure 1. Clustering of 22 genotypes using UPGMA based on SSRs data. Genotypes: G1:ORP-2, G2:ORP-3, G3:ORP-8, G4:ORP-10, G5:ORP-11, G6:ORP-12, G7:ORP-13, G8:ORP-15, G9:ORP-58, G10:ORP-D1, G11:Ns minima, G12:Green sage, G13:Oregon, G14:Provence, G15:Javor, G16:Legacy, G17:Karina, G18:Topaz, G19:Vada nunheins, G20:Bolero, G21:Lancet and G22:Rainer.

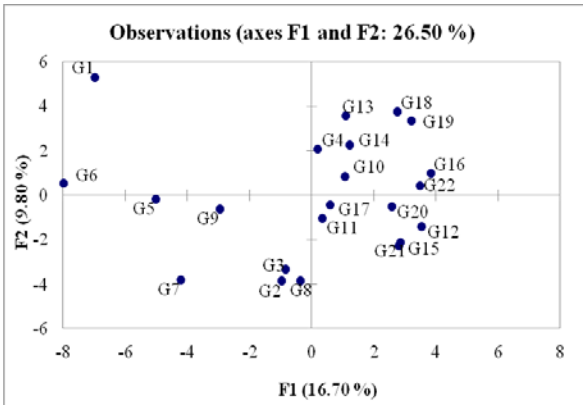

Figure 2. PCA of 22 pea genotypes based on 19 SSRs loci. The two PCA axes accounted for 16.70 and 9.80 % of the total genetic variation. Genotypes: G1: ORP-2, G2: ORP-3, G3: ORP-8, G4: ORP-10, G5: ORP-11, G6: ORP-12, G7: ORP-13, G8: ORP-15, G9: ORP-58, G10: ORP-D1, G11: Ns minima, G12: Green sage, G13: Oregon, G14: Provence, G15: Javor, G16: Legacy, G17: Karina, G18: Topaz, G19: Vada nunheins, G20: Bolero, G21: Lancet and G22: Rainer.

3.3. Analysis of Genetic Variation Among Pea Genotypes using AMOVA

For the assessment of genetic differentiation among the tested pea genotypes six populations (in term of origins), analysis of molecular variance (AMOVA) was conducted. The result exhibited that the pea genotypes were significantly distinct from their relatives at P-value of 0.002 (Table 4). The variance among the populations clarified 14% and within the populations illustrated 86% of genetic divergence. The pair-wise PhiPT value (which is corresponding to FST in the assessment of genetic differentiation) was 0.139 and denoted relatively large deal of discrimination among pea populations. This result indicated that the degree of variation within the population is higher than that obtained among populations suggesting the existence of a significant number of specific and rare alleles and a divergence in allele frequencies among the genotypes. In the current research, the variation among populations is lesser than the mean of differentiation among populations (41%) reported by Teshome *et al*. (2015) in the Ethiopian field pea. Wang *et al*. (2015) has demonstrated 41% of differentiation among populations in 266 grass pea accessions. The differentiation within the population in this study is far greater than that found by (Teshome *et al.*, 2015; Wang *et al*., 2015)

Table 4. AMOVA among six populations (origins) and within population based on 22 SSRs loci of pea genotypes.

Source	df	SS	MS	Est. Var.	%	P
Among Pops	5	78.717	15.743	1.622**	14%	0.002
Within Pops	18	180.200	10.011	10.011**	86%	0.002
Total	23	258.917		11.633	100%	

3.4. Population Structure Analysis

Based on nineteen SSRs markers, the estimation of the population structure of twenty-two pea genotypes was performed by STRUCTURE software with Pritchard correction. The structure simulation with STRUCTURE HARVESTER displayed that the L (*K)* value had the maximum peak at $K = 2$ (Figure 3), inferring that two populations can incorporate all individuals from the twenty-two genotypes with the highest likelihood. This suggested the existing of two major model populations, which were visualized in the graph (Figure 4). Based on the membership fractions, the genotypes with the probability of ≥ 80% were assigned to corresponding subgroups or subpopulations with others categorized as an admixture (Table 5). Subgroup-1 or subpopulation-1 included eight genotypes with most of the genotypes being of local origin, and subgroup-2 consisted of eleven genotypes composed of non-Iraqi genotypes. Only three genotypes: G4 (ORP-10), G11 (Ns minima), and G17 (Karina) were considered to be admixture (Table 5). The population differentiation (FST) metric for the subgroup-1 and subgroup-2 was 0.0478 and 0.267 (Table 6), respectively suggesting that the diversity in subgroup-2 genotypes is greater than the variation in subgroup-1. The heterozygosity presenting among genotypes in subpopulation-1 is higher than that existing in subpopulation-2 (Table 6). These results revealed high genetic variation. Teshome *et al.* (2015) have successfully detected nine subgroups in their study of pea population with large admixture genotypes. Wang *et al*., (2015) have identified three subpopulations among 256 pea genotypes. Based on thirty-one SSRs markers, Jain *et al*. (2014) have divided a collection of ninety-six cultivars of pea into four groups.

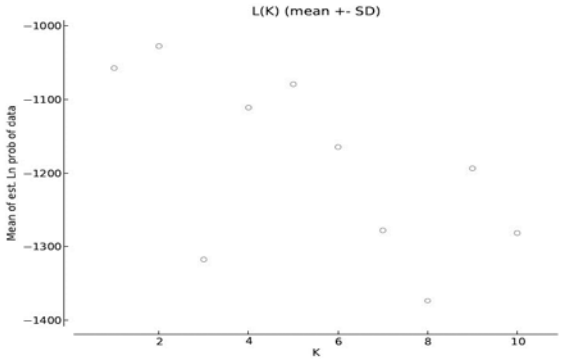

Figure 3. Graph of estimated membership fraction for the optimal value of K for 22 genotypes. The maximum of Ln prob determined by structure harvester is $K = 2$.

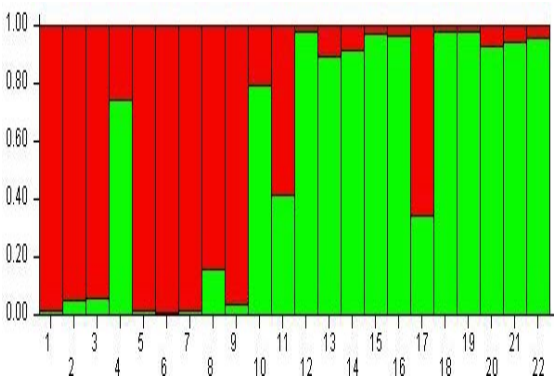

Figure 4. Population structure of 22 pea genotypes organized stand on inferred ancestry based on membership fractions, the genotypes with the probability of ≥ 80 % was set to the corresponding subpopulation, genotypes: 1: ORP-2, 2: ORP-3, 3: ORP-8, 4: ORP-10, 5: ORP-11, 6: ORP-12, 7: ORP-13, 8: ORP-15, 9: ORP-58, 10: ORP-D1, 11: Ns minima, 12: Green sage, 13: Oregon, 14: Provence, 15: Javor, 16: Legacy, 17: Karina, 18: Topaz, 19: Vada nunheins, 20: Bolero, 21: Lancet and 22: Rainer.

Table 5. Population structure cluster of pea genotypes based on inferred ancestry values.

Genotypes	Sub-polpulation-1	Sub-polpulation-2	Inferred cluster
G1	0.984	0.016	1
G2	0.945	0.055	1
G3	0.936	0.064	1
G4	0.254	0.746	1 and 2 (admixture)
G5	0.985	0.015	1
G6	0.986	0.014	1
G7	0.979	0.021	1
G8	0.836	0.164	1
G9	0.961	0.039	1
G10	0.200	0.800	2
G11	0.582	0.418	1 and 2 (admixture)
G12	0.017	0.983	2
G13	0.103	0.897	2
G14	0.084	0.916	2
G15	0.027	0.973	2
G16	0.029	0.971	2
G17	0.655	0.345	1 and 2 (admixture)
G18	0.020	0.980	2
G19	0.016	0.984	2
G20	0.071	0.929	2
G21	0.051	0.949	2
G22	0.036	0.964	2

Genotypes: G1: ORP-2, G2: ORP-3, G3: ORP-8, G4: ORP-10, G5: ORP-11, G6: ORP-12, G7: ORP-13, G8: ORP-15, G9: ORP-58, G10: ORP-D1, G11: Ns minima, G12: Green sage, G13: Oregon, G14: Provence, G15: Javor, G16: Legacy, G17: Karina, G18: Topaz, G19: Vada nunheins, G20: Bolero, G21: Lancet and G22: Rainer.

Table 6. Fixation indices (Fst) and heterozygosity within population created by structure population.

Mean value of Fst-subpopulation-1	0.0478
Mean value of Fst-subpopulation-2	0.267
Heterozygosity of subpopulation-1	0.324
Heterozygosity of subpopulation-2	0.276

4. Conclusions

Based on different statistical analyses including PCA and AHC, the current study identified three clades with 2-3 sub-clusters within twenty-two pea genotypes selected for performing the association mapping panel and breeding program. The allele information and diversity parameters have indicated the existence of a large genetic base in this collection. The output structure analysis in this investigation is not in accordance with the clustering method and principal component analysis. Thus, the results of this study indicate that the determination of genetic variation among pea genotypes by SSR markers can be useful for parental genotype selection in breeding programs.

Acknowledgment

We thank our colleagues from the Kurdistan Institution for Strategic Studies and Scientific Research who provided insight and expertise which greatly assisted with this research.

References

Ahmad S, Kaur S, Lamb-Palmer ND, Lefsrud M and Singh J. 2015. Genetic diversity and population structure of *Pisum sativum* accessions for marker-trait association of lipid content. *Crop J.*, **3**: 238–245.

Bouhadida M, Benjannet R, Madrid E, Amri M and Kharat M. 2013. Efficiency of marker-assisted selection in detection of *ascochyta* blight resistance in Tunisian chickpea breeding lines. *Phytopathol Mediterr.*, **52**: 202–211.

Burstin J, Deniot G, Potier J, Weinachter C, Aubert G and Baranger A. 2001. Microsatellite polymorphism in *Pisum sativum*. *Plant Breed.*, **120**: 311–317.

Cupic T, Tucak M, Popovic S, Bolaric S, Grljusic S, Kozumplik V and Bieb PM. 2009. Genetic diversity of pea (*Pisum sativum* L .) genotypes assessed by pedigree, morphological and molecular data. *Environment*, **7**: 343–348.

Deulvot C, Charrel H, Marty A, Jacquin F, Donnadieu C, Lejeune-Hénaut I, Burstin J and Aubert G. 2010. Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. *BMC Genomics*, **11**: 468.

Doyle J. 1991. CTAB Total DNA Isolation. Mol. Tech. Taxon. 283–293.

Hollingsworth M. 2006. **DNA Fingerprinting in Plants: Principles, Methods**, **and Applications,** 2nd Edn. Weising K, Nybom H, Wolff K, Kahl G. 2005. Boca Raton: CRC Press.

Jain S, Kumar A, Mamidi S and 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.

Kumari P, Basal N, Singh AK, Rai VP, Srivastava CP and Singh PK. 2013. Genetic diversity studies in pea (*Pisum sativum* L.) using simple sequence repeat markers. *Genet Mol Res*., **12 (3):** 3540-3550.

Lateef DD. 2015. DNA marker technologies in plants and applications for crop improvements. *J Biosci Med.*, **3**: 7–18.

Liu K and Muse SV. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, **21(9)**: 2128–2129.

Liu SR, Li WY, Long D, Hu CG and Zhang JZ. 2013. Development and characterization of genomic and expressed SSRs in citrus by genome-wide analysis M. Sun [ed.],. *PLoS One*, **8**: e75149.

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

Nasiri J, Haghnazari A and Saba J. 2010. Genetic diversity among varieties and wild species accessions of pea (*Pisum sativum* L.) based on SSR markers. *African J Biotechnol.*, **8**: 3405–3417.

Nei M and Genetics P. 1973. Analysis of gene diversity in subdivided populations. *Proc Nat Acad Sci.*,**70**: 3321–3323.

Peakall R and Smouse PE. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. **28**: 2537–2539.

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

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.

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

Reis CMG and Diogo MG. 2012. Identification of field pea cultivarsusing microsatellite molecular markers. *Plant Pysiol Genet.*, **2**: 57–63.

Sarikamiş G, Yanmaz R, Ermiş S, Bakir M and Yükse lC. 2010. Genetic characterization of pea (*Pisum sativum*) germplasm from Turkey using morphological and SSR markers. *GenetMol Res.*, **9**: 591–600.

Smýkal P, AubertG, BurstinJ, Coyne CJ, Ellis NTH, Flavell AJ, Ford R, Hýbl M, Macas J, Neumann P, McPhee KE, Redden RJ, Rubiales D, Weller JL and Warkentin TD. 2012. Pea (*Pisum sativum* L.) in the Genomic Era. *Agronomy*, **2**: 74–115.

Smýkal P, Horáček J, Dostálová R and Hýbl M. 2008. Variety discrimination in pea (*Pisum sativum* L.) by molecular, biochemical and morphological markers. *J Appl Genet.*, **49**: 155–166.

Tahir NA, Hamakareem HF and Hama-Amin BO. 2016. Differentiate of ten pea (*Pisum sativum* L) cultivars by RAPD markers and seeds storage proteins. *JJAS*, **11**: 95–107.

Tahir NA. 2010. Germination characteristics and molecular characterizations of some wheat varieties in Sulaimanyah by SSR marker. *Turkish J Biol.*, **34**: 109-117.

Teshome A, Bryngelsson T, Dagne K and GeletaM. 2015. Assessment of genetic diversity in Ethiopian field pea (*Pisum sativum* L.) accessions with newly developed EST-SSR markers. *BMC Genet.*,**16**: 102.

Wang F, Yang T, Burlyaeva M, Li L, Jiang J, Fang L, Redden R and Zong X. 2015. Genetic diversity of grasspea and its relative species revealed by SSR markers. *PLoS One*, **10**: e0118542.

XLSTAT. 2017. **Data Analysis and Statistical Solution for Microsoft Excel. Addinsoft,** Paris, France.